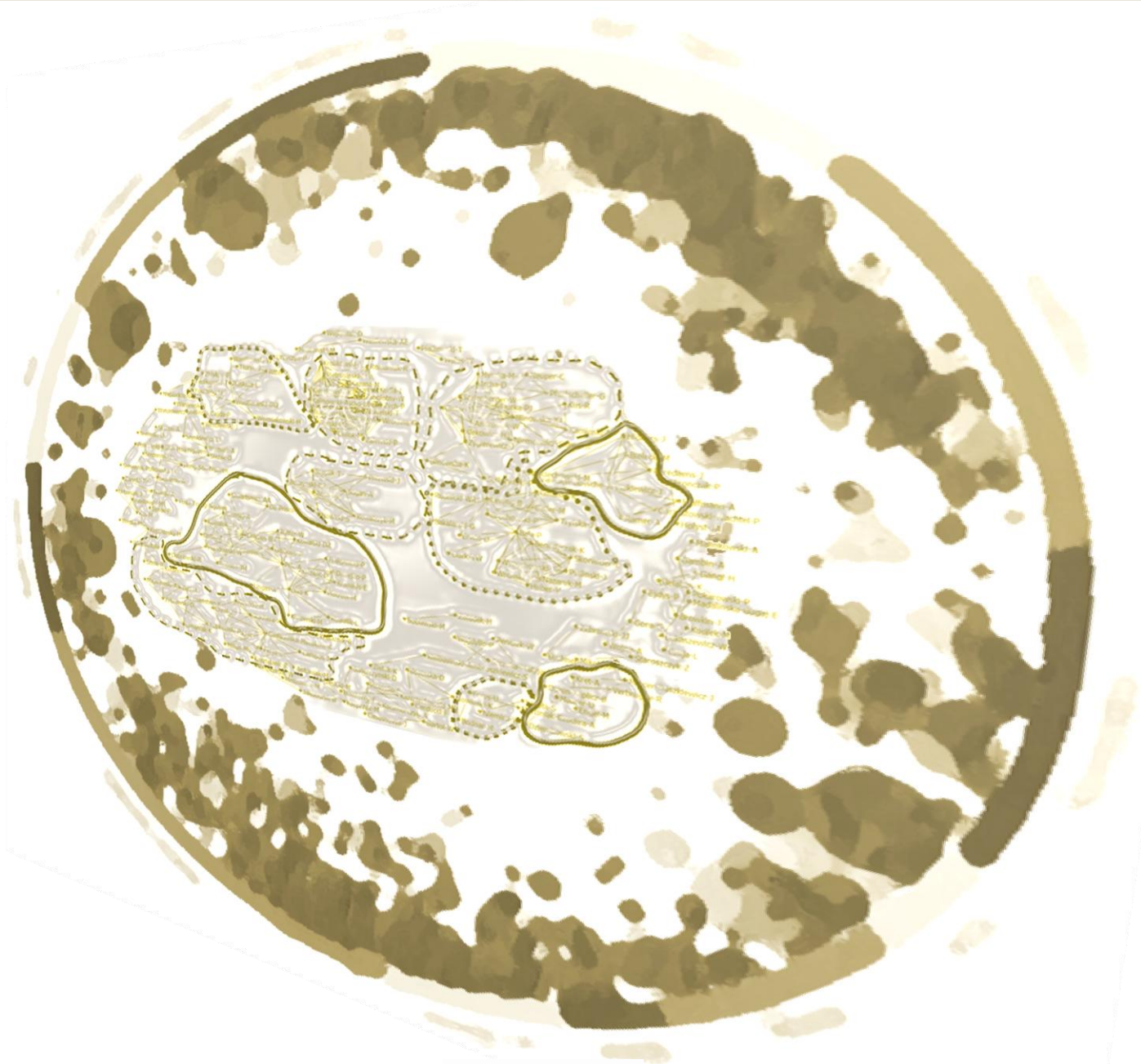


Krzysztof Klincewicz
Marcin Żemigła
Michał Mijał

Bibliometria

w zarządzaniu technologiami i badaniami naukowymi



**INNOWACYJNA
GOSPODARKA**
NARODOWA STRATEGIA SPÓJNOŚCI



UNIA EUROPEJSKA
EUROPEJSKI FUNDUSZ
ROZWOJU REGIONALNEGO



Ministerstwo Nauki i Szkolnictwa Wyższego

Krzysztof Klincewicz, Marcin Żemigła, Michał Mijał

Bibliometria

w zarządzaniu technologiami i badaniami naukowymi



Ministerstwo Nauki i Szkolnictwa Wyższego

Warszawa 2012

Autorzy:

prof. UW dr hab. Krzysztof Klincewicz (rozdziały 1-9, 11-15)

dr inż. Marcin Żemigala (rozdział 10)

dr Michał Mijał (rozdział 11)

Publikacja powstała w ramach projektu współfinansowanego ze środków UE w ramach projektu systemowego Ministerstwa Nauki i Szkolnictwa Wyższego "Wsparcie systemu zarządzania badaniami naukowymi oraz ich wynikami" (Poddziałanie 1.1.3 Programu Operacyjnego Innowacyjna Gospodarka).

Publikacja bezpłatna

© Copyright by Ministerstwo Nauki i Szkolnictwa Wyższego, Warszawa 2012

Spis treści

O czym będzie ta książka? (Krzysztof Klincewicz).....	7
1. Wprowadzenie (Krzysztof Klincewicz)	14
1.1. Podstawowe pojęcia	14
1.2. Struktura książki	14
2. Potrzeby informacyjne uczestników systemu B+R (Krzysztof Klincewicz).....	16
2.1. Znaczenie analiz bibliometrycznych dla strategicznego zarządzania nauką.....	16
2.2. Uczestnicy systemu innowacji i ich potrzeby informacyjne	16
2.3. Kierunki analiz bibliometrycznych w odpowiedzi na potrzeby interesariuszy	21
3. Analizy bibliometryczne w świetle dorobku dyscypliny zarządzania technologiami (Krzysztof Klincewicz)	22
3.1. Zarządzanie technologiami a analizy bibliometryczne.....	22
3.2. Modele tworzenia innowacji	22
3.3. Rezultaty działalności badawczo-rozwojowej.....	24
3.4. Rodzaje działalności badawczo-rozwojowej.....	26
3.5. Rozwój technologii.....	28
3.6. Wnioski do wykorzystania w analizach bibliometrycznych	32
4. Bibliometria a inne techniki analityczne (Krzysztof Klincewicz)	34
4.1. Techniki wspierające obserwację i przewidywanie rozwoju nauki i techniki.....	34
4.2. Bibliometria na tle innych technik.....	35
4.3. Wykorzystanie bibliometrii do analiz systemu B+R.....	37
4.4. Zalety i ograniczenia bibliometrii.....	39
5. Analizy bibliometryczne w świetle doświadczeń międzynarodowych (Krzysztof Klincewicz)	41
5.1. Wykorzystanie bibliometrii przez zagraniczne organizacje	41
5.2. Doświadczenia instytucji rządowych i wojskowych w Stanach Zjednoczonych	41
5.3. Doświadczenia instytucji rządowych innych krajów	43
5.4. Doświadczenia firm technologicznych.....	46
6. Dane bibliometryczne wykorzystywane do celów analitycznych (Krzysztof Klincewicz)	51
6.1. Dane bibliometryczne i bazy bibliograficzne	51
6.2. Dane bibliometryczne dostępne w bazach publikacji i streszczeń	51
6.3. Dane bibliometryczne dostępne w bazach patentów	55
6.4. Zróznicowanie aktywności badawczej poszczególnych krajów.....	57

7.	Pytania analityczne w bibliometrii (Krzysztof Klincewicz)	60
8.	Podstawowe techniki analityczne w bibliometrii (Krzysztof Klincewicz)	69
8.1.	Podstawowe techniki bibliometryczne	69
8.2.	Tworzenie zestawu danych bibliometrycznych do dalszych analiz	70
8.3.	Listy rankingowe	77
8.4.	Analiza współwystępowania słów (<i>co-word</i>)	80
8.5.	Analiza współklasyfikacji (<i>co-classification</i>)	82
8.6.	Analiza współautorstwa (<i>co-authorship</i>)	83
8.7.	Analiza cytowań (<i>citation</i>)	85
8.8.	Warianty analizy cytowań	91
8.9.	Podstawowe wskaźniki bibliometryczne	94
9.	Zaawansowane techniki analityczne w bibliometrii (Krzysztof Klincewicz)	97
9.1.	Zaawansowane techniki bibliometryczne	97
9.2.	<i>Tech mining</i>	97
9.3.	Tomografia bazodanowa	112
9.4.	<i>Literature-Based Discovery</i>	113
9.5.	<i>TRIZ</i>	118
9.6.	Identyfikacja nowych trendów i tworzenie map badań	126
9.7.	Zaawansowane wskaźniki bibliometryczne	138
10.	Bazy wykorzystywane do analiz bibliometrycznych (Marcin Żemigła)	144
10.1.	Wprowadzenie	144
10.2.	Przegląd baz bibliograficznych	144
10.2.1.	<i>Chemical Abstracts Service/ SciFinder (CAS)</i>	144
10.2.2.	<i>Compendex</i>	146
10.2.3.	<i>ACM Digital Library</i>	147
10.2.4.	<i>Scirus</i>	148
10.2.5.	<i>Google Scholar</i>	149
10.2.6.	<i>Web of Science</i>	150
10.2.7.	<i>Scopus</i>	151
10.2.8.	<i>INSPEC</i>	151
10.2.9.	<i>MEDLINE</i>	152
10.3.	Opis bazy <i>Web of Science</i>	152
10.4.	Opis bazy <i>Scopus</i>	158
10.5.	Opis bazy <i>INSPEC</i>	166

10.6.	Opis bazy <i>MEDLINE</i>	171
10.7.	Porównanie funkcjonalności baz bibliograficznych	177
10.8.	Opis bazy Europejskiego Urzędu Patentowego (<i>EPO</i>).....	182
10.9.	Opis bazy Urzędu Patentowego Stanów Zjednoczonych (<i>USPTO</i>)	188
10.10.	Opis bazy <i>Google Patents</i>	195
10.11.	Porównanie funkcjonalności baz patentów	201
10.12.	Przegląd komercyjnych baz patentów.....	204
10.12.1.	<i>Delphion</i>	204
10.12.2.	<i>Micro Patent / Patent Web</i>	207
10.12.3.	<i>Pat Base</i>	208
10.12.4.	<i>QPat</i>	210
10.12.5.	<i>WIPS Global</i>	214
10.12.6.	<i>Total Patent</i>	216
10.12.7.	<i>Thomson Innovation</i>	217
10.12.8.	<i>SureChem</i>	218
10.12.9.	<i>JP-NETe</i>	219
10.12.10.	<i>Patent Integration</i>	219
11.	Oprogramowanie wykorzystywane do analiz bibliometrycznych oraz wizualizacji ich wyników (Michał Mijał, Krzysztof Klincewicz)	222
11.1.	Wprowadzenie	222
11.2.	Rodzaje oprogramowania, wspierającego analizy bibliometryczne	223
11.3.	Opis programu <i>Matheo Analyzer</i>	225
11.4.	Opis programu <i>CiteSpace</i>	231
11.5.	Opis programu <i>VantagePoint</i>	236
11.6.	Opis programu <i>HistCite</i>	240
11.7.	Opis programu <i>PatentNet</i>	243
11.8.	Porównanie zastosowań opisanych programów	246
11.9.	Przegląd innych programów, wspierających analizy bibliometryczne	248
12.	Rekomendacje dotyczące prowadzenia analiz bibliometrycznych (Krzysztof Klincewicz)	254
12.1.	Wprowadzenie	254
12.2.	Wybór pytań badawczych.....	255
12.3.	Przydatność technik analitycznych	260
12.4.	Przydatność baz bibliograficznych i oprogramowania	264

13. Podstawowe analizy bibliometryczne – przykład analizy badań nad grafenem (Krzysztof Klincewicz)	279
14. Analiza bibliometryczna aktywności publikacyjnej polskich uczelni wyższych, 2000-2009 (Krzysztof Klincewicz)	291
15. Analiza bibliometryczna badań nad laserami półprzewodnikowymi, 1988-2005 (Krzysztof Klincewicz)	298
Literatura	306
Wykaz rysunków	323
Wykaz tabel	331
Informacja o autorach	336

O czym będzie ta książka?

Niniejsza książka została przygotowana na zamówienie Departamentu Strategii Ministerstwa Nauki i Szkolnictwa Wyższego (MNiSW) w związku z planami stworzenia „Obserwatorium badań naukowych i rozwoju technologii”. Książka prezentuje możliwości **wykorzystania bibliometrii** w zarządzaniu technologiami i badaniami naukowymi. **Bibliometria** to zbiór technik badawczych, stosowanych do ilościowych analiz publikacji, w tym publikacji naukowych i dokumentów patentowych. Bibliometria lub szerzej naukometria (ang. *scientometrics*) może być również traktowana jako odrębna dyscyplina badawcza w obszarze naukoznawstwa, zajmująca się badaniami rozwoju nauki jako procesu informacyjnego. Do typowych zastosowań praktycznych bibliometrii należy analiza działalności badawczo-rozwojowej, prowadzona w związku z potrzebami osób zarządzających podmiotami w sektorze B+R oraz formułujących politykę naukową lub politykę innowacyjną.

Rozdział 1 wyjaśnia zakres publikacji i omawia podstawowe pojęcia.

Rozdział 2 opisuje **potrzeby informacyjne** uczestników systemu B+R, wskazując na celowość podejmowania analiz bibliometrycznych oraz charakteryzując potrzeby i oczekiwania ich typowych interesariuszy. Diagnoza potrzeb interesariuszy odwołuje się do modelu *evidence-based policy*, w którym decyzje o wyborze kierunków strategicznych i alokacji środków powinny opierać się na gromadzeniu i analizie danych empirycznych. Szczegółowo omówiono potrzeby informacyjne reprezentantów resortu nauki (w tym Ministerstwa Nauki i Szkolnictwa Wyższego, Narodowego Centrum Badań i Rozwoju, Narodowego Centrum Nauki, Komitetu Ewaluacji Jednostek Naukowych i Ośrodka Przetwarzania Informacji), innych instytucji rządowych (w tym: Ministerstwa Edukacji Narodowej, Ministerstwa Finansów, Ministerstwa Gospodarki, Ministerstwa Infrastruktury, Ministerstwa Rozwoju Regionalnego, Ministerstwa Administracji i Cyfryzacji, Ministerstwa Środowiska, Ministerstwa Zdrowia, Głównego Urzędu Statystycznego, Polskiej Agencji Rozwoju Przedsiębiorczości, Polskiej Agencji Informacji i Inwestycji Zagranicznych oraz Urzędu Patentowego RP), jak również uczelni wyższych, instytutów badawczych, jednostek naukowych Polskiej Akademii Nauk i przedsiębiorstw technologicznych.

Diagnoza potrzeb informacyjnych wskazuje na możliwości podjęcia czterech szerokich kierunków działań:

1. **monitorowanie i skanowanie technologii** – gromadzenie informacji o tym, co było i/lub jest przedmiotem aktywności badawczo-rozwojowej (dalej w książce określanej jako: B+R) w poszczególnych obszarach technologicznych (monitorowanie oznacza działania nieukierunkowane na konkretne pytania analityczne, skanowanie – ukierunkowane),
2. **wywiad technologiczny** (*technology intelligence*) – zdobywanie danych o tym, kto zajmuje się poszczególnymi technologiami i obszarami badań oraz na jakich aspektach koncentruje wysiłki badawcze,
3. **prognozowanie rozwoju technologii** - w tym analiza trendów w zakresie badań naukowych i prac nad doskonaleniem technologii oraz przewidywanie ewolucji obecnie dostępnych klas rozwiązań,

4. **ocena technologii** (*technology assessment*) – ocena znaczenia indywidualnych osiągnięć oraz pozycji konkurencyjnej badaczy, instytucji, regionów i kraju.

Rozdział 3 oferuje spojrzenie na analizy bibliometryczne w świetle **dorobku dyscypliny zarządzania technologiami**, która stanowi metodologiczny i merytoryczny fundament dla obserwacji badań naukowych i rozwoju technologii. Rozdział charakteryzuje podstawowe modele tworzenia innowacji, oferuje typologię działalności badawczo-rozwojowej oraz wskazuje na typowe mechanizmy rozwoju technologii poprzez odwołanie do literatury zachodniej i rosyjskiej. Przegląd dorobku teoretyków zarządzania technologiami pozwala wskazać zastosowania technik bibliometrycznych w narodowym systemie innowacji:

- przewidywanie rozwoju określonych obszarów badań naukowych i technologii, w szczególności określanie kierunków i potencjału dalszego rozwoju,
- identyfikacja obszarów wiedzy, niezbędnych dla skutecznego prowadzenia badań naukowych i prac rozwojowych w poszczególnych obszarach,
- identyfikacja przypadków fuzji technologii oraz nowych, powstających obszarów badawczych i technologicznych (*emerging technologies*),
- identyfikacja obszarów badawczych i rozwiązań technologicznych, które mogą stopniowo wypierać dotychczas stosowane rozwiązania i ograniczać przydatność zgromadzonej wiedzy (*disruptive technologies, competence-destroying innovations*),
- gromadzenie i udostępnianie wiedzy, dotyczącej tendencji w obszarze badań naukowych i rozwoju technologii, które może służyć podnoszeniu zdolności absorpcyjnych pracowników instytucji naukowych i firm oraz stanowić inspirację do własnych badań i wysiłków wynalazczych.

Rozdział 4 zestawia techniki bibliometryczne z innymi podejściami do analizy trendów w obszarze badań naukowych i rozwoju technologii, takimi jak *foresight*, *technology roadmapping* i *competitive technology intelligence*. Zwraca również uwagę na dwa **podstawowe zastosowania bibliometrii: ewaluacyjne** (służące ocenie badaczy i jednostek naukowych) oraz **deskryptywne** (pozwalające na obserwację tendencji w rozwoju nauki i technologii, identyfikację istotnych aktorów na scenie innowacji oraz lepsze zrozumienie specyfiki poszczególnych obszarów badań). Kolejne rozdziały książki koncentrują się na deskryptywnych zastosowaniach bibliometrii.

Dalsza część rozdziału uświadamia czytelnikom **zalety i ograniczenia technik bibliometrycznych**. Bibliometria pozwala na analizy ilościowe, zobiektywizowane, odwołujące się do wymiernych rezultatów prac B+R, pozwalające na prowadzenie analiz przez osoby niezaangażowane przy wykorzystaniu nieinwazyjnych technik gromadzenia danych. Z drugiej strony, niewłaściwy dobór danych oraz nierzetelna analiza mogą w łatwy sposób doprowadzić do powstania błędów interpretacyjnych, dlatego stosowanie bibliometrii wymaga dobrego zrozumienia uwarunkowań metodologicznych i najlepszych praktyk analitycznych.

Rozdział 5 przedstawia **doświadczenia innych instytucji w zakresie obserwacji badań naukowych i rozwoju technologii**. Do pionierów systematycznych obserwacji należą Stany Zjednoczone, w szczególności Departament Obrony. Opisano powstanie najpopularniejszego rozwiązania informatycznego, wspierającego analizy bibliometryczne, początkowo

finansowanego z budżetu obronnego USA. Wskazano też na przykłady aktywności analitycznej w obszarze bibliometrii, dostępnej na militarnym portalu *DTIC (Defense Technical Information Center)*. Szczegółowo porównano też techniki analityczne, wykorzystywane przez agencje rządowe *National Science Foundation (NSF)*, *National Institute of Health (NIH)*, *Department of Energy (DOE)* oraz *National Institute of Standards and Technology (NIST)* w ramach *Advanced Technology Program (ATP)*. W kolejnym podrozdziale omówiono doświadczenia instytucji rządowych z innych krajów, w tym Francji, Holandii, Brazylii, Turcji, Australii, Indii, Malezji, Japonii, Korei Południowej, Niemiec i Finlandii. W każdym z tych krajów, opracowanie raportów analitycznych opartych o dane bibliometryczne było inspiracją do krajowych debat nad kierunkami rozwoju prac B+R oraz mechanizmami finansowania nauki. Przedstawiono również zakres tematyczny obserwacji, prowadzonych w poszczególnych krajach, w podziale na obszary technologiczne. Rozdział dokumentuje wreszcie doświadczenia konkretnych firm technologicznych oraz analizy porównawcze tendencji do wykorzystywania technik obserwacyjnych oraz źródeł informacji o otoczeniu technologicznym przez przedsiębiorstwa z poszczególnych branż.

Rozdział 6 rozpoczyna cykl **praktycznych porad, dotyczących prowadzenia analiz bibliometrycznych** i koncentruje się na **specyfice danych bibliometrycznych**. Zaprezentowane zostają bazy wykorzystywane do celów analiz bibliometrycznych (bazy publikacji, streszczeń, patentów i zgłoszeń patentowych) oraz typowa struktura danych w tych bazach. Rozdział podsumowują zestawienia wymiernych przejawów aktywności badawczej - patentów i publikacji - dla wybranych krajów, wskazujące na przydatność analiz bibliometrycznych w odniesieniu do systemów innowacji tych gospodarek.

Rozdział 7 oferuje **systematyczne zestawienie pytań analitycznych**, które mogą być zadane w projektach bibliometrycznych. Dla każdego pytania, wskazany został stopień trudności analiz - konieczność posiadania wiedzy eksperckiej w zakresie analizowanej dyscypliny oraz złożoność przetwarzania danych. Przedstawiona lista pytań jest wykorzystywana w dalszych częściach książki jako ważny punkt odniesienia i podstawa do oceny możliwości zastosowania poszczególnych technik analitycznych, baz danych i oprogramowania oraz sposób definiowania zakresu prac w projektach badawczych. Lista obejmuje następujące pytania:

- *Co?* - Identyfikacja obszarów badań i technologii w zbiorze danych bibliometrycznych dla określonego kraju, regionu, ośrodka badawczego lub dla określonych lat
- *Co dokładnie?* - Identyfikacja podobszarów badawczych w zbiorze danych bibliometrycznych, odzwierciedlających wybrany obszar badań i technologii)
- *Kto?* - Identyfikacja badaczy/wynalazców i zespołów badawczych, prowadzących badania na określony temat
- *Gdzie?* - Identyfikacja afiliacji, w których prowadzone są badania na określony temat (poziom kraju, miasta, organizacji)
- *Kiedy?* - Określenie momentu publikacji, zgłoszenia patentowego lub uzyskania patentu, jak również okresów, w których prowadzono badania na określony temat

- *Z kim?* - Analiza współautorstwa publikacji i wynalazków poprzez identyfikację osób, które wspólnie prowadziły badania
- *Z jaką organizacją?* - Analiza współautorstwa publikacji i wynalazków (*co-authorship*) poprzez identyfikację organizacji, zatrudniających osoby, które wspólnie prowadziły badania
- *Ile?* - Zliczanie publikacji, zgłoszeń patentowych lub udzielonych patentów dla wybranego obszaru lub podobszaru badań i technologii (możliwe jest także sporządzanie zestawień porównawczych liczby publikacji w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją?)
- *Ilu?* - Zestawienie liczby autorów publikacji lub zgłoszeń patentowych w danym obszarze
- *Jak zmienia się w czasie?* - Analiza szeregów czasowych, odzwierciedlających liczby publikacji lub patentów w wybranych obszarach badań i technologii (możliwe jest odzwierciedlenie zmian w czasie w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? ile? – czyli prezentacja historycznego rozwoju obszaru badań)
- *Jak będzie się rozwijać?* - Prognoza krótkookresowa, oparta na ekstrapolacji szeregów czasowych (możliwe jest przygotowanie prognoz w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? pod warunkiem posiadania dostatecznie długich serii danych historycznych)
- *Jak popularne w porównaniu do?* - Porównanie liczb publikacji, zgłoszeń patentowych lub patentów w kilku obszarach badawczych
- *Jakie są nisze?* - Identyfikacja podobszarów badawczych, które nadal cieszą się niewielkim zainteresowaniem badaczy (w oparciu o liczbę publikacji, zgłoszeń patentowych lub patentów w danym podobszarze, porównaną z ich globalną liczbą dla szerokiego obszaru badań)
- *Jakie są nowe popularne tematy?* - Identyfikacja nowych, cieszących się narastającą popularnością tematów badawczych (*hot topics*), w tym dynamicznie rozwijających się, nowych technologii (*emerging technologies*)
- *Z czym związane?* - Analiza pokrewnych obszarów badań i tematów badawczych, które równocześnie opisywane są w publikacjach lub dokumentach patentowych – na poziomie współwystępowania specjalistycznych pojęć w streszczeniach i tytułach (*co-word*) lub w słowach kluczowych i kodach klasyfikacyjnych (*co-classification*)
- *W jaki sposób związane?* - Analiza szczegółowych podejść badawczych, technik, środków i technologii, wykorzystywanych przez zidentyfikowane wcześniej obszary badań i tematy badawcze, pokrewne wobec podstawowego obszaru analiz
- *Co uzupełnia?* - Identyfikacja komplementarnych obszarów wiedzy naukowej, komplementarnych umiejętności badawczych lub komplementarnych technologii, wykorzystywanych przy prowadzeniu badań w danych obszarze
- *Jakich wariantów nie analizowano?* - Identyfikacja możliwych nowych kierunków badań i sposobów rozwiązywania problemów naukowych i praktycznych w oparciu o analogie, techniki kreatywnego rozwiązywania problemów i ocenę potencjału ewolucyjnego

- *Kto uzupełnia?* - Identyfikacja osób i organizacji, posiadających komplementarne obszary wiedzy naukowej, komplementarne umiejętności badawcze lub komplementarne technologie, wykorzystywane przy prowadzeniu badań w danym obszarze
- *Kto może być dostawcą?* - Identyfikacja osób i organizacji, które mogą być podwykonawcami określonych zleceń badawczych ze względu na dotychczasowy dorobek naukowy
- *Kto może być odbiorcą?* - Identyfikacja osób i organizacji, które mogą być zleceniodawcami projektów badawczo-rozwojowych lub nabywcami/licencjobiorcami ich rezultatów ze względu na dotychczasowe zainteresowania i potrzeby badawcze
- *Kto jest konkurentem?* - Identyfikacja osób i organizacji, prowadzących badania w analogicznych, szczegółowych obszarach badawczych, przy wykorzystaniu zbliżonych metod (podmioty zidentyfikowane jako konkurenci mogą być też potencjalnymi partnerami w konsorcjach badawczych – zwłaszcza wtedy, gdy niezbędne jest połączenie zbliżonych zasobów dla zwiększenia skali prac badawczych)
- *Kto może być konkurentem?* - Identyfikacja osób i organizacji, które mogą podjąć badania w analogicznych, szczegółowych obszarach badawczych, przy wykorzystaniu zbliżonych metod, a wskazuje na to ich profil badawczy, dotychczasowe zainteresowania i analizy statystyczne (zwłaszcza techniki analizy czynnikowej, pozwalające na identyfikację „ukrytych” podobieństw i zależności w dużych zbiorach danych)
- *Kto może być pracownikiem?* - Identyfikacja osób, które prowadzą lub mogą prowadzić badania w określonym obszarze i zgromadziły zadowalający dorobek naukowy
- *Kto może być sukcesorem?* - Identyfikacja osób, które prowadziły badania wspólnie z określonym naukowcem (*co-authorship*) i mogą w przyszłości przejąć obowiązki w zakresie zarządzania projektami badawczymi
- *Jakie X ma strategie badawcze?* - Analiza tematyki badawczej, mechanizmów współpracy autorskiej (*co-authorship*), kluczowych badaczy i zespołów badawczych, partnerów zewnętrznych i zmian w powyższych obszarach dla wybranej organizacji
- *Jak ważne?* - Zestawienie liczby cytowań dla pojedynczych publikacji lub patentów
- *Jak łącznie ważne?* - Zestawienie zagregowanej liczby cytowań dla grupy publikacji lub patentów (np. pochodzących od autorów z jednej organizacji, dotyczących jednego obszaru badawczego)
- *Kto jest ważny?* - Porównania liczby cytowań w dorobku danych badaczy oraz obliczanie wskaźników dotyczących cytowań (np. indeksów h Hirscha)
- *Do jakiej wiedzy nawiązuje?* - Identyfikacja źródeł wiedzy, ujawnionych w cytowaniach, zawartych w publikacjach naukowych i zgłoszeniach patentowych (możliwe dalsze analizy zidentyfikowanego zbioru źródeł wiedzy w oparciu o pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? ile?)
- *Do których badań nawiązuje?* - Identyfikacja naukowców i wynalazców, do prac których odwołują się cytowania w publikacjach naukowych i zgłoszeniach patentowych

Rozdział 8 omawia praktyczne aspekty stosowania **podstawowych technik bibliometrycznych**: sposoby tworzenia zestawu danych bibliometrycznych do dalszych analiz (z wykorzystaniem mechanizmów wyszukiwania, oferowanych przez bazy bibliograficzne oraz ze wskazaniem na typowe ograniczenia i problemy związane z wyszukiwaniem), budowę list rankingowych, analizy współwystępowania słów (*co-word*), analizy współklasyfikacji (*co-classification*), analizy współautorstwa (*co-authorship*) oraz warianty analizy cytowań (*citation*). Opisane zostały również typowe wskaźniki bibliometryczne. Zaprezentowane w rozdziale 8 techniki mogą być zwykle wykorzystywane nawet bez korzystania z zaawansowanych, kosztownych narzędzi informatycznych. Stanowią podstawowe sposoby przetwarzania danych bibliometrycznych, jednak ich stosowanie wiąże się z licznymi ograniczeniami. Rozdział oferuje przegląd ważnych pojęć specjalistycznych i koncepcji, wykorzystywanych w analizach bibliometrycznych, stanowiąc podstawę do lektury kolejnych części książki.

Rozdział 9 prezentuje **zaawansowane techniki bibliometryczne** - *tech mining*, tomografię bazodanową, *Literature-Based Discovery*, *TRIZ*, sposoby identyfikacji nowych trendów i tworzenia map badań oraz możliwości wykorzystywania zaawansowanych wskaźników bibliometrycznych. Opiera się na terminologii zaprezentowanej w rozdziale poprzednim i szczególnie omawia techniki, które mogą znaleźć zastosowanie w praktycznej pracy badaczy, zajmujących się bibliometrią. Szczególnie interesujące wydają się podejścia ***tech mining* (eksploracji danych technologicznych)** oraz wspierane informatycznie techniki **identyfikacji znaczących zmian w tematyce badawczej**, pozwalające na śledzenie pojawiania się i rozwoju nowych zagadnień badawczych oraz nowych technologii. Znaczący potencjał analityczny oferują też opisane **techniki wizualizacji danych**, pozwalające w graficzny sposób odzwierciedlić zależności pomiędzy badaczami, zatrudniającymi ich instytucjami i tematami badań, a nawet **"rzutować" dorobek badawczy na światową mapę nauki**, wskazując na interdyscyplinarny charakter badań i ułatwiając międzynarodowe analizy porównawcze.

Rozdział 10 oferuje rozbudowany przegląd **baz, wykorzystywanych w analizach bibliometrycznych**. Zostały one zaprezentowane w podziale na bazy bibliograficzne (publikacji i streszczeń) oraz patentowe (zgłoszeń patentowych i udzielonych patentów). Szczegółowo opisano najbardziej przydatne bazy - wśród baz bibliograficznych: *Web of Science*, *Scopus*, *INSPEC*, *MEDLINE*, a wśród baz patentowych: *Esp@cenet*, *USPTO* i *Google Patents*. Dodatkowo, zaprezentowane zostały inne komercyjne bazy: *Chemical Abstracts Service/SciFinder*, *Compendex*, *ACM Digital Library*, *Scirus*, *Google Scholar*, *Delphion*, *Micro Patent*, *Pat Base*, *QPat*, *WIPS Global*, *Total Patent*, *Thomson Innovation*, *SureChem*, *JP-NETe* oraz *Patent Integration*. Ta część książki ułatwia czytelnikowi wybór odpowiedniego źródła danych do konkretnego projektu analitycznego i wskazuje, że nie jest możliwe wskazanie jednej bazy, najlepiej zaspokajającej wszystkie możliwe potrzeby w projektach bibliometrycznych, choć poszczególne bazy będą lepiej nadawać się do udzielania odpowiedzi na konkretne pytania analityczne.

Rozdział 11 omawia dostępne **narzędzia informatyczne**, które mogą wspierać analizy bibliometryczne. Po przedstawieniu typologii aplikacji komputerowych, koncentruje się na szczegółowej analizie narzędzi: komercyjnych programów *Matheo Analyzer* i *VantagePoint*

oraz bezpłatnych narzędzi *CiteSpace*, *HistCite* i *PatentNet*. Dodatkowo, zaprezentowane zostało również zestawienie dodatkowych informacji, dotyczących kilkudziesięciu komercyjnych i bezpłatnych pakietów oprogramowania, które może wspierać specyficzne zastosowania analityczne.

Rozdział 12 oferuje **szczegółowe rekomendacje, dotyczące prowadzenia projektów bibliometrycznych**. Planowanie pracy rozpoczyna się od **wyboru konkretnych pytań badawczych**. W kolejnym kroku, wskazane zostały szczegółowe **techniki bibliometryczne**, które pomagają w udzieleniu odpowiedzi na konkretne pytania analityczne (w nawiązaniu do wcześniejszej prezentacji w rozdziałach 8 i 9). Dalsze decyzje mają bezpośrednie implikacje dla inwestycji, jakie muszą być poniesione w związku z prowadzeniem projektów bibliometrycznych i dotyczą **wyboru baz i programów komputerowych** dla celów poszczególnych analiz. Zestawienia tabelaryczne ujawniają przydatność baz i aplikacji w zależności od zadawanych pytań analitycznych.

Rozdziały 13-15 oferują **przykłady analiz bibliometrycznych**, prowadzonych w reakcji na określone potrzeby informacyjne. Pierwsza z nich odpowiada na dążenie do lepszego zrozumienia specyfiki badań nad grafenem, nowym materiałem półprzewodnikowym, w których znaczącą rolę odegrali polscy naukowcy. Druga prezentuje porównanie 24 uczelni wyższych w Polsce pod względem aktywności publikacyjnej. Trzecia odzwierciedla zaawansowane analizy badań nad laserami półprzewodnikowymi, z podziałem na miejsca prowadzenia badań i ich szczegółowe tematy. Dzięki wykorzystaniu wiedzy, zawartej w niniejszej książce, pracownicy organizacji z sektora B+R będą mogli samodzielnie prowadzić nawet bardziej złożone analizy, a zaprezentowane przykłady mogą stanowić dodatkową inspirację i uzasadnienie celowości podejmowania projektów analiz bibliometrycznych.

1. Wprowadzenie

1.1. Podstawowe pojęcia

Niniejsza książka została przygotowana na zamówienie Departamentu Strategii Ministerstwa Nauki i Szkolnictwa Wyższego (MNiSW) w związku z planami stworzenia „Obserwatorium badań naukowych i rozwoju technologii”. Jej opracowanie zostało zlecone w ramach projektu systemowego MNiSW „*Wsparcie systemu zarządzania badaniami naukowymi oraz ich wynikami*” (Poddziałanie 1.1.3 Programu Operacyjnego Innowacyjna Gospodarka).

Książka zawiera rozbudowaną analizę możliwości wykorzystania technik bibliometrycznych w zarządzaniu technologiami i badaniami naukowymi. **Bibliometria** to zbiór technik badawczych, stosowanych do ilościowych analiz publikacji, w tym publikacji naukowych i dokumentów patentowych. Bibliometria lub szerzej naukometria (ang. scientometrics) może być również traktowana jako odrębna dyscyplina badawcza w obszarze naukoznawstwa, zajmująca się badaniami rozwoju nauki jako procesu informacyjnego. Do typowych zastosowań praktycznych bibliometrii należy analiza działalności badawczo-rozwojowej, prowadzona w związku z potrzebami osób zarządzających podmiotami w sektorze B+R oraz formułujących politykę naukową lub politykę innowacyjną. Współcześnie obserwujemy zatarcie granic pomiędzy bibliometrią (tradycyjnie opartą na analizie publikacji) oraz patentometrią (pomiarem patentów) – oba podejścia stosują analogiczne techniki analityczne i narzędzia, a określenie „bibliometrii” obejmuje dziś również analizy patentów.

Bibliometria jest w Polsce mało popularną techniką badawczą, a zestawienie porównawcze dorobku polskich badaczy w obszarze bibliometrii oferuje Kozłowski (2010: 68). Prezentowana książka ma służyć jako poradnik wszystkim tym, którzy planują prowadzenie badań bibliometrycznych w obszarze zarządzania technologiami i badaniami naukowymi oraz pomoc dla osób analizujących lub tworzących politykę naukową. Rekomendowane zastosowania bibliometrii powinny koncentrować się na analizach dotyczących dorobku nauk biologicznych, chemicznych, farmaceutycznych, fizycznych, leśnych, matematycznych, medycznych, o Ziemi, rolniczych, technicznych, weterynaryjnych i wojskowych (zgodnie z podziałem, wprowadzonym przez uchwałę Centralnej Komisji do Spraw Stopni i Tytułów z dnia 24 października 2005 roku w sprawie określenia dziedzin nauki i dziedzin sztuki oraz dyscyplin naukowych i artystycznych). Opisane w książce analizy mają jedynie ograniczoną przydatność w odniesieniu do nauk humanistycznych, ekonomicznych, prawnych, teologicznych oraz sztuk.

1.2. Struktura książki

Książka posiada następującą **strukturę**:

- **Rozdział 2** opisuje potrzeby informacyjne uczestników systemu B+R, wskazując na przydatność analiz bibliometrycznych dla poszczególnych interesariuszy (**potrzeby i oczekiwania**),
- **Rozdział 3** oferuje przegląd podstawowych koncepcji teoretycznych z obszaru zarządzania technologiami, do których nawiązywać będą opisy poszczególnych technik bibliometrycznych w dalszych rozdziałach (**podstawy teoretyczne**),

- **Rozdział 4** uzasadnia celowość stosowania bibliometrii jako sposobu obserwacji badań naukowych i rozwoju technologii, wskazując jednocześnie jej ograniczenia (**argumenty na rzecz bibliometrii**),
- **Rozdział 5** prezentuje doświadczenia instytucji rządowych i przedsiębiorstw z innych krajów w zakresie wykorzystania technik bibliometrycznych do obserwacji trendów w nauce i technice (**benchmarking**),
- **Rozdział 6** rozpoczyna praktyczny opis możliwości wykorzystania bibliometrii, charakteryzując dane bibliometryczne wykorzystywane w analizach (**dane bibliometryczne**),
- **Rozdział 7** zestawia typowe pytania analityczne, zadawane w projektach bibliometrycznych (**pytania analityczne**),
- **Rozdział 8** prezentuje podstawowe techniki analityczne w bibliometrii (**podstawowe techniki**),
- **Rozdział 9** prezentuje zaawansowane techniki analityczne w bibliometrii (**zaawansowane techniki**),
- **Rozdział 10** oferuje przegląd baz publikacji i patentów oraz dyskusję możliwości ich wykorzystania dla celów analiz bibliometrycznych (**bazy**),
- **Rozdział 11** zawiera zestawienie oprogramowania, wykorzystywanego do analiz bibliometrycznych (**oprogramowanie**),
- **Rozdział 12** formułuje rekomendacje, dotyczące organizacji badań bibliometrycznych – wyboru pytań analitycznych, technik, baz i oprogramowania (**rekomendacje**),
- **Rozdziały 13-15** prezentują przykłady prostych i złożonych analiz bibliometrycznych, ilustrujące potencjał technik, omówionych w rozdziałach 8 i 9.
- **Spis literatury** zawiera wykaz pozycji, które zostały zacytowane w niniejszej książce,

2. Potrzeby informacyjne uczestników systemu B+R

2.1. Znaczenie analiz bibliometrycznych dla strategicznego zarządzania nauką

Prowadzenie systematycznych obserwacji badań naukowych i rozwoju technologii pozwoli na udoskonalenie mechanizmów strategicznego zarządzania w sektorze B+R. Umożliwi też zapobieganie negatywnym zjawiskom, gdy przy braku priorytetów oraz rezygnacji z koncentracji zasobów finansowych i ludzkich na wybranych obszarach badań, które mogłyby stanowić potencjalne przewagi krajowych badaczy, doszłoby do chaotycznego doboru tematyki projektów badawczych w poszczególnych instytucjach sektora nauki w oparciu o metodę prób i błędów. Ta negatywna sytuacja mogłaby prowadzić do uzyskiwania przeciętnych rezultatów w bardzo wielu obszarach badań, rozproszenia zasobów, tworzenia osiągnięć badawczo-rozwojowych mających znaczenie wyłącznie na poziomie regionalnym i krajowym oraz tworzenia jedynie słabej oferty współpracy dla biznesu.

Podjęcie systematycznych analiz bibliometrycznych może ułatwić strategiczne zarządzanie sektorem nauki oraz budowę strategii B+R instytucji naukowych i przedsiębiorstw poprzez dostarczenie szczegółowej wiedzy na temat tendencji w zakresie rozwoju badań naukowych i technologii. Takie podejście jest zgodne z modelem *evidence-based policy*, opierającym się na wykorzystywaniu wyników analiz i badań naukowych w procesach podejmowania decyzji politycznych, w szczególności decyzji dotyczących alokacji środków. Gromadzone informacje pozwolą na koncentrację na obszarach, w których polscy naukowcy okażą się konkurencyjni w skali światowej, a które jednocześnie są perspektywiczne dla dalszego rozwoju nauki i technologii. Obserwacja tendencji w nauce ułatwi priorytetyzację i zwiększanie nakładów na preferowane obszary badawcze. Takie działania pozwolą na promowanie osiągnięć naukowych, które są lub mają szansę być znaczącymi w skali międzynarodowej oraz stymulowanie udziału polskich badaczy w globalnym dyskursie naukowym. W dłuższym okresie czasu taka sytuacja może doprowadzić do powstawania wybitnych rezultatów w niewielkiej liczbie obszarów badawczych, przy równoczesnym umożliwieniu rozwoju innych badań, nie stanowiących kierunków priorytetowych. Świadomy wybór i koncentracja w nauce pozwolą też na skonkretyzowanie oferty środowisk naukowych dla świata biznesu, zwiększając możliwości finansowania badań naukowych ze źródeł pozabudżetowych. Takie podejście nawiązuje do koncepcji inteligentnej specjalizacji w polityce naukowej i innowacyjnej (Kardas 2010).

2.2. Uczestnicy systemu innowacji i ich potrzeby informacyjne

Systematyczna obserwacja badań naukowych i rozwoju technologii jest przydatna dla różnych grup podmiotów w narodowym systemie innowacji. Do najważniejszych z nich należą:

- resort nauki (w tym Ministerstwo Nauki i Szkolnictwa Wyższego, Narodowe Centrum Badań i Rozwoju, Narodowe Centrum Nauki, Komitet Ewaluacji Jednostek Naukowych i Ośrodek Przetwarzania Informacji),
- uczelnie wyższe, instytuty badawcze i jednostki naukowe Polskiej Akademii Nauk,

- przedsiębiorstwa technologiczne.

Poniżej zaprezentowane zostaną możliwe zastosowania informacji, zdobywanych dzięki analizom bibliometrycznym przez najważniejszych interesariuszy, w podziale na strategiczny, taktyczny i operacyjny poziom podejmowanych działań.

Zestawienie obejmuje liczne nieświadomione potrzeby informacyjne. Ich nieświadomość wynika z braku wiedzy na temat dostępnych technik analitycznych oraz ich możliwych zastosowań w praktyce zarządzania działalnością badawczo-rozwojową. Przypadki nieświadomych potrzeb uczestników polskiego systemu innowacji zaznaczono w poniższym zestawieniu symbolem {N}. Przemyślenia osób zarządzających poszczególnymi instytucjami, dotyczące możliwości wykorzystania bibliometrii w tych dotychczas nieświadomych scenariuszach, mogą stać się interesującą okazją dla doskonalenia ich działań.

Interesariusz: resort nauki (w tym Ministerstwo Nauki i Szkolnictwa Wyższego, Narodowe Centrum Badań i Rozwoju, Narodowe Centrum Nauki, Komitet Ewaluacji Jednostek Naukowych i Ośrodek Przetwarzania Informacji)

Poziom strategiczny

- planowanie strategiczne – w tym określanie priorytetów dla badań naukowych i prac rozwojowych, kierunków inwestycji oraz wydatkowania funduszy krajowych i międzynarodowych,
- *benchmarking* i identyfikacja przewag i słabości polskich osiągnięć badawczych w porównaniu do analogicznych osiągnięć innych krajów,
- identyfikacja wiodących ośrodków naukowych w kraju i za granicą, zwłaszcza w podziale na wąskie specjalizacje badawcze.

Poziom taktyczny

- ocena dokonań polskich naukowców pod względem zgodności z trendami w obszarze badań naukowych i rozwoju technologii oraz porównania międzynarodowej tematyki i osiągnięć,
- stymulowanie inicjatyw badawczych, podejmowanych przez podmioty uczestniczące w systemie B+R, w szczególności wskazywanie globalnych tendencji rozwojowych oraz pojawiających się nisz badawczych,
- stymulowanie współpracy naukowców z przedsiębiorcami oraz partnerstwa instytucjonalnego z firmami {N},
- wsparcie informacyjne przy podejmowaniu decyzji o zaangażowaniu lub rezygnacji z zaangażowania polskiej nauki w specyficzne przedsięwzięcia międzynarodowe {N},
- stymulowanie współpracy międzynarodowej z udziałem instytucji naukowych – przy świadomym wyborze określonych obszarów badań, krajów partnerskich i najbardziej adekwatnych zagranicznych kooperantów {N},
- ułatwienie udziału polskich zespołów badawczych w 7. Programie Ramowym w Zakresie Badań i Rozwoju Technologicznego poprzez identyfikację potencjalnych uczestników konsorcjów projektowych {N},

- wsparcie informacyjne w procesach wydatkowania środków budżetowych i pozabudżetowych – m.in. w programach stypendialnych i konkursach o granty badawcze {N},
- dostarczanie danych wejściowych dla procesów oceny dorobku przy awansach naukowych, procesów ewaluacji jednostek naukowych oraz wyłanianiu Krajowych Naukowych Ośrodków Wiodących (KNOW).

Interesariusze: inne instytucje rządowe (w tym: Ministerstwo Edukacji Narodowej, Ministerstwo Finansów, Ministerstwo Gospodarki, Ministerstwo Infrastruktury, Ministerstwo Rozwoju Regionalnego, Ministerstwo Administracji i Cyfryzacji, Ministerstwo Środowiska, Ministerstwo Zdrowia, Główny Urząd Statystyczny, Polska Agencja Rozwoju Przedsiębiorczości, Polska Agencja Informacji i Inwestycji Zagranicznych, Urząd Patentowy RP)

- wsparcie informacyjne dla programowania wykorzystania środków krajowych i międzynarodowych {N},
- ukierunkowanie działań, mających na celu wspieranie innowacyjności {N},
- zaspokajanie doraźnych potrzeb informacyjnych (dotyczące m.in. procedur zamówień publicznych, porównywania i oceny technologii, promowania polskich osiągnięć na forum międzynarodowym) {N}.

Interesariusze: uczelnie wyższe, instytuty badawcze i jednostki naukowe Polskiej Akademii Nauk

Poziom strategiczny

- planowanie strategiczne – w tym określanie priorytetów dla badań naukowych i prac rozwojowych, kierunków inwestycji, wydatkowania funduszy własnych oraz ubiegania się o dofinansowanie {N},
- *benchmarking* w stosunku do wybranych instytucji systemu B+R oraz określenie własnych przewag i słabości w badaniach naukowych dla poszczególnych obszarów {N},
- identyfikacja międzynarodowych tendencji badawczych, w tym popularności tematyki badań, nowych obszarów o rosnącym znaczeniu dla światowej nauki (ang. *emerging technologies*), obszarów o szczególnie intensywnej konkurencji pomiędzy badaczami jak również potencjalnych nisz badawczych {N},
- systematyczna identyfikacja wiodących ośrodków naukowych w kraju i za granicą, zwłaszcza w podziale na wąskie specjalizacje badawcze {N}.

Poziom taktyczny

- ocena dokonań pracowników pod względem innowacyjności i zgodności z trendami w obszarze badań naukowych i rozwoju technologii oraz porównania międzynarodowe tematyki i osiągnięć,
- stymulowanie inicjatyw badawczych, podejmowanych przez pracowników, w szczególności wskazywanie globalnych tendencji rozwojowych oraz pojawiających się nisz badawczych {N},

- stymulowanie współpracy badaczy z przedsiębiorcami oraz partnerstwa instytucjonalnego z firmami poprzez identyfikację możliwych partnerów biznesowych {N},
- ułatwienie udziału naukowców w programach badań międzyuczelnianych i międzynarodowych poprzez identyfikację możliwych uczestników konsorcjów projektowych, w tym podmiotów prowadzących badania na tematy pokrewne lub komplementarne wobec badań pracowników danej uczelni, instytutu badawczego lub jednostki naukowej Polskiej Akademii Nauk {N},
- wsparcie informacyjne przy podejmowaniu decyzji o zaangażowaniu w przedsięwzięcia międzynarodowe – w tym gromadzenie przesłanek dla oceny prawdopodobieństwa sukcesu przedsięwzięcia z udziałem własnych pracowników oraz zagranicznych partnerów {N},
- dostarczanie danych wejściowych dla wewnętrznych procesów oceny dorobku przy awansach naukowych, przygotowania do ewaluacji jednostek naukowych oraz ubiegania się o status Krajowego Naukowego Ośrodka Wiodącego (KNOW),
- ocena okresowa pracowników naukowych, zespołów badawczych oraz jednostek organizacyjnych, w tym identyfikacja wyróżniających się autorów badań,
- wsparcie informacyjne w procesach podziału środków na badania, przyznawania stypendiów oraz nagród dla naukowców {N}.

Poziom operacyjny

- identyfikacja wiodących publikacji naukowych, autorów tekstów naukowych oraz wynalazców,
- wsparcie informacyjne dla przygotowywania przeglądów literatury naukowej,
- pomoc w ograniczaniu podejmowania zbędnych wysiłków badawczych poprzez powielanie prac podjętych wcześniej przez innych naukowców,
- systematyczna identyfikacja potencjalnych nisz badawczych w ramach analizowanego obszaru badań {N},
- systematyczna identyfikacja nowych tendencji badawczych i dynamicznie rozwijających się obszarów technologicznych (ang. *emerging technologies*) {N},
- ocena potencjału ewolucyjnego analizowanej technologii i przewidywanie prawdopodobnych kierunków dalszego rozwoju (zgodnie z metodami, wypracowanymi przez TRIZ, por. odpowiedni podrozdział książki) {N},
- kreatywna synteza dorobku różnych obszarów badań w celu doskonalenia podejść badawczych, metod i rezultatów i stymulowania innowacyjności w badaniach (zgodnie z metodą *Literature-Based Discovery*, por. odpowiedni podrozdział książki) {N},
- systematyczna identyfikacja najbardziej adekwatnych kierunków wyjazdów badawczych i stypendialnych dla określonej tematyki badawczej {N},
- systematyczna identyfikacja najbardziej odpowiednich konferencji i czasopism naukowych, zajmujących się określoną tematyką badawczą {N},
- wsparcie informacyjne dla procesów poszukiwania możliwych sponsorów projektów badawczych, potencjalnie zainteresowanych komercjalizacją ich rezultatów {N},
- dostarczanie danych przydatnych przy analizie czystości patentowej, ocenie innowacyjności wynalazków i ich porównaniu z aktualnym stanem techniki {N},

- poszukiwanie nowych pracowników, zajmujących się pożądanymi obszarami badań, wśród badaczy polskich i zagranicznych {N},
- ocena kandydatów do pracy na stanowiskach badawczych w oparciu o ich wcześniejszy dorobek {N},
- planowanie sukcesji w ramach zespołu badawczego i jednostki organizacyjnej {N}.

Interesariusze: przedsiębiorstwa technologiczne

Poziom strategiczny

- planowanie strategiczne – w tym poszukiwanie możliwych kierunków rozwoju technologii w oparciu o dorobek nauki i dotychczasowe osiągnięcia wynalazcze,
- identyfikacja międzynarodowych tendencji badawczych, w tym popularności tematyki badań, nowych obszarów o rosnącym znaczeniu dla światowej nauki (ang. *emerging technologies*), obszarów o szczególnie intensywnej konkurencji pomiędzy badaczami jak również potencjalnych nisz badawczych {N},
- systematyczna identyfikacja wiodących twórców określonych technologii w kraju i za granicą {N}.

Poziom taktyczny

- ocena rozwijanych technologii pod względem nowoczesności i zgodności z tendencjami międzynarodowymi,
- stymulowanie projektów badawczo-rozwojowych w nawiązaniu do tendencji międzynarodowych oraz pojawiających się nisz badawczych,
- stymulowanie współpracy i partnerstwa instytucjonalnego z zespołami badawczymi z uczelni oraz jednostek badawczo-rozwojowych poprzez identyfikację możliwych partnerów i obszarów współpracy {N},
- ułatwienie udziału w programach badań krajowych i międzynarodowych poprzez identyfikację możliwych uczestników konsorcjów projektowych, w tym podmiotów prowadzących badania na tematy pokrewne lub komplementarne wobec badań danego przedsiębiorstwa {N}.

Poziom operacyjny

- identyfikacja wiodących osiągnięć naukowych i technologicznych, znaczących badaczy oraz wynalazców,
- pomoc w ograniczaniu podejmowania zbędnych wysiłków badawczych poprzez powielanie prac podjętych wcześniej przez innych naukowców i wynalazców,
- systematyczna identyfikacja potencjalnych nisz w ramach analizowanego obszaru technologicznego {N},
- systematyczna identyfikacja nowych tendencji badawczych i dynamicznie rozwijających się obszarów technologicznych (ang. *emerging technologies*) {N},
- systematyczna analiza konkurentów w zakresie działalności badawczo-rozwojowej oraz ich publikowanych osiągnięć {N},

- ocena potencjału ewolucyjnego analizowanej technologii i przewidywanie prawdopodobnych kierunków dalszego rozwoju (zgodnie z metodami, wypracowanymi przez *TRIZ*, por. odpowiedni podrozdział książki) {N},
- kreatywna synteza dorobku różnych obszarów badań w celu poszukiwania nowych rozwiązań technologicznych i stymulowania innowacyjności w procesach badawczo-rozwojowych (zgodnie z metodą *Literature-Based Discovery*, por. odpowiedni podrozdział książki) {N},
- wsparcie informacyjne dla procesów poszukiwania możliwych podwykonawców projektów badawczo-rozwojowych, w szczególności zespołów badawczych z uczelni wyższych i jednostek badawczo-rozwojowych {N},
- dostarczanie danych przydatnych przy analizie czystości patentowej, ocenie innowacyjności wynalazków i ich porównaniu z aktualnym stanem techniki {N},
- poszukiwanie nowych pracowników do działów badawczo-rozwojowych, zajmujących się pożądanymi obszarami badań, wśród badaczy i wynalazców polskich i zagranicznych {N},
- ocena kandydatów do pracy w działach badawczo-rozwojowych w oparciu o ich wcześniejszy dorobek {N}.

2.3. Kierunki analiz bibliometrycznych w odpowiedzi na potrzeby interesariuszy

Zaprezentowane na poprzednich stronach zestawienie prezentuje niezwykle bogaty zbiór deklarowanych i nieuświadomionych potrzeb informacyjnych oraz możliwych zastosowań analiz bibliometrycznych. Próba uogólnienia tych potrzeb i zastosowań pozwala na wyodrębnienie czterech podstawowych kierunków analizy światowych trendów w obszarze badań naukowych i rozwoju technologii:

1. **monitorowanie i skanowanie technologii** – gromadzenie informacji o tym, co było i/lub jest przedmiotem aktywności badawczo-rozwojowej w poszczególnych obszarach technologicznych (monitorowanie oznacza działania nieukierunkowane na konkretne pytania analityczne, skanowanie – ukierunkowane),
2. **wywiad technologiczny** (ang. *technology intelligence*) – zdobywanie danych o tym, kto zajmuje się poszczególnymi technologiami i obszarami badań oraz na jakich aspektach koncentruje wysiłki badawcze,
3. **prognozowanie rozwoju technologii** - w tym analiza trendów w zakresie badań naukowych i prac nad doskonaleniem technologii oraz przewidywanie ewolucji obecnie dostępnych klas rozwiązań,
4. **ocena technologii** (ang. *technology assessment*) – ocena znaczenia indywidualnych osiągnięć oraz pozycji konkurencyjnej badaczy, instytucji, regionów i kraju.

3. Analizy bibliometryczne w świetle dorobku dyscypliny zarządzania technologiami

3.1. Zarządzanie technologiami a analizy bibliometryczne

Analiza trendów w obszarze badań naukowych i rozwoju technologii powinna opierać się na dobrym zrozumieniu specyfiki procesów badawczo-rozwojowych oraz ekonomicznych uwarunkowań osiągnięć badawczych i technologicznych czyli zagadnień, które stanowią przedmiot zainteresowań dyscypliny badawczej zarządzania technologiami. Poniższy podrozdział oferuje krótki przegląd wybranych zagadnień z obszaru zarządzania technologiami, które mogą okazać się niezbędne dla poprawnego zrozumienia rekomendowanych sposobów prowadzenia analiz bibliometrycznych, ich zastosowań oraz ewentualnych ograniczeń.

Wbrew potocznym przekonaniom, wynalazki nie zawsze są wynikiem szczęśliwego zbiegu okoliczności, a działania mające na celu tworzenie innowacji mogą podlegać procesom planowania i koordynacji. Od lat 40-tych XX wieku badacze z różnych krajów prowadzili intensywne poszukiwania zależności, rządzących pracami badawczo-rozwojowymi. Dyskurs naukowy w Stanach Zjednoczonych i Europie Zachodniej koncentrował się na poszukiwaniu ilościowych zależności w rozwoju poszczególnych parametrów rozwiązań technologicznych, które pozwoliłyby na przygotowywanie prognoz za pośrednictwem modelowania ekonometrycznego. W Związku Radzieckim powstała alternatywna perspektywa badawcza *TRIZ* (teoria rozwiązywania zadań wynalazczych, ros. *Теория решения изобретательских задач*), oparta na jakościowych analizach zawartości zgłoszeń patentowych i identyfikacji prawidłowości, rządzących rozwiązywaniem problemów technicznych (Altszuller 1972). Po upadku ZSRR, wielu badaczy zajmujących się *TRIZ*-em zdecydowało się na emigrację, dorobek *TRIZ* był aktywnie wykorzystywany m.in. przez wiodące koncerny elektroniczne w Stanach Zjednoczonych, Japonii i Korei, a zachodni teoretycy zarządzania technologiami dopiero teraz odkrywają zalety tego podejścia. Poniższy podrozdział w syntetyczny sposób zaprezentuje dorobek obu wzajemnie uzupełniających się podejść, wskazując na teoretyczne podstawy dla prezentowanych w dalszych rozdziałach technik analitycznych.

3.2. Modele tworzenia innowacji

Związki pomiędzy nauką a techniką (badaniami naukowymi a rozwojem technologii) w wielu przypadkach nie są jednoznaczne, a badacze w obszarach ekonomii i zarządzania technologiami od lat interesują się modelowym procesem powstawania nowych rozwiązań (por. Coombs, Saviotti, Walsh 1987: 96-103). Klasyczny model „*technology push*” (dosł. pchania przez technologie) odnosi się do sytuacji, gdy przełomowe odkrycia naukowe są bezpośrednią podstawą do tworzenia przydatnych komercyjnie rozwiązań technicznych, a działające w tym modelu firmy określa się jako firmy opierające się na badaniach naukowych (ang. *science-based firms*). Są to przypadki rzadkie, ograniczone do firm w kilku obszarach technologicznych, m.in. farmacji, biotechnologii i zaawansowanych materiałów. Wyniki pojedynczego projektu badawczego mogą stać się podstawą do założenia firmy i stworzenia obiecującego produktu.

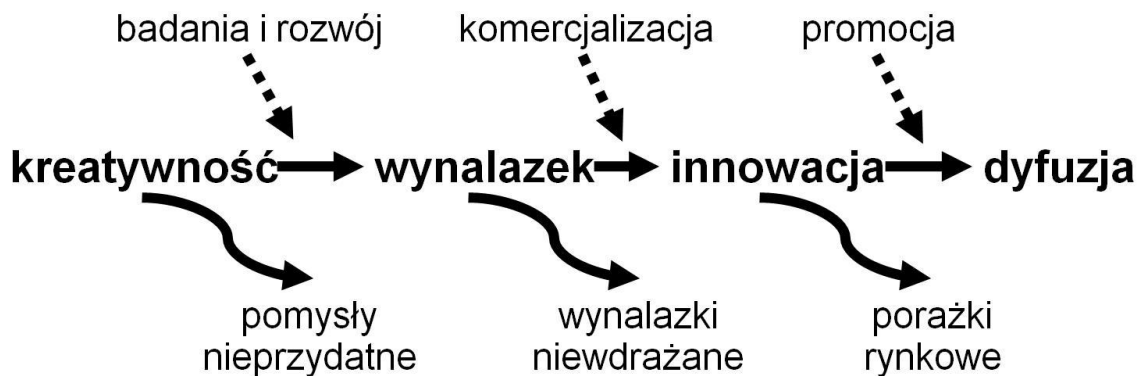
Częściej zaobserwować można sytuacje, gdy działalność badawczo-rozwojowa, mająca na celu udoskonalenie produktu, podejmowana jest w odpowiedzi na sprecyzowane oczekiwania klientów lub świadomość niedoskonałości własnych produktów. Taki model określany jest mianem „*market pull*” (dosł. ciągnięcia przez rynek). W tym przypadku, badania naukowe i prace rozwojowe są zamawiane u zewnętrznych wykonawców lub zlecane pracownikom firmy dopiero po określeniu potrzeb rynkowych, co stanowi odwrócenie wcześniej opisywanego modelu „*science-based firms*”.

Współcześnie większość firm zaawansowanych technologii działa w modelach mieszanych, których odmiany określane są jako modele łańcuchowe lub sieciowe: w zależności od nadarzających się okazji biznesowych i technologicznych, w niektórych przypadkach osiągnięcia badań naukowych stają się inspiracją dla rozwoju lub doskonalenia produktów, a w innych sytuacjach to oczekiwania klientów i zdiagnozowane niedoskonałości oferty rynkowej motywują do podejmowania projektów badawczo-rozwojowych. Warto pamiętać, że zależności pomiędzy osiągnięciami badawczymi a korzyściami ekonomicznymi z komercjalizacji nie są jednoznaczne, a możliwości bezpośredniego zastosowania wiedzy naukowej w rozwoju technologii są zwykle ograniczone.

Wiedza o wynikach badań naukowych przydaje się jednak również tym podmiotom, które samodzielnie nie prowadzą analogicznych badań lub prac nad rozwojem technologii i nie znajdują bezpośrednich zastosowań praktycznych dla poznawanych rezultatów badawczych. Teoretycy zarządzania technologiami zidentyfikowali zjawisko zdolności absorpcyjnych (ang. *absorptive capacities*), pozwalających firmom na lepsze przyswajanie rozwiązań technicznych i sprawniejsze wykorzystywanie technologii dzięki pogłębianiu wiedzy na ich temat (Cohen, Levinthal 1990: 128). Obserwacja badań naukowych i rozwoju technologii może pomóc w lepszym planowaniu inwestycji technologicznych, dokonywaniu trafnych zakupów i ocenie dostawców, czy wreszcie udanym wdrażaniu zakupionych rozwiązań.

Dodatkowo, poznawanie obszarów nauki i techniki, w których nie prowadzi się bezpośrednio własnych badań lub prac rozwojowych, może stać się źródłem cennych inspiracji wynalazczych i oferować analogie dla rozwiązania pokrewnych problemów naukowych (takie podejście wspierają techniki oparte na *TRIZ* i *LBD*, które zostaną opisane w dalszej części książki).

Rysunek 1 prezentuje modelowy proces komercjalizacji pomysłów. Obejmuje on przechodzenie od kreatywności, będącej cechą indywidualną, do generowania pomysłów rozwiązań problemów naukowych i technicznych, w tym wynalazków, a w przypadku decyzji o ich wdrożeniu – do tworzenia innowacji, które następnie podlegają procesom dyfuzji-rozpowszechniania na rynku. W praktyce badawczej, nie każdy pomysł prowadzi do powstania wynalazku, nie każdy wynalazek jest wdrażany czy przekształcany w innowacyjny produkt, nie każda wreszcie innowacja odnosi sukces rynkowy.



Rysunek 1: Proces komercjalizacji pomysłów. Źródło: Klineciewicz (2010: 27).

Wiedza dostarczana w wyniku analiz bibliometrycznych może być przydatna na różnych etapach omawianego procesu. Możliwe jest wykorzystanie informacji o osiągnięciach naukowych i technologicznych dla wspierania kreatywności, w tym poszukiwania oryginalnych tematów badań i nisz badawczych. Na etapie generowania odkryć lub wynalazków, bibliometria może wspierać badaczy poprzez analizę dorobku innych osób oraz problemów, jakie napotykali we własnej pracy B+R oraz dostęp do baz danych, odzwierciedlających aktualny stan techniki. Przydatne mogą też okazać się specyficzne techniki kreatywne, znane obecnie tylko nielicznym polskim naukowcom i przedsiębiorstwom - w tym *TRIZ*, stymulujący tworzenie wynalazków oraz *LBD*, wspierające odkrycia naukowe. Specyficzną odmianą wynalazczości jest generowanie „kreatywnych imitacji” (Klineciewicz 2010: 31-32). Na rynkach zaawansowanych technologii dążenie do wytworzenia imitacji niekoniecznie oznacza bezpośrednie kopiowanie, naruszenia praw własności przemysłowej czy kradzież cudzych rozwiązań. Częściej chodzi o czerpanie inspiracji z cudzych pomysłów i takie ukierunkowanie własnych prac badawczo-rozwojowych, by stworzyć odpowiednik cudzego wynalazku, zawierający dodatkowe ulepszenia i nie naruszający patentów. W takich przypadkach imitatorzy wykorzystują techniki, wspierające generowanie wynalazków, mogą więc uznać analizy bibliometryczne za przydatną w swojej działalności. Na etapie prac rozwojowych, mających na celu opracowanie nadającego się do komercjalizacji produktu lub innej formy rozwiązania technicznego, a następnie wprowadzenie go na rynek, przydatna będzie znajomość tendencji rynkowych oraz wypracowanych przez innych badaczy sposobów rozwiązywania praktycznych problemów technicznych, występujących na etapie przygotowywania produktu (często związanych z dorobkiem innych dyscyplin nauki i techniki niż wynalazek, stanowiący podstawę produktu). Prezentowane w książce techniki analityczne mogą więc towarzyszyć badaczom w kolejnych krokach procesu komercjalizacji pomysłów.

3.3. Rezultaty działalności badawczo-rozwojowej

Obserwacja badań naukowych i rozwoju technologii dotyczy zwykle rezultatów prac B+R oraz innych przejawów działalności innowacyjnej, do których należą przede wszystkim:

- wnioski o dofinansowanie projektów badawczo-rozwojowych,

- raporty z realizacji projektów badawczo-rozwojowych,
- publikacje naukowe,
- wystąpienia konferencyjne,
- wywiady udzielane dziennikarzom,
- informacje prasowe,
- zgłoszenia patentowe,
- udzielone patenty,
- produkty,
- dokumentacja produktów,
- materiały promocyjne o produktach,
- raporty analityczne o rynkach technologicznych,
- informacje publikowane na prywatnych stronach internetowych, blogach i serwisach społecznościowych,
- wnioski do komitetów standaryzacyjnych,
- dokumenty sądowe, dotyczące naruszeń patentów, konfliktów pomiędzy partnerami w projektach badawczo-rozwojowych lub zgłoszeń do organów regulacyjnych.

System informacyjny może gromadzić informacje z wszystkich wymienionych powyżej źródeł, tylko niektóre z nich pozwalają jednak na systematyczne obserwacje, regularnie dostarczając nowych dokumentów, relatywnie łatwo dostępnych i poddających się analizom (publikacje naukowe, wystąpienia konferencyjne, zgłoszenia patentowe i udzielone patenty).

Nie wszystkie rezultaty projektów badawczo-rozwojowych można poznać z pozycji zewnętrznego obserwatora. Liczne projekty prowadzone przez firmy bywają utajnione, osiągane wyniki nie stają się przedmiotem publikacji naukowych, a czasami nie są też patentowane. Motywacją do takiego zaniechania jest fakt, iż ochrona patentowa jest ograniczona w czasie, a po jej wygaśnięciu wszyscy uczestnicy rynku uzyskują możliwość wykorzystania szczegółowego opisu technicznego opatentowanego rozwiązania. Z tego powodu firmy rezygnują z patentowania w odniesieniu do niektórych wynalazków, dbając o ich utajnienie i uniemożliwienie konkurentom poznania unikatowych sposobów rozwiązywania problemów technicznych, wypracowanych przez dane przedsiębiorstwo.

Wobec opisanych powyżej ograniczeń, dogodnym sposobem analiz zasobów informacyjnych, dotyczących badań naukowych i rozwoju technologii, okazuje się bibliometria, obejmująca ilościowe analizy kolekcji publikacji naukowych i baz patentowych. W dalszych częściach książki opisane zostaną możliwości zastosowania, uwarunkowania, korzyści oraz praktyczne sposoby prowadzenia analiz bibliometrycznych. Ich cechą charakterystyczną jest ilościowe podejście do danych, pobranych z baz publikacji i patentów, pozwalające na przetwarzanie nawet bardzo obszernych zbiorów danych i identyfikację tendencji badawczych lub rynkowych, co zwykle nie jest możliwe w przypadkach samodzielnej lektury i analizy aspektów technicznych publikacji lub zgłoszeń patentowych.

Alternatywą wobec analiz bibliometrycznych jest odwołanie się do wiedzy i doświadczeń ekspertów oraz informacji zbieranych doraźnie z różnorodnych źródeł. Większość analiz

eksperymentalnych posiada jednak istotne ograniczenia, związane z subiektywnym charakterem, oparciem na indywidualnych zasobach wiedzy i na zainteresowaniach ekspertów.

Badania naukowe i działania mające na celu rozwój technologii charakteryzować może zróżnicowany poziom nowatorstwa. Niektóre prowadzą do **racjonalizacji** (drobnej modyfikacji stosowanych metod), inne – do **modernizacji** (bardziej złożonych udoskonaleń). Doniosłość prac badawczo-rozwojowych wzrasta wraz z udanym określeniem **nowej zasady** (zwykle w odniesieniu do problemów naukowych i praktycznych, które wcześniej nie były odpowiednio ustrukturalizowane). Jeszcze bardziej ambitnym osiągnięciem będzie **synteza** – kreatywne połączenie już istniejącej wiedzy z różnych, wcześniej odrębnych obszarów, w celu wprowadzenia znaczącej zmiany. Najbardziej ambitnym rezultatem prac B+R będzie **odkrycie** – stworzenie wynalazku lub nowej, oryginalnej wiedzy naukowej lub technicznej (Orloff 2006: 17). Przypadki racjonalizacji i modernizacji nie są zwykle opisywane w publikacjach naukowych, gdyż ich autorzy nie sądzą, że tego typu osiągnięcia wpisują się w dyskurs naukowy – często znajdują jednak odzwierciedlenie w zgłoszeniach patentowych, zwłaszcza wtedy, gdy stanowią sposób udoskonalenia właściwości produktów lub procesów produkcyjnych.

3.4. Rodzaje działalności badawczo-rozwojowej

Teoretycy i praktycy zarządzania technologiami często nawiązują do podziału działalności badawczo-rozwojowej na trzy rodzaje (OECD 2002: 30):

- **badania podstawowe** – ukierunkowane na rozwiązywanie fundamentalnych problemów naukowych i pogłębianie podstawowej wiedzy teoretycznej w ramach określonej dyscypliny,
- **badania stosowane** – wykorzystujące dorobek badań podstawowych do rozwiązywania konkretnych problemów praktycznych i technicznych, zwykle prowadzące do rezultatów mających potencjalne zastosowania komercyjne,
- **prace rozwojowe** – oparte na wynikach badań podstawowych i stosowanych oraz doświadczeniach praktycznych, prowadzące do tworzenia rozwiązań technicznych oraz doskonalenia rozwiązań istniejących.

Popularny podział na badania na podstawowe i stosowane oraz prace rozwojowe okazuje się trudny do odzwierciedlenia w przypadku analiz, polegających na obserwacji wymiernych rezultatów prac badawczo-rozwojowych. Watts i Porter (1997) sugerowali, że dane zgromadzone w poszczególnych bazach publikacji mogą odzwierciedlać określone rodzaje aktywności badawczo-rozwojowej - np. badania podstawowe miałyby znaleźć odzwierciedlenie w bazie *Web of Science*, a badania stosowane - w bazie *Compendex*. Aktualne analizy empiryczne nie potwierdzają jednak tych założeń, a bazy okazują się dokumentować badania różnego rodzaju, uniemożliwiając jednoznaczne rozróżnienie na wyniki badań podstawowych, stosowanych i prac rozwojowych (Järvenpää, Mäkinen, Seppänen 2011).

Z perspektywy analiz bibliometrycznych, bardziej adekwatny wydaje się podział na:

- **wyniki badań ogólnodostępne** - należące do domeny publicznej (ang. *public domain*), w tym dostępne dla wszystkich zainteresowanych publikacje naukowe (czasami dostęp do nich wymaga oczywiście nabycia czasopisma, monografii lub subskrypcji do elektronicznej kolekcji publikacji, jednak może to zrobić każda osoba zainteresowana) oraz zgłoszenia patentowe, które nie doprowadziły do udzielenia patentów (zwykle dotyczą one rozwiązań, które nie były dostatecznie oryginalne, jednak zawarte w nich opisy mogą zawierać cenne wskazówki dla badaczy i technologów),
- **wyniki badań odtajnione i jednocześnie chronione** – przede wszystkim wynalazki objęte ochroną patentową, stanowiące zasób wiedzy technicznej dostępny dla wszystkich zainteresowanych i mogący stymulować i ukierunkowywać dalsze prace badawcze (choć samej istoty opatentowanego wynalazku nie można bezpośrednio powielić w okresie obowiązywania ochrony patentowej),
- **wyniki badań nieujawnione** – nie zostały opisane w publikacjach naukowych ani zgłoszeniach patentowych, przejawiają się we wprowadzanych na rynek produktach i metodach produkcji, można je poznać za pośrednictwem technik szpiegostwa przemysłowego i odwrotnej inżynierii produktów (w większości przypadków te działania są nielegalne).

Można podjąć próbę powiązania typologii prac B+R z powyższym zestawieniem możliwych wyników badań i uznać, że w wielu przypadkach występują poniższe zależności:

- wyniki badań ogólnodostępne są zwykle przejawami badań podstawowych, rzadziej badań stosowanych (niektóre zorientowane praktycznie publikacje oraz zgłoszenia patentowe, które nie doprowadziły do udzielenia patentów),
- wyniki badań odtajnione i jednocześnie chronione to rezultaty badań stosowanych lub prac rozwojowych,
- wyniki badań nieujawnione to najczęściej przejawy prac rozwojowych.

	Badania podstawowe	Badania stosowane	Prace rozwojowe
Wyniki badań ogólnodostępne	++	+	
Wyniki badań odtajnione i chronione		++	++
Wyniki badań nieujawnione		+	++

Tabela 1: Zależność pomiędzy badaniami podstawowymi, stosowanymi i pracami rozwojowymi a rodzajami wyników badań. Źródło: opracowanie własne.

Analizy publikacji i patentów pozwalają określić, jakie koncepcje teoretyczne, wyniki badań, wynalazki czy modele stanowią podstawy dla kolejnych osiągnięć naukowych i wynalazczych. Identyfikację takich zależności wspierają techniki analizy cytowań –

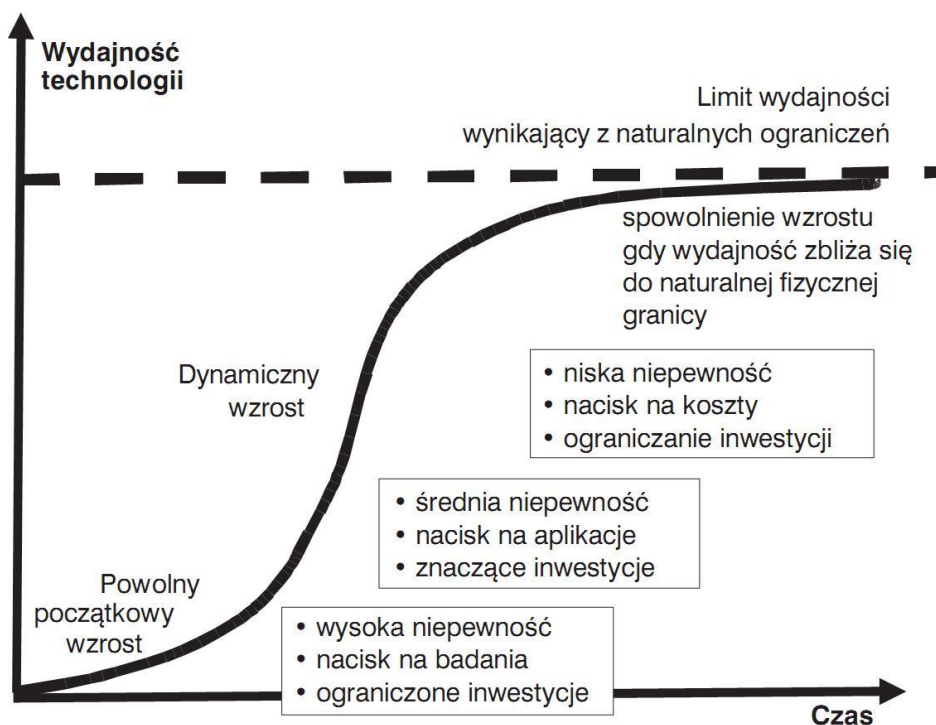
przeglądu źródeł bibliograficznych (artykułów, wystąpień konferencyjnych i patentów), cytowanych w konkretnym artykule naukowych lub zgłoszeniu patentowym. Osiągnięcia szczególnie często cytowane są zwykle interpretowane jako posiadające większą doniosłość naukową i znaczenie dla rozwoju technologii.

Współcześnie coraz częściej obserwowane są zjawiska fuzji technologii – zlewania się wcześniej niezależnych dyscyplin i obszarów technologicznych (Kodama 1992). W taki sposób powstały nowe dyscypliny badawcze, np. biotechnologie czy nanotechnologie, jak również wąskie, interdyscyplinarne specjalności naukowe. Występowanie fuzji technologii stanowi dla badaczy dodatkową motywację do zdobywania wiedzy z obszarów, które wcześniej nie były przedmiotem ich zainteresowań naukowych. Bibliometria pozwoli na identyfikację zlewających się obszarów oraz wiedzy, niezbędnej dla podejmowania badań w nowych podobszarach nauki i techniki.

3.5. Rozwój technologii

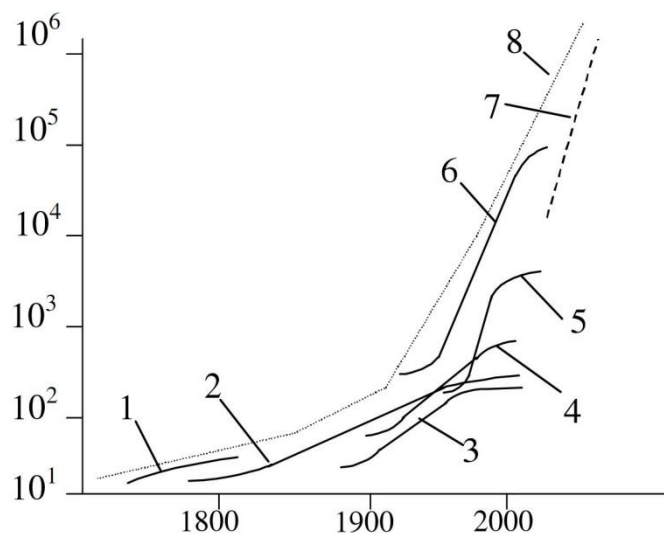
Obserwacja technologii dotyczyć może nie tylko samego faktu stworzenia określonej technologii, ale też jej dalszego doskonalenia. Rozwój każdej technologii można opisać jako systematyczne doskonalenie wybranego parametru, charakteryzującego określony aspekt wydajności technologii.

Przykładowo, stopniowej poprawie ulegać mogą: prędkość, wydajność, pojemność, transfer danych, czas pracy na baterii, liczba operacji na minutę, liczba komponentów przypadających na 1 cm² powierzchni, zużycie paliwa na 100 km, bezpieczeństwo lub niezawodność. Zależność pomiędzy poziomem tego parametru a czasem (lub wielkością poniesionych nakładów na prace badawczo-rozwojowe) zwykle odpowiada krzywej S-kształtnej. Krzywa S prezentuje początkowo powolny wzrost, po nim dynamiczny rozwój, aż do zahamowania w momencie zbliżenia się do granicy fizycznych możliwości technologii. Każdy punkt krzywej odpowiadać będzie wprowadzonej innowacji, która pozwala udoskonalić technologię. Rysunek 2 przedstawia model krzywej S-kształtnej (nazywany też modelem cyklu życia technologii), który odpowiada idealnemu wzorcowi rozwojowi i częściej wykorzystywany jest do analizy danych historycznych niż do prognozowania przyszłego rozwoju nowej technologii, wskazuje jednak na oczywiste zależności i uświadamia istnienie ograniczeń technologicznych.



Rysunek 2: Model rozwoju technologii w oparciu o krzywą S-kształtną. Źródło: Klincewicz (2010: 41).

Firma doradcza *McKinsey* wykorzystała powyższy model do opisu mechanizmów substytucji – zastępowania dotychczasowych rozwiązań przez kolejne generacje technologii. W okresie, gdy dojrzała technologia osiąga szczyt swoich możliwości, rozpoczyna się rozwój technologii alternatywnej, początkowo mniej doskonałej od poprzedniczki, posiadającej jednak znaczący potencjał rozwoju. Model *McKinsey*'a uzmysławia korzyści, związane z obserwacją rozwoju nauki i technologii – pozwala w odpowiednim momencie podjąć decyzję o podjęciu prac nad alternatywą wobec starzejącej się technologii tak, by przygotować się na jej zastąpienie rozwiązaniami nowej generacji. Punkt rozpoczęcia nowej krzywej S-kształtnej odpowiada wprowadzeniu przełomowej innowacji, która uruchamia cykl dalszego doskonalenia technologii. Najprostszym sposobem identyfikacji takich przypadków jest systematyczna obserwacja osiągnięć naukowych i zgłoszeń patentowych. Rysunek 3 prezentuje uproszczone wyniki długookresowej obserwacji substytucji kolejnych generacji technologii transportowych.



Rysunek 3: Doskonalenie systemów transportowych poprzez substytucję kolejnych technologii. Źródło: Orloff (2006: 200). Legenda: (1) wóz konny, (2) kolej, (3) samochód, (4) samolot śmigłowy, (5) samolot odrzutowy, (6) rakietą z paliwem chemicznym, (7) rakietą z paliwem nuklearnym, (8) krzywa sumaryczna.

Specyficznym przypadkiem nowych rozwiązań są technologie zaburzające (ang. *disruptive technologies*) (Christensen 2000). Charakteryzują się one początkowo wyjątkowo niekorzystnym podstawowym parametrem technicznym, który interesuje głównych odbiorców rozwiązań technologicznych – może być to przykładowo wspomniana już prędkość, wydajność lub pojemność. Ta niedoskonałość prowadzi do ograniczonego zainteresowania nową technologią przez dotychczasowych graczy rynkowych oraz ich klientów, którzy nie są skłonni porzucić sprawdzonych produktów, zaspokajających ich oczekiwania. Technologie zaburzające charakteryzują jednak jeszcze dwie istotne cechy. Po pierwsze, wspomniany parametr techniczny może być dalej doskonalony i należy spodziewać się, że w ciągu kilku lat osiągnie wartości przewyższające charakterystyki dotychczasowej technologii, a co za tym idzie, będzie odpowiadał oczekiwaniom masowego nabywcy. Po drugie, technologia zaburzająca oferuje dodatkowe korzyści, których nie posiada poprzedni wariant technologii, a które mogą zainteresować nowe grupy klientów. Przykładowo, częstym scenariuszem jest wzbogacanie tradycyjnych rozwiązań technologicznych o element mobilności – możliwość przenoszenia urządzeń, pracy w dowolnym miejscu, także bez stałego zasilania z sieci elektrycznej (np. odtwarzacze MP3, przenośne komputery, telefonia internetowa, wykorzystanie telefonu komórkowego do korespondencji e-mail, fotografia cyfrowa). W odniesieniu do technologii produkcyjnych, znaczące udoskonalenie może dotyczyć większej łatwości stosowania - radykalnie niższych kosztów, wymaganych temperatur, ciśnień, dodatkowych substancji, prostoty montażu. Opisana kombinacja właściwości doprowadza zwykle do znaczącego zaburzenia rynku (stąd nazwa: technologia zaburzająca), zmiany preferencji odbiorców i zaskoczenia firm, które skoncentrowały się na rozwoju tradycyjnego wariantu technologii i nie dostrzegały zagrożenia ze strony początkowo niedoskonałej, ale systematycznie ulepszanej alternatywy. Nie każda nowa technologia posiada cechy technologii zaburzającej, jednak ryzyko pojawienia się odkryć naukowych i rozwiązań technologicznych, które mogą wprowadzić istotne zaburzenia rynku, motywuje do systematycznego monitorowania trendów w zakresie badań naukowych i rozwoju technologii.

Szczególnie niebezpieczne są te innowacje, które mogą doprowadzić do zmniejszenia praktycznej przydatności wypracowanej przez firmę wiedzy i umiejętności (ang. *competence-destroying innovations*) (Tushman, Anderson 1986: 442).

Rozwój technologii rządzi się określonymi prawidłowościami, których poszukiwaniem zajmowali się przez kilkadziesiąt lat radzieccy teoretycy, reprezentujący perspektywę badawczą *TRIZ* (por. Clarke 2000; Mann 2002; Orloff 2006; Cavallucci, Rousselot, Zanni 2009; Mizrachi 2010). Ich osiągnięcia są zbliżone do analogicznych analiz, prowadzonych w krajach zachodnich, podejście *TRIZ* wydaje się jednak systemowym ujęciem zagadnień rozwoju technologii, bardziej kompleksowym od fragmentarycznych modeli i teorii, tworzonych na Zachodzie. Jedną z ujawnionych prawidłowości opisuje wzajemne następstwo faz eskalacji i de-eskalacji w rozwoju technologii i znana jest jako „metamodel fal ewolucji” (Orloff 2006: 213). Eskalacja to wzrost liczby i złożoności elementów systemu technicznego, a de-eskalacja stanowi jej odwrotność, wynikającą m.in. z integracji komponentów, miniaturyzacji i uproszczenia obsługi. Analizy historycznego rozwoju poszczególnych technologii wskazują na wielokrotne powielanie cyklu eskalacji i de-eskalacji w odniesieniu do rozwiązań technologicznych w danym obszarze pozwalając na przyjęcie założenia, że naturalną tendencją w rozwoju rozwiązań skomplikowanych będzie ich upraszczanie, a najbardziej prawdopodobnym kierunkiem zmian systemów uproszczonych będzie ponowny wzrost stopnia złożoności, zwykle poprzez integrację dotychczasowej technologii z innymi, wcześniej odrębnymi rozwiązaniami albo poszukiwanie możliwych udoskonaleń parametrów technicznych lub kosztów produkcji poprzez zmiany materiałów i komponentów systemu. Klasyczne modele krzywej S-kształtnej oraz substytucji technologii nie pozwalają niestety na odzwierciedlenie opisywanych zależności, co stanowi istotne ograniczenie ilościowego opisu rozwoju technologii.

W wielu obszarach technologicznych zaobserwować można równoczesny rozwój kilku alternatywnych wariantów danej technologii oraz inwestycje w każdy z nich do czasu, aż nie okaże się, który z wariantów stanie się dominującą formą technologii (ang. *dominant design*), przyjętą przez rynek (Tushman, Anderson 1990). Przebieg procesu akceptowania dominującej formy technologii przez uczestników rynku jest trudny do przewidzenia, a na jego wynik wpływają zwykle nie tylko argumenty techniczne, ale też działania dostawców, w tym promocja produktów, tworzenie koalicji firm, wspierających określony wariant technologiczny i przekonywanie własnych partnerów do porzucenia rozwiązań konkurencyjnych (Klincewicz 2005).

Rozwój alternatywnych i udoskonalonych wariantów technologii nie zawsze oznacza całkowitą wzajemną substytucję. Często zaobserwować można współzystowanie konkurencyjnych rozwiązań technologicznych, służących zaspokajaniu tych samych potrzeb klientów. Przykładowo, użytkowników mogą znaleźć jednocześnie telefony stacjonarne i komórkowe, kino i telewizja, rowery, motory i samochody. W perspektywie badawczej opartej na *TRIZ* ta regularność prezentowana jest jako „prawo przechodzenia w super-system” (Orloff 2006: 204-206).

Ewolucja technologii może być więc interpretowana jako proces społeczny a nie tylko techniczny (Tushman, Anderson 1990). Jej przewidywanie powinno opierać się nie tylko na

założeniu systematycznego doskonalenia istniejących rozwiązań – można przykładowo wyobrazić sobie sytuację, w których uczestnicy rynku wybiorą rozwiązanie gorsze funkcjonalnie i technicznie, które będzie jednak aktywnie promowane przez renomowanego dostawcę. Nakłady na działalność badawczo-rozwojową mogą też okazać się nietrafne, prowadząc firmę w „ślepej uliczce” (ang. *technological lockout*) (Schilling 1998), gdy wybrany przez nią wariant rozwiązania technologicznego zostanie zastąpiony inną formą technologii. Systematyczna obserwacja rozwoju badań naukowych i technologii pozwala więc na zabezpieczenie inwestycji, świadomość prac badawczo-rozwojowych podejmowanych przez konkurentów i odpowiednio wczesne korekty własnych planów. Mniejszym firmom-dostawcom komponentów technologicznych pozwala wreszcie na dopasowanie własnej oferty do oczekiwań odbiorców-dużych graczy na rynku zaawansowanych technologii.

Niezależnie od społecznych uwarunkowań procesów doskonalenia technologii, udało się też zidentyfikować powtarzalne wzorce zmian technicznych. Ich przykładem jest „metamodel kolejności rozwoju zasobów”, stworzony w ramach podejścia *TRIZ* (Orloff 2006: 219-221). W odniesieniu do wielu przykładów empirycznych można zaobserwować następujące po sobie zainteresowanie odmiennymi sposobami udoskonalenia danej technologii, z których każdy czerpał z dorobku innej dyscypliny badawczej i angażował inne grupy badaczy. Terminologia *TRIZ* opisująca te transformacje wydaje się skomplikowana i posiada różne warianty dla poszczególnych rodzajów złożonych systemów technicznych, narzędzi i materiałów, jednak opisywane przez nią zależności okazują się łatwe do zaobserwowania. Przykładowo, narzędzia chirurgiczne ewoluowały od metalowego skalpela poprzez skalpel z wymiennym elementem, wykorzystanie fal o wysokiej częstotliwości, wody pod ciśnieniem i wreszcie lasera, co dokładnie odzwierciedla modelowy schemat *TRIZ* (Orloff 2006: 220). „Metamodel kolejności rozwoju zasobów” nie powinien być postrzegany jako przejaw bezwzględnych praw, rządzących rozwojem technologii – może jednak być wartościową inspiracją dla badaczy, zastanawiających się nad prawdopodobnym, dalszym rozwojem określonej technologii lub zmianami zainteresowań naukowców. Dzięki jego zastosowaniu, można zidentyfikować pokrewne obszary badań naukowych, które już niedługo mogą okazać się przydatne dla prac nad dalszym rozwojem określonej technologii, choć nie jest oczywiście możliwe bezbłędne przewidywanie kolejnych form technologii, które zostaną opracowane przez badaczy (sam *TRIZ* jest najczęściej wykorzystywany jako wsparcie w procesach tworzenia nowych technologii i stymulowania wynalazczości, a nie analizowania i przewidywania działań innych osób czy instytucji, tworzących te technologie).

3.6. Wnioski do wykorzystania w analizach bibliometrycznych

W oparciu o powyższy przegląd wybranych pojęć i modeli z obszaru zarządzania technologiami można wskazać podstawowe obszary, które mogą złożyć się na projekty badawcze i analityczne, oparte na wykorzystaniu bibliometrii:

- przewidywanie rozwoju określonych obszarów badań naukowych i technologii, w szczególności określanie kierunków i potencjału dalszego rozwoju,
- identyfikacja obszarów wiedzy, niezbędnych dla skutecznego prowadzenia badań naukowych i prac rozwojowych w poszczególnych obszarach,

- identyfikacja przypadków fuzji technologii oraz nowych, powstających obszarów badawczych i technologicznych (ang. *emerging technologies*),
- identyfikacja obszarów badawczych i rozwiązań technologicznych, które mogą stopniowo wypierać dotychczas stosowane rozwiązania i ograniczać przydatność zgromadzonej wiedzy (ang. *disruptive technologies, competence-destroying innovations*),
- gromadzenie i udostępnianie wiedzy, dotyczącej tendencji w obszarze badań naukowych i rozwoju technologii, które może służyć podnoszeniu zdolności absorpcyjnych pracowników instytucji naukowych i firm oraz stanowić inspirację do własnych badań i wysiłków wynalazczych.

4. Bibliometria a inne techniki analityczne

4.1. Techniki wspierające obserwację i przewidywanie rozwoju nauki i techniki

Obserwacja i przewidywanie rozwoju nauki i techniki są możliwe przy wykorzystaniu wielu alternatywnych technik. Porter i in. (2004: 290-291) oferują charakterystykę i porównanie aż 51 szczegółowych metod, wykorzystywanych do omawianych celów. Zestawienie możliwych sposobów pomiaru i obserwacji rozwoju nauki, techniki i innowacji zaprezentował Kozłowski (2010). Reger (2001: 535) zwraca uwagę na najważniejsze zadania: analizę technologii, monitorowanie istniejących rozwiązań, prognozowanie tendencji rozwojowych i skanowanie otoczenia rynkowego. Przegląd literatury dotyczącej prognozowania rozwoju technologii oferują Mishra, Deshmukh i Vrat (2002: 3). Klasycznym podręcznikiem metodologicznym w omawianym obszarze jest książka Portera i in. (1991). Decydenci, opracowujący strategie technologiczne, potrzebują informacji wspierających procesy decyzyjne i sugerujących, które obszary badawcze i technologiczne rozwijać, czy dążyć w tych obszarach do przywództwa technologicznego czy pozycji naśladowcy oraz w jakim zakresie korzystać z licencjonowania technologii (Wyk 1997: 22).

Wśród najbardziej popularnych metod, szczególnie często wymieniany jest *foresight* czyli "systematyczny proces budowania wspólnego postrzegania długookresowych trendów" (Héraud, Cuhls 1999: 56). Chodzi o systemowe działania, mające wspomagać "patrzenie w przyszłość" i podejmowanie bardziej efektywnych decyzji i działań (Grupp, Linstone 1999: 86). Praktyczne zastosowanie tej metody oznacza wykorzystanie konsultacji z ekspertami dziedzinowymi, prowadzenie wywiadów indywidualnych, zogniskowanych wywiadów grupowych oraz sesji analitycznych opartych o metodę delfijską (strukturalizująca dyskusję i wspierająca poszukiwanie konsensusu między zróżnicowanymi opiniami ekspertów). *Foresight* dotyczy zwykle długiej perspektywy czasowej i przewidywania rozwoju zjawisk, który obarczony jest istotnym poziomem niepewności. Tak pojmowane projekty typu *foresight* były podejmowane w Stanach Zjednoczonych już w latach 60-tych XX wieku przez *RAND Corporation*, a w Japonii w latach 70-tych (Grupp, Linstone 1999: 88; Kuwahara 1999: 6-7) i do dziś stanowią popularny motyw przewodni inicjatyw rządów wielu krajów.

Niektórzy rozumieją *foresight* inaczej – jako systematyczną identyfikację i obserwację "słabych sygnałów" (ang. *weak signals*) dotyczących nowych technologii lub istniejących rozwiązań (Reger 2001: 535). Takie pojmowanie tej metody jest zbliżone do celów, które pozwalają zrealizować analizy bibliometryczne.

Zestawienia porównawcze poszczególnych metod obserwacji i prognozowania technologii oferują m.in. Porter i in. (1991), Watt i Porter (1997: 26), Mishra, Deshmukh i Vrat (2002: 4) oraz Slocum i Lundberg (2001: 139-143). Można podzielić je na metody jakościowe i ilościowe, odwołujące się do subiektywnej oceny lub zobiektywizowanej interpretacji danych. Subiektywnych ocen dostarczać będą m.in. wywiady z naukowcami, menedżerami firm technologicznych, handlowcami, odbiorcami technologii oraz badania ankietowe uczestników rynku. Do metod bardziej ustrukturalizowanych i pozwalających na intersubiektywność zaliczyć można wspomnianą wcześniej metodę delfijską i inne metody eksperckie, oparte na

dążeniu do konsensusu. Analiza scenariuszowa pozwala na tworzenie alternatywnych wizji przyszłego rozwoju otoczenia (Postma, Liebl 2005: 162). Modelowanie zgodne z podejściem dynamiki systemów opiera się na dogłębnych analizach relacji przyczynowo-skutkowych pomiędzy zjawiskami w otoczeniu. Metoda analogii nawiązuje do innych zdarzeń, zjawisk lub procesów historycznych, fizycznych lub biologicznych (Slocum, Lundberg 2001: 141). Analiza morfologiczna stymuluje kreatywność poprzez dekompozycję problemów i identyfikację nowych wariantów. Na drugim krańcu kontinuum, rozciągającego się pomiędzy intuicją a „twardymi” danymi, znajdują się metody ilościowe, w tym ekstrapolacja trendów, modelowanie regresji czy wielozmiennowa analiza danych. W tych przypadkach wyzwaniem stanowi jednak zdobycie danych, które mogą być podstawą do budowy prognoz. Mogą być to dane ekonomiczne lub bibliometryczne, które muszą być wcześniej zdobyte przy wykorzystaniu odpowiednich technik obserwacyjnych. Niniejsza książka koncentruje się na sposobach zdobywania tych danych, natomiast w ramach projektów analitycznych wykorzystujących zebrane dane bibliometryczne, możliwa będzie również budowa prognoz i wykorzystanie zaawansowanych technik statystycznych.

Jak sugerują Coates i in. (2001: 9), "prognozowanie technologii opartej na badaniach naukowych może przypominać prognozowanie dzisiejszej pogody - długookresowe trendy klimatyczne, alternatywne scenariusze albo panele ekspertów są mniej efektywne niż uzyskanie bogatego, kontekstowego obrazu pogody (np. z telewizyjnej prognozy pogody) i przyjrzenie się najbardziej aktualnym trendom, w tym kierunkom i prędkości przesuwania się frontów pogodowych". Prognozowanie rozwoju nauki i techniki przysparza wielu problemów praktycznych, zwłaszcza gdy oczekuje się przewidywania rozwoju nowych technologii (ang. *emerging technologies*) (Watts, Porter 1997: 27). Analizy dopuszczają pewien poziom niepewności, ale nie może on być nadmiernie wysoki (Postma, Liebl 2005: 165). To oznacza, że nie da się przykładowo przewidzieć przełomowych innowacji lub istotnych zaburzeń technologicznych, bo samo ich przewidzenie oznaczałoby przecież... stworzenie zarysu przewidywanej innowacji.

Oprócz prognozowania, ważnym elementem obserwacji i przewidywania zmian w otoczeniu technologicznym jest *technology roadmapping* – przewidywanie rozwoju głównych komponentów i technik wytwarzania, składających się na konkretną technologię tak, by zagwarantować poprawną realizację harmonogramu projektu jej rozwoju (Coates i in. 2001: 9). W praktyce naukowej i gospodarczej przydatna okazuje się też *competitive technology intelligence* - identyfikacja szans i zagrożeń dla technologii, pochodzących z otoczenia organizacji (Coates i in. 2001: 10). W tych działaniach również może pomóc odwołanie się do bibliometrii – ilościowych analiz publikacji naukowych i patentów.

4.2. Bibliometria na tle innych technik

Techniki bibliometryczne są często stosowane przez naukowców, instytucje publiczne i przedsiębiorstwa, jednak relatywnie mało znane w Polsce. Można je podzielić na techniki ewaluacyjne i opisowe (deskryptywne). Techniki **ewaluacyjne** służą ocenie badaczy i ośrodków naukowych, identyfikacji najważniejszych osiągnięć i porównaniu dorobku naukowego. Ta odmiana bibliometrii nie stanowi przedmiotu zainteresowania niniejszej książki. Ważniejsze z perspektywy zarządzania technologiami są **deskryptywne**

zastosowania bibliometrii, pozwalające na obserwację tendencji w rozwoju nauki i technologii, identyfikację istotnych aktorów na scenie innowacji oraz lepsze zrozumienie specyfiki poszczególnych obszarów badań.

Bibliometria okazuje się szczególnie przydatna wtedy, gdy niezbędna jest analiza obszarów słabo rozpoznanych lub nieznanymi analitykom oraz gdy nie ma możliwości zdobycia szczegółowych informacji od organizacji, prowadzącymi badania w danym obszarze. Techniki bibliometryczne oferują dogodne sposoby analizy ze względu na odwołanie do wiedzy skodyfikowanej, wykorzystanie mierzalnych danych oraz dostępność danych (Tijssen 2004: 701). Tabela 2 prezentuje porównanie różnych mechanizmów transferu wiedzy technicznej, potwierdzając przydatność analiz publikacji naukowych i patentów.

Mechanizm	Typ wiedzy	Możliwość pomiaru	Dostępność danych
Kontakty i sieci nieformalne	nieskodyfikowana	-	-
Współdzielona infrastruktura techniczna	nieskodyfikowana	+	-
Kontakty i sieci formalne	nieskodyfikowana	+	+
Edukacja i szkolenia	nieskodyfikowana i skodyfikowana	+	+
Badania zlecone, usługi doradcze	nieskodyfikowana i skodyfikowana	++	+
Współpraca B+R pomiędzy instytucjami sektorów publicznego i prywatnego	nieskodyfikowana i skodyfikowana	++	++
Mobilność pracowników	nieskodyfikowana	++	++
Firmy odpryskowe i startup	nieskodyfikowana	+++	++
Patenty, cytowania patentów	skodyfikowana	+++	+++
Publikacje naukowe, cytowania publikacji	skodyfikowana	+++	+++

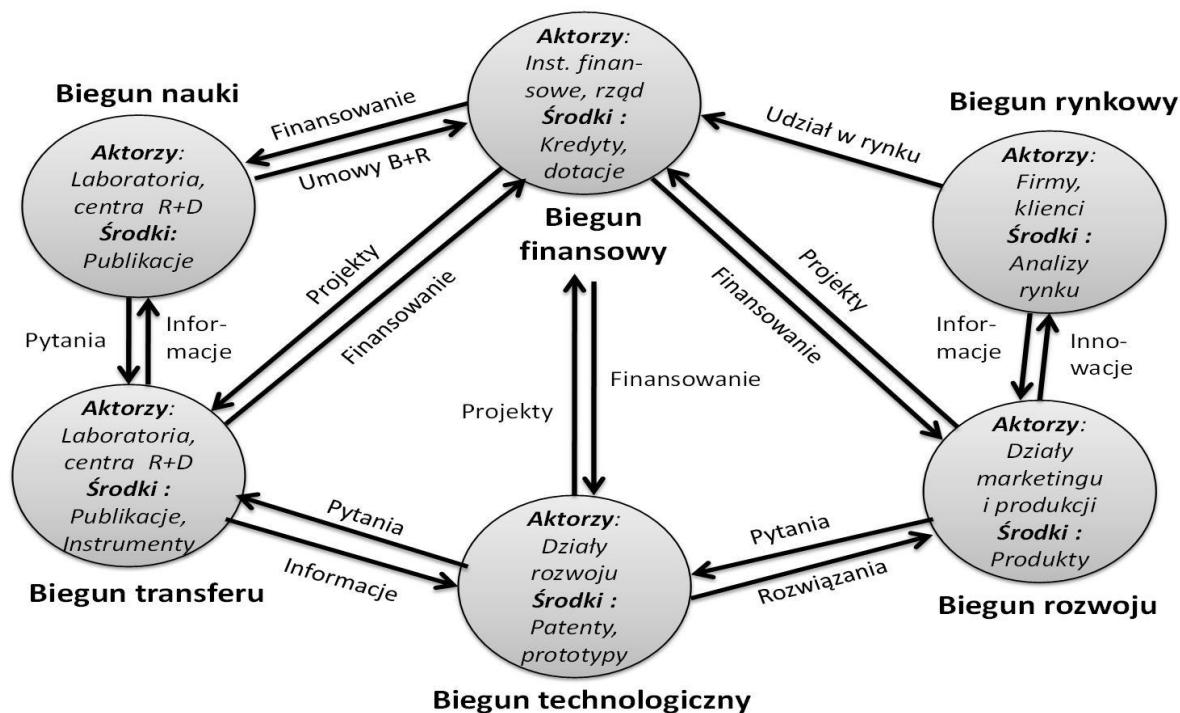
Tabela 2: Porównanie mechanizmów transferu wiedzy technicznej. Źródło: Tijssen (2004: 701).
Legenda:

- +++ - dogodne możliwości pomiaru, wystandaryzowane wskaźniki międzynarodowe,
- ++ - satysfakcjonujące możliwości pomiaru, nieliczne wystandaryzowane wskaźniki,
- +- ograniczone możliwości pomiaru,
- -- dostępne jedynie dane o charakterze jakościowym.

4.3. Wykorzystanie bibliometrii do analiz systemu B+R

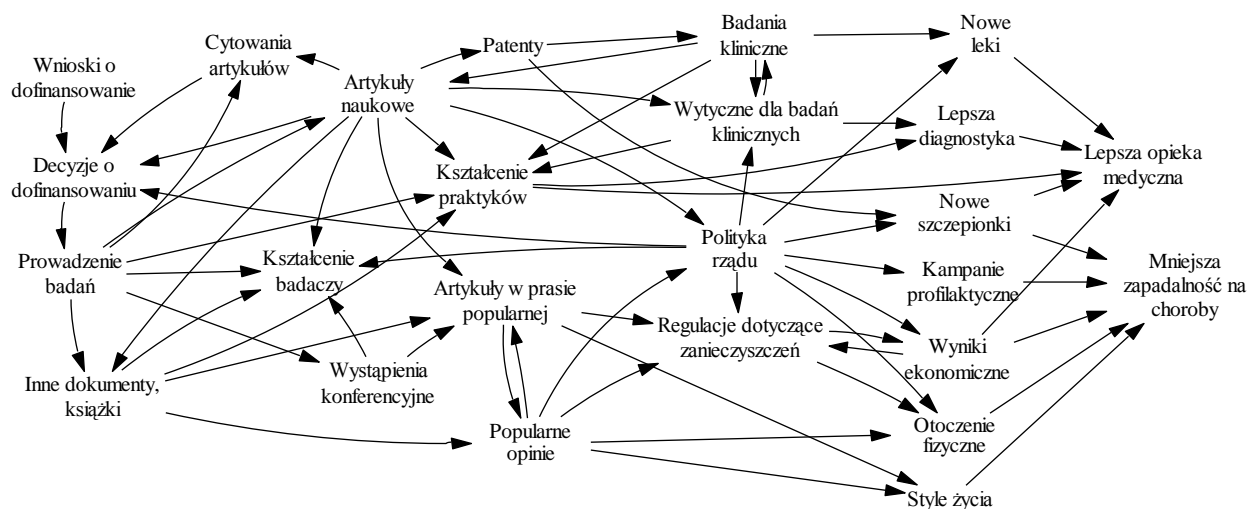
Przydatność bibliometrii przy analizie systemu B+R pozwala lepiej zrozumieć schemat analityczny sieci techno-ekonomicznej (Rysunek 4), zaproponowany dla potrzeb OECD przez Bella i Callona (1994), a później wykorzystywany jako podstawa analiz systemów innowacji przez wielu badaczy (por. dyskusja w: Klincewicz 2008). Wyniki różnych „biegunów” sieci mogą być zmierzone przy wykorzystaniu odpowiednich wskaźników i technik badawczych. Bibliometria okazuje się szczególnie przydatna do analiz biegunów nauki, technologii i transferu, odzwierciedla więc główne obszary zainteresowania sektora nauki, a jednocześnie może uzupełniać analizy ekonomiczne i marketingowe. W nawiązaniu do modelu sieci

techno-ekonomicznej, OECD opublikowało poradnik metodologiczny dla osób, prowadzących analizy bibliometryczne (Okubo 1997).



Rysunek 4: Sieć techno-ekonomiczna jako podstawa analiz bibliometrycznych. Źródło: Bell i Callon (1994: 75).

Analizy bibliometryczne, prowadzące do obserwacji badań naukowych i rozwoju technologii, funkcjonują w szerszym systemie zależności. Lewison (2004: 459) wskazuje, że efekty prac B+R, odzwierciedlone przez bibliometrię, są tylko elementem bardziej złożonej sieci przyczynowo-skutkowej. Rysunek 5 prezentuje przykładowe zależności w obszarze medycyny, w której badania naukowe i rozwój technologii wykazują liczne pośrednie i bezpośrednie związki z poprawą zdrowia publicznego.



Rysunek 5: Sieć zależności pomiędzy badaniami medycznymi a poprawą zdrowia publicznego. Źródło: Lewison (2004: 459).

4.4. Zalety i ograniczenia bibliometrii

Specyficzny charakter technik bibliometrycznych jest źródłem ich licznych zalet metodologicznych, gdyż prowadzone analizy są (Hicks i in. 2002: 45):

- ilościowe - zobiektywizowane, trudne do manipulacji, precyzyjne i spójne;
- dające się znormalizować - pozwala porównać obszary badawczych lub ośrodki o zróżnicowanej produktywności;
- bezpośrednie - proste w interpretacji;
- oparte na publikacjach i cytowaniach - czyli wymiernych rezultatach badań a nie trudnych do weryfikacji opiniach eksperckich na temat osiągnięć;
- charakteryzujące się niewielkim odstępem czasowym – pomiędzy momentem prowadzenia analiz a uzyskania ich wyników;
- skalowalne - pozwalają na analizy zarówno niewielkich, jak i bardzo dużych zbiorów danych;
- pozwalające na prowadzenie analiz przez osoby niezaangażowane - w większości przypadków analizy mogą prowadzić analitycy, nie prowadzący własnych badań naukowych w poddanym analizom obszarze i dlatego nie reprezentujący interesów żadnego z ocenianych podmiotów;
- nieinwazyjne - nie wymagają gromadzenia danych w drodze badań ankietowych lub wywiadów, mogą być prowadzone wielokrotnie w oparciu o dostępne bazy.

Techniki bibliometryczne, które zostaną omówione w dalszych częściach książki, zawierają również wiele ograniczeń, które dyskutowane będą w odniesieniu do poszczególnych technik.

Podstawowe ograniczenie bibliometrii dotyczy przedmiotu prowadzonych analiz: najważniejszymi produktami działalności naukowej są idee a nie obiekty, podczas gdy to właśnie obiekty (m.in. publikacje i patenty) można mierzyć przy wykorzystaniu technik bibliometrycznych (Okubo 1997: 6). Przy prowadzeniu analiz bibliometrycznych możliwe jest dodatkowo popełnienie:

- błędu pierwszego rodzaju (ang. *false positive*) – prowadzący do pominięcia zjawisk i tendencji, które nie znajdują dostatecznego odzwierciedlenia w analizowanym podziorze danych bibliometrycznych,
- błędu drugiego rodzaju (ang. *false negative*) – w tym identyfikacji zjawisk i tendencji, które nie występują w rzeczywistości.

Przykładowo, istotny problem związany z analizami bibliometrycznymi stanowią decyzje niektórych firm, które unikają publikowania swoich osiągnięć badawczych. Analiza bibliometryczna dotycząca wykorzystania ceramiki w konstrukcji silników nie wykazała żadnej aktywności ani zainteresowań tym obszarem ze strony firm General Motors i Chrysler, choć obie firmy w rzeczywistości podejmowały takie działania (Watts, Porter 1997: 44), czyli wyniki analizy nie odzwierciedlały dostatecznie dobrze realiów technologicznych i popełniony został błąd pierwszego rodzaju. Analogicznie, wykorzystanie zaawansowanych metod statystycznych do przetwarzania zbioru danych może doprowadzić do powstania artefaktów - niepożądanych rezultatów, odzwierciedlających nieistniejące w rzeczywistości

zjawiska. Szczególnie łatwo popełnić błąd drugiego rodzaju przy poszukiwaniu „ukrytych zmiennych” (ang. *latent variables*), zależności pomiędzy obiektami w zbiorze danych przy pomocy technik analizy czynnikowej, gdyż odnalezione zmienne (czynniki, składowe) mogą nie posiadać logicznej interpretacji.

Warto być świadomym opisanych ograniczeń, jednocześnie pamiętając o istotnych korzyściach, wynikających z praktycznych zastosowań bibliometrii w odniesieniu do obserwacji badań naukowych i rozwoju technologii.

5. Analizy bibliometryczne w świetle doświadczeń międzynarodowych

5.1. Wykorzystanie bibliometrii przez zagraniczne organizacje

Obserwacja badań naukowych i rozwoju technologii jest praktykowana przez organizacje na całym świecie. Poniżej przedstawiony zostanie przegląd wybranych doświadczeń międzynarodowych i najlepszych praktyk, które mogą okazać się przydatne dla polskich instytucji.

Techniki analityczne, opisane w książce, opierają się na praktycznych doświadczeniach administracji publicznej, instytucji naukowych oraz przedsiębiorstw. Wiele publikacji naukowych, dotyczących metod obserwacji technologii, w szczególności prowadzenia analiz bibliometrycznych, nawiązuje do projektów, których podstawowym celem było uzyskanie wiedzy, wspierającej zarządzanie działalnością badawczo-rozwojową oraz podejmowanie decyzji o alokacji środków, w szczególności o inwestycjach w projekty dotyczące nowych technologii, jednak teksty publikowane zwykle z oczywistych względów nie zawierają bezpośrednich informacji na temat pierwotnych, praktycznie zorientowanych celów podjęcia analiz.

5.2. Doświadczenia instytucji rządowych i wojskowych w Stanach Zjednoczonych

Do pionierów systematycznej obserwacji rozwoju nauki i technologii należały organizacje z sektora obronności Stanów Zjednoczonych, który znaczącą część swojego budżetu przeznacza na dofinansowanie prac badawczo-rozwojowych przez podmioty zewnętrzne (m.in. za pośrednictwem programu *DARPA*, finansującego projekty B+R wyłaniane w trybie konkursowym dla poszczególnych tematów badawczych, w drodze przekazywania grantów dla instytucji naukowych oraz wsparcia małych firm technologicznych poprzez fundusz *SBIR*). Współczesna wojskowość wykorzystuje zaawansowane technologie w szczególnie aktywny sposób, opierając się na badaniach z różnorodnych obszarów, obejmujących m.in. zaawansowane materiały, biotechnologie, optoelektronikę, robotykę, informatykę czy badania podstawowe, pozwalające na lepsze zrozumienie zjawisk fizycznych, chemicznych i biologicznych. Przy znaczących środkach budżetowych, przeznaczanych na prace B+R oraz dążeniu do zwiększania efektywności wydatkowania tych środków, reprezentanci sektora obronności prowadzili samodzielne analizy, pozwalające na: identyfikację nowych tendencji w badaniach naukowych i rozwoju technologii, selekcję tych osiągnięć, które mogą znaleźć potencjalne zastosowania w sferze militarnej oraz poszukiwanie potencjalnych wykonawców badań zleconych (indywidualnych badaczy, instytucji akademickich oraz przedsiębiorstw).

Projekty, finansowane przez Departament Obrony Stanów Zjednoczonych, przyczyniły się do znaczącego udoskonalenia technik analitycznych, pozwalających na obserwację rozwoju technologii przy wykorzystaniu bibliometrii.

Z budżetu obronnego Stanów Zjednoczonych wydano na przestrzeni kilkunastu lat znaczące środki na sfinansowanie rozwoju programu komputerowego *TechOASIS*, służącego do analiz

bibliometrycznych zgodnie z podejściem, które w dalszej części książki zostanie zaprezentowane jako „*tech mining*”. Rozwój *TechOASIS* był finansowany na przestrzeni kilkunastu lat przez: *Information Technology Office DARPA (Defense Advanced Research Projects Agency)*, *TACOM (Army Tank-Automotive and Armaments Command)* oraz *Office of Naval Research*. *TechOASIS* rozwijany był przez zespół specjalistów z uczelni *GeorgiaTech*, która wyodrębniła później spółkę odpryskową *Search Technology, Inc.*, przenosząc do niej własność kluczowych technologii bibliometrycznych. Firma oferuje obecnie również komercyjną wersję produktu *TechOASIS*, sprzedawaną pod nazwą *VantagePoint* (oba produkty oferują zbliżoną funkcjonalność i interfejsy użytkownika), udostępniając uczelniom wyższym i firmom techniki, wypracowane w oparciu o najlepsze praktyki wojskowe. Co ciekawsze, amerykański sektor obronności udostępnia publicznie wiele raportów i wyników analiz rozwoju nauki i technologii. Część tych dokumentów odnaleźć można na ogólnodostępnym portalu *DTIC (Defense Technical Information Center)*, który oferuje funkcje wyszukiwania raportów technicznych pod adresem: www.dtic.mil/dtic/search/tr/index.html

Przykładowe analizy, wykonywane w oparciu o techniki bibliometryczne i program *TechOASIS*, obejmowały m.in.:

- systematyczne analizy dorobku naukowego wybranych krajów (m.in. Brazylia, Chiny, Indie, Finlandia) ze szczególnym uwzględnieniem ośrodków badawczych, których prace mogą mieć zastosowania militarne,
- przeglądy dziedzin nauki i wybranych obszarów badawczych,
- raporty przygotowywane w odpowiedzi na doraźne potrzeby informacyjne (np. analiza badań nad wąglikiem, pozwalająca na identyfikację zespołów badawczych, zajmujących się omawianą tematyką i mających dostęp do szczepów bakterii tej śmiertelnej choroby).

Ze względu na łatwość dostępu do dorobku analitycznego sektora obronności USA oraz wykorzystanie zaawansowanych technik bibliometrycznych, opartych na dedykowanym oprogramowaniu, można uznać omawiane doświadczenia amerykańskie za najlepszy dobrze udokumentowany przykład systematycznych obserwacji i analiz rozwoju badań naukowych i technologii. Program komputerowy *TechOASIS* wraz z jego wersją komercyjną *VantagePoint* jest obecnie najbardziej wszechstronnym narzędziem do analiz bibliometrycznych, pozwalającym na łatwe dostosowanie do współpracy z dowolnymi bazami publikacji i patentów, a jednocześnie wykorzystującym algorytmy analityczne, opatentowane przez armię Stanów Zjednoczonych. Dodatkową zaletą podejścia jest jego uniwersalność – możliwość prowadzenia analiz, dotyczących różnych obszarów naukowych, w przeciwieństwie do wielu narzędzi, skoncentrowanych na węższych specjalizacjach (np. naukach medycznych).

Ruegg i Jordan (2007: 3-4) prezentują porównanie zróżnicowanych technik ewaluacji, wykorzystywanych przez amerykańskie agencje rządowe: *National Science Foundation (NSF)*, *National Institute of Health (NIH)*, *Department of Energy (DOE)* oraz *National Institute of Standards and Technology (NIST)* w ramach *Advanced Technology Program (ATP)* w okresie do roku 2002 (Tabela 3). Wśród projektów bibliometrycznych, *Department of Energy* prowadził przykładowo analizy wpływu finansowania przez *DOE* projektów

badawczych na rozwój nowych, przydatnych komercyjnie technologii (Ruegg, Jordan 2007: 41), w tym m.in. związki pomiędzy rozwojem nanotechnologii a inicjatywami *DOE* (Ruegg, Jordan 2007: 45-46). Przykładem rezultatów analiz bibliometrycznych, prowadzonych przez amerykańskie instytucje militarne, jest obszerna, licząca ponad 500 stron analiza osiągnięć i uwarunkowań instytucjonalnych nauki chińskiej (Kostoff i in. 2006). Kostoff (2002: 55) opisuje też przykłady wykorzystania technik bibliometrycznych do ewaluacji propozycji badawczych, zgłaszanych do instytucji publicznej (najprawdopodobniej w sektorze obronności USA), gdzie analizy bibliometryczne pomogły ocenić poziom naukowy oraz związek proponowanych badań z misją sponsora.

Metody / instytucje	<i>National Science Foundation</i>	<i>National Institute of Health</i>	<i>Department of Energy</i>	<i>National Institute of Standards and Technology</i>
Badania ankietowe	+	+	+	+
Studia przypadku	+	+	+	+
Panele ekspertów	-	+	+	+
Obliczanie wskaźników	-	+	+	+
Bibliometria	+	+	+	+
Analizy historyczne	+	+	+	-
Modelowanie ekonometryczne	-	+	-	+
Benchmarking	-	+	+	+
Analiza sieci	-	-	+	+
Foresight	-	-	+	-

Tabela 3: Metody ewaluacji wykorzystywane przez instytucje publiczne w Stanach Zjednoczonych w okresie do roku 2002. Źródło: Ruegg, Jordan (2007: 4).

National Science Board przygotowuje regularne raporty „*Science and Engineering Indicators*”, wykorzystujące m.in. wskaźniki bibliometryczne oparte na bazach firmy *Thomson Reuters* i zasobach rodzimego urzędu patentowego (*National Science Board* 2010) – jest to najbogatsze ogólnodostępne źródło aktualnych danych na temat systemu B+R Stanów Zjednoczonych.

5.3. Doświadczenia instytucji rządowych innych krajów

Większość rządów krajów rozwiniętych posiada własne zespoły analityków lub zleca regularne badania, mające na celu odzwierciedlanie rozwoju nauki i technologii, ze szczególnym uwzględnieniem dorobku krajowych badaczy. W wielu przypadkach, takie analizy ograniczają się do zliczania publikacji w poszczególnych obszarach badawczych oraz

identyfikacji najczęściej cytowanych badaczy oraz zatrudniających ich instytucji. Warto zauważyć, że współcześnie wymienione informacje zdobyć można za pośrednictwem prostych poleceń wyszukiwania oraz odpowiedniego sortowania wyników z baz takich jak *Web of Science* lub *Scopus*. Dostawcy tych narzędzi systematycznie je doskonalą, ułatwiając użytkownikom obsługę i oferując dodatkowe funkcje. Przykładowo, prace analityczne wykonywane w latach 90-tych XX wieku przez zespół pod kierownictwem prof. Benoîta Godina dla rządu Kanady obecnie mogą być w łatwy sposób zrealizowane w drodze prostych kwerend baz, wykonywanych samodzielnie przez pracowników instytucji rządowych, pod warunkiem posiadania dostępu do tych baz oraz umiejętności obsługi klasycznego interfejsu wyszukiwarki internetowej.

Rządy niektórych krajów zdecydowały się na prowadzenie bardziej zaawansowanych analiz przy wykorzystaniu dedykowanych narzędzi informatycznych. Nie są dostępne szczegółowe informacje o instytucjach administracji publicznej - użytkownikach poszczególnych systemów, gdyż w wielu przypadkach rządy poszczególnych krajów traktują omawianą działalność jako specyficzną odmianę „białego wywiadu gospodarczo-technologicznego”, nie decydując się na ujawnianie faktu prowadzenia tego typu analiz ani prezentowanie wykorzystywanych technik.

Historyczne przykłady wykorzystania analizy patentów obejmują projekty *National Science Foundation* z lat 70-tych i 80-tych XX wieku (Hicks i in. 2002: 34). W tym samym okresie Michel Callon z *École des Mines de Paris* wypracował technikę analizy *co-word*, stosowaną m.in. do oceny rezultatów interwencji rządowych w sferę badań naukowych m.in. w obszarach chemii makromolekularnej, akwakultury i biotechnologii (Kostoff 2003: 6-14). W latach 80-tych XX wieku w polityce naukowej niektórych krajów zaczęto wykorzystywać analizę cytowań jako element ewaluacji naukowców i instytucji (por. przegląd publikacji, dokumentujących projekty rządowe w: Hicks 1987: 296). Do pionierów podejścia analitycznego w Europie należało założone w 1990 roku francuskie *Observatoire des Sciences et Techniques (OST)* oraz powstałe w 1992 roku przy Uniwersytecie w Lejdzie *NOWT* (Holdenderskie Centrum Obserwacji Nauki i Technologii) (Leeuven 2004: 375-376, Delanghe, Sloan, Muldur 2010). Francuskie analizy są regularnie publikowane w postaci zestawień najważniejszych wskaźników bibliometrycznych dla Francji, jej regionów oraz innych krajów świata (OST 2008). Aktywne analizy rozwoju technologii zlecają współpracującym podmiotom m.in. rządy Brazylii (*CGEE*, Centrum Zarządzania i Studiów Strategicznych, por. Miranda Santo i in. 2006: 1015), Turcji (studium typu *foresight "Vision 2023"* z 2003 roku, powiązane z prognozami ilościowymi, opartymi na danych bibliometrycznych, por. Bengisu, Nekhili 2006), Australii (analiza cytowań patentów australijskich, ujawniająca niski poziom narodowej innowacyjności i stymulująca debatę na temat systemów wsparcia, por. Hicks i in. 2002: 140-141), Indii i Malezji. Japońskie Ministerstwo Edukacji, Nauki i Kultury (jap. *Monbukagakusho*) zleciło już w 1987 roku bibliometryczne analizy porównawcze siedmiu krajów w celu lepszego zrozumienia specyfiki japońskich badań naukowych (Okubo 1997: 12), a od tego czasu regularnie wykorzystuje omawiane techniki. Szczególnie interesujące wydają się analizy prowadzone przez koreańską instytucję *KISTI (Korea Institute of Science and Technology Information)*, która wykorzystuje amerykański program *VantagePoint* do kompleksowej analizy wszystkich obszarów badań w

światowej nauce w celu systematycznej identyfikacji nowych podobszarów, w których mogą podjąć działalność koreańscy naukowcy (Porter, Newman 2011).

Systematyczne prace nad mapowaniem kompetencji naukowych i technologicznych w zestawieniu z tendencjami w globalnej nauce podejmowane są przez większość krajów Europy Zachodniej. Do nowszych przykładów należą analizy, prowadzone na zlecenie niemieckiego resortu nauki *Bundesministerium für Bildung und Forschung* i publikowane na stronach resortu (np. zestawienie międzynarodowych prognoz dotyczących rozwoju zaawansowanych technologii: Holtmannspötter i in. 2010 oraz przeprowadzona przez *Fraunhofer-Institut für Systemtechnik und Innovationsforschung* analiza konkurencyjności niemieckiej nauki na tle międzynarodowym: Schmoch 2007). Rząd Finlandii zlecił przygotowanie analogicznych analiz rodzimej instytucji badawczej *Academy of Finland* w 2006 roku (Lehvo, Nuutinen 2006).

Również Komisja Europejska przygotowuje regularne raporty, dotyczące nauki i techniki krajów członkowskich, zlecając też dodatkowe analizy bibliometryczne, służące identyfikacji wiodących zespołów badawczych w poszczególnych obszarach badań i mające na celu wskazanie potencjalnych uczestników konsorcjów badawczych, mogących wystąpić o środki w ramach 7. Programu Ramowego w Zakresie Badań i Rozwoju Technologicznego (takie analizy w odniesieniu do Polski były realizowane na zlecenie Komisji w 2009 r. przez jednego ze współautorów niniejszej książki).

Interesująca wydaje się również szczegółowa tematyka analiz, które były prowadzone na zlecenie instytucji rządowych poszczególnych krajów – jej zestawienie oferuje Tabela 4, wskazując jednocześnie na skalę stosowania *evidence-based policy* w polityce naukowej i innowacyjnej poszczególnych krajów.

Obszar badawczy	Stany Zjednoczone	Japonia	Hiszpania	Francja	Wielka Brytania	Indie
Energia						
Ochrona Środowiska						
Zdrowie i żywienie						
Technologie informacyjne						
Transport i logistyka						
Biotechnologie i nauki przyrodnicze						
Wojskowość i bezpieczeństwo						
Budownictwo i mieszkalnictwo						
Techn. produkcyjne i procesowe						
Technologie materiałowe						
Nanotechnologie i mikrosystemy						
Lotnictwo i loty kosmiczne						
Technologie morskie						
Technologie optyczne						
Usługi						
Elektronika						

Tabela 4: Zestawienie analiz badań naukowych i rozwoju technologii, prowadzonych na zlecenie instytucji publicznych w wybranych krajach w latach 2004-2010. Źródło: Holtmannspötter i in. (2010: 173). Legenda:

- (■) - obszar analizowany intensywnie,
- (□) - obszar analizowany doraźnie,
- () - obszar niepoddany analizom.

5.4. Doświadczenia firm technologicznych

Analogicznie, tylko ograniczone informacje można zdobyć na temat prac, prowadzonych przez przedsiębiorstwa. Duże firmy technologiczne prezentują na konferencjach oraz w magazynach branżowych ogólne informacje dotyczące własnych działań w zakresie *technology intelligence*, jednak przekazywane informacje ograniczają się zwykle do bardzo ogólnych założeń, utrudniających bezpośrednio powielenie działań analitycznych przez konkurentów. Również badania naukowe, dotyczące metod analizy i prognozowania trendów technologicznych przez przedsiębiorstwa, ograniczają się do zdobycia wiedzy o tym, które metody są wykorzystywane, bez opisu szczegółowych sposobów wykorzystywania tych

metod. Przykładami są opisy działań, podejmowanych przez firmy *Kodak* (Hicks i in. 2002: 35; Mortara i in. 2010), *Deutsche Telekom* (Thom, Rohrbeck 2009), *Procter & Gamble* (Calcagno 2008), *TetraPak* (Fattori, Pedrazzi, Turra 2003), *Bristol-Meyers Squibb* (Yang i in. 2008), *Syngenta Crop Protection AG* (Fischer, Lalyre 2006), *Air Products and Chemicals, Inc.* (Brenner 2005), duży koncern lotniczy (Arman, Foden 2010), niemiecką firmę przemysłową *Wittenstein AG* (Glaser, Miecznik 2009), jak również specjalistyczne usługi oferowane dla firm europejskich przez firmę doradczą *FutureManagementGroup AG* (Mićić 2010) sposoby pracy analityków z *Georgia Institute of Technology* przy pracy z klientami zewnętrznymi (Porter, Detampel 1995), czy przegląd metod, wykorzystywanych przez największe amerykańskie i europejskie przedsiębiorstwa z branż farmaceutycznej, elektronicznej i motoryzacyjnej (Lichtenthaler 2004). Przekazywane w wymienionych tekstach informacje są wprawdzie interesujące poznawczo, dostarczają nowej wiedzy na temat zainteresowań analitycznych firm, jednak nie zawierają wskazówek praktycznych, które pozwoliłyby na powielenie podejść, stosowanych przez opisywane firmy.

Badania Lichtenthalera (2004), dotyczące 27 liderów trzech branż technologicznych w Stanach Zjednoczonych i Europie, pozwoliły na ujawnienie dominujących tendencji w zakresie obserwacji rozwoju nauki i technologii przez pracujących dla tych firm specjalistów. Wyniki tych badań prezentuje Tabela 5. Interesujące wydaje się ograniczone zastosowanie analiz cytowań publikacji i patentów przez sektor przedsiębiorstw (z wyjątkiem firm farmaceutycznych).

Metoda	Firmy farmaceutyczne	Firmy elektroniczne	Firmy motoryzacyjne
Analiza liczby publikacji naukowych	+++	++	+
Analiza cytatów publikacji naukowych	+++	-	-
Analiza liczby wystąpień konferencyjnych	++	+++	+
Analiza liczby patentów	++	+++	+++
Analiza cytowań patentów	-	-	+
Modelowanie rozwoju technologii (krzywe S-kształtne)	-	-	-
Benchmarking / obserwacje działań konkurentów	+++	+++	+++
Studia delfickie (foresight)	-	-	-
Panele ekspertów	+++	+	++
Symulacje	++	-	-
Analizy scenariuszowe	+++	+++	+++

Tabela 5: Intensywność wykorzystywania metod w zakresie technologii intelligence w wybranych branżach. Źródło: Lichtenthaler (2004: 130). Legenda:

- +++ - metoda wykorzystywana często,
- ++ - metoda wykorzystywana czasami,
- + - metoda wykorzystywana rzadko,
- - - metoda nie wykorzystywana.

Wielu badaczy, specjalizujących się w bibliometrii, właśnie cytowania uważają za najważniejszą podstawę analiz – jednak praktyka gospodarcza wskazuje na ograniczoną praktyczną przydatność studiów, przy znacznie większym nacisku na analizy liczby publikacji

dotyczących poszczególnych zagadnień. Nie są również stosowane w praktyce metody, oparte na modelowaniu cyklu życia technologii, zgodnie z modelem krzywej S-kształtnej, a zaawansowane techniki analiz statystycznych i symulacji znajdują zastosowanie tylko w branży farmaceutycznej. Interesujące są też intensywne analizy wystąpień konferencyjnych przez firmy z branży elektronicznej – w elektronice regularnie pojawiają się nowe rozwiązania technologiczne, a konferencje naukowe są znacznie szybszym sposobem na przekazanie wiedzy o wynikach badań niż publikacje w czasopiśmie recenzowanych. Analiza wystąpień konferencyjnych z kolei jest rzadko stosowana przez badaczy, zajmujących się bibliometrią, nie wspiera jej też baza *Web of Science*, ograniczająca się do artykułów z najbardziej znanych magazynów naukowych.

Ciekawym wynikiem analiz sektora przedsiębiorstw jest wskazanie, że procesy obserwacji rozwoju technologii nie są zwykle delegowane do wyodrębnionych działów, a realizowane przy zaangażowaniu pracowników, zajmujących się też innymi zadaniami (Lichtenthaler 2004: 122) – zdobycie podstawowych umiejętności analitycznych i znajomości wykorzystywanych technik wydaje się przydatne wielu pracownikom, a nie tylko ściśle wyodrębnionej grupie „obserwatorów” lub zewnętrznych doradców. Często zaobserwować można działalność tzw. *technology scouts* (dosł. harcerzy technologicznych) – pracowników, utrzymywanych przez firmy technologiczne w zagranicznych oddziałach, którym zlecono zadania skanowania otoczenia oraz gromadzenia informacji o badaniach naukowych i rozwoju technologii. Przegląd praktycznych doświadczeń firm niemieckich w tym zakresie oferuje Reger (2001: 546).

Badania Lichtenthalera (2004: 130) wskazują też na najważniejsze źródła informacji, wykorzystywanych przez firmy technologiczne z poszczególnych branż, co odzwierciedla Tabela 6.

Źródło	Firmy farmaceutyczne	Firmy elektroniczne	Firmy motoryzacyjne
Publikacje	+++	+++	++
Patenty	++	+++	+++
Fundusze <i>venture capital</i>	+++	++	+
Targi dla firm <i>startup</i>	+++	++	+
Kontakty na uczelniach	+++	+++	++
Konsultanci i zewnętrzn eksperci	+++	++	++
Konferencje naukowe	+++	+++	+++
Targi technologiczne	-	+	++
Konferencje biznesowe	++	+++	+++
Partnerzy	+++	+++	+++
Dostawcy	-	++	+++
Agencje dostarczające informacje na temat technologii	+++	+++	+++
Raporty spółek giełdowych	+++	+	+
Rządowe studia typu <i>foresight</i>	+	+	++
Komitety branżowe i normalizacyjne	+	+++	++
Udział w krajowych komitetach ds. badań naukowych	+++	+++	+++
Finansowanie badań doktorskich	+++	++	+
Urlopy pracowników, podejmowane w celu pracy na uczelni	+++	+	-
Okresowe zatrudnianie naukowców ze stopniami doktora	+++	++	+
Rozmowy z klientami	-	+++	+++

Tabela 6: Intensywność wykorzystywania źródeł informacji w procesach *technology intelligence* w wybranych branżach. Źródło: Lichtenthaler (2004: 130). Legenda:

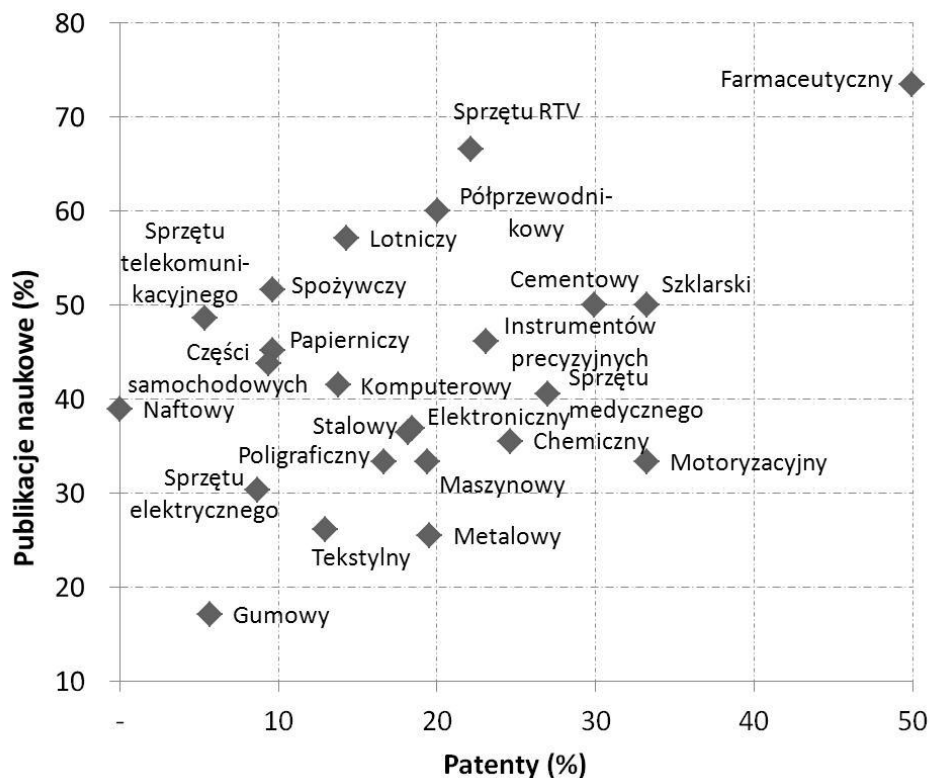
- +++ - źródło wykorzystywane często,
- ++ - źródło wykorzystywane czasami,
- + - źródło wykorzystywane rzadko,
- - - źródło nie wykorzystywane.

Cohen, Nelson i Walsh (2002) przeprowadzili szerzej zakrojone badanie, identyfikujące sposoby wykorzystywania publicznie dostępnej wiedzy naukowej przez firmy z poszczególnych branż. Ich analizy nie nawiązują wprawdzie do ewentualnego wykorzystywania technik bibliometrycznych, ale ujawniają zróżnicowane zainteresowania przedsiębiorstw wynikami badań naukowych w artykułach, wystąpieniach konferencyjnych i dokumentach patentowych.

Stopień wykorzystania wyników finansowanych publicznie badań naukowych przez firmy jest zróżnicowany dla poszczególnych branż - relatywnie wysoki w przemysłach półprzewodnikowym, telekomunikacyjnym, petrochemicznym i farmaceutycznym, a bardzo niski w poligrafii, produkcji cementu, stali i metalurgii (Cohen, Nelson, Walsh 2002: 9).

Aktywne wykorzystywanie informacji pochodzących z dokumentów patentowych deklarują tylko firmy farmaceutyczne i motoryzacyjne (przez aktywne wykorzystywanie rozumie się sytuację, w której co najmniej 1/3 ankietowanych firm z branży udzieliła pozytywnych odpowiedzi) (Cohen, Nelson, Walsh 2002: 11). Na tym tle, publikacje naukowe i wystąpienia

konferencyjne okazały się ważne dla reprezentantów większości branż (Cohen, Nelson, Walsh 2002: 11), co może stanowić potwierdzenie przydatności analiz bibliometrycznych publikacji także dla podmiotów z sektora przedsiębiorstw. Rysunek 6 prezentuje wyniki analiz, przeprowadzonych przez Cohena, Nelsona i Walsha (2002: 11) w podziale na poszczególne branże, ujawniając zróżnicowane wykorzystanie publicznie dostępnych źródeł wiedzy.



Rysunek 6: Ocena ważności publicznie dostępnych źródeł wiedzy dla firm z wybranych przemysłów w oparciu o badania ankietowe. Źródło: opracowanie własne na podstawie Cohen, Nelson, Walsh (2002: 11).

6. Dane bibliometryczne wykorzystywane do celów analitycznych

6.1. Dane bibliometryczne i bazy bibliograficzne

Zanim omówione zostaną techniki analizy danych bibliometrycznych, warto wyjaśnić, jakiego rodzaju dokumenty mogą być wykorzystywane jako podstawa tych analiz. Najczęściej są nimi streszczenia artykułów naukowych i wystąpień konferencyjnych oraz patenty lub zgłoszenia patentowe. Technicznie możliwa jest też analiza pełnych treści artykułów naukowych, jednak w większości baz publikacji, wykorzystywanych przez analityków, dostępne są tylko streszczenia i opisy tekstów.

Poszczególne bazy bibliograficzne oferują zróżnicowane zestawy danych metryczkowych, opisujących indeksowane w bazie publikacje. Szczegółowa prezentacja najpopularniejszych baz znajduje się w dalszej części książki, a dyskusję ich praktycznej przydatności do analiz bibliometrycznych znaleźć można np. w publikacji OECD (1994). Warto jednak zaprezentować przykładowe dane metryczkowe w celu uświadomienia, jakie analizy można prowadzić w oparciu o poszczególne bazy.

6.2. Dane bibliometryczne dostępne w bazach publikacji i streszczeń

Kolejne rysunki prezentują typowy wygląd streszczeń (abstraktów) artykułów w bazach:

- *Web of Science* firmy *Thomson Reuters* (dostępna za pośrednictwem platformy *Web of Knowledge*),
- *Scopus* firmy *Elsevier*,
- *MEDLINE* rozwijana przez amerykańską instytucję rządową *National Library of Medicine* (baza znana też jako *PubMed*),
- *INSPEC* utrzymywana przez *Institution of Engineering and Technology (IET)*.

Wybór odpowiedniej bazy do konkretnego projektu analitycznego zależy będzie od obszaru badań oraz planowanych sposobów analiz. Przykładowo, analizy cytowań są możliwe w oparciu o bazy *Web of Science* i *Scopus*, zbiory bazy *MEDLINE* wydają się szczególnie przydatne w projektach analitycznych, dotyczących medycyny, a bazy *INSPEC* – w analizach dorobku nauk ścisłych i inżynieryjnych. Warto zwrócić uwagę na zróżnicowaną budowę danych metryczkowych w poszczególnych bazach. *Web of Science* gromadzi streszczenia artykułów, nie oferuje jednak dodatkowych kodów, opisujących zawartość tekstu. *Scopus* udostępnia dodatkowo listy słów kluczowych, a bazy *MEDLINE* i *INSPEC* oferują rozbudowane taksonomie, pozwalające na szczegółowe odzwierciedlenie zawartości tekstu, a przez to łatwą identyfikację wszystkich prac badawczych z danego obszaru. Przykładowo, w bazie *INSPEC* możliwe jest wyszukiwanie publikacji nie tylko na podstawie specjalistycznych pojęć z danego obszaru nauki i techniki, ale też w oparciu o indeksy związków chemicznych i obiektów astronomicznych, do których nawiązują badania, dodatkowo różnicując też indeksowane teksty na badania podstawowe, stosowane, prace eksperymentalne i przeglądy teorii.

Baza *Web of Science* nie oferuje własnej taksonomii dla poszczególnych publikacji, przypisuje jednak każde czasopismo do określonego obszaru badań naukowych.

Wykorzystanie tego podziału może niestety prowadzić do mylnych rezultatów poszukiwań, gdyż odnosi się on do całego periodyku a nie pojedynczych publikacji, utrudniając poprawną identyfikację treści. Panel ekspertów amerykańskiego *National Science Board* uznał, że baza *Web of Science* nie nadaje się do analiz interdyscyplinarnych obszarów badawczych, gdyż przypisanie poszczególnych czasopisma do określonych dyscyplin uniemożliwia identyfikację publikacji dotyczących kilku obszarów (National Science Board 2010: 5-35).

Warto zwrócić uwagę na możliwości wyszukiwania wystąpień konferencyjnych, które nie są przedmiotem publikacji w czasopismach naukowych, zwykle jednak przechodzą przez proces recenzji naukowej (ang. *peer-review*) i mogą okazać się cennym źródłem wiedzy na temat nowych tendencji w nauce i technice oraz zainteresowań badawczych reprezentantów określonych ośrodków naukowych. Baza *Web of Science* nie zawiera wystąpień konferencyjnych, podczas gdy w bazie *INSPEC* stanowią one istotną część wszystkich indeksowanych zasobów.

Pregnancy and delivery after liver transplantation

Full Text NCB! Print E-mail Add to Marked List Save to EndNote Web Save to EndNote, RefMan, ProCite Save to RefWorks more options

Author(s): Jabiry-Zieniewicz Z, Cyganek A, Luterek K, Bobrowska K, Kaminski P, Ziolkowski J, Zieniewicz K, Krawczyk M

Source: TRANSPLANTATION PROCEEDINGS **Volume:** 37 **Issue:** 2 **Pages:** 1197-1200 **Published:** MAR 2005

Times Cited: 4 **References:** 7 [Citation Map](#)

Conference Information: 20th International Congress of the Transplantation-Society
Vienna, AUSTRIA, SEP 05-10, 2004
Transplant Soc

Abstract: Aim. According to statistics, women constitute one-third of all liver recipients and approximately 75% of female recipients are of reproductive age. Successful liver transplantation in these patients results in the restoration of menstrual function and fertility. The aim of this study was to assess the course of pregnancy and delivery in liver-transplanted women.

Materials and methods. We retrospectively analyzed data of 138 liver-transplanted women, aged from 18 to 63 years, who underwent regular gynecological evaluations. Among 77 patients of reproductive age, 11 women conceived and delivered babies.

Results. All patients have successfully delivered. The mean gestation age at delivery was 36.5 weeks. All neonates were delivered in a good state with no congenital abnormalities. Common pregnancy complications were preterm birth, anemia, intrahepatic cholestasis, and infection. In 1 case, graft rejection was observed due to willful discontinuation of immunosuppressive therapy. Two spontaneous vaginal deliveries and 9 caesarean sections were performed. All caesarean sections were performed for obstetrical indications: fetal intrauterine asphyxia (n = 4), breech presentation (n = 2), threatening intrauterine infection (n = 2), and preterm twin delivery (n = 1).

Conclusion. High-risk pregnancies in liver-transplanted women are generally associated with good outcomes, although an increased rate of preterm labor, intrauterine infections, anemia, and cholestasis were observed. Pregnancy did not seem to impair graft function or accelerate rejection in patients receiving immunosuppressive therapy.

Document Type: Proceedings Paper

Language: English

Reprint Address: Kaminski, P (reprint author), Med Univ Warsaw, Dept Obstet & Gynecol 1, Pl Starynkiewicza 1-3, PL-02015 Warsaw, Poland

Addresses:

1. Med Univ Warsaw, Dept Obstet & Gynecol 1, PL-02015 Warsaw, Poland
2. Med Univ Warsaw, Dept Gen & Liver Surg, PL-02015 Warsaw, Poland

E-mail Addresses: pawel.kaminski@amwaw.edu.pl

Publisher: ELSEVIER SCIENCE INC, 360 PARK AVE SOUTH, NEW YORK, NY 10010-1710 USA

IDS Number: 918SW

ISSN: 0041-1345

DOI: 10.1016/j.transproceed.2005.01.011

Rysunek 7: Streszczenie artykułu w bazie *Web of Science*. Źródło: *Web of Science*.

Bioelectrochemistry

Volume 80, Issue 1, November 2010, Pages 73-80

ISSN: 15675394
CODEN: BIOEFC
DOI: 10.1016/j.bioelechem.2010.06.003
Document Type: Article
Source Type: Journal

View at publisher | [Full Text Elsevier](#) |

Derivatization of single-walled carbon nanotubes with redox mediator for biocatalytic oxygen electrodes

Sadowska, K.^a, Stolarczyk, K.^b, Biernat, J.F.^a, Roberts, K.P.^c, Rogalski, J.^d, Bilewicz, R.^b

^a Chemical Faculty, Gdansk University of Technology, Narutowicza 11/12, 80-233 Gdansk, Poland

^b Department of Chemistry, University of Warsaw, Pasteura 1, 02-093, Warsaw, Poland

^c Department of Chemistry and Biochemistry, University of Tulsa, 800 S. Tucker Dr., Tulsa, OK 74104, United States

^d Department of Biochemistry, Maria Curie Skłodowska University, Akademicka 19 Str., Lublin 20-031, Poland

Abstract

Single-walled carbon nanotubes (SWCNTs) were covalently modified with a redox mediator derived from 2,2'-azino-bis-(3-ethylbenzothiazoline)-6-sulfonic acid (ABTS), and impen electrodes for biocatalytic oxygen reduction. The procedure is based on: covalent bonding of mediator to nanotubes, placing the nanotubes directly on the carbon electrode surface nanostructured electrode with a Nafion film containing laccase as the biocatalyst. The modified electrode is stable and the problem of mediator (ABTS) leaking from the film is eliminated. Three different synthetic approaches were used to obtain ABTS-modified carbon nanotubes. Nanotubes were modified at ends/defect sites or on the nanotube side. Raman spectroscopy, TGA and electrochemistry. The accessibility of differently located ABTS units by the laccase active center and mediation of electron transfer were studied. Concentrations of ABTS groups electrically connected with the electrode were compared for each of the electrodes based on the charges of the voltammetric peaks recorded in the nanotube modification procedure giving the best parameters of the catalytic process was selected. © 2010 Elsevier B.V.

Language of original document

English

Author keywords

Biocathode; Bioelectrocatalysis; Biofuel cell; Laccase; Modified carbon nanotubes; Oxygen reduction

Index Keywords

Biocathodes; Bioelectrocatalysis; Biofuel cell; Laccases; Modified carbon; Oxygen reduction

Engineering controlled terms: Acids; Biofuels; Biological fuel cells; Cyclic voltammetry; Electrodes; Electrolytic reduction; Oxygen; Raman spectroscopy

Engineering main heading: Single-walled carbon nanotubes (SWCNT)

EMTREE drug terms: 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid); biofuel; laccase; single-walled nanotube

EMTREE medical terms: article; biocatalysis; covalent bond; cyclic potentiometry; derivatization; electrochemistry; oxidation-reduction reaction; oxygen electrode; Raman spectroscopy

Chemicals and CAS Registry Numbers

2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid), 28752-68-3; laccase, 80498-15-3

References (43) View in table layout

Rysunek 8: Streszczenie artykułu w bazie *Scopus*. Źródło: *Scopus*.

Bull Med Libr Assoc., 1990 Jan;78(1):29-37.

Medical literature as a potential source of new knowledge.

Swanson DR.

Graduate Library School, Center for Information Studies, University of Chicago, IL 60637.

Abstract

Specialized biomedical literatures have been found that are implicitly linked by arguments that they respect conclusions that cannot be drawn from the separate literatures. One such analysis identified one set of art similar changes might benefit patients with Raynaud's syndrome. Yet these two literatures had no articles in might benefit Raynaud patients. Two years after publication of that analysis, the first clinical trial demonstrating eleven indirect connections, led to an inference that magnesium deficiency might be a causal factor in migr somatomedins, a potentially fruitful but neglected area of research with implications for the decline with age noninteractive literatures is described. Such structures are probably not rare and may provide the foundation

PMID: 2403828 [PubMed - indexed for MEDLINE] PMID: PMC225324 Free PMC Article

Publication Types, MeSH Terms, Substances

Publication Types:

Research Support, U.S. Gov't, Non-P.H.S.

MeSH Terms:

[Abstracting and Indexing as Topic/methods](#)
[Arginine/pharmacology](#)
[Artificial Intelligence](#)
[Fish Oils/therapeutic use](#)
[Forecasting](#)
[Humans](#)
[Insulin-Like Growth Factor I/metabolism](#)
[MEDLARS](#)
[Magnesium Deficiency/complications](#)
[Migraine Disorders/etiology](#)
[Online Systems*/organization & administration](#)
[Periodicals as Topic*](#)
[Raynaud Disease/diet therapy](#)
[United States](#)

Rysunek 9: Streszczenie artykułu w bazie *MEDLINE*. Źródło: *MEDLINE*.

In Database: Inspec 2010/07-2010/09
 TITLE: Modulation of atomic positions in $\text{CaCu}/\text{sub } x/\text{Mn}/\text{sub } 7-x/\text{O}/\text{sub } 12/$ (x les 0.1)
 AUTHOR: Słotkowski W.; Przeniosłowski R.; Sosnowska I.; Bieringer M.; Margiolaki I.; Suard E.
 AUTHOR AFFILIATION: Słotkowski W. Przeniosłowski R. Sosnowska I. : Institute of Experimental Physics, **University of Warsaw**, Hozdota 69, **Warsaw** 00-681, Poland; Bieringer M. : Department of Chemistry, University of Manitoba, Winnipeg, MB R3T 2N2, Canada; Margiolaki I. : European Synchrotron Radiation Facility, BP220, Grenoble F-38043, France; Suard E. : Institut Laue-Langevin, 6 rue Jules Horowitz, BP-156X, Grenoble 38 042, France
 SOURCE: Acta Crystallographica, Section B-Structural Science. 2009; B65(5): 535-42
 PUBLISHER: Published for the International Union of Crystallography by Munksgaard International Booksellers and Publishers Ltd.
 COUNTRY OF PUBLICATION: Denmark
 RECORD TYPE: Journal-Paper
 DIGITAL OBJECT IDENTIFIER: doi:10.1107/S0108768109025300
 LANGUAGE: English
 ABSTRACT: The modulation of atomic positions in $\text{CaCu}/\text{sub } x/\text{Mn}/\text{sub } 7-x/\text{O}/\text{sub } 12/$ ($x = 0$ and 0.1) was studied using synchrotron radiation powder diffraction below 250 and 220 K, respectively. The copper-rich member $\text{CaCu}/\text{sub } x/\text{Mn}/\text{sub } 7-x/\text{O}/\text{sub } 12/$ ($x = 0.23$) does not show any modulation of the atomic positions at temperatures as low as 10 K. Using low-temperature neutron powder diffraction the modulation of the magnetic moments of Mn ions in $\text{CaCu}/\text{sub } x/\text{Mn}/\text{sub } 7-x/\text{O}/\text{sub } 12/$ ($x = 0, 0.1$ and 0.23) has been investigated. Long-range modulated magnetic ordering in $\text{CaCu}/\text{sub } x/\text{Mn}/\text{sub } 7-x/\text{O}/\text{sub } 12/$ ($x = 0, 0.1$ and 0.23) is observed below 90.4, 89.2 and 78.1 K. $(0, 0, q/\text{sub } p)$ and $(0, 0, q/\text{sub } m)$ are the propagation vectors describing the modulations of the atomic positions and the magnetic moments. For $\text{CaCu}/\text{sub } x/\text{Mn}/\text{sub } 7-x/\text{O}/\text{sub } 12/$ ($x = 0$ and 0.1) the magnetic modulation and atomic modulation lengths are related by a factor of 2 satisfying the relation $(1 - q/\text{sub } p) = 2(1 - q/\text{sub } m)$.

NUMBER OF REFERENCES: 33
 DESCRIPTORS: calcium-compounds; crystal-structure; magnetic-moments; neutron-diffraction; synchrotron-radiation
 IDENTIFIERS: atomic-position-modulation; synchrotron-radiation-powder-diffraction; copper-rich-member; low-temperature-neutron-powder-diffraction; magnetic-moments; long-range-modulated-magnetic-ordering; propagation-vectors; atomic-modulation-lengths; CaMn-sub-7-O-sub-12; CaCu-sub-0.1-Mn-sub-6.9-O-sub-12; CaCu-sub-0.23-Mn-sub-6.77-O-sub-12
 CLASSIFICATIONS: A6160- Crystal-structure-of-specific-inorganic-compounds; A7530C- Magnetic-moments-and-susceptibility-in-magnetically-ordered-materials
 CLASSIFICATION CODES: A6160; A61; A6; A7530C; A75; A7
 SUBFILE: Physics-
 TREATMENT CODE: X (Experimental)
 CHEMICAL INDEXING: CaMn7O12-ss; O12-ss; Mn7-ss; Ca-ss; Mn-ss; O-ss; CaCu0.1Mn6.9O12-ss; Mn6.9-ss; Cu0.1-ss; O12-ss; Ca-ss; Cu-ss; Mn-ss; O-ss; CaCu0.23Mn6.77O12-ss; Mn6.77-ss; Cu0.23-ss; O12-ss; Ca-ss; Cu-ss; Mn-ss; O-ss
 ISSN: 0108-7681
 CODEN: ASBSDK
 PUBLICATION YEAR: 2009
 ACCESSION NUMBER: 11425435
 SORT KEY: 0108-768100B65000005000535
 URLs etc.: DOI: Digital-object-identifier
 COPYRIGHT STATEMENT: Copyright 2010, The Institution of Engineering and Technology
 UPDATE CODE: 2010031

Rysunek 10: Streszczenie artykułu chemicznego w bazie *INSPEC*. Źródło: *INSPEC*.

In Database: Inspec 2011/01 Week 1 2011/01/02-2011/01/08
 TITLE: Millimagnitude photometry for transiting extrasolar planetary candidates. V. Follow-up of 30 OGLE transits. New candidates
 AUTHOR: Pietrukowicz P.; Minniti D.; Diacutecz R.F.; Fernaacutez JM.; Zoccali M.; Gieren W.; Pietrzynacuteski G.; Ruiaacutez MT.; Udalski A.; Szeffert T.; Hempel M.
 AUTHOR AFFILIATION: Pietrukowicz P. Minniti D. Fernaacutez JM. Zoccali M. Hempel M. : Departamento de Astronomia y Astrofisica, Pontificia Universidad Catolica de Chile, Av. Vicunilla MacKenna 4860, Casilla 306, Santiago, Chile; Diacutecz R.F. : CONICET-UBA, Instituto de Astronomia y Fisica del Espacio, Buenos Aires, Argentina; Gieren W. Pietrzynacuteski G. : Departamento de Astronomia, Universidad de Concepcion, Casilla 160-C, Concepcion, Chile; Udalski A. : **Warsaw University** Observatory, Al. Ujazdowskie 4, Warszawa 00-478, Poland; Ruiaacutez MT. : Departamento de Astronomia, Universidad de Chile, Casilla 36-D, Santiago, Chile; Szeffert T. : European Southern Observatory, Casilla 19001, Vitacura, Santiago, Chile
 SOURCE: Astronomy and Astrophysics. 2010; 509: A4 (10 pp.)
 PUBLISHER: EDP Sciences S.A.
 COUNTRY OF PUBLICATION: France
 RECORD TYPE: Journal-Paper
 DIGITAL OBJECT IDENTIFIER: doi:10.1051/0004-6361/200912141
 LANGUAGE: English
 ABSTRACT: Aims. We used VLT/VIMOS images in the V band to obtain light curves of the extrasolar planetary transits OGLE-TR-111 and OGLE-TR-113 and the candidate planetary transits: OGLE-TR-82, OGLE-TR-86, OGLE-TR-91, OGLE-TR-106, OGLE-TR-109, OGLE-TR-110, OGLE-TR-159, OGLE-TR-167, OGLE-TR-170, OGLE-TR-171. Methods. Using difference imaging photometry, we were able to achieve millimagnitude errors in the individual data points. We present the analysis of the data and the light curves by measuring transit amplitudes and ephemerides and by calculating geometrical parameters for some of the systems. Results. We observed nine OGLE objects at the predicted transit moments. Two other transits were shifted in time by a few hours. For another seven objects we expected to observe transits during the VIMOS run, but they were not detected. Conclusions. The stars OGLE-TR-111 and OGLE-TR-113 are probably the only OGLE objects in the observed sample to host planets, with the other objects being very likely eclipsing binaries or multiple systems. In this paper we also report on four new transiting candidates which we have found in the data.

NUMBER OF REFERENCES: 41
 DESCRIPTORS: astronomical-ephemerides; extrasolar-planets; stellar-motion; stellar-photometry; transits
 IDENTIFIERS: millimagnitude-photometry; transiting-extrasolar-planetary-candidates; VLT-VIMOS-images; light-curves; planetary-transits; imaging-photometry; ephemerides; geometrical-parameters; eclipsing-binaries; multiple-systems; OGLE-TR-111; OGLE-TR-113; OGLE-TR-86; OGLE-TR-91; OGLE-TR-106; OGLE-TR-109; OGLE-TR-110; OGLE-TR-159; OGLE-TR-167; OGLE-TR-170; OGLE-TR-171; OGLE-TR-82
 CLASSIFICATIONS: A9780M- Extrasolar-planetary-systems; A9710W- Stellar-space-motions-proper-motions-radial-velocities-and-orbits; A9510G- Eclipses-transits-and-occultations; A9510E- Orbit-determination-and-improvement; A9710R- Stellar-radiation-and-spectra; A9580J- Photographic-region-astronomical-observations
 CLASSIFICATION CODES: A9780M; A97; A9; A9710W; A9510G; A95; A9510E; A9710R; A9580J
 SUBFILE: Physics-
 TREATMENT CODE: X (Experimental)
 ASTRONOMICAL OBJECT: OGLE-TR-111; OGLE-TR-113; OGLE-TR-86; OGLE-TR-91; OGLE-TR-106; OGLE-TR-109; OGLE-TR-110; OGLE-TR-159; OGLE-TR-167; OGLE-TR-170; OGLE-TR-171; OGLE-TR-82
 INDEXING: OGLE-TR-170; OGLE-TR-171; OGLE-TR-82
 ISSN: 0004-6361
 CODEN: AAJAF
 PUBLICATION YEAR: 2010
 ACCESSION NUMBER: 11682790
 SORT KEY: 0004-6361509000000000A400000000(1000000000pp.)
 URLs etc.: DOI: Digital-object-identifier
 COPYRIGHT STATEMENT: Copyright 2011, The Institution of Engineering and Technology
 UPDATE CODE: 2011001

Rysunek 11: Streszczenie artykułu astronomicznego w bazie *INSPEC*. Źródło: *INSPEC*.

6.3. Dane bibliometryczne dostępne w bazach patentów

Kolejną grupę dokumentów, często wykorzystywanych w analizach bibliometrycznych, stanowią patenty. Najczęściej wykorzystywanym źródłem wiedzy jest baza patentów *United States Patent and Trademark Office (USPTO)*, oferująca bogate dane metryczkowe oraz pełny, bezpłatny dostęp do zawartości zgłoszenia patentowego. Baza *USPTO* pozwala również na wyszukiwanie opublikowanych zgłoszeń patentowych, które nie stały się jeszcze przedmiotem decyzji o udzieleniu patentu albo zostały odrzucone na etapie rozpatrywania – ich analizy pozwalają obserwować zainteresowania badawcze poszczególnych instytucji, a zgłoszenia patentowe udostępniane są na wiele miesięcy przed udzieleniem patentu, co ułatwia identyfikację najnowszych tendencji w badaniach stosowanych. Firma *Google* podpisała umowę z *USPTO*, na podstawie której udostępnia zawartość tej samej bazy patentów w swoim dedykowanym serwisie internetowym, oferując zbliżone możliwości wyszukiwania przy większej prędkości działania bazy. Umieszczone na kolejnych stronach rysunki prezentują fragmenty stron z opisem patentów amerykańskich w serwisach *USPTO* i firmy *Google*. Dane metryczkowe patentów obejmują m.in. streszczenia, wskazanie wynalazcy (ang. *inventor*, zgodnie z prawem amerykańskim może być nim tylko osoba prywatna) i właściciela (ang. *assignee*, którym często jest firma, zatrudniająca wynalazcę), daty zgłoszenia i udzielenia patentu oraz numery klasyfikacji patentowej, pozwalające na identyfikację grup rozwiązań technologicznych. Dokumenty patentowe zawierają też odwołania do wcześniejszych patentów i publikacji, a bazy pozwalają dodatkowo zidentyfikować wszystkie patenty, które cytują interesujący nas patent.

United States Patent	5,440,481	
Kostoff, et al.	August 8, 1995	
System and method for <i>database tomography</i>		
Abstract		
A Process for analyzing full-text is provided for identifying often-repea, high user interest, word phrases in a database. Often-repeated, high user interest, word phrases are defined as pervasive theme areas (PTAs). The process also allows the relationship defined as connectivity among the various PTAs to be identified. In addition, phrases that are in proximity to the PTAs and which are strongly supportive of the PTAs are identified. Numerical indices, figure of merit, and user defined thresholds are used to quantify relations between PTAs and among PTAs and phrases.		
Inventors: Kostoff; Ronald N. (Falls Church, VA), Miles; David L. (Ridgecrest, CA), Eberhart; Henry J. (Ridgecrest, CA)		
Assignee: The United States of America as represented by the Secretary of the Navy (Washington, DC)		
Appl. No.: 07/967,341		
Filed: October 28, 1992		
Current U.S. Class:	1/1 ; 707/999.005; 707/E17.084	
Current International Class:	G06F 17/30 (20060101); G06F 17/27 (20060101); G06F 017/20 (); G06F 017/30 ()	
Field of Search:	364/419.08,419.07,419.13,419.19 395/600	
References Cited [Referenced By]		
U.S. Patent Documents		
4839853	June 1989	Deerwester et al.
4849898	July 1989	Adi
4942526	July 1990	Okajima et al.
4992972	February 1991	Brooks et al.
4994967	February 1991	Asakawa

Rysunek 12: Opis patentu w bazie *USPTO*. Źródło: *USPTO*.

Method for data and text mining and literature-based discovery Ronald N. Kostoff

Text searching is achieved by techniques including phrase frequency analysis and phrase-c matrix analysis is also advantageously applied to select high technical content phrases to be The described...

Inventor: Ronald N. Kostoff
Assignee: The United States of America as represented by the Secretary of the Navy
Primary Examiner: Jean M. Corriellus
Attorneys: Barry A. Edelberg, Charles H. Harris, Thomas E. McDonnell

Read this patent
 Download PDF

U.S. Classification
 707/3; 707/4; 707/6; 707/10

International Classification
 G06F017/30

[View patent at USPTO](#)

Citations

Patent Number	Title
4744050	Method for automatically registering frequently used phrases
5440481	System and method for database tomography
5745602	Automatic method of selecting multi-word key phrases from a document

Rysunek 13: Opis patentu w bazie *Google Patents*. Źródło: *Google Patents*.

Serwisy *USPTO* i *Google* gromadzą i udostępniają wyłącznie patenty, które udzielone zostały przez Urząd Patentowy Stanów Zjednoczonych. Nie wszystkie istotne wynalazki są zgłaszane do *USPTO*, choć wielu badaczy zajmujących się bibliometrią przyjmuje, że zagraniczne wynalazki o największym potencjale komercyjnym będą podlegały takiemu zgłoszeniu. Baza *USPTO* z oczywistych przyczyn zawiera nadreprezentację patentów, udzielonych wynalazcom amerykańskim – jednak podmioty zagraniczne z rozwiniętych gospodarek zgłaszają swoje najważniejsze wynalazki w Stanach Zjednoczonych, dążąc do zapewnienia ochrony na terytorium tego kraju.

W warunkach polskich można podjąć próbę wykorzystania bazy Urzędu Patentowego RP w celu analizy wynalazków chronionych na terenie Polski (Rysunek 14), jednak baza zawiera ubogi zestaw danych metryczkowych i ograniczone możliwości wyszukiwania.

(54) Mechanizm wychyłu ramienia kombajnu górniczego

Kategoria, numer i data zgłoszenia:
WYN: (21) 358109, (22) 02-01-2003

Kategoria i numer ochrony:
WYN: (11) 199858

<<Powrót

Osoby				
Rodzaj osoby	Nazwa	Miasto	Adres	Kraj
(72) Twórca	KNYĆ JÓZEF	KATOWICE		PL
(72) Twórca	SOK HENRYK	CZECHOWICE-DZIEDZICE		PL
(72) Twórca	STYRSKI BOGUSŁAW	TYCHY		PL
(72) Twórca	PŁONKA RUDOLF	GLIWICE		PL
(73) Zgłaszający/Uprawniony	FABRYKA MASZYN FAMUR SPÓŁKA AKCYJNA	KATOWICE		PL
Pełnomocnik	Kułaczk Marek			

Klasyfikacje		
Typ klasyfikacji	Symbol	Edycja
(51) MKP	E21C31/08	8

Publikacje		
Typ	Data publikacji	Numer
BUP	12-07-2004	14/2004
WUP	28-11-2008	11/2008

Decyzje	
Data	Typ decyzji
13-05-2008	Prawo wyłączne udzielone

Rysunek 14: Opis patentu w bazie *UPRP*. Źródło: *UPRP*.

W przypadku analiz technologii na rynku europejskim, wskazane wydaje się korzystanie z bazy Europejskiego Urzędu Patentowego (*European Patent Office, EPO*), oferowanej bezpłatnie pod marką *esp@cenet* i gromadzącej dane o patentach europejskich, zgłaszanych w trybie odrębnym od krajowego i chronionych w 38 krajach członkowskich Europejskiej Organizacji Patentowej. Baza udostępnia również dane patentów z innych, wybranych obszarów geograficznych (por. Rysunek 15).

METHOD AND PROCESS FOR PRODUCING YOUTHFUL-APPEARING, SMALL-PORED, AND SMOOTH SKIN

Bibliographic data	Description	Claims	Mosaics	Original document	INPADOC legal status
<p>Publication number: YUP60904 (A) Publication date: 2006-05-25 Inventor(s): PARSONS DIANA J [US] + Applicant(s): PARSONS DIANA J [US] + Classification: - international: A61B17/00; A61B17/50; A61B18/00; A61B18/20; A61K31/203; A61K31/74; A61N5/06; A61B; A61Q17/00; (IPC1-7): A61B17/00; A61B17/50; A61K31/74; A61K7/42 - European: A61B18/20H; A61K31/203; A61K8/67C; A61Q19/08 Application number: YUP60904 20020828 Priority number(s): US20020047335 20020114; WO2002US27570 20020828</p> <p>View INPADOC patent family View list of citing documents</p> <p style="text-align: right;">Report a data error here</p>					
<p>Abstract of YU P60904 (A) Translate this text</p> <p>The method and process for producing youthful appearing, small-pored, and smooth skin comprises pretreating the skin topically with retinoic acid, applying a contaminate in oil to the skin, and removing the contaminate by exploding or rupturing it with laser light. When treating the face, removing the contaminate requires only about 4 minutes. The laser light itself does not injure the skin but the rupture of the contaminant produces thermal injury in the high dermis leaving the epidermis intact. Repeated intermittent laser treatments and applications of topical retinoic acid produces the desired results in the skin indefinitely as long as the treatments and applications are continued.</p> <p style="text-align: center;">Data supplied from the <i>espacenet</i> database — Worldwide</p>					

Rysunek 15: Opis patentu w bazie *EPO esp@cenet*. Źródło: *Esp@cenet*.

6.4. Zróżnicowanie aktywności badawczej poszczególnych krajów

Analizy bibliometryczne powinny uwzględniać specyfikę objętych analizami krajów i obszarów badawczych. W wielu przypadkach decyzja o koncentracji na określonych krajach wpływa na wybór odpowiedniej bazy – dzieje się tak m.in. w przypadku analizy patentów. Tabela 7 prezentuje zestawienie aktywności patentowej wynalazców z poszczególnych krajów w urzędach europejskim i amerykańskim. W zależności od tematyki analiz, należy rozważyć, która z baz okaże się najlepiej odzwierciedlać interesujące nas tendencje.

Kraj	Zgłoszenia patentowe do EPO (2006)	Patenty udzielone przez USPTO (2003)
EU-27	52.612	15.988
Stany Zjednoczone	31.403	86.574
Niemcy	22.675	7.258
Japonia	19.990	29.598
Francja	7.891	2.085
Włochy	4.736	1.226
Wielka Brytania	4.691	1.925
Szwajcaria	3.024	809

Kraj	Zgłoszenia patentowe do EPO (2006)	Patenty udzielone przez USPTO (2003)
Holandia	2.900	927
Szwecja	2.200	546
Austria	1.451	403
Belgia	1.365	394
Hiszpania	1.333	249
Finlandia	1.190	425
Dania	1.011	219
Norwegia	457	127
Irlandia	251	117
Turcja	154	18
Portugalia	129	13
Polska	122	30
Grecja	116	25
Luksemburg	107	29
Słowenia	102	19
Czechy	97	42
Węgry	96	38
Słowacja	30	6
Rumunia	29	9
Chorwacja	27	25
Islandia	25	18
Liechtenstein	24	13
Łotwa	22	3
Bulgaria	20	3
Cypr	17	2
Malta	13	0
Litwa	11	12
Estonia	6	1

Tabela 7: Liczba zgłoszeń patentowych do *EPO* (2006) i patentów udzielonych przez *USPTO* (2006) podmiotom z poszczególnych krajów. Źródło: Eurostat (2010: 608).

Warto pamiętać również o zróżnicowaniu systemów ochrony własności przemysłowej w poszczególnych krajach, gdyż może mieć ono istotny wpływ na możliwości prowadzenia analiz. Różnice dotyczą m.in. możliwości patentowania programów komputerowych i organizmów żywych, co wpływa na przydatność analiz patentów do obserwacji rozwoju dyscyplin takich jak informatyka czy biotechnologie.

Kraj	Liczba publikacji w <i>Web of Science</i>
Stany Zjednoczone	209.695
Chiny	56.806
Japonia	52.896
Wielka Brytania	47.121
Niemcy	44.408
Francja	30.740
Kanada	27.799
Włochy	26.544
Hiszpania	20.981

Kraj	Liczba publikacji w <i>Web of Science</i>
Korea Południowa	18.467
Indie	18.194
Australia	17.831
Holandia	14.210
Rosja	13.953
Tajwan	12.742
Brazylia	11.885
Szwecja	9.914
Szwajcaria	9.190
Turcja	8.638
Polska	7.136
Belgia	7.071
Izrael	6.622
Dania	5.236
Finlandia	4.989
Grecja	4.980
Świat	758.142

Tabela 8: Artykuły naukowe opublikowane w roku 2007, zgromadzone w bazie *Web of Science*.
Źródło: National Science Board (2010: 5-31).

Tabela 8 prezentuje zestawienie publikacji naukowych, zestawiając aktywność badawczy z poszczególnych krajów i wskazując na bogactwo danych bibliometrycznych, które mogą być podstawą do prowadzenia opisywanych w niniejszej książce analiz.

7. Pytania analityczne w bibliometrii

Analizy bibliometryczne mają zwykle na celu znalezienie odpowiedzi na konkretne pytania analityczne (pytania badawcze), które powinny zostać sformułowane zanim podjęta zostanie decyzja o wyborze określonych technik analizy oraz baz danych. Tabela 9 oferuje zestawienie najbardziej typowych pytań, które mogą stanowić punkt wyjścia do prowadzenia projektu bibliometrycznego, wraz z przykładami ich zastosowań.

Lista tych pytań będzie w dalszych rozdziałach książki wykorzystywana jako podstawa do porównań konkretnych technik analitycznych, baz oraz narzędzi, wspierających analizy.

Warto zauważyć zróżnicowane wymogi, dotyczące wiedzy i umiejętności osób, prowadzących analizy bibliometryczne. Niektóre z podanych poniżej przykładów analiz nie wymagają dogłębnej znajomości analizowanej dyscypliny naukowej czy obszaru technologii, a więc mogą być wykonywane przez osoby nie będące ekspertami dziedzinowymi. Gromadzenie odpowiedzi na poszczególne pytania różnić się będzie też stopniem trudności przetwarzania danych i jak wynika z zaprezentowanego zestawienia, wiele analiz może być prowadzonych samodzielnie przez pracowników zainteresowanych instytucji, bez potrzeby korzystania z zewnętrznych doradców. W XXI wieku dostępne są łatwe w obsłudze interfejsy użytkownika baz danych oraz oprogramowanie, co sprawia, że większość typowych analiz nie wymaga specjalistycznej wiedzy informatycznej, ani dogłębnej znajomości analizowanego obszaru badawczego. W literaturze jako przejaw dobrych praktyk prezentowane są przypadki, gdy analizy bibliometryczne prowadzone są przez osoby, które bezpośrednio będą korzystać z ich rezultatów (Kostoff, Geisler 2004: 8). Obsługa dostępnych narzędzi analitycznych może być opanowana przez każdego użytkownika komputerów, a udział w prowadzeniu analiz pozwala decydentom sprawdzenie dodatkowych wariantów, które mogą być istotne przy podejmowaniu decyzji związanych z planowaniem prac badawczo-rozwojowych.

Pytanie analityczne	Wyjaśnienie	Przykłady	Znajomość analizowanej dyscypliny	Stopień trudności przetwarzania danych
Co?	Identyfikacja obszarów badań i technologii w zbiorze danych bibliometrycznych dla określonego kraju, regionu, ośrodka badawczego lub dla określonych lat	<i>Co stanowi przedmiot zainteresowań badaczy z Malezji? Czym zajmują się badacze z Wydziału Fizyki Uniwersytetu Warszawskiego? Jakie tematy badań interesują naukowców z województwa opolskiego? Czego dotyczyły polskie badania w obszarze nauk medycznych, których wyniki opublikowano w roku 2008?</i>	+	*
Co dokładnie?	Identyfikacja podobszarów badawczych w zbiorze danych bibliometrycznych, odzwierciedlających wybrany obszar badań i technologii	<i>Jakie podobszary badawcze można wyróżnić w ramach embriologii? Jakimi dokładnie zagadnieniami zajmują się polscy naukowcy, badający nanotechnologie?</i>	++	**
Kto?	Identyfikacja badaczy/wynalazców i zespołów badawczych, prowadzących badania na określony temat	<i>Kto na świecie/w Polsce/w Krakowie prowadzi badania nad metodami eksploatacji gazu łupkowego?</i>	+	*
Gdzie?	Identyfikacja afiliacji, w których prowadzone są badania na określony temat (poziom kraju, miasta, organizacji)	<i>Badacze z jakich krajów/ośrodków akademickich prowadzą badania nad wykorzystaniem komórek macierzystych w medycynie?</i>	+	*
Kiedy?	Określenie momentu publikacji, zgłoszenia patentowego lub uzyskania patentu, jak również okresów, w których prowadzono badania na określony temat	<i>Od kiedy publikowane są wyniki badań nad wykorzystaniem azotku galu jako półprzewodnikowa? Kiedy zgłoszony został do opatentowania pierwszy wynalazek, dotyczący konstrukcji papierosa elektronicznego? Kiedy firma Google uzyskała swój pierwszy patent?</i>	+	*
Z kim?	Analiza współautorstwa publikacji i wynalazków (co-	<i>Kto współpracował z prof. Janem Lubińskim w</i>	+	**

Pytanie analityczne	Wyjaśnienie	Przykłady	Znajomość analizowanej dyscypliny	Stopień trudności przetwarzania danych
	<i>authorship</i>) poprzez identyfikację osób, które wspólnie prowadziły badania	<i>badaniach nad genetyczną diagnostyką nowotworów i kto był współwynalazcą zgodnie ze zgłoszeniami patentowymi?</i>		
Z jaką organizacją?	Analiza współautorstwa publikacji i wynalazków (<i>co-authorship</i>) poprzez identyfikację organizacji, zatrudniających osoby, które wspólnie prowadziły badania	<i>Z jakimi ośrodkami naukowymi i firmami współpracował wynalazca dr Robert Dwiliński w badaniach nad metodą wytwarzania kryształów azotku galu?</i>	+	**
Ile?	Zliczanie publikacji, zgłoszeń patentowych lub udzielonych patentów dla wybranego obszaru lub podobszaru badań i technologii (możliwe jest także sporządzanie zestawień porównawczych liczby publikacji w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją?)	<i>Jak popularnym obszarem badań w Polsce jest protetyka? Ile publikacji naukowych autorów zatrudnionych w Polsce dotyczy robotów mobilnych?</i>	+	*
Ilu?	Zestawienie liczby autorów publikacji lub zgłoszeń patentowych w danym obszarze	<i>Ilu badaczy w Polsce przygotowuje publikacje na temat katalizy przeniesienia międzyfazowego?</i>	+	**
Jak zmienia się w czasie?	Analiza szeregów czasowych, odzwierciedlających liczby publikacji lub patentów w wybranych obszarach badań i technologii (możliwe jest odzwierciedlenie zmian w czasie w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? ile? – czyli prezentacja historycznego rozwoju obszaru badań)	<i>Jak na przestrzeni lat zmieniała się polska aktywność badawcza w zakresie biotechnologii?</i>	+	*
Jak będzie się rozwijać?	Prognoza krótkookresowa, oparta na ekstrapolacji szeregów czasowych (możliwe jest przygotowanie prognoz w powiązaniu z odpowiedziami na pytania:	<i>Jak prawdopodobnie kształtować się będzie w najbliższych 3 latach liczba polskich publikacji w obszarze nauk matematycznych?</i>	+	***

Pytanie analityczne	Wyjaśnienie	Przykłady	Znajomość analizowanej dyscypliny	Stopień trudności przetwarzania danych
	co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? pod warunkiem posiadania dostatecznie długich serii danych historycznych)			
Jak popularne w porównaniu do?	Porównanie liczb publikacji, zgłoszeń patentowych lub patentów w kilku obszarach badawczych	<i>Jakie specjalności są najpopularniejszymi w obszarze inżynierii środowiskowej w oparciu o liczbę publikacji? Czy wśród polskich fizyków, większą popularnością cieszy się fizyka kwantowa czy astrofizyka?</i>	+	*
Jakie są nisze?	Identyfikacja podobszarów badawczych, które nadal cieszą się niewielkim zainteresowaniem badaczy (w oparciu o liczbę publikacji, zgłoszeń patentowych lub patentów w danym podobszarze, porównaną z ich globalną liczbą dla szerokiego obszaru badań)	<i>Jakie nisze badawcze istnieją współcześnie w obszarze chemii analitycznej?</i>	++	*
Jakie są nowe popularne tematy?	Identyfikacja nowych, cieszących się narastającą popularnością tematów badawczych (<i>hot topics</i>), w tym dynamicznie rozwijających się, nowych technologii (<i>emerging technologies</i>)	<i>Jakie nowe zagadnienia cieszą się szczególną popularnością wśród badaczy, zajmujących się tematyką programowania baz danych? Jaka nowa grupa rozwiązań technologicznych cieszy się szczególnym zainteresowaniem wynalazców w obszarze technologii przetwarzania odpadów niebezpiecznych?</i>	++	**
Z czym związane?	Analiza pokrewnych obszarów badań i tematów badawczych, które równocześnie opisywane są w publikacjach lub dokumentach patentowych – na poziomie współwystępowania specjalistycznych pojęć w	<i>Jakie obszary badawcze są powiązane z badaniami nad rozwojem niebieskiego lasera półprzewodnikowego? Z jakich pokrewnych obszarów badawczych korzystają naukowcy, poszukujący sposobów</i>	++	**

Pytanie analityczne	Wyjaśnienie	Przykłady	Znajomość analizowanej dyscypliny	Stopień trudności przetwarzania danych
	streszczeniach i tytułach (<i>co-words</i>) lub w słowach kluczowych i kodach klasyfikacyjnych (<i>co-classification</i>)	<i>leczenia choroby Raynauda?</i>		
W jaki sposób związane?	Analiza szczegółowych podejść badawczych, technik, środków i technologii, wykorzystywanych przez zidentyfikowane wcześniej obszary badań i tematy badawcze, pokrewne wobec podstawowego obszaru analiz	<i>Jakie grupy związków chemicznych wykorzystywane są w badaniach, dotyczących niebieskiego lasera półprzewodnikowego?</i> <i>Jakie techniki epitaksjalne wykorzystywane są w badaniach, dotyczących niebieskiego lasera półprzewodnikowego?</i>	++	**
Co uzupełnia?	Identyfikacja komplementarnych obszarów wiedzy naukowej, komplementarnych umiejętności badawczych lub komplementarnych technologii, wykorzystywanych przy prowadzeniu badań w danych obszarze	<i>Jaką dodatkową wiedzę powinien posiadać zespół badawczy, zainteresowany pracą nad metodami rekultywacji odpadów ropopochodnych?</i> <i>Jakie komplementarne technologie należy licencjonować, by móc stworzyć własny wariant ogniwa paliwowego?</i>	++	**
Jakich wariantów nie analizowano?	Identyfikacja możliwych nowych kierunków badań i sposobów rozwiązania problemów naukowych i praktycznych w oparciu o analogie, techniki kreatywnego rozwiązywania problemów i ocenę potencjału ewolucyjnego	<i>Jakie mogą być hipotetyczne, przyszłe warianty określonej technologii?</i> <i>W którym kierunku prawdopodobnie ewoluować będzie określona technologia?</i> <i>W jakich (nieoczywistych) obszarach badań można szukać potencjalnych rozwiązań określonego problemu naukowego/praktycznego?</i>	+++	****
Kto uzupełnia?	Identyfikacja osób i organizacji, posiadających komplementarne obszary wiedzy naukowej, komplementarne umiejętności badawcze lub komplementarne technologie, wykorzystywane przy	<i>Z jakimi ośrodkami badawczymi można podjąć współpracę przy danym projekcie badawczym, w oparciu o ich unikalną wiedzę, uzupełniającą nasze doświadczenia?</i>	++	**

Pytanie analityczne	Wyjaśnienie	Przykłady	Znajomość analizowanej dyscypliny	Stopień trudności przetwarzania danych
	prowadzeniu badań w danym obszarze			
Kto może być dostawcą?	Identyfikacja osób i organizacji, które mogą być podwykonawcami określonych zleceń badawczych ze względu na dotychczasowy dorobek naukowy	<i>Kto może być kompetentnym podwykonawcą konkretnych prac badawczo-rozwojowych?</i>	++	**
Kto może być odbiorcą?	Identyfikacja osób i organizacji, które mogą być zleceniodawcami projektów badawczo-rozwojowych lub nabywcami/licencjobiorcami ich rezultatów ze względu na dotychczasowe zainteresowania i potrzeby badawcze	<i>Jakie firmy zajmują się pracami badawczo-rozwojowymi, w których wykorzystywany jest polioksymetylen?</i>	++	**
Kto jest konkurentem?	Identyfikacja osób i organizacji, prowadzących badania w analogicznych, szczegółowych obszarach badawczych, przy wykorzystaniu zbliżonych metod (podmioty zidentyfikowane jako konkurenci mogą być też potencjalnymi partnerami w konsorcjach badawczych – zwłaszcza wtedy, gdy niezbędne jest połączenie zbliżonych zasobów dla zwiększenia skali prac badawczych)	<i>Jakie zespoły badawcze prowadzą badania dotyczące zagadnień zbliżonych do tematu naszego projektu badawczego?</i>	++	**
Kto może być konkurentem?	Identyfikacja osób i organizacji, które mogą podjąć badania w analogicznych, szczegółowych obszarach badawczych, przy wykorzystaniu zbliżonych metod, a wskazuje na to ich profil badawczy, dotychczasowe zainteresowania i analizy statystyczne (zwłaszcza techniki analizy czynnikowej, pozwalające na identyfikację „ukrytych” podobieństw i zależności w dużych zbiorach danych)	<i>Jakie zespoły badawcze prowadzą badania dotyczące zagadnień pokrewnych do tematu naszego projektu badawczego?</i>	+++	***

Pytanie analityczne	Wyjaśnienie	Przykłady	Znajomość analizowanej dyscypliny	Stopień trudności przetwarzania danych
Kto może być pracownikiem?	Identyfikacja osób, które prowadzą lub mogą prowadzić badania w określonym obszarze i zgromadziły zadowalający dorobek naukowy	<i>Jacy badacze zajmowali się interesującym nas, szczegółowo określonym podobszarem badań?</i>	++	**
Kto może być sukcesorem?	Identyfikacja osób, które prowadziły badania wspólnie z określonym naukowcem (<i>co-authorship</i>) i mogą w przyszłości przejąć obowiązki w zakresie zarządzania projektami badawczymi	<i>Kto był regularnym współautorem publikacji naukowych prof. Mariusza Ratajczaka, dotyczących właściwości szpiku kostnego i krwi pepowinowej?</i>	++	**
Jakie X ma strategie badawcze?	Analiza tematyki badawczej, mechanizmów współpracy autorskiej (<i>co-authorship</i>), kluczowych badaczy i zespołów badawczych, partnerów zewnętrznych i zmian w powyższych obszarach dla wybranej organizacji	<i>Jaki jest profil badawczy Collegium Medicum Uniwersytetu Jagiellońskiego?</i>	++	**
Jak ważne?	Zestawienie liczby cytowań dla pojedynczych publikacji lub patentów	<i>Ile razy cytowany był w latach 1996-2010 tekst zespołu badaczy z Instytutu Wysokich Ciśnień PAN pt. „Lattice parameters of gallium nitride”, opublikowany w „Applied Physics Letters”, 1996, vol. 69, no. 1?</i> <i>Ile razy cytowany był w latach 1996-2010 patent nr 5637531, udzielony Instytutowi Wysokich Ciśnień PAN przez USPTO w 1996 roku?</i>	+	*
Jak łącznie ważne?	Zestawienie zagregowanej liczby cytowań dla grupy publikacji lub patentów (np. pochodzących od autorów z jednej organizacji, dotyczących jednego obszaru badawczego)	<i>Ile razy łącznie cytowane były patenty, udzielone Instytutowi Wysokich Ciśnień PAN przez USPTO?</i>	+	**
Kto jest ważny?	Porównania liczby cytowań w dorobku danych badaczy oraz obliczanie wskaźników dotyczących cytowań (np. indeksów h Hirscha)	<i>Który z badaczy w zespole, zajmującym się badaniami nad azotkiem galu i niebieskim laserem półprzewodnikowym</i>	+	*

Pytanie analityczne	Wyjaśnienie	Przykłady	Znajomość analizowanej dyscypliny	Stopień trudności przetwarzania danych
		<i>Instytutu Wysokich Ciśnień PAN jest autorem najczęściej cytowanych publikacji? Którego z pracowników Wydziału Chemii Uniwersytetu Warszawskiego charakteryzuje najwyższy poziom indeksu h Hirscha?</i>		
Do jakiej wiedzy nawiązuje?	Identyfikacja źródeł wiedzy, ujawnionych w cytowaniach, zawartych w publikacjach naukowych i zgłoszeniach patentowych (możliwe dalsze analizy zidentyfikowanego zbioru źródeł wiedzy w oparciu o pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? ile?)	<i>Do osiągnięć badawczych jakich organizacji odwołują się patenty, udzielone przez USPTO firmie Ammono?</i>	++	**
Do czyich badań nawiązuje?	Identyfikacja naukowców i wynalazców, do prac których odwołują się cytowania w publikacjach naukowych i zgłoszeniach patentowych	<i>Kto jest autorem najczęściej cytowanym przez publikacje dotyczące morfogenezy tkanek?</i>	++	**

Tabela 9. Zestawienie pytań analitycznych w projektach bibliometrycznych. Źródło: opracowanie własne. Legenda:

Symbole użyte w kolumnie „Znajomość analizowanej dyscypliny”:

- + - odpowiedzi na pytania wymagają jedynie ogólnej orientacji w specyfice i terminologii analizowanej dyscypliny, mogą być wykonywane przez niespecjalistów pod warunkiem wstępnych konsultacji z ekspertami dziedzinowymi i/lub wstępnych studiów literaturowych,
- ++ - odpowiedzi na pytania wymagają wsparcia ze strony ekspertów dziedzinowych na etapie analizy i interpretacji danych, często także bezpośredniego zaangażowania ekspertów w prace analityczne,
- +++ - odpowiedzi na pytania mogą być udzielone tylko przez ekspertów dziedzinowych, dostatecznie dobrze rozumiejących specyfikę poddanego analizom obszaru badań naukowych i technologicznych.

Symbole użyte w kolumnie „Stopień trudności przetwarzania danych”:

- - odpowiedzi na pytania nie wymagają posiadania specjalistycznego oprogramowania, mogą być udzielone na podstawie kwerend (poleceń wyszukiwania) w bazach danych,
- ** - odpowiedzi na pytania wymagają dodatkowego przetwarzania podzbiorów danych, pobranych z bibliometrycznych baz danych, przy wykorzystaniu specjalistycznego oprogramowania,
- *** - odpowiedzi na pytania wymagają statystycznej obróbki podzbiorów danych, pobranych z bibliometrycznych baz danych, przy wykorzystaniu specjalistycznego oprogramowania w oparciu o wybór odpowiednich algorytmów,
- **** - odpowiedzi na pytania wymagają jakościowej analizy podzbiorów danych, pobranych z bibliometrycznych baz danych w oparciu o wiedzę ekspercką z analizowanej dyscypliny oraz sprawność w wykorzystywaniu zaawansowanych technik analitycznych odwołujących się do *LBD*, *KBMA* i/lub *TRIZ*.

Warto zróżnicować analizy bibliometryczne w oparciu o pożądaną dogłębność czy stopień szczególności. Porter i in. (1991: 128-129) proponują podzielić monitorowane obszary naukowo-technologiczne na obszary „zimne”, „ciepłe” i „gorące” w zależności od obecnej ich znajomości i potrzeby precyzyjnych informacji. W obszarach „zimnych” wystarczy zdobycie ogólnej orientacji dotyczącej ich funkcjonowania, podczas gdy obszary „gorące” są zwykle na tyle ważne dla prowadzącej analizy organizacji, że niezbędne okazuje się dogłębne rozpoznanie podmiotów prowadzących działalność B+R oraz specyfiki podejmowanych prac badawczych.

8. Podstawowe techniki analityczne w bibliometrii

8.1. Podstawowe techniki bibliometryczne

Niniejszy rozdział prezentuje najczęściej wykorzystywane, a jednocześnie najmniej zaawansowane techniki analityczne, opierające się na danych pobranych ze specjalistycznych baz.

Najprostszą formą analiz bibliometrycznych jest opracowywanie zestawień liczby publikacji, wystąpień konferencyjnych, patentów lub zgłoszeń patentowych. Proste wskaźniki liczbowe, uzyskane w wyniku doprecyzowanych poleceń wyszukiwania w bazach, pozwalają na zgromadzenie serii danych, pozwalających m.in. na: porównanie skali aktywności badawczej i/lub wynalazczej w wybranym obszarze dla poszczególnych krajów, instytucji i osób, jak również wygenerowanie zestawień dla określonych lat. Omawiane analizy nie wymagają posiadania licencji specjalistycznego oprogramowania bibliometrycznego, a jedynie dostępu do jednej z baz publikacji lub patentów, wykonywania kolejnych poleceń wyszukiwania i spisywania liczby uzyskanych wyników.

Taki scenariusz analizy nie wiąże się z koniecznością poniesienia dodatkowych kosztów, może być łatwo wykonany nawet przez osobę nie znającą specjalistycznych technik analizy bibliometrycznych czy statystycznych. Jedynym wymogiem jest posiadanie dostępu do jednej z baz – w odniesieniu do baz publikacji i wystąpień konferencyjnych, zwykle pozwalają na to biblioteki uczelni publicznych także gościom spoza uczelni, a bazy patentów i zgłoszeń patentowych są ogólnodostępne w Internecie.

Podstawowe analizy bibliometryczne posiadają jednak istotne ograniczenia, które podważają rzetelność uzyskanych wyników badań:

- dane pobrane z baz nie są poddane obróbce, mającej na celu oczyszczenie zbioru danych z pozycji, które znalazły się tam w wyniku błędu klasyfikacyjnego (np. literówki, niewłaściwie przypisany obszar badań, kraj, instytucja, duplikaty dla niektórych pozycji),
- algorytmy prezentacji wyników wyszukiwania niektórych baz danych okazują się zawodne (np. informacja o łącznej liczbie rezultatów, zaprezentowana osobie wyszukującej po wykonaniu polecenia wyszukiwania, nie odpowiada rzeczywistej liczbie uzyskanych rezultatów),
- uzyskane dane mogą być analizowane w oparciu o jedną zmienną (lub – pod warunkiem spisywania dodatkowych danych – dwie zmienne), co nie pozwala na uchwycenie wielu istotnych aspektów rozwoju badań naukowych i technologii.

Z powyższych względów najprostsze analizy bibliometryczne, prowadzone bez wykorzystania specjalistycznego oprogramowania i zaawansowanych technik analitycznych, nie są rekomendowane jako źródło wiedzy na temat działalności badawczo-rozwojowej. Ich przydatność ogranicza się do gromadzenia ogólnych danych na temat aktywności badawczej krajów lub liczby publikacji i patentów w szeroko określonych obszarach nauki i technologii. Potrzeby informacyjne, które miałyby zaspokoić rozwiązania proponowane w niniejszej książce, wymagają bardziej zaawansowanych prac analitycznych. Jednocześnie warto

pamiętać, że analizy bardziej zaawansowane, które będą opisane w kolejnych podrozdziałach, korzystają z list rezultatów, wygenerowanych przy wykorzystaniu tych samych wyszukiwań w bazach danych – istotną różnicą jest jedynie fakt pobrania tych list rezultatów, zapisania na dysku lokalnym komputera i dalszego przetwarzania przy wykorzystaniu dedykowanych programów komputerowych. Z tego powodu poniżej omówione zostaną szczegółowo sposoby wyszukiwania oraz podstawowych analiz danych, które będą przydatne także przy bardziej zaawansowanych projektach analitycznych.

8.2. Tworzenie zestawu danych bibliometrycznych do dalszych analiz

Proces wyszukiwania rozpoczyna się od określenia zakresu danych, które mają być podstawą dla dalszych analiz. Ten zakres danych znajdzie odzwierciedlenie w poleceniu wyszukiwania, a zebrane dane powinny pozwalać na udzielenie odpowiedzi na sprecyzowane wcześniej pytania badawcze.

Bazy bibliograficzne oferują możliwości wyszukiwania w oparciu o dwa mechanizmy:

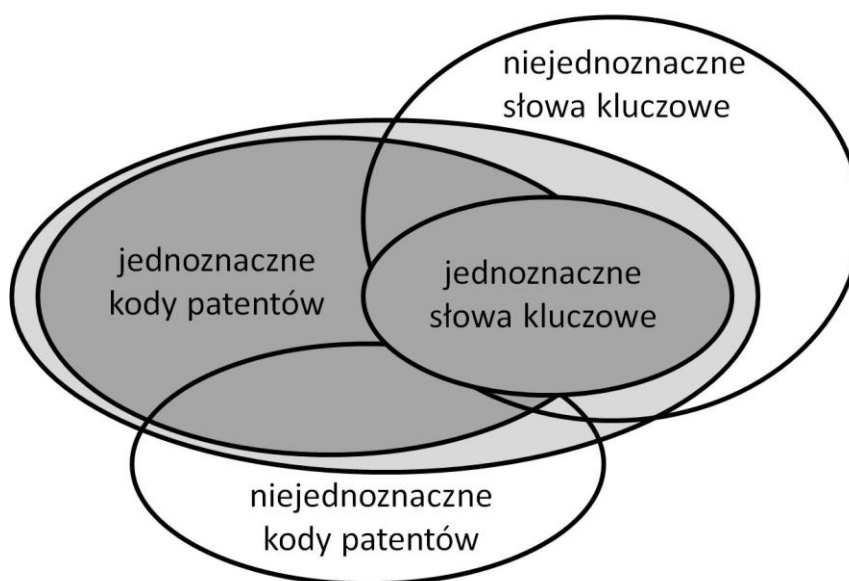
- wyszukiwanie pełnotekstowe – czyli wyszukiwanie w oparciu o słowa lub kombinacje słów, występujących w treści danej pozycji (w zależności od bazy, takie wyszukiwanie dotyczy zawartości streszczenia lub pełnego tekstu),
- wyszukiwanie w oparciu o indeksy bazy – pozwalające na określenie szczegółowych kryteriów, dotyczących np. lat, zakresu geograficznego lub obszaru badawczego.

Szczególnie przydatne okazuje się wykorzystanie indeksów, które umożliwiają zawężenie list rezultatów w oparciu o m.in. następujące dane:

- rok publikacji,
- autorów,
- ich afiliacje,
- kody, opisujące zawartość tekstu (przypisywane przez wyspecjalizowanych redaktorów, pracujących dla wydawcy bazy, w oparciu o przyjętą przez wydawcę taksonomię,
- słowa kluczowe (wybierane przez autorów publikowanych tekstów, co nie gwarantuje spójności klasyfikacji ani rzetelnego odzwierciedlenia zawartości indeksowanych publikacji – albo przez redaktorów bazy, co gwarantuje porównywalność wyników),
- dane identyfikujące pismo akademickie lub zbiór wystąpień konferencyjnych, z którego pochodzi dany tekst.

Bazy patentów, stosowane do analiz bibliometrycznych, opisują zwykle autorów wynalazku oraz ich miejsce pracy. Warto pamiętać, że wynalazca (ang. *inventor*) nie musi być jednocześnie właścicielem patentu (ang. *assignee*) – jest to szczególnie widoczne w przypadku dużych firm międzynarodowych. Bazy patentowe stosują własne schematy klasyfikacyjne, w które wpisywane są poszczególne wynalazki na podstawie deklaracji zgłaszających. Praktyka gospodarcza wskazuje, że firmy zgłaszające patenty czasami świadomie dążą do przypisania własnego wynalazku do niewłaściwej klasy, aby zmylić czujność konkurentów, śledzących zgłoszenia w szczególnie interesującej ich klasie. To może stanowić istotne ograniczenie w przypadku ilościowej analizy patentów (Hinze, Schmoeh

2004: 231). Rysunek 16 ilustruje wyzwania, związane z identyfikacją podzbioru patentów, odpowiadających analizowanej technologii.



Rysunek 16: Problemy z identyfikacją patentów w oparciu o kody i słowa kluczowe. Źródło: Hinze, Schmoch (2004: 231).

Niektóre firmy zgłaszają wynalazki za pośrednictwem swoich spółek-córek, co utrudnia identyfikację podmiotów faktycznie kontrolujących daną technologię. O ile autorzy publikacji naukowych dążą do przekazania zrozumiałego i pełnego obrazu dokonanych odkryć, o tyle dokumenty patentowe podporządkowane są odmiennemu dążeniu do utrudnienia zrozumienia istoty wynalazku, sposobów jego uzyskania lub wykorzystania.

Szczegółowe zasady formułowania poleceń wyszukiwania są różne dla poszczególnych baz, a wydawca każdej bazy oferuje rozbudowane wskazówki i poradniki, dotyczące budowy zapytań. Możliwe jest też formułowanie zapytań złożonych przy wykorzystaniu warunków logicznych (*AND*, *OR*, *NOT*) oraz zagnieżdżonych struktur, w których nawiasy spełniają takie funkcje jak przy zapisywaniu równań matematycznych. Niektóre bazy umożliwiają wykorzystywanie symboli, odpowiadających dowolnemu ciągowi znaków (np. „*” odpowiadać może dowolnemu ciągowi znaków, a „B*” - ciągowi znaków, który rozpoczyna się od litery B i zawiera dowolne inne symbole alfanumeryczne). W poszczególnych bazach, różnie mogą być interpretowane polecenia wyszukiwania, obejmujące kilka wyrazów – przykładowo, w niektórych przypadkach wykorzystanie określenia *heart attack* może prowadzić do wyszukiwania jednocześnie występujących kombinacji słów *heart* oraz *attack*, w których oba słowa znajdują się bezpośrednio obok siebie, w innych przypadkach – do poszukiwań rekordów, które zawierają oba słowa, ale niekoniecznie w bliskim sąsiedztwie, wreszcie w pozostałych przypadkach mogą być też brane pod uwagę przypadki, w których w danym rekordzie występuje jedno lub drugie słowo (w związku z czym zalecane byłoby wprowadzenie wyszukiwanych słów w cudzysłowach). Składnia poleceń wyszukiwania każdej bazy jest inna, dlatego zalecane jest staranne zapoznanie się z jej opisem, oferowanym przez dostawcę bazy oraz przeprowadzenie wstępnych testów w celu weryfikacji poprawności skonstruowanego polecenia.

Przykłady wyszukiwań w bazie *INSPEC*:

2007 in PY

- publikacje pochodzące z roku 2007 (*PY* = „*PUBLICATION YEAR*”)

(2007 in PY) AND (Poland in AA)

- publikacje pochodzące z roku 2007, których autorzy wskazali jako afiliację adres w Polsce (*AA* = “*AUTHOR AFFILIATION*”)

((2007 in PY) OR (2008 in PY)) AND (Poland in AA)

- publikacje pochodzące z roku 2007 lub 2008, których autorzy wskazali jako afiliację adres w Polsce

(2007 in PY) AND (Poland in AA) NOT (A* in CC)

- publikacje pochodzące z roku 2007, których autorzy wskazali jako afiliację adres w Polsce i które nie zostały przez redaktorów bazy przypisane do obszaru badawczego „A” (fizyka)

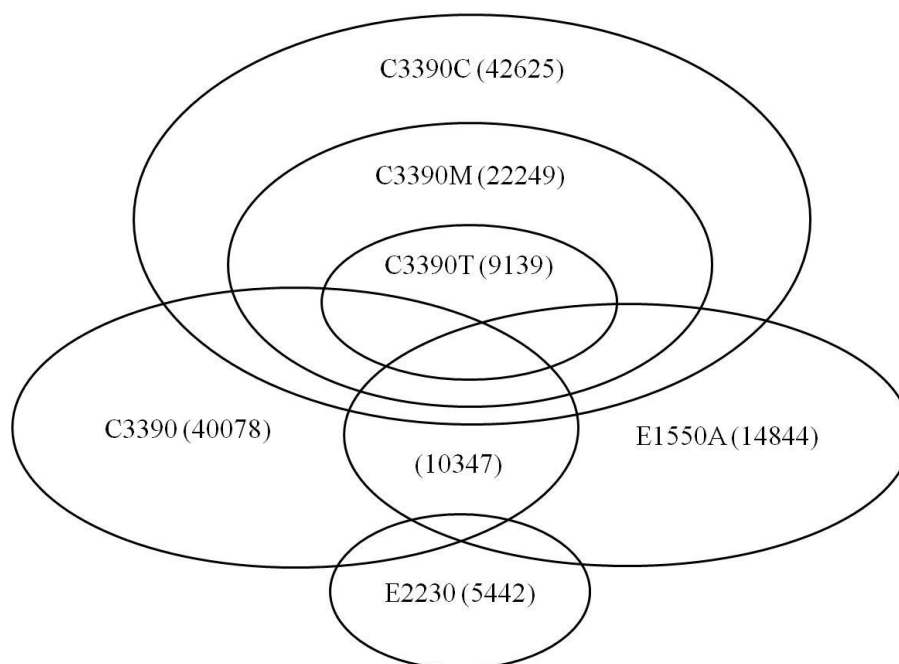
(2007 in PY) AND (Poland in AA) AND ((C3390 in CC) OR (E1550A in CC))

- publikacje pochodzące z roku 2007, których autorzy wskazali jako afiliację adres w Polsce i które zostały przez redaktorów bazy przypisane do podobszarów badawczych „C3390” lub „E1550A” (robotyka)

Niektóre bazy oferują rozbudowane systemy klasyfikacji obszarów badawczych oraz indeksy słów kluczowych. Wykorzystanie takich klasyfikacji może okazać się bardzo przydatne w pracach analitycznych, pozwalając w łatwy sposób wyodrębnić poszukiwane pozycje.

Szczegółowe taksonomie, stosowane przez wybrane bazy bibliograficzne do klasyfikacji publikacji i patentów, zostaną opisane w dalszej części książki. Warto rozważyć ich wykorzystanie w pracach badawczych zwłaszcza wtedy, gdy istotne jest wyodrębnienie specjalistycznego podzbioru rekordów. W obszarze nauk technicznych, najbardziej rozbudowany sposób opisu publikacji oferuje baza *INSPEC*, korzystająca z tzw. *Classification Codes* - taksonomii opracowanej przez *Institution of Engineering and Technology* (IET), jak również dodatkowych zbiorów kontrolowanych słów kluczowych (identyfikujących m.in. nowe obszary badań technologicznych, które nie dają się jednoznacznie wpisać w tradycyjny podział dyscyplin naukowych), indeksu chemicznego (wskazującego wzory związków chemicznych, które występują w danej publikacji), indeksu astronomicznego oraz innych ustrukturalizowanych metod klasyfikacji treści. Analogiczne możliwości oferuje baza *MEDLINE*, korzystająca z tezaury *MeSH* (*Medical Subject Headings*), obejmującego szerokie spektrum tematów badawczych, organów, substancji, zjawisk i jednostek chorobowych. Praktyczne wskazówki dotyczące zaawansowanych technik wyszukiwania publikacji w bazie *MEDLINE* oferują Claverie i Notredame (2007: 40-41).

Wybór odpowiednich elementów taksonomii nie jest zadaniem łatwym i powinien być wynikiem starannej analizy wyników różnych wariantów zapytań. W wielu przypadkach niezbędne okazuje się połączenie w jednym poleceniu wyszukiwania kryteriów, odnoszących się do fragmentu taksonomii bazy (pozycji przypisanych do danego obszaru badań lub technologii przez redaktorów bazy), a jednocześnie wyszukiwania pełnotekstowego (pozycji, w których treści występują wybrane słowa). Taksonomie baz są pomocne w identyfikacji poszczególnych obszarów tematycznych, nie można jednak zakładać nieomyślności redaktorów baz – niektóre pozycje mogły zostać przypisane do niewłaściwego obszaru merytorycznego.



Rysunek 17: Liczba publikacji dotyczących robotyki zidentyfikowanych w bazie *INSPEC* w oparciu o wybrane słowa kluczowe. Źródło: opracowanie własne w oparciu o badania Gity Ghiasi Hafezi (2011), stan na 27 lutego 2011. Wyjaśnienia kodów klasyfikacyjnych bazy *INSPEC*, zaprezentowanych na schemacie:

- *C3390: Robotics,*
- *C3390C: Mobile robots,*
- *C3390M: Manipulators,*
- *C3390T: Telerobotics,*
- *E1550A: Robotics,*
- *E2230: Robot and manipulator mechanics.*

Liczby w nawiasach odzwierciedlają liczbę publikacji przypisanych wyłącznie do danego kodu klasyfikacyjnego.

Analiza kodów klasyfikacyjnych nie jest w wielu przypadkach zadaniem prostym, co ilustruje przykład prób wyodrębnienia publikacji, dotyczących robotyki, co prezentuje Rysunek 17. Przedstawione na rysunku elipsy odpowiadają podzbiorom publikacji, opisanych poszczególnymi kodami klasyfikacyjnymi bazy *INSPEC*. Jak się okazuje, znacząca liczba publikacji opisanych kodami niższego poziomu nie jest jednocześnie opisana kodem dla wyższego poziomu (np. znaleziono 42.625 publikacji zaklasyfikowanych jako dotyczących robotów mobilnych – *C3390C* – a jednocześnie nie przypisanych do szerszej kategorii *C3390*, odpowiadającej robotyce). Poprawna identyfikacja wszystkich publikacji, dotyczących

robotyki, wymaga więc pracochłonnych analiz różnych kodów poprzez złożone wyszukiwania, które ujawnią liczby publikacji dotyczących robotów, a nie przypisanych do najbardziej oczywistych kategorii. W efekcie zbudowane zostanie odpowiednie polecenie wyszukiwania, które w zaprezentowanym powyżej (uproszczonym) przykładzie robotyki powinno obejmować następujące elementy:

- **(C3390 in CC) OR (E1550A in CC) OR (E2230 in CC) OR (C3390C in CC) OR (C3390M in CC) OR (C3390T in CC)**

W rzeczywistości publikacje dotyczące zagadnień związanych z robotyką można znaleźć również w kategoriach *C3355 (Control applications in manufacturing processes)* oraz *C3355F (Control applications in assembling)*, choć dotyczy to tylko niewielkiej części prac przypisanych do tych kategorii, co skłania do poszukiwania odpowiedniej metody selekcji odpowiednich prac. Dodatkowo, znacząca liczba publikacji w bazie *INSPEC* zawiera słowo „robot” w tytule lub treści abstraktu, choć nie jest przypisana do żadnej z wymienionych powyżej kategorii. Część streszczeń publikacji, uzyskanych jako wyniki wyszukiwania, powinna być starannie przeczytana po to, by zidentyfikować sytuacje, gdy w wynikach znajdują się publikacje nie związane z tematem analiz – może być to punkt wyjścia do poszukiwania sposobów przeformułowania poleceń wyszukiwania. Analiza różnych scenariuszy wyszukiwania jest więc niezbędna przed podjęciem decyzji o ostatecznym kształcie polecenia wyszukiwania w bazie, które zostanie wykorzystane do wygenerowania podzbioru danych, wykorzystywanych w dalszych analizach bibliometrycznych.

Należy pamiętać o zjawisku systematycznego uzupełniania baz nowymi danymi. Może to wynikać z dwóch powodów:

- systematycznego indeksowania nowych pozycji,
- podpisania przez dostawcę bazy umowy z kolejnymi wydawnictwami, które wyrażają zgodę na uzupełnienie bazy licznymi pozycjami z wcześniejszych lat (problem nie dotyczy baz patentów i zgłoszeń patentowych).

Konsekwencją ich wystąpienia jest to, że wyniki tego samego polecenia wyszukiwania, uzyskane w dwóch różnych dniach, mogą się istotnie różnić. Pierwszy problem można wyeliminować, zawężając analizy do pozycji starszych niż 1 rok (np. gromadząc dane w kwietniu 2011 roku, analizować wyłącznie wyniki dla okresu do grudnia 2010 roku). Drugi problem nie jest możliwy do rozwiązania, gdyż po uzupełnieniu bazy o nowe publikacje z ubiegłych lat nie jest możliwa eliminacja niedawno dodanych pozycji, które mogą być w rzeczywistości publikacjami sprzed wielu lat. Jego skala zależy od strategii produktowej, przyjętej przez dostawcę bazy. Przykładowo, bazę *Web of Science* charakteryzuje relatywna stabilność składu czasopism, indeksowanych w bazie, co jest wynikiem polityki wydawcy, w tym decyzji o ograniczeniu liczby indeksowanych czasopism i koncentracji na najważniejszych pozycjach dyscypliny (choć w ostatnich latach baza została uzupełniona o wiele pozycji chińskich). Odmienne jest podejście dostawcy bazy *Scopus*, dążącego do zaprezentowania w niej jak najszerszego spektrum publikacji naukowych i regularnie uzupełniającego indeksy bazy o nowe periodyki.

Ze względu na opisane powyżej ograniczenia, zaleca się zastosowanie do trzech poniższych rekomendacji:

- Wszystkie wyszukiwania w bazie bibliometrycznej, które mają być ze sobą bezpośrednio porównywane, powinny być przeprowadzone tego samego dnia, co ograniczy niebezpieczeństwo zmian zawartości bazy w trakcie gromadzenia danych.
- Wyszukiwania powinny być ograniczone czasowo i nie uwzględniać pozycji z ostatnich kilku miesięcy, gdyż można spodziewać się, że nadal trwa proces dodawania i indeksowania najnowszych pozycji, a więc wyniki dla ostatnich miesięcy byłyby zaburzone.
- Osoba prowadząca analizy powinna starannie udokumentować wykorzystane polecenie wyszukiwania, datę jego przeprowadzenia oraz ewentualne inne kryteria, ograniczające wyszukiwanie, co pozwoli w przyszłości odtworzyć wyniki analiz w przypadku zmian zawartości bazy.

Fakt systematycznego uzupełniania baz nowymi publikacjami wiąże się z inną zależnością, dotyczącą danych bibliometrycznych. Informacja o roku publikacji, wystąpienia lub zgłoszenia, przypisany do danego obiektu w bazie, powinna być poddana dalszej interpretacji przez analityka. Należy pamiętać o przesunięciach czasowych wynikających z cyklu publikacji w renomowanych czasopismach oraz procedur rozpatrywania zgłoszeń patentowych. Gdyby dane dotyczące tych różnych grup obiektów były wykorzystywane do analiz, w których znaczenie odgrywa chronologia - wskazane może być uwzględnienie przesunięć pomiędzy artykułami i patentami a wystąpieniami konferencyjnymi. Warto nie zapominać również, że informacja o konkretnym roku publikacji lub wystąpienia nie wskazuje jednoznacznie momentu rozpoczęcia badań – zwykle zaczynają się one na długo przed uzyskaniem wymiernych rezultatów, podlegających upublicznieniu.

Bazy publikacji zawierają niestety wiele błędów, które powstały na etapie indeksowania albo zostały przejęte z oryginalnych publikacji drukowanych. Typowe są literówki w obco brzmiących nazwiskach autorów (np. zamiast *Porowski* - *Porowowski*, *Porowshi*), różne warianty nazwy tej samej instytucji (np. *Institute of Physics*, *Department of Physics*, *Faculty of Physics*) oraz błędne przypisania geograficzne. Proste polecenia wyszukiwania mogą więc prowadzić do przeoczenia istotnych przypadków lub zależności. Jeśli poszukujemy tekstów, które zostały stworzone przez polskich autorów, zapewne pominiemy te publikacje, które zostały błędnie przypisane do innych krajów (np. pochodzące zgodnie z zawartością bazy z "*Warsaw, Russia*"). Nawet nazwy krajów bywają w bazach podawane w błędny sposób (Archambault i in. 2009: 1321). Sposobem na uniknięcie opisanych błędów byłby przegląd artykułów ze wszystkich krajów, jednak byłby on niezwykle pracochłonny. Na szczęście błędy indeksowania tekstów z ostatnich lat zdarzają się rzadko – stanowiły istotny problem w odniesieniu do tekstów pochodzących z krajów Europy Środkowo-Wschodniej w pierwszej połowie lat 90-tych XX-ego wieku. W przypadku zaawansowanych analiz, wybrany podzbiór z bazy należy poddać operacji ręcznego czyszczenia danych, która pozwoli na eliminację niektórych błędnie sklasyfikowanych tekstów oraz ujednoczenie nazewnictwa instytucji i nazwisk autorów. Taka możliwość nie pojawia się oczywiście w odniesieniu do analiz podstawowych, prowadzonych bez dodatkowego wsparcia programu komputerowego.

W kolejnym kroku, podstawowa analiza bibliometryczna sprowadza się do spisania liczby rekordów, która jest wynikiem danego polecenia wyszukiwania. Powtarzając polecenia wyszukiwania ze zmienionym kryterium, można w ten sposób otrzymać przykładowo dane, odzwierciedlające liczby publikacji w kolejnych latach, napisanych przez autorów z różnych krajów i ośrodków badawczych lub odnoszących się do różnych obszarów badań. Dane można następnie zebrać w arkuszu kalkulacyjnym i poddać dalszej obróbce i wizualizacji. Należy jednak pamiętać o wszystkich opisanych wcześniej ograniczeniach takich podstawowych analiz.

Podjęciem gwarantującym rzetelność badawczą będzie wykorzystanie pełnego zbioru danych, uzyskanego w wyniku przeprowadzonego wyszukiwania, do dalszej analizy przy wykorzystaniu specjalistycznego oprogramowania. Może bowiem okazać się, że niektóre rekordy były błędnie przypisane, a liczba rezultatów rzeczywiście uzyskanych jest różna od wartości wyświetlanej po wykonaniu wyszukiwania przez bazę bibliometryczną. W takich przypadkach wyszukiwanie jest punktem wyjścia do wygenerowania pliku, zawierającego wybrane pola rekordów bazy, który zostanie zapisany na dysku lokalnym komputera użytkownika, a później przetworzony przy pomocy odpowiedniej aplikacji. Większość baz oferuje opcję zapisu na dysku wybranych rekordów w określonym formacie. Te możliwości są różne w zależności od bazy, a najważniejsze różnice dotyczą:

- konieczności ręcznego zaznaczania rekordów, które powinny być ściągnięte lub możliwości wskazania przedziału liczbowego rekordów do ściągnięcia (np. od rekordu 1 do rekordu 8534),
- ograniczeń maksymalnej liczby rekordów, które mogą być ściągnięte (takie ograniczenia występują m.in. w bazach *Scopus* oraz *Web of Science*, a sposobem na ich obejście jest przeformułowanie zapytań do bazy w taki sposób, by każde z nich generowało znacznie krótszą listę wyników, które następnie mogą być połączone w plikach, zapisanych na komputerze; najbardziej dogodnym kryterium dla takiego podziału są kolejne lata publikacji),
- wybór sposobu formatowania danych wyjściowych, pól zapisywanych w rekordach, formatu etykiet identyfikujących dane czy rodzaju pliku wyjściowego.

Bezpłatne bazy patentów takie jak *USPTO*, *esp@cenet* oraz *Google Patents*, w ogóle nie oferują możliwości masowego ściągania na dysk lokalny wyników wyszukiwania. Istnieją jednak liczne specjalistyczne programy, które pozwalają takie operacje zapisywania zautomatyzować. Zwykle jest możliwe dzięki temu, że adresy stron internetowych, na których znajdują się opisy poszczególnych patentów, zawierają w swojej składni numer patentu – czyli po identyfikacji listy dokumentów patentowych z określonymi numerami, do odpowiedniego programu wystarczy wprowadzić listę interesujących nas numerów. Przykładem darmowego programu, oferującego wsparcie w ściąganiu tego typu plików, jest aplikacja *HTTrack Website Copier* (www.httrack.com). Istnieje też wiele programów i skryptów, dedykowanych do ściągania dokumentów patentowych z określonych baz. Warto pamiętać o możliwości istnienia określonych przez dostawcę bazy ograniczeń licencyjnych – zwykle nie jest dozwolone masowe ściąganie bardzo dużych liczb dokumentów w krótkich odstępach czasu. Takie ograniczenie wprowadziło przykładowo *USPTO*, obawiając się

nadmiernego obciążenia serwera bazy danych, a próba seryjnego ściągnięcia tysięcy dokumentów patentowych może hipotetycznie doprowadzić nawet do zablokowania dostępu danego komputera do serwera bazy. W praktyce operacje pobierania danych nie dotyczą zwykle tak rozbudowanych zbiorów dokumentów, mogą też być prowadzone w regularnych odstępach czasowych, nie stanowiąc nadmiernego obciążenia dla serwera dostawcy.

Warto przypomnieć o istotnym rozróżnieniu pomiędzy analizą udzielonych patentów oraz nierozpatrzonych lub odrzuconych zgłoszeń patentowych. Zgłoszenia odzwierciedlają prace badawczo-rozwojowe i niezależnie od ewentualnego kwestionowania poziomu wynalazczego, mogą być źródłem cennych informacji na temat rozwoju technologii. *USPTO* od 1999 roku rozpoczął publikację zgłoszeń patentowych w wyodrębnionej części swojego serwisu internetowego (Hinze, Schmoch 2004: 218).

Techniki bibliometryczne można zastosować również do przetwarzania danych, pochodzących z baz niedostępnych publicznie. Przykładowo, mogą to być bazy, zawierające dokumentację finansowanych projektów badawczych, opisy prac, na podstawie których uzyskano stopnie naukowe, ankiety jednostek naukowych, wykorzystywane przy ich ocenie parametrycznej lub wnioski o dofinansowanie projektów B+R z funduszy budżetowych lub międzynarodowych. Instytucją zarządzającą niektórymi spośród wymienionych baz w polskim systemie nauki i szkolnictwa wyższego jest Ośrodek Przetwarzania Informacji, publicznie dostępne są tylko wybrane serie danych, zgromadzonych w bazach OPI. Przykładami analogicznych analiz, prowadzonych w innych krajach, było przetwarzanie danych pobranych z bazy *Korean Research Foundation*, obejmujących ponad 200 tysięcy wniosków o dofinansowanie projektów badawczych z lat 2002-2006 (Yoon, Lee, Lee 2010) oraz analizy bibliometryczne oparte na bazie *CRISP*, stworzonej przez *US National Institute of Health* i obejmującej dokumentację projektów badawczych w obszarze nauk medycznych prowadzonych w Stanach Zjednoczonych, jak również analogicznej, szerszej bazie *US National Science Foundation* (Mogoutov 2008: 343-344).

Przetwarzanie rekordów, pobranych z bazy, w specjalistycznym oprogramowaniu pozwala na ich oczyszczenie (w tym eliminację danych błędnie zaklasyfikowanych), ujednoczenie nazewnictwa (np. poprawki w nazwiskach autorów i nazwach instytucji lub przypisanie różnych wariantów nazwy tej samej instytucji lub różnych jednostek organizacyjnych tej samej instytucji do jednej, wspólnej grupy) oraz prowadzenie bardziej zaawansowanych analiz. Wiele programów komputerowych, wspierających analizy bibliometryczne, opiera się na strukturze bazy danych, co oznacza, że użytkownik ma możliwość prowadzenia analiz w oparciu o kombinację różnych kryteriów. W odniesieniu do zgromadzonego zbioru rekordów możliwa jest budowa szczegółowych zapytań w celu znalezienia np. rekordów, napisanych przez autorów z danego kraju w określonym roku, albo należących do określonych podobszarów badawczych. Możliwości takich analiz zostaną opisane w dalszej części książki – stanowią one podstawę dla technik *tech mining*.

8.3. Listy rankingowe

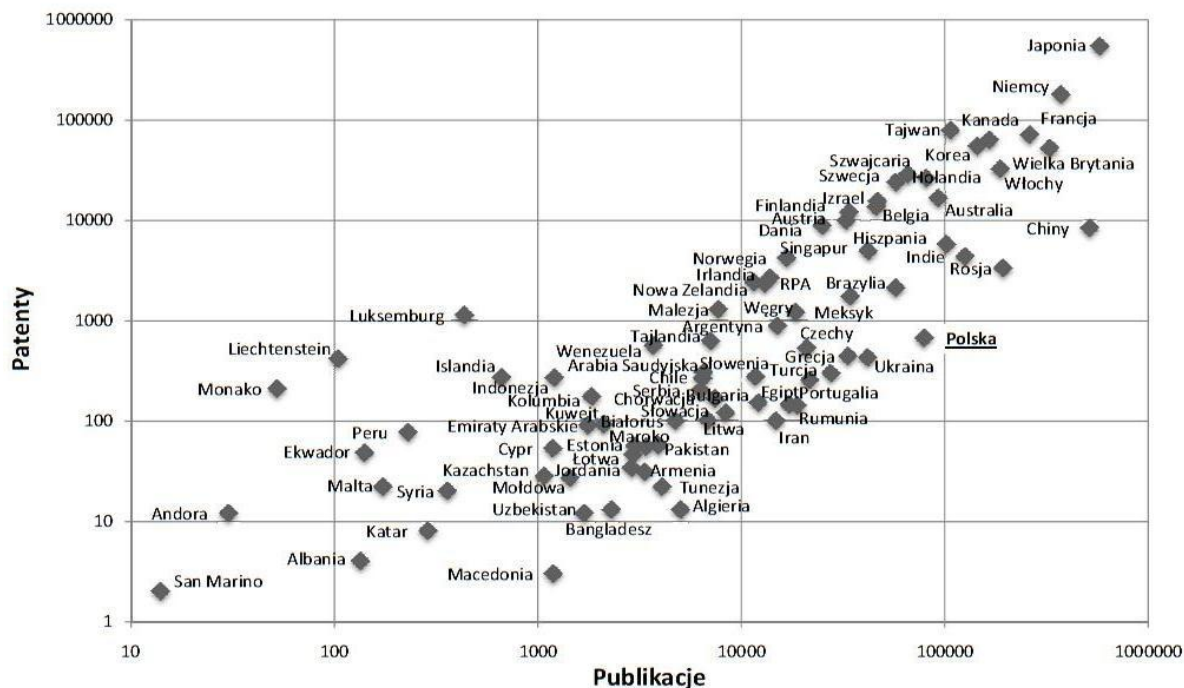
Poniżej zaprezentowane zostaną podstawowe scenariusze analiz, które były wykorzystywane jeszcze zanim badacze w obszarze bibliometrii uzyskali dostęp do specjalistycznych narzędzi

informatycznych. Ich praktyczne zastosowania zostaną zaprezentowane w rozdziale, opisującym *tech mining*, jednak w tym miejscu wskazana wydaje się prezentacja ich najważniejszych założeń. Pięć podstawowych scenariuszy analitycznych to:

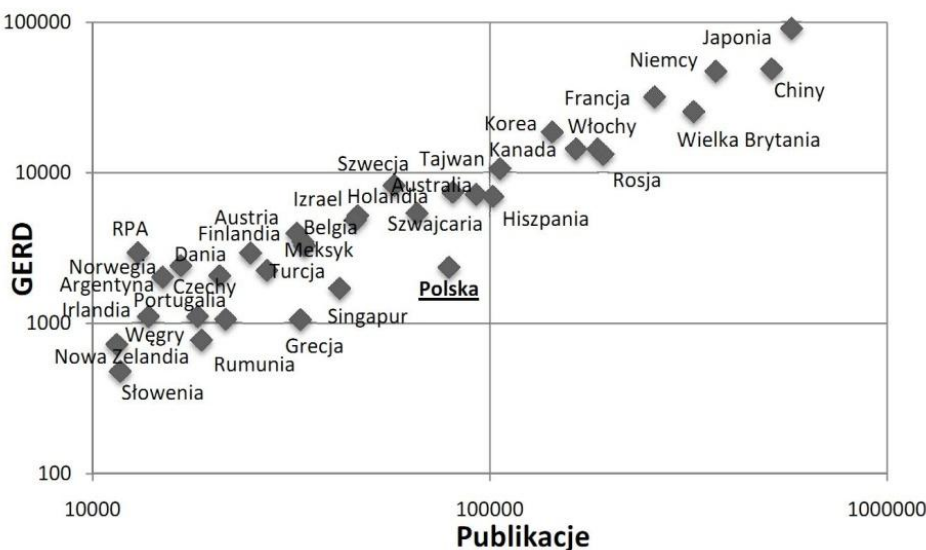
1. budowa list rankingowych w oparciu o wybrane kryteria – elementy opisu rekordów w bazie bibliometrycznej,
2. analiza współwystępowania słów (*co-word analysis*),
3. analiza współklasyfikacji (*co-classification analysis*),
4. analiza współautorstwa (*co-authorship analysis*),
5. analiza cytowań (*citation analysis*).

Listy rankingowe mogą prezentować liczby publikacji lub patentów, które zostały stworzone przez określonych autorów, pracowników określonych organizacji, należą do wybranych podobszarów badań, powstały w określonym roku itd. Odpowiednie listy można wygenerować dla każdego z pól indeksu, opisującego rekord w bazie danych, a programy komputerowe ułatwiają sortowanie wyników i tworzenie rankingów. Dzięki ich wykorzystaniu, możliwe jest udzielenie odpowiedzi na pytania takie jak: *kto jest najbardziej produktywnym autorem w danym obszarze badań? które uczelnie niemieckie są najbardziej zaangażowane w badania w tym obszarze? jakie podobszary badawcze cieszą się największą lub najmniejszą popularnością?*

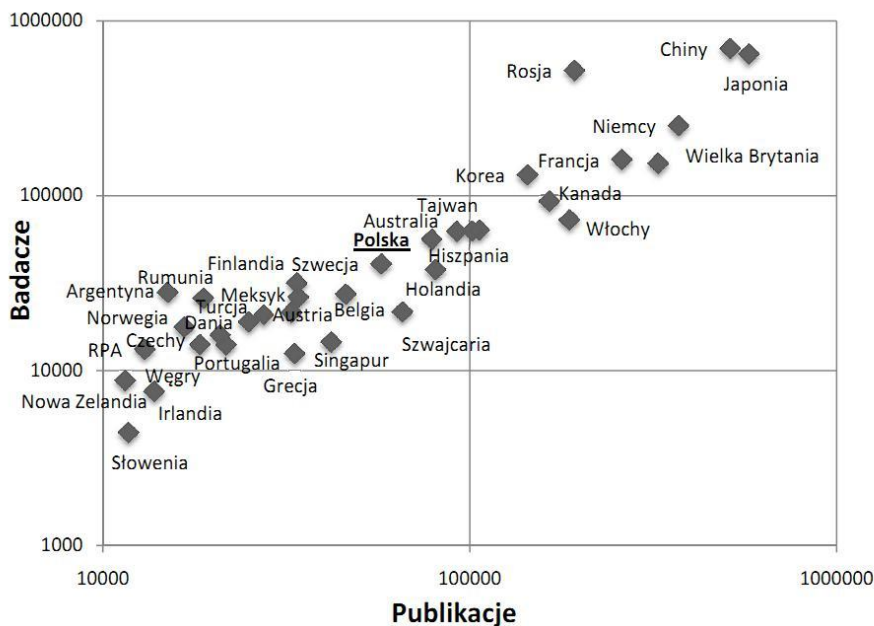
Wygenerowane zestawienia pozwalają na porównanie skali prowadzonej działalności badawczo-rozwojowej. Oprócz list rankingowych, dogodną formą prezentacji danych są diagramy rozprożeń (ang. *scatterplots*), zestawiające dwie zmienne bibliometryczne lub zmienną bibliometryczną z innymi danymi socjoekonomicznymi. Możliwe jest przykładowo zestawienie liczby publikacji i patentów, pozwalające na określenie orientacji systemu B+R na zastosowania teoretyczne lub praktyczne (Rysunek 18). Często prowadzone są też analizy produktywności badań, oparte na zestawieniach liczby publikacji lub patentów z nakładami na B+R (Okubo 1997: 42) (por. Rysunek 19, prezentujący wartość całkowitych nakładów na B+R w gospodarce narodowej czyli *GERD*), jak również porównaniu liczby rezultatów badań z liczbą pracujących badaczy (Rysunek 20).



Rysunek 18: Liczba publikacji naukowych w naukach ścisłych i technicznych a liczba patentów, udzielonych wynalazcom z poszczególnych krajów dla lat 1990-2007 w oparciu o bazy INSPEC i USPTO. Źródło: Klincewicz (2008: 57).



Rysunek 19: Średnie roczne całkowite wydatki gospodarek na B+R (GERD) w latach 1989-2006 a liczba publikacji naukowych w naukach ścisłych i technicznych w latach 1990-2007 w oparciu o bazy INSPEC i OECD-Stat. Źródło: Klincewicz (2008: 60).



Rysunek 20: Średnia liczba badaczy w sektorze B+R w latach 1989-2006 a liczba publikacji naukowych w naukach ścisłych i technicznych w latach 1990-2007 w oparciu o bazy INSPEC i OECD-Stat. Źródło: Klincewicz (2008: 64).

8.4. Analiza współwystępowania słów (*co-word*)

Analiza *co-word* opiera się na przypadkach współwystępowania określonych słów. Może być stosowana na poziomie tytułów, listy słów kluczowych, streszczenia lub całego rekordu. Korzenie teoretyczne tej techniki sięgają dorobku francuskich badań leksykograficznych i lingwistycznych oraz socjologicznej teorii aktora-sieci *ATN* (ang. *actor-network theory*) (Courtial 1994). Prekursorem techniki był Michel Callon z *École des Mines de Paris*, podejmujący pierwsze próby badawcze w latach 70-tych XX wieku (Kostoff 2003: 10). Przykłady projektów, opartych na technice *co-word*, obejmują m.in. analizę zmian zainteresowań badawczych w obszarze chemii polimerów na przestrzeni lat (Callon, Courtial, Laville 1991) oraz wykorzystanie omawianej techniki do analizy wynalazków w oparciu o tytuły zgłoszeń patentowych (Courtial, Callon, Sigogneau 1993).

Pierwsze podejścia opierały się na analizie słów kluczowych, przypisanych poszczególnym publikacjom przez osoby prowadzące badania. W latach 80-tych narastała krytyka takiego sposobu analiz, odwołująca się do występowania tzw. *indexer effect*, konsekwencji subiektywnego indeksowania analizowanych publikacji poprzez przypisywanie im określonych słów kluczowych przez osoby, które nie prowadziły samodzielnych badań w analizowanym obszarze i w związku z tym nie potrafiły zidentyfikować nowych tendencji badawczych, a słowa kluczowe przypisywały w sposób niezwykle konserwatywny (Kostoff 2003: 11).

W przypadku przetworzenia całego korpusu tekstów (czyli przykładowo zbioru rekordów, pobranych z bazy), możliwa jest jednak identyfikacja najczęściej występujących związków frazeologicznych lub innych regularnie współwystępujących słów. Niektóre z tych

8.5. Analiza współklasyfikacji (*co-classification*)

Analiza *co-classification* odwołuje się do przypadków współwystępowania określonych słów kluczowych lub symboli alfanumerycznych, pochodzących z tezauryusa określonej bazy danych. Opracowanie tego typu zestawienia pozwala na identyfikację szczegółowych podobszarów badań, przypadków badań interdyscyplinarnych oraz tych, które interdyscyplinarnymi nie są (tj. pozostają przypisane wyłącznie do jednego obszaru). Przykładowo, możliwy jest podział wyodrębnionego zbioru publikacji z obszaru nanotechnologii na bardziej szczegółowe podgrupy związane z określonymi zagadnieniami fizycznymi, chemicznymi i biologicznymi. Badania informatyczne można podzielić na prace skoncentrowane na tematyce *software* i *hardware*. Dorobek badań w obszarze robotyki podzielić w oparciu o zastosowania robotów lub szczegółowe aspekty ich konstrukcji (m.in. ruch, identyfikację obiektów, manipulowanie obiektami, komunikację). Prowadzenie tego typu analiz jest możliwe tylko wtedy, gdy określona baza oferuje rozbudowaną taksonomię, która może być wykorzystana przez analityka. Do tych celów szczególnie dobrze nadają się bazy publikacji *INSPEC* i *MEDLINE*. Bazy patentów również opierają się na standardowych systemach klasyfikacyjnych, których wykorzystanie umożliwia prowadzenie analiz – przykładowo, Leydesdorff (2008) poddał analizom współklasyfikacji zbiór ponad 136 tysięcy zgłoszeń patentowych w trybie *PCT*, odzwierciedlając organizację wiedzy i techniki (związki pomiędzy obszarami badawczymi, w tym interdyscyplinarność badań). Leydesdorff zauważa, że praca ekspertów z urzędów patentowych gwarantuje trafniejszą klasyfikację patentów niż analogiczne działania redaktorów baz publikacji, gdyż stosowanie i doskonalenie systemu klasyfikacyjnego należy do najważniejszych obszarów działalności każdego urzędu patentowego (Leydesdorff 2008: 1592).

	Podobszar 1	Podobszar 2	Podobszar 3
Podobszar 1	8211	1345	0
Podobszar 2	1345	6387	28
Podobszar 3	0	28	2096

Tabela 10: Przykład fragmentu wyników analizy współklasyfikacji. Źródło: opracowanie własne.

Tabela 10 prezentuje przykładowy fragment wyników analizy współklasyfikacji. Wyodrębniono w nim trzy podobszary badawcze, które w oryginalnym zbiorze danych, pobranym z bazy, były opisane trzema różnymi kodami klasyfikacyjnymi (rzeczywiście analizy opierają się oczywiście na znacznie obszerniejszych tabelach, obejmujących setki lub tysiące takich kodów). W analizowanym zbiorze danych nie zidentyfikowano żadnych przypadków współwystępowania kodów, oznaczających Podobszar 1 i Podobszar 3. Można to zinterpretować jako prawdopodobną rozłączność tematyczną tych podobszarów badawczych, brak wspólnych zagadnień, metod i technik. Warto wspomnieć jednak, że poddany analizom zbiór danych mógł nie zawierać wielu istotnych publikacji ze względu na sposób sformułowania polecenia wyszukiwania – ostateczne wnioski o rozłączności obszarów badawczych powinny być dodatkowo weryfikowane w oparciu o kompletną bazę

bibliometryczną. Podobszary 1 i 2 zazębiają się, o czym świadczy aż 1345 publikacji, do których przypisano oba kody. Słabsze zależności można zaobserwować pomiędzy podobszarami 2 i 3, które łączy jedynie 28 publikacji. W ewentualnych dalszych krokach analitycznych warto sprawdzić, w jakim okresie powstały te publikacje – może jesteśmy świadkami narodzin nowego interdyscyplinarnego kierunku badań, a po niewielkiej początkowej liczbie publikacji wkrótce możemy spodziewać się intensyfikacji rozwoju tych badań. Tabełacyjne zestawienia przypadków współklasyfikacji pozwalają też na identyfikację przypadków publikacji, które są przypisane tylko do jednego podobszaru, choć we współczesnej literaturze naukowej takie przypadki zdarzają się niezwykle rzadko. Dane, zgromadzone dzięki zastosowaniu technik *co-classification*, mogą służyć również porównaniu orientacji badawczych różnych instytucji i krajów oraz przemian dyscypliny naukowej na przestrzeni lat.

8.6. Analiza współautorstwa (*co-authorship*)

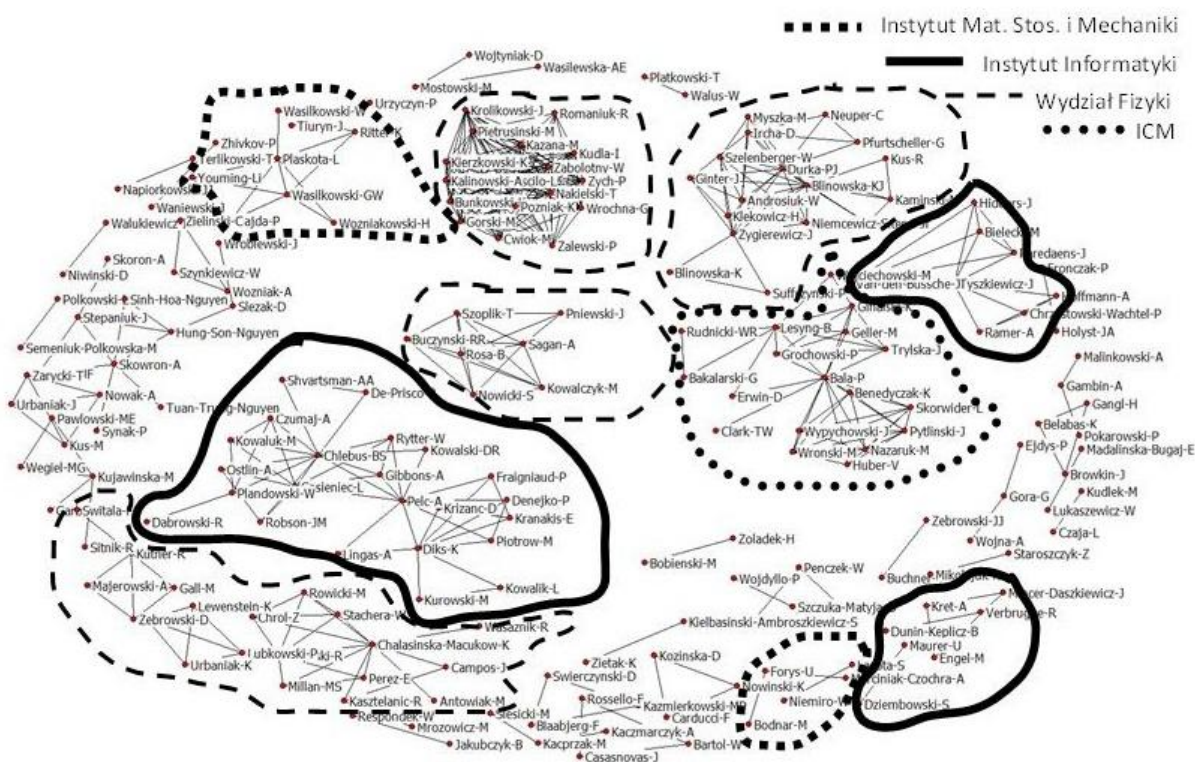
Analiza współautorstwa (*co-authorship*) pozwala na identyfikację schematów współpracy w ramach zespołów badawczych poprzez ustalenie częstotliwości autorstwa publikacji lub wynalazków przez dwie lub więcej osób. Ze względu na ograniczenia technik wizualizacji, najczęściej przeprowadza się te analizy w poszukiwaniu przypadków wspólnego autorstwa dwóch osób (umieszczanych odpowiednio w wierszach i kolumnach tabeli), jednak możliwe są też analizy oparte o bardziej rozbudowane algorytmy (np. identyfikacja zespołów badaczy-współautorów, liczących co najmniej 5 osób).

Tabela 11 prezentuje uproszczony przykład analizy współautorstwa dla wybranej dziedziny badań, ograniczony do czterech autorów. Autor 1 nie angażuje się we współpracę badawczą ani publikacyjną z żadnym innym naukowcem, co może być podejrzane w przypadku badań z obszaru nauk eksperymentalnych. Autor 2 jest z kolei jednocześnie produktywnym badaczem (65 publikacji) i aktywnym współpracownikiem (we współpracy z innymi autorami powstała liczba publikacji równa różnicy pomiędzy łączną liczbą jego publikacji: 65, a liczbą publikacji napisanych indywidualnie: 33). Autor 3 i 4 ograniczają się do udziału w pisaniu prac zespołowych, nie stworzyli żadnego indywidualnego tekstu. Zestawienie pozwala też dostrzec, że 2 spośród 32 tekstów, współtworzone przez Autora 3, powstały przy współpracy z Autorem 2 i Autorem 4, a 30 tekstów było napisanych tylko z Autorem 2. Interpretacja wyników analiz pozwala na ustalenie najbliższych współpracowników określonych badawczy. W omawianym przypadku można przykładowo przyjąć, że Autor 3 jest dobrze zaznajomiony z częścią dorobku badawczego Autora 2 – zapewne był jego „prawą ręką” w przypadku niektórych projektów, potencjalnie może też stać się jego sukcesorem.

	Autor 1	Autor 2	Autor 3	Autor 4	SUMA
Autor 1	47	0	0	0	47
Autor 2	0	33	32	3	65
Autor 3	0	32	0		32
Autor 4	0	3	2	0	3
SUMA	47	65	32	3	

Tabela 11: Przykład fragmentu analizy współautorstwa. Źródło: opracowanie własne.

Analogiczne analizy można przeprowadzić w odniesieniu do instytucji badawczych, stanowiących afiliacje autorów – jednak wymagają one zwykle odpowiedniego przetworzenia zbioru danych, ujednoczenia nazewnictwa instytucji i wygenerowania sumarycznych zestawień rekordów.



Rysunek 22: Przykład analizy współautorstwa: sieć autorów publikacji informatycznych z Uniwersytetu Warszawskiego. Źródło: Klineciewicz (2008: 183).

Warto pamiętać o ograniczeniach niektórych baz, utrudniających prowadzenie analizy współautorstwa. Baza *INSPEC* dopiero kilka lat temu rozpoczęła podawanie danych dotyczących afiliacji kolejnych współautorów tekstu, a jej wcześniejsze rekordy zawierają tylko informacje o afiliacji pierwszego autora. Baza *MEDLINE* podaje maksymalnie dziesięciu autorów publikacji - nazwiska autorów na miejscach dalszych nie są dostępne, co

może stanowić problem w odniesieniu do niektórych obszarów badań, przykładowo genomiki (Claverie, Notredame 2007: 41).

8.7. Analiza cytowań (*citation*)

Analiza cytowań (ang. *citation analysis*) odwołuje się do związków merytorycznych pomiędzy publikacjami lub patentami. Fakt cytowania innej publikacji można interpretować jako oparcie się na jej treści czy pośredni transfer wiedzy. Tego typu analizy są szczególnie często podejmowane przez autorów tekstów naukowych z obszaru bibliometrii – znajdują jednak tylko ograniczone zastosowania w odniesieniu do analizy rozwoju nauki i technologii. Podstawowym celem analizy cytowań jest identyfikacja najważniejszych pozycji literaturowych, ich autorów lub zatrudniających ich instytucji. W języku polskim pojęcie „cytowanie” można stosować zamiennie z „odsyłaczem”, a określenie „liczba cytowań” można wyrazić też jako „liczba odsyłaczy”.

Podstawy metodologiczne dla analizy cytowań literatury naukowej stworzył w latach 60-tych XX wieku Eugene Garfield, założyciel *Institute for Scientific Information (ISI*, obecnie *Thomson Reuters*) i twórca *Science Citation Index*, który później przekształcony został w bazę *Web of Science*, a następnie - *Web of Knowledge*. Analizy te opierają się na założeniu, że większa liczba cytowań jest powiązana z „jakością” czy naukową doniosłością. Niestety, to założenie nie zawsze jest prawdziwe. Współcześnie dominuje przekonanie, że analiza cytowań pozwala uzyskać informacje na temat wpływu badań, a nie ich jakości (Moed 2005: 37; 81).

Cytowania odzwierciedlają tendencje w zakresie czytania i przywoływania literatury naukowej oraz zwyczaje, panujące w środowiskach akademickich poszczególnych krajów. Socjolog Robert Merton sugerował istnienie tzw. efektu Mateusza w nauce: samonapędzającej się pętli cytowań, w której najczęściej cytowani autorzy będą też najchętniej czytani przez kolejnych badaczy, a reprezentanci mniej istotnych ośrodków naukowych na zawsze pozostawać będą na peryferiach światowej nauki (Merton 1968). Czasami „wypada” cytować określonego autora lub ośrodek badawczy ze względów pragmatycznych, warunkujących publikację tekstu naukowego. Z wymienianych powodów ograniczenie się do cytowań w przypadku analiz szczegółowych obszarów badawczych prowadzi do przeoczenia wielu istotnych kierunków badań, które nie należą do głównego nurtu dyscypliny. Analiza cytowań nie nadaje się też do identyfikacji nowych tendencji technologicznych (ang. *emerging technologies*) – nowe zjawiska w nauce i technologii w wielu przypadkach nie stały się nawet jeszcze przedmiotem publikacji w czasopiśmie recenzowanych, a nieistniejące publikacje nie mogą być cytowane. Dodatkowo, istotne teksty naukowe dostrzegane są jako przełomowe dopiero po pewnym czasie, a długi cykl publikacji odsuwa w czasie moment pojawienia się pierwszych cytowań – dlatego analiza cytowań wydaje się bezużyteczna w odniesieniu do nowych, dynamicznie rozwijających się obszarów badań. Z drugiej strony, właśnie cytowania pozwalają zidentyfikować najważniejsze, najczęściej przywoływane pozycje, stanowiące kanon danej dyscypliny naukowej (ang. *research fronts*).

Cytowanie nie musi jednak oznaczać uznania dla tekstu naukowego lub badacza. Niektóre cytowania mają na celu krytykę tekstu i wykazanie błędów, czego przykładem było częste powoływanie się fizyków na tekst, opisujący prawdopodobnie nieprawdziwe odkrycie tzw. zimnej fuzji (Nowak 2006: 103-104). Analizy empiryczne wskazują jednak, że cytowania negatywne stanowią niewielki odsetek wszystkich cytowań (Moed 2005: 84). Zasady oceny naukowców, pracujących w wielu ośrodkach badawczych i rywalizujących o środki na finansowanie swoich projektów często odwołują się do wskaźników, opartych o liczbę cytowań publikacji danej osoby. Taka sytuacja skłania do podejmowania działań, mających w aktywny sposób wpływać na podnoszenie liczby cytowań własnych tekstów. Ujawniają się one zwłaszcza wtedy, gdy występuje silna zawodowa lub finansowa motywacja. Może wtedy dochodzić do „skłaniania” innych do cytowania własnych prac i budowy „koleżeńskich układów”, których uczestnicy dodają nazwiska współautorów, którzy nie wnosili rzeczywistego wkładu w opracowanie badań i publikacji, lub wzajemnie cytują prace znajomych (Nowak 2006: 104). W niektórych kulturach i instytucjach można zaobserwować niepisany obowiązek cytowania prac zwierzchników akademickich lub dopisywania przełożonych jako współautorów tekstów akademickich. Zhou i Leydesdorff (2006) pokazali, że chińskim badaczom z obszaru chemii udało się awansować w światowych rankingach cytowań dzięki intensyfikacji wzajemnych, „grzecznościowych” cytatów (Zhou, Leydesdorff 2006: 89-91), co może świadczyć nie tyle o naukowej doniosłości ich prac badawczych, co o występowaniu planowanych działań, mających doprowadzić do wzrostu prestiżu chińskich ośrodków naukowych. Dostępne techniki bibliometryczne nie oferują sposobów łatwego „oczyszczenia” danych z cytatów publikacji pracowników tego samego ośrodka naukowego (cytowania publikacji własnych nie są zwykle uwzględniane w bazach), choć bywa, że właśnie one stanowią znaczącą część cytatów. Kolejnym problemem jest nieetyczne (jednak prawdopodobnie relatywnie częste) zjawisko kopiowania referencji z innych prac do swojego tekstu bez zapoznawania się z treścią w ten sposób „cytowanych” publikacji (Nowak 2006: 104). Analizy empiryczne wskazują dodatkowo na częste przypadki przytaczania w artykułach wiedzy, możliwej do zdobycia wyłącznie w wyniku lektury określonej publikacji, bez jej jednoczesnego cytowania. Skala tego zjawiska w niektórych obszarach naukowych pozwala niestety na kwestionowanie podstawowych założeń, przyświecających analizie cytowań (MacRoberts, MacRoberts 2010: 5).

Warto pamiętać, że analiza cytowań oferuje jedynie pośrednie dowody współzależności między obszarami badawczymi, a prowadzone w jej ramach prace analityczne abstrahują od specyfiki poddanych analizie badań i ograniczają się do niewielkiego podzbioru dostępnych publikacji (Kostoff 2003: 6), co może być źródłem błędów. Ze względu na to, że reguły dotyczące cytowania są w świecie nauki „zrytualizowane” (Cockburn, Henderson 1998: 165), niektórzy uważają, że współautorstwo badań przez reprezentantów instytucji naukowych i przedsiębiorstw jest lepszym wyznacznikiem transferu wiedzy z uczelni do biznesu niż cytowania (Cockburn, Henderson 1998: 165).

Typowe problemy analizy cytowań obejmują (MacRoberts, MacRoberts 1989: 343-346):

- unikanie cytowania formalnych źródeł wpływu (zwłaszcza dotyczących fundamentalnych odkryć naukowych i podstaw teoretycznych dyscypliny),

- tendencje do cytowania określonych autorów przy jednoczesnej rezygnacji z cytowania innych,
- unikanie cytowania nieformalnych źródeł wpływu (zwłaszcza informacji zdobytych w drodze komunikacji między naukowcami),
- skłonność do auto-cytowania, zaburzająca wyniki analiz (choć współcześnie bazy eliminują auto-cytowania z zestawień),
- zróżnicowanie cytowań na pozytywne i negatywne,
- istotne różnice we wzorcach cytowania pomiędzy dyscyplinami naukowymi i rodzajami tekstów (np. prace dotyczące metod badawczych cytowane częściej niż prace, dotyczące rozwoju teorii), co uniemożliwia porównania,
- problemy interpretacyjne wynikające z błędnego rozróżnienia lub utożsamienia nazwisk i imion autorów.

Potencjał zastosowań analizy cytowań ogranicza też dostępność odpowiednich baz danych. Spośród znanych baz publikacji naukowych i streszczeń, cytowania gromadzą bazy *Web of Science* oraz *Scopus*. Ściągnięcie cytowań dla większej liczby publikacji na lokalny dysk komputera stwarza dodatkowe problemy techniczne, podobnie jak ich dalsze przetwarzanie. Relatywnie łatwa jest identyfikacja cytowań dotyczących pojedynczego artykułu naukowego (por. Rysunek 23), w przeciwieństwie do przetwarzania większej liczby danych o cytowaniach. Należy pamiętać, że liczba cytowań wykazywana przez daną bazę bibliometryczną jest zależna od zakresu publikacji, jakie obejmuje dana baza (ang. *coverage*) (Bar-Ilan, Levene, Lin 2007: 26).

Przykładowo, zgodnie z analizami empirycznymi (Moed 2005: 3; 138), baza *Web of Science* zawiera bardzo dobrą reprezentację cytowań w obszarach fizyki, chemii, biologii molekularnej i biochemii, nauk o człowieku i medycyny klinicznej, dobrą w obszarach nauk stosowanych i inżynieryjnych, nauk biologicznych o zwierzętach i roślinach, nauk o ziemi, matematyki, psychologii i nauk społecznych związanych z medycyną i zdrowiem, a jedynie umiarkowaną w innych naukach społecznych, politycznych, pedagogicznych i humanistycznych, co oznacza zróżnicowany potencjał prowadzenia analizy cytowań w zależności od badanego obszaru.

Citing Articles

Title: [Lattice parameters of gallium nitride](#)
 Author(s): Leszczynski, M
 Source: **APPLIED PHYSICS LETTERS** Volume: 69 Issue: 1 Pages: 73-75 Published: JUL 1 1996
[Citation Map](#)

The above article has been cited by the articles listed below.
Note: The Times Cited count is calculated across all *Web of Science* editions. [More information](#).

Results: **166** Page 1 of 17 Go Sort by: Latest Date

Print E-mail Add to Marked List Save to EndNote@Web Save to EndNote, RefMan, ProCite Save to RefWorks more options Analyze Results

Refine Results

Search within results for

 Search

Subject Areas Refine

- PHYSICS, APPLIED (66)
- MATERIALS SCIENCE, MULTIDISCIPLINARY (47)
- PHYSICS, CONDENSED MATTER (40)
- CRYSTALLOGRAPHY (19)
- PHYSICS, MULTIDISCIPLINARY (12)
- [more options / values...](#)

Document Types Refine

- ARTICLE (105)
- PROCEEDINGS PAPER (50)
- REVIEW (9)
- LETTER (2)
- [more options / values...](#)

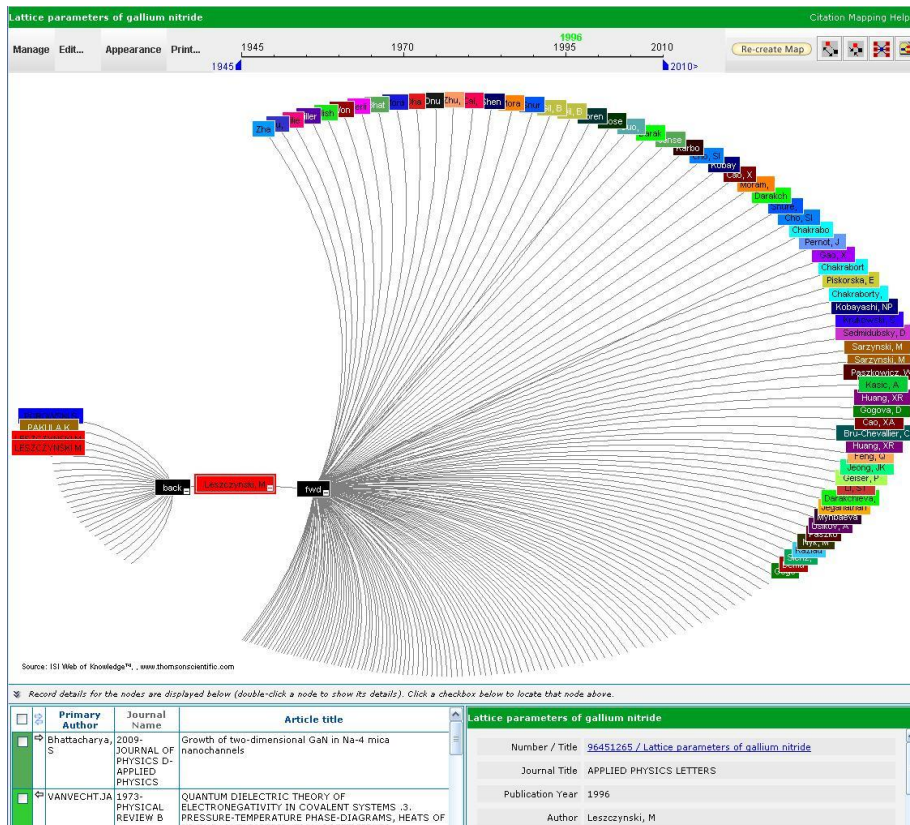
Authors

Source Titles

Publication Years

- 1. Title: [Indium distribution and light emission in wurtzite InGaN alloys: Several-atom In-N clusters](#)
 Author(s): Zhang SA, Shi JJ, Zhu SG, et al. Source: **PHYSICS LETTERS A** Volume: 374 Issue: 47 Pages: 4767-4773 Published: OCT 25 2010
 Times Cited: 0 [Full Text](#)
- 2. Title: [Basal-plane Stacking Faults in Non-polar GaN Studied by Off-axis Electron Holography](#)
 Author(s): Liu LZY, Rao DVS, Kappers MJ, et al.
 Conference Information: 16th International Conference on Microscopy of Semiconducting Materials, MAR 17-20, 2009 Univ Oxford, Oxford, ENGLAND Source: **16TH INTERNATIONAL CONFERENCE ON MICROSCOPY OF SEMICONDUCTING MATERIALS** Book Series: Journal of Physics Conference Series Volume: 209 Pages: - Published: 2010
 Times Cited: 0 [Full Text](#)
- 3. Title: [Extended defects in bulk GaN and III-nitrides grown on this substrate](#)
 Author(s): Liliental-Weber Z
 Conference Information: 6th International Workshop on Bulk Nitride Semiconductors, AUG 23-28, 2009 Galindia, POLAND Source: **JOURNAL OF CRYSTAL GROWTH** Volume: 312 Issue: 18 Pages: 2599-2606 Published: SEP 1 2010
 Times Cited: 0

Rysunek 23: Lista publikacji cytujących wybrany artykuł naukowy. Źródło: *Web of Science*.



Rysunek 24: Graficzna forma prezentacji publikacji cytujących wybrany artykuł naukowy (mapa cytowań). Źródło: *Web of Science*.

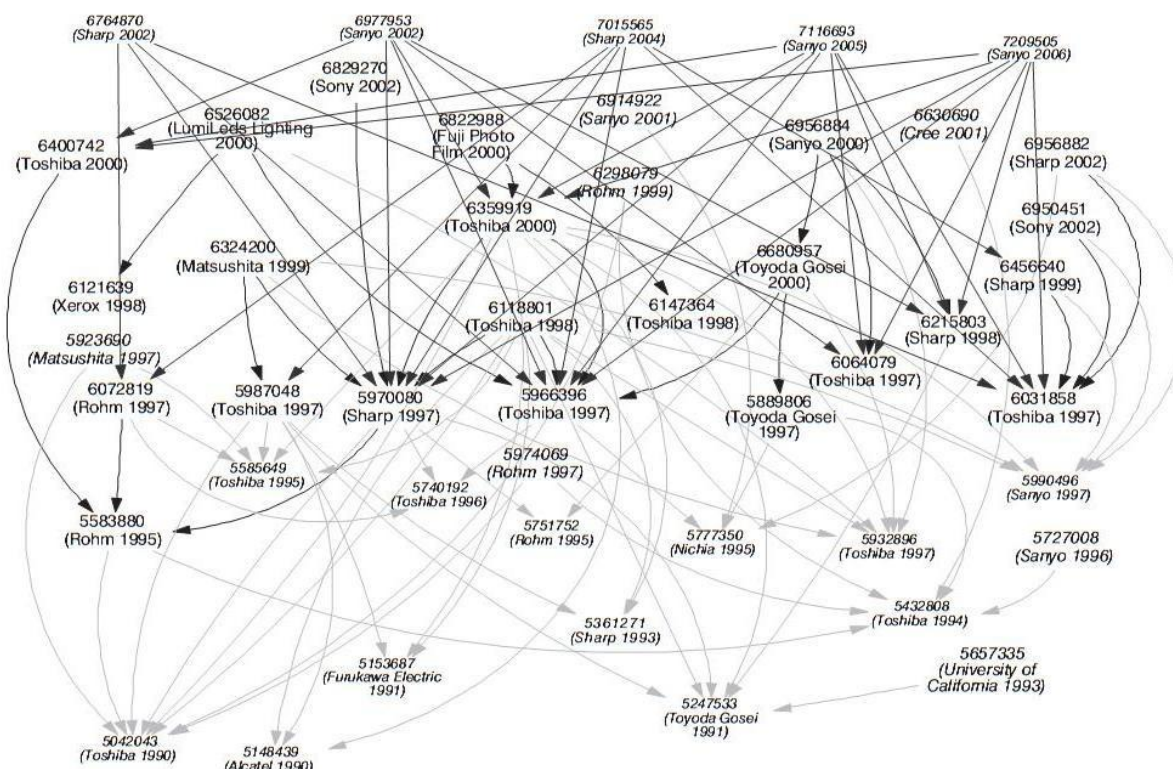
Szczególnie trudny problem analityczny stwarza zaobserwowane zjawisko częstego wskazywania w cytacie tylko pierwszego wymienionego autora zamiast wszystkich autorów tekstu - poprawne prowadzenie analizy cytowań wymagałoby weryfikacji danych metryczkowych cytowanych publikacji (Hicks i in. 2002: 44), co jest niezwykle czasochłonne.

Dodatkowym problemem przy porównywaniu danych dotyczących cytowań są nieproporcjonalne zmiany liczby odwołań do danej publikacji wraz z upływem czasu. Długość okresu, jaki minął od momentu publikacji tekstu naukowego do momentu gromadzenia danych do analiz bibliometrycznych powinna być brana pod uwagę w przypadku analiz - starsze teksty mają większe szanse być dostrzeżone niż teksty niedawne. Uwzględnienie tego wymogu znacznie utrudnia agregację danych o cytowaniach dla większej liczby tekstów tego samego autora, pochodzących z tej samej instytucji, kraju lub obszaru badawczego. Poprawne analizy wymagałyby ujednoczenia danych poprzez przekształcenia, które byłyby różne w odniesieniu do pojedynczych analizowanych obiektów. Najprostszym sposobem na ujednoczenie danych o cytowaniach jest podział liczby cytowań w danym roku przez liczbę lat, które upłynęły od momentu publikacji (Guo, Huang, Porter 2010: 201), co w praktyce okazuje się niezwykle czasochłonnym zadaniem w odniesieniu do dużych zbiorów danych.

Sposobem na udoskonalenie analizy cytowań może być też uszczegółowienie jednostki analizy - zamiast śledzenia cytowań całych artykułów, można skoncentrować się na fragmentach tekstu, które są cytowane, co pozwala na ustalenie kontekstu cytowania i odzwierciedlenie "siły" cytowania (Callan, Hockema, Eysenbach 2010: 1130). Tego typu analizy są niestety bardzo pracochłonne i nie dają się w łatwy sposób zautomatyzować. Kostoff (2002: 68) zauważa też, że "wbrew większości popularnych poglądów, wiedza i doświadczenia techniczne osoby analizującej cytowania mogą mieć znaczący wpływ na jakość rezultatów", wskazując m.in. na subiektywny charakter decyzji o przypisaniu obiektów do obszarów tematycznych.

Większe możliwości wykorzystania w analizach rozwoju nauki i technologii oferuje analiza cytowań patentów. Zgłoszenia patentowe zawierają odwołania do stanu techniki (ang. *prior art*) czyli wcześniejszych wynalazków i publikacji naukowych z pokrewnych obszarów. Jeśli zgłaszający przeoczy lub przemilczy fakt istnienia pokrewnego wynalazku i nie wykaże oryginalności swojej propozycji, jego zgłoszenie może zostać odrzucone przez urząd patentowy. Silna ekonomiczna motywacja do starannego przygotowania kompletnego wniosku sprawia, że cytaty w bazach patentowych stały się wartościowym źródłem informacji. Analizy cytowań patentów pozwalają na wyodrębnienie tzw. nisz innowacyjnych (Podolny, Stuart 1995: 1229) z kluczowymi wynalazkami, stanowiącymi punkt odniesienia dla pokrewnych zgłoszeń patentowych. Dokumenty patentowe udostępniane przez *USPTO* i serwis *Google Patents* zawierają informacje o cytowaniach. Analogiczne informacji nie są niestety łatwe do identyfikacji w europejskich dokumentach patentowych, udostępnianych nieodpłatnie przez *EPO*. Istnieją istotne różnice w wykorzystaniu cytowań w zgłoszeniach patentowych do *USPTO* i *EPO*. *USPTO* wymaga podania w zgłoszeniu wszystkich istotnych odwołań do stanu techniki, motywując zgłaszających do opracowywania długich list cytowań, obejmujących niekoniecznie adekwatne patenty i publikacje, podczas gdy w przypadku

zgłoszeń do *EPO*, odwołania do wcześniejszych patentów wprowadzają pracownicy Urzędu, ograniczając się wyłącznie do niewielkiej liczby bezpośrednio powiązanych i zbliżonych wynalazków (Breschi, Lissoni 2004: 618). Analizy porównawcze wykazały, że zgłoszenia patentowe w *USPTO* cytują średnio 13 wcześniejszych patentów oraz 3 źródła nie będące patentami (w tym publikacje naukowe), podczas gdy zgłoszenia w *EPO* odwołują się przeciętnie do 4 patentów i 1 innego źródła (Breschi, Lissoni 2004: 618). Patenty w *USPTO* zawierają bardziej starannie dobrane cytowania niż w przypadku dokumentów w *EPO* m.in. ze względu na rygorystyczne wymogi prawne, obowiązujące w Stanach Zjednoczonych (Hsieh 2011: 386). Szczegółową dyskusję różnic w cytowaniach zawartych w dokumentach patentowych *USPTO* i *EPO* oferuje Meyer (2000).

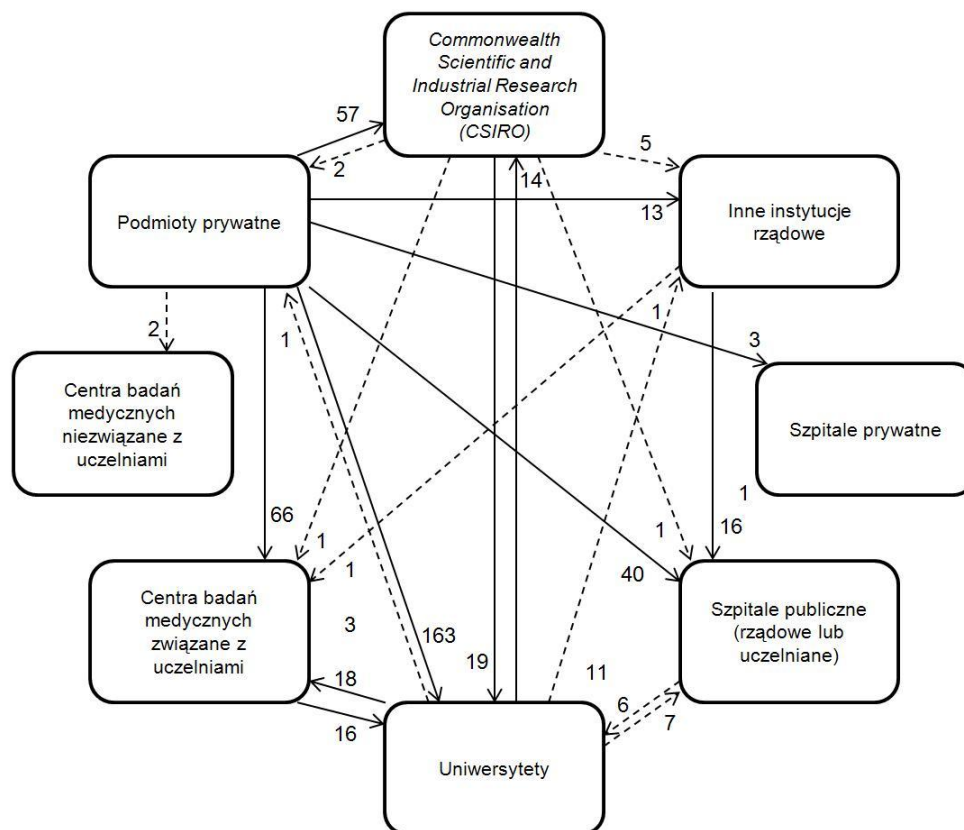


Rysunek 25: Sieć cytowań patentów z *USPTO*, dotyczących laserów półprzewodnikowych opartych na azotku galu. Źródło: Klineciewicz (2010: 148).

Dodatkowym obszarem do możliwych eksploracji, nadal rzadko wykorzystywanym przez badaczy, jest analiza cytowań publikacji naukowych w zgłoszeniach patentowych. Fakt docenienia praktycznej przydatności konkretnego tekstu naukowego przez wynalazcę wydaje się dobrym kryterium wyodrębnienia bardziej wartościowych i przydatnych praktycznie prac badawczych.

Możliwa jest też analiza cytowań między grupami instytucji - przykładem może być analiza cytowań patentów udzielonych przez *USPTO* podmiotom z Australii, która ujawniła dominującą rolę rządowych instytucji badawczych (w tym *Australian Research Council* i *Commonwealth Scientific and Industrial Research Organisation - CSIRO*) oraz niewielkie znaczenie australijskich wynalazków, co odzwierciedla ograniczona liczba cytowań (Rysunek

26). Analiza przygotowana przez rząd Australii w 2002 roku była podstawą do publicznej debaty na temat sposobów stymulowania innowacyjności (Hicks i in. 2002: 140-141).



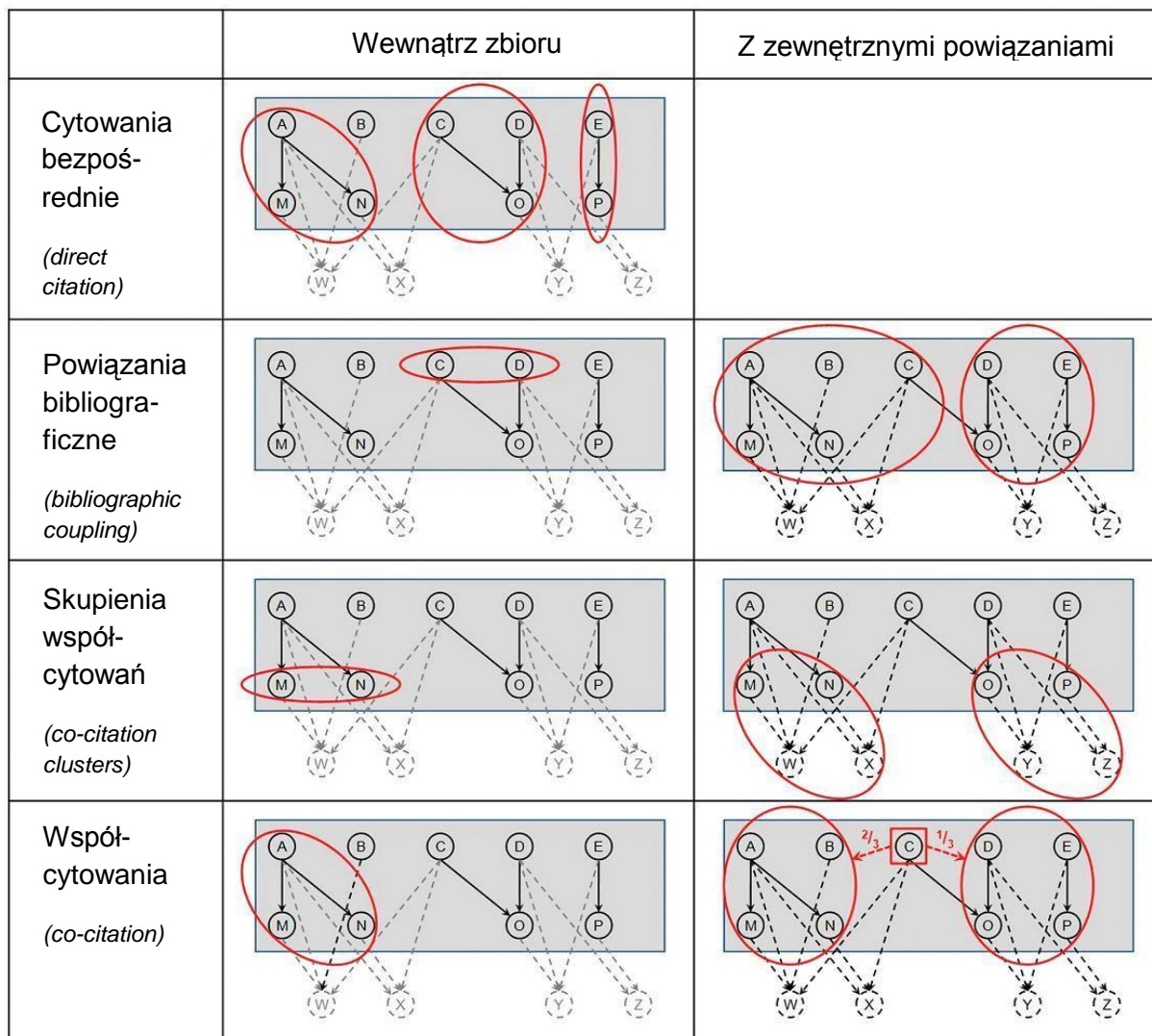
Rysunek 26: Sieć cytowań patentów pomiędzy grupami organizacji australijskich. Źródło: Hicks (2002: 141).

8.8. Warianty analizy cytowań

Pojęcie „analiza cytowań” może dotyczyć czterech wariantów analiz, które różnią się zakresem rozpatrywanych danych. Te cztery warianty obejmują:

- analizę cytowań bezpośrednich (ang. *direct citation*) – czyli budowę macierzy z danymi, odzwierciedlającymi przypadki bezpośrednich cytowań tekstów jednego autora przez innych;
- analizę powiązań bibliograficznych (ang. *bibliographic coupling*) – czyli identyfikację publikacji, które cytują ten sam artykuł, czyli odwołują się do tych samych źródeł wiedzy; ten wariant analizy pozwala zwykle wyodrębnić najnowsze publikacje w danej sieci cytowań;
- analizę skupień współcytowań (ang. *co-citation clustering*) – czyli wskazanie, do jakich publikacji-źródeł wiedzy odwołują się wybrane artykuły, co zwykle pozwala wyodrębnić najstarsze publikacje w danej sieci cytowań;
- analizę współcytowań (ang. *co-citation analysis*) – obejmującą wybrane publikacje wraz z przywoływanymi przez nie źródłami wiedzy.

Rozróżnienie tych rodzajów analiz omawia Small (1973: 265), a graficznie są one zaprezentowane na Rysunek 27. W publikacjach dotyczących bibliometrii niektóre z podanych powyżej nazw wariantów analizy cytowań bywają czasem używane w odmiennym niż zaprezentowane znaczeniu.



Rysunek 27: Warianty analizy cytowań. Źródło: Boyack, Klavans (2010: 2390).


Najprostszym scenariuszem analizy cytowań pozostaje odzwierciedlenie bezpośrednich zależności dla wybranej publikacji – cytowanych przez nią źródeł wiedzy oraz publikacji odwołujących się do niej. Forma graficzna takiej analizy określana jest jako „historiograf” (Small 1999: 800) i przypomina drzewo z korzeniami (publikacjami cytowanymi), pniem (główną poddaną analizom publikacją) i konarami (publikacjami cytującymi).

Zestawienie wariantów analiz przynoszą istotnie różne wyniki w zależności od tego, czy analityk ograniczy się do przetwarzania publikacji z wybranego wcześniej zbioru (wyników określonego polecenia wyszukiwania), czy też zdecyduje się na włączenie do analiz innych pozycji, które wykraczają poza ten zbiór i zostały zidentyfikowane jako artykuły cytowane lub cytujące publikacje ze zbioru. Scenariusz analiz opartych na szerszym zbiorze obiektów pozwala na lepsze odzwierciedlenie tendencji w badaniach i rozwoju technologii, może


jednak okazać się trudny do praktycznej realizacji (wymagać serii szczegółowych wyszukiwań w bazie). Interesujący wydaje się potencjał prowadzonych na dużą skalę analiz skupień współcytowań jako sposobu identyfikacji pokrewnych prac badawczych i wyodrębniania obszarów badań, co pokazali m.in. Boyack i Klavans (2010). W literaturze pojawiają się jednak również głosy krytykujące takie analizy jako prowadzące do wyodrębnienia artefaktów - grup publikacji, które w wielu przypadkach nie odzwierciedlają rzeczywistych podziałów w nauce (Hicks 1987). Prowadzenie analiz skupień współcytowań wykracza jednak poza podstawowe techniki analityczne, wymaga zgromadzenia i systematycznej analizy bardzo dużych zbiorów danych przy wykorzystaniu dedykowanego oprogramowania i zostanie opisane w dalszej części niniejszej książki.

Klasyczna analiza cytowań nie wprowadza rozróżnienia między cytowaniami "lepszymi" i "gorszymi". Przypadek cytowania artykułu przez tekst z mało znanego czasopisma może być zgodnie z opisanym powyżej podejściem interpretowany tak samo, jak odwołanie do tego artykułu przez światowej klasy naukowca - pod warunkiem, że wszystkie te czasopisma będą indeksowane w tej samej bazie bibliometrycznej (Ding i in. 2009: 2229). Sposobem na rozwiązanie tego problemu jest wykorzystanie algorytmu różnicującego "jakość" cytowań poprzez sprawdzenie, czy sam tekst cytujący jest również wielokrotnie cytowany. Algorytm dokonuje oceny ważonej cytowań przy wykorzystaniu liczby cytowań tekstu cytującego. Internetowa odmiana tego algorytmu jest wykorzystywana przez firmę *Google* do ustalania kolejności wyświetlania wyników wyszukiwań stron internetowych i nosi nazwę *PageRank*. *PageRank* został dostosowany do specyfiki baz stosowanych do celów analiz bibliometrycznych (Ding i in. 2009), jednak należy pamiętać, że prowadzenie tego typu analiz wymaga znaczących nakładów pracy i zgromadzenia ogromnego zbioru danych, zawierającego cytowania publikacji cytujących na wielu zagłębionych poziomach. Można przypuszczać, że w najbliższych latach na wykorzystanie analogicznego algorytmu zdecyduje się jeden z dostawców baz bibliograficznych – dopiero takie działanie pozwoli na wykorzystanie zalet algorytmu *PageRank* w praktycznych projektach analitycznych. W literaturze pojawiają się również alternatywne propozycje, np. metoda ważenia cytowań bezpośrednich Perssona (2010).

Analiza cytowań jest uważana za sposób odzwierciedlenia popularności i znaczenia poszczególnych publikacji. W ostatnich latach pojawiają się też nowe tendencje w badaniach bibliometrycznych, związane z istotnymi zmianami w branży wydawniczej, wynikającymi z dostępności elektronicznych wersji artykułów naukowych w Internecie oraz wykorzystywania przez naukowców baz zawierających treść artykułów. Niektórzy uważają, że wskaźniki dotyczące liczby pobrań artykułu z bazy (ang. *usage metrics*) mogą być lepszymi wyznacznikami jego popularności i znaczenia w środowiskach badawczych niż cytowania (Schloegl, Gorraiz 2011: 161). Liczby pobrań łatwiej jest interpretować i aktualizować w bazach, a analizy dotyczące artykułów z wybranych obszarów nauk medycznych wykazały bardzo wysokie współczynniki korelacji pomiędzy liczbami pobrań i cytowań (Schloegl, Gorraiz 2011: 165-166). Można przewidywać, że stopniowo upowszechni się w bibliometrii tendencja do wykorzystywania omawianych danych do celów analitycznych - tym bardziej, że analogiczne algorytmy, porządkujące dokumenty według częstości pobrań, stosowane są już przez firmy, zajmujące się wyszukiwaniem treści w Internecie.

- 3 [Supporting RFID-based item tracking applications in Oracle DBMS using a bitmap datatype](#)
Ying Hu, Seema Sundara, Timothy Chorma, Jagannathan Srinivasan
August 2005 **VLDB '05**: Proceedings of the 31st international conference on Very large data bases
Publisher: VLDB Endowment
Full text available:  Pdf (200.93 KB)
Bibliometrics: Downloads (6 Weeks): 221, Downloads (12 Months): 2581, Downloads (Overall): 4725, Citation Count: 12

Radio Frequency Identification (RFID) based item-level tracking holds the promise of revolutionizing supply-chain, retail store, and asset management applications. However, the high volume of data generated by item-level tracking poses challenges to ...

- 4 [TrACS: transceiver architecture and wireless channel simulator](#)
Chithrupa Ramesh, Ana Rusu, Mohammed Ismail, Mikael Skoglund
September 2007 **SBCCI '07**: Proceedings of the 20th annual conference on Integrated circuits and systems design
Publisher: ACM
Full text available:  Pdf (981.79 KB)
Bibliometrics: Downloads (6 Weeks): 135, Downloads (12 Months): 212, Downloads (Overall): 326, Citation Count: 1

This paper presents the design of a system-level simulator for radio receivers, including receiver circuits, in Matlab. The system level outlook offers a better characterization of circuit design, as the signal processing in the digital receiver is asymmetric ...

Rysunek 28: Wskaźniki dotyczące liczby pobrań artykułu z bazy jako sposób na określenie popularności publikacji. Źródło: ACM.

8.9. Podstawowe wskaźniki bibliometryczne

Częstym sposobem wykorzystania wyników analizy cytowań jest odwoływanie się do typowych wskaźników, obliczanych i publikowanych przez bazy bibliograficzne. Wspomniany wcześniej Eugene Garfield zaproponował wskaźnik wpływu czasopisma (ang. *journal impact factor*) jako podstawę do budowy indeksu cytowań literatury naukowej i wyodrębniania najbardziej wartościowych publikacji (Garfield 1955). W latach 60-tych XX wieku Garfield wspólnie z Irvingiem H. Sherem stworzyli *Science Citation Index*, zawierający m.in. wartości wskaźnika wpływu dla poszczególnych czasopism naukowych (Garfield 2005: 1). *Impact factor* obliczany jest jako liczba cytowań publikacji z danego okresu podzielona przez liczbę publikacji z tego samego okresu, przy czym analiz są zwykle prowadzone w odniesieniu do dwóch lat poprzedzających rok obliczania wskaźnika (Garfield 2005: 5). Przykładowo, obliczenie wskaźnika dla roku 2009 oznaczać więc będzie zestawienie opublikowanych w roku 2009 cytowań artykułów z lat 2007-2008 z łączną liczbą artykułów opublikowanych w latach 2007-2008. Analizy empiryczne wskazują, że tylko nieliczne artykuły są "odkrywane" po wielu latach i dopiero wtedy narastają ich cytowania (tzw. efekt Mendla), a większość znaczących dla danego obszaru badań artykułów uzyskuje liczne cytowania w ciągu pierwszych lat po opublikowaniu (Moed 2005: 81). Przekonanie o tym, że naukowcy mają tendencję do częstszego cytowania nowszych publikacji (Price 1965: 513), stało się podstawą konstrukcji wskaźnika wpływu. Poniżej zaprezentowany jest przykładowe obliczanie wartości wskaźnika dla roku 2009:

$$\text{IF 2009} = \frac{\text{Cytowania artykułów z lat 2007–8, pochodzące z roku 2009}}{\text{Liczba artykułów opublikowanych w latach 2007–8}},$$

Dodatkowe przykłady oraz informacje dotyczące obliczeń oferują m.in. Garfield (1972), Moed (2005: 92) i Coleman (2007). Wskaźnik wpływu może być również obliczany dla indywidualnego badacza, instytucji lub kraju, jednak zgromadzenie danych do takich obliczeń jest znacznie bardziej skomplikowane. Garfield i jego kontynuatorzy próbowali też zmodyfikować *impact factor* w oparciu o dłuższy okres analiz, w tym 5 i 15 lat, jednak dwuletnie ramy czasowe okazały się najbardziej dogodnym rozwiązaniem (Leydesdorff 2009: 1327). Platforma *Web of Knowledge* oferuje regularnie aktualizowane zestawienia wskaźników wpływu dla czasopism naukowych, indeksowanych w tej bazie. Dla konkurencyjnej bazy *Scopus*, odpowiednikiem *impact factor* jest *SCImago Journal Rank (SJR)*, publikowany przez hiszpańską grupę badawczą *SCImago* niezależną od firmy *Elsevier*, wydawcy bazy *Scopus*. Konstrukcja wskaźnika *impact factor*, jego oparcie na arbitralnie wybranej grupie czasopism w bazie *Web of Science* oraz brak przejrzystości kalkulacji stanowią podstawy do jego częstej krytyki (Garcia-Perez 2010: 2070).

Wydawca bazy *Scopus* korzysta z udoskonalonego wskaźnika *SNIP* (ang. *source normalized impact per paper*), który jest interpretowany jako „miara cytowań w kontekście” (ang. *contextual citation measure*). Sposób obliczania wskaźnika *SNIP* zaproponował holenderski badacz Henk Moed z jednego z najważniejszych ośrodków bibliometrycznych Europy na uniwersytecie w Lejdzie. *SNIP* odwołuje się do pojęcia potencjału cytowań (ang. *citation potential*), który odzwierciedla częstotliwość cytowania publikacji w danym obszarze naukowym (Moed 2010: 267). Poszczególne dyscypliny nauki różnią się skłonnością do cytowania i długością przeciętnych spisów literatury cytowanej w artykule. Skala różnic między obszarami badań może wpływać na wyniki ewaluacji - przykładowo, artykuły z obszarów biochemii i biologii molekularnej były na przestrzeni lat średnio cytowane nawet 10-krotnie częściej niż artykuły z obszaru inżynierii lądowej, a prawie 3-krotnie częściej niż artykuły z obszaru fizyki stosowanej (Hicks i in. 2002: 41). *SNIP* wykorzystuje dane określone jako "względny potencjał bazy do cytowania" (ang. *relative database citation potential*) do przekształcenia danych wejściowych, odzwierciedlających liczbę cytowań publikacji z danego pisma naukowego. Potencjał cytowań w ramach danej bazy zależy też od zakresu tematycznego, jaki pokrywa baza oraz doboru indeksowanych czasopism. Moed (2010) oferuje szczegółowy opis sposobów obliczania *SNIP* oraz porównanie wskaźnika z alternatywnymi metodami oceny wpływu w bibliometrii. W ramach intensywnej konkurencji między firmami *Thomson Reuters* i *Elsevier*, każda z komercyjnych baz cytowań opiera się obecnie na odmiennym wskaźniku wpływu – *Web of Science* wykorzystuje *impact factor*, a *Scopus* - *SNIP*. Należy zaznaczyć, że *SNIP* spotyka się z krytyką niektórych znanych specjalistów z zakresu bibliometrii z przyczyn metodologicznych (Leydesdorff, Opthof 2010) - uzyskany wskaźnik nie pozwala przykładowo na obliczenie standardowego odchylenia zmiennej, nie może być też wykorzystywany do testów istotności statystycznej (Leydesdorff, Opthof 2011), co zdecydowanie ogranicza jego zastosowania w pracach badawczych z obszaru bibliometrii i sprawia, że podobnie jak *impact factor* posiada istotne ograniczenia i nie jest rozwiązaniem optymalnym.

Wskaźniki *impact factor* i *SNIP* dla czasopisma naukowego nie powinny być wykorzystywane jako podstawa do oceny dorobku badawczego osoby, która opublikowała tekst w tym piśmie (Garcia-Perez 2010: 2070). Relatywnie niedawno zaproponowanym

wskaźnikiem bibliometrycznym jest indeks h (indeks Hirscha), pozwalający na prostą ocenę dorobku pojedynczego naukowca. Indeks h opiera się na prostej kalkulacji: naukowiec, którego dorobek charakteryzuje wartość indeksu na poziomie x , opublikował x artykułów naukowych, a każdy z nich był cytowany co najmniej x razy (Leydesdorff 2009: 1328). Szczegółową dyskusję uwarunkowań teoretycznych i metod obliczania indeksu h oferuje Hirsch (2005), a sam indeks jest obecnie automatycznie obliczany dla każdego autora przez bazy *Web of Science* i *Scopus*.

Wskaźniki oferują szybki sposób oceny dorobku naukowego, należy jednak pamiętać o ich poważnych ograniczeniach – nie powinny być wykorzystywane jako jedyna podstawa do oceny osiągnięć konkretnego badacza. Ograniczone są też korzyści z wykorzystania omawianych wskaźników do obserwacji badań naukowych i rozwoju technologii – służą one raczej syntezie i ewaluacji niż odzwierciedlaniu specyfiki obszarów badawczych (tj. analizom o charakterze eksploracyjnym). W szczególności, nie wydaje się poprawne metodologicznie wykorzystywanie pozbawionych kontekstu wskaźników cytowań przy podejmowaniu decyzji o zatrudnieniu lub promocji naukowca (Moed 2005: 29).

Zastosowania bibliometrii do celów ewaluacji badań nie są przedmiotem niniejszej publikacji. Indywidualni naukowcy, ośrodki badawcze i obszary geograficzne mogą podlegać ocenie w oparciu o dane bibliometryczne i obliczane:

- wskaźniki produktywności (liczby publikacji),
- wskaźniki wpływu (cytowania tych publikacji),
- wskaźniki specjalizacji (odzwierciedlające udział publikacji na określony temat we wszystkich publikacjach),
- wskaźniki konkurencyjności (porównujące względne przewagi różnych osób lub grup, wynikające ze specjalizacji w danym obszarze badawczym).

Ewaluacja bibliometryczna odwołuje się do specyficznych technik, które nie są wykorzystywane w projektach mających na celu obserwację badań naukowych i rozwoju technologii, dlatego nie zostaną szczegółowo opisane w niniejszej książce. Przykładowo, istotne znaczenie dla oceny indywidualnego dorobku będzie miała decyzja o sposobie przypisywania wag do publikacji, autorstwa kilku osób. Jeśli do dorobku każdego ze współautorów doliczana jest kolejna publikacja, mówimy o tzw. *whole counting*, podczas gdy obliczanie wartości ułamkowej, proporcjonalnej do liczby współautorów, lub odpowiadającej kolejności wymieniania nazwisk współautorów, to tzw. *fractional counting* (Okubo 1997: 21). *Whole counting* wydaje się bardziej logiczne przy porównaniach, jest też łatwiejsze do stosowania i interpretacji, a jednocześnie jest najczęściej wykorzystywane w publikowanych analizach bibliometrycznych. Istotną wadą *fractional counting* jest zniekształcanie skali współpracy między organizacjami – przykładowo, im bardziej umiędzynarodowione badania, tym mniejsze wartości liczbowe otrzymują partnerzy badawczy z kolejnych krajów (Okubo 1997: 21). Opisywane w kolejnych rozdziałach techniki opierają się na wykorzystaniu *whole counting*, możliwa jest jednak transformacja zbioru danych w programach komputerowych w celu zastosowania innych algorytmów przypisywania wag poszczególnym autorom.

9. Zaawansowane techniki analityczne w bibliometrii

9.1. Zaawansowane techniki bibliometryczne

W niniejszym rozdziale omówione zostaną złożone techniki analityczne, wykorzystywane do przetwarzania danych bibliometrycznych. Opisywanych analiz nie da się zwykle przeprowadzić bez odpowiedniego wsparcia narzędziowego, oferowanego przez wyspecjalizowane oprogramowanie. Punktem wyjścia do prowadzenia tych analiz jest wygenerowanie podzbioru danych bibliometrycznych - wyników wyszukiwania w bazie oraz jego import do odpowiedniej aplikacji komputerowej.

W dalszej części rozdziału omówione zostaną techniki:

- *tech mining*,
- tomografii bazodanowej,
- *Literature-Based Discovery*,
- TRIZ,
- identyfikacji nowych trendów i tworzenia map badań,
- wykorzystywania zaawansowanych wskaźników bibliometrycznych.

9.2. *Tech mining*

Pojęcie *tech mining* można przetłumaczyć jako „eksplorację danych o technologiach”, a samo podejście opiera się na wykorzystaniu znanych w informatyce technik *data mining* (eksploracji danych) i *text mining* (eksploracji danych tekstowych) w odniesieniu do działalności badawczo-rozwojowej (danych bibliometrycznych). Termin został spopularyzowany jako tytuł książki Portera i Cumminghama (2005). Analizy oparte na podejściu *tech mining* zakładają pracę ze zbiorem danych, pobranych z bazy w celu generowania zestawień, porównań, odzwierciedlaniu wzajemnych zależności między kilkoma zmiennymi, w tym przy wykorzystaniu zaawansowanych metod statystycznych. Umowną granicą między podstawowymi analizami bibliometrycznymi a zastosowaniem *tech mining*-u może stanowić konieczność wykorzystania dodatkowego, specjalistycznego oprogramowania i pracy ze zbiorem danych, pobranym z bazy, a nie tylko wykorzystywaniem informacji o zbiorczej liczbie rezultatów wyszukiwania w bazie. Najpopularniejszym i najbardziej rozbudowanym narzędziem wspierającym *tech mining* jest program komputerowy *VantagePoint*, pierwotnie opracowany jako program *TechOASIS* na zamówienie amerykańskiego sektora obronnego. Współautor wspomnianej wcześniej książki o *tech mining*-u, Alan Porter, był jednym z twórców tego programu w trakcie pracy dla *Georgia Institute of Technology*.

Do praktycznych sposobów wykorzystania wyników analiz *tech mining* należą:

- dobre zrozumienie specyfiki danego obszaru badawczego - wiedzy i umiejętności niezbędnych do prowadzenia badań w tym obszarze, technologii składowych (niezbędnych do stworzenia własnego rozwiązania), technologii komplementarnych (mogących to rozwiązanie uzupełnić i podnieść jego wartość dla odbiorców),

- identyfikacja aktorów na scenie innowacji - potencjalnych partnerów, konkurentów, dostawców i odbiorców,
- obserwacja zmian w czasie, nowopowstających zjawisk i obszarów, które osiągnęły dojrzałość i mogą nie być atrakcyjne dla kolejnych badaczy,
- podstawa do prowadzenia analizy *SWOT* dla własnego zespołu badawczego, instytucji lub kraju - jak również przygotowania analogicznych analiz *SWOT* dla innych podmiotów (*SWOT* wymaga zrozumienia własnych przewag i słabości w porównaniu do pozostałych aktorów na scenie innowacji, jak również świadomości szans i zagrożeń, wynikających z rozwoju badań naukowych i technologii),
- okazja do identyfikacji obiecujących obszarów badań, w tym obszarów interesujących wielu konkurentów oraz nisz tematycznych, na razie jeszcze nie zagospodarowanych przez znaczących graczy (odpowiednik "strategii błękitnego oceanu", w której sukces jest wynikiem inteligentnego wyboru specjalizacji, która do tej pory nie była rozważana przez innych uczestników rynku, a stwarza szansę na wypracowanie silnej pozycji rynkowej).

Przykłady zastosowań *tech mining* obejmują m.in.:

- wywiad technologiczny (*technology intelligence*) firmy lub instytucji naukowej - zrozumienie własnej pozycji, identyfikacja innych graczy oraz obserwacja ich badań,
- analizy strategiczne dla instytucji badawczej – pozwalające na świadomy wybór kierunków badawczych,
- wkład do polityki naukowej kraju - zwłaszcza programowania sposobów wydatkowania funduszy oraz podejmowania decyzji o alokacji środków na poszczególne obszary badawcze.

Należy zaznaczyć, że większość omawianych w niniejszej książce zastosowań technik bibliometrycznych dla celów wspierania polityki naukowej oraz działalności badawczo-rozwojowej może być zrealizowana przy wykorzystaniu technik należących do obszaru *tech mining* oraz standardowych, dostępnych narzędzi informatycznych.

Tech mining jest przejawem stosowania bibliometrii w celu rozwiązywania problemów praktycznych, podczas gdy niektóre omawiane wcześniej techniki są w literaturze stosowane głównie w ramach badań naukowych, dotyczy to w szczególności analizy cytowań i formułowania w oparciu o dane bibliometryczne „praw”, dotyczących rozwoju aktywności naukowej.

Analitycy stosujący *tech mining* mogą w łatwy sposób zidentyfikować w zbiorze danych bibliometrycznych między innymi:

- publikacje, dotyczące poszczególnych materiałów i technik produkcyjnych,
- instytucje zaangażowane w badania w określonym obszarze,
- indywidualnych badaczy (w tym: potencjalnych pracowników, zajmujących się badaniami i rozwojem),
- potencjalnych partnerów - dostawców poszukiwanych technologii (prowadzących badania komplementarne), bezpośrednich konkurentów (prowadzących zbliżone

badania i niezależnie od badawczej rywalizacji, mogących stać się również uczestnikami konsorcjum projektowego), potencjalnych odbiorców (zajmujących się badaniami, które mogą wykorzystać nasze rezultaty),

- uczelnie, z którymi warto nawiązać współpracę badawczą lub wybrać jako miejsce odbywania stypendium.

Dla zidentyfikowanych badaczy, zespołów badawczych lub organizacji, możliwa jest budowa w oparciu o dane bibliometryczne profilu, ujawniającego m.in.:

- szczegółowe zainteresowania badawcze,
- koncentrację na jednym obszarze badań lub rozproszenie tematów badawczych,
- liczbę pracowników zaangażowanych w badania oraz wielkość poszczególnych zespołów badawczych,
- historyczne zmiany aktywności badawczej w obszarze podstawowym i pokrewnych.

Ważną cechą *tech mining* jest indukcyjny charakter analiz (ang. *inductive patterning*) (Porter, Newman 2011): zamiast przetwarzania danych w oparciu o wcześniej ustalone kategorie, możliwa jest eksploracja zbioru danych, identyfikacja "ukrytych" zmiennych i stopniowe określanie kolejnych zagadnień wraz z postępowaniem analiz.

Specjalistyczny program komputerowy, wspierający *tech mining*, pozwala na stworzenie własnej wielowymiarowej bazy danych, odzwierciedlającej relacje między zmiennymi. Ta baza będzie dostępna *offline* a nie *online*, czyli nie będzie wymagała ponownego korzystania z oryginalnej bazy bibliograficznej, umożliwiając dalsze, lokalne przetwarzanie danych. Ważnym elementem jest czyszczenie danych – w tym wprowadzanie poprawek oczywistych błędów redakcyjnych, eliminacja niepotrzebnych rekordów i duplikatów i korekta nazw instytucji-afiliacji. Szczególnie to ostatnie zadanie stwarza poważne wyzwanie dla analityków, zajmujących się bibliometrią. Ta sama uczelnia może być opisana przy wykorzystaniu nawet kilkudziesięciu wariantów jej nazwy (obejmujących m.in. skróty niektórych części nazwy, warianty nazewnictwa i literówki).

Dodatkowym problemem jest „poziom przypisania” afiliacji – czasami ten sam autor jest identyfikowany jako pracownik uczelni, w innych przypadkach – wydziału, instytutu, katedry lub centrum badawczego. Jeśli w danym projekcie badawczym ważne jest analizowanie aktywności poszczególnych organizacji (uczelni, instytutów badawczych i firm), niezbędne jest ujednoczenie listy afiliacji. Program *VantagePoint* pozwala zrobić to w oparciu o algorytm, korzystający z rozbudowanego tezaury nazw i skrótów nazw jednostek badawczo-rozwojowych. W wyniku jego zastosowania udaje się znacząco skrócić często bardzo długie listy afiliacji w zbiorze danych. W analogiczny sposób można także eliminować synonimy w określeniach naukowych i technologicznych oraz ujednoczyć nazewnictwo – przykładowo, w bazie mogły występować dwa warianty tego samego słowa kluczowego, jeden w liczbie pojedynczej drugi mnogiej. W przypadku operacji czyszczenia danych, podjęte działania powinny być starannie dokumentowane, by było możliwe powtórzenie tych samych procedur w oparciu o nowe dane.

Informatycy prowadzą badania nad algorytmami i metodami heurystycznymi, pozwalającymi na czyszczenie danych bibliometrycznych (D'Angelo i in. 2011), jednak najwygodniejszym rozwiązaniem wydaje się wykorzystanie gotowych pakietów oprogramowania. Frey, Minsk i Porter (2005) opisują procedury czyszczenia danych bibliometrycznych w programie *VantagePoint*, oparte na wykorzystaniu wyspecjalizowanych tezaurusów oraz metod heurystycznych. Rysunek 29, Rysunek 30 i Rysunek 31 pokazują przykładowe rezultaty operacji czyszczenia listy terminów specjalistycznych, występujących w streszczeniach artykułów naukowych. Przez eliminację różnorodnych form słownikowych i grupowanie synonimów, udało się zbudować bardziej przydatne w dalszych analizach zestawienia.

# Records	# Instances	Abstract (NLP) (Phrases) (Cleaned)
10	16	haptic interface
9	10	user
7	9	system
6	6	development
6	7	haptic device
6	10	operator
5	6	device
5	5	effectiveness
5	6	force
5	7	method
5	6	motion
5	5	one
4	4	applications
4	4	combination
4	4	dynamics
4	4	experimental results
4	5	friction
4	4	haptic rendering
4	4	presented
4	4	virtual environment
4	5	virtual objects
		(... 885 items in all)

Rysunek 29: Przykładowa operacja oczyszczania danych bibliometrycznych – krok 1. Źródło: Frey, Minsk, Porter (2005: 18).

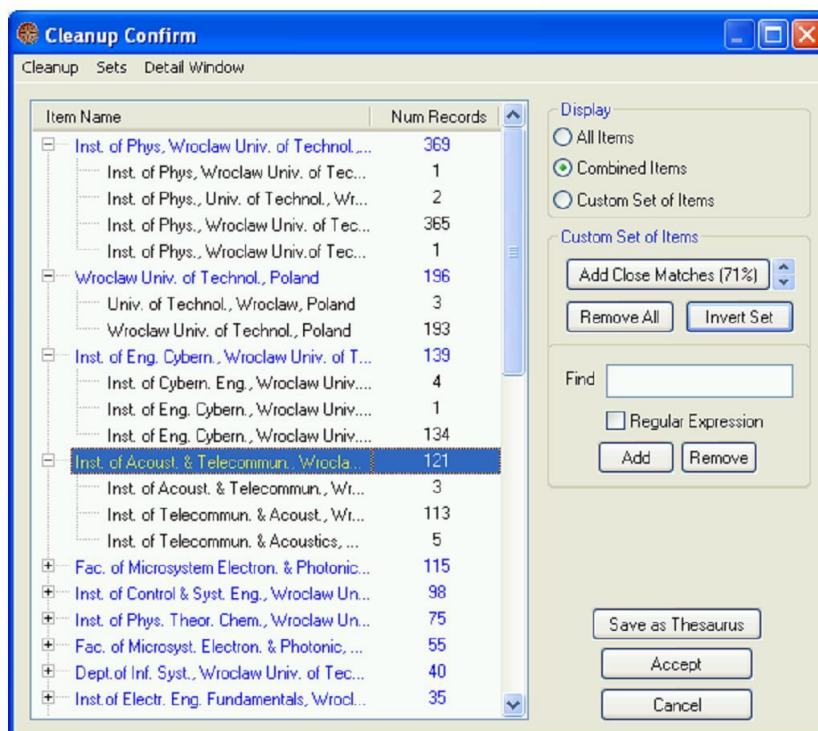
# Records	# Instances	Abstract (NLP) (Phrases) (2) (Cleaned)
33	95	Haptic
23	38	Device
22	58	Force
22	44	System
17	31	Interface
16	29	Model
15	38	Objects
14	22	User
14	27	Virtual
12	14	Two
11	25	Interaction
10	23	Based
10	10	Development
10	20	Display
10	12	Environment
10	20	Method
10	16	Motion
10	11	New
10	18	Operator
		(... 752 items in all)

Rysunek 30: Przykładowa operacja oczyszczania danych bibliometrycznych – krok 2. Źródło: Frey, Minsk, Porter (2005: 18).

# Records	# Instances	Abstract (NLP) (Phrases) (2) (Cleaned) (Group Names)
34	99	haptic tactile HapticFlow
26	72	force effectiveness effector
23	38	Device
23	59	force pull
23	59	force push
23	61	force strength
23	50	system scheme
22	59	force power
21	45	model simulation
18	32	interface port
17	31	model framework
16	31	model pattern
16	31	model posture
15	17	new novel
15	38	Objects
14	18	one 1 10 100 1D single
14	24	results effectiveness effector
14	20	two 2 25 2D
14	22	User
14	28	virtual practical
13	28	contact touch
13	21	motion movement
13	16	results resolution solution
11	13	environment surroundings
11	20	information data
		(... 571 items in all)

Rysunek 31: Przykładowa operacja oczyszczania danych bibliometrycznych – krok 3. Źródło: Frey, Minsk, Porter (2005: 20).

Rysunek 32 prezentuje przykładowe oczyszczanie zbioru danych bibliometrycznych przy wykorzystaniu programu *VantagePoint* w odniesieniu do różnych wariantów nazw Politechniki Wrocławskiej w publikacjach naukowych.



Rysunek 32: Czyszczenie listy afiliacji autorów publikacji naukowych w programie *VantagePoint*. Źródło: Myszk (2010: 18).

Zbiór danych bibliometrycznych może być następnie podstawą do opracowania zestawień: list, rankingów, macierzy lub map, wizualizujących zależności między obiektami. Szczególnie częste analizy, prowadzone przy wykorzystaniu narzędzi do eksploracji danych o technologiach, polegają na generowaniu macierzy, zestawiających wartości różnych zmiennych, obserwowanych w zbiorze danych. Niektóre warianty takich zestawień były omówione już wcześniej, np. analiza współautorstwa tekstów naukowych, czyli zestawienie zmiennej dotyczącej autorów tekstów z samą sobą. Analogicznie, analiza *co-classification* pozwala na stworzenie macierzy, w której zarówno wiersze, jak i kolumny będą odpowiadały kodom klasyfikacyjnym lub specjalistycznym terminom, a poszczególne komórki macierzy zawierać będą liczby publikacji, spełniających oba kryteria.

Możliwe jest jednak krzyżowanie dowolnych zmiennych, pobranych z bazy – często zestawiane są m.in.:

- serie danych dotyczących roku oraz innych zmiennych – jako sposób odzwierciedlenia tendencji czasowych i zmian na przestrzeni lat,
- serie danych dotyczących autorów, afiliacji lub krajów z danymi dotyczącymi słów kluczowych lub kodów klasyfikacyjnych – w celu porównania zainteresowań badawczych i specjalizacji,
- serie danych odpowiadające kodom klasyfikacyjnym z danymi dotyczącymi słów kluczowych – dla identyfikacji specyficznych orientacji badawczych oraz odpowiadających im kodów i haseł,
- nazwy czasopism lub konferencji z kodami klasyfikacyjnymi lub słowami kluczowymi – w celu określenia, które formy upubliczniania wiedzy w danym obszarze mogą okazać się najbardziej efektywne, bo dotrą do odbiorców najlepiej znających obszar badawczy.

Możliwe scenariusze zestawień krzyżowych prezentuje Tabela 12.

	Autor	Afiliacja	Kraj	Słowo kluczowe, kod	Słowo w treści	Pismo, konferencja	Rok	Cytowane dokumenty	Cytujące dokumenty
Autor	Współpraca w zespołach badawczych	Miejsca pracy autorów	Kraje autorów	Specjalizacja autorów	Specjalizacja autorów	Sposoby upowszechniania wybierane przez autorów	Okres aktywności autora / aktualność dorobku	Źródła wiedzy dla autorów	Uznanie, znaczenie dorobku
Afiliacja	Miejsca pracy autorów	Współpraca instytucjonalna	Kraje instytucji	Specjalizacja instytucji	Specjalizacja instytucji	Sposoby upowszechniania wybierane przez instytucje	Okres zaangażowania instytucji / aktualność dorobku	Źródła wiedzy dla instytucji	Uznanie, znaczenie dorobku
Kraj	Kraje autorów	Kraje instytucji	Współpraca międzynarodowa	Specjalizacja kraju	Specjalizacja kraju	Sposoby upowszechniania wybierane przez kraje	Okres zaangażowania kraju / aktualność dorobku	Źródła wiedzy dla krajów	Uznanie, znaczenie dorobku
Słowo kluczowe, kod	Specjalizacja autorów	Specjalizacja instytucji	Specjalizacja kraju	Współwystępowanie obszarów badawczych	Współwystępowanie zagadnień badawczych / weryfikacja poprawności klasyfikacji	Czasopisma i konferencje dotyczące obszarów badawczych	Okres występowania zagadnień badawczych / aktualność tematyki	Źródła wiedzy dla obszarów badawczych	Cytowania obszaru badawczego
Słowo w treści	Specjalizacja autorów	Specjalizacja instytucji	Specjalizacja kraju	Współwystępowanie zagadnień badawczych / weryfikacja poprawności klasyfikacji	Współwystępowanie zagadnień badawczych	Czasopisma i konferencje dotyczące zagadnień badawczych	Okres występowania zagadnień badawczych / aktualność tematyki	Źródła wiedzy dla zagadnień badawczych	Cytowania zagadnienia badawczego

Pismo, konferencja	Sposoby upowszechniania wybierane przez autorów	Sposoby upowszechniania wybierane przez instytucje	Sposoby upowszechniania wybierane przez kraje	Czasopisma i konferencje dotyczące obszarów badawczych	Czasopisma i konferencje dotyczące zagadnień badawczych	-	Okres publikacji czasopism i organizacji konferencji / zmiany znaczenia czasopism i konferencji	Źródła wiedzy typowe dla czasopism i konferencji	Wskaźnik wpływu (<i>impact factor</i>)
Rok	Okres aktywności autora / aktualność dorobku	Okres zaangażowania instytucji / aktualność dorobku	Okres zaangażowania kraju / aktualność dorobku	Okres występowania zagadnień badawczych / aktualność tematyki	Okres występowania zagadnień badawczych / aktualność tematyki	Okres publikacji czasopism i organizacji konferencji / zmiany znaczenia czasopism i konferencji	-	Okres powstania źródeł wiedzy	Okres powstania cytowań
Cytowane dokumenty	Źródła wiedzy dla autorów	Źródła wiedzy dla instytucji	Źródła wiedzy dla krajów	Źródła wiedzy dla obszarów badawczych	Źródła wiedzy dla zagadnień badawczych	Źródła wiedzy typowe dla czasopism i konferencji	Okres powstania źródeł wiedzy	Skupienia współcytowań / identyfikacja pokrewnych badań	Współcytowania / identyfikacja pokrewnych badań
Cytujące dokumenty	Uznanie, znaczenie dorobku	Uznanie, znaczenie dorobku	Uznanie, znaczenie dorobku	Cytowania obszaru badawczego	Cytowania zagadnienia badawczego	Wskaźnik wpływu (<i>impact factor</i>)	Okres powstania cytowań	Współcytowania / identyfikacja pokrewnych badań	Powiązania bibliograficzne / identyfikacja pokrewnych badań

Tabela 12: Możliwe scenariusze analiz krzyżowych przy wykorzystaniu technik *tech mining*. Źródło: opracowanie własne.

Stworzony zbiór danych może być wykorzystywany do zadawania specyficznych pytań, które są istotne w danym projekcie analitycznym. Ważne może być zdobycie informacji o tym, w którym roku pojawiły się publikacje dotyczące danego podobszaru, spełniające określone, dodatkowe kryterium. Przykładowo, pierwszy rok wystąpienia publikacji z przypisanym określonym kodem klasyfikacyjnym, napisanej przez autorów z danego kraju lub instytucji, będzie sygnałem wejścia na rynek badawczy analizowanego „gracza”. Innym przykładem jest weryfikowanie stopnia dojrzałości technologii lub dyscypliny badawczej – o dojrzałości świadczyć mogą: brak nowych, publicznie dostępnych publikacji lub wystąpień konferencyjnych (m.in. ze względu na utajnianie dalszych wyników badań przez firmy) lub wzrastająca liczba słów kluczowych, opisujących badania w danym obszarze, odpowiadająca postępującej specjalizacji badań (Watts, Porter 1997: 41).

Przy wykorzystaniu odpowiedniego oprogramowania, możliwe jest też grupowanie obiektów – łączenie w jedną grupę kilku autorów (np. członków tego samego zespołu badawczego), instytucji (np. wydziału, instytutu i centra badawcze jednej uczelni; uczelnie w jednym mieście), podobszarów technologicznych (np. połączenie dwóch prawie identycznych słów kluczowych, na różne sposoby opisujących tę samą technologię lub zbliżone obszary badań). Analityk musi podejmować indywidualne decyzje o zaliczeniu lub nie danego obiektu z listy do wybranej grupy. Taka ingerencja w listę obiektów poddanych dalszej analizie wymaga oczywiście świadomych decyzji, a w przypadku operacji na słowach kluczowych lub kodach klasyfikacyjnych – zrozumienia danej dyscypliny badawczej. Jej wykorzystanie pozwala jednak uprościć analizy i lepiej odzwierciedlić tendencje, interesujące badacza.

Oprócz „ręcznego” przydziału obiektów do grup, możliwe jest też odwołanie się do technik statystycznych, w tym analizy korelacji i analizy czynnikowej.

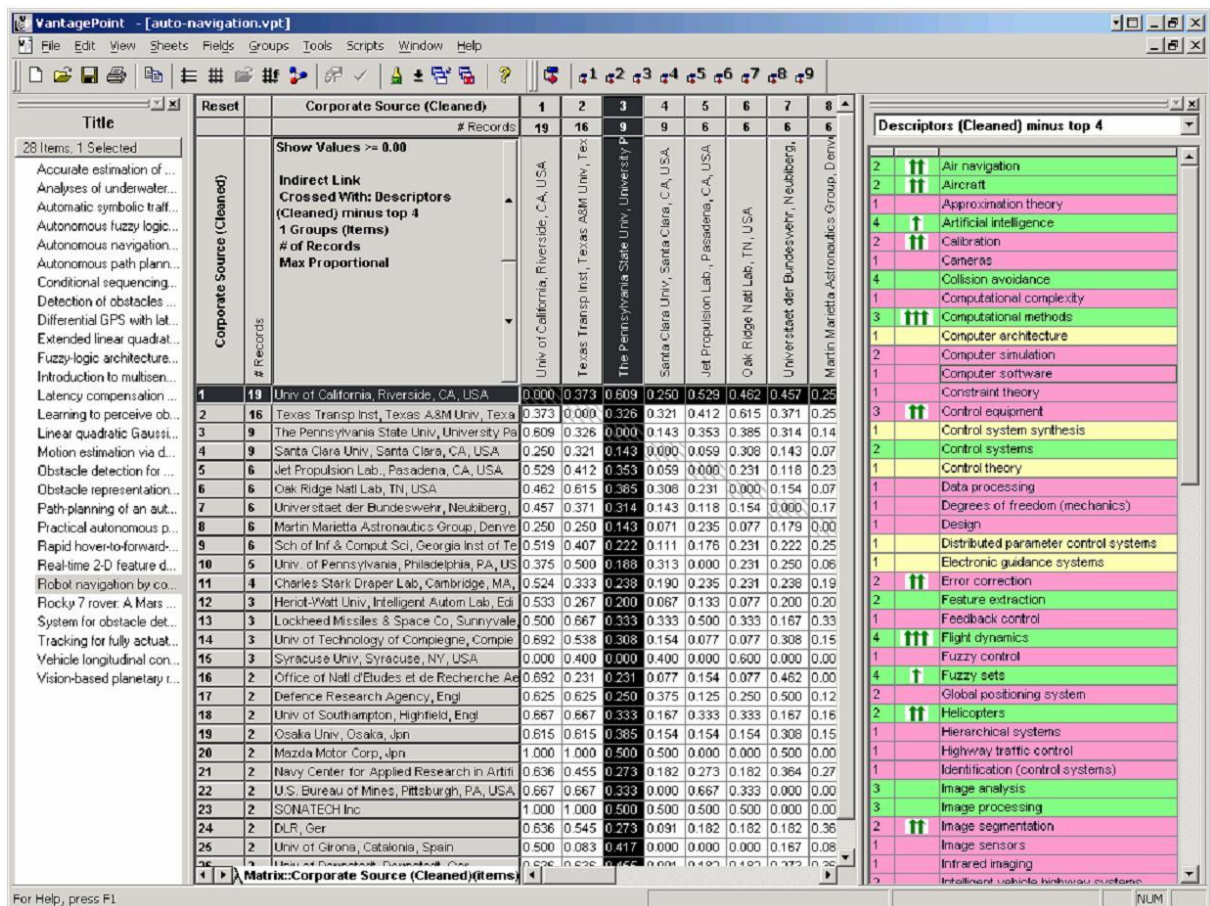
Autokorelacja jest miarą relacji pomiędzy obiektami z tej samej listy (np. autorami, afiliacjami, słowami kluczowymi). Posiada istotne przewagi w stosunku do analiz współwystępowania (takich jak w przypadku *co-authorship*, *co-word*, *co-classification*), gdyż odzwierciedla częstotliwość współwystępowania, a nie tylko bezwzględne liczby przypadków wzajemnego wystąpienia. Obliczenie współczynników autokorelacji pozwala na matematyczne i graficzne określenie częstotliwości współpracy lub współwystępowania obiektów. Porównanie normalizuje łączne liczby publikacji - np. zespół badawczy, którego autorzy są mało produktywni, ale prawie zawsze piszą razem, uzyska zbliżony wynik do bardzo produktywnego zespołu, którego członkowie również regularnie ze sobą współpracują. W przypadku klasycznej analizy współwystępowania, mniejszy zespół nie zostałby zapewne dostrzeżony.

Reset	Mesh Terms	10	11	12	13	14	15	16	17	18	19	20	21	22	23
	# Records	65	54	46	46	43	43	39	38	37	36	35	33	33	33
	Show Values >= 0.00														
	Auto-Correlation # of Records Cosine	Adult	Rats, Sprague-Dawley	Insulin/blood	Dose-Response Relationship, Drug	Blood Glucose/metabolism	Insulin Resistance	Aged	Mice, Inbred C57BL	Blood Glucose/analysis	Cells, Cultured	Time Factors	Research Support, General	Molecular Sequence	Glucose/metabolism
1	46	Insulin/blood	0.219	0.140	0.196	0.270	0.045	0.142	0.191	0.242	0.098	0.125	0.077	0.000	0.100
2	211	Male	0.470	0.393	0.305	0.273	0.189	0.353	0.313	0.306	0.149	0.279	0.168	0.072	0.200
3	7	Leptin/blood	0.000	0.051	0.279	0.000	0.115	0.000	0.184	0.062	0.063	0.000	0.132	0.000	0.000
4	43	Blood Glucose/metabolism	0.076	0.104	0.270	0.112	0.093	0.073	0.223	0.000	0.076	0.103	0.027	0.000	0.100
5	288	Animals	0.022	0.433	0.261	0.313	0.279	0.225	0.019	0.363	0.165	0.314	0.259	0.185	0.236
6	9	Organ Size/drug effects	0.000	0.091	0.246	0.098	0.000	0.000	0.054	0.110	0.000	0.056	0.000	0.000	0.000
7	133	Female	0.549	0.130	0.243	0.153	0.159	0.159	0.472	0.127	0.200	0.087	0.117	0.151	0.030
8	37	Blood Glucose/analysis	0.245	0.112	0.242	0.024	0.000	0.075	0.158	0.160	0.000	0.111	0.057	0.000	0.000
9	7	Dietary Supplements	0.047	0.000	0.223	0.056	0.058	0.058	0.061	0.000	0.000	0.000	0.000	0.066	0.000
10	4	Glycogen/metabolism	0.000	0.068	0.221	0.000	0.076	0.000	0.162	0.082	0.000	0.000	0.000	0.000	0.000
11	4	Liver Glycogen/metabolism	0.000	0.068	0.221	0.000	0.229	0.000	0.162	0.082	0.000	0.000	0.000	0.000	0.000
12	65	Adult	0.000	0.000	0.219	0.073	0.076	0.076	0.536	0.020	0.245	0.021	0.105	0.022	0.043
13	2	Muscle, Skeletal/enzymology/metabo	0.000	0.096	0.209	0.000	0.108	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.200
14	2	Adipose Tissue/drug effects/growth	0.000	0.000	0.209	0.104	0.216	0.108	0.000	0.229	0.000	0.118	0.120	0.000	0.000
15	2	Dehydroepiandrosterone Sulfate/blo	0.088	0.000	0.209	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
16	2	Hypoglycemic Agents/chemical synt	0.000	0.096	0.209	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
17	2	Anticholesteremic Agents/administra	0.000	0.000	0.209	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
18	2	Drinking/drug effects	0.000	0.096	0.209	0.104	0.216	0.000	0.000	0.000	0.000	0.120	0.000	0.000	0.100
19	5	Eating	0.111	0.061	0.198	0.000	0.000	0.000	0.073	0.074	0.000	0.000	0.000	0.000	0.000
20	127	Rats	0.011	0.652	0.196	0.340	0.189	0.162	0.000	0.158	0.117	0.281	0.210	0.108	0.124
21	46	Dose-Response Relationship, Drug	0.073	0.281	0.196	0.000	0.112	0.022	0.094	0.072	0.024	0.197	0.249	0.077	0.026
22	286	Humans	0.477	0.072	0.192	0.148	0.180	0.262	0.369	0.115	0.204	0.177	0.120	0.268	0.247
23	38	Mice, Inbred C57BL	0.020	0.066	0.191	0.072	0.223	0.148	0.026	0.000	0.160	0.108	0.110	0.028	0.085
24	102	Mice	0.025	0.081	0.190	0.190	0.211	0.181	0.032	0.610	0.130	0.264	0.167	0.155	0.207
25	76	Comparative Study	0.356	0.125	0.186	0.220	0.175	0.157	0.220	0.112	0.151	0.057	0.252	0.120	0.040

Rysunek 33: Przykład macierzy autokorelacji dla podobszarów badań medycznych w programie VantagePoint. Źródło: materiały marketingowe firmy Search Technology, Inc.

Analiza korelacji krzyżowych (ang. *cross-correlations*) dotyczy przypadków, w których jedna zmienna porównywana jest w oparciu o wartości drugiej zmiennej. Przykładowo, badacze mogą być porównywani pod względem kodów klasyfikacyjnych, odpowiadających tematowi badań, które prowadzą, a wysoki poziom współczynnika korelacji krzyżowej pomiędzy dwoma badaczami odpowiadać będzie podobieństwu ich zainteresowań badawczych. W analogiczny sposób porównać można instytucje badawcze, kraje oraz warianty technologii (jako zbiory wybranych słów kluczowych lub kodów klasyfikacyjnych).

Rysunek 34 prezentuje zestawienie danych w programie VantagePoint - macierz, opartą na obliczeniach współczynników korelacji oraz dodatkowe informacje na temat poszczególnych obiektów bibliometrycznych.



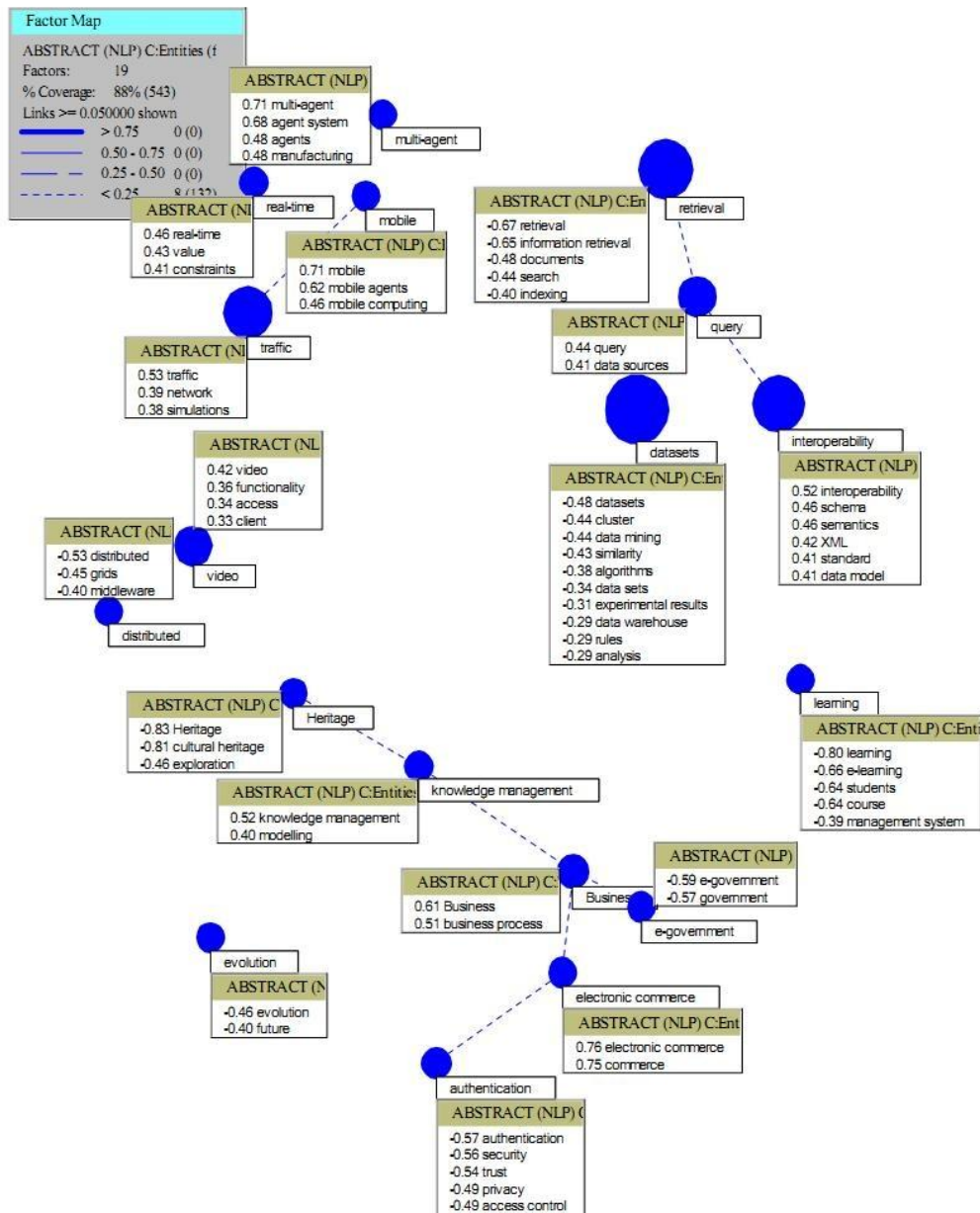
Rysunek 34: Przykład zaawansowanych analiz w programie *VantagePoint*. Źródło: Frey, Minsk, Porter (2005: 36).

Techniki *tech mining* pozwalają również na identyfikację tzw. *invisible colleges* – zbiorów badaczy, których charakteryzują zbliżone zainteresowania badawcze, podejścia badawcze, stosowane techniki i rozwiązywane problemy. Tytułowa "niewidzialność" odnosi się do braku bezpośrednich związków między badaczami lub zespołami, pracującymi w różnych instytucjach, a nawet krajach. Chociaż nie utrzymują oni formalnej współpracy badawczej, ani nie publikują wspólnych artykułów, można jednak dostrzec podobieństwo między ich dokonaniem. Do ich identyfikacji wykorzystane mogą być techniki statystyczne, oparte na różnych scenariuszach analizy czynnikowej – przykładowo, oprogramowanie *VantagePoint* korzysta z analizy głównych składowych *PCA* (ang. *Principal Component Analysis*). Statystyczna obróbka dużego zbioru danych pozwala na ujawnienie „ukrytych zmiennych” (ang. *latent variables*) i nieoczywistych powiązań. Przy interpretacji danych dąży się do nadania znaczenia każdej zidentyfikowanej w ten sposób zmiennej – może być to przykładowo pomiar „intensywności” danej orientacji badawczej – należy jednak pamiętać, że techniki analizy czynnikowej bywają czasami kwestionowane przez metodologów badań, a uzyskane w drodze zastosowania algorytmów obliczeniowych czynniki lub składowe niekoniecznie będą posiadały logiczne interpretacje. Zastosowanie takich analiz do zbioru danych o badaczach w wybranym obszarze nauki powinno ułatwić odnalezienie „ukrytych” podobieństw między naukowcami, a także identyfikację podmiotów, z którymi warto nawiązać bliską, bezpośrednią współpracę. W analogiczny sposób można poszukiwać

podobieństw i ukrytych zależności pomiędzy instytucjami lub krajami (np. poszukiwanie potencjalnych źródeł sukcesu badawczego, które mogą tkwić w połączeniu nieoczywistych kompetencji i obszarów badań), jak również profilować obszary badań i technologie.

Analiza *PCA* pozwala na redukcję liczby wymiarów w większych zbiorach danych. Przykładowo, zamiast kilkuset specyficznych kodów klasyfikacyjnych lub słów kluczowych, poszukiwanych jest kilka "ukrytych zmiennych" (ang. *latent variables*) – głównych składowych, które dostatecznie dobrze opisują zbiór danych, a jednocześnie zmniejszają liczbę danych do dalszego przetwarzania. Logiczna interpretacja uzyskanych składowych pozwala w wielu przypadkach uznać niektóre z nich za reprezentacje określonych podejść do badań lub rozwoju technologii. Stosowanie technik redukcji danych przynosi praktyczne korzyści dla analityków. Przykładowo, jeśli zbiór danych wykorzystuje 300 kodów klasyfikacyjnych i zawiera 5000 publikacji, oznacza to, że macierz zestawiająca kody i publikacje ma łącznie 1.500.000 pól (300 x 5.000). Praca analityczna z tak rozbudowaną macierzą nie jest łatwa, a każde jej przetworzenie w drodze obliczeń okazuje się niezwykle czasochłonne. Wskazana jest redukcja liczby kodów klasyfikacyjnych. Nie wystarczy jedynie ograniczyć się w dalszych analizach do najpopularniejszych, najczęściej występujących kodów - rezygnacja z pozostałych kodów może prowadzić do przeoczenia istotnych tendencji w badaniach. Techniki *PCA* pozwalają na redukcję liczby zmiennych przy jednoczesnym odzwierciedleniu różnic między oryginalnymi danymi. W wyniku obliczeń powstają nowe zmienne, nazywane składowymi, oparte na przekształconych matematycznie danych wejściowych. Każda składowa jest wynikiem agregacji i statystycznej obróbki licznych, szczegółowych zmiennych. Właściwe zastosowanie techniki *PCA* wymaga niestety w większości przypadków podejmowania wielokrotnych prób analitycznych w oparciu o różny dobór zmiennych - nie jest to działanie w pełni zalgorytmizowane, zakłada dobre zrozumienie technik statystycznych i umiejętność interpretacji otrzymanych wyników przez odwołanie do znajomości różnic między poszczególnymi zmiennymi, poddawanych agregacji (np. rozumienia różnic i podobieństw między specjalistycznymi terminami, które są słowami kluczowymi, opisującymi poszczególne publikacje naukowe). W dalszych analizach, możliwe jest zestawienie składowych z innymi zmiennymi, m.in. afiliacjami i autorami, co pozwala na ocenę orientacji badawczych poszczególnych instytucji i indywidualnych badaczy. Wykrywanie istotnych statystycznie podobieństw i różnic w tych orientacjach nie byłoby możliwe w oparciu o kilkaset serii danych, odpowiadających liczbie publikacji dotyczących szczegółowych zagadnień badawczych, natomiast jest łatwiejsze w przypadku redukcji liczby zmiennych dzięki analizie *PCA* i sprowadzeniu tych różnic lub podobieństw do kilku wymiarów (składowych w technice *PCA*). Techniką oferującą zbliżone do *PCA* korzyści jest skalowanie wielowymiarowe *MDS* (ang. *multidimensional scaling*), również pozwalające na redukcję wymiarów w zbiorze zmiennych. *MDS* było często stosowane we wcześniejszych publikacjach bibliometrycznych ze względu na prostotę obliczeniową i możliwość wykorzystania do analiz standardowych statystycznych aplikacji komputerowych, jednak współcześnie jest zastępowane przez techniki bardziej złożone, odwołujące się do analizy czynnikowej.

Wyniki analiz są zwykle prezentowane graficznie jako mapa powiązanych obiektów. Rysunek 35 prezentuje przykładową mapę czynników, stworzoną dzięki wykorzystaniu techniki PCA w odniesieniu do terminów specjalistycznych, występujących w streszczeniach publikacji naukowych. W analogiczny sposób zestawiać można inne zmienne, w tym badaczy, instytucje, słowa kluczowe i kody klasyfikacyjne.



Rysunek 35: Graficzna forma prezentacji wyników analizy PCA w programie *VantagePoint*. Źródło: Frey, Minsk, Porter (2005: 51).

Możliwe jest wreszcie wykorzystanie analizy skupień – grupowanie obiektów w zbiorze danych w oparciu o wyniki analiz statystycznych, czyli podział zbioru na rozłączne skupienia (klastry) badaczy, instytucji, krajów lub obszarów badawczych. Algorytmy analizy skupień poszukują podobnych obiektów w oparciu o bogaty zestaw atrybutów, opisujących dany obiekt.

Techniki służące redukcji danych – takie jak *PCA* czy *MDS* – nie pozwalają na przejrzyste wyodrębnienie istotnie różnych statystycznie grup obiektów, koncentrując się raczej na wskazywaniu znaczących różnic między tymi obiektami i identyfikacji „ukrytych zmiennych”. Analiza skupień pozwala z kolei na podział obiektów na odrębne grupy, ale nie odzwierciedla dodatkowych relacji między tymi obiektami (Kopcsa, Schiebel 1998: 7-8). Wyniki analizy skupień będą więc zapewne mniej przydatne dla analityka, zajmującego się obserwacją badań naukowych i rozwoju technologii – nie otrzyma on bowiem zbioru danych zredukowanego do mniejszej liczby zmiennych, jak w przypadku techniki *PCA*, a jedynie grupy podobnych obiektów. W publikacjach, dotyczących wykorzystania bibliometrii do obserwacji nauki i techniki, analiza skupień wykorzystywana jest rzadko. Algorytmy komputerowej analizy skupień są czasochłonne i wymagają starannego przemyślenia parametrów wejściowych. Należy zadbać również o możliwość interpretacji wyników (zrozumienia zmiennych, wykorzystywanych w obliczeniach).

Mimo dostępnych mocy obliczeniowych komputerów osobistych oraz możliwości oprogramowania komputerowego, w wielu przypadkach może okazać się, że wygenerowany podzbiór danych bibliometrycznych uniemożliwia dalsze przetwarzanie ze względu na zbyt duży rozmiar. Już zbiór 20.000 streszczeń publikacji lub dokumentów patentowych może przekraczać możliwości oprogramowania. W takich przypadkach warto rozważyć zastosowanie techniki próbkowania w celu analizy wybranego losowo podzbioru danych (Jahangirian i in. 2011).

Tech mining może być podstawą do dalszych analiz, prowadzonych przy wykorzystaniu dodatkowych narzędzi informatycznych. Programy takie jak *VantagePoint* pozwalają bowiem na oczyszczenie, analizę i przetworzenie danych, które następnie można eksportować do innego programu, w tym do specjalistycznych aplikacji, służących wizualizacji danych albo budowie modeli prognostycznych.

Do możliwych przykładów dalszej obróbki ilościowej należą:

- sortowanie list według złożonych kryteriów,
- obliczanie wskaźników, pozwalających na porównania poszczególnych podmiotów,
- wykorzystanie serii danych w modelach ekonometrycznych,
- analiza trendów,
- prognozowanie dalszego rozwoju liczby publikacji w krótkiej i średniej perspektywie czasowej (same dane ilościowe w technikach *tech mining* nie pozwalają jednak na przewidywanie kierunków rozwoju badań naukowych lub technologicznych),
- powiązanie pomiędzy seriami danych bibliometrycznych z różnych źródeł (np. seria liczb publikacji i patentów dla porównywanych organizacji),
- powiązanie z innymi seriami danych, nie mającymi charakteru bibliometrycznego - zwł. dane ekonomiczno-społeczne (np. skala wydatków na B+R, liczba badaczy zatrudnionych, wyniki finansowe),
- normalizacja danych, przekształcenia matematyczne, obliczanie wskaźników i indeksów (istotne jest zwłaszcza obliczenie wartości względnych, biorących pod uwagę różnice wynikające z odmiennej produktywności lub liczby badaczy),

- wielowymiarowa wizualizacja danych.

Warto być świadomym kilku praktycznych ograniczeń analiz w obszarze *tech mining*-u, wynikających ze specyfiki baz i dostępnych algorytmów ich przetwarzania. Pierwszym problemem jest właściwe odwzorowanie afiliacji badawczy. Dostawcy baz dopiero niedawno rozpoczęli przypisanie afiliacji do wszystkich autorów publikacji, przykładowo starsze dokumenty pobrane z baz *INSPEC* i *MEDLINE* zawierają tylko afiliację pierwszego autora, co uniemożliwia analizy współpracujących instytucji. Bazy *Web of Science* i *Scopus* nie mają tych problemów. Niektóre analizy mogą wymagać uzupełnienia danych pobranych z bazy o samodzielnie odnalezione lub zweryfikowane dane dotyczące miejsc pracy autorów publikacji, co może okazać się bardzo czasochłonnym zadaniem.

Drugim wyzwaniem analitycznym jest brak możliwości bezpośredniego powiązania danych, pochodzących z różnych baz danych. Każda baza posiada własną strukturę rekordu i odmienny sposób opisu publikacji lub patentu. Program *VantagePoint* umożliwia import dodatkowego zbioru danych i połączenie z dotychczas wykorzystywanym, co stwarza możliwości jednoczesnej analizy publikacji z dwóch różnych baz, jednak dane te nie będą posiadały jednolitej struktury, a wykrycie ewentualnych duplikatów może być bardzo trudne. Zastosowanie różnych baz danych wydaje się wskazane w niektórych przypadkach. Przykładowo, baza *INSPEC* posiada najlepsze możliwości wyszukiwania specjalistycznych podobszarów badań dzięki rozbudowanej taksonomii tematycznej - ale w odniesieniu do starszych danych, brakuje w niej afiliacji drugiego i kolejnych autorów, nie ma też danych dotyczących cytowań tekstów. Idealnym rozwiązaniem byłoby powiązanie danych z *INSPEC*-a z danymi z baz *Scopus* lub *Web of Science*. Z drugiej strony, znacząca część rekordów w bazie *INSPEC* to wystąpienia konferencyjne, których nie ma w dwóch wspomnianych bazach, a które doskonale nadają się do identyfikacji nowych tendencji w nauce i technice.

Trzecią przeszkodą na drodze do zdobycia wszechstronnej wiedzy na temat danego obszaru badań lub technologii jest trudność powiązania danych o liczbie publikacji z liczbą ich cytowań. Nie ma prostego sposobu na przetworzenie tych danych w jednym zbiorze ze względu na ich różny charakter, brakuje też łatwych w obsłudze mechanizmów prowadzenia rozbudowanych analiz, łączących dane obu rodzajów. Omawiany wcześniej przegląd analiz bibliometrycznych, prowadzonych przez inne instytucje, w tym firmy prywatne, sugeruje, że koncentrują się one na analizach liczby publikacji a nie cytowań – na takich zastosowaniach *tech mining*-u skupia się również funkcjonalność popularnych narzędzi informatycznych. Ewentualnym sposobem na rozwiązanie omawianego problemu jest identyfikacja w bazach cytowań kilku - kilkudziesięciu najczęściej cytowanych autorów tekstów z danego obszaru i dodanie informacji o nich do przetwarzanego zbioru danych (np. odznaczenie tych autorów jako należących do grupy autorów "najważniejszych" lub uzupełnienie danych o odpowiednie współczynniki oparte na cytowaniach). To sprawi, że w dalszych analizach nie zostaną pominięte osoby, które mogły być mało produktywne, jednak ważne – bez zastosowania proponowanego podejścia badacz, który napisał tylko jeden artykuł w danym obszarze badawczym, nie zostałby zidentyfikowany jako ważny aktor sceny innowacji w przypadku analiz opartych na liczbie publikacji nawet wtedy, gdyby jego pojedynczy tekst był najczęściej cytowany w danej dyscyplinie. Nie wydaje się natomiast niezbędne przetwarzanie

wszystkich informacji o cytowaniach analizowanych publikacji - ich gromadzenie jest na ogół niezwykle czasochłonne, wymaga licznych operacji wyszukiwania i dodawania nowych danych do opracowywanego zbioru publikacji, jak również odpowiedniego oprogramowania procesu analizy tych danych. W odniesieniu do prac analitycznych, mających na celu obserwację rozwoju badań naukowych i technologii, analizy cytowań stosowane są relatywnie rzadko, co wiąże się z omawianymi trudnościami realizacyjnymi.

9.3. Tomografia bazodanowa

Tomografia bazodanowa (ang. *database tomography*) jest podejściem analitycznym, opracowanym przez zespół badaczy, pracujących dla armii amerykańskiej, pod kierunkiem Ronalda N. Kostoffa. Szczegółowy opis technik stosowanych w tomografii bazodanowej, uwarunkowań metodologicznych oraz różnic w stosunku do wcześniejszych podejść analitycznych oferuje Kostoff (2003). Podejście opiera się na analizach częstotliwości oraz miejsc występowania wieloczłonowych terminów specjalistycznych w tekstach naukowych – przy wykorzystaniu algorytmów matematycznych, pozwala na identyfikację wiodących tematów badawczych i związanych z nimi tendencji. Podejście historycznie poprzedza rozwój technik *tech mining*-u - powstało na początku lat 90-tych XX wieku, w sytuacji ograniczonych możliwości obliczeniowych komputerów osobistych i niedoskonałości dostępnych wówczas baz.

Sekretarz Marynarki Wojennej Stanów Zjednoczonych jest posiadaczem patentu nr 5.440.481, udzielonego w 1995 roku przez *USPTO*, zgłoszonego wcześniej przez Kostoffa i opisującego sposób prowadzenia analiz w oparciu o tomografię bazodanową (ochrona patentowa obowiązuje tylko na terytorium Stanów Zjednoczonych). Warto wspomnieć, że twórcy tomografii bazodanowej zlecieli później rozwój dedykowanego oprogramowania *TechOASIS*, którego komercyjna wersja dostępna jest jako opisany w podrozdziale o *tech mining*-u program *VantagePoint*. Oferowane przez ten program funkcje mogą być wykorzystane do prowadzenia analiz w oparciu o *database tomography*.

Pojęcie tomografii pochodzi z obszaru diagnostyki medycznej i oznacza uzyskiwanie przekrojowego obrazu organu lub ciała. Analogicznie, tomografia bazodanowa pozwala na przekrojowe spojrzenie na dorobek badawczy danej dyscypliny lub obszaru nauki. Powstała jako próba udoskonalenia technik analizy *co-words*, opiera się na zbliżonych założeniach.

Przetwarzając zbiór tekstów (artykułów naukowych, opisów wynalazków, wystąpień konferencyjnych, wniosków o dofinansowanie badań), analityk dąży do identyfikacji często powtarzających się słów i ich kombinacji (przypadków pojęć dwuwyrazowych, trzywyrazowych itd.). Przy identyfikacji kombinacji słów istotne znaczenie odgrywa ich wzajemna bliskość położenia w analizowanym tekście.

Po otrzymaniu komputerowego zestawienia tych kombinacji, posortowanego według częstości ich wystąpień, analityk arbitralnie wybiera te, które uważa za istotne dla analizowanego obszaru badań. Ryzyko błędu decyzyjnego obniża włączenie w ten wybór zespołu ekspertów dziedzinowych. Wybrane kombinacje słów odpowiadają *PTA* - dominującym obszarom tematycznym (ang. *pervasive theme areas*). Dla poszczególnych

PTA, prowadzona jest następnie analiza częstotliwości występowania publikacji lub patentów i obliczane odpowiednie wskaźniki. Dokument patentowy, opisujący algorytm *database tomography*, liczy 118 stron, jest dostępny w bazie *USPTO* i zawiera m.in. rozbudowany kod źródłowy algorytmu (Kostoff i in. 1995). Praktyczne wykorzystanie algorytmu tomografii bazodanowej jest współcześnie możliwe w ramach analiz *tech mining*, jednak współcześnie badacze korzystają zwykle z bardziej wyrafinowanych metod analitycznych, odwołując się do kombinacji słów kluczowych, kodów klasyfikacyjnych oraz technik statystycznych.

Kostoff i in. (2001) opisują pierwotne zastosowania tomografii bazodanowej, która była wykorzystywana w latach 90-tych przez amerykańską marynarkę do systematycznej analizy dorobku naukowego badaczy byłego Związku Radzieckiego, w szczególności w obszarach mogących znaleźć zastosowania militarne. Argumentując na rzecz wykorzystania komputerowych technik przetwarzania literatury naukowej Kostoff wskazuje wymierne korzyści na przykładzie badaczy, zajmujących się konstrukcjami lotniczymi. Ponad 700 czasopism naukowych zajmuje się samolotami, trzy spośród tych czasopism należą do grupy najczęściej czytanej i skoncentrowanej głównie na lotnictwie, odpowiadając za około 18-20% artykułów z omawianego obszaru badawczego. Badacze i osoby zarządzające projektami badawczymi powinny jednak wiedzieć, jakie nowe trendy pojawiają się w pokrewnych obszarach badawczych. W rzeczywistości ich czas pracy pozwala na lekturę co najwyżej 10% publikacji podstawowych i niewielu pozycji z obszarów pokrewnych (Kostoff i in. 2001: 240-241). Technika tomografii bazodanowej miała pomóc w przezwycięzeniu tych ograniczeń, a liczne publikacje zespołu Kostoffa wskazują na to, że rzeczywiście była ona aktywnie wykorzystywana przez sektor obronny Stanów Zjednoczonych. Wraz z rozwojem doskonalszych technik *tech miningu* i dostępnością dedykowanych narzędzi informatycznych, klasyczna forma tomografii bazodanowej uległa stopniowemu zapomnieniu. Zespół Kostoffa korzystał w nowym tysiącleciu z innych technik analitycznych, choć w niektórych publikacjach nadal wykorzystywał pojęcie "*database tomography*" - tym razem nie do opisu analiz, opartych o opatentowany algorytm, a do opisu projektów zgodnych z szerszym podejściem *tech mining-u*.

9.4. Literature-Based Discovery

Literature-Based Discovery jest nazwą podejścia, w którym stymulowany jest proces odkryć naukowych w oparciu o analizę literatury naukowej. Zbliżone pojęcia to *LAD* (*Literature-Assisted Discovery*) i *LRD* (*Literature-Related Discovery*) (Kostoff 2008a: 165) oraz *CBD* (*Complementary but Disjoint literatures*) (Stegmann, Grohmann 2003: 111). *LBD* łączy dwa dotychczas niepowiązane pojęcia lub obszary badawcze, tworząc w ten sposób nową, przydatną wiedzę (Kostoff 2008a: 165), a w praktyce: zwracając uwagę naukowców na potencjalnie interesujące zależności i zjawiska, które dotąd nie były przedmiotem ich zainteresowań. Zgodnie z humorystycznym powiedzeniem, "miesiąc pracy w laboratorium może zaoszczędzić godzinny pobyt w bibliotece" (Banville 2009: 3).

Cel *LBD* różni się istotnie od dotychczas prezentowanych zastosowań technik bibliometrycznych. Prowadzi bowiem do tworzenia nowej wiedzy naukowej i technicznej, a nie tylko zdobywania informacji o tym, kto i w jaki sposób prowadzi badania naukowe.

Warto zaznaczyć, że wykorzystanie *LBD* wykracza poza zakres prac analitycznych, opisywanych w niniejszej książce, a metody analityczne *LBD* mogą być umiejętnie wykorzystywane tylko przez osoby bardzo dobrze znające określony podobszar badawczy. *LBD* demonstruje jednocześnie potencjał zastosowania technik bibliometrycznych w celu wspierania kreatywności i generowania odkryć naukowych.

Istotą *LBD* jest odwołanie się do dorobku dyscypliny istotnie różnej od obszary badawczego, który jest przedmiotem zainteresowań naukowca, w celu rozwiązania konkretnych problemów badawczych. *LBD* wykorzystuje dwa zbiory publikacji, dotychczas pozostające bez związków w literaturze czyli przypadki zbiorów tekstów, które nie są cytowane pomiędzy zbiorami, a ich autorzy nie tworzą jednocześnie publikacji należących do obu obszarów. Czytelnik tekstów naukowych prawdopodobnie nigdy nie zapoznałby się z zawartością jednego z tych zbiorów, koncentrując swoją uwagę na drugim. Pierwszy zbiór można oznaczyć jako "AB", a drugi jako "BC". Dla zbioru "AB", dominującym wątkiem tematycznym jest "A", podczas gdy "B" pozostaje jedynie wątkiem pobocznym. W przypadku "BC", dominują badania nad "B", a "C" jest drugoplanowym zagadnieniem. Jak widać, zagadnienie "B" występuje w obu analizowanych zbiorach. Nie istnieją jednak żadne publikacje, które łączyłyby jednocześnie tematy "A", "B" i "C" - zbiory "AB" i "BC" są rozłączne. Badacz zajmując się *LBD* zainteresuje się publikacjami dotyczącymi tematu "B" jako potencjalnego łącznika między dwoma, dotychczas odrębnymi obszarami badań. Może okazać się, że sposobem na rozwiązanie problemów obszaru "A" będzie wykorzystanie wyników badań z obszaru "C", choć do tej pory nikt nie zastanawiał się nad ich ewentualnymi związkami.

Swanson sugerował w oparciu o zastosowanie *LBD*, że *EPA* czyli kwas eikozapentaenowy ("C" zgodnie z powyższym opisem) może zmniejszyć lepkość krwi ("B") , a przez to złagodzić symptomy zespołu Raynauda ("X") (Kostoff 2008a: 168). Wcześniejsze badania nad związkami między *EPA* a lepkością krwi to obszar "BC", a dorobek badań nad lepkością krwi u osób cierpiących na zespół Raynauda to obszar "AB". Nigdy wcześniej nie powiązано wyników odrębnych obszarów badawczych, a wykorzystanie *LBD* pozwoliło wykorzystać omawiany kwas, zawarty m.in. w oleju ryb morskich, do celów łagodzenia omawianej dolegliwości.

LBD jest najczęściej wykorzystywana w naukach medycznych, a na przestrzeni lat powstało wiele badań, odwołujących się do bazy *MEDLINE* w celu generowania odkryć literaturowych. Twórca *LBD*, Donald R. Swanson, prowadził m.in. badania nad związkami pomiędzy migreną a magnezem, arginina a sometomedyną oraz potencjalnym wykorzystaniem substancji biologicznych do celów militarnych. Przegląd dorobku *LBD* można znaleźć w pozycjach: (Swanson, Smalheiser 1996; Ganiz, Pottenger, Janneck 2005: 40-41; Kostoff 2008a: 171-172). *LBD* przydaje się w szczególności do identyfikacji nowych zastosowań dla istniejących środków farmakologicznych (ang. *drug repurposing*) (Banville 2009: 4).

Realizacja analiz wymaga zgromadzenia podzbiorów artykułów na tematy "AB" i "BC" - w rzeczywistości, pierwszym krokiem *Literature-Based Discovery* jest identyfikacja obszaru

tematycznego "B". Badacze nie wiedzą początkowo, jak ukierunkować swoje poszukiwania - gdyby od początku wiedzieli, że mają interesować się przykładowo czynnikami wpływającymi na lepkość krwi, ich dalsza praca nie byłaby w rzeczywistości odkryciem naukowym. W praktyce, w celu identyfikacji właściwego kierunku dalszych analiz, należy rozpocząć od zgromadzenia literatury dotyczącej podstawowego obszaru badań "A" - czyli w omawianym powyżej przykładzie, zespołu Raynauda. Następnie analizowane są pokrewne obszary badawcze, powiązane z tematem "A". Najprościej można to zrobić przy wykorzystaniu narzędzi, wspierających analizy współklasyfikacji (ang. *co-classification*), takie jak omawiany wcześniej program *VantagePoint*. Przy jego wykorzystaniu wygenerowane zostaną zestawienia różnych kombinacji tematu "A" z innymi tematami - np. "AX", "AY", "AZ", ..., wśród nich także "AB". Dla każdego zidentyfikowanego związku należy poszukać dalszych kombinacji - tym razem w innym podzbiórce tekstów naukowych. W kolejnym kroku niezbędne jest więc niezależne wyszukiwanie w bazie bibliometrycznej takiej jak *MEDLINE*, oparte o temat "X" oraz analiza wszystkich kombinacji współklasyfikacji (np. "XY", "XZ", "XV, ..., "XA"). To samo należy powtórzyć w odniesieniu do następnych tematów. W ten sposób można wygenerować setki "kandydatów do odkryć" - potencjalnych sposobów rozwiązania problemu "A" - a dla każdego z nich należy dodatkowo zapoznać się z zawartością literatury przedmiotu i ocenić możliwość wykorzystania w odniesieniu do stawianego problemu badawczego.

Proces *LBD* może być zautomatyzowany przy wykorzystaniu oprogramowania *Arrowsmith*, stworzonego przez Swansona i Smalheisera (Smalheiser 2005: 1-2), od lat aktywnie rozwijających *LBD*. Program *Arrowsmith* dostępny jest nieodpłatnie jako serwis internetowy (http://arrowsmith.psych.uic.edu/arrowsmith_uic), pozwalający na przeprowadzenie wyszukiwań w bazie *MEDLINE* dla obszarów „A” i „C”, a następnie wygenerowanie listy zagadnień badawczych, które mogą stanowić potencjalne odkrycie literaturowe („B”) oraz przeprowadzenie dalszych analiz. Rysunek 36, Rysunek 37 i Rysunek 38 prezentują przykładowe analizy w programie *Arrowsmith*, wskazujące na możliwe oparte na literaturze odkrycia, dotyczące zjawisk wspólnych dla migreny i cukru.

Start **A-Literature** C-Literature B-list Filter Literature

Search PubMed for migraine Search

History 1: Search migraine : 24219 Select from History

Use this query for ARROWSMITH? Yes

Items 1 - 20 of 24219 Page 1 of 1211

1:Headache in the Emergency Department.
Friedman BW, Lipton RB.
Curr Pain Headache Rep. 2011 Mar 15; [Epub ahead of print]
PMID: 21400252 [PubMed - as supplied by publisher]

2:Waterloo Eye Study. Data Abstraction and Population Representation.
Machan CM, Hrynchak PK, Irving EL.
Optom Vis Sci. 2011 Mar 10; [Epub ahead of print]
PMID: 21399556 [PubMed - as supplied by publisher]

3:Migraine features, associated symptoms and triggers: A principal component analysis in the Women's Health Study.
SchÅYarks M, Buring JE, Kurth T.
Cephalalgia. 2011 Mar 11; [Epub ahead of print]
PMID: 21398421 [PubMed - as supplied by publisher]

Rysunek 36: Wyszukiwanie literatury dotyczącej podstawowego analizowanego zagadnienia („A”) zgodnie z podejściem *Literature-Based Discovery* w programie *Arrowsmith*. Źródło: strona internetowa programu *Arrowsmith*.

Arrowsmith Project Home Page

Job Id: 2355

B-list

Filters:

- Semantic

Undo last edit
Undo all edits

View search history for printing

Clipboard

Start A-Literature C-Literature **B-list** Filter Literature

A: 24219 C: 50000 490
A-query: migraine
C-query: sugar

The B-list contains title words and phrases (terms) that appeared in both the A and the C literature. **490** articles appeared in both literatures and were not included in the process of computing the B-list but can be viewed [here](#). The results of this search are saved under id # **2355** and can be accessed from the start page after you leave this session. There are **18470** terms on the current B-list (**3343** are predicted to be relevant), which is shown ranked according to predicted relevance. The list can be further trimmed down using the filters listed in the left margin.

To assess whether there appears to be a biologically significant relationship between the AB and BC literatures for specific B-terms, please select one or more B-terms and then click the button to view the corresponding AB and BC literatures. Use Ctrl to select multiple B-terms.

Rank	Prob	B-term
1	0.97	adiponectin
2	0.97	diffusion weighted
3	0.97	cox-2
4	0.97	liquid chromatography tandem
5	0.97	chromatography tandem
6	0.97	optical coherence tomography
7	0.97	neuroimaging
8	0.97	transcranial magnetic stimulation
9	0.97	transcranial magnetic
10	0.97	pylori eradication
11	0.97	cardiovascular risk factor
12	0.97	helicobacter pylori eradication
13	0.97	antiphospholipid syndrome
14	0.97	atherosclerosis risk community
15	0.97	contrast enhanced
16	0.97	helicobacter pylori
17	0.97	weighted imaging
18	0.97	antiphospholipid antibody
19	0.97	helicobacter
20	0.97	electrospray

Restrict by semantic categories? Yes

Rysunek 37: Lista zagadnień („B”) występujących wspólnie z dwoma wskazanymi terminami („A” i „C”) zgodnie z podejściem *Literature-Based Discovery* w programie *Arrowsmith*. Źródło: strona internetowa programu *Arrowsmith*.

Arrowsmith Project Home Page
Job Id: 2355

Start	A-Literature	C-Literature	B-list	Filter	Literature
AB literature		B-term		BC literature	
migraine		coffee caffeine		sugar	
1: Potential teratogenic and neurodevelopmental consequences of coffee and caffeine exposure: a review on human and animal data. 1994 Add to clipboard		1: Coffee and caffeine ameliorate hyperglycemia, fatty liver, and inflammatory adipocytokine expression in spontaneously diabetic KK-Ay mice. 2010 Add to clipboard			

Rysunek 38: Zestawienie literatury dotyczącej potencjalnego odkrycia („B”) zgodnie z podejściem *Literature-Based Discovery* w programie *Arrowsmith*. Źródło: strona internetowa programu *Arrowsmith*.

Oprócz najpopularniejszego programu *Arrowsmith*, dostępne są inne, specjalistyczne narzędzia, wspierające *LBD* (Weeber, Kors, Mons 2005). Pracujący dla sektora obronności USA Ronald N. Kostoff zaproponował własny wariant *LBD*, wykorzystując oprogramowanie *VantagePoint* i zaawansowane analizy statystyczne, a swój pomysł opatentował (Kostoff 2005). Szeroki przegląd zastosowań *LBD* w medycynie, nauce i technice oferuje specjalny numer pisma "*Technological Forecasting & Social Change*", zawierający m.in. wprowadzający artykuł Kostoffa (2008a) z obszerną analizą wcześniejszych badań i publikacji. Opis proponowanych metod gromadzenia i analizy danych oferują Kostoff i in. (2008c), uzupełniając dotychczasowe przeszukiwanie literatury naukowej o analizy patentów. Debra L. Banville, pracująca w dziale B+R firmy *AstraZeneca*, prezentuje praktyczne spojrzenie na potencjał *LBD*, omawiając ograniczenia techniczne tradycyjnego podejścia oraz możliwości wykorzystania wyszukiwarek internetowych, w tym narzędzi oferowanych przez firmę *Google* oraz sposobów opisu związków chemicznych (Banville 2009: 5-6).

Literature-Based Discovery posiada ogromny potencjał naukowy i pozwala na skuteczne stymulowanie kreatywności w nauce i technologii. Zastosowania tego podejścia są ograniczone do rozwiązywania konkretnych problemów badawczych w poszczególnych dyscyplinach - nie nadaje się do analiz przekrojowych. Nie będą też potrafili wykorzystać go zarządzający pracami badawczymi lub decydujący o strategicznych priorytetach instytucji badawczych, gdyż *LBD* pozostaje narzędziem dla specjalistów dziedzinowych, poszukujących możliwości wykorzystania nowej wiedzy z odrębnych obszarów badań. Warto też wspomnieć o dużej czasochłonności analiz *LBD*. W publikacjach naukowych, dokumentujących wyniki *LBD*, prezentowane są te kombinacje zjawisk, które zidentyfikowano jako obiecujące - jednak analizy obejmują zwykle setki innych wariantów, stopniowo eliminowanych po lekturze literatury przedmiotu. Kostoff i in. (2008b) przytaczają przykład analizy dwóch dotychczas niezwiązanych obszarów badań, która wymaga około 6 osobomiesięcy na lekturę i interpretację specjalistycznej literatury w celu oceny przydatności dorobku innego obszaru nauki (Kostoff i in. 2008b: 293). Ze względu na uwarunkowania finansowe i czasowe, analizy *LBD* nadają się więc do wykorzystania w odniesieniu do wąskich, starannie zdefiniowanych wyzwań badawczych.

9.5. TRIZ

Nazwa *TRIZ* pochodzi od rosyjskiego określenia teorii rozwiązywania problemów wynalazczych (ros. *Теория решения изобретательских задач*). Przegląd historii i założeń *TRIZ* oferują m.in. Orloff (2006) i Mizrachi (2010). Twórca teorii, Henryk Altszuller, poszukiwał uniwersalnego sposobu stymulowania wynalazczości i w oparciu o analizy potężnego zbioru amerykańskich patentów, opracował „algorytm wynalazku”. Systematycznie stosowany algorytm rozwiązywania zadań wynalazczych miał zdaniem Altszullera zastąpić dotychczasowe, chaotyczne poszukiwania sposobów rozwiązania pojawiających się w nauce i technice wyzwań (Altszuller 1972: 17-19), które nazywano "metodą prób i błędów", ale błędy w poszukiwaniach wynalazczych były bardzo kosztowne, a każda kolejna podejmowana próba w rzeczywistości odwracała uwagę wynalazcy od właściwego kierunku poszukiwań, co Altszuller określił jako "wektor inercji" (Altszuller 1972: 19).

Badacze reprezentujący nurt *TRIZ* określają wynalazek jako całkowite lub częściowe przewyciężenie napotkanej sprzeczności technologicznej (Altszuller 1972: 86-87). Jak wyjaśniał Altszuller, często próby udoskonalenia określonego parametru maszyny negatywnie wpływają na inną jej właściwość (Altszuller 1972: 87). To oznacza powstanie sprzeczności technologicznej: wprowadzenie zamierzonej zmiany wywoła przyszłe problemy techniczne, dlatego niezbędne jest znalezienie oryginalnego rozwiązania, które pogodzi sprzeczne oczekiwania.

Dorobek *TRIZ* pozwala na uproszczenie każdego problemu wynalazczego przez jego przedefiniowanie i wyrażenie jako sprzeczności między wybranymi właściwościami, których lista jest ściśle określona i zawiera 39 pozycji, w tym: ciężar, długość, powierzchnię, objętość, szybkość, przyspieszenie, trwałość, temperaturę, moc, wydajność, dokładność, dogodność wytwarzania, komfort pracy, dogodność napraw i stopień zautomatyzowania (por. wczesna wersja w: Altszuller 1972). Altszuller stworzył „tablicę sprzeczności”, w której komórkach na przecięciu poszczególnych wierszy i kolumn znajdują się pola, symbolizujące sprzeczności technologiczne między konkretnymi właściwościami. W każdej takiej komórce znajdują się konkretne rekomendacje, dotyczące możliwości rozwiązania danej sprzeczności - zasady, oznaczane liczbami. Po kilkudziesięciu latach analiz, prowadzonych przez specjalistów z wielu obszarów, obecnie lista możliwych zasad rozwiązywania sprzeczności technologicznych liczy 40 pozycji i obejmuje m.in.: zasadę rozdrobnienia, zasadę ujednolicenia, zasadę "matrioszki", zasadę wstępnego naprężenia, zasadę kopiowania czy zasadę samoobsługi (Altszuller 1972: 137-170). Terminologia stosowana w obszarze *TRIZ* wymaga oczywiście odpowiedniego zrozumienia, aby we właściwy sposób definiować istotę konkretnej sprzeczności technologicznej, a następnie wybrać rekomendowane rozwiązanie.

Właściwość pogarszająca się → Właściwość doskonalona ↓	Ciężar	Długość	Powierzchnia	Objętość	Szybkość	...
Ciężar →		15 8 29 34	29 30 8 34	29 34 6 9	2 8 12 15	
Długość	8 14 15 29		4 14 15 17	7 17 14	13 14 34	
Powierzchnia	2 14 29 30	14 15 35		7 14 17	29 30 34	Zasady rozwiązania sprzeczności technologicznej
Objętość	2 14 29 8	1 7 4	1 7 34 17		29 18 34	
Szybkość	8 31 13	18 13 14	29 30 34	7 29 34		
...						

Rysunek 39: Fragment „tablicy sprzeczności” TRIZ. Źródło: opracowanie własne w oparciu o: Altszuller (1972).

Sprzeczność powinna zostać rozwiązana zgodnie z rekomendacjami *TRIZ*, odczytanymi w "tablicy sprzeczności". Przykładowo, jeśli sprzeczność dotyczy właściwości "ciężar" i "szybkość" - czyli w konkretnej sytuacji dążymy do zwiększenia wagi obiektu, jednak wpłynie to negatywnie na jego prędkość - z "tabeli sprzeczności" wyczytamy rekomendacje, dotyczące zastosowania zasad 2, 8, 12 lub 15 jako sposobu rozwiązania problemu wynalazczego. Zasada *TRIZ* o numerze 2 to "zasada separacji" czyli sugestia wyodrębnienia z obiektu jego części, która w danej sytuacji wydaje się szkodliwa. Zasada nr 8 to "zasada antycieżaru" - kompensacji nadmiernej wagi obiektu poprzez połączenie go z innymi obiektami, które mogą go unosić w powietrze, albo poprzez wykorzystanie układów i sił fizyki, które zmniejszą jego ciężar. W wielu sytuacjach korzystanie z rekomendacji *TRIZ* może stać się źródłem znaczących odkryć naukowych i technologicznych, w innych - systematycznie ukierunkowuje poszukiwania rozwiązania w obszarach, które są najbardziej obiecujące. Należy pamiętać, że poprawne zdefiniowanie istoty sprzeczności technologicznej, na jaką napotyka badacz, nie jest rzeczą prostą ani oczywistą (Orloff 2006: 93), podobnie jak nie będzie zadaniem banalnym interpretacja zasad *TRIZ* w celu rozwiązania tej konkretnej sprzeczności. Poprawna diagnostyka, stosowana przez specjalistów *TRIZ*, jest też oczywiście o wiele bardziej rozbudowana niż wynikać to może z powyższego opisu.

Algorytm stosowany w ramach *TRIZ* może wydawać się z pozoru arbitralnym zestawieniem rekomendacji, a może nawet "czarną magią". W rzeczywistości jest on oparty na kilkudziesięcioletnim dorobku radzieckich technik analitycznych, odwołujących się do

bibliometrii. "Tabela sprzeczności" wraz z rekomendacjami wykorzystania konkretnych zasad powstała w oparciu o analizę 2 milionów dokumentów patentowych Stanów Zjednoczonych, pobranych z *USPTO*. Interesujący jest fakt prowadzenia w ZSRR na przestrzeni kilkudziesięciu lat tak systematycznych analiz dorobku amerykańskiej nauki i techniki.

Badacze poszukiwali ogólnych prawidłowości, abstrahując od specyfiki poszczególnych wynalazków. W ten sposób doszli do wniosków, że analogiczne sposoby rozwiązywania dotyczą problemów konstrukcji maszyn górniczych, urządzeń medycznych, pojazdów samochodowych, lotnictwa czy procesów chemicznych. Identyfikacja podobieństw i generalnych reguł, rządzących rozwojem nauki i techniki oraz tworzeniem wynalazków wydaje się niespotykanym na skalę światową osiągnięciem. Okazało się, że mimo ogromnej różnorodności pojawiających się zadań wynalazczych, lista sposobów ich rozwiązania jest zamknięta i relatywnie krótka (Orloff 2006: 35). Zawartość "tabeli sprzeczności" i lista zasad ich rozwiązywania ulegały na przestrzeni lat dalszemu doskonaleniu - ich pierwszą wersję Altszuller opublikował w roku 1961, ostateczny wariant powstał w roku 1985 (Orloff 2006: 124), a w ciągu ostatnich 25 lat rosyjscy kontynuatorzy badań nad *TRIZ*-em wprowadzali systematyczne uzupełnienia, uwzględniając nowe sposoby rozwiązywania problemów wynalazczych w dobie rozwoju biotechnologii, nanotechnologii i optoelektroniki.

Zachodnie firmy technologiczne odkryły dorobek *TRIZ* dopiero po upadku Związku Radzieckiego. Wielu badaczy z byłego ZSRR zdecydowało się na emigrację, zasilając działy badawczo-rozwojowe wielkich amerykańskich korporacji, choć do wiernych "wyznawców" *TRIZ* należy też np. firma *Samsung*, zatrudniająca w Korei duży zespół rosyjskojęzycznych badaczy. Publikacje dotyczące *TRIZ* były w Polsce dostępne od lat 70-tych XX wieku, jednak nie cieszyły się popularnością ani w okresie PRL-u ani obecnie. Tymczasem praktyczne zastosowania *TRIZ* w odniesieniu do działalności wynalazczej uczyniły z tego podejścia niezwykle dochodowy obszar bibliometrycznych analiz nauki i technologii. Przykładowo, zespół białoruskich badaczy pod kierunkiem Valery'ego Tsurikowa, w schyłkowym okresie Związku Radzieckiego pracujący nad rozwojem specjalistycznego oprogramowania, które miało ułatwiać praktyczne stosowanie *TRIZ*, stał się załączkiem jednej z najważniejszych obecnie firm oferujących oprogramowanie wspierające analizy technologii *Invention Machine*, mającej obecnie w Bostonie swoją światową centralę, ale utrzymującą dział B+R w białoruskim Mińsku.

Do pierwszych dużych użytkowników oprogramowania *Invention Machine* należały firmy *Motorola* i *Mitsubishi* (Orloff 2006: 289), obecnie zaś materiały marketingowe firmy wymieniają wśród klientów m.in. przedsiębiorstwa *Boeing*, *Caterpillar*, *General Electric*, *Honeywell*, *Nestle*, *Samsung*, *Shell*, *Sony*, *Statoil*, *Unilever*, *Veolia Environment* i *Whirlpool*. Firma nie informuje obecnie w materiałach marketingowych o związkach ze szkołą *TRIZ*, prawdopodobnie ze względu na obawy o ewentualne roszczenia finansowe fundacji, dbającej o kontynuację tradycji Altszullera i kontrolującej odpowiednie znaki towarowe.

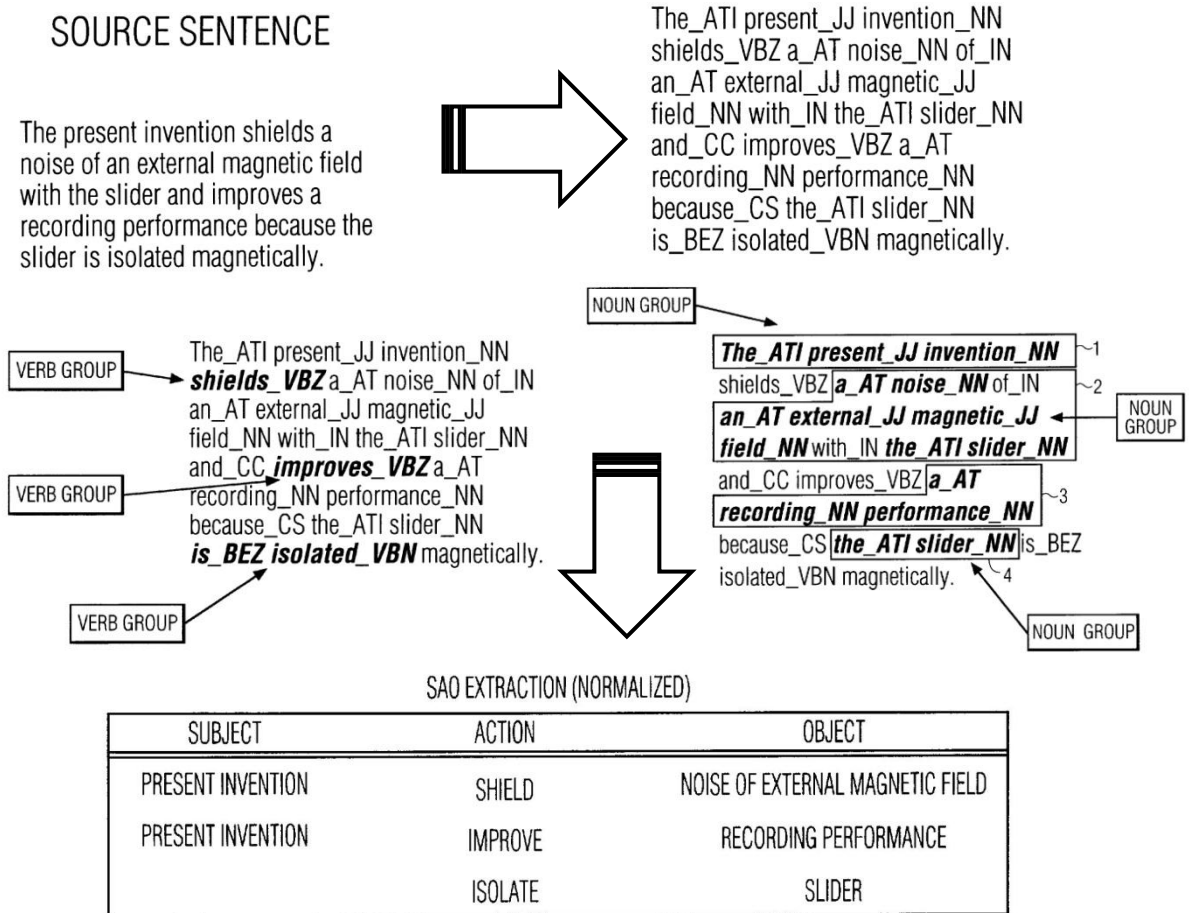
Program *Invention Machine Goldfire* opiera się na wykorzystaniu algorytmów sztucznej inteligencji do przetwarzania dużej liczby dokumentów patentowych i publikacji naukowych.

Stanowią one znaczące udoskonalenie klasycznych sposobów analiz bibliometrycznych - program potrafi interpretować znaczenie zdań, a nie tylko częstotliwości występowania czy współwystępowania zwrotów, jest też możliwe formułowanie w języku naturalnym zapytań do bazy światowej wiedzy naukowej i technologicznej. Wynikiem wieloletnich prac badawczych firmy *Invention Machine* jest też budowa listy specjalistycznych efektów technologicznych, które są uszczegółowieniem 40 zasad *TRIZ* - po identyfikacji podstawowych sposobów rozwiązania określonej sprzeczności technologicznej, wynalazca może skorzystać z dodatkowych wskazówek, w tym klasyfikacji ponad czterech tysięcy efektów chemicznych, fizycznych, biologicznych i geometrycznych, które zostają przez oprogramowanie dobrane odpowiednio do specyfiki analizowanej sprzeczności.

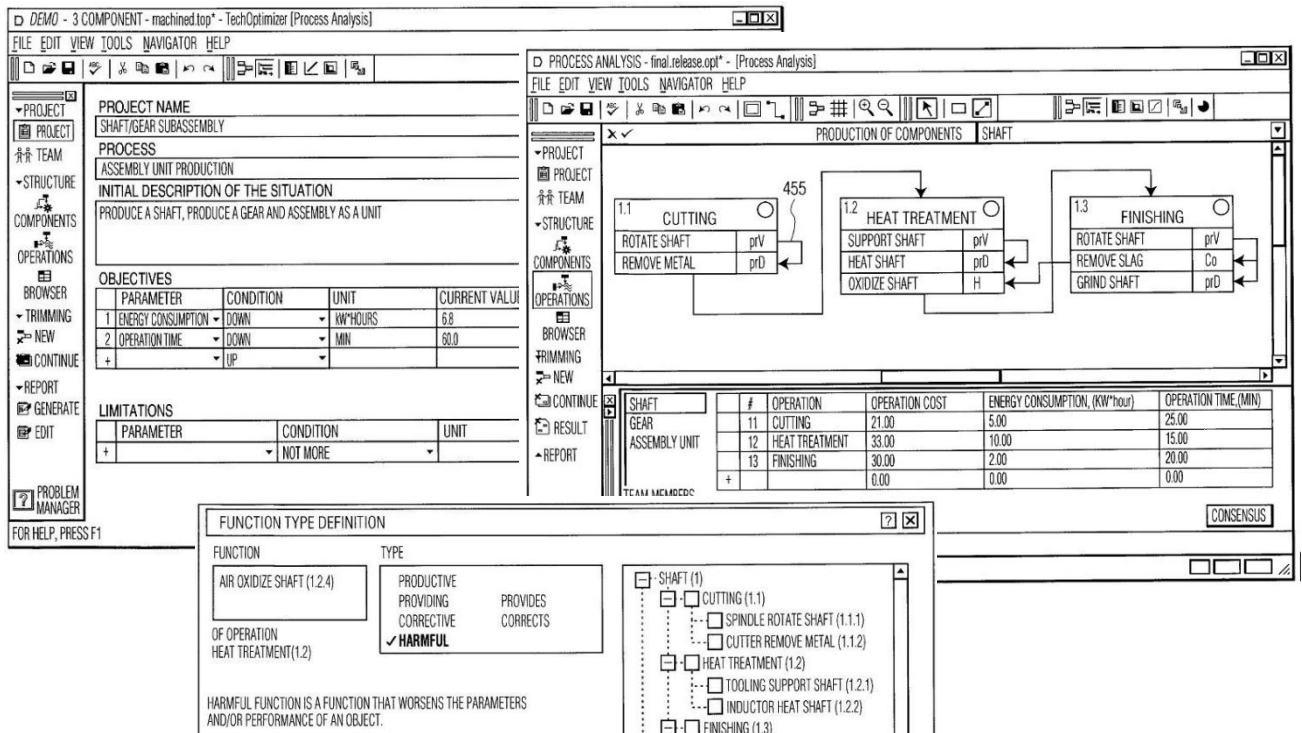
Można pokusić się o porównanie skali przedsięwzięcia *Invention Machine* do działalności *Google* - o ile *Google* koncentruje się na "zrozumieniu" struktury i zawartości ogólnodostępnych stron internetowych, o tyle *Invention Machine* dąży do poznania i interpretacji wiedzy naukowej z różnorodnych dyscyplin, przetworzenia jej przy wykorzystaniu technik sztucznej inteligencji i algorytmów maszynowego uczenia się tak, by następnie udzielać trafnych odpowiedzi na pytania badaczy i wynalazców. Partnerami-dostawcami treści do analiz są m.in. wydawcy baz, rząd Stanów Zjednoczonych jako wydawca bazy *MEDLINE* oraz organizacja techniczna *IEEE*, firma wykorzystuje też ogólnodostępne bazy patentów.

Metody programistyczne, wykorzystywane w rozwoju oprogramowania *Invention Machine Goldmine*, zostały szczegółowo udokumentowane w zbiorze kilkunastu patentów, udzielonych przez *USPTO* dla *Invention Machine Corporation*. Większość opatentowanych metod odnosi się do technik wydobywania użytecznej praktycznie wiedzy z publikacji naukowych i dokumentów patentowych. Rysunek 40 ilustruje przykład rozkładu zdania z wybranego dokumentu patentowego przez algorytm, wykorzystywany przez oprogramowanie *Invention Machine* (Tsourikov i in. 2000) i pozwala zrozumieć różnice pomiędzy opisywanymi wcześniej technikami interpretacji danych bibliometrycznych a podejściem, łączącym bibliometrię z przetwarzaniem języka naturalnego. Alternatywny sposób lingwistycznego rozkładu treści patentów, wykorzystujący bezpłatne oprogramowanie do statystycznej analizy tekstów *Stanford Parser* (<http://nlp.stanford.edu/software/lex-parser.shtml>), prezentują Yoon, Choi i Kim (2011).

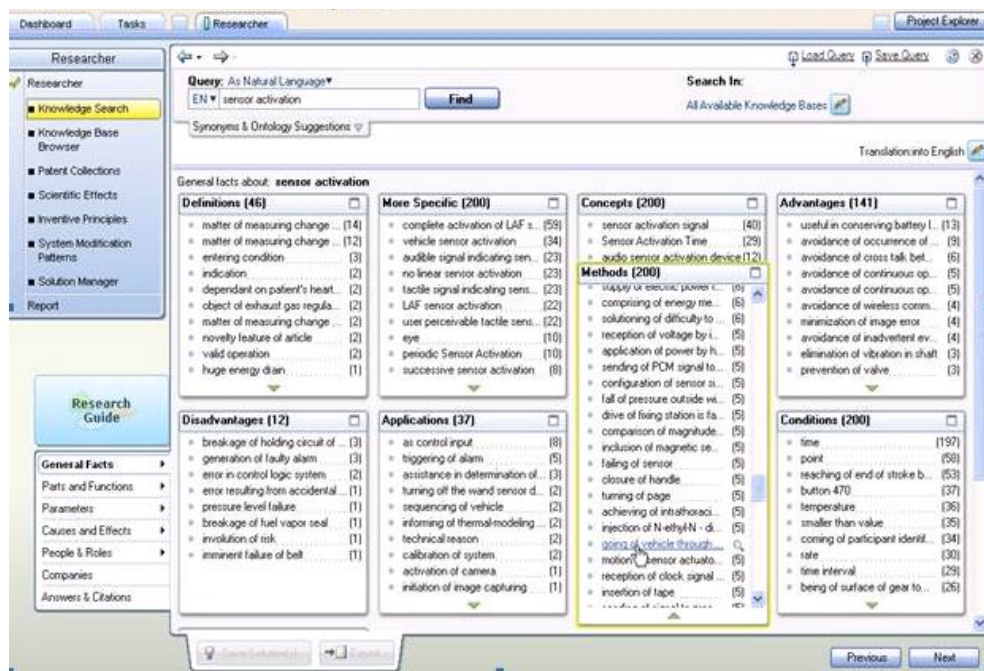
Szczegółowy opis architektury oprogramowania *Invention Machine* oraz zasad jego wykorzystywania znaleźć można w odpowiednich dokumentach patentowych (Devoino i in. 2001; Todhunter 2009; Todhunter i in. 2010). Rysunek 41 prezentuje możliwości wykorzystania oprogramowania do rozwiązywania konkretnych problemów wynalazczych i pochodzi z przykładu, omówionego szczegółowo w dokumencie patentowym (Devoino i in. 2001).



Rysunek 40: Algorytm komputerowego przetwarzania tekstów naukowych, wykorzystywany przez program firmy *Invention Machine*. Źródło: Tsourikov i in. (2000).



Rysunek 41: Przykładowe możliwości poszukiwania rozwiązania problemu wynalazczego w oprogramowaniu firmy *Invention Machine*. Źródło: Devoino i in. (2001).



Rysunek 42: Możliwości przeszukiwania zasobów wiedzy naukowej w programie *Invention Machine Goldfire*. Źródło: strona internetowa firmy *Invention Machine*.

Programy takie jak opisany *Invention Machine Goldfire* pozwalają na wykorzystanie zgromadzonych zasobów wiedzy naukowej, przetwarzanie ich przy wykorzystaniu algorytmów rozpoznawania języka naturalnego oraz identyfikację najbardziej pożądanых sposobów rozwiązywania określonych problemów technicznych. Analogiczne rozwiązania istnieją również dla obszaru nauk medycznych - niekoniecznie oparte na podejściu TRIZ,

jednak odwołujące się do zaawansowanych lingwistycznie i merytorycznie analiz publikacji. Alterovitz i Ramoni (2010: 157) oferują przykładowo przegląd rozwiązań, przetwarzających rekordy z bazy *MEDLINE* w celu ujawniania dorobku badań nad poszczególnymi genami, białkami, lekami czy gatunkami.

Techniki oparte o *TRIZ* są wykorzystywane przez wiele znanych firm technologicznych. Moehrle (2005) oferuje analizę porównawczą 45 studiów przypadku takich organizacji z Europy Zachodniej i Stanów Zjednoczonych. Moehrle i Lessing (2004) prezentują możliwości wykorzystania opartych o *TRIZ* analiz patentów do porównania kompetencji technologicznych firm i indywidualnych badaczy. Ich badania opierały się jednak na analizach zawartości tysięcy dokumentów patentowych, wykonywanych przez ekspertów, co znacząco ogranicza możliwe replikacje tego podejścia w innych projektach. Bardziej obiecującą inicjatywą bibliometryczną, inspirowaną *TRIZ*-em, jest podjęta przez badaczy z *Department of Mechanical Engineering, National University of Singapore* próba stworzenia nowej klasyfikacji patentów w oparciu o 33 z 40 zasad *TRIZ* (Cong, Tong 2007). Dla 7 pozostałych zasad *TRIZ* nie udało się zautomatyzowane przypisanie poszczególnych patentów do klasyfikacji, jednak badaczom udało się zaproponować 22 klasy patentów, odzwierciedlające wynalazki oparte o poszczególne zasady (niektóre z zasad zostały połączone w jedną grupę). Dokładnie udokumentowany przez badaczy algorytm wyszukiwania i klasyfikacji tekstu opiera się na występowaniu w tekstach poszczególnych słów lub ich kombinacji (Tong, Cong, Lixiang 2006; Cong, Tong 2007; Cong, Tong 2008) i może być samodzielnie wykorzystany przez innych badaczy do klasyfikacji danych tekstowych przy wykorzystaniu narzędzi takich jak wspomniane wcześniej narzędzie *VantagePoint*, wspierające *tech mining*. Potencjalne zastosowania algorytmu obejmują identyfikację dominujących sposobów rozwiązywania problemów technicznych przez określone firmy lub przegląd najczęściej stosowanych rozwiązań w danym obszarze technologicznym. Literatura opisuje też inną metodę zautomatyzowanej identyfikacji wynalazków, odwołujących się do ośmiu wybranych zasad *TRIZ* w bazach patentów (Glaser, Miecznik 2009: 92). Częstsze są jednak jakościowe analizy patentów zgodne z *TRIZ* (lektura i kodowanie zawartości), w których techniki bibliometryczne pomagają tylko na etapie identyfikacji odpowiednich dokumentów. Przykładem takich analiz może być klasyfikacja unowocześnień, wprowadzonych w konstrukcji pralki automatycznej w oparciu o zasady *TRIZ* (Jones i in. 2001: 7).

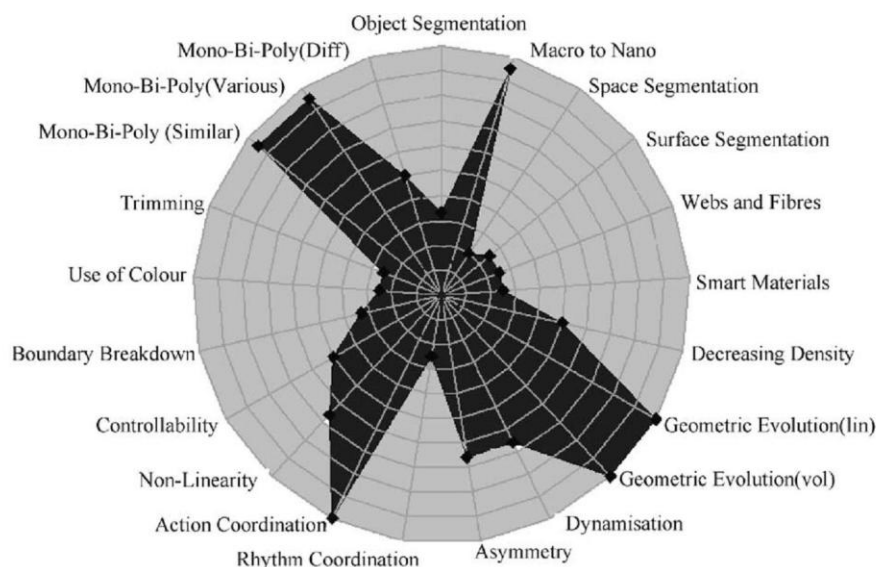
Interesujące możliwości stwarza wykorzystanie dorobku *TRIZ* do przewidywania kierunków dalszego rozwoju technologii. Rosyjski autor Urazajew prezentuje przegląd opublikowanych prognoz opartych o *TRIZ* (Urazajew 2006: 225-226). Obserwacja rozwoju wynalazków w poszczególnych obszarach technologicznych pozwoliła na określenie "naturalnych" praw, rządzących rozwojem systemów technicznych, których zastosowanie nie pozwala wprowadzić na opracowanie trafnej prognozy ilościowej, dotyczącej przyszłego rozwoju technologii, jednak pozwala na wskazanie prawdopodobnych kierunków jej ewolucji, a nawet przewidywanych przyszłych form (Orloff 2006: 192-221).

Wang, Chang i Kao (2010) zastosowali założenia *TRIZ* do przewidywania dalszego rozwoju technologii *MRAM* (pamięci magnetycznych, wykorzystywanych w elektronice). Po

przetworzeniu zawartości dokumentów patentowych przez program komputerowy, analizujący komunikaty w języku naturalnym, opracowali zestawienia dominujących cech technologii w poszczególnych okresach historycznych oraz opisali na tej podstawie trendy ewolucyjne technologii i zidentyfikowali podobszary i właściwości *MRAM*, w których należy spodziewać się dalszych zmian. Wykorzystana metoda badawcza nie została jednak dostatecznie dobrze udokumentowana w artykule, co uniemożliwia jej bezpośrednie powielenie; można też mieć podejrzenie, że opisywana analiza danych bibliometrycznych obejmowała wiele arbitralnych decyzji i interpretacji, nie wynikających bezpośrednio z uzyskanych danych - autorzy piszą bowiem o konieczności oparcia znaczącej części analiz na pracy ekspertów (Wang, Chang, Kao 2010: 506), choć wcześniejsze opisy mogłyby sugerować znaczącą automatyzację wspieranych komputerowo analiz.

Yoon i Park (2004) proponują zbliżoną technikę *KBMA* - analizy morfologicznej opartej na słowach kluczowych (ang. *keyword-based morphology analysis*). Analiza morfologiczna jest techniką wspierającą kreatywność, zaproponowaną kilkadziesiąt lat temu przez Fritza Zwicky'ego. Polega na rozpisaniu złożonego problemu na poszczególne zagadnienia składowe i zebraniu dla każdego z nich listy możliwych wariantów. W tak stworzonym zbiorze można następnie zaznaczyć kombinacje wariantów, które już istnieją (np. są dostępne w oferowanych przez konkurencyjnych dostawców produktach), aby następnie zidentyfikować takie kombinacje, których nikt jeszcze nie wdrożył, a które wydają się być atrakcyjne od strony technicznej lub komercyjnej. Analizę morfologiczną można zastosować do obszaru badań naukowych i rozwoju technologii w celu identyfikacji nowych, dotychczas nie rozpatrywanych wariantów technicznych, właściwości i zależności. Badacze skoncentrowali się na analizie podzbioru patentów z bazy *USPTO*, wybierając zbiór słów kluczowych, reprezentujących istotne cechy technologii ekranów ciekłokrystalicznych (*TFT-LCD*), a następnie porównując za ich pośrednictwem osiągnięcia poszczególnych konkurentów, wskazując na nowe, dotychczas nie rozpatrywane kombinacje parametrów technicznych. Yoon, Phal i Probert (2008) opisują udoskonalony proces analiz opartych o *KBMA*, wykorzystując zbiór patentów i podręczników użytkownika dla telefonów komórkowych firmy Nokia. Podejście *KBMA* jest zbliżone do opisanych powyżej technik analizy literatury patentowej, jednak nie korzysta z teoretycznych założeń *TRIZ*, dotyczących prawidłowości w rozwoju technologii.

Warto przypomnieć, że oryginalne analizy patentów, prowadzone przez Altszullera i jego współpracowników, były analizami o charakterze jakościowym a nie ilościowym i polegały na uważnej lekturze każdego z dokumentów oraz przypisaniu go do określonej grupy, odpowiadającej zastosowaniu jednej z zasad *TRIZ*. Wsparcie dla analiz ilościowych oferuje kosztowne oprogramowanie *Invention Machine*, jak również europejskie rozwiązanie *CREAX Creation Suite*. Jego przykładowe możliwości wykorzystania do oceny potencjału ewolucyjnego technologii ilustruje artykuł Manna (2003) oraz Rysunek 43, przedstawiający przykładową ocenę stopnia dojrzałości wynalazku, opisanego w dokumencie patentowym *USPTO* nr 4.174.358.



Rysunek 43: Przykład analizy potencjału ewolucyjnego technologii zgodnie z podejściem *TRIZ*. Źródło: Mann (2003: 788).

Dopiero w ostatnich latach rozpoczęto próby wsparcia analiz ewolucji zgodnie z założeniami *TRIZ* przy wykorzystaniu łatwiej dostępnego oprogramowania i ilościowych technik bibliometrycznych. Teksty badawcze dotyczące tych zagadnień pojawiają się w międzynarodowych czasopismach naukowych od niedawna i można przypuszczać, że prawdziwy renesans *TRIZ* i powiązanie radzieckiej teorii z dorobkiem zachodniej ilościowej bibliometrii dopiero nadejdzie.

Jednocześnie nie wydaje się, by założenia *TRIZ* dało się bezpośrednio przełożyć na oczekiwania i cele, przyświecające typowym projektom analiz bibliometrycznych. Algorytmy stworzone w ramach *TRIZ* nadają się bardziej do wsparcia indywidualnych prac wynalazczych, do tych celów służy również dedykowane oprogramowanie, omawiane powyżej. Z kolei techniki ilościowego przetwarzania dużych zbiorów danych na temat badań naukowych i rozwoju technologii nie pozwalają jeszcze na poważne zastosowania dorobku *TRIZ* do kształtowania polityki naukowej lub strategii badawczych organizacji sektora B+R.

9.6. Identyfikacja nowych trendów i tworzenie map badań

Kolejny podrozdział będzie opisywał grupę technik, wspierających analizy i wizualizację danych bibliometrycznych. Ta grupa nie posiada w literaturze jednej wspólnej nazwy, jednak poszczególne techniki charakteryzują się podobieństwami w sposobie wykorzystywania lub zastosowaniach. Elementem wspólnym jest też wsparcie dla tej grupy technik przez opisane w podrozdziale narzędzia informatyczne, w tym program *CiteSpace*, korzystający ze złożonych algorytmów obliczeniowych.

Wykorzystanie zaawansowanych technik statystycznych w odniesieniu do danych bibliometrycznych pozwala na realizację dwóch rodzajów analiz:

- budowy modeli prognostycznych, opartych o liczbę publikacji i patentów w wybranym obszarze badań, a tym samym przewidywanie dalszego tempa jego rozwoju (przy świadomości, że modele ilościowe nie uwzględniają ewentualnych zmian jakościowych, wynikających z możliwości pojawienia się przełomowych odkryć i wynalazków),
- wizualizacji danych bibliometrycznych, pozwalającej na obserwację podobieństw i różnic pomiędzy obszarami badań, pracami konkretnych naukowców, czy orientacjami ośrodków badawczych i krajów.

Zgromadzone dane historyczne, odzwierciedlające rozwój badań mierzonych liczbami publikacji lub patentów mogą być wykorzystywane do tworzenia krótkookresowych prognoz i przewidywania aktywności badawczej w kilku nadchodzących latach. Warto pamiętać o ograniczonych zastosowaniach prognoz w bibliometrii - aktywność badawcza nie jest zależna jedynie od wcześniejszych prac, ale także od nowych osiągnięć i odkryć, zmian zainteresowań i instytucjonalnych decyzji o inwestycjach w określone obszary badań.

Podstawą dla analizy szeregów czasowych mogą być zestawienia rocznych wartości albo wartości skumulowanych w czasie (np. skumulowanej liczby publikacji lub patentów dla kolejnych lat). Mogą być one punktem wyjścia do poszukiwania odpowiedniej funkcji regresji, która opisze rozwój analizowanego zjawiska. Przy modelowaniu wzrostu istotny jest odpowiedni dobór analizowanych obiektów - analiza szczegółowych podobszarów badawczych wydaje się podejściem lepszym od ilościowych zestawień szerokiego dorobku dyscypliny, gdyż uzyskane zestawienia pomagają zidentyfikować tendencje, oparte na łatwych do zaobserwowania zjawiskach w danym podobszarze. Oprócz budowy modelu w oparciu o liczbę artykułów, interesujące możliwości oferuje też modelowanie przyrostu liczby słów kluczowych, odzwierciedlające wzrastającą specjalizację obszaru badań (Porter, Cummingham 2005: 206-207).

W odniesieniu do danych bibliometrycznych, szczególnie przydatne okazują się cztery modele (Porter, Cummingham 2005: 194-195):

- model wzrostu liniowy - oparty na równaniu $y = ax + b + e$ (e - składnik losowy),
- model wzrostu wykładniczego - w którym analizowane wielkości są interpretowane jako logarytm naturalny zmiennej y czyli $\ln(y)$,
- model wzrostu Gompertza - $\ln(l/y)$, l = górny limit wzrostu,
- model wzrostu Fishera-Pry'a - $\ln\{(l-y)/y\}$, l = górny limit wzrostu.

Przydatny może okazać się również model Lotki-Volterry - bardziej złożony, obejmujący sprzężenie zwrotne czyli wzajemne zależności między x i y (Porter i in. 1991: 187-196). Wybór odpowiedniego modelu powinien opierać się na weryfikacji dopasowania danych empirycznych. Model wzrostu wykładniczego szczególnie dobrze oddaje dynamiczny rozwój popularnych obszarów badawczych. Krzywe Gompertza i Fishera-Pry'a odpowiadają obserwowanej w wielu obszarach nauki i techniki, S-kształtnej tendencji wzrostowej, czyli początkowo powolnym przyrostom, po których następuje dynamiczny rozwój, który stopniowo zaczyna zwalniać wraz ze zbliżaniem się do wartości l czyli górnego limitu

wzrostu (Porter i in. 1991: 175-187). Porter i Cummingham (2005: 195) podpowiadają sposób określenia tego limitu w przypadku analiz bibliometrycznych: wartość graniczna nie wynika w tych przypadkach z ograniczeń fizycznych, może być jednak obliczona w oparciu o założenia, dotyczące maksymalnej liczby artykułów, które mogą być opublikowane rocznie, w zależności np. od liczby naukowców lub czasopism w danej dziedzinie badań. Warto podjąć próby podstawienia do modelu kilku różnych wartości l i porównania uzyskanych wyników (Porter, Cummingham 2005: 195).

Drugi obok modelowania wzrostu sposób wykorzystania technik statystycznych w bibliometrii stanowi zaawansowana wizualizacja. Wizualizacja danych bibliometrycznych służy graficznemu odzwierciedleniu wyników analiz, które były opisywane we wcześniejszych podrozdziałach - zestawień dotyczących współautorstwa publikacji, współwystępowania w nich określonych zagadnień (*co-word*, *co-classification*) oraz wzajemnych cytowań.

Najprostszy wariant reprezentacji graficznej opierać się może na danych bezpośrednio pobranych z macierzy współwystępowania, bez dodatkowych przekształceń. Wskazana jest jednak normalizacja danych tak, by uzależnić wyniki współwystępowania od całkowitej aktywności danego obiektu - przykładowo, w odniesieniu do autora który napisał tylko 2 teksty, ale oba wspólnie ze współpracownikiem, należy odzwierciedlić fakt 100% zaangażowania w pracę zespołową, podczas gdy zwykła analiza współwystępowania ujawniłaby jedynie dwa współautorskie teksty, ginące w zestawieniach, w których pojawiają się także dane dotyczące innych, zdecydowanie bardziej produktywnych badaczy. Normalizacja pozwala na identyfikację podobieństw między jednostkami niezależnie od indywidualnej produktywności.

Przy omawianiu technik stosowanych w *tech mining*-u, pojawiły się opisy podstawowych sposobów normalizacji danych w oparciu o współczynniki autokorelacji i korelacji krzyżowych – zdaniem wielu badaczy, o poprawności analiz decydować będzie wykorzystany sposób przekształcania danych o współwystępowaniu (Eck, Waltman 2009: 1645). Oprócz algorytmów obliczeniowych dostępnych w narzędziach do *tech mining*-u, badacze zainteresowani wykorzystaniem alternatywnych technik normalizacji mogą jednak pobrać z programu macierz z danymi o współwystępowaniu, a następnie dokonać odpowiednich transformacji w innym programie, m.in. w arkuszu kalkulacyjnym. Istnieje wiele innych szczegółowych metod, których wykorzystanie możliwe jest po eksporcie danych z programu takiego jak *VantagePoint* do specjalistycznych programów matematycznych lub nawet arkusza kalkulacyjnego, w którym opisano odpowiednie wzory obliczeniowe. Leydesdorff (2008) oferuje obszerną dyskusję wad i zalet poszczególnych metod normalizacji danych bibliometrycznych dotyczących współwystępowania na przykładzie analizy cytowań, omawiając metodę Saltona (ang. *Salton's cosine*), współczynnik korelacji oraz indeks Jaccarda (Leydesdorff 2008). Inny przegląd sposobów normalizacji danych oferują Eck i Waltman (2009).

Oprócz prób przewidywania dalszego rozwoju zaobserwowanych trendów, szczególne znaczenie praktyczne ma identyfikacja znaczących odchyłeń od tendencji wzrostowych, które

mogą być postrzegane jako odzwierciedlenie zjawisk cieszących się nagłą popularnością (ang. *hot topics*), w tym również nowych technologii (ang. *emerging technologies*).

Do identyfikacji wymienionych zjawisk nie wystarczy prosta obserwacja trendów wzrostowych. Nowe zagadnienia badawcze nie znajdują zwykle odzwierciedlenia w tradycyjnych taksonomiach i słowach kluczowych, co wynika bezpośrednio z ich nowości. Redaktorzy baz mogą początkowo nie potrafić dostrzec, odpowiednio nazwać i zaklasyfikować tych zjawisk. Można założyć, że przynajmniej część z nich zostanie błędnie przypisana do innego obszaru badawczego w wyniku niezrozumienia lub niedocenienia doniosłości nowych prac badawczych, a dane bibliometryczne dotyczące owych *emerging technologies* charakteryzuje znaczący szum informacyjny. Co więcej, w typowym scenariuszu analiz, osoba gromadząca i przetwarzająca dane nie wie jeszcze o istnieniu określonego zjawiska, a analizy powinny pomóc w jego identyfikacji. W takiej sytuacji, dużym wyzwaniem będzie nawet określenie obszaru, który miałby podlegać szczegółowym analizom. Konsultacja z ekspertami dziedzinowymi zwykle nie wystarcza - identyfikacja tendencji opierałaby się wówczas głównie ich subiektywnych opiniach i z oczywistych powodów ograniczonej wiedzy, a nie na systematycznej, bibliometrycznej analizie dyscypliny, a opinie różnych ekspertów dotyczące nowych zjawisk w nauce i technice mogą być istotnie rozbieżne (Tseng i in. 2009: 74).

Jednym z możliwych podejść do identyfikacji nowych trendów badawczych jest rezygnacja z wykorzystania klasycznych, dotychczas opisanych sposobów określania obszarów badań i technologii na rzecz przetwarzania danych masowych, potężnych zbiorów opisujących różne obszary badań i technologii i podejmowania prób identyfikacji wyróżniających się zjawisk przy wykorzystaniu technik statystycznych. W praktyce, takie podejście oznacza konieczność pobrania dużej liczby obiektów z baz danych w oparciu o szerokie kryteria, takie jak rok publikacji, kraj pochodzenia badaczy lub szeroko zdefiniowana dyscyplina nauki (np. fizyka), a nie słowa kluczowe czy elementy klasyfikacji (Mogoutov i in. 2008: 343). W analogiczny sposób prowadzone są klasyczne projekty eksploracji danych (ang. *data mining*), należy jednak zauważyć, że narzędzia dedykowane do przetwarzania danych bibliometrycznych nie są zwykle w stanie pracować z tak rozbudowanymi zbiorami danych.

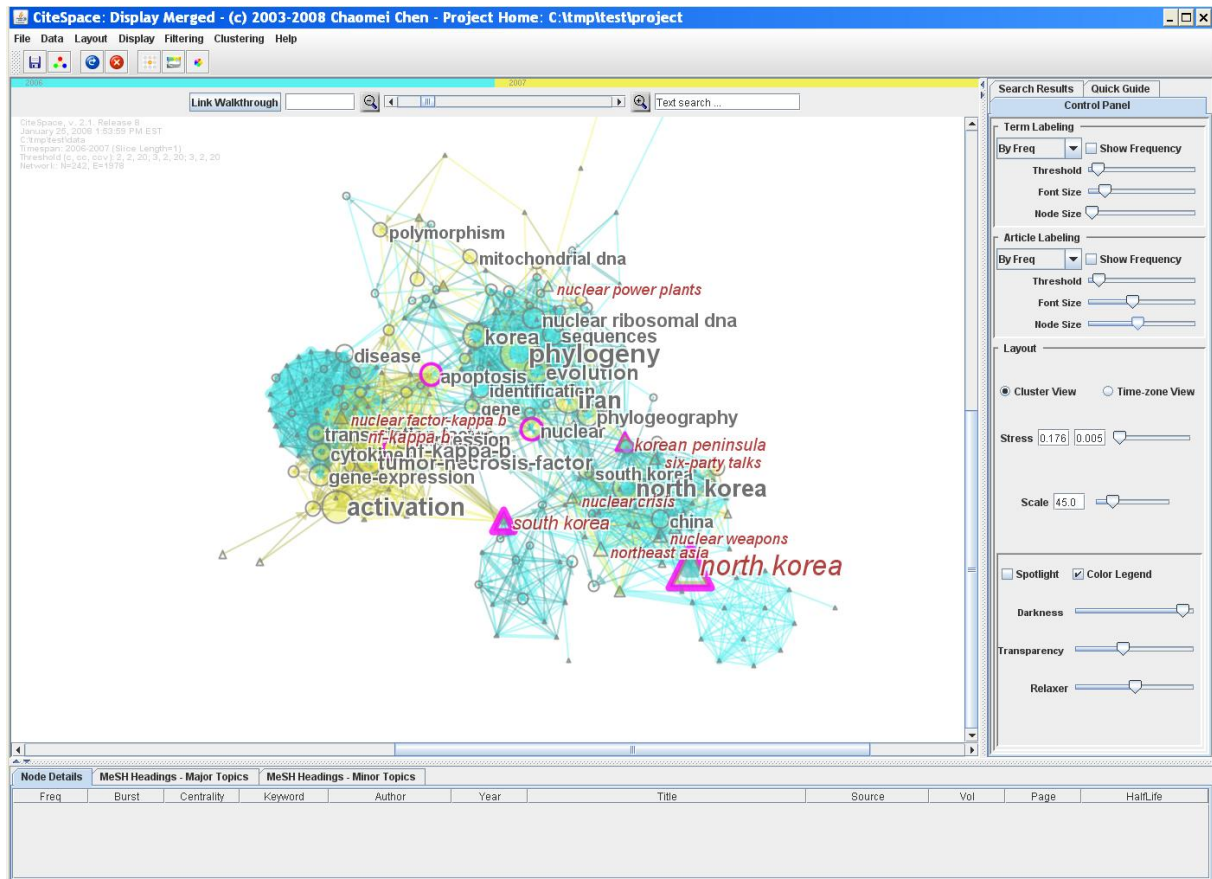
Tseng i in. (2009: 74-75) opisują próby, podejmowane przez teoretyków oraz japoński *National Institute of Science & Technology Policy*, oparte na wielokrotnym porównywaniu dwóch części serii danych historycznych, pozwalającym na identyfikację nieproporcjonalnie wysokiego procentowego przyrostu lub spadku liczby publikacji w poszczególnych latach czyli analizę wycinków czasowych (ang. *time slice*). Badacze z Tajwanu zaproponowali lepszą metodę, odwołującą się do zaawansowanych technik statystycznych. Jej szczegółowy opis oferuje Tseng i in. (2009). Analizowane szczegółowe tematy badań (słowa kluczowe, kody klasyfikacyjne) są najpierw grupowane przy wykorzystaniu analizy skupień w hierarchicznie ułożone klastry, reprezentujące obszary badawcze. Dla każdego poziomu hierarchii gromadzone są następnie serie danych historycznych i obliczane indeksy, pozwalające na zróżnicowanie średnich odstępów czasowych między publikacjami oraz funkcje trendu dla każdego z analizowanych obszarów. Po zebraniu danych o liczbie publikacji w macierzach są one poddawane dekompozycji, która ujawnia indywidualne

parametry funkcji trendu, dotyczące poszczególnych obszarów i pozwala na łatwą identyfikację tych obszarów, które podlegają zmianom ilościowym istotnie przewyższającym zmiany w pozostałych obszarach (Tseng i in. 2009: 75-78). Ta prosta metoda stwarza niestety wiele wyzwań praktycznych, gdyż wymaga dużych mocy obliczeniowych i oprogramowania, mogącego przetwarzać potężne zbiory danych. Dane wejściowe do analiz statystycznych, mających na celu identyfikację trendów, mogą być opracowane przy wykorzystaniu narzędzi do *tech mining*-u takich jak program *VantagePoint*, samo szacowanie funkcji trendu i dekompozycja macierzy musi być jednak wykonana w specjalistycznym oprogramowaniu.

Alternatywne podejście do identyfikacji nowych zjawisk w nauce i technice przy wykorzystaniu danych bibliometrycznych może opierać się na analizie cytowań. Tseng i in. (2009: 80) byli zmuszeni do rezygnacji z analizy cytowań ze względu na zbyt duże rozmiary analizowanego zbioru danych i ograniczenia funkcjonalne programów do analiz statystycznych. Chen (2006a) opisuje jednak uwieńczony sukcesem projekt analityczny oraz oferuje dyskusję uwarunkowań metodologicznych tego podejścia i wcześniejszych przykładów zbliżonych badań. Chen stworzył pakiet oprogramowania *CiteSpace*, służący do przetwarzania sieci cytowań - w tym analizy wycinków czasowych (ang. *time slices*) i wykrywania znaczących odchyłeń od trendów rozwojowych, określanymi jako "wykrywanie wybuchów" (ang. *burst detection*). Algorytm *burst detection* opiera się na dorobku wcześniejszych badań innych naukowców, a jego podstawową zaletą jest umiejętność identyfikacji znaczących tendencji niezależnie od tego, jak często cytowane są publikacje dominujące w określonym nurcie badawczym (Chen 2006a: 364), co jest korzystne w odniesieniu do nowatorskich tematów i podejść badawczych. Naukowe podstawy *burst detection* stanowią metody identyfikacji nieproporcjonalnych przyrostów występowania określonych słów lub natężenia zjawisk, znane w różnych obszarach badań (por. przegląd w: Chen 2006b: 272-274).

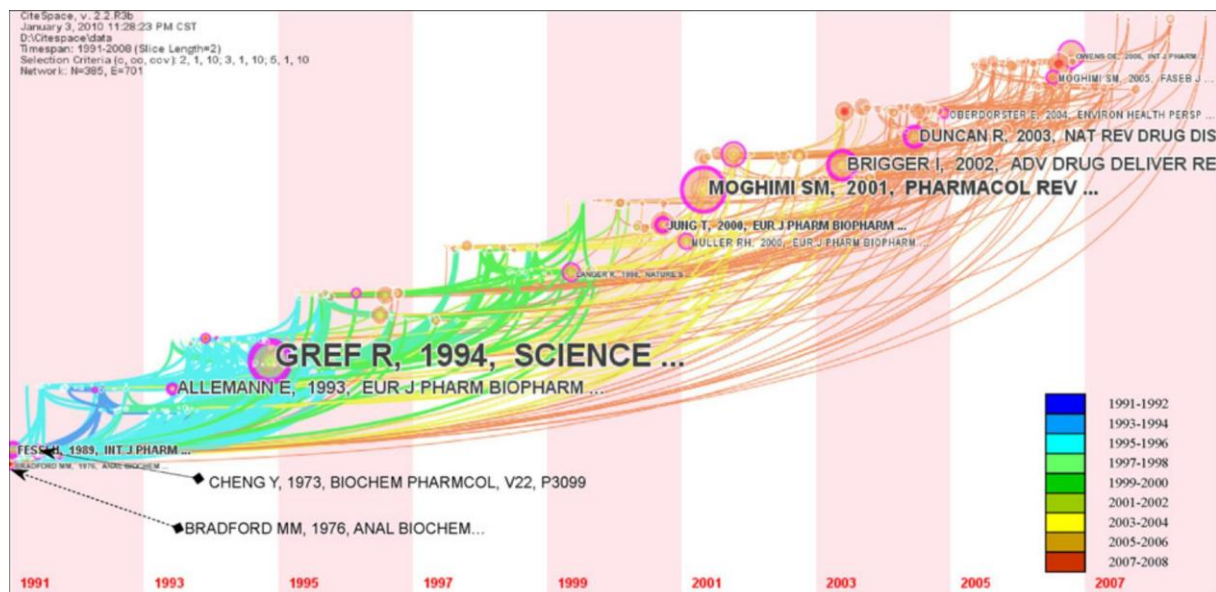
Program *CiteSpace* jest bezpłatnie dostępny na stronie <http://cluster.cis.drexel.edu/~cchen/citespace> i pozwala na samodzielne prowadzenie w oparciu o bazę *Web of Science* zestawień rankingowych i tabelarycznych, jak również zaawansowanych analiz sieci cytowań.

Chaomei Chen zajmuje się badaniami nad wizualizacją informacji (por. Chen 2006b), a opracowany przez niego program oferuje bogate możliwości prezentacji wyników analiz, m.in. poprzez łatwe do interpretacji oznaczanie "wybuchów" publikacji kolorem czerwonym oraz prezentację "cieni" aktywności publikacyjnej we wcześniejszych okresach niż stanowiącym podstawę analiz. Program *CiteSpace* wykonuje analizy statystyczne, pozwalające na grupowanie sieci według współwystępowania pojęć, autorów, instytucji i krajów, może więc być wykorzystywany do realizacji niektórych zadań, opisanych w podrozdziale dotyczącym *tech mining*-u. Rysunek 44 prezentuje przykładowe możliwości omawianego programu.



Rysunek 44: Przykładowa mapa zależności pomiędzy pojęciami, występującymi w publikacjach naukowych, stworzona w programie *CiteSpace*. Źródło: strona internetowa programu *CiteSpace*.

Wypracowane przez Chena metody analityczne pozwalają również na wizualizację sieci publikacji i cytowań w przekroju czasowym, czego przykładem jest Rysunek 45. Bogaty przegląd możliwych zastosowań programu *CiteSpace* oferują Chen i in. (2008) – obok identyfikacji nowych tendencji badawczych oraz wizualizacji danych bibliometrycznych, możliwe są również niektóre analizy, omawiane w podrozdziale dotyczącym techniki *tech mining*. Opis możliwości wykorzystywania *CiteSpace* do analizy i prezentacji graficznej danych oferuje Chen (2006b: 287-290), sam program zostanie też opisany w dalszej części niniejszej książki.



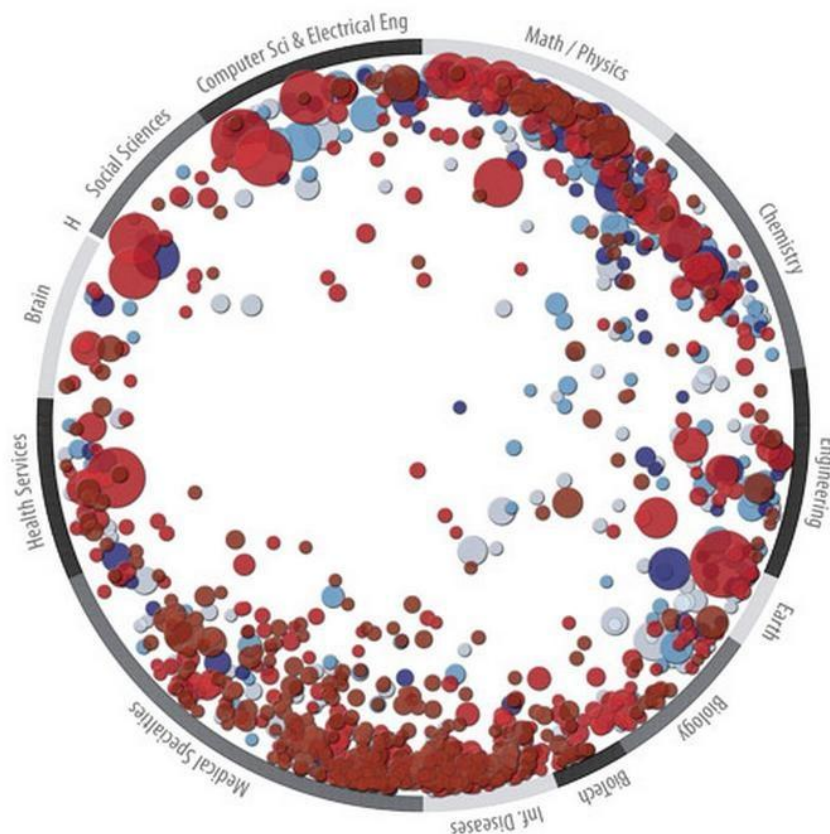
Rysunek 45: Wizualizacja sieci cytowań w przekroju czasowym przy wykorzystaniu programu CiteSpace. Źródło: Chen, Guan (2011: 245).

Zbliżone rozwiązania technologiczne i algorytmny mogą być wykorzystywane nie tylko do identyfikacji nowych tendencji badawczych, ale również do budowy kompleksowych map dyscyplin naukowych i całego dorobku światowej nauki. Opisywane dotąd narzędzia i metody koncentrowały się na wizualizacji sieci zależności pomiędzy wybranym podzbiorem danych z bazy. Możliwe jest jednak także "zmapowanie" pełnej struktury badań naukowych albo zestawienie wybranego podzbioru danych bibliometrycznych z taką mapą ("rzut na mapę"). Jest to obecnie możliwe dzięki projektom badawczym, których przeprowadzenie wymagało znaczących nakładów pracy i mocy obliczeniowych komputerów. Co ciekawsze, wyniki jednego z tych projektów stały się podstawą do stworzenia rozwiązania komercyjnego, podczas gdy rezultaty drugiego, porównywalnego przedsięwzięcia, są dostępne bezpłatnie dla wszystkich zainteresowanych. Prezentowane wcześniej analizy wycinkowe wymagały arbitralnych decyzji przy wyodrębnianiu skupień lub podzbiorów słów kluczowych, podczas gdy tworzenie map nauki, oparte na przetwarzaniu potężnych zbiorów danych bibliometrycznych, jest pozbawione tej wady (Chen, Zhang, Vogeley 2010: 673).

Do tradycyjnych sposobów wizualnego odwzorowania różnic i podobieństw między obiektami należy technika skalowania wielowymiarowego *MDS* (ang. *multidimensional scaling*), wspierana przez większość programów statystycznych. Istotą skalowania wielowymiarowego oraz alternatywnych technik jest redukcja liczby wymiarów w zbiorze danych, co pozwala na łatwiejsze dostrzeżenie niektórych zależności, niestety prowadzi również do utraty części informacji (Okubo 1997: 31). *MDS* bywa czasem krytykowane za niedostateczne uwypatnienie różnic między obiektami. Szczegółową dyskusję korzyści i ograniczeń związanych z zastosowaniem poszczególnych podejść analitycznych w bibliometrii oferują Börner, Chen i Boyack (2003), omawiając m.in. techniki analizy czynnikowej, *MDS*, dekompozycji wartości własnych (*eigenvalue*), sieci *pathfinder* i samoorganizujących się map (ang. *self-organizing maps*). Szczegółowa dyskusja uwarunkowań metodologicznych, matematycznych algorytmów i sposobów rzutowania

danych na wykresy wykracza poza zakres niniejszej publikacji. Zainteresowani czytelnicy mogą zapoznać się z naukowymi podstawami wizualizacji danych bibliometrycznych w książce Chena (2006b), autora programu *Citespace*, w którym wdrożył najważniejsze propozycje badawcze. To podejście jest zbliżone do bezpłatnego programu *VOSviewer* (www.vosviewer.com), który koncentruje się na wizualizacji sieci bibliometrycznych i opiera na algorytmie nazywanym "wizualizacją podobieństw" - czyli *VOS* (ang. *visualization of similarities*), zaproponowanym przez Ecka i in. (2010). Podobne były też zainteresowań badawcze zespołu z udziałem Kevina W. Boyacka z *Sandia National Laboratories*, który stworzył narzędzie do wizualizacji danych bibliometrycznych *VxInsight* (Boyack, Wylie, Davidson 2002).

Analiza skupień współcytowań (ang. *co-citation cluster analysis*) oferuje interesujące możliwości identyfikacji pokrewnych badań i wyodrębniania obszarów badawczych, co pokazali m.in. Small (1999; 2006), Klavans i Boyack (2006) oraz Upham i Small (2010). Na dużą skalę te techniki wykorzystali Boyack i Klavans (2010), tworząc przy wykorzystaniu programu *VxInsight* mapy poszczególnych dyscyplin badawczych, których uproszczona wersja jest dostępna na stronie www.mapofscience.com (por. Rysunek 46, prezentujący przykładową mapę słabości badawczych Stanów Zjednoczonych). Obecnie ich firma - *SciTech Strategies, Inc.* - prowadzi dla klientów zamawiane prace analityczne oparte na stworzonych mapach. Wyniki map nauki, metodę analityczną i technologie wizualizacyjne wykorzystwała też firma *Elsevier* przy tworzeniu komercyjnego serwisu *SciVal Spotlight* i subskrypcja tego serwisu jest najprostszym sposobem na zamówienie map, interesujących daną instytucję.

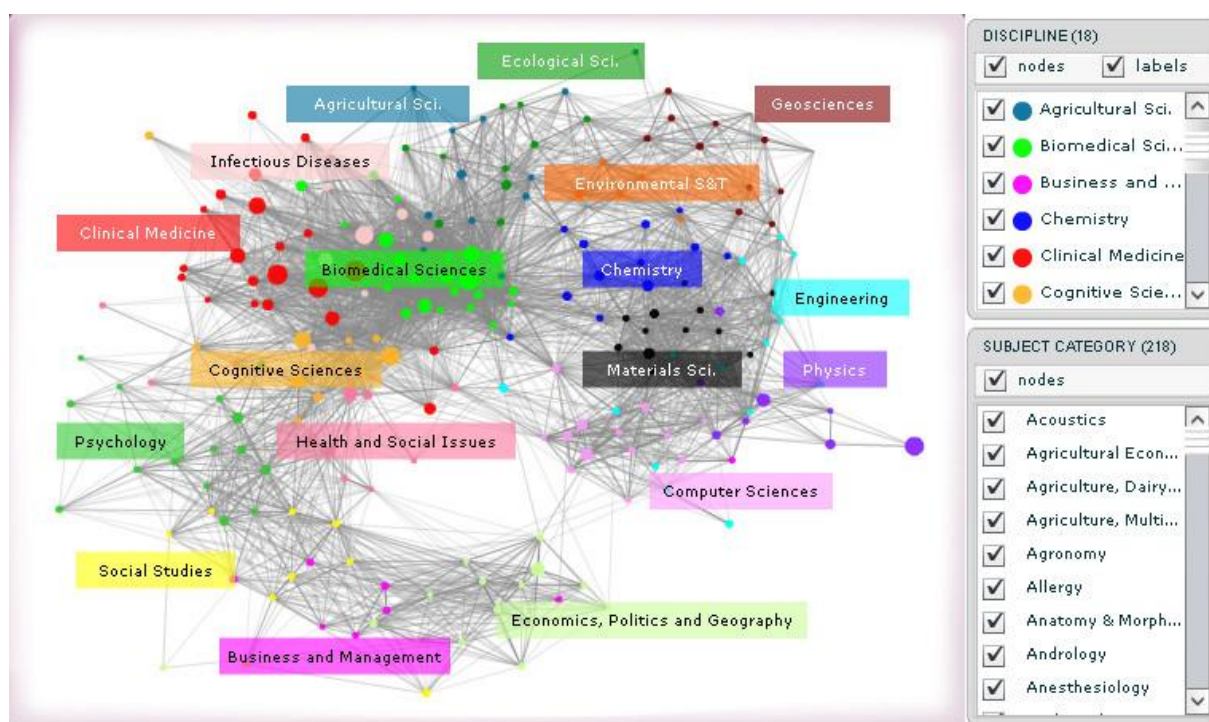


Rysunek 46: Mapa obszarów badawczych, które są mało konkurencyjne i niedostatecznie rozwinięte w Stanach Zjednoczonych. Źródło: MapOfScience.com

Alternatywą dla płatnego serwisu firmy *Elsevier* jest rozwiązanie bezpłatne, stworzone przy współpracy ważnych uczelnianych ośrodków badań nad nauką i technologią z Sussex, Amsterdamu i Atlanty. Ismael Rafols z centrum badawczego *Science Policy Research Unit* na Uniwersytecie Sussex zrealizował przy współpracy Loeta Leydesdorffa oraz Alana L. Portera (współtwórcy podejścia *tech mining* i programu *VantagePoint*) przedsięwzięcie analityczne, którego ogólnodostępne wyniki pozwalają na opracowanie własnych analiz dla kraju, instytucji lub obszaru badawczego poprzez rzutowanie wybranego podzbioru danych bibliometrycznych na globalną mapę nauki, odzwierciedlającą istniejące powiązania między podobszarami badań. Szczegółowy sposób wygenerowania mapy nauki światowej został opisany w artykułach naukowych (Leydesdorff, Rafols 2009; Porter, Rafols 2009). Mapa opiera się na próbkach wszystkich publikacji naukowych zawartych w bazie *Web of Science* autorstwa naukowców zatrudnionych w Stanach Zjednoczonych (Porter, Rafols 2009: 724). Wyniki projektu stanowi "science overlay map", możliwa do wykorzystania we własnych analizach i udostępniona na stronach www.idr.gatech.edu oraz www.leydesdorff.net/overlaytoolkit.

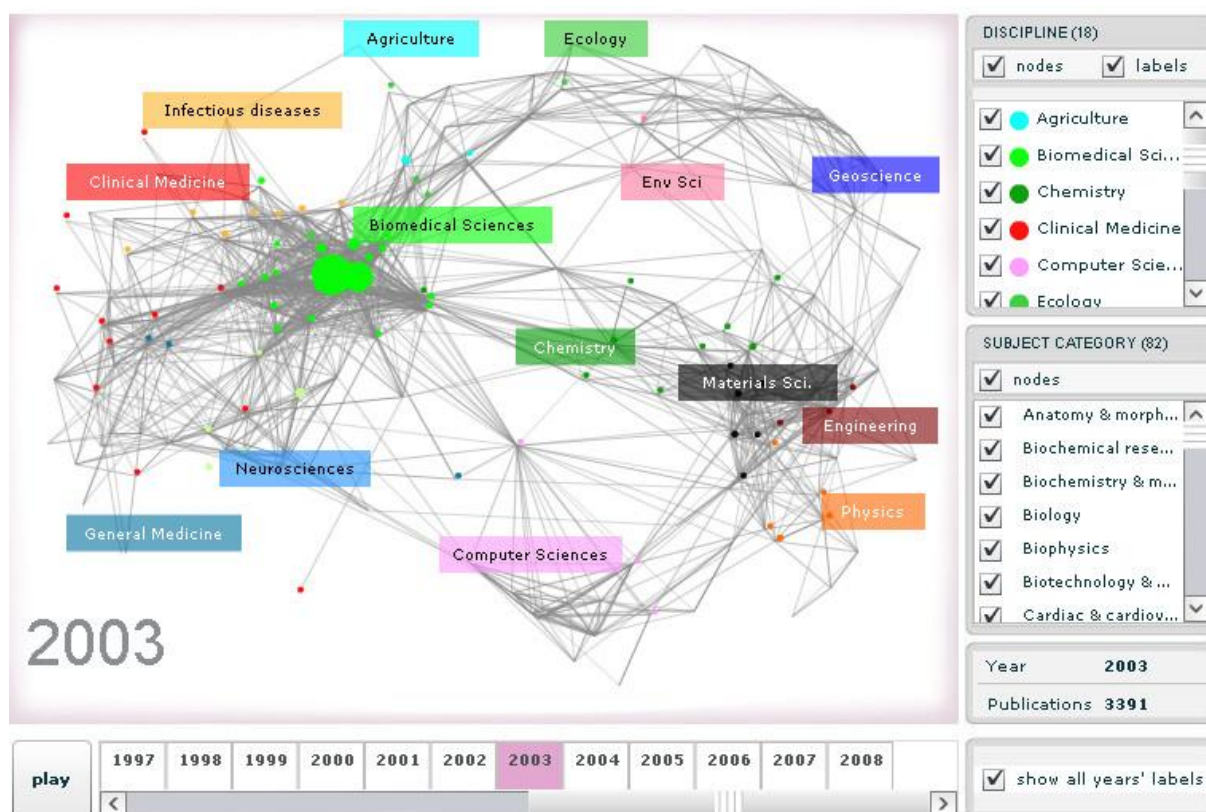
Szczegółowy sposób prowadzenia analiz przy wykorzystaniu udostępnionych w Internecie narzędzi opisują Rafols, Porter i Leydesdorff (2010). Serwis internetowy pozwala nawet użytkownikom przeprowadzić "rzut na mapę" pobranego z bazy *Web of Science* zbioru danych bezpośrednio na serwerze dostawcy strony, bez potrzeby instalacji oprogramowania na własnych komputerach. Analiza umożliwia zrozumienie interdyscyplinarności

prowadzonych badań oraz posiadanych lub niezbędnych kompetencji i wiedzy. Rysunek 47 przedstawia przykład analiz - rzut na mapę nauki dorobku Uniwersytetu w Amsterdamie, pozwalający na obserwację obszarów badań i szczególnych zainteresowań badaczy pracujących na uczelni. Mapa pozwala na wyróżnianie wybranych obszarów badań w oparciu o kilkanaście dyscyplin naukowych oraz ponad 200 szczegółowych podkategorii badawczych. Po najechnięciu wskaźnikiem myszy komputerowej na którykolwiek z przedstawionych na mapie węzłów, wyświetlane są dodatkowe informacje dotyczące podzbioru publikacji w tym obszarze.



Rysunek 47: Dorobek badawczy Uniwersytetu w Amsterdamie zaprezentowany na globalnej mapie wiedzy. Źródło: www.idr.gatech.edu/detail.php?tab=2&id=3

W zbliżony sposób można zdobyć wiedzę na temat wybranego obszaru badawczego i jego interdyscyplinarności. Rysunek 48 prezentuje dorobek światowych badań nad białkiem kinezyną. Badacz zainteresowany podjęciem studiów nad kinezyną może dzięki zaprezentowanej mapie zrozumieć, jakie obszary badawcze i dyscypliny naukowej współtworzą dorobek tych badań na przestrzeni lat. Możliwa jest obserwacja zmian zainteresowań badaczy (np. zmiany dominujących pozycji badań z obszarów biologii komórki i biochemii), co dodatkowo ułatwia "suwak", prezentujący kolejne lata analiz. Mapowanie obszarów badawczych pozwala na realistyczną ocenę perspektyw powodzenia nowych przedsięwzięć badawczych - uświadamia, w których obszarach badań zebrano już "masę krytyczną" i jakie dodatkowe obszary wiedzy powinny być uwzględniane przy planowaniu składu zespołu badawczego.




Rysunek 48: Badania nad kinezyną zaprezentowane na globalnej mapie wiedzy. Źródło: www.idr.gatech.edu/detail.php?show=dyn&d=1

Wykorzystanie bezpłatnie dostępnego i prostego w obsłudze programu komputerowego lub serwisu internetowego, powstałych w projekcie Rafolsa, jest najbardziej efektywnym sposobem stworzenia mapy badań na tle nauki światowej. Bardziej rozbudowane możliwości oferuje płatny serwis firmy *Elsevier*, jednak do większości zastosowań analitycznych wystarczy wykorzystanie prostego narzędzia, które dodatkowo pozwala na samodzielny wybór analizowanego podzbioru publikacji i wykonania dowolnej liczby analiz.

Oprócz omawianych w niniejszym podrozdziale technik analitycznych, prostym i jednocześnie kosztownym sposobem na poznanie aktualnych trendów jest wykorzystanie gotowych analiz, opracowywanych przez wydawców baz. Firma *Thomson Reuters* oferuje produkt informacyjny pod nazwą *Essential Science Indicators*, obejmujący zestawienia najważniejszych publikacji naukowych (ang. *research fronts*) i istotnych zagadnień dla 22 obszarów badań, aktualizowane co 2 miesiące i dostępne jako zestawienia danych dla okresu 10 lat. Szczegółowy opis metod, stosowanych przez firmę *Thomson Reuters* do wyodrębniania *research fronts* opisują Upham i Small (2010). Analogiczne metody stosują Ohniwa, Hibino i Takeyasu (2010: 114) do analizy bazy *MEDLINE*.

Thomson Reuters oferuje też bezpłatne zestawienia analityczne dla wybranych obszarów badawczych w ramach serwisu internetowego <http://sciencewatch.com> (w sekcji "*Data & Rankings*"). Ich lektura może być przydatna dla osób odpowiadających za wspieranie polityki naukowej oraz dla badaczy w poszczególnych dyscyplinach, jednak poziom ogólności może okazać się niewystarczający do wielu zastosowań. Pod hasłem "*Top Topics*" na stronie

publikowane są co dwa miesiące zestawienia dominujących zagadnień w poszczególnych obszarach badawczych. Rysunek 49 oferuje przykładowe zestawienie zagadnień, pochodzące z raportu z lutego 2011 r., a Rysunek 50 – mapę światowej wiedzy, wygenerowaną w maju 2010 r.

Field	Research Front Name
Agricultural Sciences	 TOTAL PHENOLIC CONTENTS; ALGERIAN MEDICINAL PLANTS EXTRACTS CONTAINING PHENOLIC COMPOUNDS; PHENOLIC COMPOUNDS MEASURED; ANTIOXIDANT ACTIVITY; ANTIOXIDANT CAPACITY Research Front Map: PLANT EXTRACTS WITH ANTIOXIDANTS
Biology & Biochemistry	ABNORMAL SERUM VITAMIN D; HEMODIALYSIS PATIENTS RECEIVING DIFFERENT VITAMIN D ANALOGS; CHRONIC KIDNEY DISEASE; SERUM FIBROBLAST GROWTH FACTOR (FGF)-23; VITAMIN D LEVELS
Chemistry	MELT MIXED POLYCARBONATE-MULTIWALLED CARBON NANOTUBE COMPOSITES; MELT PROCESSED POLYCARBONATE-MULTIWALLED CARBON NANOTUBE COMPOSITES; POLYETHYLENE MULTIWALLED CARBON NANOTUBE COMPOSITES; CARBON NANOTUBE EPOXY COMPOSITES
Clinical Medicine	AFRICA'S NEGLECTED TROPICAL DISEASES; GREAT NEGLECTED TROPICAL DISEASES; HUMAN HELMINTH INFECTIONS; NEGLECTED DISEASES; SOIL-TRANSMITTED HELMINTH INFECTIONS
Computer Science	ORDERED CONE METRIC SPACES; COMMON FIXED POINT THEOREMS; PARTIALLY ORDERED METRIC SPACES; ASSAD-KIRK-TYPE FIXED POINT THEOREMS; COUPLED FIXED POINT THEOREMS

Rysunek 49: Przykładowa lista popularnych zagadnień badawczych „Top Topics”, oferowana przez serwis *ScienceWatch*. Źródło: <http://sciencewatch.com/dr/tt/2011/11-febtt>

innowacji. We wcześniejszych podrozdziałach omówione zostały **wskaźniki produktywności**, oparte na liczbie publikacji lub patentów oraz **wskaźniki wpływu**, odwołujące się do cytowań. Obserwacja badań naukowych i rozwoju technologii może również wiązać się z wykorzystaniem **wskaźników konkurencyjności** i **wskaźników specjalizacji**, pozwalających na porównania dorobku krajów, instytucji badawczych lub naukowców.

Najczęściej wykorzystywanym wskaźnikiem konkurencyjności jest **indeks ujawnionej przewagi technologicznej RTA** (ang. *revealed technological advantage*).

Pozwala on na bezpośrednie porównania poprzez identyfikację przewag i słabości w wybranych obszarach. Indeks RTA opiera się na zależnościach, zaobserwowanych w teorii przewag komparatywnych w handlu międzynarodowym, która rekomenduje koncentrację krajów-partnerów wymiany na obszarach, stanowiących ich względne przewagi w porównaniu z produkcją światową. W przypadku badań naukowych i rozwoju technologii, analizie podlegają publikacje lub patenty. Wzór pozwalający na obliczanie RTA to:

$$RTA_{it} = \frac{(P_{it} / \sum_t P_{it})}{(\sum_i P_{it} / \sum_i \sum_t P_{it})} ,$$

gdzie: P_{it} - liczba publikacji lub patentów w obszarze i stworzonych w kraju t .

Wartość RTA powyżej 1,0 oznacza istnienie względnej przewagi – z kolei wartość indeksu poniżej jedności odzwierciedla słabość analizowanego podmiotu w danym obszarze. Sposoby obliczania RTA oraz przykładowe zastosowania opisują Nesta i Patel (2004: 536-537) oraz Klincewicz (2008: 184-195). Alternatywny wobec RTA sposób obliczania wskaźników konkurencyjności i koncentracji oferują m.in. Brusoni i Geuna (2004). W niektórych publikacjach RTA określone jest jako indeks aktywności (ang. *activity index*) (Chen, Guan 2011: 238).

RTA okazuje się szczególnie przydatnym sposobem analiz dlatego, że proste zestawienia liczby publikacji lub patentów nie uwzględniają zróżnicowanych tendencji do korzystania z ochrony patentowej w różnych krajach lub produktywności badaczy i wynalazców. Względna konkurencyjność uzależniona jest od skali prowadzonej działalności B+R, a normalizacja wyników dla różnych krajów lub ośrodków badawczych pozwoli na bezpośrednie porównania i odzwierciedlenie różnic niezależnie od skali produkcji naukowej.

W odniesieniu do nowych obszarów badawczych, przydatny może okazać się inny wskaźnik: **FGSI** (ang. *fast growing specialization index*), odzwierciedlający udział danego kraju (instytucji) w wybranym podobszarze technologii, podzielony przez udział tego samego kraju (instytucji) w ilościowych osiągnięciach dla całego, szeroko zdefiniowanego obszaru badań lub technologii (Nesta, Patel 2004: 537). FGSI stanowi analogię dla RTA, jednak zamiast odwołania się do wszystkich publikacji lub patentów, jako punkt odniesienia przy obliczaniu względnych przewag wykorzystuje łączny dorobek dyscypliny nauki lub obszaru techniki.

$$FGSI_{it} = \frac{(F_{it} / \sum_t F_{it})}{(P_{it} / \sum_t P_{it})},$$

gdzie: P_{it} - liczba publikacji lub patentów w obszarze t stworzonych w kraju j ;

F_{it} - liczba publikacji lub patentów w podobszarze t stworzonych w kraju j .

Kolejnym wskaźnikiem, ułatwiającym porównania międzynarodowe i międzyorganizacyjne, jest **indeks *RII*** (ang. *relative impact index*), odwołujący się do cytowań publikacji lub patentów (Nesta, Patel 2004: 538). Również on stanowi analogię dla indeksu *RTA*, jednak ze względu na konieczność zgromadzenia danych o cytowaniach, analizy *RII* są trudniejsze i bardziej czasochłonne niż obliczanie opisanych wcześniej wskaźników *RTA* i *FGSI*, odwołujących się jedynie do danych o produktywności.

$$RII_{it} = \frac{(C_{it} / \sum_t C_{it})}{(P_{it} / \sum_t P_{it})},$$

gdzie: C_{it} - liczba cytowań publikacji lub patentów w obszarze t stworzonych w kraju j ;

P_{it} - liczba publikacji lub patentów w obszarze t stworzonych w kraju j .

Możliwa jest też agregacja wyników analiz różnych obszarów nauki i technologii dla poszczególnych krajów lub instytucji - obliczenie stopnia specjalizacji (koncentracji) prac badawczych, oznaczanego symbolem *CV*.

$$CV_i = \frac{\sigma_{RTA_{it}}}{\mu_{RTA_{it}}},$$

gdzie: $\sigma_{RTA_{it}}$ – odchylenie standardowe wartości wskaźnika *RTA* dla obszaru t i kraju j ;

$\mu_{RTA_{it}}$ - średnia arytmetyczna wartości wskaźnika *RTA* dla obszaru t i kraju j .

Wysoka wartość CV_i oznacza koncentrację kraju lub instytucji na wąskich obszarach badań, podczas gdy wartość niska odzwierciedla szeroki rozwój badań w wielu obszarach, bez dodatkowej specjalizacji (Nesta, Patel 2004: 538). Obliczenie wartości CV_i wymaga oczywiście wcześniejszej kalkulacji indeksów *RTA* dla wszystkich zidentyfikowanych obszarów badawczych, którymi zajmują się porównywane kraje lub instytucje.

Możliwy jest też alternatywny sposób obliczania wskaźnika *CV* poprzez analizę zmienności względem t (obszaru badawczego), co pozwala na odzwierciedlenie geograficznej koncentracji badań. Wysoka wartość CV_i odpowiadać będzie przypadkom, w których badania nad określoną technologią lub podobszarem nauki prowadzone są w niewielu krajach lub instytucjach, podczas gdy niskie CV_i reprezentuje badania prowadzone powszechnie w różnych miejscach (Nesta, Patel 2004: 538-539).

Innym sposobem przetworzenia danych będzie ilościowa ocena koncentracji w określonym obszarze badawczym - wskaźnik bibliometryczny, pozwalający określić stopień zróżnicowania badań i potencjał dla poruszania oryginalnych tematów badawczych (Yoon, Lee, Lee 2010: 810-811). Jak się okazuje, większość obszarów badawczych przypomina

"oligopole" ze względu na dominację maksymalnie kilku wiodących tematów badawczych (Yoon, Lee, Lee 2010: 814).

Analiza wskaźników bibliometrycznych pozwala również na dodatkową charakterystykę specjalizacji badawczej w wymiarze czasowym. Brusoni i Geuna (2004: 744) proponują porównania instytucji badawczych w oparciu o:

- liczbę podobszarów badań, w których dana instytucja jest aktywna,
- liczbę nowych podobszarów, w których instytucja rozpoczęła publikowanie lub patentowanie w analizowanym okresie (wymaga porównania z okresem wcześniejszym),
- liczbę podobszarów, z których instytucja wycofała się (wymaga porównania z okresem wcześniejszym).

Obliczanie wskaźników jest dogodnym sposobem porównania dokonań badawczych krajów, regionów lub instytucji. Dzięki wykorzystaniu prostych wskaźników, pozwala ocenić obecną sytuację badawczą oraz zidentyfikować najbardziej obiecujące nisze badawcze. W wielu przypadkach zestawienia wyników analiz wskaźnikowych są bardziej przejrzyste niż graficzne mapy. Tabela 13 prezentuje przykład analiz RTA i CV_t , przeprowadzonych dla poszczególnych obszarów technologicznych różnych krajów OECD. Wskaźniki RTA pozwalają na identyfikację krajów-ważnych graczy na scenie innowacji – przykładowo, Francja okazuje się liderem w badaniach nad technologiami nuklearnymi i lotnictwem, Wielka Brytania jest relatywnie silna w obszarze inżynierii chemicznej i procesowej, a Niemcy – w chemii organicznej. Jednocześnie łatwo można dostrzec słabości badawcze niektórych krajów, np. szczególnie niski poziom wskaźnika RTA dla Japonii w odniesieniu do badań nad technologiami, wykorzystywanymi w lotnictwie. Z kolei kalkulacja CV_t pozwala na ocenę geograficznego rozproszenia badań w każdym z obszarów. Badania nad technologiami nuklearnymi są skoncentrowane w jednym lub kilku krajach (wysoka wartość CV_t), podczas gdy inżynieria chemiczna i procesowa interesuje w zbliżonym stopniu naukowców i technologów z wielu krajów (niski poziom CV_t).

Obszar technologiczny	Niemcy	Francja	Wielka Brytania	Japonia	...	CV _t x 100
Chemia nieorganiczna	1,39	2,04	0,92	0,70	...	74,3
Chemia organiczna	1,83	1,10	1,05	0,85	...	56,0
Inżynieria chemiczna i procesowa	1,25	1,25	1,43	0,86	...	38,4
Farmacja i bioinżynieria	1,06	1,93	2,28	0,57	...	59,2
Technologie nuklearne	1,32	4,49	0,64	0,72	...	180,1
Lotnictwo	1,71	4,99	2,84	0,15	...	122,7
...

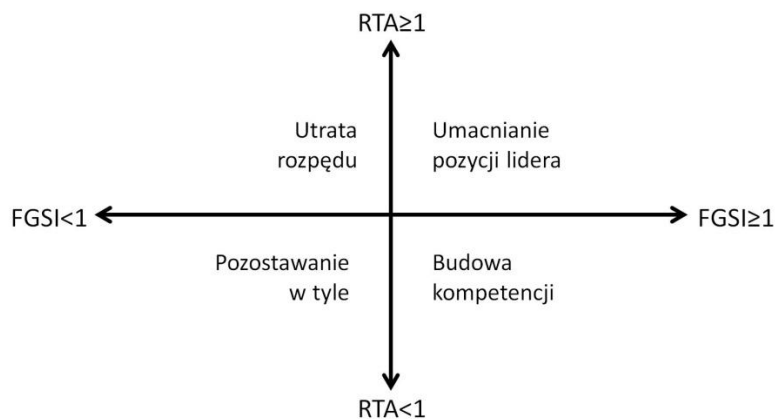
Tabela 13: Wskaźniki RTA obliczone w oparciu o patenty w poszczególnych obszarach technologicznych dla wybranych krajów OECD dla lat 1991-2000. Źródło: Nesta, Patel (2004: 539).

Kolejnym popularnym wskaźnikiem jest tzw. *affinity index* (angielskie słowo *affinity* oznacza: podobieństwo, pokrewieństwo, powinowactwo, koligację). Indeks mierzy intensywność współpracy pomiędzy dwoma graczami (autorami, instytucjami, krajami) w określonym obszarze badań i przedziale czasowym (Okubo 1997: 28).

$$\text{Affinity index} = \frac{\text{COP (A+B)}}{\text{COP (A+WD)}} \times 100,$$

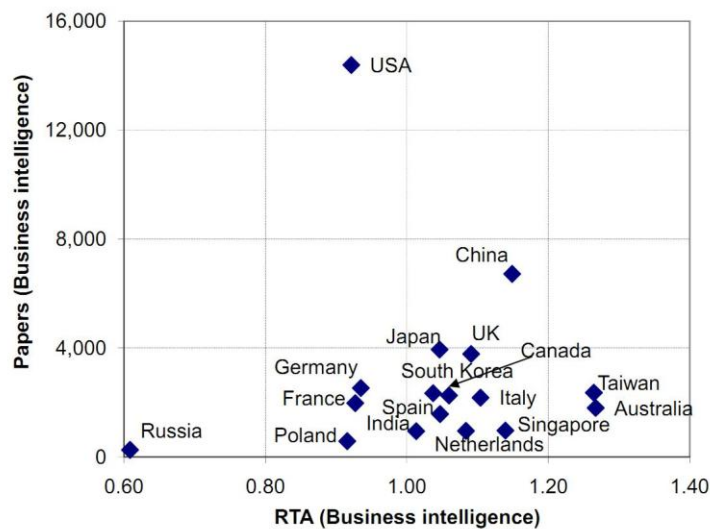
gdzie: COP (A+B) - liczba publikacji lub patentów współtworzonych przez podmioty A i B;
COP (A+WD) - liczba publikacji lub patentów współtworzonych przez A i wszystkie inne podmioty.

Wartości poszczególnych wskaźników można również odzwierciedlić na wykresie. Interesującą formą porównania dorobku badawczego jest dwuwymiarowy wykres, zestawiający ogólny dorobek w szeroko zdefiniowanej dyscyplinie naukowej (RTA) z osiągnięciami w wybranym podobszarze badawczym, przykładowo odzwierciedlającym nowy nurt badań lub popularną technologię (FGSI). Takie zestawienie prezentuje Rysunek 51. Wysoka wartość indeksu RTA przy jednoczesnym niskim wyniku FGSI oznacza stopniową utratę przez kraj lub instytucję roli, odgrywanej w danym obszarze badawczym, mimo znaczącego dorobku w analizowanej dyscyplinie (np. dużej liczby publikacji naukowych). Wysokie FGSI przy niskim RTA oznacza z kolei dążenia do uzyskania znaczącej pozycji w dziedzinie badań, choć na razie osiągnięcia danego kraju lub instytucji ograniczają się do wybranego, dynamicznie rozwijającego się podobszaru.



Rysunek 51: Mapa przewag technologicznych krajów lub instytucji. Źródło: Nesta, Patel (2004: 244).

Rysunek 52 prezentuje przykład zestawienia wartości indeksu *RTA* dla porównywanych krajów.



Rysunek 52: Zestawienie RTA i liczby publikacji dla badań nad *business intelligence*. Źródło: Klincewicz, Miyazaki (2011: 181).

10. Bazy wykorzystywane do analiz bibliometrycznych

10.1. Wprowadzenie

Poniższy rozdział zawiera charakterystykę baz, wykorzystywanych do analiz bibliometrycznych – kolekcji streszczeń publikacji naukowych oraz baz patentów. W odniesieniu do każdej z tych grup, początkowo przedstawione zostaną przeglądy najpopularniejszych baz, a następnie szczegółowej analizie zostaną poddane bazy, które wydają się szczególnie przydatne w typowych projektach bibliometrycznych.

Rozdział zawierać będzie szczegółowe omówienie kilku wybranych baz streszczeń publikacji naukowych oraz ogólnodostępnych baz patentów. Dodatkowo, przedstawione zostaną inne, mniej popularne bazy, w tym komercyjne bazy patentowe, których funkcjonalność zostanie porównana z bazami bezpłatnymi. Warto pamiętać, że analizy bibliometryczne mogą opierać się również na bazach, zawierających dane o przekazywanych grantach i realizowanych projektach badawczych – przykładami mogą być bazy National Science Foundation (www.nsf.gov/awardsearch), National Institute of Health (baza CRISPER, www.sunshine-project.org/crisper), Komisji Europejskiej (baza CORDIS, <http://cordis.europa.eu/search/index.cfm?fuseaction=proj.advSearch>). Przydatna może również okazać się baza danych patentowych, stworzona przez *National Bureau of Economic Research* w oparciu o zasoby *USPTO* i bazę *Compustat*, dostępna do ściągnięcia na własny dysk i prowadzenia kompleksowych analiz cytowań dla lat 1976-2006 (www.nber.org/patents/, <https://sites.google.com/site/patentdataproyect/Home>). Polska platforma bazodanowa Index Copernicus może być dodatkowym źródłem danych do analiz bibliometrycznych, ze szczególnym uwzględnieniem nauk medycznych (www.indexcopernicus.com).

10.2. Przegląd baz bibliograficznych

10.2.1. Chemical Abstracts Service/ SciFinder (CAS)

Baza CAS, prowadzona przez Amerykańskie Towarzystwo Chemiczne (*American Chemical Society*) jest bazą bibliograficzną abstraktów dziedzinowych, obejmującą chemię i dyscypliny pokrewne (nauki biomedyczne, inżynierię, materiałoznawstwo, nauki rolnicze). Jeżeli chodzi o formę dostępnych danych pochodzą one zarówno z literatury (artykuły z czasopism naukowych w wymiarze ponad 10000 tytułów) i patentów (z 61 państw) (CAS 2011a), a także ze źródeł internetowych (CAS 2011b). CAS oferuje indeksowanie za pomocą CAS *REGISTRY*, taksonomii substancji chemicznych (organicznych i nieorganicznych) polegającą na numerycznym identyfikatorze substancji (np. kofeina ma identyfikator numeryczny 58-08-2) (CAS 2011c). W szczególności CAS *REGISTRY* zawiera następujące kategorie substancji chemicznych (klasyfikacja zawiera ponad 56 milionów substancji i jest uzupełniana o około 12000 nowych substancji dziennie) (CAS 2011c):

- Związki organiczne i nieorganiczne,
- Metale i ich stopy oraz metale organiczne,
- Minerale,
- Izotopy,

- Cząstki jądrowe,
- Białka i kwasy nukleinowe,
- Polimery.

Wyszukiwanie w bazie jest możliwe poprzez kryterium tematu, nazwiska autora (możliwe jest zaznaczenie alternatywnej pisowni nazwiska), nazwę organizacji, czasopismo (uwzględnione są skróty tytułów), patent (nazwisko wynalazcy, nr patentu, nazwę organizacji, datę publikacji), strukturę chemiczną substancji (można narysować wzór chemiczny w dostępnym edytorze) lub jej część (sub-strukturę) a także struktury podobne (punktowo określając podobieństwo), formułę molekularną substancji (np. H_2SO_4), nazwę substancji lub jej identyfikator z rejestru CAS, reakcje chemiczne (wykorzystując do tego dostępny edytor reakcji umożliwiający ich dokładne rozpisanie). Możliwe jest przeglądanie, analizowanie, kategoryzowanie otrzymanych wyników poszukiwań. Są linki dostępu do pełnych tekstów. Można przeglądać dokumenty zacytowane i cytowania. Stosując operatory logiczne można wykonywać poszukiwania kombinowane. Użytkownik może komentować znalezione informacje a także dzielić się nimi z innymi użytkownikami a także stosować swoje tagi (maksymalnie 50 na jednego użytkownika). Dostępna jest także opcja drukowania, zapisywania wyników (maksymalnie 20.000) oraz eksport w różnych formatach plików (*ris*, *akx*, *pdf*, *rtf*, *txt*, *xls*) przy czym można eksportować jednorazowo do 500 wyników skróconych i do 100 kompleksowych. Jeżeli chodzi o pracę z wynikami dotyczącymi reakcji chemicznych i substancji chemicznych można analogicznie uzyskane wyniki przeglądać i analizować, drukować, zapisywać i eksportować. Dodatkowo jeżeli chodzi o substancje chemiczne można uzyskać źródło (producentów) i przeglądać właściwości chemiczne (CAS 2011d).

Ujednolicając porównanie baz bibliograficznych można skupić się na następujących kryteriach: szacunkowa liczba abstraktów, rodzaj źródeł, reprezentowane dyscypliny naukowe, indeksowanie.

Kryterium	Baza bibliograficzna: CAS
Szacunkowa liczba abstraktów	Ponad 33 mln abstraktów i patentów, ponad 56 mln substancji chemicznych, ponad 43 mln reakcji chemicznych.
Rodzaj źródeł	Czasopisma naukowe, patenty, materiały konferencyjne, raporty techniczne, książki, rozprawy, recenzje, czasopisma elektroniczne, zasoby Internetu.
Reprezentowane dyscypliny naukowe	Chemia i dziedziny pokrewne m.in.: proteomika, genomika, biochemia, genetyka biochemiczna, chemia organiczna, makromolekularna, stosowana, fizyczna, nieorganiczna, analityczna.
Indeksowanie	Dostępne indeksowanie według rejestrów substancji chemicznych CAS.

Tabela 14: Podsumowanie informacji na temat bazy CAS. Źródło: opracowanie własne.

10.2.2. Compendex

Compendex jest inżynierską bazą bibliograficzną obejmującą 190 dyscyplin inżynierskich. W bazie jest wykorzystywane indeksowanie inżynierskie (*Engineering Index Thesaurus*). Główne obszary tematyczne zawierają inżynierię chemiczną, budownictwo lądowe i wodne, elektrotechnikę, mechanikę i górnictwo. W bazie zgromadzone są dane od roku 1970, jednak dostępne są także dane archiwalne z lat 1884 – 1969 w ramach bazy *Engineering Index Backfile* (EI 2011a). *Compendex* umożliwia także powiadomienia alertowe (przez e – mail) oraz personalizowanie profili użytkowników ułatwiające ich wzajemną współpracę i zarządzanie wyszukiwaniem (EI 2011b). Prócz abstraktów jest możliwość dostępu do pełnego tekstu artykułu. *Compendex* zawiera materiały z czasopism naukowych, materiałów konferencyjnych i raportów technicznych.

Kryterium	Baza bibliograficzna: <i>Compendex</i>
Szacunkowa liczba abstraktów	12 milionów rekordów (od 1970). Dodatkowo ponad 1,7 mln abstraktów archiwalnych (1884 - 1969)
Rodzaj źródeł	Czasopisma naukowe, materiały konferencyjne, raporty techniczne.
Reprezentowane dyscypliny naukowe	190 dyscyplin inżynierskich: inżynieria chemiczna, budownictwo lądowe i wodne, elektrotechnika, mechanika i górnictwo
Indeksowanie	Klasyfikacja <i>Engineering Index Thesaurus</i>

Tabela 15: Podsumowanie informacji na temat bazy *Compendex*. Źródło: opracowanie własne.

10.2.3.ACM Digital Library

Baza ACM zawiera pełne teksty artykułów oraz dane bibliograficzne z zakresu technologii informacyjnej i informatyki. Zawartość bazy stanowią publikacje ACM (*Association for Computing Machinery* – Stowarzyszenie Maszyn Komputerowych), w szczególności czasopisma, książki, materiały konferencyjne, magazyny, *newsletters* i wydawnictwa multimedialne (Librarians ACM 2011). Dodatkowo dostępne są także materiały specjalnych grup tematycznych i wywiady, można także znaleźć propozycje konferencji tematycznych (nadchodzących i tych, które się już odbyły). Można wyszukiwać informacje w oparciu o typ publikacji, afiliację, wydawcę oraz klasyfikację ACM CCS (*ACM Computing Classification System*). ACM CCS obejmuje 11 dziedzin (Portal ACM 2011):

- Literatura ogólna,
- *Hardware*,
- Organizacja systemów komputerowych,
- *Software*,
- Dane,
- Teoria obliczeń,
- Matematyka informatyczna,
- Systemy informacyjne,
- Metody obliczeniowe,
- Aplikacje komputerowe,
- Środowiska obliczeń.

W wyszukiwanie można także rozszerzyć o słowa kluczowe autora, daty publikacji, sponsora i miejsce konferencji. Wyniki poszukiwań można przesyłać e-mailem, zapisywać na swoim koncie w bazie a także eksportować (dostępne formaty to *BibTex*, *EndNote* i *ACMRef*). Dodatkową opcją jest możliwość dzielenia się wynikami przez konta np. w portalu *Google* czy *Facebook*. Pełne teksty artykułów są dostępne w formacie pdf natomiast materiały multimedialne w formacie *Windows Media Player* lub *Real Player*.

Kryterium	Baza bibliograficzna: <i>ACM DL</i>
Szacunkowa liczba abstraktów	990647 abstraktów, 1639151 rekordów.
Rodzaj źródeł	Czasopisma, książki, materiały konferencyjne, magazyny, newslettery i wydawnictwa multimedialne, grupy tematyczne.
Reprezentowane dyscypliny naukowe	Technologia informacyjna i informatyka
Indeksowanie	Klasyfikacja <i>ACM CCS (ACM Computing Classification System)</i>

Tabela 16: Podsumowanie informacji na temat bazy *ACM DL*. Źródło: opracowanie własne.

10.2.4.Scirus

Wyszukiwarka *Scirus* zawiera dane naukowe, techniczne i medyczne zarówno z recenzowanych artykułów, raportów, książek, patentów, firmowych stron internetowych, naukowych stron internetowych, konferencji, recenzji, abstraktów, reprintów i czasopism identyfikowanych przez inne wyszukiwarki z różnych źródeł sieciowych (np. *BioMed Central*, *Medline*, *MIT OpenCourseWeb*, *NASA*, urzędy patentowe). *Scirus* filtruje przeszukiwane strony internetowe i odrzuca te, które zawierają nienaukowe informacje, przez co ogranicza chaos informacyjny przy wyszukiwaniu wieloznacznych terminów. Wyszukiwanie odbywać się może przez słowa w całym dokumencie, tytule (artykułu i czasopisma), nazwisko i afiliację autora czy słowa kluczowe. Można stosować operatory logiczne w procesie poszukiwań danych. Dane są dostępne w kilku formatach plików – *pdf*, *ppt*, *TeX*, *html*, *doc*, *PS* (SCIRUS 2011a). Dostępne jest wyszukiwanie przez wskazanie odpowiedniej dziedziny nauki (SCIRUS 2011a):

- Rolnictwo i nauki biologiczne,
- Astronomia,
- Chemia i inżynieria chemiczna,
- Nauki informatyczne,
- Nauki o ziemi i innych planetach,
- Ekonomia, biznes i zarządzanie,
- Inżynieria, energia i technologia,
- Nauki o środowisku,
- Języki i lingwistyka,
- Prawo,
- Nauki przyrodnicze,
- Materiałoznawstwo,
- Matematyka,
- Medycyna,
- Neurologia,
- Farmakologia,

- Fizyka,
- Psychologia,
- Nauki społeczne,
- Nauki o zachowaniu,
- Socjologia.

Sortowanie wyników przez *Scirus* odbywa się domyślnie względem istotności, można także zastosować kryterium daty. Lista wyników może zostać zapisana, wysłana e-mailem oraz eksportowana. Przy czym są do wyboru dwie opcje, eksport odwołań oraz odwołań, abstraktów i słów kluczowych. Dane mogą być eksportowane w formacie tekstowym lub *RIS* (*SCIRUS* 2011b).

Kryterium	Wyszukiwarka: <i>Scirus</i>
Szacunkowa liczba przeszukiwanych stron	380 milionów stron naukowych i związanych z nauką
Rodzaj źródeł	Recenzowane artykuły, raporty, książki, patenty, firmowe strony internetowe, naukowe strony internetowe, konferencje, recenzje, abstrakty, reprinty
Reprezentowane dyscypliny naukowe	Rolnictwo i nauki biologiczne, astronomia, chemia i inżynieria chemiczna, nauki informatyczne, nauki o ziemi i innych planetach, ekonomia, biznes i zarządzanie, inżynieria, energia i technologia, nauki o środowisku, języki i lingwistyka, prawo, nauki przyrodnicze, materiałoznawstwo, matematyka, medycyna, neurologia, farmakologia, fizyka, psychologia, nauki społeczne, nauki o zachowaniu, socjologia.
Indeksowanie	Klasyfikacja według wyżej wymienionych obszarów nauki

Tabela 17: Podsumowanie informacji na temat wyszukiwarki *Scirus*. Źródło: opracowanie własne.

10.2.5. Google Scholar

Wyszukiwarka *Google Scholar* jest mniej ustrukturyzowana od poprzednio omawianych baz, choć obejmuje szerszy (nie ograniczony dziedzinowo) obszar nauk. Materiały, jakich można szukać w *Google Scholar*, to artykuły i prace naukowe, książki, abstrakty, materiały z wydawnictw naukowych, towarzystw naukowych i repozytoriów materiałów, które zostały zgłoszone do publikacji (*Google Scholar* 2011a). Podobnie jak *Scirus*, *Google Scholar* sortuje

wyniki według istotności, czyli rangi jaką ma dany wynik (artykuł, autor). Możliwe jest także stosowanie operatorów logicznych przy wyszukiwaniu.

Jeżeli chodzi o kryteria wyszukiwania są one mniej ustrukturyzowane niż w *Scirusie*. Można szukać słów i fraz w tytule lub tekście artykułu, nazwiska autora, tytułu źródła i zakresu dat. Nie są wprowadzone żadne deskryptory ułatwiające kategoryzowanie i zawężanie poszukiwań.

Zaletą *Google Scholar* jest możliwość wyszukiwania w innych językach niż angielski, funkcjonuje również polski interfejs użytkownika. Dodatkowo można wprowadzić nazwy do 3 bibliotek, których zasoby będą przeszukiwane. Jeżeli chodzi o eksport danych, *Google Scholar* umożliwia wyświetlanie linków umożliwiających eksport danych z wyszukiwarki do programów *BibTex*, *EndNote*, *RefMan*, *RefWorks* i *WenXianWang*. W wynikach wyszukiwania wyświetlane są prócz liczby cytowań danego źródła także artykuły powiązane oraz możliwość pobrania pełnego tekstu np. w formacie *pdf*.

Kryterium	Wyszukiwarka: <i>Google Scholar</i>
Szacunkowa liczba przeszukiwanych stron	Brak danych (por: Meho, Yang 2007: 2109)
Rodzaj źródeł	Artykuły i prace naukowe, książki, abstrakty, materiały z wydawnictw naukowych, towarzystw naukowych i repozytoriów materiałów
Reprezentowane dyscypliny naukowe	Brak wyspecyfikowanych dziedzin
Indeksowanie	Brak określonych taksonomii

Tabela 18: Podsumowanie informacji na temat wyszukiwarki *Google Scholar*. Źródło: opracowanie własne.

10.2.6. *Web of Science*

Ponieważ w dalszych częściach tekstu baza *Web of Science* zostanie przeanalizowana bardziej szczegółowo, w tym miejscu tytułem wprowadzenia zostanie przedstawione jedynie zestawienie tabelaryczne.

Kryterium	Baza bibliograficzna: <i>Web of Science</i>
Szacunkowa liczba abstraktów	Ponad 46,1 mln rekordów
Rodzaj źródeł	Artykuły, patenty, strony internetowe, sprawozdania pokonferencyjne, materiały ogólnodostępne
Reprezentowane dyscypliny naukowe	Nauki ścisłe, nauki społeczne, nauki humanistyczne i sztuki

Indeksowanie	Prosta taksonomia nauk (przypisanie do niej czasopism, a nie poszczególnych artykułów)
--------------	--

Tabela 19: Podsumowanie informacji na temat bazy *Web of Science*. Źródło: opracowanie własne.

10.2.7.Scopus

Ponieważ w dalszych częściach tekstu baza *Scopus* zostanie przeanalizowana bardziej szczegółowo, w tym miejscu tytułem wprowadzenia zostanie przedstawione jedynie zestawienie tabelaryczne.

Kryterium	Baza bibliograficzna: <i>Scopus</i>
Szacunkowa liczba abstraktów	41 mln rekordów
Rodzaj źródeł	Recenzowana literatura naukowa i badawcza jak i źródła internetowe, książki, materiały konferencyjne, raporty komercyjne (<i>trade publications</i>)
Reprezentowane dyscypliny naukowe	Nauki ścisłe, nauki techniczne, nauki medyczne, nauki społeczne, nauki humanistyczne i sztuki
Indeksowanie	Prosta taksonomia nauk

Tabela 20: Podsumowanie informacji na temat bazy *Scopus*. Źródło: opracowanie własne.

10.2.8.INSPEC

Ponieważ w dalszych częściach tekstu baza *INSPEC* zostanie przeanalizowana bardziej szczegółowo, w tym miejscu tytułem wprowadzenia zostanie przedstawione jedynie zestawienie tabelaryczne.

Kryterium	Baza bibliograficzna: <i>INSPEC</i>
Szacunkowa liczba abstraktów	11 mln rekordów
Rodzaj źródeł	Artykuły z czasopism naukowych, materiałów konferencyjnych, książki, raporty, rozprawy, patenty
Reprezentowane dyscypliny naukowe	Nauki ścisłe i inżynierskie
Indeksowanie	Taksonomia inżynierska zawierająca 5 głównych obszarów dziedzinowych.

	Rozbudowana klasyfikacja obszarów badawczych, indeks chemiczny i astronomiczny oraz tezaurus
--	--

Tabela 21: Podsumowanie informacji na temat bazy INSPEC. Źródło: opracowanie własne.

10.2.9. MEDLINE

Ponieważ w dalszych częściach tekstu baza *MEDLINE* zostanie przeanalizowana bardziej szczegółowo, w tym miejscu tytułem wprowadzenia zostanie jedynie przedstawione zestawienie tabelaryczne.

Kryterium	Baza bibliograficzna: <i>MEDLINE</i>
Szacunkowa liczba abstraktów	18 mln rekordów
Rodzaj źródeł	Artykuły w większości z czasopism naukowych, w mniejszej części z materiałów prasowych (z gazet, magazynów i <i>newsletterów</i>)
Reprezentowane dyscypliny naukowe	Nauki przyrodnicze z naciskiem na biomedycynę i zdrowie
Indeksowanie	Taksonomia <i>MeSH</i> . Wyszukiwanie w oparciu o 12 poziomów porządkowych, ponad 26 tys. deskryptorów, dostępnych 177 tysięcy słów, 199 tysięcy nagłówek uzupełniających z osobnym tezaurem

Tabela 22: Podsumowanie informacji na temat bazy MEDLINE. Źródło: opracowanie własne.

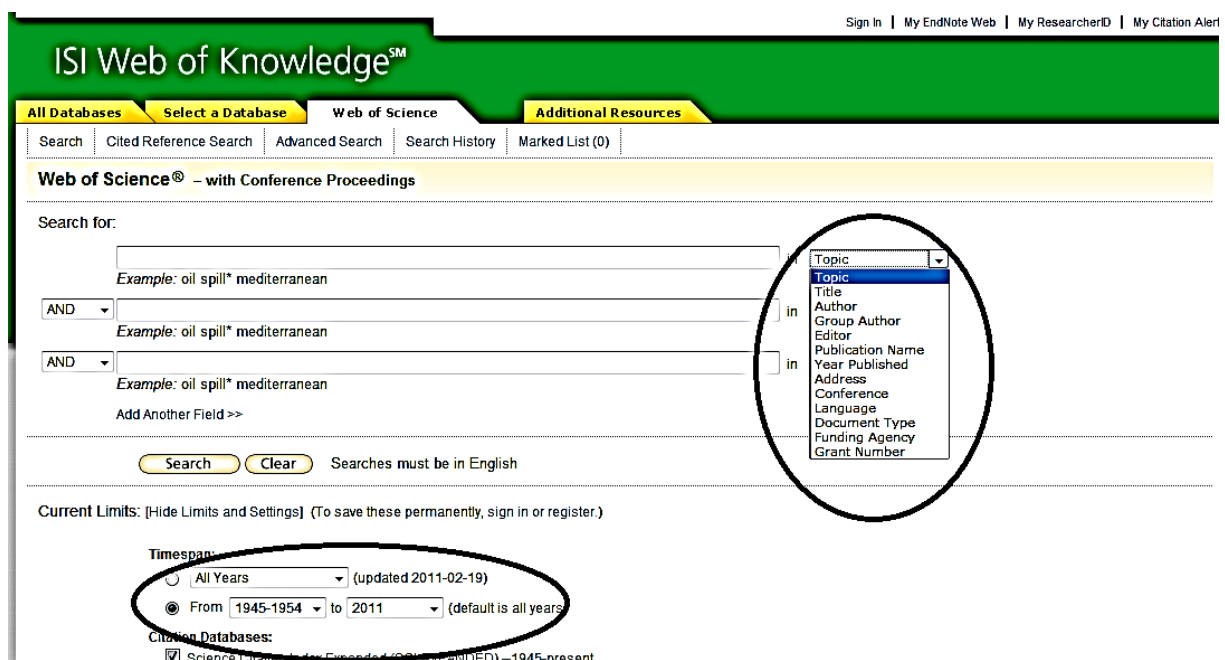
10.3. Opis bazy *Web of Science*

Baza *Web of Science*, utrzymywana przez firmę *Thomson Reuters* (wcześniej: *ISI*, tzw. Instytut Filadelfijski), oferuje dostęp do światowych baz cytowań z zakresu nauk ścisłych, nauk społecznych, nauk humanistycznych i sztuk. Baza zawiera ponad stu letni zbiór danych archiwalnych i cytowań dotyczących artykułów, patentów, stron internetowych, sprawozdań pokonferencyjnych, materiałów ogólnodostępnych. Indeksowanie w bazie odbywa się w obie strony, można uzyskać indeks cytowań odwołujących się do danej pracy jak i cytowań zawartych w danej pracy. W bazie dostępne są następujące narzędzia podstawowe (WoK 2011a):

- Wyszukiwanie cytowanych odwołań między innymi w czasopismach, sprawozdaniach, raportach technicznych, rozprawach, filmach, dziełach sztuki.
- Identyfikacja autora.

- Wyszukiwanie trendów i wzorów także z uwzględnieniem trendów graficznych i mapowania cytowań, co umożliwia identyfikację powiązań między artykułami i ich cytowaniami.
- Kluczowe wskaźniki – Indeks Wpływu (ang. *Impact Factor*) używany do określania wpływu badań naukowych; Indeks h Hirscha (ang. *h – index*) pozwalający na ocenę indywidualnego wpływu badacza; *Times Cited* pokazujący wpływ artykułu poprzez liczbę jego cytowań.
- Ewaluacja czasopism używana do pomiaru wpływu czasopisma.
- Ewaluacja dokonań w ujęciu indywidualnym, instytucjonalnym, regionalnym i narodowym.
- Zintegrowane zarządzanie bibliografią – użytkownicy mają dostęp do sieci *Thomson Reuters EndNote Web*, która pozwala na organizowanie pracy własnej i współpracy z innymi w trybie *online*. Można skorzystać także z aplikacji *Reference Manager*, użytecznej dla grup, sieci i wspólnych projektów.
- Alert Cytowań pozwalający na przesyłanie wiadomości e-mailowej do użytkownika o nowych rekordach dodanych do bazy lub nowych cytowaniach, które zostały uprzednio przez niego wybrane.

Wyszukiwanie artykułu w bazie *Web of Science* odbywać się może przez użycie kilku opcji: tematu, tytułu, autora, nazwy wydawcy, roku wydania, miejsca wydania, języka i typu dokumentu (WoK 2011b).



Rysunek 53: Interfejs użytkownika w bazie *Web of Science*- wyszukiwanie. Źródło: *Web of Science*.

Kryterium	Web of Science
Adres	+
Agencja finansująca	+
Autor	+
Daty	+
Grupa autorów / Autorzy/autor korporacyjny/grupowy	+
Język	+
Konferencja	+
Nazwa publikacji / tytuł źródła	+
Numer grantu	+
Rok opublikowania	+
Temat	+
Typ dokumentu	+
Tytuł	+
Wydawca	+
SUMA:	14

Tabela 23: Kryteria wyszukiwania w bazie Web of Science. Źródło: opracowanie własne.

Dostępne są także bardziej złożone narzędzia i produkty usprawniające i poszerzające zakres bazy Web of Science w ramach platformy Web of Knowledge:

- multidyscyplinarną platformę wyszukiwania umożliwiającą przeszukiwanie wielu baz jednocześnie (*Web of Science, Chinese Science Citation Database, CurrentContents Connect, DerwentInnovations Index*) (WoK 2011c),
- multidyscyplinarną platformę wyszukiwania umożliwiającą przeszukiwanie wielu obszarów naukowych jednocześnie (*BiologicalAbstracts, BiosisCitation Index, BiosisPreviews, CAB Abstracts, FSTA – Food Science and Technology Abstracts, Global Health form CABI, INSPEC, Medline, ZoologicalRecord*) (WoK 2011d),
- analizę zasobów na dwa sposoby. Ilościowy wskaźnik wyników badań i trendów naukowych (*Essential Science Indicators*) i wskaźnik oceny czasopism (*JournalCitationReports*).

Podsumowując charakterystykę bazy Web of Science można stwierdzić (WoK 2011e):

1. Wyszukiwanie odbywa się według tematyki, autora, zespołu autorów, tytułu, roku wydania i miejsca wydania, języka publikacji, rodzaju dokumentu.
2. Tytuły w językach obcych są tłumaczone na język angielski.
3. Wszyscy autorzy są indeksowani z nazwiska i inicjału imienia (maksymalnie może być pięć inicjałów).
4. Indeksowane są także słowa kluczowe wskazane przez autorów oraz słowa i zwroty z tytułu (dodatkowa funkcja *KeyWords Plus*).
5. Wszystkie abstrakty autorów są indeksowane, wyszukać je można przez temat, dostępne jest także przekierowanie na stronę wydawcy i dostęp do pełnego tekstu artykułu.
6. Można przeszukiwać wszystkie cytowane odnośniki (dodatkowo odnośniki zaznaczone kolorem niebieskim są linkami do innych rekordów źródłowych w bazie; odnośniki w kolorze czarnym dotyczą dokumentów nieindeksowanych w bazie, nie objętych subskrypcją użytkownika lub zacytowane błędnie).

7. W przypadku odnośników jedynie pierwszy autor jest indeksowany, wszystkie cytowane prace są indeksowane, tytuł jest pełny, indeksowany jest rok, opcjonalnie tom (maksymalnie cztery znaki) i strony (maksymalnie do pięciu znaków).
8. Można doprecyzować wyszukiwanie i posortować wyniki według daty, liczby cytowań, trafności, roku wydania, tytułu i pierwszego autora.
9. Można przesłać wyniki do programu *MS Excel* i tworzyć własne wykresy.
10. Można tworzyć własne zakładki, zapisywać rekordy w *EndNote Web*, wysyłać e-mailem, zapisywać zakładki tymczasowe (do 500 rekordów) lub stałe w *EndNote Web* (do 10.000 rekordów).
11. Można eksportować dane do *EndNote*, *Reference Manager*, *ProCite*.
12. Można tworzyć graficzne raporty cytatów z artykułów w wynikach wyszukiwania.
13. Użytkownik może spersonalizować swój profil, przeglądać historię wyszukiwania, tworzyć swoje kombinacje wyszukiwań i powiadomienia o zacytowaniu.

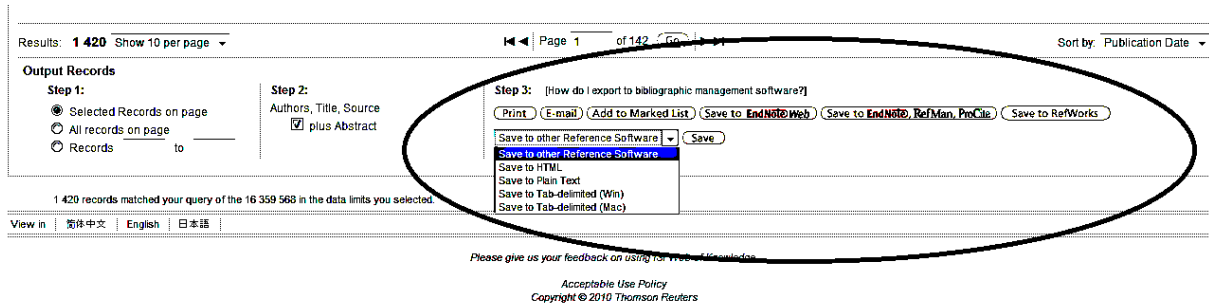
Na szczególną uwagę zasługują ciekawe opcje bazy *Web of Science*, wymienić należy wizualizację, analizę i integrację. Opcja wizualizacji powiązań cytowań umożliwia tworzenie map cytowań i śledzenie cytowań w artykułach (*Citation Map*). Możliwe jest śledzenie w czasie w przód i w tył związków cytowań, wykorzystanie barwnych kodów, rekonfigurowanie i organizowanie mapy w taki sposób, aby pokazywała trendy w cytowaniach (WoK 2011f).

Analiza (*AnalyzeTool*) pozwala na odkrywanie trendów i wzorów (wcześniej niewidocznych) i umożliwia redefiniowanie wyszukiwania z uwzględnieniem filtrów wyników i tworzenie wyizolowanych zakresów zainteresowań. Analiza może być przeprowadzana wielokrotnie z uwzględnieniem różnych kryteriów filtracyjnych (autor, typ dokumentu, nazwa instytucji, język, rok wydania, tytuł, obszar tematyczny).

Integracja natomiast jest możliwa dzięki *Endnote Web*, który spaja szukanie, pisanie i tworzenie bibliografii w jeden proces (WoK 2011g). Jeżeli chodzi o eksport wyników dostępne są następujące opcje zestawione w poniższej tabeli. Istnieje jednak ograniczenie do 500 rekordów jednocześnie, we wszystkich poniższych opcjach.

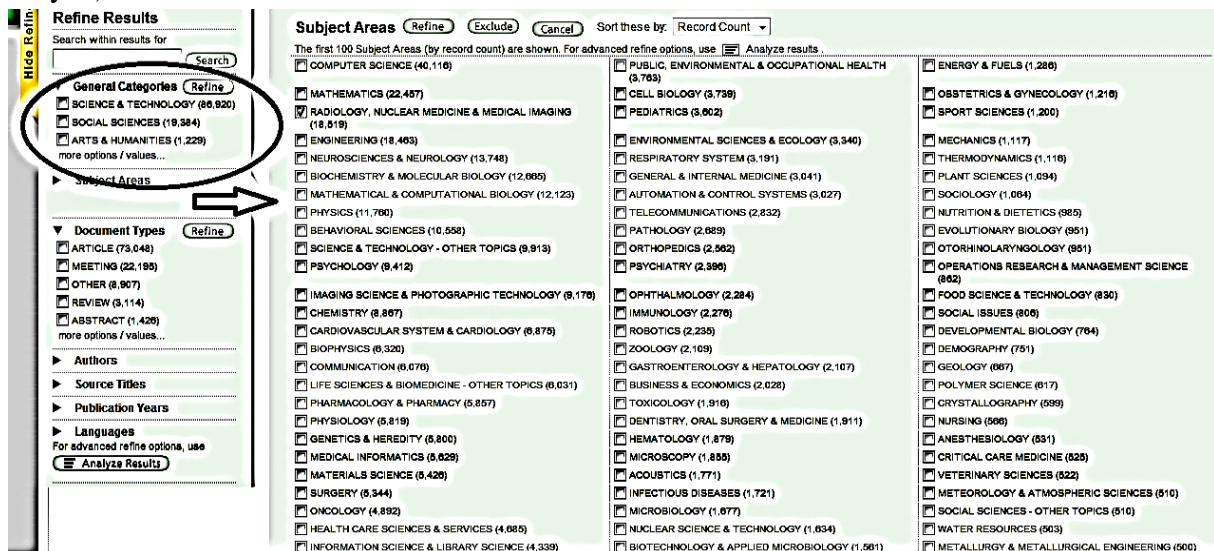
Opcja	<i>Web of Science</i>
Dodanie do listy	+
Przesłanie e-mailem	+
W formacie dla systemu Mac	+
W formacie dla systemu Windows	+
Wydruk	+
Zapis tekstowy	+
Zapis w <i>EndNote</i>	+
Zapis w <i>EndNote Web</i>	+
Zapis w <i>HTML</i>	+
Zapis w <i>ProCite</i>	+
Zapis w <i>RefMan</i>	+
Zapis w <i>RefWorks</i>	+
SUMA:	12

Tabela 24: Opcje eksportu wyników w bazie *Web of Science*. Źródło: opracowanie własne.



Rysunek 54: Interfejs użytkownika w bazie *Web of Science* – opcje eksportowania wyników wyszukiwania. Źródło: *Web of Science*.

Baza *Web of Science* prezentuje także wyniki z podziałem na ogólne kategorie nauk (nauka i technologia, nauki społeczne, nauki humanistyczne i sztuki) a także ze względu na obszary tematyczne (np. matematyka, nauki informatyczne, inżynieria, biologia komórkowa i wiele innych).



Rysunek 55: Interfejs użytkownika w bazie *Web of Science* –ogólne kategorie nauk i obszary tematyczne. Źródło: *Web of Science*.

Główna dziedzina nauki	Hasło po polsku	Hasło po angielsku	Wyniki w bazie: Web of Science
Med, B, Ch	Alergia	<i>allergy</i>	20.465
Med, B, Ch	Złamanie otwarte	<i>“compound fracture”</i>	19
T, F, Mat	System operacyjny	<i>“operating system”</i>	3.644
T, F, Mat	Internet	<i>Internet</i>	67.067
T, F, Mat	Konstrukcja nośna	<i>„supporting structure”</i>	318
T, F, Mat	Wytrzymałość materiałów	<i>„strength of materials”</i>	324
Ch, B	Całkowita zawartość węgla organicznego	<i>„totalorganic carbon content”</i>	156
B, Ch	Sosna śródziemnomorska	<i>„stone pine”</i>	120
B, Ch	Dolina Nilu	<i>„Nile Valley”</i>	111
Ch, B	Destylacja ropy naftowej	<i>“petroleum distillation”</i>	10
E	Bezrobocie	<i>unemployment</i>	6.960
E	Teoria zarządzania	<i>“theory of management”</i>	109
H	Filologia klasyczna	<i>„classical philology”</i>	26
H	Edukacja młodzieży	<i>“education of youth”</i>	55
S	Przedstawienie teatralne	<i>„theater performance”</i>	7
S	Obraz olejny	<i>„oil painting”</i>	53

Tabela 25: Przykładowe wyniki wyszukiwania w *Web of Science* dla kilku haseł z wybranych obszarów nauki (w polu: temat) dla okresu 2005 – 2010. Źródło: opracowanie własne. Legenda: skróty literowe przy hasłach odpowiadają dziedzinom nauki z polskiej klasyfikacji nauk (CK 2005), która na potrzeby niniejszego opracowania została nieznacznie zmodyfikowana:

- B – nauki biologiczne,
- Ch – nauki chemiczne,
- E – nauki ekonomiczne,
- F – nauki fizyczne,
- H – nauki humanistyczne,
- Mat – nauki matematyczne,
- Med – nauki medyczne,
- T – nauki techniczne,
- S – sztuka.

Dodatkowo, po otrzymaniu listy wyników wyszukiwania, możliwa jest ich prosta lecz szybka analiza za pomocą narzędzia *AnalyzeResults*.

198 records. Topic=(oil painting) AND Year Published=(2005 - 2010)

Rank the records by this field:	Analyze:	Set display options:	Sort by:
Language Publication Year Source Title Subject Area	Up to 500 Records.	Show the top 500 Results. Minimum record count (Threshold): 20	<input checked="" type="radio"/> Record count <input type="radio"/> Selected field

Analyze

Use the checkboxes below to view the records. You can choose to view those selected records, or you can exclude them (and view the others).
 Note: The number of records displayed may be greater than the listed Record Count if the original set contained more records than the number of records analyzed.

<input checked="" type="checkbox"/> View Records <input checked="" type="checkbox"/> Exclude Records	Field: Subject Area	Record Count	% of 198	Bar Chart	Save Analysis Data to File
<input type="checkbox"/>	CHEMISTRY	58	29.2929 %		
<input type="checkbox"/>	ART	53	26.7677 %		
<input type="checkbox"/>	SPECTROSCOPY	29	14.6465 %		
<input type="checkbox"/>	PHYSICS	27	13.6364 %		
<input type="checkbox"/>	ARCHITECTURE	23	11.6162 %		
<input type="checkbox"/>	MATERIALS SCIENCE	23	11.6162 %		
<input type="checkbox"/>	ARCHAEOLOGY	20	10.1010 %		
<input type="checkbox"/>	BIOCHEMISTRY & MOLECULAR BIOLOGY	20	10.1010 %		

(88 Subject Area value(s) outside display options.)

Rysunek 56: Interfejs użytkownika w bazie *Web of Science* – narzędzie analizy wyników.
 Źródło: *Web of Science*.

Otrzymane wyniki można zestawić rangowo ze względu na autora, typ dokumentu, kategorie ogólne, język, rok publikacji, źródło i obszar tematyczny (na rysunku powyżej). Maksymalnie analizować można do 500 rekordów. Można określić minimalną liczbę punktów, od której będzie tworzone zestawienie analityczne (na powyższym rysunku ustalono wartość 20). Wyniki analizy można zapisać w zewnętrznym pliku.

10.4. Opis bazy *Scopus*

Baza abstraktów i cytowań *Scopus*, udostępniana przez firmę *Elsevier*, dotyczy zarówno recenzowanej literatury badawczej, jak i źródeł internetowych z obszarów nauk ścisłych, nauk technicznych, nauk medycznych, nauk społecznych oraz nauk humanistycznych i sztuki. Według prowadzących bazę danych, do jej funkcjonalnych zalet należy zaliczyć między innymi (Sciverse 2011a):

- Prosty i intuicyjny w obsłudze interfejs użytkownika.
- Linki do tekstów artykułów (nie tylko streszczeń) a także do zasobów bibliotecznych.
- Identyfikację Autora (*Author Identifier*), która automatycznie zestawia jego prace badawcze (wykorzystywany jest także indeks *h* Hirscha).
- Śledzenie Cytowań (*Citation Tracker*), które ma za zadanie uprościć poszukiwania, sprawdzanie i śledzenie cytowań w czasie rzeczywistym.
- Identyfikację Afiliacji (*Affiliation Identifier*), która w sposób automatyczny identyfikuje i zestawia organizacje i jej wszystkie wyniki badawcze.
- Analizera Czasopism (*Journal Analyzer*), który ma za zadanie dostarczyć szybki wgląd w wyniki odnośnie danego czasopisma
- Alert pozwalający pozostać użytkownikowi na bieżąco (wykorzystuje *RSS* i *HTML*).
- Menedżera Pobierania Dokumentów (*Document Download Manager*), który pozwala w łatwy sposób pobrać i organizować pracę nad wieloma pełno tekstowymi artykułami jednocześnie.

- Współpracę między innymi z *SciVerseScienceDirect*, *Reaxis* oraz *ProQuest's CSA Illumina*.
- Możliwość eksportowania danych przez oprogramowanie do zarządzania bibliograficznego (*RefWorks*, *EndNote* i *BibTeX*).

Charakteryzując bazę *Scopus*, należy zwrócić uwagę między innymi na (SciVerse Scopus 2010):

- Możliwość wyszukiwania podstawowego za pomocą słów kluczowych, autora, afiliacji, ram czasowych, typu dokumentu i obszaru tematycznego.
- Wyniki wyszukiwania obejmują pięć typów źródeł (recenzowanych czasopism, list referencyjnych, sieci internetowej (przez Scirusa), biur patentowych (przez Scirusa), instytucjonalnych repozytoriów i specjalnych zbiorów tematycznych (przez Scirusa)).
- Wyszukane rezultaty można przeglądać i sortować ze względu na datę, istotność dokumentu, autorów, tytuł źródła oraz cytowania.
- Dodatkowo dla każdego znalezionej rezultatu można przejść do abstraktu, przypisów i cytowań, do strony wydawcy dokumentu, do pełnego tekstu (pod warunkiem, że jest dostępny w zasobach bibliotecznych) a także pobrać tekst w formacie „pdf” (do 50 artykułów na raz).
- W szybki sposób można zdefiniować wyszukiwanie za pomocą funkcji „*refine results*”, gdzie można wyszukiwanie zawęzić lub poszerzyć według tytułu, autora, daty, typu dokumentu i obszaru tematycznego.
- Można dodać dodatkowe terminy do szukanego zapytania przez korzystanie z funkcji „*search within results*”.
- Wyszukane informacje można eksportować do programów służących do zarządzania cytowaniami obsługującymi formaty RIS oraz ASCII.
- Można także wyniki wydrukować, przesłać e-mailem lub utworzyć bibliografię.
- Nazwiska autorów są linkowane, co umożliwi szybkie wyszukanie wszystkich publikacji autora, dostępne są także szczegółowe dane o autorze.
- Dwa ostatnie cytowania tekstu przez innych autorów są na monitorze wraz z linkiem do kompletnej listy cytowań. Cytowania w źródłach internetowych i patentowych są dostępne w odpowiedniej zakładce.
- Przy wyszukiwaniu autorów dostępne są także warianty nazwiska, które są zgromadzone w profilu autora. Czasem jednak niektóre dokumenty napisane przez jednego autora nie są zestawiane z powodu braku dostatecznych informacji w bazie. Brak ten uzupełnia specjalna opcja, która wyszukuje wszystkich autorów o wprowadzonym nazwisku bez ich zestawiania. Pozwala to na uniknięcie pominięcia np. autorów o tym samym nazwisku i zmienionym inicjale imienia.
- W bazie dostępne są następujące informacje o autorach: ostatnia podana afiliacja, liczba odwołań w bazie, indeks h Hirscha, liczba wyników z Internetu za pośrednictwem *Scirusa*, wizualizacja wyników autora (*Author Evaluator*), liczba dokumentów w bazie, liczba dokumentów jakie autor zacytował, liczba współautorów, obszar tematyczny autora, można także uzyskać informacje historyczne.
- Śledzenie cytowań generuje wynik w postaci liczby i czasu cytowania danego artykułu, rok po roku od 1996. Możliwe jest sortowanie wyników w myśl tych dwóch

kryteriów w dowolnym porządku. Można także uzyskać samo cytowania danego autora i wyłączyć je z wygenerowanego wyniku. Możliwe jest także wyeksportowanie przeglądu cytowań do pliku w formacie *csv*, jego wydrukowanie lub zapisanie w profilu użytkownika.

- Wykorzystanie indeksu *h* umożliwia ocenę (ewaluację) osiągnięć autora od 1996 za pomocą graficznych wykresów. Jest także możliwość wygenerowania graficznego zestawienia dokumentów autora i jego cytowań w poszczególnych latach.
- Przy wyszukiwaniu afiliacji automatycznie brane są także pod uwagę warianty nazw instytucji, co ujednotacza proces wyszukiwania i generuje listę z liczbą dokumentów, miastem i krajem.
- W ustawieniach profilowych można także ustawić alert o nowych wynikach ze zdefiniowanego zakresu (autora, dokumentu) przesyłany w określonym czasie (dzienny, tygodniowy, miesięczny).

Interfejs użytkownika w bazie *Scopus* jest bogatszy niż w *Web of Science* i wydaje się bardziej czytelny/przyjazny użytkownikowi. Przede wszystkim oferuje jednak więcej funkcji i kryteriów wyszukiwania.

Kryterium	Scopus
Abstrakt	+
Afiliacja	+
Autor	+
<i>CODEN</i>	+
Daty	+
<i>DOI</i>	+
Grupa autorów / Autorzy/autor korporacyjny/grupowy	+
<i>ISSN</i>	+
Język	+
Konferencja	+
Nazwa chemiczna	+
Nazwa publikacji / tytuł źródła	+
Numer CAS	+
Odwołania	+
Pierwszy autor	+
Rok opublikowania	+
Słowa kluczowe	+
Typ dokumentu	+
Tytuł	+
Tytuł/abstrakt/słowa kluczowe	+
SUMA:	20

Tabela 26: Kryteria wyszukiwania w bazie *Scopus*. Źródło: opracowanie własne.

The screenshot shows the Scopus search interface. At the top, there are logos for SciVerse and Scopus, and navigation links for Hub, ScienceDirect, Scopus, SciTopics, and Applications. Below this is a search bar and navigation links for Search, Sources, Analytics, My alerts, My list, and My settings. A notification states: "Scopus will be unavailable due to scheduled maintenance for approximately 12 hours from 1:00PM C". The main search area is titled "Document search" and includes tabs for Author search, Affiliation search, and Advanced search. The search criteria are defined by a search box (with an example: "heart attack" AND stress) and a "Limit to:" section. The "Limit to:" section includes:

- Date Range (inclusive):**
 - Published All years to Present
 - Added to Scopus in the last 7 days
- Subject Areas:**
 - Life Sciences (> 4,300 titles)
 - Health Sciences (> 6,800 titles. 100% Medline coverage)

 A dropdown menu is open, showing search criteria options:

- Article Title, Abstract, Keywords
- All Fields
- Article Title, Abstract, Keywords
- Authors
- First Author
- Source Title
- Article Title
- Abstract
- Keywords
- Affiliation
- Language
- ISSN
- CODEN
- DOI
- References
- Conference
- Article Title, Abstract, Keywords, Authors
- Chemical Name
- CAS Number

Rysunek 57: Interfejs użytkownika w bazie *Scopus* – kryteria wyszukiwania. Źródło: *Scopus*.

Jeżeli chodzi o eksport wyników dostępne są następujące opcje zestawione w poniższej tabeli. Istnieje jednak ograniczenie do 2.000 rekordów jednocześnie. Podobne ograniczenia

obowiązują jeżeli chodzi o tworzenie bibliografii (2.000), wysyłanie e-maila (200) i drukowanie (200).

Opcja	Scopus
Dodanie do listy	+
Przesłanie e-mailem	+
Wydruk	+
Zapis tekstowy	+
Zapis w <i>EndNote</i>	+
Zapis w formacie csv (<i>Excel</i>)	+
Zapis w pdf	+
Zapis w <i>ProCite</i>	+
Zapis w <i>RefMan</i>	+
Zapis w <i>RefWorks</i>	+
SUMA:	10

Tabela 27: Opcje eksportu wyników w bazie *Scopus*. Źródło: opracowanie własne.

i Please note that there is a maximum number of documents for certain output types.

If you have selected more, only the first documents up to the limit will be in the final output.

The output limits are

Export: 2,000 E-mail: 200
Bibliography: 2,000 Print: 200

1 Select the desired output type for the 1,536 selected documents.

Export Print E-mail Bibliography

2 **Export:** Choose your preferences and click **Export**.

Export format: RIS format (Reference Manager, ProCite, EndNote)

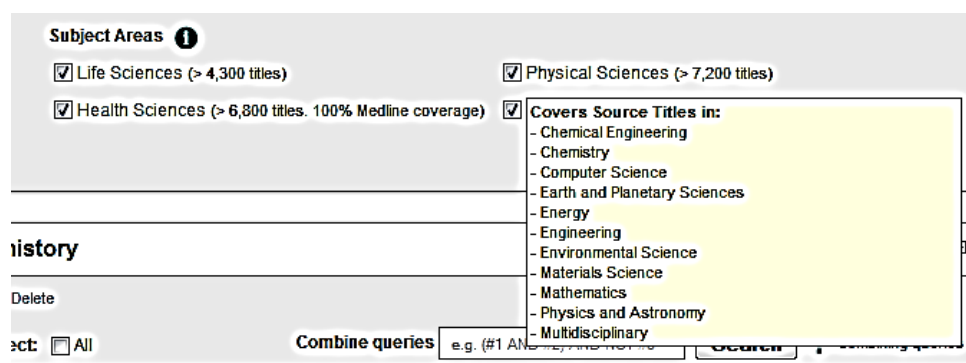
Output: Text (ASCII format)
RefWorks direct export
RIS format (Reference Manager, ProCite, EndNote)
BibTeX
Comma separated file, .csv (e.g. Excel)

Selected output includes:

<p>Citation information</p> <ul style="list-style-type: none"> • Author(s) • Document title • Year 	<p>Abstract and Keywords</p> <ul style="list-style-type: none"> • Abstract • Author Keywords • Index Keywords
--	---

Rysunek 58: Interfejs użytkownika w bazie *Scopus* – opcje eksportowania wyników wyszukiwania. Źródło: *Scopus*.

Baza *Scopus* (podobnie jak *Web of Science*) prezentuje także wyniki z podziałem na ogólne kategorie nauk (nauki przyrodnicze, nauki o zdrowiu, nauki fizyczne, nauki społeczne i humanistyczne) a w każdej z kategorii prezentuje dodatkowe podziały (np. matematyka, nauki informatyczne, inżynieria, biologia komórkowa i wiele innych).



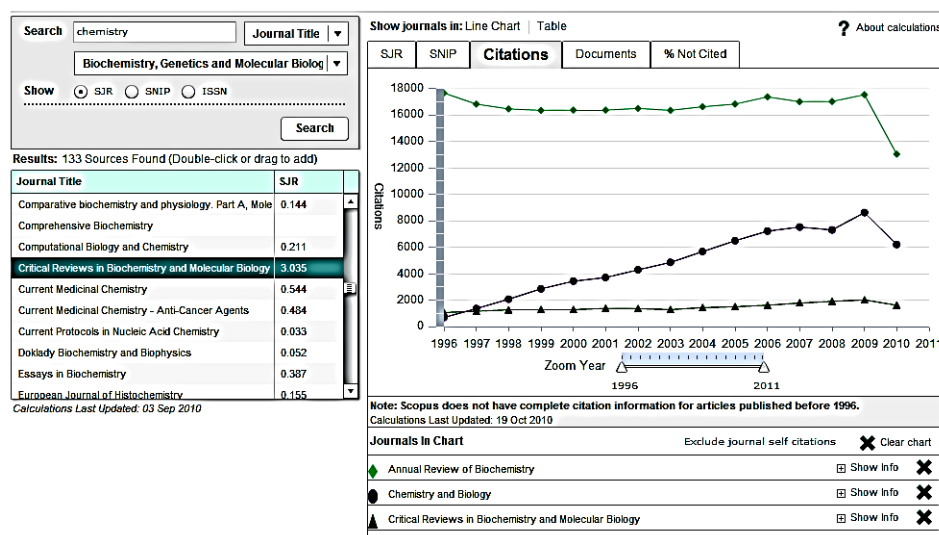
Rysunek 59: Interfejs użytkownika w bazie Scopus – ogólne kategorie nauk i obszary tematyczne. Źródło: Scopus.

Główna dziedzina nauki	Hasło po polsku	Hasło po angielsku	Wynik w bazie: Scopus
Med, B, Ch	Alergia	<i>allergy</i>	31.179
Med, B, Ch	Złamanie otwarte	<i>“compound fracture”</i>	227
T, F, Mat	System operacyjny	<i>“operating system”</i>	13.243
T, F, Mat	Internet	<i>Internet</i>	124.030
T, F, Mat	Konstrukcja nośna	<i>„supporting structure”</i>	918
T, F, Mat	Wytrzymałość materiałów	<i>„strength of materials”</i>	16.780
Ch, B	Całkowita zawartość węgla organicznego	<i>„total organic carbon content”</i>	170
B, Ch	Sosna śródziemnomorska	<i>„stone pine”</i>	112
B, Ch	Dolina Nilu	<i>„Nile Valley”</i>	129
Ch, B	Destylacja ropy naftowej	<i>“petroleum distillation”</i>	102
E	Bezrobocie	<i>unemployment</i>	8.222
E	Teoria zarządzania	<i>“theory of management”</i>	33
H	Filologia klasyczna	<i>„classical philology”</i>	9
H	Edukacja młodzieży	<i>“education of youth”</i>	18
S	Przedstawienie teatralne	<i>„theater performance”</i>	12
S	Obraz olejny	<i>„oil painting”</i>	140

Tabela 28: Przykładowe wyniki wyszukiwania w bazie Scopus dla kilku haseł z wybranych obszarów nauki (w polu: tytuł artykułu, abstrakt, słowa kluczowe) dla okresu 2005 – 2010. Źródło: opracowanie własne. Legenda: wyniki dotyczą tylko abstraktów dostępnych w bazie Scopus, a nie udostępnianych przez interfejs wyszukiwania zasobów internetowych. Skróty literowe przy hasłach odpowiadają dziedzinom nauki z polskiej klasyfikacji nauk (CK 2005), która na potrzeby niniejszego opracowania została nieznacznie zmodyfikowana:

- B – nauki biologiczne,
- Ch – nauki chemiczne,
- E – nauki ekonomiczne,
- F – nauki fizyczne,
- H – nauki humanistyczne,
- Mat – nauki matematyczne,
- Med – nauki medyczne,
- T – nauki techniczne,
- S – sztuka.

Jeżeli chodzi o narzędzia analizy otrzymanych wyników wyszukiwania w bazie *Scopus* dostępny jest analizator czasopism (*Journal Analyzer*), który porównuje czasopisma (maksymalnie 10 tytułów) i prezentuje wykres graficzny SJR (*SCImagoJournalRank*), określający wartość czasopisma biorąc pod uwagę tematykę, jakość, reputację i ich wpływ (liczbę cytowań). Można ocenić wartość czasopisma także z użyciem wykresu graficznego SNIP (*Source Normalized Impact per Paper*), który mierzy ważony wpływ cytowań w kontekście wszystkich cytowań w danym obszarze tematycznym. Kolejną opcją analizatora czasopism jest wykres cytowań, który prezentuje liczbę cytowań w bazie w ciągu każdego roku a także wykres dokumentów, który pokazuje liczbę artykułów publikowanych przez wybrane i porównywane czasopisma według roczników. Można także sporządzić wykres artykułów, które nie zostały zacytowane (*% Not Cited*). Wyniki analiz można przejrzeć także w formie tabelarycznej. Analizę można wydrukować lub wysłać e-mailem, nie ma możliwości eksportowania.

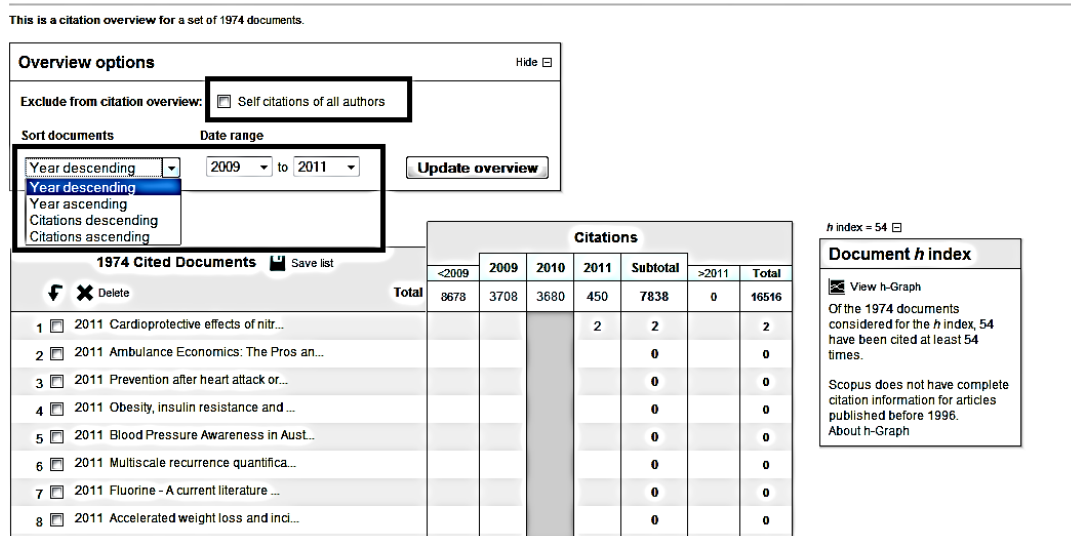


Rysunek 60: Interfejs użytkownika w bazie *Scopus* – analizator czasopism. Źródło: *Scopus*.

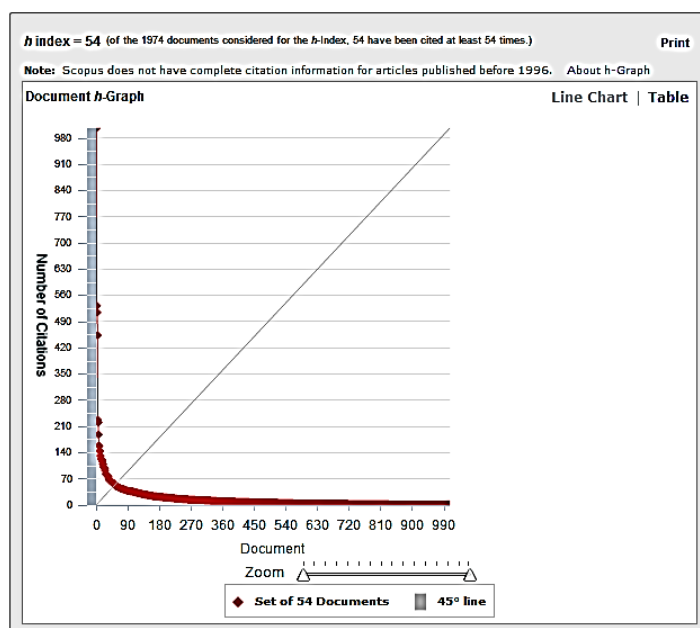
Można także wykonać przegląd cytowań dla wyszukanego zbioru informacji z podziałem na lata, w wyznaczonym porządku a także z wykluczeniem auto-cytowań autora. Można także sporządzić wykres indeksu h dla przeglądu cytowań.

Citation overview

Citations received since 1996



Rysunek 61: Przykładowy przegląd cytowań w bazie *Scopus* – wersja tabelaryczna. Źródło: *Scopus*.



Rysunek 62: Przykładowy przegląd cytowań w bazie *Scopus* – wersja graficzna. Źródło: *Scopus*.

Można także otrzymać listę dokumentów cytowanych i cytujących dla określonego zakresu wyszukanych wyników.

The screenshot shows a search results page for Scopus with 4,281 results. The interface includes a toolbar with options like 'Download PDF', 'Export', 'Print', 'Email', 'Create bibliography', 'Add to My List', 'View citation overview', and 'View citations'. A table lists four documents with their titles, authors, dates, source titles, and citation counts.

Document title	Author(s)	Date	Source title	Citations
1 Cardioprotective effects of nitrite during exercise	Calvert, J.W.	2011	Cardiovascular Research 89 (3), pp. 499-506	2
2 Ambulance Economics: The Pros and Cons of Fiscal Stimuli	Corden, W.M.	2011	Open Economies Review, pp. 1-11 Article in Press	0
3 Prevention after heart attack or stroke - Neither omega 3 fatty acids nor B vitamins protect [Prävention nach Herzinfarkt oder Schlaganfall - Weder Omega-3-Fettsäuren noch B-Vitamine schützen]	[No author name available]	2011	Klinikerzt 40 (1), pp. 8	0
4 Obesity, insulin resistance and free fatty acids	Boden, G.	2011	Current Opinion in Endocrinology, Diabetes and Obesity Article in Press	0

Rysunek 63: Opcje generowania listy cytowań w bazie Scopus. Źródło: Scopus.

10.5. Opis bazy INSPEC

INSPEC jest bibliograficzną bazą z zakresu nauk ścisłych i inżynierskich zawierająca dane z czasopism zarówno drukowanych jak i *online*, sprawozdań konferencyjnych, książek, raportów i rozpraw. Dodatkowo zastosowane jest eksperckie indeksowanie, dzięki któremu możliwe jest uniknięcie niewłaściwych informacji. Około 60% zasobów bazy posiada bezpośrednie linki do pełnotekstowych artykułów. Zakres dostępnych danych obejmuje pięć głównych obszarów (IET 2011f: 4):

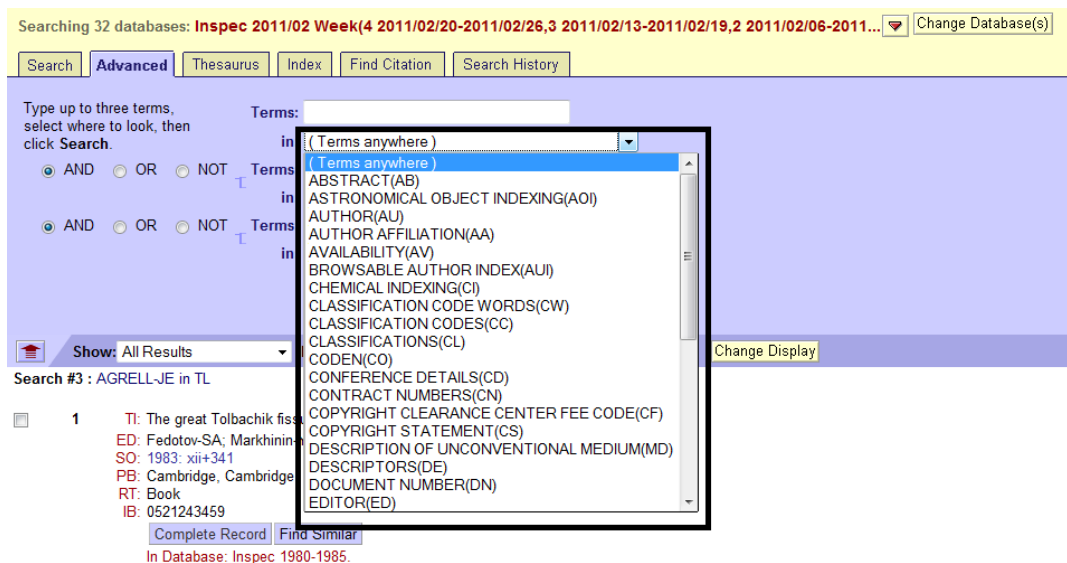
- Sekcja A – fizyka,
- Sekcja B – elektrotechnika i elektronika,
- Sekcja C – komputery i systemy kontrolne,
- Sekcja D – technologia informacyjna dla biznesu,
- Sekcja E – mechanika i inżynieria produkcji,

Główna klasyfikacja jest detalicznie podzielona na węższe obszary tematyczne w każdej z pięciu wyżej wymienionych sekcji. Każdy rekord jest przypisany odpowiednim kodem do właściwej (lub kilku właściwych) pozycji z klasyfikacji *INSPEC*.

Prócz klasyfikacji zastosowane jest także indeksowanie numeryczne (wartości liczbowych), chemiczne (substancje chemiczne), astronomiczne (obiektów astronomicznych) i słowa kluczowe i kodowanie wskazujące podejście autora do tematu (np. praktyczne, ekonomiczne, bibliograficzne, produktowe, teoretyczne/matematyczne, eksperymentalne) a także tezaurus ułatwiający wychwytywanie relacji między terminami i datami. Tytuł artykułu i abstrakt są w języku angielskim, wyświetlane są także prócz nazwiska i inicjałów imienia autora(ów) jego afiliacja, tytuł źródła, wydawca, data wydania, miejsce wydania, typ dokumentu i oryginalny język artykułu (IET 2011f: 6).

Dostęp do bazy *INSPEC* można uzyskać przez serwer *OVID* a także przez *WebSPIRS*. Zalecany jest dostęp przez drugą z wymienionych platform z uwagi na pewne ograniczenia *OVID*, np. dotyczące eksportowania danych czy też uboższe wyniki wyszukiwania w samej bazie. Dla przykładu, dla hasła „Internet” (w polu „*terms anywhere*” z limitem dat 2005 – 2010), *OVID* zwraca 6747 wyników zaś *WebSPIRS* aż 82958, co stanowi ogromną różnicę.

Baza *INSPEC* jest najbogatszą z dotychczas analizowanych jeżeli chodzi o kryteria wyszukiwania, zawiera ich ponad 40.



Rysunek 64: Przegląd opcji wyszukiwania w bazie *INSPEC*. Źródło: *INSPEC*.

Kryterium	<i>INSPEC</i>
Abstrakt	+
Afiliacja	+
Autor	+
<i>CODEN</i>	+
Czasopismo	+
Data priorytetu patentu	+
Data publikacji	+
Daty	+
Deskryptory <i>INSPEC</i>	+
<i>DOI</i>	+
Dostępność	+
Identyfikatory	+
Indeks astronomiczny	+
Indeks chemiczny	+
Indeks numeryczny	+
<i>ISBN</i>	+
<i>ISSN</i>	+
Język	+
Klasyfikacja <i>INSPEC</i>	+
Klasyfikacje (kod oryginalnej klasyfikacji)	+
Klucz sortowania	+
Kod <i>Copyright Clearance Center</i>	+
Kod uaktualnienia	+
Konferencja	+
Kraj publikacji	+

Kryterium	INSPEC
Liczba odwołań	+
Liczba stron	+
Nagłówek tematyczny <i>INSPEC</i>	+
Nota <i>Copyright</i>	+
Numer dokumentu	+
Numer kontraktu	+
Numer raportu	+
Opis źródła elektronicznego	+
Słowa w klasyfikacji <i>INSPEC</i>	+
Szczegóły dokumentu patentowego	+
Szczegóły patentu	+
Temat/słowa kluczowe	+
Tłumacz	+
Typ dokumentu	+
Tytuł	+
<i>URL</i>	+
Właściciel patentu	+
Wszystkie pola	+
Wydawca	+
Źródło dokumentu	+
SUMA	45

Tabela 29: Kryteria wyszukiwania w bazie *INSPEC*. Źródło: opracowanie własne.

Mimo wielu kryteriów wyszukiwania, interfejs ma strukturę przejrzystą i intuicyjną w obsłudze w przeciwieństwie do bazy *MEDLINE*, gdzie wielość kryteriów niekiedy komplikuje i utrudnia proces wyszukiwania. Warto podkreślić, że w bazie dostępne są także abstrakty patentów i wyszukiwanie ze względu na właściwości dokumentów patentowych, dotyczy to jednak patentów historycznych, pochodzących sprzed kilkadziesiąt lat. *INSPEC* zawiera abstrakty patentów z baz brytyjskich i amerykańskich (IET 2011f: 3).

W bazie *INSPEC* dostępne są dla użytkowników interesujące produkty i usługi usprawniające pracę:

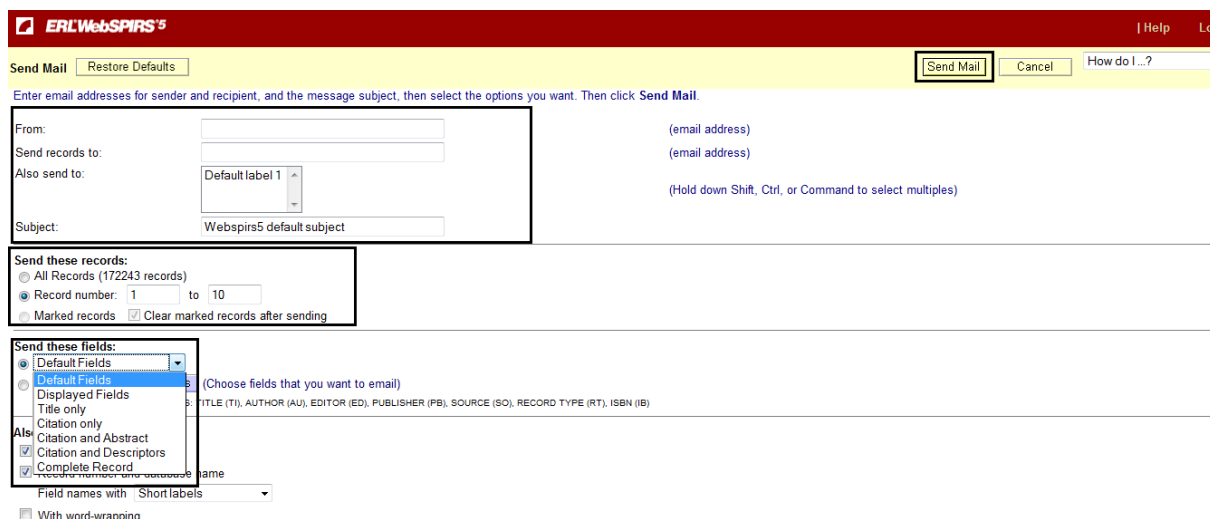
- *Inspec Direct* – platforma zwiększająca możliwości wyszukiwania w bazie i ułatwiająca przeciwdziałanie przeładowaniu informacyjnemu. W *Inspec Direct* można wyszukiwać używając wymienionych już kryteriów i sposobów indeksowania. Możliwe jest także sortowanie w odniesieniu do tychże kryteriów oraz eksport danych z użyciem wielu dostępnych formatów plików. Zapisywanie wyników i historii wyszukiwania oraz przesyłanie e-mailem także są dostępne. Możliwy jest także dostęp do pełnych tekstów przez odpowiedni link oraz spersonalizowane alerty. Dla szukających odpowiednich danych (*professional searcher and data miner*), *Inspec Direct* umożliwia analizę wyników wyszukiwania przez odpowiednie procedury rangowania i prezentacji ich w czasie oraz pobrania danych w formacie CSV. Analizowane wyniki można także wizualizować za pomocą wykresów kołowych czy słupkowych. Dodatkowe usługi dostępne są specjalnie dla bibliotekarzy, np. prowizje od statystyk użytkowania, linki do pełnych tekstów i zasobów bibliotecznych oraz

MetaSearch, które pozwala na przeszukiwania bazy *INSPEC* wraz z innymi bazami przy użyciu odpowiedniego systemu (np. *MetaLib*) (InspecDirect 2011).

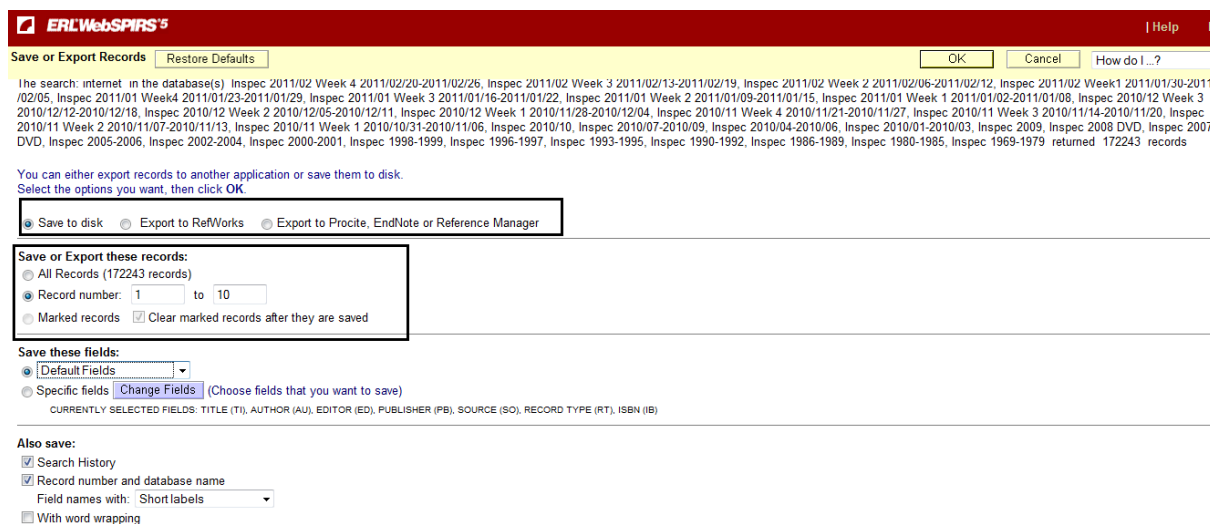
- *Inspec Direct data* – narzędzie dostarczające w XML cotygodniowy plik z danymi, dla klientów, którzy chcą pracować w domu na swoim własnym oprogramowaniu (IET 2011f).
- *Inspec Archive* – dostęp do archiwalnych zasobów z czasopism z lat 1898 – 1968. Umożliwia to dostęp do prac np. A. Einsteina, M. Plancka czy M. Curie – Skłodowskiej (IET 2011f).
- *Inspec Pay-as-you-go* – dla użytkowników nie potrzebujących stałego i nieograniczonego dostępu do bazy (IET 2011c).
- *Inspec Print Journals* – drukowane, krótkie podsumowania materiałów dla użytkowników w różnych konfiguracjach (IET 2011d).
- *Inspec Search Aids* – pakiety zawierające elektroniczne wersje klasyfikacji, tezauryusa, listę czasopism indeksowanych i dokumentację użytkownika, które mają za zadanie usprawnić pracę (IET 2011e).

The screenshot shows the ERLWebSPIRS 5 interface. At the top, there is a red header with the logo and 'Database Guide | Help | Log'. Below the header, there is a yellow bar with 'Print Options', 'Restore Defaults', and 'Preview' buttons. A 'Print' button is highlighted. Below this, there is a text area showing search results for 'internet in the database(s)'. Below the search results, there are instructions: 'Select the print options you want, then click Print. Restore Defaults restores the print options your administrator set. Preview lets you see the records you are going to print.' Below the instructions, there is a 'Records to be printed:' section with radio buttons for 'All Records (172243 records)', 'Record number: 1 to 10', 'Marked Records Only', and 'Clear marked records after printing'. There is also a checkbox for 'Print one record per page'. Below this, there is a 'Print these fields:' section with a dropdown menu showing 'Default Fields' selected. Below the dropdown, there is a list of fields: 'Title only', 'Citation only', 'Citation and Abstract', 'Citation and Descriptors', and 'Complete Record'. There is also a checkbox for 'Record number and database name' which is checked.

Rysunek 65: Opcja wydruku wyników wyszukiwania w bazie *INSPEC*. Źródło: *INSPEC*.



Rysunek 66: Opcja przesłania e-mailem wyników wyszukiwania w bazie INSPEC. Źródło: INSPEC.



Rysunek 67: Opcja przesłania eksportu i zapisu wyników wyszukiwania w bazie INSPEC. Źródło: INSPEC.

Jeżeli chodzi o eksport wyników, dostępne są następujące opcje zestawione w poniższej tabeli. Wyniki wyszukiwania można wydrukować, przesłać e-mailem i eksportować bez żadnych ograniczeń, co do liczby rekordów.

Opcja	INSPEC
Przesłanie e-mailem	+
Wydruk	+
Zapis w <i>EndNote</i>	+
Zapis w <i>ProCite</i>	+
Zapis w <i>RefMan</i>	+
Zapis w <i>RefWorks</i>	+
Zapis w formacie .dat	+
SUMA:	7

Tabela 30: Opcje eksportu wyników w bazie INSPEC. Źródło: opracowanie własne.

Główna dziedzina nauki	Hasło po polsku	Hasło po angielsku	Wynik w bazie: <i>INSPEC</i>
Med, B, Ch	Alergia	<i>allergy</i>	179
Med, B, Ch	Złamanie otwarte	<i>“compound fracture”</i>	7
T, F, Mat	System operacyjny	<i>“operating system”</i>	8396
T, F, Mat	Internet	<i>Internet</i>	82958
T, F, Mat	Konstrukcja nośna	<i>„supporting structure”</i>	340
T, F, Mat	Wytrzymałość materiałów	<i>„strength of materials”</i>	1103
Ch, B	Całkowita zawartość węgla organicznego	<i>„total organic carbon content”</i>	31
B, Ch	Sosna śródziemnomorska	<i>„stone pine”</i>	3
B, Ch	Dolina Nilu	<i>„Nile Valley”</i>	11
Ch, B	Destylacja ropy naftowej	<i>“petroleum distillation”</i>	1
E	Bezrobocie	<i>unemployment</i>	414
E	Teoria zarządzania	<i>“theory of management”</i>	498
H	Filologia klasyczna	<i>„classical philology”</i>	0
H	Edukacja młodzieży	<i>“education of youth”</i>	9
S	Przedstawienie teatralne	<i>„theater performance”</i>	3
S	Obraz olejny	<i>„oil painting”</i>	34

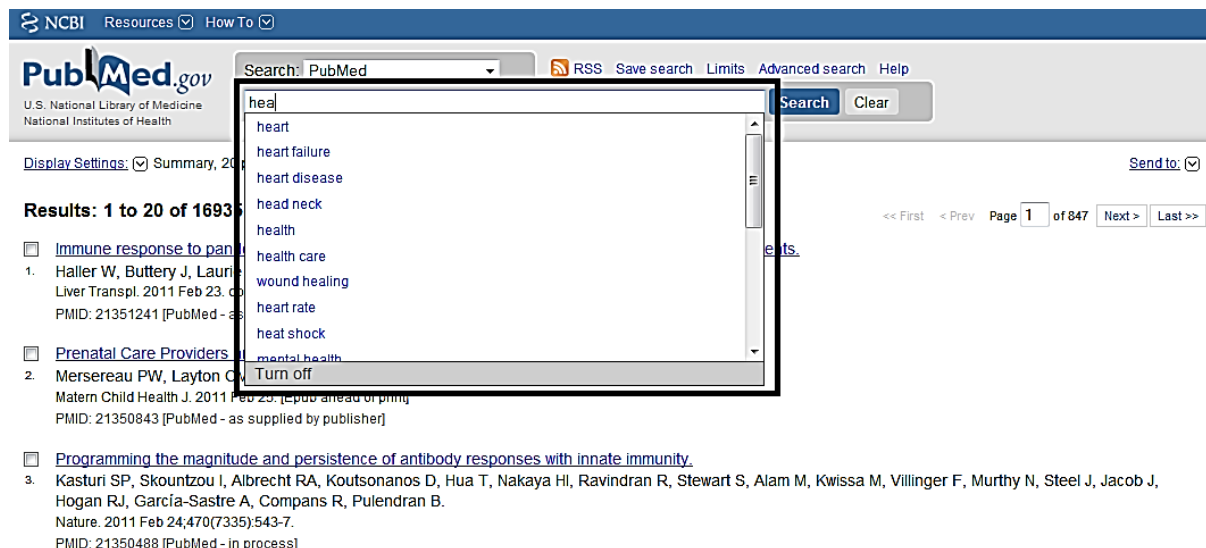
Tabela 31: Przykładowe wyniki wyszukiwania w bazie *INSPEC* dla kilku haseł z wybranych obszarów nauki (w polu: „*terms anywhere*”) dla okresu 2005 – 2010. Źródło: opracowanie własne. Legenda: Skrót literowy przy hasłach odpowiada dziedzinom nauki z polskiej klasyfikacji nauk (CK 2005), która na potrzeby niniejszego opracowania została nieznacznie zmodyfikowana:

- B – nauki biologiczne,
- Ch – nauki chemiczne,
- E – nauki ekonomiczne,
- F – nauki fizyczne,
- H – nauki humanistyczne,
- Mat – nauki matematyczne,
- Med – nauki medyczne,
- T – nauki techniczne,
- S – sztuka.

10.6. Opis bazy *MEDLINE*

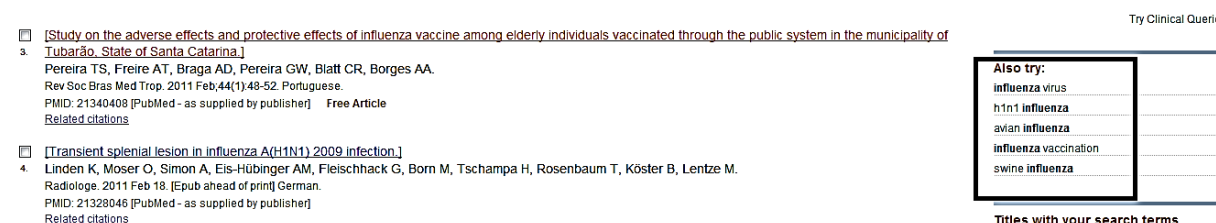
MEDLINE (znana też pod nazwą *PubMed*) jest bazą bibliograficzną opisów artykułów z zakresu nauk przyrodniczych z naciskiem na biomedycynę. Rekordy w bazie *MEDLINE* są indeksowane za pomocą taksonomii *MeSH* (*Medical Subject Headings*). *MeSH* zawiera zbiór terminów (taksonów zwanych w bazie deskryptorami), które pozwalają na wielowarstwowe wyszukiwanie w oparciu o 12 poziomów porządkowych, ogółem w 2011 jest dostępnych ponad 26 tys. szczegółowych deskryptorów. Dodatkowo jest dostępnych 177 tysięcy słów kluczowych ułatwiających wyszukiwanie, które to słowa przekierowują użytkownika

automatycznie do danej kategorii opisywanej przez deskryptory. Ponadto istnieje 199 tysięcy nagłówków uzupełniających z osobnym tezaurem zwanych *Supplementary Concept Records* (NLM 2011a). Po wpisaniu terminu, w oknie wyszukiwarki dzięki klasyfikacji *MeSH* sugerowane są odpowiednie terminy.



Rysunek 68: Wyszukiwanie z pomocą klasyfikacji *MeSH* w bazie *MEDLINE*. Źródło: *MEDLINE*.

Dodatkowo, po wyświetleniu listy wyników dla szukanego terminu tezaurek, działający według klasyfikacji *MeSH*, podpowiada pokrewne terminy.



Rysunek 69: Sugerowane przez tezaurek powiązane terminy w bazie *MEDLINE*. Źródło: *MEDLINE*.

Wyszukiwanie można uszczegółowić biorąc pod uwagę bogaty zestaw kryteriów. Obok przeszukiwanych pól, których jest prawie 40.

Kryterium	<i>MEDLINE</i>
Afiliacja	+
Autor	+
Badacz	+
Czasopismo	+
Data dodania do bazy	+
Data <i>MeSH</i> (data dodania do klasyfikacji)	+
Data publikacji	+
Data stworzenia rekordu	+

Kryterium	MEDLINE
Daty	+
Działania farmakologiczne	+
Filtry używane przez <i>LinkOut</i>	+
Główny temat wg <i>MeSH</i>	+
Grupa autorów / Autorzy/autor korporacyjny/grupowy	+
Identyfikator lokalizacyjny (np. <i>DOI</i>)	+
Inne źródło danych	+
<i>ISBN</i>	+
Język	+
Książka	+
Nr czasopisma	+
Nr <i>EC/RN</i> (klasyfikacja enzymów/numer rejestru <i>CAS</i>)	+
Numer grantu	+
Ostatni autor	+
Paginacja	+
Pełna nazwa autora	+
Pełna nazwa badacza	+
Pierwszy autor	+
Podtemat wg <i>MeSH</i>	+
Rocznik (czasopisma)	+
Słowa w tekście	+
Termin <i>MeSH</i>	+
Typ dokumentu	+
Tytuł	+
Tytuł oryginału	+
Tytuł/abstrakt	+
Wszystkie pola	+
Wydawca	+
SUMA:	36

Tabela 32: Kryteria wyszukiwania w bazie *MEDLINE*. Źródło: opracowanie własne.

Na uwagę zasługują rozbudowane kryteria związane z klasyfikacją *MeSH* i nazwą/nazwiskiem autora, a także rozróżnienie ze względu na autora i badacza. Niektóre z pozostałych kryteriów w większości przypadków będą się pokrywać (pełna data wydania zamiast tylko roku wydania, tytuł/abstrakt zamiast tytuł/abstrakt/słowa kluczowe, książka i czasopismo zamiast tytułu źródła lub nazwy publikacji).

PubMed Advanced Search

Rysunek 70: Interfejs użytkownika w bazie *MEDLINE* – kryteria wyszukiwania. Źródło: *MEDLINE*.

Dodatkowo wyniki wyszukiwania można uszczegółowić ze względu na takie kryteria jak:

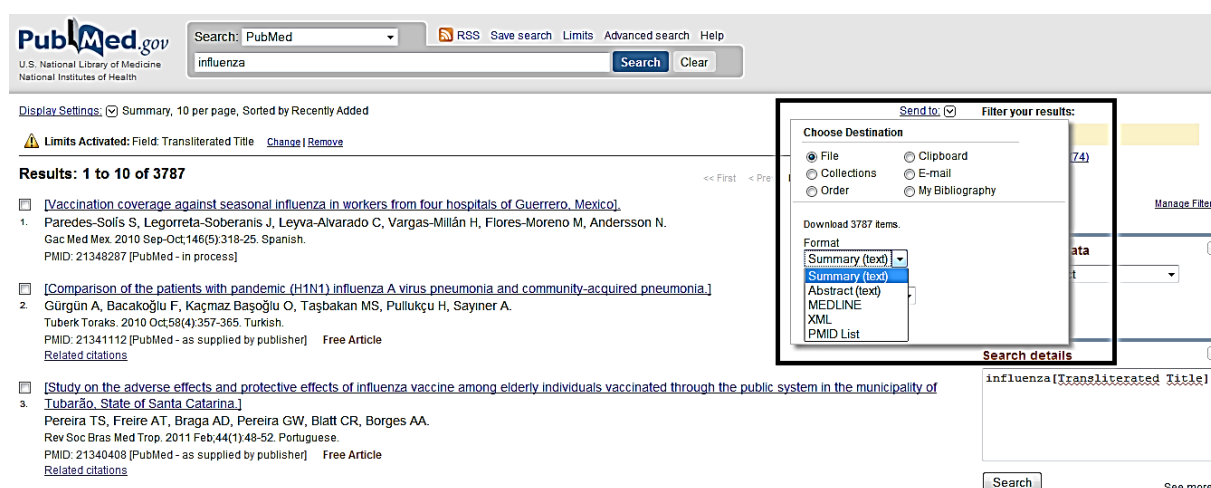
- opcje dostępu do tekstu (abstraktu, pełnego tekstu, pełnego tekstu płatnego),
- płeć,
- wiek,
- gatunek (ludzie, zwierzęta).

Mimo wielu szczegółowych kryteriów wyszukiwania, możliwości także wielokryterialnego selekcjonowania wyników i używania tezauryśa *MeSH*, interfejs wyszukiwania pozostaje skomplikowany, a sama klasyfikacja *MeSH* może wielu użytkownikom wydawać się trudna w użyciu. Mimo niekwestionowanych zalet w postaci możliwości wyszukiwania złożonych medycznych terminów i odkrywania zależności dzięki *MeSH*, interfejs bazy *MEDLINE* nie należy do najłatwiejszych w obsłudze.

Jeżeli chodzi o eksport wyników, dostępne są następujące opcje zestawione w poniższej tabeli. Można zapisać wyniki w pliku (nie ma ograniczeń, co do liczby rekordów) tekstowym (tylko podsumowanie lub także abstrakty), w formacie *MEDLINE*, *XML*, listy *PMID*. Jeżeli chodzi o tworzenie listy (kolekcji) istnieje ograniczenie do 1000 rekordów, kopiowanie do schowka (500 rekordów), tworzenie bibliografii (500 rekordów), wysyłanie e-mailem (200 rekordów od wybranego numeru).

Opcja	<i>MEDLINE</i>
Dodanie do listy	+
Lista <i>PMID</i>	+
Przesłanie e-mailem	+
Format <i>MEDLINE</i>	+
Format <i>XML</i>	+
Format tekstowy	+
SUMA:	6

Tabela 33: Opcje eksportu wyników w bazie *MEDLINE*. Źródło: opracowanie własne.



Rysunek 71: Interfejs użytkownika w bazie MEDLINE – opcje eksportowania wyników wyszukiwania. Źródło: MEDLINE.

Baza MEDLINE zawiera dane od 1967 oraz dodatkowe zasoby starszych dokumentów. Głównym obszarem tematycznym jest jak już wspomniano biomedycyna i zdrowie, z uwzględnieniem obszarów użytecznych dla opieki klinicznej, zdrowia publicznego rozwoju polityki zdrowotnej, biomedyków, badaczy i dydaktyków, w tym przede wszystkim (NLM 2011b):

- Nauk przyrodniczych,
- Nauk humanistycznych,
- Nauk chemicznych,
- Niektórych działów bioinżynierii,
- Elementów biologii,
- Nauk o ochronie środowiska,
- Biologii morskiej,
- Nauk o roślinach i zwierzętach,
- Biofizyki i chemii.

Materiały w bazie MEDLINE pochodzą w większości z czasopism naukowych, w mniejszej części dotyczą materiałów prasowych (z gazet, magazynów i newsletterów), które są używane w środowisku adresatów bazy. Wynikiem wyszukiwania w bazie jest lista cytowań zawierających autorów, tytuł, źródło i często abstrakt, wskazana jest także możliwość wolnego dostępu do elektronicznej wersji pełnego tekstu, jeżeli jest w archiwum bazy. Można się także przekierować do strony wydawcy danego tekstu i w ten sposób uzyskać dostęp do jego pełnej wersji. Niekiedy można zamówić pełną wersję tekstu przez Narodową Sieć Bibliotek Medycznych (*National Network of Libraries of Medicine*) za pomocą narzędzia *LoansomeDoc* (samo narzędzie jest bezpłatne, natomiast ceny poszczególnych artykułów mogą się różnić w zależności od poszczególnych taryfikatorów bibliotecznych (NLM 2011c)). Dostęp do bazy MEDLINE jest bezpłatny (NLM 2011b).

Dane w bazie MEDLINE gromadzone są na zamówienie rządu USA, lecz mogą być chronione przez prawa autorskie, co ogranicza możliwości ich wykorzystywania. Abstrakty w bazie mogą być chronione zarówno przez amerykańskie, jak i inne prawa autorskie, zależnie

od publikacji, z której zostały pobrane. Narodowa Biblioteka Medyczna nie ponosi odpowiedzialności za użytkowanie chronionych prawem autorskim materiałów - dotyczy to także przekazywania, powielania i rozpowszechniania, także w celach komercyjnych. Można pobierać i wykorzystywać dane na własny użytek lub rozpowszechniać pojedyncze kopie, w innym przypadku zalecana jest konsultacja prawna. Z tego względu, dla indywidualnych użytkowników nie jest wymagana odpowiednia umowa licencyjna (NLM 2011d). Dostępne są także pewne warianty bazy *MEDLINE* np. *MedlinePlus* (zawierająca głównie informacje z zakresu zdrowia (NLM 2011e)) czy *OLDMedline* (zawierająca dane z okresu od 1946 do 1965 (NLM 2011f)). Należy zaznaczyć, że wszyscy autorzy danego artykułu są w bazie indeksowani, dostępne są też dane o afiliacjach, mogą być widoczne także nazwiska poszczególnych członków określonego zespołu badawczego, komitetu, konsorcjum czy grupy studyjnej (NLM 2011g). Prócz nazwisk i nazw autorów, użytkownik dostaje także standardowe informacje o miejscu i dacie wydania, dokładnym tytule i numerze czasopisma, języku publikacji i informacje na temat praw autorskich oraz kategoriach *MeSH* (NLM 2011h).

Główna dziedzina nauki	Hasło po polsku	Hasło po angielsku	Wynik w bazie: <i>MEDLINE</i>
Med, B, Ch	Alergia	<i>allergy</i>	64.233
Med, B, Ch	Złamanie otwarte	<i>“compound fracture”</i>	11
T, F, Mat	System operacyjny	<i>“operating system”</i>	162
T, F, Mat	Internet	<i>Internet</i>	24.833
T, F, Mat	Konstrukcja nośna	<i>„supporting structure”</i>	36
T, F, Mat	Wytrzymałość materiałów	<i>„strength of materials”</i>	0
Ch, B	Całkowita zawartość węgla organicznego	<i>„total organic carbon content”</i>	51
B, Ch	Sosna śródziemnomorska	<i>„stone pine”</i>	16
B, Ch	Dolina Nilu	<i>„Nile Valley”</i>	26
Ch, B	Destylacja ropy naftowej	<i>“petroleum distillation”</i>	0
E	Bezrobocie	<i>unemployment</i>	2.119
E	Teoria zarządzania	<i>“theory of management”</i>	0
H	Filologia klasyczna	<i>„classical philology”</i>	2
H	Edukacja młodzieży	<i>“education of youth”</i>	0
S	Przedstawienie teatralne	<i>„theater performance”</i>	2
S	Obraz olejny	<i>„oil painting”</i>	10

Tabela 34: Przykładowe wyniki wyszukiwania w bazie *MEDLINE* dla kilku haseł z wybranych obszarów nauki (w polu: „*all fields*”) dla okresu 2005 – 2010. Źródło: opracowanie własne. Legenda: Skrót literowy przy hasłach odpowiada dziedzinom nauki z polskiej klasyfikacji nauk (CK 2005), która na potrzeby niniejszego opracowania została nieznacznie zmodyfikowana:

- **B** – nauki biologiczne,
- **Ch** – nauki chemiczne,
- **E** – nauki ekonomiczne,
- **F** – nauki fizyczne,
- **H** – nauki humanistyczne,
- **Mat** – nauki matematyczne,
- **Med** – nauki medyczne,
- **T** – nauki techniczne,
- **S** – sztuka.

10.7. Porównanie funkcjonalności baz bibliograficznych

Poniższy podrozdział zawiera tabelaryczne zestawienia funkcjonalności poszczególnych baz bibliograficznych, które mogą ułatwić wybór odpowiedniego narzędzia dla potrzeb analiz bibliometrycznych. Warto pamiętać, że o wyborze bazy decydują przede wszystkim:

- specjalizacja tematyczna,
- dostępność niezbędnych danych (np. danych o cytowaniach),
- możliwość identyfikacji danych (taksonomie, ułatwiające wyszukiwanie),
- łatwość eksportu danych w celu przetwarzania w dedykowanym oprogramowaniu.

Kryterium	<i>Web of Science</i>	<i>Scopus</i>	<i>INSPEC</i>	<i>MEDLINE</i>
Abstrakt		+	+	
Adres	+			
Afiliacja		+	+	+
Agencja finansująca	+			
Autor	+	+	+	+
Badacz				+
<i>CODEN</i>		+	+	
Czasopismo			+	+
Data dodania do bazy				+
Data <i>MeSH</i> (data dodania do klasyfikacji)				+
Data priorytetu patentu			+	
Data publikacji			+	+
Data stworzenia rekordu				+
Daty	+	+	+	+
Deskryptory <i>INSPEC</i>			+	
Dodatkowe klasyfikacje (inne niż <i>MeSH</i>)				
<i>DOI</i>		+	+	
Dostępność			+	
Działania farmakologiczne				+
Filtry używane przez <i>LinkOut</i>				+
Główny temat wg <i>MeSH</i>				+
Grupa autorów / Autorzy/autor korporacyjny/grupowy	+	+		+

Kryterium	<i>Web of Science</i>	<i>Scopus</i>	<i>INSPEC</i>	<i>MEDLINE</i>
Identyfikator lokalizacyjny (np. DOI)				+
Identyfikatory			+	
Indeks astronomiczny			+	
Indeks chemiczny			+	
Indeks numeryczny			+	
Inne źródło danych				+
ISBN			+	+
ISSN		+	+	
Język	+	+	+	+
Klasyfikacja <i>INSPEC</i>			+	
Klasyfikacje (kod oryginalnej klasyfikacji)			+	
Klucz sortowania			+	
Kod <i>Copyright Clearance Center</i>			+	
Kod uaktualnienia			+	
Konferencja	+	+	+	
Kraj publikacji			+	
Książka				+
Liczba odwołań			+	
Liczba stron			+	
Nagłówek tematyczny <i>INSPEC</i>			+	
Nazwa chemiczna		+		
Nazwa publikacji / tytuł źródła	+	+		
Nota <i>Copyright</i>			+	
Nr czasopisma				+
Nr <i>EC/RN</i> (klasyfikacja enzymów/numer rejestru CAS)				+
Numer CAS		+		
Numer dokumentu			+	
Numer grantu	+			+
Numer kontraktu			+	
Numer raportu			+	
Odwołania		+		
Opis źródła elektronicznego			+	
Ostatni autor				+
Paginacja				+
Pełna nazwa autora				+
Pełna nazwa badacza				+
Pierwszy autor		+		+
Podtemat wg <i>MeSH</i>				+
Rocznik (czasopisma)				+
Rok opublikowania	+	+		

Kryterium	<i>Web of Science</i>	<i>Scopus</i>	<i>INSPEC</i>	<i>MEDLINE</i>
Słowa kluczowe		+		
Słowa w klasyfikacji <i>INSPEC</i>			+	
Słowa w tekście				+
Szczegóły dokumentu patentowego			+	
Szczegóły patentu			+	
Temat	+			
Temat/słowa kluczowe			+	
Termin <i>MeSH</i>				+
Tłumacz			+	
Typ dokumentu	+	+	+	+
Tytuł	+	+	+	+
Tytuł oryginału				+
Tytuł/abstrakt				+
Tytuł/abstrakt/słowa kluczowe		+		
<i>URL</i>			+	
Właściciel patentu			+	
Wszystkie pola			+	+
Wydawca	+		+	+
Źródło dokumentu			+	
SUMA:	14	20	46	36

Tabela 35: Kryteria wyszukiwania – zestawienie zbiorcze. Źródło: opracowanie własne.

Opcja	Web of Science	Scopus	INSPEC	MEDLINE
Dodanie do listy	+	+		+
Lista <i>PMID</i>				+
Przesłanie e-mailem	+	+	+	+
W formacie dla systemu Mac	+			
W formacie dla systemu Windows	+			
W formacie <i>MEDLINE</i>				+
Wydruk	+	+	+	
XML				+
W formacie tekstowym	+	+		+
Zapis w <i>EndNote</i>	+	+	+	
Zapis w <i>EndNote Web</i>	+			
Zapis w formacie <i>csv</i> (Excel)		+		
Zapis w formacie <i>dat</i>			+	
Zapis w <i>HTML</i>	+			
Zapis w <i>pdf</i>		+		
Zapis w <i>ProCite</i>	+	+	+	
Zapis w <i>RefMan</i>	+	+	+	
Zapis w <i>RefWorks</i>	+	+	+	
SUMA:	12	10	7	6

Tabela 36: Opcje eksportu wyników w bazach – tabela zbiorcza. Źródło: opracowanie własne.

Główna dziedzina nauki	Hasło po polsku	Hasło po angielsku	WoK	Scopus	INSPEC	MEDLINE
Med, B, Ch	Alergia	<i>allergy</i>	20.465	31.179	179	64.233
Med, B, Ch	Złamanie otwarte	<i>“compound fracture”</i>	19	227	7	11
T, F, Mat	System operacyjny	<i>“operating system”</i>	3.644	13.243	8.396	162
T, F, Mat	Internet	<i>Internet</i>	67.067	124.030	82.958	24.833
T, F, Mat	Konstrukcja nośna	<i>„supporting structure”</i>	318	918	340	36
T, F, Mat	Wytrzymałość materiałów	<i>„strength of materials”</i>	324	16.780	1.103	0
Ch, B	Całkowita zawartość węgla organicznego	<i>„total organic carbon content”</i>	156	170	31	51
B, Ch	Sosna śródziemnomorska	<i>„stone pine”</i>	120	112	3	16
B, Ch	Dolina Nilu	<i>„Nile Valley”</i>	111	129	11	26
Ch, B	Destylacja ropy naftowej	<i>“petroleum distillation”</i>	10	102	1	0
E	Bezrobocie	<i>unemployment</i>	6.960	8.222	414	2.119
E	Teoria zarządzania	<i>“theory of management”</i>	109	33	498	0
H	Filologia klasyczna	<i>„classical philology”</i>	26	9	0	2
H	Edukacja młodzieży	<i>“education of youth”</i>	55	18	9	0
S	Przedstawienie teatralne	<i>„theater performance”</i>	7	12	3	2
S	Obraz olejny	<i>„oil painting”</i>	53	140	34	10
Liczba najwyższych wyników wyszukiwania:			3	11	1	1
Bezwzględny wynik wyszukiwań:			99.444	195.324	93.987	91.501
Miejsce ze względu na wynik bezwzględny:			II	I	III	IV
Udział procentowy (100% = całkowita suma wyników = 480.256)			20,71%	40,67%	19,57%	19,05%

Tabela 37: Przykładowe wyniki wyszukiwania w bazach dla kilku haseł z wybranych obszarów nauki (w najszerszej określonej opcji wyszukiwania) dla okresu 2005 – 2010. Źródło: opracowanie własne. Legenda: Skróty literowe przy hasłach odpowiadają dziedzinom nauki z polskiej klasyfikacji nauk (CK 2005), która na potrzeby niniejszego opracowania została nieznacznie zmodyfikowana:

- B – nauki biologiczne,
- Ch – nauki chemiczne,
- E – nauki ekonomiczne,
- F – nauki fizyczne,
- H – nauki humanistyczne,
- Mat – nauki matematyczne,
- Med – nauki medyczne,
- T – nauki techniczne,
- S – sztuka.

10.8. Opis bazy Europejskiego Urzędu Patentowego (EPO)

Baza Europejskiego Urzędu Patentowego *esp@cenet* wyszukuje patenty spośród około 60 milionów dokumentów. Dostępny jest polski interfejs bazy (*esp@cenet PL*), który pozwala na przeszukiwanie zarówno zasobów polskich jak i zagranicznych z *EPO*. W bazie można wyszukiwać dokumenty patentowe według następujących kryteriów (Espacenet PL 2011a):

- Słów kluczowych w tytule lub skrócie (abstrakcie),
- Numeru publikacji,
- Numeru zgłoszenia,
- Numeru pierwszeństwa,
- Daty publikacji,
- Zgłaszającego patent,
- Międzynarodowej Klasyfikacji Patentowej (*IPC*) – zawiera ok. 70 tys. grup,
- Klasyfikacji Europejskiej (*ECLA*) – bazuje na *IPC*, ale jest bardziej szczegółowa i zawiera ok. 129 tys. grup.

Kryteria różnią się nieznacznie w zależności od wariantu przeszukiwania wybranego przez użytkownika: baza patentów europejskich, baza ogólnoświatowa, baza Światowej Organizacji Ochrony Własności Intelektualnej *WIPO*. W każdym polu można wprowadzić maksymalnie 4 terminy przy czym w polu skrót koniecznie po angielsku. Można stosować operatory logiczne (*and*, *or*, *not*), pytania zagnieżdżone (z użyciem nawiasów okrągłych), istnieją operatory domyślne (w odpowiednich polach wyszukiwania przypisywane automatycznie) i symbole wieloznaczne (*, ?, #), nie można wyszukiwać ze względu na zakres dat.

Kryterium	EPO
Data publikacji	+
<i>ECLA</i> (Europejska Klasyfikacja Patentowa)	+
<i>IPC</i> (Międzynarodowa Klasyfikacja Patentowa)	+
Numer pierwszeństwa	+
Numer publikacji	+
Numer zgłoszenia	+
Słowa kluczowe w tekście	+
Słowa kluczowe w tytule/tytuł	+
Słowa kluczowe w tytule lub abstrakcie	+
Wynalazca	+
Zgłaszający	+
SUMA:	11

Tabela 38: Kryteria wyszukiwania w bazie *EPO*. Źródło: opracowanie własne.

Advanced Search

1. Database

Select patent database:

- Worldwide - full collection of published patent applications from 80+ countries
- EP - complete collection including full text of European published applications
- Worldwide - full collection of published patent applications from 80+ countries
- WIPO - complete collection including full text of PCT published applications

2. Search terms

Enter keywords in English - ctrl-enter expands the field you are in

Keyword(s) In title:	plastic and bicycle
Keyword(s) In title or abstract:	hair
Publication number:	WO2008014520
Application number:	DE19971031696
Priority number:	WO1995US15925
Publication date:	yyyyymmdd
Applicant(s):	Institut Pasteur
Inventor(s):	Smith
European Classification (ECLA):	F03G7/10
International Patent Classification (IPC):	H03M1/12

SEARCH CLEAR

Rysunek 72: Formularz wyszukiwania w bazie EPO. Źródło: Esp@cenet.

In my patents list | [Print](#) [Return to WO2009128702 \(A1\)](#)

BICYCLE PEDAL

Bibliographic data	Description	Claims	Mosaics	Original document	INPADOC legal status
Publication number: EP2282928 (A1) Publication date: 2011-02-16 Inventor(s): VAN EIJK STEFAN NICOLEY [NL] + Applicant(s): STRALA B V [NL] + Classification: - International: B62M3/08 - European: B62M3/08 Application number: EP20080873964 20081205 Priority number(s): WO2008NL50774 20081205; US20080046407P 20080419	Also published as: WO2009128702 (A1) US2011041648 (A1) Cited documents: WO2004098986 (A1) EP1533220 (A1) EP0485956 (A1) FR2861651 (A1) View all				
View INPADOC patent family View list of citing documents View document in the European Register	Report a data error here				
Abstract not available for EP 2282928 (A1) Abstract of corresponding document: WO 2009128702 (A1)					
<p>Bicycle pedal (1) comprising a pedal axle (2) for connection to a crank of a bicycle, a pedal body (6) comprising at least one actuation surface (8) for actuating the pedal, and connection means (12) connect the pedal body (6) to the pedal axle (2), such that the pedal body (6), while cycling, can freely rotate with respect to the pedal axle (2) about a first axis (A1) extending substantially parallel to the pedal axle (2), and a second axis (A2) extending substantially perpendicular to the pedal axle (2) and substantially parallel to the actuation surface (8), and wherein the second axis (A2) is located at a distance from, and in use above, the actuation surface (8).</p>					
Translate this text					
FIG. 1c					
Data supplied from the espacenet database — Worldwide					

Rysunek 73: Przykład widoku dokumentu w bazie EPO. Źródło: Esp@cenet.

Po wyświetleniu wybranego rekordu dostępne są szczegółowe dane dotyczące patentu. Oprócz wymienionych wyżej parametrów wyszukiwania, widoczne są patenty cytujące i cytowane, rysunki techniczne, charakterystyki, roszczenia, oryginalne formaty dokumentacji patentowej i status prawny udokumentowanego patentu.

W bazie nie można wykonywać analiz statystycznych, ani zapisywania historii poszukiwań, można jednak tworzyć listę ulubionych zakładek (Espacenet PL 2011b). Możliwe jest także eksportowanie danych do programu *MS Excel* (w formacie *csv*), a także wklejanie do plików *MS Word* (Espacenet PL 2011c) i drukowanie. Można także pobrać oryginał dokumentu w formacie *pdf* (Espacenet PL 2011d). Można także przejrzeć dany patent w Europejskim Rejestrze Patentowym, z którego wyciąg można wydrukować, zapisać w formacie *XML* lub *ST36*. Dostępne są także różne opcje przeglądania: ogólne informacje o dokumencie, status prawny, historia, cytowania, rodzina patentów.

Rysunek 74: Opcja eksportu dokumentu w bazie EPO. Źródło: Esp@cenet.

Jeżeli chodzi o listę wyszukiwania, można ją wydrukować lub eksportować do programu MS Excel. Pamiętać należy, że tylko pierwszych 500 wyników jest wyświetlanych - aby wyświetlić pozostałe wyniki, należy zawęzić kryteria wyszukiwania, co może stanowić utrudnienie wielu analiz.

Rysunek 75: Opcja eksportu listy wyszukiwania w bazie EPO. Źródło: Esp@cenet.

Wyszukiwanie za pomocą ECLA opiera się na systemie 9 indeksów (Espacenet 2011a):

- A. Potrzeby ludzkie,
- B. Operacje i transport,
- C. Chemia i metalurgia,
- D. Tekstylnia i papier,
- E. Konstrukcje stałe,
- F. Mechanika, oświetlenie, ogrzewanie, broń, silniki spalinowe i pompy,
- G. Fizyka,

H. Elektryczność,

Y. Ogólne oznaczenia dotyczące rozwoju nowych technologii.

Każda z tych grup jest rozbita na podgrupy, które również można wykorzystać w procesie wyszukiwania.

How do I use the Classification search? Get assistance

Search the European classification

Find classification(s) for keywords **syringe injection**

Index A B C D E F G H Y Previous: C40B99/00 Next: D01

TEXTILES; PAPER	D	<input type="checkbox"/>
NATURAL OR ARTIFICIAL THREADS OR FIBRES; SPINNING (metal threads B21 ; fibres or filaments of softened glass, minerals, or slag C03B37/00 ; yarns D02)	D01	<input type="checkbox"/>
YARNS; MECHANICAL FINISHING OF YARNS OR ROPES; WARPING OR BEAMING	D02	<input type="checkbox"/>
WEAVING	D03	<input type="checkbox"/>
BRAIDING; LACE-MAKING; KNITTING; TRIMMINGS; NON-WOVEN FABRICS	D04	<input type="checkbox"/>
SEWING; EMBROIDERING; TUFTING	D05	<input type="checkbox"/>
TREATMENT OF TEXTILES OR THE LIKE; LAUNDERING; FLEXIBLE MATERIALS NOT OTHERWISE	D06	<input type="checkbox"/>
	D07	<input type="checkbox"/>
	D21	<input type="checkbox"/>

show notes

Simply enter the code you were given, eg. B60N2/28, in the text field next to the index and click on GO. It is also possible to search at various levels in the hierarchy, eg. B60, B60N, etc.

Rysunek 76: Przykład wyszukiwania wg klasyfikacji ECLA w bazie EPO. Źródło: Esp@cenet.

Główna dziedzina nauki	Hasło po polsku	Hasło po angielsku	Wynik w bazie: EPO
Med, B, Ch	Alergia	allergy	374
Med, B, Ch	Złamanie otwarte	“compound fracture”	0
T, F, Mat	System operacyjny	“operating system”	1.825
T, F, Mat	Internet	Internet	6.443
T, F, Mat	Konstrukcja nośna	„supporting structure”	845
T, F, Mat	Wytrzymałość materiałów	„strength of materials”	1
Ch, B	Całkowita zawartość węgla organicznego	„total organic carbon content”	3
B, Ch	Sosna śródziemnomorska	„stone pine”	1
B, Ch	Dolina Nilu	„Nile Valley”	0
Ch, B	Destylacja ropy naftowej	“petroleum distillation”	2
E	Bezrobocie	unemployment	3
E	Teoria zarządzania	“theory of management”	0
H	Filologia klasyczna	„classical philology”	0
H	Edukacja młodzieży	“education of youth”	0
S	Przedstawienie teatralne	„theater performance”	0
S	Obraz olejny	„oil painting”	17

Tabela 39: Przykładowe wyniki wyszukiwania w bazie EPO dla kilku haseł z wybranych obszarów nauki (w polu: słowa kluczowe w tytule i abstrakcie) w roku 2010. Źródło: opracowanie własne. Legenda: Skrót literowy przy hasłach odpowiada dziedzinom nauki z polskiej klasyfikacji nauk (CK 2005), która na potrzeby niniejszego opracowania została nieznacznie zmodyfikowana:

- B – nauki biologiczne,
- Ch – nauki chemiczne,
- E – nauki ekonomiczne,
- F – nauki fizyczne,
- H – nauki humanistyczne,
- Mat – nauki matematyczne,
- Med – nauki medyczne,
- T – nauki techniczne,
- S – sztuka.

Kryterium	EPO
Interfejs	Polski, angielski i w innych językach
Obszar, z którego pochodzą dane	Europa (patenty europejskie), świat (patenty <i>WIPO</i>)
Dane nie-patentowe	Dostępne (oznaczenie <i>XP</i>)
Indeksowanie	<i>ECLA</i> (szersza niż <i>IPC</i>)
Cytowania	Dokumenty cytujące i cytowane
Status prawny danych	<i>INPADOC</i> (International Patent Documentation Center)
Klasyfikacje danych	<i>IPC</i> i <i>ECLA</i>
Maksymalna liczba rekordów na liście	500
Eksport danych	Do 500 rekordów
Format danych eksportowanych	<i>csv</i> (<i>MS Excel</i>), <i>pdf</i>
Narzędzia analizy danych	Brak
Koszty	Darmowe
Słabe strony	Limity zapisu i eksportu danych, brak historii wyszukiwania, brak funkcji wyszukiwania pełnych tekstów
Silne strony	Dane ogólnosiwiatowe, darmowe pobieranie dokumentacji patentowej, określony prawny status danych, wiarygodne źródła danych

Tabela 40: Podsumowanie bazy EPO. Źródło: opracowanie własne, przy wykorzystaniu danych z: Intellogist (2011a).

10.9. Opis bazy Urzędu Patentowego Stanów Zjednoczonych (*USPTO*)

Baza patentowa *USPTO* (*The United States Patent and Trademark Office* – Urząd Patentowy Stanów Zjednoczonych) umożliwia wyszukiwanie dwóch rodzajów dokumentów patentowych.

Podstawowy sposób wyszukiwania oferuje baza *PatFT* (*Patent Full-Text and Image Database* – pełnotekstowa i pełnoobrazkowa baza patentów), w której zgromadzona jest pełna dokumentacja patentowa od 1976 a także zeskanowane wersje dokumentów starszych od roku 1790. Patenty z okresu 1970 – 1976 mogą być wyszukiwane za pomocą kryteriów daty, numeru patentowego i klasyfikacji patentowej Stanów Zjednoczonych. Jeżeli chodzi o patenty od 1976, dostępny jest dodatkowo szerszy zakres kryteriów: tytuł, abstrakt, data, międzynarodowa klasyfikacja patentowa i wiele innych (*USPTO 2011a*), łącznie 31 kryteriów (*USPTO 2011b*). Możliwe jest użycie operatorów logicznych, symboli wieloznacznych (\$) a także określonych fraz (z użyciem znaków „”) (*USPTO 2011c*). Nie są indeksowane tzw. *stopwords*, czyli słowa, które najczęściej pojawiają się w abstraktach, tytułach, opisach i innych polach wyszukiwania. Chodzi o ciągi znaków takie jak: „*a*”, „*the*”, „*that*”, „*of*” itp. (*USPTO 2011d*).

Z kolei przez bazę *AppFT* (*Patent Application Full-Text and Image Database*) można wyszukiwać opublikowane zgłoszenia patentowe od 2001 roku. Wyszukiwanie może odbywać się w oparciu o 24 kryteria, z wykorzystaniem tych samych metod co powyżej.

Kryterium	USPTO
Abstrakt	+
Asystujący egzaminator	+
Dane powiązane	+
Dane powiązane ze zgłoszeniem (USA)	+
Data publikacji	+
Data reedycji	+
Data zgłoszenia	+
Informacje <i>PCT</i>	+
Inne źródła cytowane	+
Interesy rządowe	+
<i>IPC</i> (Międzynarodowa Klasyfikacja Patentowa)	+
Klasyfikacja US	+
Kraj właściciela	+
Kraj wynaleźnia	+
Miasto właściciela	+
Miasto wynaleźnia	+
Nazwisko właściciela	+
Numer patentu	+
Numer zgłoszenia	+
Opis/specyfikacja	+
Patenty cytowane	+
Patenty cytujące	+
Pierwszy egzaminator	+
Prawnik/agent	+
Priorytet roszczeń w innych krajach	+
Roszczenia	+
Stan właściciela	+
Stan wynaleźnia	+
Typ zgłoszenia/patentu	+
Tytuł	+
Wynalazca	+
SUMA:	31

Tabela 41: Kryteria wyszukiwania w bazie *USPTO*. Źródło: opracowanie własne.

Patents from 1790 through 1975 are searchable only by Issue Date, Patent Number, and Current US Classification.
When searching for specific numbers in the Patent Number field, patent numbers must be seven characters in length, excluding commas, which are optional.

Field Code	Field Name	Field Code	Field Name
PN	Patent Number	IN	Inventor Name
ISD	Issue Date	IC	Inventor City
TTL	Title	IS	Inventor State
ABST	Abstract	ICN	Inventor Country
ACLM	Claim(s)	LREP	Attorney or Agent
SPEC	Description/Specification	AN	Assignee Name
CCL	Current US Classification	AC	Assignee City
ICL	International Classification	AS	Assignee State
APN	Application Serial Number	ACN	Assignee Country
APD	Application Date	EXP	Primary Examiner
PARN	Parent Case Information	EXA	Assistant Examiner
RLAP	Related US App. Data	REF	Referenced By
REIS	Reissue Data	FREF	Foreign References
PRIR	Foreign Priority	OREF	Other References
PCT	PCT Information	GOVT	Government Interest
APT	Application Type		

Rysunek 77: Kryteria wyszukiwania w bazie USPTO. Źródło: USPTO.

Kryterium	USPTO
Zawartość bazy (tekstowa/graficzna)	W bazie dostępne są zasoby tekstowe i graficzne
Pokrycie obszarów nauki	Baza zawiera szerokie zasoby dokumentacji patentowej z możliwością selekcjonowania rodzaju dokumentacji
Jakość danych	Dane pochodzą z zasobów urzędowych (USA)
Struktura	Baza posiada jasną i prostą strukturę, choć interfejs nie jest przyjazny dla użytkownika

Tabela 42: Charakterystyka bazy patentów USPTO. Źródło: opracowanie własne.

USPTO PATENT FULL-TEXT AND IMAGE DATABASE

(1 of 22129)

United States Patent D633,416
Teal, et al. March 1, 2011

Elliptical *bicycle* frame

Claims

CLAIM The ornamental design for an elliptical *bicycle* frame, as shown and described.

Inventors: Teal; Brent (Solana Beach, CA), Pate; Bryan (Atherton, CA)
Assignee: PT Motion Works, Inc. (Solana Beach, CA)
Appl. No.: D/357,779
Filed: March 17, 2010

Current U.S. Class: D12/111

Rysunek 78: Opcja przeglądania rysunków w bazie USPTO. Źródło: USPTO.

Baza USPTO udostępnia możliwość przeglądania zawartości dokumentów patentowych, natomiast pobieranie kopii dokumentacji patentowej jest możliwe dzięki narzędziu PAIR

(*Patent Application Information Retrieval*) (USPTO 2011e) i jest to usługa płatna. Należy zaznaczyć, że korzystanie z niej nie wydaje się niezbędne w przypadku analiz bibliometrycznych, gdyż dotyczy głównie papierowych wersji pliku, podczas gdy wersje *pdf* można nieodpłatnie pobrać z serwisu *Google Patents*.

Rysunek 79: Opcja odpłatnego pobrania dokumentacji w bazie USPTO. Źródło: USPTO.

Utility Patent					
A utility patent is granted to the inventor or discoverer of any new and useful process, machine, manufacture, composition of matter, or any new and useful improvement thereof. A utility patent is the most common type of patent issued by the U.S. Patent and Trademark Office.					
Certification Statement	Document Media	Processing Time	Delivery Method	Cost	Notes
	Paper	10 days	Mail	\$25.00 per copy	
	Paper	10 days	PSF Box	\$25.00 per copy	
	PDF File	1 day	Internet	\$3.00 per copy	
	Paper	5 days	Mail	\$3.00 per copy	
	Paper	5 days	PSF Box	\$3.00 per copy	

Rysunek 80: Opcje cen dokumentacji w bazie USPTO. Źródło: USPTO.

Nie ma możliwości eksportu listy wyszukiwania ani dokumentacji patentowej, sam interfejs bazy - choć prosty - jest bardzo ubogi, a wyszukiwania zaawansowanego można dokonywać jedynie z użyciem specjalistycznej składni, której użycie jest nieintuicyjne. Nie można także w łatwy sposób wyszukiwać, zawężając zakresy dat albo wskazując rok publikacji patentu, chyba że wykorzystana zostanie specjalistyczna składnia zapytań. Wielość kryteriów wyszukiwania w połączeniu z nieintuicyjną metodą wyszukiwania może powodować pewne trudności w wykorzystaniu całkowitego potencjału bazy.

USPTO PATENT FULL-TEXT AND IMAGE DATABASE

[Home](#) [Quick](#) [Advanced](#) [Pat Num](#) [Help](#)

[View Cart](#)

Data current through March 1, 2011.

Query [\[Help\]](#)

Select Years [\[Help\]](#)

1976 to present [full-text] ▼

1976 to present [full-text]

1790 to present [entire database]

Examples:

tfl/(tennis and (racquet or racket))

isd/1/8/2002 and motorcycle

in/newmar-julie

Patents from 1790 through 1975 are searchable only by Issue Date, Patent Number, and Current US Classification.
When searching for specific numbers in the Patent Number field, patent numbers must be seven characters in length, excluding commas, which are optional.

Rysunek 81: Formularz wyszukiwania zaawansowanego w bazie USPTO. Źródło: USPTO.

Pomoc w bazie tylko w nieznacznym sposób ułatwia poszukiwania (por. Rysunek 82).

Application Date (APD)

This field contains the date when a complete application was received by the US Patent and Trademark Office, following receipt of all filing material requirements.

TIP: You can use one of three formats to search in any of the date fields :

1. <four digit year><two digit month><two digit day>. For example, use 19960103 to retrieve documents filed on January 3, 1996.
2. <Month>-<Day>-<Year>.

Month can be the number of the month, the full name of the month, or an abbreviation (i.e. 1, January or Jan).

Day must be a number between 1 and 31 or a \$.

Year must be a four digit year (1996).

Example: Use 1-3-1996, Jan-3-1996 or January-3-1996 to retrieve documents filed on January 3, 1996.

3. <Month>/<Day>/<Year>.

Month can be the number of the month, the full name of the month, or an abbreviation (i.e. 1, January or Jan).

Day must be a number between 1 and 31 or a \$.

Year must be a four digit year (1996).

Example: Use 1/3/1996 or Jan/3/1996 to retrieve documents filed on January 3, 1996.

You can search for a range of dates using the -> operator.

Example: To search for patents with application dates between January 1, 1995, and Feb 14, 1995, use APD/1/1/1995->2/14/1995.

Rysunek 82: Przykład okna pomocy w bazie USPTO. Źródło: USPTO.

Chcąc przykładowo wyszukać patenty dla słowa w tytule „alergia” z datą publikacji 2010 należy wpisać w oknie wyszukiwania zaawansowanego formułę:

TTL/(allergy)AND ISD/1/\$/2010

USPTO PATENT FULL-TEXT AND IMAGE DATABASE

[Home](#) [Quick](#) [Advanced](#) [Pat Num](#) [Help](#)
[View Cart](#)

Data current through March 1, 2011.

Query [\[Help\]](#)

Term 1: in Field 1:
AND

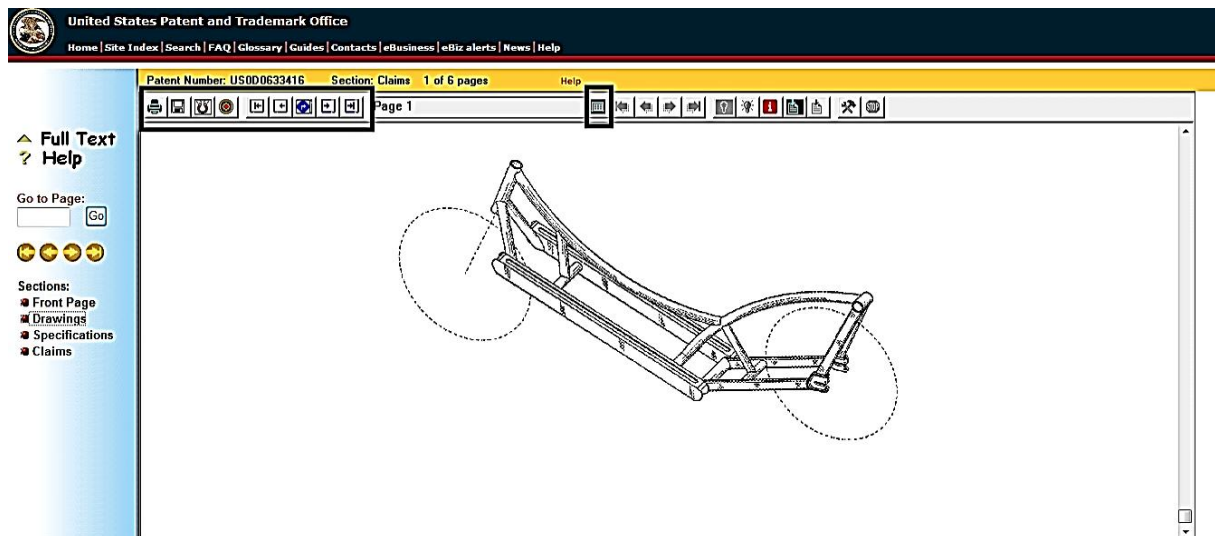
Term 2: in Field 2:

Select years [\[Help\]](#)

Patents from 1790 through 1975 are searchable only by Issue Date, Patent Number, and Current US Classification.
When searching for specific numbers in the Patent Number field, patent numbers must be seven characters in length, excluding commas, which are optional.

Rysunek 83: Przykład wyszukiwania w bazie USPTO. Źródło: USPTO.

Problematyczne jest także przeglądanie oryginałów dokumentacji patentowej i rysunków. Sugerowane przez stronę USPTO wtyczki do przeglądarek nie zawsze działają (na niektórych komputerach nie działa przykładowo przeglądarka do popularnych plików pdf), inne wymagają korzystania z programu Microsoft Internet Explorer i nie działają pod innymi przeglądarkami internetowymi. Po prawidłowej instalacji wtyczki (np. Internetiff), dostępny jest podgląd rysunków a także pewne dodatkowe opcje: druku, zapisu (w formacie TIFF lub pdf) oraz przeglądania listy wyników.



Rysunek 84: Opcja przeglądania pełnej dokumentacji w bazie USPTO. Źródło: USPTO.

Główna dziedzina nauki	Hasło po polsku	Hasło po angielsku	Wynik w bazie: USPTO
Med, B, Ch	Alergia	<i>allergy</i>	67
Med, B, Ch	Złamanie otwarte	<i>“compound fracture”</i>	0
T, F, Mat	System operacyjny	<i>“operating system”</i>	1076
T, F, Mat	Internet	<i>Internet</i>	2404
T, F, Mat	Konstrukcja nośna	<i>„supporting structure”</i>	94
T, F, Mat	Wytrzymałość materiałów	<i>„strength of materials”</i>	10
Ch, B	Całkowita zawartość węgla organicznego	<i>„total organic carbon content”</i>	3
B, Ch	Sosna śródziemnomorska	<i>„stone pine”</i>	0
B, Ch	Dolina Nilu	<i>„Nile Valley”</i>	0
Ch, B	Destylacja ropy naftowej	<i>“petroleum distillation”</i>	1
E	Bezrobocie	<i>unemployment</i>	3
E	Teoria zarządzania	<i>“theory of management”</i>	0
H	Filologia klasyczna	<i>„classical philology”</i>	0
H	Edukacja młodzieży	<i>“education of youth”</i>	0
S	Przedstawienie teatralne	<i>„theater performance”</i>	1
S	Obraz olejny	<i>„oil painting”</i>	2

Tabela 43: Przykładowe wyniki wyszukiwania w bazie USPTO dla kilku haseł z wybranych obszarów nauki (w polu: „*all fields*”) w roku 2010. Źródło: opracowanie własne. Legenda: Skróty literowe przy hasłach odpowiadają dziedzinom nauki z polskiej klasyfikacji nauk (CK 2005), która na potrzeby niniejszego opracowania została nieznacznie zmodyfikowana:

- B – nauki biologiczne,
- Ch – nauki chemiczne,
- E – nauki ekonomiczne,
- F – nauki fizyczne,
- H – nauki humanistyczne,
- Mat – nauki matematyczne,
- Med – nauki medyczne,
- T – nauki techniczne,
- S – sztuka.

Kryterium	USPTO
Interfejs	Angielski
Obszar, z którego pochodzą dane	USA
Dane nie-patentowe	Brak
Indeksowanie	<i>US Classification, IPC</i>
Cytowania	Dokumenty cytujące i cytowane
Status prawny danych	Dane właściciela patentu
Klasyfikacje danych	<i>IPC, USC</i>
Maksymalna liczba rekordów na liście	Brak ograniczeń
Eksport danych	brak opcji, zapis plików po podglądzie w opcji „ <i>images</i> ” lub po zakupie (możliwy zapis strony <i>HTML</i> z patentem)
Format danych eksportowanych	<i>pdf, TIFF</i>
Narzędzia analizy danych	Brak
Koszty	Darmowe przeglądanie i wyszukiwanie, pełna dokumentacja płatna
Słabe strony	Częściowo płatny dostęp do dokumentów, niejasny i skomplikowany składniowo interfejs wyszukiwania, niekompatybilność bazy w konfrontacji z użyciem opcji przeglądania graficznej wersji dokumentacji
Silne strony	Dane z obszaru USA, wiarygodność zapewniona przez afiliację <i>USPTO</i>

Tabela 44: Podsumowanie bazy *USPTO*. Źródło: opracowanie własne.

10.10. Opis bazy *Google Patents*

Firma *Google* na podstawie umowy z *USPTO* stworzyła bazę *Google Patents* (www.google.com/patents). Wszystkie prezentowane w niej dokumenty patentowe pochodzą z *USPTO*. Do bazy wprowadzono ponad siedem milionów patentów amerykańskich, nie ma patentów międzynarodowych. Dokumenty można przeglądać i pobierać w formacie *HTML* oraz *pdf*. Możliwe jest wyszukiwanie w oparciu o następujący zestaw kryteriów: numer patentu, wynalazca, nabywca (właściciel praw patentowych), patentowa klasyfikacja amerykańska i międzynarodowa i daty. Możliwe jest także używanie operatorów „-” (wyłączenie), „or” („lub”), a także wyszukiwanie fraz (z użyciem cudzysłówów) (*Google Patents 2011a*). Można także sortować wyniki wyszukiwania według dat lub istotności (*Google Patents 2011b*).

Kryterium	Google Patents
Data publikacji	+
Data zgłoszenia	+
Dokładna fraza w tekście	+
IPC (Międzynarodowa Klasyfikacja Patentowa)	+
Klasyfikacja US	+
Nazwisko właściciela	+
Numer patentu	+
Słowa kluczowe w tekście	+
Słowa kluczowe w tytule/tytuł	+
Typ zgłoszenia/patentu	+
Wynalazca	+
SUMA:	11

Tabela 45: Kryteria wyszukiwania w bazie *Google Patents*. Źródło: opracowanie własne.Rysunek 85: Formularz wyszukiwania podstawowego w bazie *Google Patents*. Źródło: *Google Patents*.

The image shows the Google Patents Advanced Patent Search interface. It features a search bar with four options: "with all of the words", "with the exact phrase", "with at least one of the words", and "without the words". There are input fields for each option. To the right of the search bar, it says "10 results" and has a "Google Search" button. Below the search bar, there are several sections for filtering results: "Patent number", "Title", "Inventor", "Assignee", "U.S. Classification", "International Classification", "Document status" (with checkboxes for "Issued patents" and "Applications"), "Patent type" (with a dropdown menu showing options like "All types", "Utility", "Design (D)", "Plant (PP)", "Defensive publication (T)", "Statutory invention registration (H)", "Reissue design (RD)", "Reissue utility (RE)", "X patents (X)", "X patent reissue (RX)", and "Additional improvements (AI)"), "Issue date" (with radio buttons for "Return patents issued anytime" and "Return patents issued between" and date input fields), and "Filing date" (with radio buttons for "Return patents filed anytime" and "Return patents filed between" and date input fields). The copyright notice "©2011 Google" is visible at the bottom.

Rysunek 86: Formularz wyszukiwania zaawansowanego w bazie *Google Patents*. Źródło: *Google Patents*.

Kryterium	<i>Google Patents</i>
Zawartość bazy (tekstowa/graficzna)	W bazie dostępne są zasoby tekstowe i graficzne
Pokrycie obszarów nauki	Baza zawiera szerokie zasoby dokumentacji patentowej z możliwością selekcjonowania rodzaju dokumentacji
Jakość danych	Dane pochodzą z zasobów urzędowych (z USPTO)
Struktura	Baza posiada jasną i prostą strukturę, interfejs przyjazny dla użytkownika

Tabela 46: Charakterystyka bazy patentów *Google Patents*. Źródło: opracowanie własne.

Proces wyszukiwania w bazie *Google Patents* jest prosty i intuicyjny, oparty na zasadach przeszukiwania zasobów sieci internetowej. Wyszukiwanie zaawansowane jest proste i jasne a wskazówki pomocy w bazie zrozumiałe. *Google Patents* zawiera zasoby bazy *USPTO* w bardziej przystępnej formie, choć kryteriów wyszukiwania jest mniej. Mimo tego, sam proces wyszukiwania może być bardziej efektywny ze względu na łatwość obsługi wyszukiwarki. Lista wyników wyszukiwania wygląda jak w zwykłej wyszukiwarce internetowej.

Google patents

Advanced Patent Search

Patents

- > **List view**
 - [Cover view](#)
- > **Sort by relevance**
 - [Sort by date \(new first\)](#)
 - [Sort by date \(old first\)](#)
- > **Any status**
 - [Issued patents](#)
 - [Applications](#)

[APPLICATION] **Bicycle saddle assembly**
 US Pat. 11521555 - Filed Sep 15, 2006
 Field of the Invention [0002] The present invention relates generally to bicycle saddles and more particularly to a bicycle saddle assembly that includes a ...

[APPLICATION] **Bicycle improved with two front wheels**
 US Pat. 11253335 - Filed Oct 19, 2005
 BACKGROUND OF THE INVENTION [0004] This invention pertains to a bicycle with two front wheels instead of a single wheel in front; and it includes all other ...

Bicycle
 US Pat. 4053173 - Filed Mar 23, 1976
 ... Sheridan & Sprinkle [57] ABSTRACT A bicycle of the type comprising a frame and a front and rear wheel rotatably mounted to the frame is provided with ...

Bicycle headlamp
 US Pat. D597237 - Filed Sep 13, 2007 - Princeton Tectonics, Inc.
 cited by examiner Primary Examiner — Clare E Heflin (74) Attorney, Agent, or Firm — Howson & Howson LLP (57) CLAIM The ornamental design for a bicycle ...

Bicycle lock
 US Pat. 5092142 - Filed Feb 11, 1991 - Kryptonite Corporation
 Benson [57] An improved ABSTRACT bicycle lock, featuring a U-shaped shackle, a cross bar and means for securing one to the other, is disclosed. ...

Rysunek 87: Przykład listy wyników wyszukiwania w bazie Google Patents. Źródło: Google Patents.














Można także uzyskać wyniki w postaci zestawienia rysunków zawartych w dokumentach patentowych, a także sortować rezultaty według istotności i daty, jak również wyświetlać tylko zgłoszenia lub tylko patenty.

Google patents

Advanced Patent Search

Patents Patents 1 - 30 on bicycle. (0.15 seconds)

- List view
- Cover view
- Sort by relevance
- Sort by date (new first)
- Sort by date (old first)
- Any status
- Issued patents
- Applications

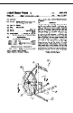
 Bicycle saddle assembly Filed 2008APPLICATION	 Bicycle improved with two front wheels Filed 2005APPLICATION	 Bicycle Issued 1977	 Bicycle headlamp Assignee: Princeton Tectonics, Inc. Issued 2009	 Stay up to date on these results using the patents RSS feed on bicycle .
 Bicycle lock Assignee: Kryptonite Corporation Issued 1992	 Bicycle headlight Assignee: Cateye Co., Ltd. Issued 2002	 Bicycle light Assignee: Cateye Co., Ltd. Issued 2010	 Bicycle wheel Assignee: Shimano, Inc. Issued 2000	
 Bicycle front derailleur Assignee: Shimano Inc. Filed 2004APPLICATION	 Bicycle pedal cover Issued 2009	 Road Bicycle Handlebar Filed 2009APPLICATION	 Bicycle pedal Assignee: Shimano, Inc. Issued 1999	

Rysunek 88: Przykład graficznej listy wyników w bazie Google Patents. Źródło: Google Patents.

Nie można eksportować listy wyników, można ją tylko przeglądać. Po wyświetleniu poszukiwanego patentu, widok dokumentu jest jasny i intuicyjny. Dostępne są wszystkie informacje zawarte w kryteriach wyszukiwania, można przeglądać oryginalną dokumentację patentową, zapisać plik w formacie pdf lub przejrzeć dokument w bazie USPTO. Wyświetla się także lista dokumentów cytowanych i cytujących. Warto zaznaczyć, że w Google Patents wyszukiwane są jednocześnie zgłoszenia i patenty, podczas gdy w USPTO służą do tego odrębne bazy i nie można przeszukiwać tych dwóch rodzajów dokumentacji jednocześnie.

Bicycle Douglas Chase et al

Overview
 Abstract
 Drawing
 Description
 Claims



Patent number: 4953173
 Filing date: Mar 23, 1976
 Issue date: Oct 11, 1977

Go

A bicycle of the type comprising a frame and a front and rear wheel rotatably mounted to the frame is provided with improved means for rotatably driving the rear wheel. The driving means comprises a right and left foot lever pivotally connected at one end to the frame and adapted to be alternately...

Inventors: Douglas Chase, Sr.
 Primary Examiner: Terrance L. Siemens

Read this patent
 Download PDF

U.S. Classification
 280/253; 280/238; 280/257

International Classification
 B62M 104

View patent at USPTO

Citations

Patent Number	Title	Issue date
891872	PORATION	Jun 30, 1908
1503878	CHILD S VEHICLE	Aug 5, 1924
1808332	CHILD S VEHICLE	Nov 9, 1926
1844838	PEDAL-PROPELLED SCOOTER	Oct 11, 1927

Referenced by

Patent Number	Title	Issue date
5527246	Mobile exercise apparatus	Jun 18, 1996
5529555	Crank assembly for an exercising device	Jun 25, 1996
5540837	Stationary exercise apparatus having a preferred foot platform orientation	Jul 30, 1996
5549528	Stationary exercise apparatus	Aug 27, 1996
5573480	Stationary exercise apparatus	Nov 12, 1996

Rysunek 89: Widok przykładowego dokumentu w bazie *Google Patents*. Źródło: *Google Patents*.

Główna dziedzina nauki	Hasło po polsku	Hasło po angielsku	Wynik w bazie: Google Patents
Med, B, Ch	Alergia	<i>allergy</i>	560
Med, B, Ch	Złamanie otwarte	<i>“compound fracture”</i>	3
T, F, Mat	System operacyjny	<i>“operating system”</i>	560
T, F, Mat	Internet	<i>Internet</i>	582
T, F, Mat	Konstrukcja nośna	<i>„supporting structure”</i>	589
T, F, Mat	Wytrzymałość materiałów	<i>„strength of materials”</i>	32
Ch, B	Całkowita zawartość węgla organicznego	<i>„total organic carbon content”</i>	11
B, Ch	Sosna śródziemnomorska	<i>„stone pine”</i>	1
B, Ch	Dolina Nilu	<i>„Nile Valley”</i>	0
Ch, B	Destylacja ropy naftowej	<i>“petroleum distillation”</i>	13
E	Bezrobocie	<i>unemployment</i>	69
E	Teoria zarządzania	<i>“theory of management”</i>	0
H	Filologia klasyczna	<i>„classical philology”</i>	0
H	Edukacja młodzieży	<i>“education of youth”</i>	0
S	Przedstawienie teatralne	<i>„theater performance”</i>	17
S	Obraz olejny	<i>„oil painting”</i>	10

Tabela 47: Przykładowe wyniki wyszukiwania w bazie *Google Patents* dla kilku haseł z wybranych obszarów nauki (w polu: „*exact phrase*”) w 2010 roku. Źródło: opracowanie własne. Legenda: Skrótów literowe przy hasłach odpowiadają dziedzinom nauki z polskiej klasyfikacji nauk (CK 2005), która na potrzeby niniejszego opracowania została nieznacznie zmodyfikowana:

- B – nauki biologiczne,
- Ch – nauki chemiczne,
- E – nauki ekonomiczne,
- F – nauki fizyczne,
- H – nauki humanistyczne,
- Mat – nauki matematyczne,
- Med – nauki medyczne,
- T – nauki techniczne,
- S – sztuka.

Kryterium	Google Patents
Interfejs	Angielski
Obszar, z którego pochodzą dane	USA
Dane nie-patentowe	Brak
Indeksowanie	<i>US Classification, IPC</i>
Cytowania	Dokumenty cytujące i cytowane
Status prawny danych	Dane właściciela patentu, statusu prawnego dokumentów brak
Klasyfikacje danych	<i>IPC, USC</i>
Maksymalna liczba rekordów na liście	Około 600
Eksport danych	brak opcji, jedynie zapis plików po podglądzie w opcji w formacie pdf
Format danych eksportowanych	<i>pdf</i>
Narzędzia analizy danych	Brak
Koszty	Darmowe
Słabe strony	Stosowana w bazie klasyfikacja <i>USC</i> nie jest wersją najnowszą, brak bezpośredniego licznika wszystkich wyszukanych rekordów (konieczność przewijania do końca listy, aby sprawdzić całkowitą liczbę wyszukanych danych), maksymalnie tylko około 600 rekordów jest wyświetlanych, nie można zapisać wyników wyszukiwania ani ich bezpośrednio eksportować.
Silne strony	Dane z obszaru USA, wiarygodność zapewniona przez afiliację <i>USPTO</i> , intuicyjny i łatwy w obsłudze interfejs i pobieranie plików dokumentacji w pdf.

Tabela 48: Podsumowanie bazy Google Patents. Źródło: opracowanie własne przy wykorzystaniu Intellogist (2011b).

10.11. Porównanie funkcjonalności baz patentów

Poniższy podrozdział zawiera tabelaryczne zestawienia funkcjonalności poszczególnych baz patentów, które mogą ułatwić wybór odpowiedniego narzędzia dla potrzeb analiz bibliometrycznych. Warto pamiętać, że o wyborze bazy decydują przede wszystkim:

- zasięg terytorialny,
- możliwość identyfikacji danych (taksonomie, ułatwiające wyszukiwanie),
- łatwość eksportu danych w celu przetwarzania w dedykowanym oprogramowaniu.

Kryterium	<i>EPO</i>	<i>USPTO</i>	<i>Google Patents</i>
Abstrakt		+	
Asystujący egzaminator		+	
Dane powiązane		+	
Dane powiązane ze zgłoszeniem (USA)		+	
Data publikacji	+	+	+
Data reedycji		+	
Data zgłoszenia		+	+
Dokładna fraza w tekście			+
<i>ECLA</i> (Europejska Klasyfikacja Patentowa)	+		
Informacje <i>PCT</i>		+	
Inne źródła cytowane		+	
Interesy rządowe		+	
<i>IPC</i> (Międzynarodowa Klasyfikacja Patentowa)	+	+	+
Klasyfikacja US		+	+
Kraj właściciela		+	
Kraj wynaleźnia		+	
Miasto właściciela		+	
Miasto wynaleźnia		+	
Nazwisko właściciela		+	+
Numer patentu		+	+
Numer pierwszeństwa	+		
Numer publikacji	+		
Numer zgłoszenia	+	+	
Opis/specyfikacja		+	
Patenty cytowane		+	
Patenty cytujące		+	
Pierwszy egzaminator		+	
Prawnik/agent		+	
Priorytet roszczeń w innych krajach		+	
Roszczenia		+	
Słowa kluczowe w tekście	+		+
Słowa kluczowe w tytule/tytuł	+		+
Słowa kluczowe w tytule lub abstrakcie	+		
Stan właściciela		+	
Stan wynaleźnia		+	
Typ zgłoszenia/patentu		+	+
Tytuł		+	

Kryterium	EPO	USPTO	Google Patents
Wynalazca	+	+	+
Zgłaszający	+		
SUMA:	11	31	11

Tabela 49: Kryteria wyszukiwania – zestawienie zbiorcze. Źródło: opracowanie własne.

Główna dziedzina nauki	Hasło po polsku	Hasło po angielsku	EPO	USPTO	Google Patents
Med, B, Ch	Alergia	<i>allergy</i>	374	67	560
Med, B, Ch	Złamanie otwarte	<i>“compound fracture”</i>	0	0	3
T, F, Mat	System operacyjny	<i>“operating system”</i>	1.825	1076	560
T, F, Mat	Internet	<i>Internet</i>	6.443	2404	582
T, F, Mat	Konstrukcja nośna	<i>„supporting structure”</i>	845	94	589
T, F, Mat	Wytrzymałość materiałów	<i>„strength of materials”</i>	1	10	32
Ch, B	Całkowita zawartość węgla organicznego	<i>„total organic carbon content”</i>	3	3	11
B, Ch	Sosna śródziemnomorska	<i>„stone pine”</i>	1	0	1
B, Ch	Dolina Nilu	<i>„Nile Valley”</i>	0	0	0
Ch, B	Destylacja ropy naftowej	<i>“petroleum distillation”</i>	2	1	13
E	Bezrobocie	<i>unemployment</i>	3	3	69
E	Teoria zarządzania	<i>“theory of management”</i>	0	0	0
H	Filologia klasyczna	<i>„classical philology”</i>	0	0	0
H	Edukacja młodzieży	<i>“education of youth”</i>	0	0	0
S	Przedstawienie teatralne	<i>„theater performance”</i>	0	1	17
S	Obraz olejny	<i>„oil painting”</i>	17	2	10
Liczba najwyższych wyników wyszukiwania:			5	0	8
Bezwzględny wynik wyszukiwań:			9.514	3.661	2.447
Miejsce ze względu na wynik bezwzględny:			I	II	III
Udział procentowy (100% = całkowita suma wyników = 15.622)			60,90%	23,43%	15,66%

Tabela 50: Przykładowe wyniki wyszukiwania w bazach dla kilku haseł z wybranych obszarów nauki (w najszerszym polu wyszukiwania) w 2010. Źródło: opracowanie własne. Legenda: Skrót literowy przy hasłach odpowiada dziedzinom nauki z polskiej klasyfikacji nauk (CK 2005), która na potrzeby niniejszego opracowania została nieznacznie zmodyfikowana:

- B – nauki biologiczne,
- Ch – nauki chemiczne,
- E – nauki ekonomiczne,

- F – nauki fizyczne,
- H – nauki humanistyczne,
- Mat – nauki matematyczne,
- Med – nauki medyczne,
- T – nauki techniczne,
- S – sztuka.

10.12. Przegląd komercyjnych baz patentów

W tej części książki zostaną krótko scharakteryzowane niektóre bazy patentów, dostępne na zasadach komercyjnych. Charakterystyka dotyczy następujących baz:

- *Delphion*,
- *Micro Patent / Patent Web*,
- *Pat Base*,
- *QPat*,
- *WIPS Global*,
- *Total Patent*,
- *Thomson Innovation*,
- *SureChem*,
- *JP-NETe*,
- *Patent Integration*.

Analiza powyższych baz dotyczy przede wszystkim trzech kryteriów:

- Źródła danych,
- Koszty użytkowania,
- Silne strony i ograniczenia bazy w porównaniu do baz bezpłatnych.

10.12.1. *Delphion*

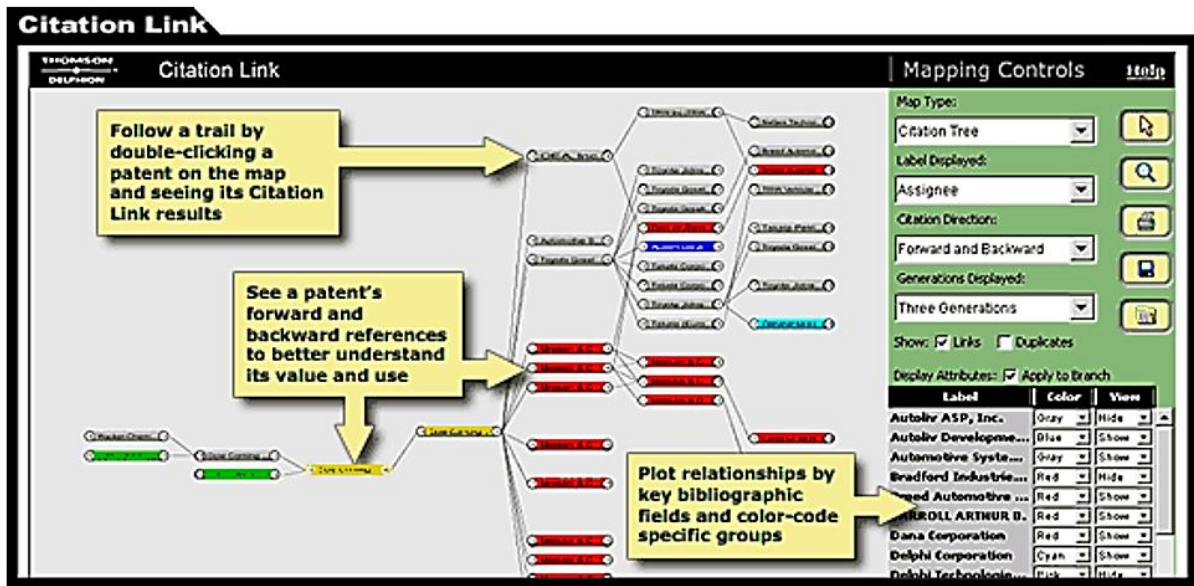
Źródła danych: *Delphion* jest amerykańską bazą patentów afiliowaną przez *Thomson Reuters*, interfejs wyszukiwania jest w języku angielskim oraz japońskim. W bazie dostępne są zasoby 4 urzędów patentowych – *USA*, *EPO*, *WIPO/PCT* i *German Patent and Trademark Office*. Dostępne są także abstrakty z *Japan Patent Information Organization* oraz dokumenty graficzne z *Swiss Patent Office*.

Koszty użytkowania: Dostępne są dwa rodzaje subskrypcji:

1. *Unlimited* – 308\$ miesięcznie
2. *Premier* – 154\$ miesięcznie

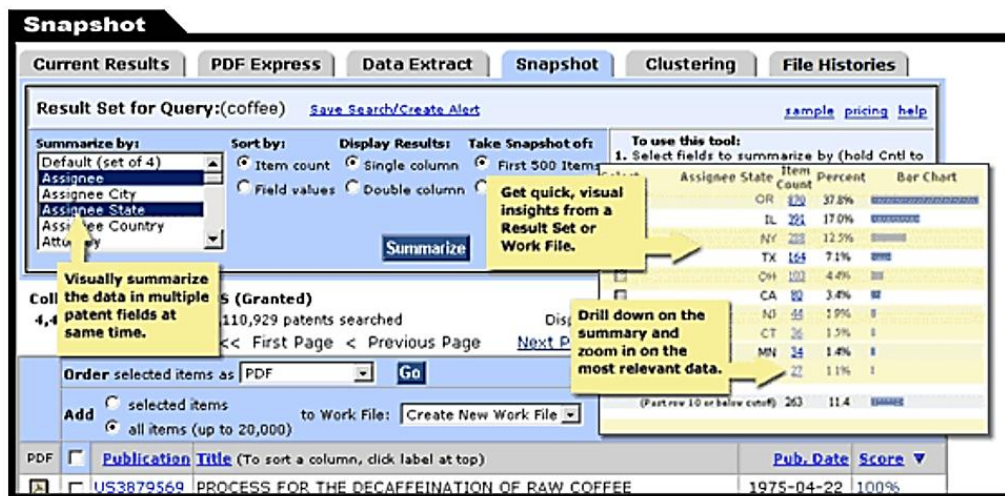
Silne strony i ograniczenia bazy w porównaniu do baz bezpłatnych: Baza *Delphion* oferuje zestaw narzędzi analizy danych, przede wszystkim:

- Graficzne mapowanie odnośników cytujących i cytowanych (*Citation Link*)



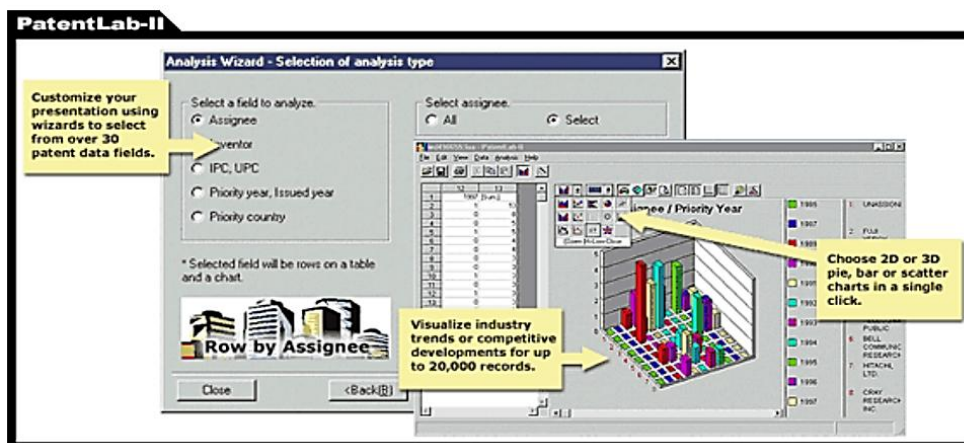
Rysunek 90: Mapowanie cytowań w bazie *Delphion*. Źródło: strona internetowa *Delphion*.

- Analizę wyników wyszukiwania za pomocą wykresów słupkowych (*Snapshot*)



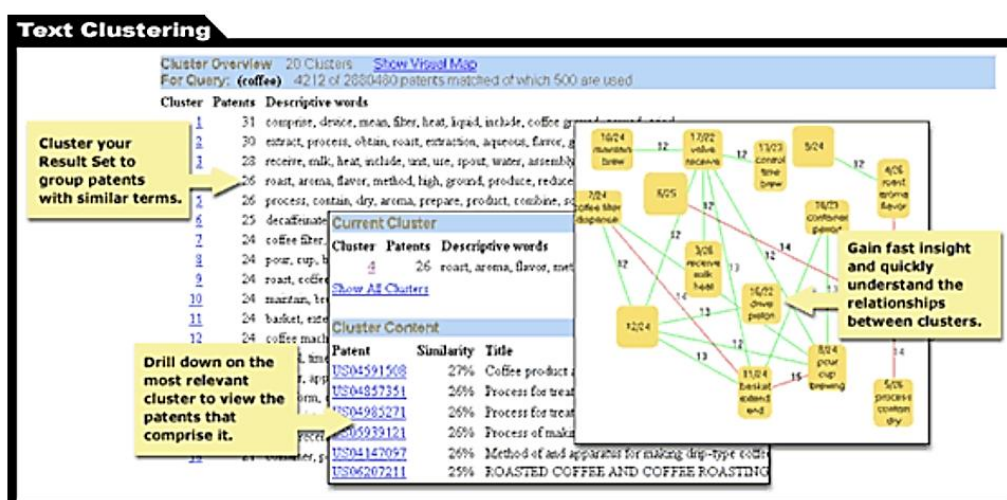
Rysunek 91: Sporządzanie wykresów w bazie *Delphion*. Źródło: strona internetowa *Delphion*.

- Analizę wyników wyszukiwania w trybie *offline* z wykorzystaniem grafów i wykresów trójwymiarowych (*PatentLab II*)



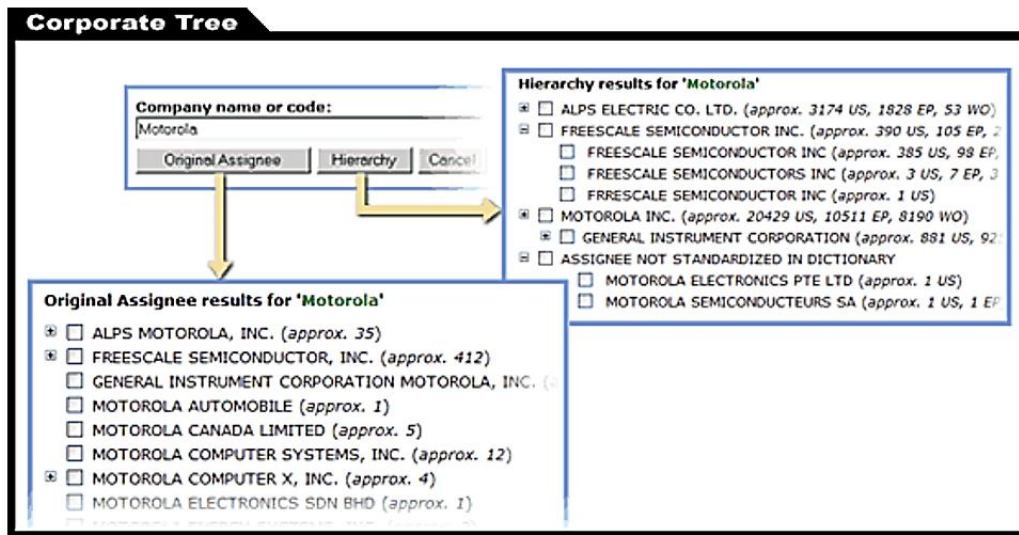
Rysunek 92: PatentLab II w bazie Delphion. Źródło: strona internetowa Delphion.

- Analizę opartą o słowa kluczowe (*TextClustering*)



Rysunek 93: Przykład analizy słów kluczowych w bazie Delphion. Źródło: strona internetowa Delphion.

- Drzewa korporacyjne – diagramy ułatwiające wyszukiwanie ze względu na właściciela.



Rysunek 94: Przykład drzewa korporacyjnego w bazie *Delphion*. Źródło: strona internetowa *Delphion*.

Do mocnych stron bazy *Delphion* analitycy z portalu internetowego www.intellogist.com zaliczają także (Intellogist 2011c):

- Przeszukiwanie zasobów bazy *Derwent WPI*,
- Szczegółową listę wyników usprawniającą proces wyszukiwania,
- Możliwość wyszukiwania przez kryteria niestandardowe, nieodstępne w innych bazach.

Baza jest raczej przeznaczona do wyszukiwania szybkiego i okazjonalnego, dla profesjonalnych użytkowników jest raczej niewystarczająca, chyba że brak alternatywnego dostępu do zasobów *Derwent World Patents Index* (Intellogist 2011c).

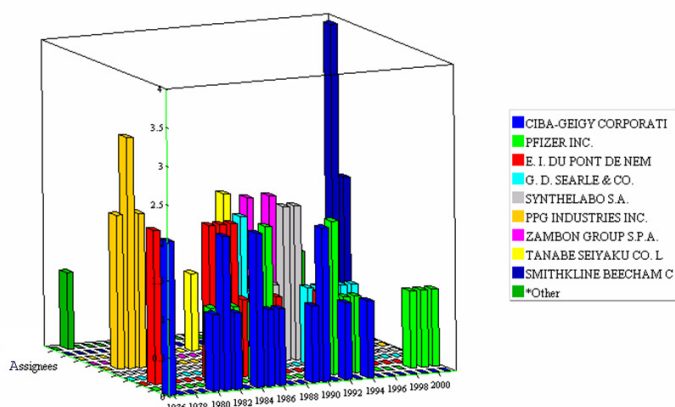
10.12.2. *Micro Patent / Patent Web*

Źródła danych: *Micro Patent* jest amerykańską bazą patentów afiliowaną przez *Thomson Reuters*, interfejs wyszukiwania jest w języku angielskim. W bazie dostępne są zasoby 6 urzędów patentowych – *USPTO*, *EPO*, *WIPO/PCT*, niemieckiego, francuskiego i brytyjskiego. Dostępne są także abstrakty z *Japan Patent Information Organization*.

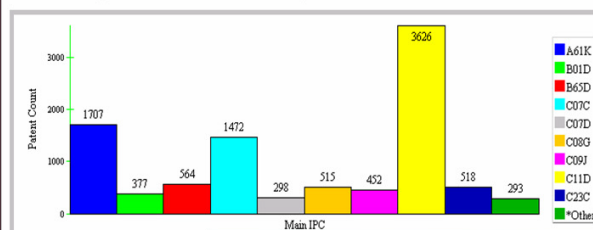
Koszty użytkowania: Początkowa opłata to 50\$ (za założenie konta), dostępne są dzienne i roczne opcje cenowe lecz brak ogólnodostępnych informacji cenowych.

Silne strony i ograniczenia bazy w porównaniu do baz bezpłatnych: W bazie można przeprowadzać proste analizy danych. Poza prostym sortowaniem, czy alertami, śledzeniem odwołań (cytowanych i cytujących), wyszukiwaniu patentów powiązanych, sprawdzaniem statusu prawnego można też przedstawiać wyniki wyszukiwania za pomocą reprezentacji graficznej (2D i 3D).

3D Bar Chart (Patent count vs. Assignee vs. Publication Date)

[Close Graph Window](#)
[Print Graph Window](#)


2D Bar Chart (Patent count vs. Main IPC class)

[Click here for IPC Class Titles](#)
[Close Graph Window](#)
[Print Graph Window](#)


* All other IPCs are represented by the "Other" b

Left-clicking on graph shows the
 Right-clicking on a particular bar

File Edit View Favorites Tools Help

Main IPC:
C07d

WO0198276A1 20011227 PARA-AMINOPHENOL DERIVATIVES AND THEIR USE AS DEVELOPERS FOR PRODUCING OXIDATION COLORATIONS HENKEL KOMMANDITGESELLSCHAFT AUF AKTIEN (C07D)
 WO0110840A1 20010215 QUATERNIZED AZAFLORENONES AND THEIR USE IN AGENTS FOR DYING FIB CONTAINING KERATIN HENKEL KOMMANDITGESELLSCHAFT AUF AKTIEN (C07D)
 DE19937301A1 20010215 Quaternierte Azaflorenone und deren Verwendung in Mitteln zum Färben von keratin Fasern Henkel KGaA (C07D)
 WO9830175A1 19980803 NOVEL 1,4-DIAZACYCLOHEPTANE DERIVATIVES AND THEIR USE IN HAIR OXID DYES HENKEL KOMMANDITGESELLSCHAFT AUF AKTIEN (C07D)
 EP066449A1 19941224 NOVEL 1,4-DIAZACYCLOHEPTANE DERIVATIVES AND THEIR USE IN HAIR OXID

Rysunek 95: Przykład wykresów w bazie *Micro Patent*. Źródło: strona internetowa *Micropat*.

Do mocnych stron bazy *Micro Patent* analitycy z portalu internetowego www.intellogist.com zaliczają (Intellogist 2011d):

- Dobry zbiór dokumentacji graficznej,
- Dobrze zaprojektowany formularz wyszukiwania,
- Dobre narzędzia do zarządzania historią wyszukiwania.

Ze względu na interfejs i zakres wyszukiwanych danych, baza nie dorównuje innym produktom komercyjnym. Poza tym, interfejs nie jest nowoczesnie skonstruowany, na wyświetlenie wyszukiwanych dokumentów oczekuje się długo, niektóre kryteria wyszukiwania są uciążliwe w zastosowaniu (np. śledzenie cytowań), zdarzają się przypadki utraty danych (Intellogist 2011d).

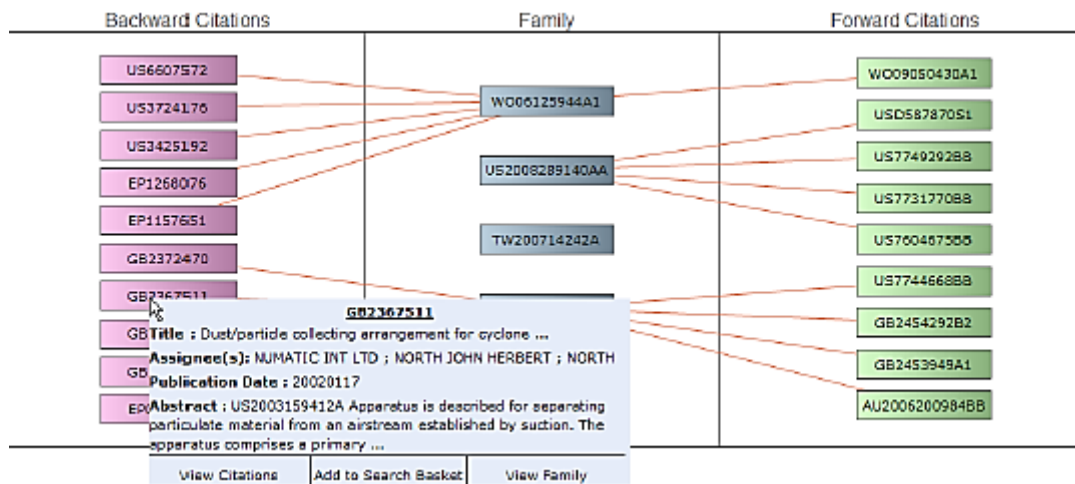
10.12.3. Pat Base

Źródła danych: *Pat Base* jest brytyjską bazą patentów, afiliowaną przez *Minesoft Ltd.*, interfejs wyszukiwania jest dostępny w językach angielskim, francuskim, niemieckim i

japońskim. W bazie dostępne są zasoby 11 urzędów patentowych – *USPTO*, *WIPO*, *EPO*, Wielkiej Brytanii, Francji, Chin, Japonii, Niemiec, Indii, Południowej Korei i Tajwanu. Do tego dochodzą niepełne dokumentacje (np. abstrakty) z innych państw, ogółem z 95 urzędów.

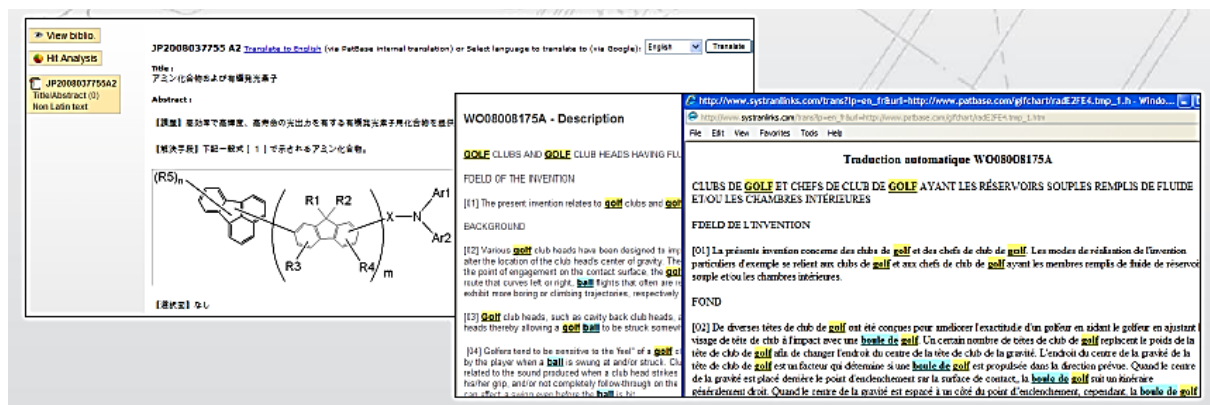
Koszty użytkowania: Opłata roczna, ceny różnicowane dla poszczególnych klientów.

Silne strony i ograniczenia bazy w porównaniu do baz bezpłatnych: Możliwe są dodatkowe analizy danych. Przykładowo, analizę dokumentów cytowanych i cytujących można wykonywać w przystępnej formie graficznej.



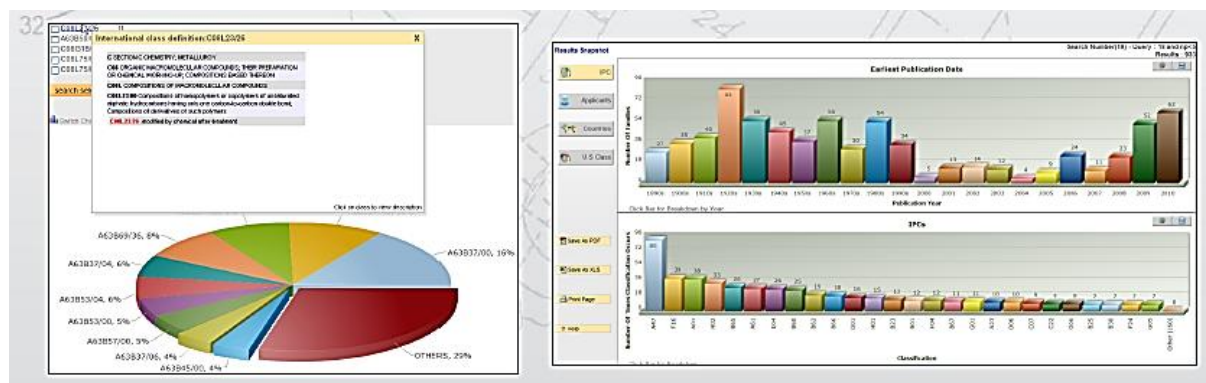
Rysunek 96: Przykład analizy cytowań w bazie *PatBase*. Źródło: materiały marketingowe *Patbase*.

Przy dokumentach obcojęzycznych (w szczególności azjatyckich) pomocny jest wbudowany translator.



Rysunek 97: Przykład użycie translatora w bazie *PatBase*. Źródło: materiały marketingowe *Patbase*.

Dostępne są narzędzia analizy graficznej wyszukanych danych oraz ich eksport, prócz takich formatów jak *XML* czy *csv*, a także tworzenia raportów w *HTML*-u, *Wordzie* czy *Excelu* można wykorzystać programy takie jak *Minesoft Insight Pro*, *VantagePoint*, *BizInt*.



Rysunek 98: Przykład analizy graficznej danych w bazie *PatBase*. Źródło: materiały marketingowe *Patbase*.

Do mocnych stron bazy *PatBase* analitycy z portalu internetowego www.intellogist.com zaliczają (Intellogist 2011e):

- Dobrze ustrukturalizowane wyświetlanie dokumentów z powiązanych dziedzin, co ułatwia agregowanie danych,
- Dokumenty z zasobów azjatyckich i przydatny przy pracy z nimi translator,
- Linki do zewnętrznych źródeł, udostępniających dokumentację patentową nieodpłatnie.

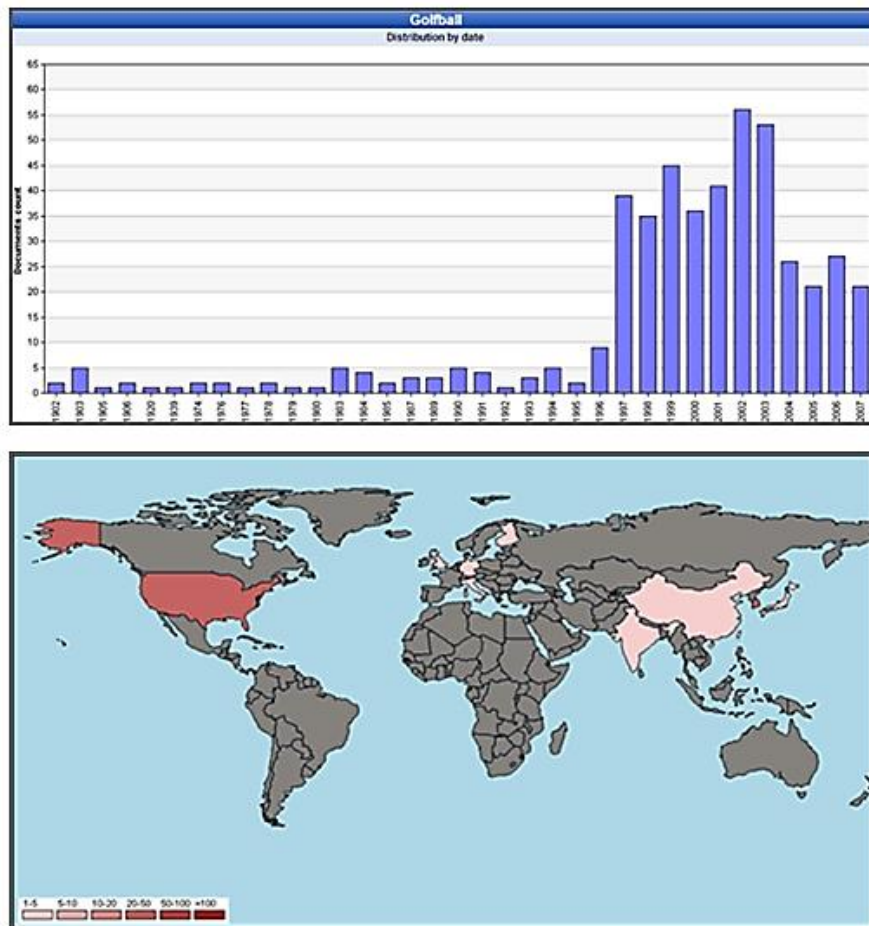
Podczas procesu wyszukiwania mogą zdarzyć się pominięcia lub niewłaściwe przyporządkowania niektórych dokumentów, historia wyszukiwania może również niekiedy zostać utracona, co analitycy *Intellogist* uznają za ograniczenia bazy (Intellogist 2011e).

10.12.4. QPat

Źródła danych: *QPat* jest amerykańską bazą patentów, oferowaną przez *Questel-Orbit*, której interfejs wyszukiwania jest dostępny w językach angielskim, francuskim, niemieckim i japońskim. W bazie zgromadzone są zasoby 22 urzędów patentowych – *EPO*, *WIPO*, *USA*, *Austrii*, *Belgii*, *Brazylia*, *Kanady*, *Chin*, *Danii*, *Finlandii*, *Francji*, *Niemiec*, *Indii*, *Japonii*, *Korei*, *Rosji*, *Hiszpanii*, *Szwecji*, *Szwajcarii*, *Tajwanu*, *Wielkiej Brytanii* i byłego Związku Radzieckiego.

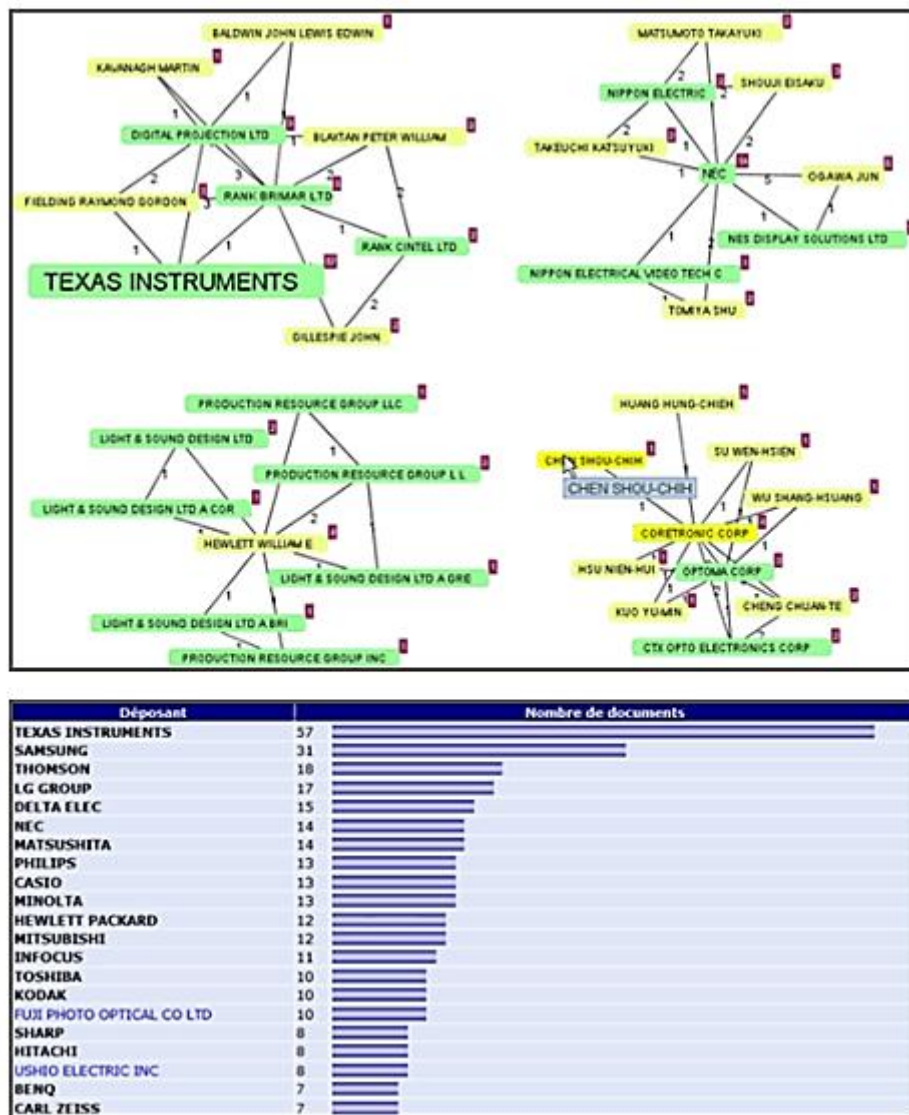
Koszty użytkowania: Dla klienta korporacyjnego roczna opłata wynosi 1.000 euro, z dodatkowymi opłatami za generowanie zestawień wyników wyszukiwania - szczegółowe tabele cenowe są dostępne na stronie internetowej bazy: www.questel.com/becomeauser.htm#trademark

Silne strony i ograniczenia bazy w porównaniu do baz bezpłatnych: Baza oferuje dodatkowe narzędzia analizy statystycznej i graficznej wyszukanych danych. Można tworzyć wykresy geograficzne w oparciu o geograficzne kryteria wyszukiwania.



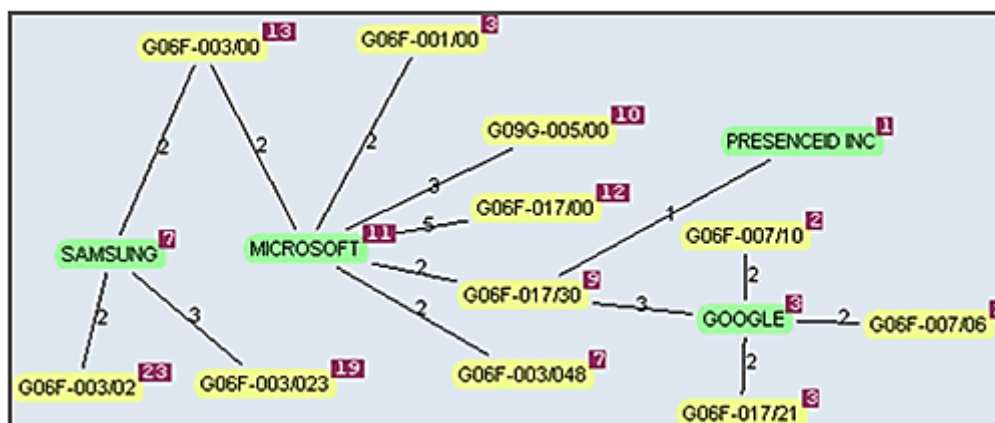
Rysunek 99: Przykład wykresów geograficznych w bazie *QPat*. Źródło: strona internetowa *Questel*.

Można przygotować w bazie wykresy słupkowe i sieciowe, ilustrujące właścicieli patentów i ich wzajemne relacje.



Rysunek 100: Przykład relacyjnych wykresów sieciowych i słupkowych w bazie *QPat*. Źródło: strona internetowa *Questel*.

Możliwe jest sporządzenie wykresów ilustrujących działalność korporacji badawczych i ich podział według jednej z klasyfikacji patentowych (*IPC* lub *ECLA*).



Rysunek 103: Przykład krzyżowego wykresu relacyjnego w bazie *QPat*. Źródło: strona internetowa *Questel*.

Do mocnych stron bazy *QPat* analitycy z portalu internetowego www.intellogist.com zaliczają (Intellogist 2011f):

- Dobrze strukturalizowane powiązania cytowań w obszarach tematycznych, a system grupowania rodzin patentów (*PatFam*) jest unikalny, podobny do *INPADOC* lecz bardziej szczegółowy,
- Dostępny wbudowany translator na język angielski z 35 języków,
- Ciekawe narzędzia analizy danych i ich ilustracji graficznej.

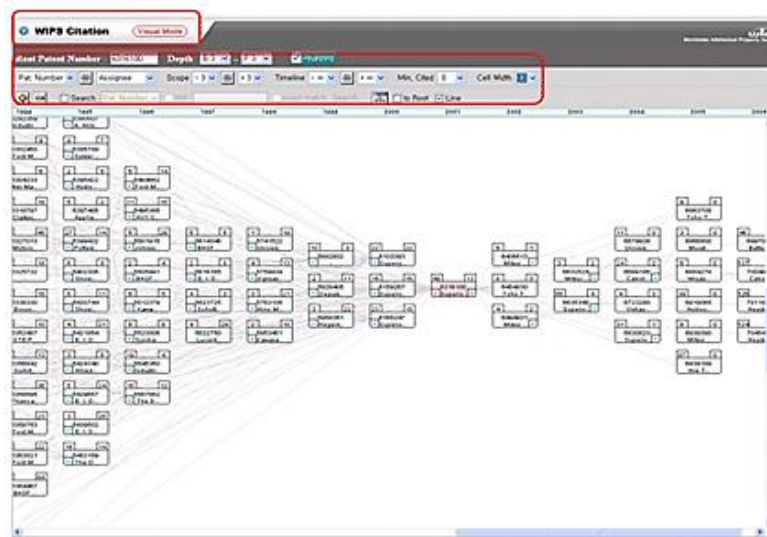
Do ograniczeń bazy należy jednak zaliczyć grupowanie wyników wyszukiwania dla maksymalnie 5 tysięcy dokumentów, podział pełnych tekstów dokumentacji na mniejsze elementy, powolne wyświetlanie rezultatów. Dodatkowo, wyświetlane bibliograficzne kryteria wyboru mogą być niekiedy mylące, a historia wyszukiwania może zostać czasami przypadkowo utracona (Intellogist 2011f).

10.12.5. *WIPS Global*

Źródła danych: *WIPS Global* jest południowo-koreańską bazą patentów, oferowaną przez firmę o tej samej nazwie. Interfejs wyszukiwania jest oferowany w językach angielskim, koreańskim i japońskim. W bazie dostępne są zasoby dwóch urzędów patentowych – *USPTO* i *EPO*. Częściowo znalazły się w niej także dokumenty z *WIPO*, szwajcarskie, niemieckie, francuskie, brytyjskie, japońskie, koreańskie i chińskie.

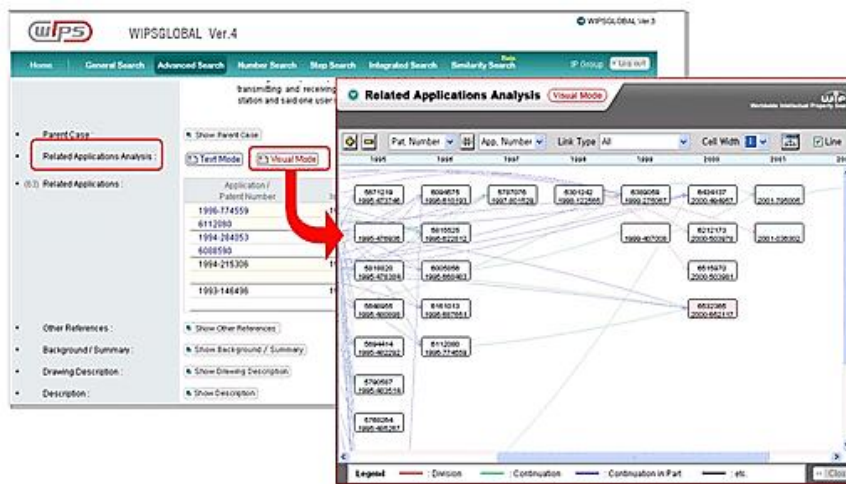
Koszty użytkowania: W bazie *WIPS* przewidziane są subskrypcje miesięczne i roczne w opcji podstawowej i specjalnej, w zależności od wyboru subskrypcji ceny wynoszą: 250\$ (miesięczna, podstawowa), 350\$ (miesięczna, specjalna), 2200\$ (roczna, podstawowa), 2500\$ (roczna, specjalna).

Silne strony i ograniczenia bazy w porównaniu do baz bezpłatnych: W zakresie analiz bibliometrycznych, w bazie *WIPS* jest możliwa analiza cytowań w trybie tekstowym i graficznym.



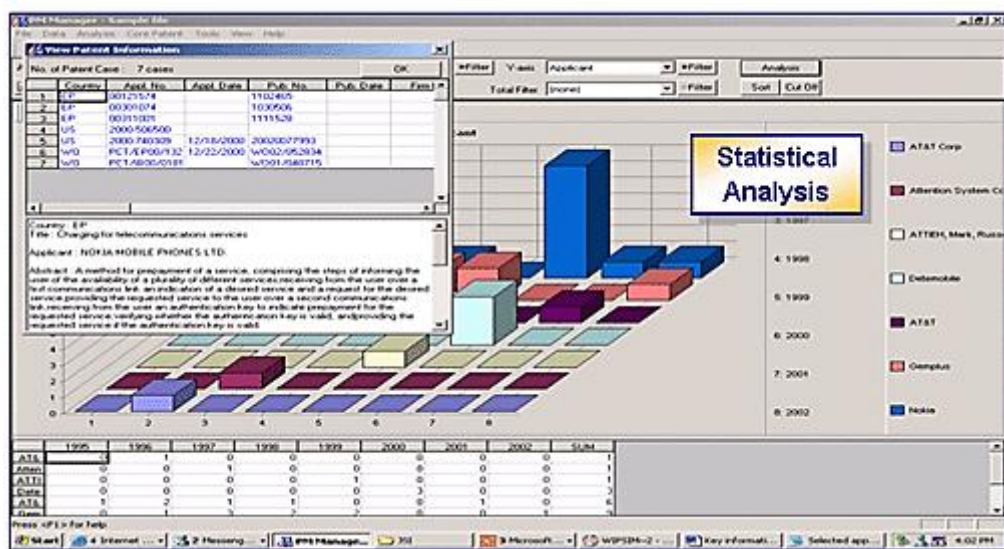
Rysunek 104: Przykład analitycznego wykresu cytowań w bazie WIPs. Źródło: strona internetowa WIPs.

Można także analizować tekstowo i graficznie dokumenty powiązane.



Rysunek 105: Przykład analitycznego wykresu relacyjnego w bazie WIPs. Źródło: strona internetowa WIPs.

W bazie jest zawarte specjalne narzędzie (*PM Manager*), służące do analizy statystycznej, edycji i zarządzania wynikami oraz tworzenia grafów i raportów.



Rysunek 106: Przykładowy widok ekranu *PM Manager* w bazie *WIPS*. Źródło: strona internetowa *WIPS*.

Rozszerzone analizy można prowadzić za pomocą dedykowanego programu, powiązanego z bazą o nazwie *Thinklear*.

Do mocnych stron bazy *WIPS* analitycy z portalu internetowego www.intellogist.com zaliczają (Intellogist 2011g):

- Indeksowanie według Patentowego Indeksu Koreańskiego (*KPI, Korean Patent Index*),
- Oryginalne narzędzia analizy danych i sposoby prezentacji graficznej.

Do ograniczeń bazy należy jednak zaliczyć niejasny interfejs (zbyt wiele osobnych okien) oraz brak możliwości zapisu pełnej historii wyszukiwania. Zmiany w kryteriach wyszukiwania urzędu patentowego mogą powodować utratę historii wyszukiwania. Baza dobrze pokrywa patentowe zasoby koreańskie, ma jednak liczne braki w odniesieniu do innych grup patentów (Intellogist 2011g).

10.12.6. Total Patent

Źródła danych: *Total Patent* jest amerykańską bazą patentów oferowaną przez *Lexis-Nexis*, interfejs wyszukiwania jest w języku angielskim. W bazie dostępne są zasoby 29 urzędów patentowych – Austrii, Australii, Belgii, Brazylii, Kanady, Szwajcarii, Chin, Niemiec, Danii, *EA (Eurasian Patent Office)*, *EPO*, Hiszpanii, Finlandii, Francji, Wielkiej Brytanii, Irlandii, Indii, Włoch, Japonii, Luksemburga, Monako, Meksyku, Holandii, Portugalii, Rosji, Szwecji, dawnego Związku Radzieckiego, USA, *WIPO*.

Koszty użytkowania: W zależności od działań w bazie, ceny wahają się od kilku do kilkuset dolarów za konkretne wyszukiwanie. Ostateczna cena jest kalkulowana indywidualnie.

Silne strony i ograniczenia bazy w porównaniu do baz bezpłatnych: Według autorów bazy, możliwe jest tworzenie map graficznych i wykresów, analiz i wizualizacji 3D, lecz brak informacji o szczegółach tych narzędzi. Do mocnych stron bazy *Total Patent* analitycy z portalu internetowego www.intellogist.com zaliczają (Intellogist 2011h):

- Pełne dane z wielu krajów,
- Dostępność technologii semantycznych, służących do generowania słów kluczowych,

- Przyjazny użytkownikowi interfejs i opcje przeglądania wyników,
- Narzędzia analiz w bazie,
- Opcje dostarczania dokumentów.

Do ograniczeń bazy należy jednak zaliczyć limit 3.000 rekordów uzyskanych w wyniku wyszukiwania, jak również problemy w wyszukiwaniu i tłumaczeniu dokumentów w językach innych niż angielski (Intellogist 2011h).

10.12.7. Thomson Innovation

Źródła danych: Thomson Innovation jest amerykańską bazą patentów afiliowaną przez Thomson Reuters, interfejs wyszukiwania jest w języku angielskim i japońskim. W bazie dostępne są zasoby 8 urzędów patentowych – USA, EPO, WIPO, Japonii, Niemiec, Wielkiej Brytanii, Francji i Korei Południowej. Dodatkowo, dostępne są niektóre dokumenty chińskie.

Koszty użytkowania: Istnieje wiele opcji cenowych, które są dopasowywane do klienta, brak jednak szczegółowych ogólnodostępnych informacji na ten temat.

Silne strony i ograniczenia bazy w porównaniu do baz bezpłatnych: W bazie dostępne są interesujące narzędzia analizy danych, warto zwrócić uwagę na mapy tematyczne, które w graficzny sposób hierarchizują terminy zarówno dotyczące patentów, jak i literatury nie-patentowej w postaci dwuwymiarowych wizualizacji.



Rysunek 107: Przykład mapy tematycznej w bazie Thomson Innovation. Źródło: strona internetowa Thomson Innovation.

Można prowadzić lingwistyczne analizy tekstu, uwidaczniać relacje między terminami, tworzyć strategie wyszukiwania. Dodatkowo także dostępne są opcje tworzenia różnych typów wykresów (słupkowych, kolumnowych, kołowych, 2D i 3D) a także mapy cytowań (cytowanych i cytujących).

Do mocnych stron bazy Thomson Innovation analitycy z portalu internetowego www.intellogist.com zaliczają (Intellogist 2011i):

- Wsparcie dla *Derwent World Patent Index*,

- Wbudowany translator języka japońskiego,
- Częściowo tłumaczone teksty z języka chińskiego,
- Zawartość zarówno danych patentowych jak i literatury nie-patentowej.

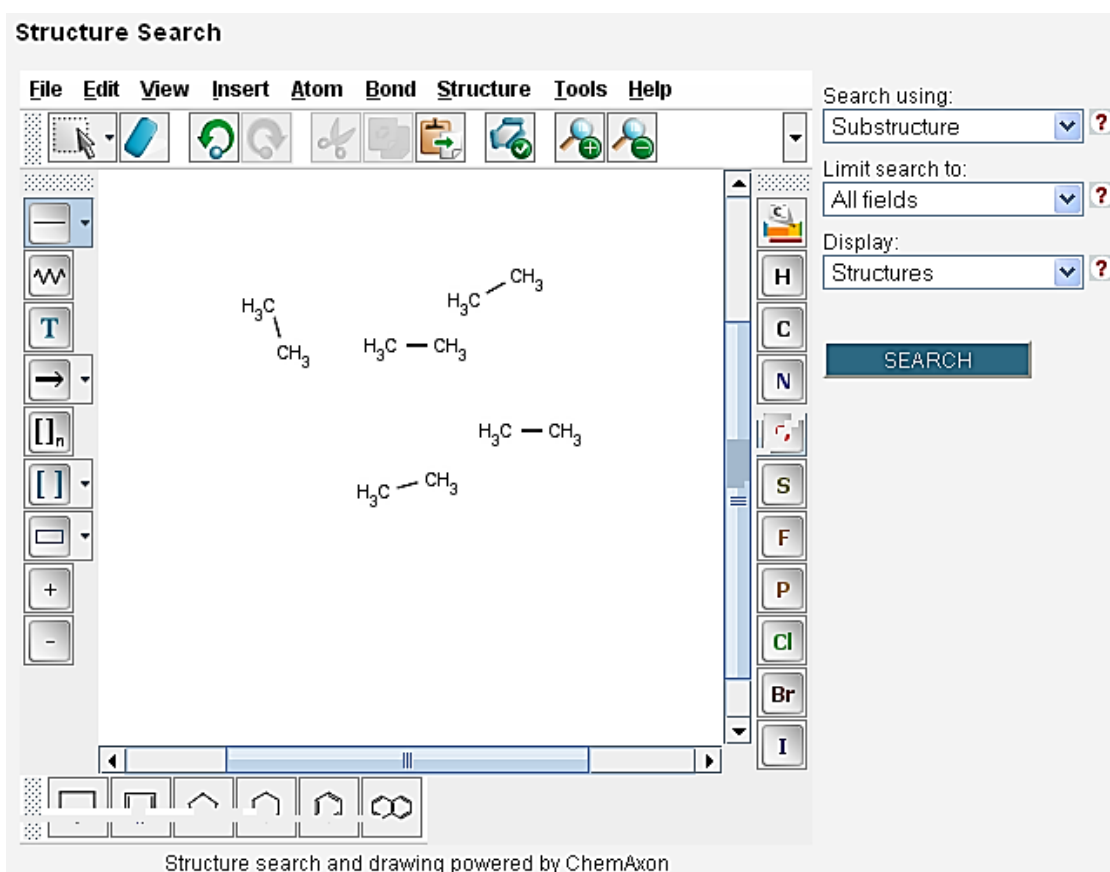
Do ograniczeń bazy należy jednak zaliczyć niedoskonałości wyszukiwania przez słowa kluczowe (*Keywords In Context*) oraz niewyświetlanie niektórych elementów graficznych z *esp@cenet* (Intellogist 2011i).

10.12.8. SureChem

Źródła danych: *SureChem* jest brytyjską bazą patentów firmy *Macmillan Publishers Ltd.*, interfejs wyszukiwania jest w języku angielskim. W bazie dostępne są zasoby 3 urzędów patentowych – USA, *EPO* i *WIPO* oraz częściowa dokumentacja japońska.

Koszty użytkowania: Są dwie opcje subskrypcji – miesięczna za 300\$ i roczna za 2.400\$. Te opcje dotyczą indywidualnych użytkowników, inne licencje są udzielane po przeprowadzeniu indywidualnej kalkulacji. Brak szczegółowych informacji.

Silne strony i ograniczenia bazy w porównaniu do baz bezpłatnych: W odniesieniu do analiz bibliograficznych, nie ma opcji graficznych, można przeprowadzać jedynie analizy statystyczne. Istotną korzyścią jest jednak możliwość wyszukiwania ze względu na graficzne wzory strukturalne substancji chemicznych.



Rysunek 108: Widok formularza wyszukiwania struktur chemicznych w bazie *SureChem*. Źródło: strona internetowa *SureChem*.

Do mocnych stron bazy *SureChem* analitycy z portalu internetowego www.intellogist.com zaliczają (Intellogist 2011j):

- Wyszukiwanie ze względu na strukturę chemiczną,
- Dane pokrywające najważniejsze urzędy patentowe a także bazę *MEDLINE*.

Do ograniczeń bazy należy jednak zaliczyć brak historii wyszukiwania. Baza jest też mało kompleksowa i poza wyszukiwaniem struktur chemicznych trudno znaleźć jej inne przewagi (Intellogist 2011j).

10.12.9. *JP-NETe*

Źródła danych: *JP-NETe* jest japońską bazą patentów oferowaną przez *Japan Patent Data Service Co., Ltd.*, interfejs wyszukiwania jest w języku angielskim i japońskim. Baza nie udostępnia pełnotekstowych dokumentów patentowych, jedynie częściową zawartość dokumentów japońskich i amerykańskich.

Koszty użytkowania: Za patenty japońskie: 96.000 jenów za jeden identyfikator, 360.000 jenów za 5 identyfikatorów. Za patenty amerykańskie odpowiednio: 60.000 i 240.000 jenów, a za materiały konferencyjne: 36.000 jenów za jeden identyfikator. W wypadku licencji na więcej niż 5 identyfikatorów, przewidziane są ustalane indywidualnie rabaty.

Silne strony i ograniczenia bazy w porównaniu do baz bezpłatnych: W bazie nie można bezpośrednio prowadzić żadnych analiz bibliometrycznych, ani tekstowych, ani statystycznych ani graficznych. Do mocnych stron bazy *JP-NETe* analitycy z portalu internetowego www.intellogist.com zaliczają (Intellogist 2011k):

- Tłumaczenie danych japońskich,
- Kompletnie zestawy rysunków z dokumentacji patentowej.

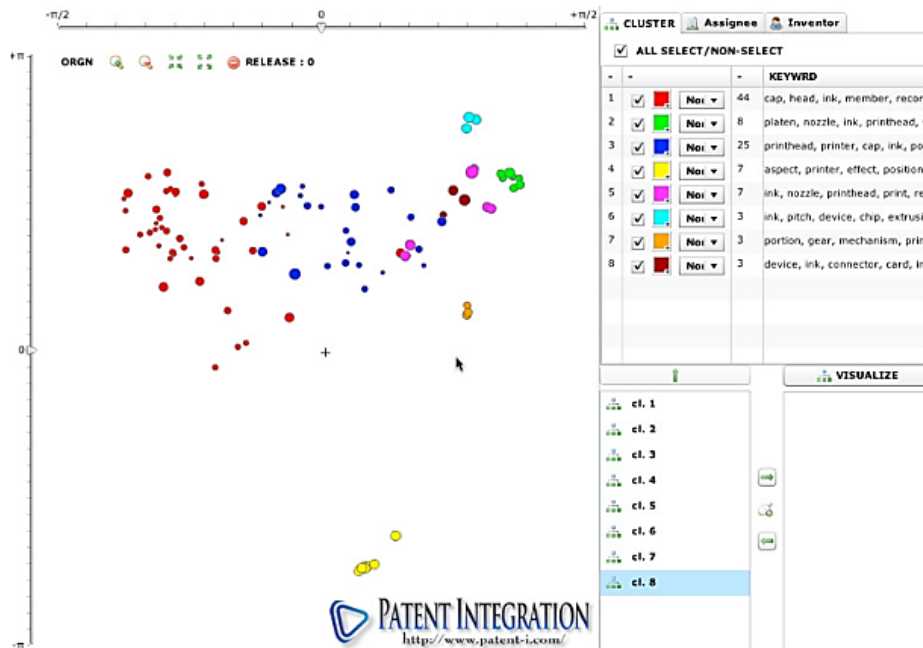
Baza ma jednak wiele ograniczeń – restrykcyjne opcje wyszukiwania, ograniczenia listy wyników, niemożliwość zapisywania uzyskanych wyników. Baza może służyć jako dodatkowa pomoc przy poszukiwaniu japońskich źródeł, jednak do profesjonalnych analiz i wyszukiwania jej przydatność wydaje się mocno ograniczona (Intellogist 2011k).

10.12.10. *Patent Integration*

Źródła danych: *Patent Integration* jest japońską bazą patentów firmy *Patent Integration Co. Ltd.* Interfejs wyszukiwania jest w języku angielskim, niemieckim, francuskim, szwedzkim, fińskim, włoskim, hiszpańskim, portugalskim, hinduskim, rosyjskim, hebrajskim, arabskim, japońskim, chińskim, koreańskim. W bazie dostępne są zasoby ośmiu urzędów patentowych – *USPTO*, *EPO*, *WIPO*, chińskie, koreańskie, tajwańskie, meksykańskie i japońskie.

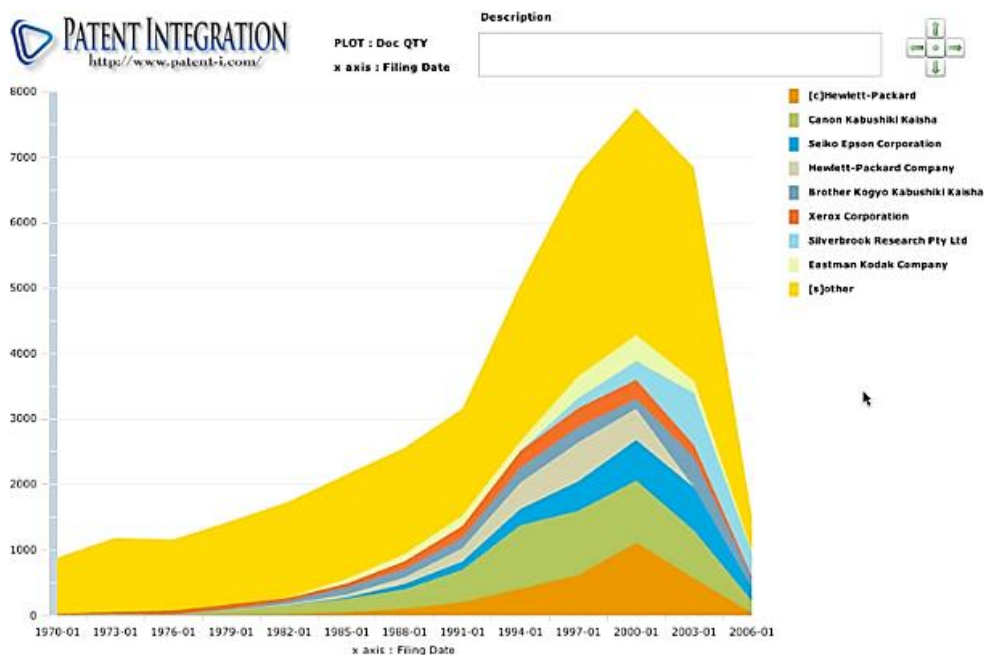
Koszty użytkowania: Licencje można wykupić na rok lub miesiąc w trzech opcjach. Licencja standardowa (indywidualna) miesięczna kosztuje 80\$, roczna 800\$. Koszt licencji akademickiej to miesięcznie 48\$ lub rocznie 480\$. Natomiast licencję korporacyjną (dla nieograniczonej liczby pracowników) można nabyć jedynie w opcji rocznej i jest to związane z wydatkiem 24.000\$.

Silne strony i ograniczenia bazy w porównaniu do baz bezpłatnych: W bazie dostępne są interesujące narzędzia analizy danych. Obok funkcji, dostępnych w większości komercyjnych rozwiązań, warto zwrócić uwagę na możliwość wizualizacji patentów i ich automatycznego grupowania.



Rysunek 109: Przykład wizualizacji i grupowania patentów w bazie *Patent Integration*. Źródło: strona internetowa *Patent Integration*.

Można również tworzyć mapy patentów z wykorzystaniem różnych typów wykresów.



Rysunek 110: Przykład wykresu prezentującego liczby patentów w bazie *Patent Integration*. Źródło: strona internetowa *Patent Integration*.

Za pomocą tych samych narzędzi mapowania można tworzyć wykresy relacji pomiędzy właścicielami patentów, generować i samodzielnie edytować ich zestawienia, a także korzystać z funkcji osobistego słownika, który ułatwia budowę własnych zestawień oraz dzielenie się nimi z innymi użytkownikami.

Do mocnych stron bazy należy zaliczyć:

- Gromadzenie danych z wielu obszarów geograficznych, w tym obszarów niedostępnych w innych bazach,
- Rozbudowane narzędzia analizy i mapowania,
- Uporządkowany, przejrzysty interfejs użytkownika,
- Najatrakcyjniejszą cenę spośród wszystkich omawianych baz komercyjnych, z dodatkowymi upustami dla instytucji akademickich.

Baza *Patent Integration* jest relatywnie nowym rozwiązaniem i można spodziewać się dalszego rozwoju jej funkcjonalności. Obecnie istotnym zaobserwowanym ograniczeniem jest niestety brak wsparcia dla analizy cytowań patentów.

11. Oprogramowanie wykorzystywane do analiz bibliometrycznych oraz wizualizacji ich wyników

11.1. Wprowadzenie

Ze względu na różnorodność analiz bibliometrycznych i eksploracji danych tekstowych (ang. *text mining*), nie wydaje się możliwe wskazanie jednego narzędzia jako zdecydowanie najlepszego czy najdoskonalszego dla zróżnicowanych zastosowań. W zależności od podmiotu przeprowadzającego analizę, źródłowej bazy, przedmiotu i celu analizy, a także stopnia zaawansowania użytkownika, wskazane jest stosowanie różnych aplikacji. Można jednocześnie wyróżnić najważniejsze uwarunkowania, determinujące przydatność konkretnego narzędzia do określonych zastosowań.

Niniejszy rozdział rozpoczyna się od przeglądu dostępnych na rynku narzędzi. Następnie została w nim dokonana analiza porównawcza – najpierw pomiędzy szeroko określonymi rodzajami oprogramowania, następnie prezentacja kilku wybranych programów, których różnice i podobieństwa z punktu widzenia użytkownika końcowego zaprezentowano w formie tabelarycznej. W końcowej części rozdziału znalazło się zestawienie kilkudziesięciu narzędzi z każdej z omawianych wcześniej grup. Taka konstrukcja rozdziału ma służyć łatwiejszemu zrozumieniu podstawowych różnic pomiędzy różnymi rodzajami programów, a także pozwalać w razie potrzeby na samodzielne odnalezienie dodatkowych informacji na temat konkretnego narzędzia.

Najprostszym z punktu widzenia użytkownika końcowego podziałem narzędzi do analizy bibliometrycznej jest podział na programy komercyjne i bezpłatne. Najpopularniejsze programy komercyjne to: *VantagePoint* i *Matheo*, zaś bezpłatne: *CiteSpace*, *HistCite* i *PatentNet*.

Innym sposobem podziału jest określenie podstawowych funkcji oprogramowania (Porter, Cunnigham 2005: 363):

- narzędzia stworzone wyłącznie do eksploracji danych tekstowych (np. *ClearResearch*),
- narzędzia wzbogacone o moduł do eksploracji danych tekstowych (np. *SPSS*, *SAS Text Miner/Enterprise Miner*, *WordStat*),
- narzędzia zintegrowane z określonymi bazami danych (np. *Aureka*, *Delphion*),
- narzędzia skoncentrowane na przetwarzaniu informacji (np. *Dataview*, *Focust*, *VantagePoint* – sprzedawany także jako *TechOASIS* i *Derwent Analytics*),
- narzędzia służące do wizualizacji wyników (np. *Anacubis*, *VXInsight*),
- narzędzia pozwalające na tworzenie rozwiązań, dostosowanych do specyficznych potrzeb klienta (np. *Intelligent Miner for Text*).

Nieco inny sposób podziału proponują Yang i in. (2008: 281), wyróżniając programy wykonujące:

- analizę słów kluczowych (wyjęte z kontekstu słowa użyte w teście),

- analizę statystyczną (częstość występowania słów w tekście),
- analizę lingwistyczną (analiza semantyczna oparta na języku naturalnym).

Obydwa podziały nie ujmują najważniejszych z punktu widzenia użytkownika końcowego kwestii, jak: przydatność programu do przeprowadzenia konkretnych rodzajów analiz, cena, dostępność interfejsu, czy sposób wizualizacji wyników. Ponadto, drugi z powyższych podziałów nie odnosi się do większości funkcji, oferowanych przez omawiane oprogramowanie. Dlatego w dalszej części tekstu kategorie te będą jedynie krótko scharakteryzowane, a szczegółowo omówione zostaną jedynie te programy, których przydatność do analiz wskazanych w raporcie wydaje się największa.

11.2. Rodzaje oprogramowania, wspierającego analizy bibliometryczne

W niniejszym podrozdziale znajduje się zestawienie podstawowych cech i funkcjonalności opisywanych grup programów komputerowych. Zestawienie to nie ma na celu kompletnego porównania wszystkich modułów i funkcji omawianych narzędzi – nie byłoby to celowe ze względu na duże zróżnicowanie zarówno interfejsu użytkownika, jak i podstawowych nawet sposobów działania. Po drugie, tego typu zestawienie byłoby nieczytelne i nie pozwalało na łatwą kategoryzację różnych aspektów funkcjonalności i możliwości wykorzystywania tych programów. Zanim zostaną omówione poszczególne narzędzia analityczne, warto przyjrzeć się cechom charakterystycznym szerzej zdefiniowanych grup programów.

Narzędzia wzbogacone o moduł do eksploracji danych tekstowych (np. dedykowane narzędzia z rodziny *IBM SPSS*, *SAS Text Miner*) mają ograniczoną przydatność do analiz bibliometrycznych ze względu na dużą złożoność interfejsu i mnogość funkcji niezwiązanych bezpośrednio z analizami bibliometrycznymi. Powoduje to, że przeciętny użytkownik musi opanować wiele niepotrzebnych z jego punktu widzenia funkcji, zanim może przystąpić do analizy bibliometrycznej. Co gorsza, często wymagana jest też umiejętność programowania w odpowiednim języku skryptowym. Poza tym, są to zwykle programy stosunkowo drogie i wymagające przeszkolenia przed użytkowaniem.

Zalety	Wady
Duże możliwości	Złożony interfejs
Współpraca z różnymi modułami i programami tych samych producentów	Nieprzystępny dla przeciętnego użytkownika
Zwykle dobra pomoc techniczna	Wysoki koszt

Tabela 51: Najważniejsze cechy oprogramowania wzbogaconego o moduł do eksploracji danych tekstowych. Źródło: opracowanie własne na podstawie (Porter, Cunningham 2005) oraz materiałów producentów.

Programy zintegrowane z bibliometrycznymi bazami danych (np. *Aureka, Delphion*) zostały zwykle stworzone przez dostawców baz i służą przede wszystkim do obsługi danych zawartych w tych właśnie bazach – zwłaszcza bazach patentowych. Niektóre z nich były opisywane w poprzednim rozdziale, przy okazji przeglądu komercyjnych baz patentów. Z jednej strony programy takie są doskonale dopasowane do pracy z jedną konkretną bazą danych, ale z drugiej strony ta specjalizacja odbija się niekorzystnie na ich uniwersalności. W ograniczonym zakresie lub w ogóle nie współpracują z innymi rodzajami danych czy baz i z tego powodu w przypadku potrzeby poszerzenia analizy o informacje pochodzące z innej bazy, zwykle potrzebne jest dodatkowe oprogramowanie.

Zalety	Wady
Doskonała współpraca z konkretną bazą danych	Ograniczony zakres stosowania
Łatwo dostępne z poziomu bazy danych	Nie oferują tak wielu możliwości analitycznych, jak programy uniwersalne
Zwykle dość prosty interfejs	Konieczność korzystania z innych programów w celu uzupełnienia analizy

Tabela 52: Najważniejsze cechy oprogramowania zintegrowanego z bibliometrycznymi bazami danych. Źródło: opracowanie własne na podstawie (Porter, Cunningham 2005) oraz materiałów producentów.

Programy o uniwersalnym zastosowaniu (np. *VantagePoint, Matheo, CiteSpace*) są stosunkowo popularne i zwykle oferują dużą elastyczność w zakresie obróbki danych (na przykład możliwość zawężenia wyszukiwań bez konieczności powtórnego ściągania danych), łączenie wyników z różnych baz danych oraz możliwość automatyzacji powtarzalnych czynności (Porter, Cunningham 2005: 132). Można za ich pomocą przeprowadzić zarówno proste kwerendy dotyczące częstotliwości występowania publikacji na konkretny temat, jak i złożone analizy, uwzględniające współzależności pomiędzy różnymi autorami z różnych krajów i współpracującymi ze sobą przy okazji konkretnego projektu. We wcześniejszych rozdziałach książki omówione zostały założenia podejścia *tech mining*, które wspiera tego typu oprogramowanie.

Zalety	Wady
Pozwalają analizować dane z różnych baz	Importowanie danych z baz bywa skomplikowane i wymaga wiedzy
Stosunkowo duże możliwości	Korzystanie z niektórych baz jest mniej wygodne niż przy użyciu

	wyspecjalizowanych narzędzi
Łatwość analizy po opanowaniu (zwykle dość prostego) interfejsu	Rozbudowane opcje mogą powodować względną trudność w obsłudze (szczególnie na początku)
Duża popularność tych programów ułatwia znalezienie rozwiązania konkretnego problemu	

Tabela 53: Najważniejsze cechy oprogramowania o charakterze uniwersalnym. Źródło: opracowanie własne na podstawie (Porter, Cunningham 2005) oraz materiałów producentów.

Grupa narzędzi służących do wizualizacji wyników nakłada się na pozostałe grupy oprogramowania, gdyż możliwości wizualizacji danych bibliometrycznych są współcześnie standardową funkcją oprogramowania. Z tego względu grupa ta nie zostanie odrębnie omówiona – w przypadku poszczególnych programów, zaprezentowane zostaną moduły służące do wizualizacji.

Zalety	Wady
Często bardziej rozbudowane możliwości wizualizacji wyników niż w przypadku zwykłych programów bibliometrycznych	Oferują bardziej ograniczony zakres analiz niż narzędzia dedykowane
Zwykle prostota obsługi	Wymagają dodatkowego oprogramowania do przygotowania danych wsadowych

Tabela 54: Najważniejsze cechy oprogramowania stworzonego do wizualizacji wyników. Źródło: opracowanie własne na podstawie (Porter, Cunningham 2005) oraz materiałów producentów.

Najbardziej rozpowszechnionymi i najpopularniejszymi programami do eksploracji danych bibliometrycznych są wspomniane wcześniej *VantagePoint*, *Matheo Analyzer*, *CiteSpace* oraz *VXInsight*. Ten ostatni nie jest już jednak rozwijany przez producenta i ze strony internetowej można jedynie pobrać *Sandia Toolkit*, czyli zestaw narzędzi programistycznych wymagających znajomości środowiska *Python*, dlatego program nie został ujęty w poniższej analizie. Pozostałe trzy aplikacje zostały opisane ze szczególnym uwzględnieniem oferowanych funkcji, wygody użytkownika, dostępnych wersji oraz kosztów. Ponadto, każdy opis został wzbogacony o zrzuty ekranowe, obrazujące programy w działaniu, aby uwidocznić podstawowe elementy interfejsu użytkownika i najważniejsze funkcje.

11.3. Opis programu *Matheo Analyzer*

Producent: *IALE Technologia* – hiszpańska forma doradczo-informatyczna, powstała w 1998 roku jako *spin-off* przy Politechnice Katalonii w Barcelonie.

Adres internetowy: www.matheo-analyzer.com

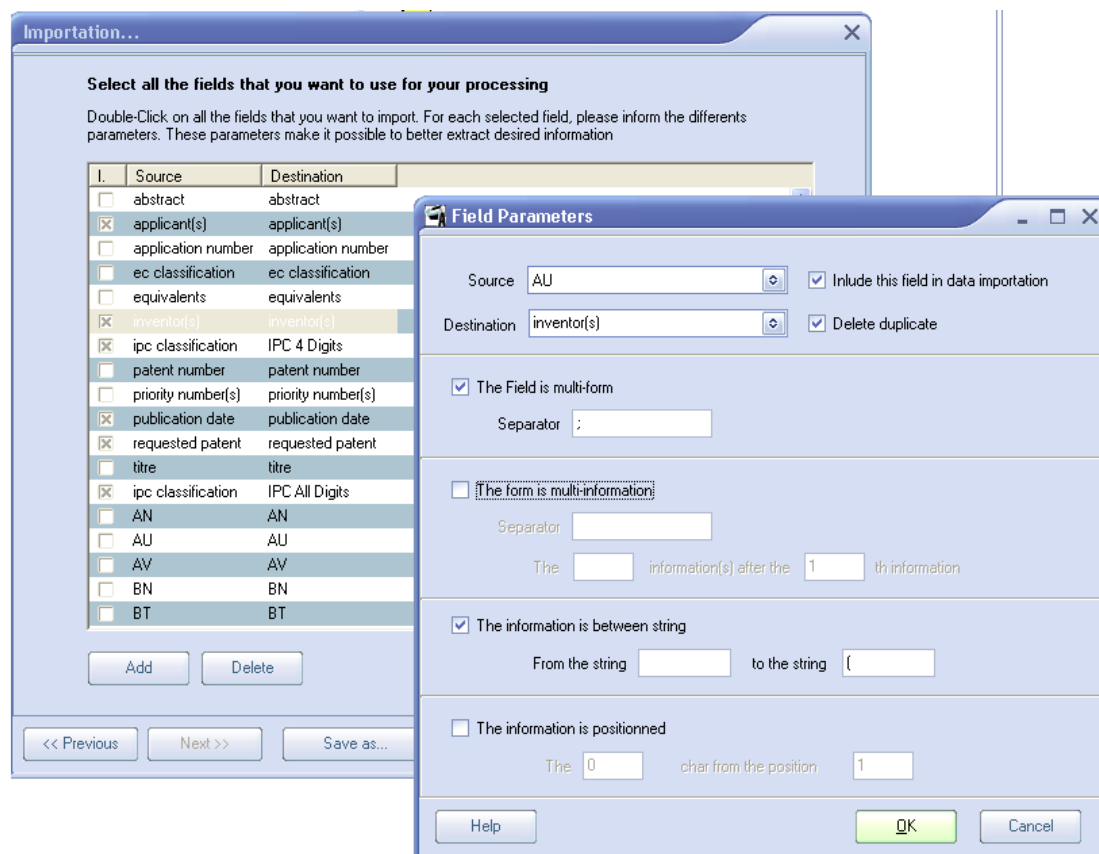
Cechy charakterystyczne: program oferuje zarówno przetwarzanie danych, mapowanie informacji, jak i rozbudowane narzędzia wizualizacyjne, kładąc nacisk na stosunkowo prostą obsługę. Stanowi relatywnie tanią, europejską alternatywę wobec najpopularniejszego narzędzia wspierającego *tech mining* – amerykańskiego programu *VantagePoint*.

Podstawowe cechy:

- Wielojęzyczny interfejs (francuski, angielski, hiszpański i niemiecki), jednak tłumaczenie poszczególnych funkcji pozostawia wiele do życzenia – nie wszystkie pojęcia i opcje zostały przetłumaczone i obsługa programu wymaga podstawowej chociaż znajomości języka francuskiego.
- Prosty interfejs, przystępny nawet dla początkujących użytkowników.
- Stosunkowo szybkie działanie.
- Dobra pomoc techniczna.

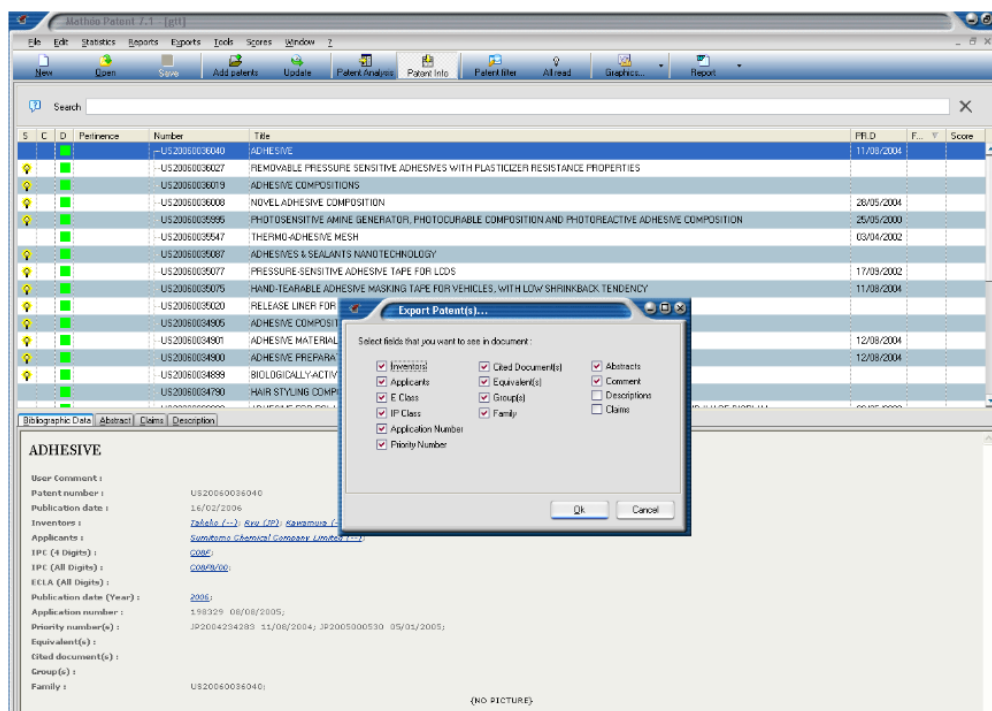
Podstawowa funkcjonalność:

- Obsługuje dane na temat patentów pochodzące z baz *USPTO* i *esp@cenet*.
- Wiele predefiniowanych sposobów importu, obróbki i eksportu danych.
- Pozwala na samodzielne zdefiniowanie nowego formatu importu danych (innych baz danych).
- Możliwość eksportu danych do formatów tekstowego i *XML*.
- Eksport danych może być przeprowadzany na kilka sposobów:
 - wszystkie informacje z wybranego pola,
 - wszystkie odwołania do jednej kwerendy,
 - wszystkie/wybrane pola równocześnie.
- Oferuje także wiele predefiniowanych sposobów importu (wyłącznie z formatu *.txt*), obróbki i eksportu danych poprzez obsługę predefiniowanych lub tworzonych przez użytkownika filtrów (ang. *pre-set import rules*). Pozwalają one uprościć i przyspieszyć proces wyciągania danych z zewnętrznych baz.
- Na każdym etapie importowania danych pozwala na filtrowanie pobieranych treści, co znacznie przyspiesza całą procedurę.



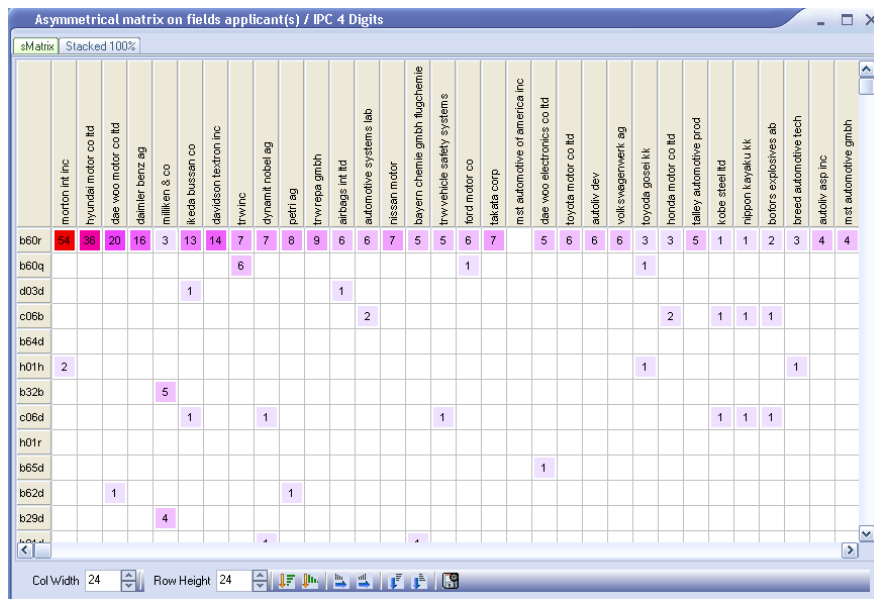
Rysunek 111: Program *Matheo Analyzer* – okna opcji importu danych. Źródło: strona internetowa *Matheo Analyzer*.

- Posiada dużo funkcji syntetyzujących informacje z baz danych.
- Bogate opcje wizualizacji wyników w formie histogramów, sieci i macierzy.
- Wszystkie mapy są w pełni interaktywne i pozwalają na dowolną modyfikację np. metodą przeciągnij-i-upuść.
- Współpracuje z *Matheo Patent* (programem do pobierania, przetwarzania i analizy informacji z baz danych patentowych *USPTO* i *esp@cenet*) oraz *Matheo Web* (programem do wyszukiwania i analizy informacji dostępnych w Internecie).



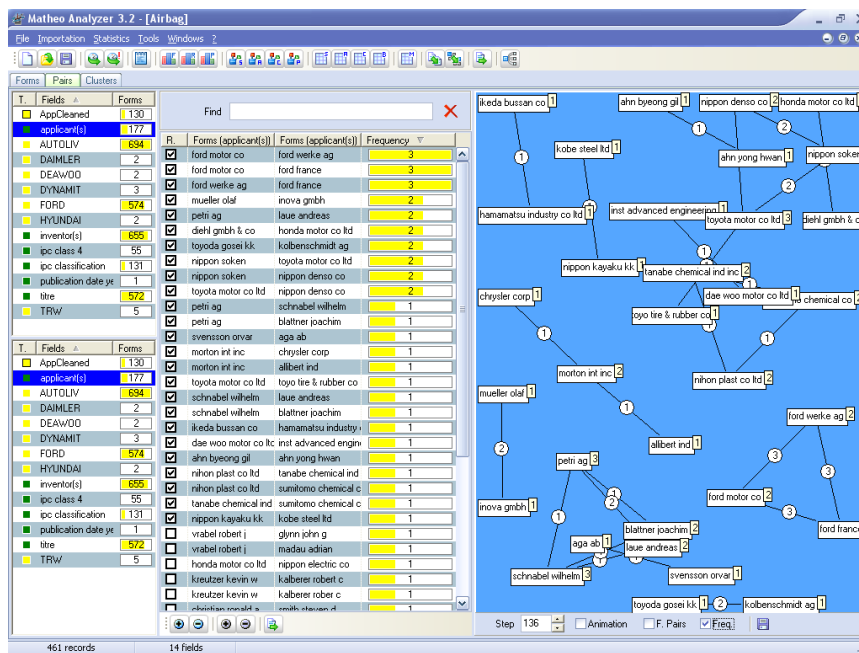
Rysunek 112: Program Matheo Analyzer – okno prezentujące wyniki pobierania danych z programu Matheo Patent. Źródło: strona internetowa Matheo Analyzer.

- Oczyszczanie danych może być przeprowadzone na kilka sposobów:
 - usuwanie powtarzających się rekordów,
 - łączenie powtarzających się rekordów,
 - usuwanie i łączenie powtarzających się rekordów, które nie są dokładnie identyczne,
 - usuwanie i łączenie powtarzających się rekordów, które zawierają określenia synonimiczne lub podobne (także według własnoręcznie zdefiniowanych przez użytkownika list synonimów).
- Prezentacja wyników analizy w formie wykresu jest dostępna na trzy sposoby:
 - wykresy kolumnowe (pionowe),
 - wykresy słupkowe (poziome),
 - wykresy kołowe.
- Macierze czynnikowe (*factor matrix*) pozwalają na prezentację wyników analizy czynnikowej w formie graficznej.



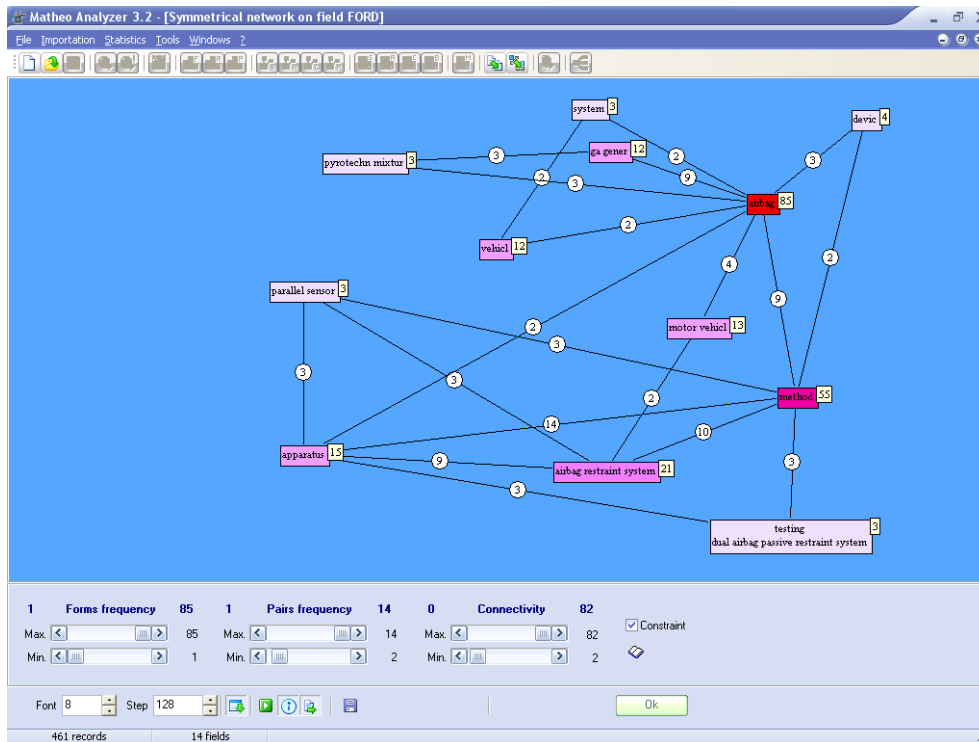
Rysunek 113: Program *Matheo Analyzer* – przykładowa macierz czynnikowa. Źródło: strona internetowa *Matheo Analyzer*.

- Prezentacja w postaci map pozwala na ukazanie złożonych zależności pomiędzy zmiennymi. Wskazuje przykładowo grupy różnych patentów z tej samej dziedziny i ich wzajemne powiązania.



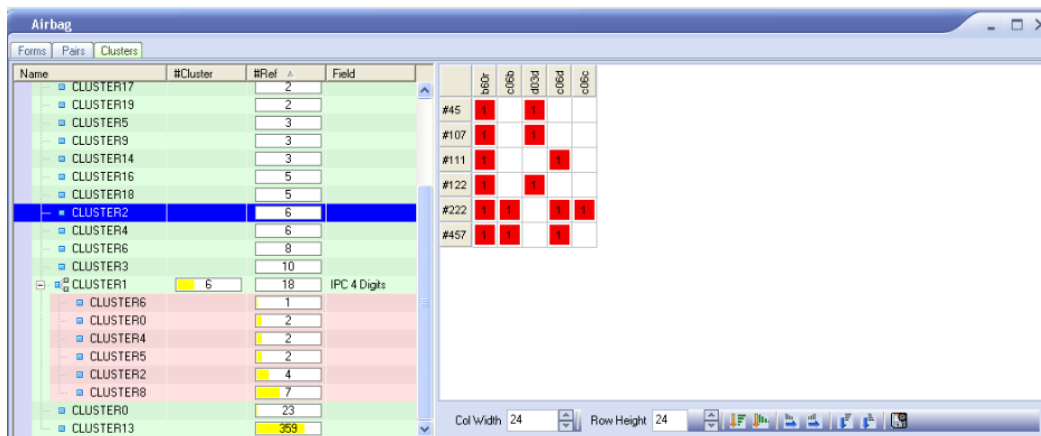
Rysunek 114: Program *Matheo Analyzer* – przykładowa mapa prezentująca powiązania pomiędzy organizacjami-właścicielami patentów. Źródło: strona internetowa *Matheo Analyzer*.

- Analogiczne możliwości pojawiają się w przypadku map asymetrycznych, które pozwalają na ukazanie powiązań pomiędzy dwoma różnymi zmiennymi.



Rysunek 115: Program *Matheo Analyzer* – przykładowa mapa prezentująca powiązania pomiędzy danymi w obrębie dwóch zmiennych. Źródło: strona internetowa *Matheo Analyzer*.

- Program pozwala też na łączenie danych w klastry (skupienia, które łączy wspólna cecha), dzięki czemu można operować na mniejszych pakietach danych i obserwować zależności na większym poziomie agregacji.



Rysunek 116: Program *Matheo Analyzer* – przykładowy klaster danych. Źródło: strona internetowa *Matheo Analyzer*.

Wersje:

- Wersja demonstracyjna umożliwia zapoznanie z funkcjonowaniem programu – jest pozbawiona niektórych funkcji, ale daje dobre możliwości ćwiczenia obsługi. Na stronie producenta jest także 18-minutowa prezentacja, wprowadzająca do pracy z programem i jego podstawowe funkcje.

- Wersja pełna płatna.

Cena: 3.450 EUR za wersję dla jednego stanowiska. Import danych patentowych wymaga dodatkowej subskrypcji programu *Matheo Patent* (600 EUR rocznie za jedno stanowisko).

11.4. Opis programu *CiteSpace*

Producent: Chaomei Chen z *Drexel University*

Adres internetowy: <http://cluster.cis.drexel.edu/~cchen/citespace>

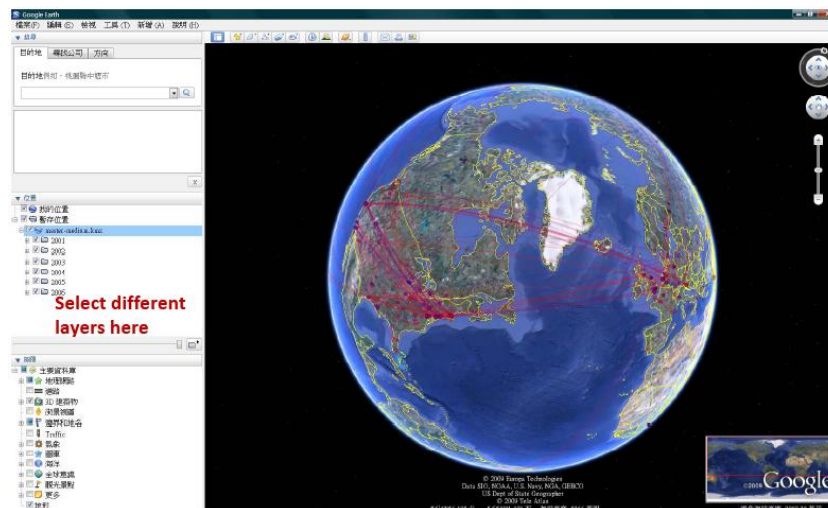
Cechy charakterystyczne: oparta na środowisku *Java* aplikacja do wizualizacji i analizy wzorców i trendów w literaturze naukowej, skoncentrowana na wyszukiwaniu punktów krytycznych dla poszczególnych dziedzin. Program został stworzony w roku 2007 i jest rozwijany przez jedną osobę. Użytkowany głównie w USA i w Chinach.

Podstawowe cechy:

- Program wymaga, by na komputerze było zainstalowane środowisko *Java* w wersji 1.6 lub późniejszej.
- Do działania wymaga połączenia z Internetem lub instalacji aplikacji, która jest aktywna przez 3-6 miesięcy (Chaomei Chen tłumaczy to chęcią zapewnienia, żeby każdy egzemplarz programu dostępny na rynku był możliwie aktualny). Autor dostarcza na życzenie wersję bez ograniczonego terminu działania.
- Korzysta głównie z bazy *Web of Science*, ale można go używać także do danych pochodzących z innych źródeł – baz *MEDLINE*, *arXiv*, *ADS* czy *NSF Award Abstract*. Dla danych pobranych z bazy *Scopus*, możliwa jest konwersja do formatu zgodnego z formatem plików wynikowych bazy *Web of Science* przy wykorzystaniu programu *Scop2ISI.exe*, dostępnego na stronie www.leydesdorff.net/software.htm.
- Główny nacisk położono na wizualizację cytowań i związków pomiędzy autorami oraz słowami kluczowymi.
- Pomoc techniczna jest dostępna w dość ograniczonym zakresie - autor zastrzega, że każdy używa jego programu na własną odpowiedzialność.
- Całość dokumentacji jest dostępna *online*; konieczność samodzielnego zdobycia wiedzy na temat użytkowania programu.

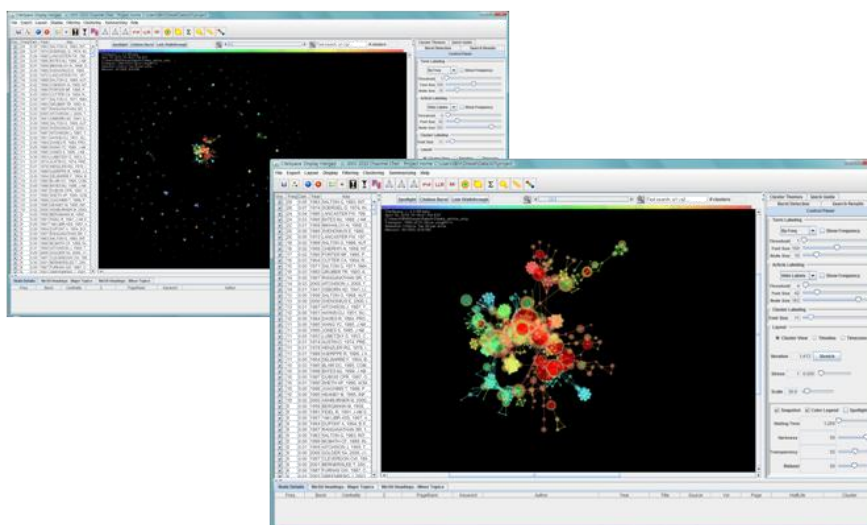
Podstawowa funkcjonalność:

- Generuje infografiki w postaci map, łącząc dane bibliometryczne z serwisem *Google Earth*.



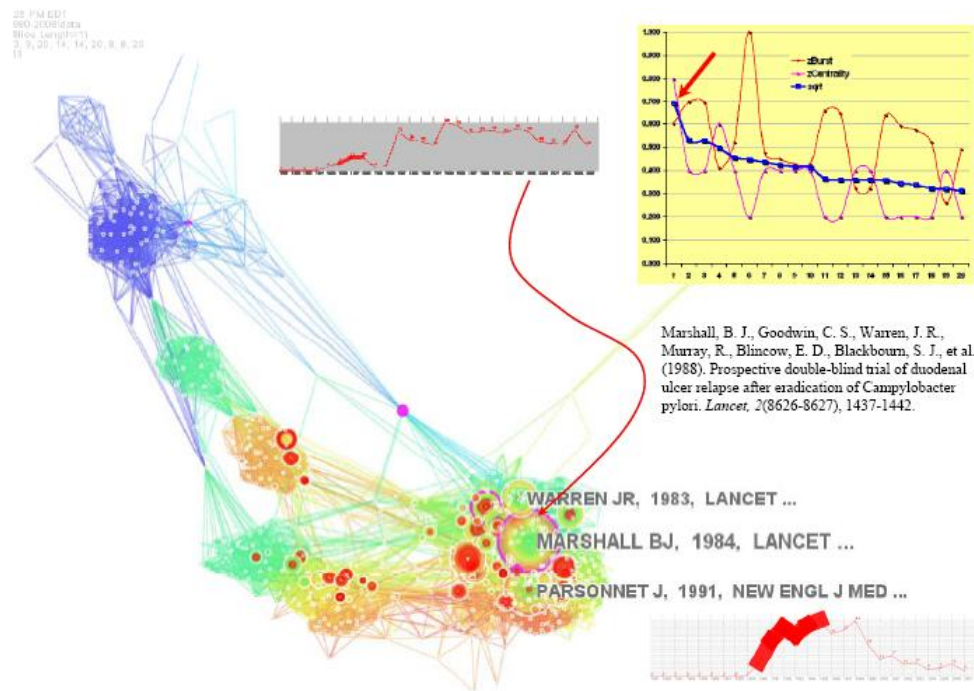
Rysunek 117: Program *CiteSpace* – mapa wygenerowana przy wykorzystaniu serwisu *Google Earth*. Źródło: Chen (2009).

- Po pobraniu danych na komputer, pozwala na ich dalsze analizowanie, zapisywanie i eksportowanie – w tym do programów do obróbki statystycznej.
- Interfejs umożliwia umieszczenie na ekranie równocześnie kilku okien do analizy danych.



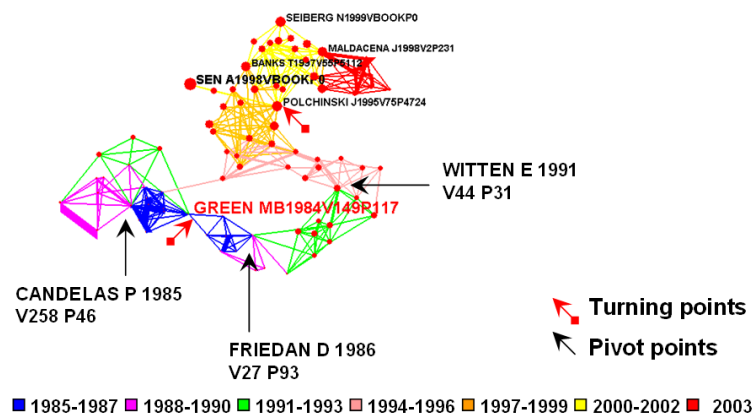
Rysunek 118: Program *CiteSpace* – przykładowe okna obrazujące równocześnie różne poziomy tej samej analizy. Źródło: strona internetowa *CiteSpace*.

- Przy pomocy analizy cytowań umożliwia identyfikację autorów, którzy stanowią „węzły” w sieci cytowań, do których odnosi się największa liczba innych tekstów.



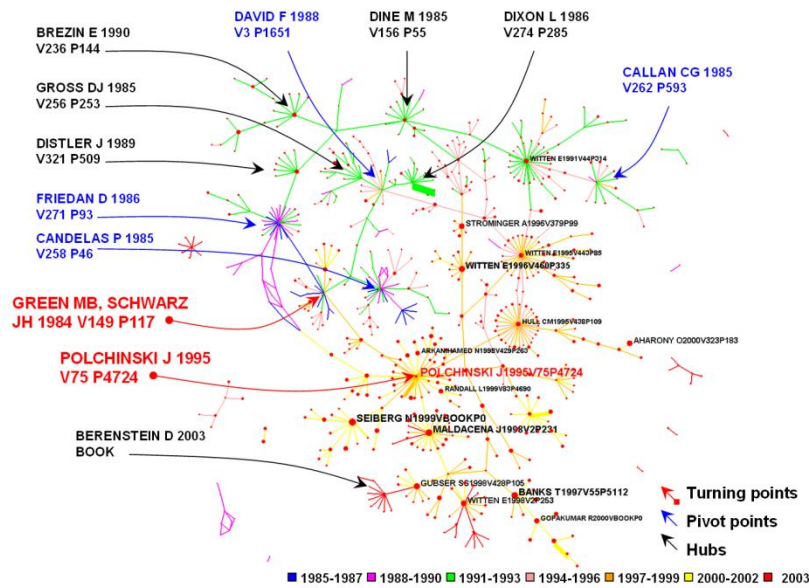
Rysunek 119: Program *CiteSpace* – schemat prezentujący kluczowych autorów w analizowanym obszarze badań. Źródło: Chen (2009).

- Program umożliwia:
 - grupowanie tematyczne analizowanych tekstów,
 - odnajdywanie przełomowych momentów w rozwoju badań,
 - określanie zmian w dominujących tematach badań w czasie,
 - identyfikację nagłych zmian wzorców publikowania w danej tematyce (ang. *burst*).



Rysunek 120: Program *CiteSpace* – schemat prezentujący analizę zmian publikacji w czasie. Źródło: strona internetowa *CiteSpace*.

- Na jednym schemacie można jednocześnie obserwować równocześnie: daty, tematy, autorów i miejsca publikacji.

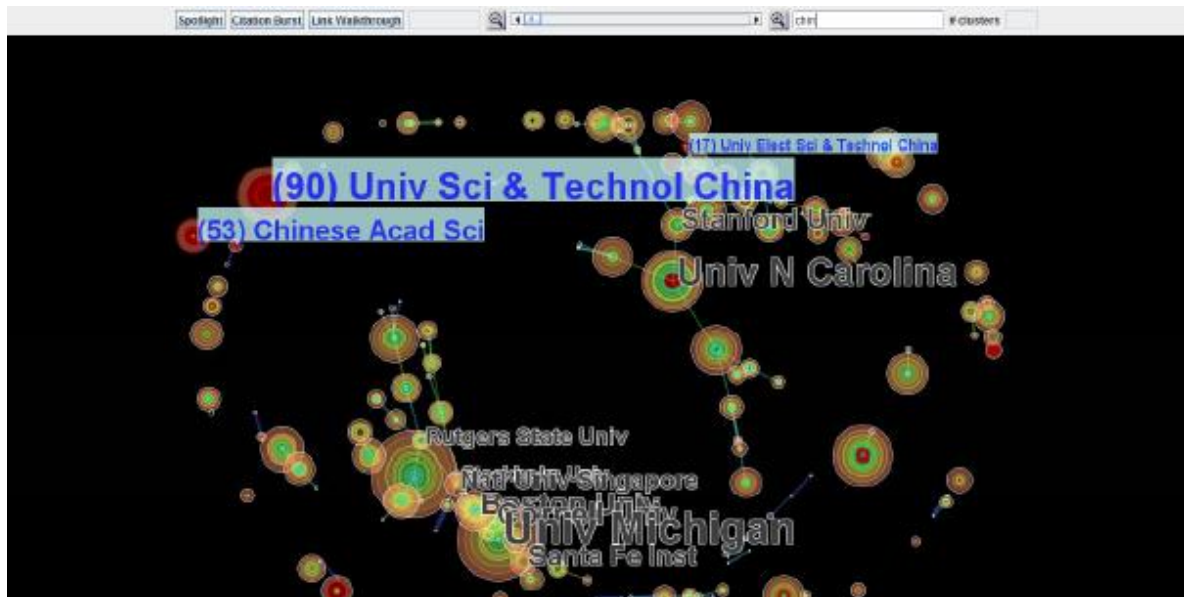


Rysunek 121: Program *CiteSpace* – schemat prezentujący zestawienie kilku serii danych. Źródło: strona internetowa *CiteSpace*.

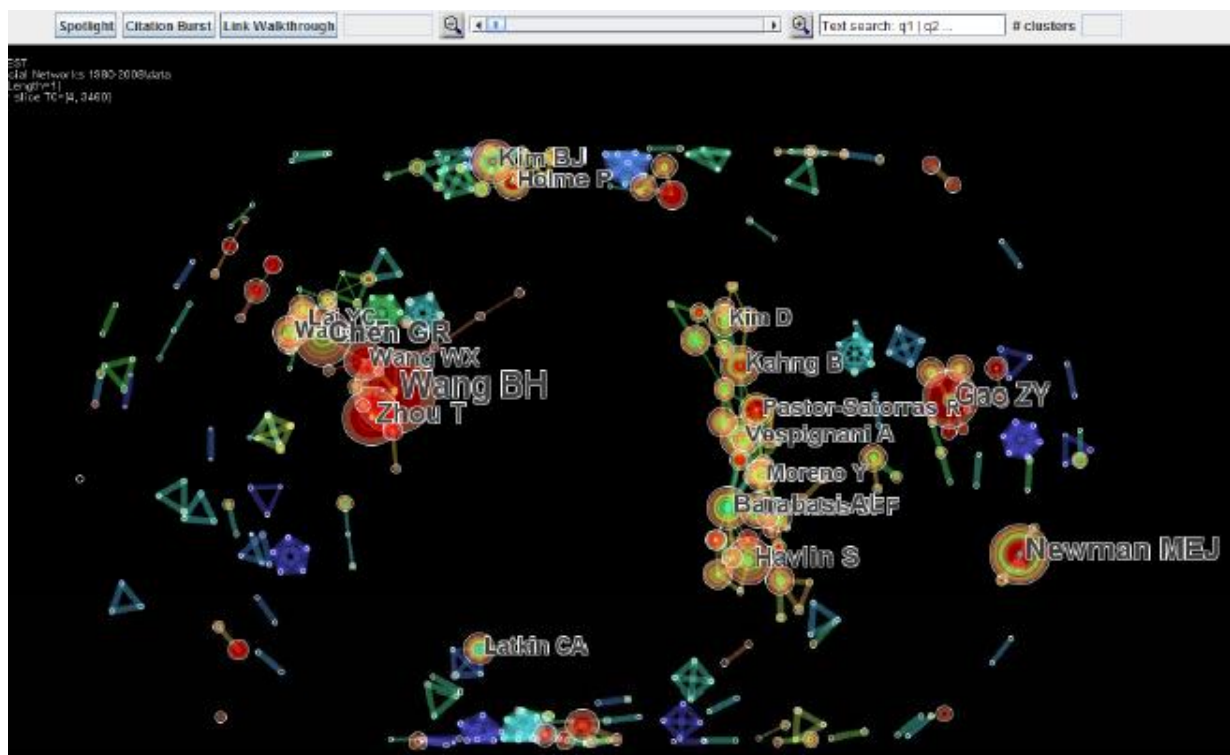
- Dzięki możliwości zdefiniowania jako węzłów na schemacie różnych pól (autor, słowo kluczowe, kategoria, instytucja, kraj), program pozwala z tych samych danych uzyskać różne przekroje i analizować różne aspekty zbioru danych bibliometrycznych.



Rysunek 122: Program *CiteSpace* – schemat prezentujący wyniki analizy z użyciem krajów jako węzłów sieci. Źródło: Chen (2009).



Rysunek 123: Program *CiteSpace* – schemat prezentujący wyniki analizy z użyciem ośrodków badawczych jako węzłów sieci. Źródło: Chen (2009).



Rysunek 124: Program *CiteSpace* – schemat prezentujący wyniki analizy z użyciem nazwisk badaczy jako węzłów sieci. Źródło: Chen (2009).

- Program pozwala na automatyzację powtarzalnych czynności, dokonywanych na zbiorach danych, np. tworzenia map lub oczyszczania danych. Program zawiera bibliotekę predefiniowanych skryptów i kwerend w języku SQL, którą można dowolnie rozbudowywać. Ze względu na wygodę użytkownika, skrypty wymagające dużej ilości czasu na wykonanie są specjalnie oznaczone jako obciążające komputer.

Wersje:

- Darmowa, dostępna na stronie autora.

Cena: Program darmowy.

11.5. Opis programu *VantagePoint*

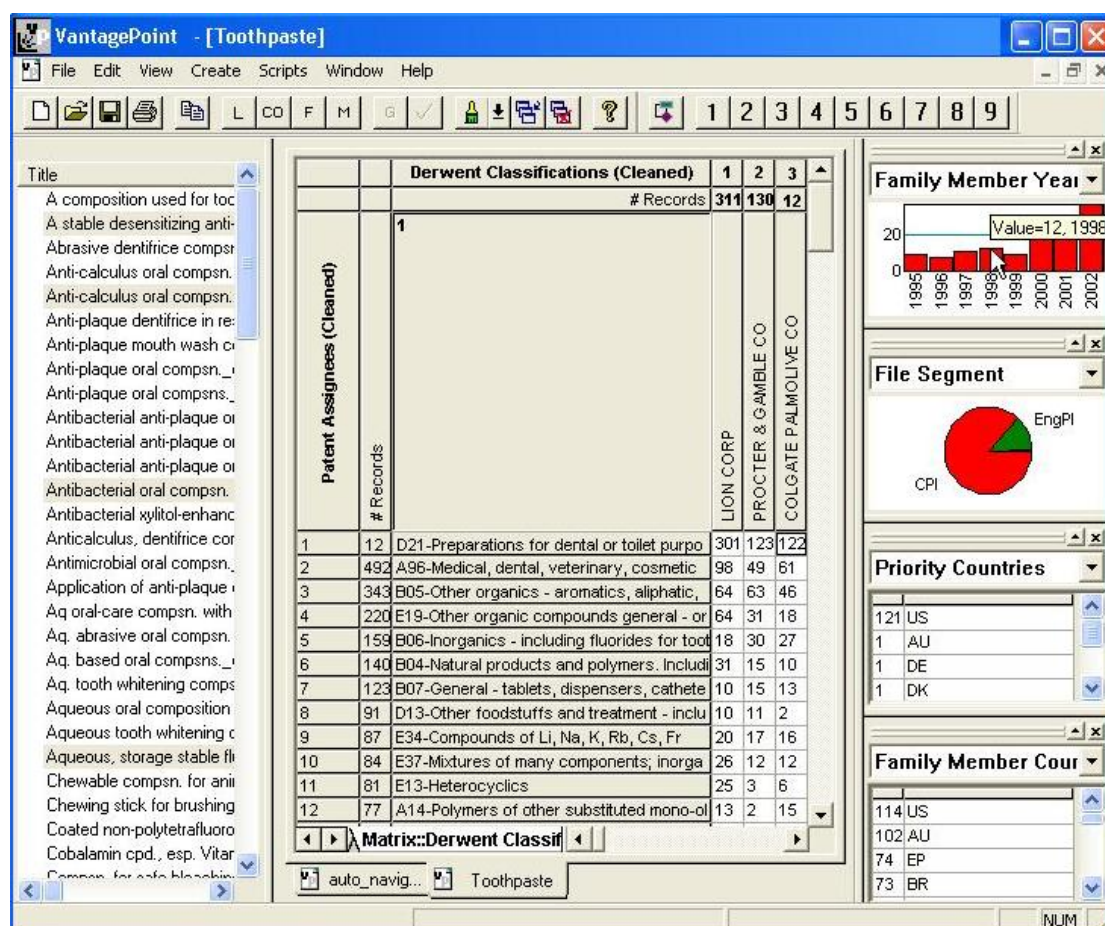
Producent: *Search Technology, Inc.*

Adres internetowy: www.thevantagepoint.com

Cechy charakterystyczne: uniwersalne narzędzie do zaawansowanej analizy zagadnień związanych z patentami i badaniami naukowymi. Duży nacisk położony na uniwersalność. Program napisany początkowo dla armii amerykańskiej, współtworzony przez badaczy, rozwijających podejścia *tech mining*, tomografię bazodanową i *Literature-Based Discovery*, obecnie stanowi *de facto* standard dla wspieranych komputerowo analiz bibliometrycznych.

Podstawowe cechy:

- Współpracuje z wieloma różnymi bazami danych i programami dzięki łatwym do zdefiniowania filtrom do importowania danych (*import filters*), których bogatą bibliotekę udostępnia producent.
- Analizy mogą być prowadzone na tekstach pisanych w języku naturalnym.
- Interfejs dostosowywany do potrzeb użytkownika – kilka okien pozwala równolegle obserwować kilka różnych aspektów danych, a także prowadzić kilka różnych analiz.



Rysunek 125: Program *VantagePoint* – interfejs użytkownika. Źródło: plik pomocy programu *VantagePoint*.

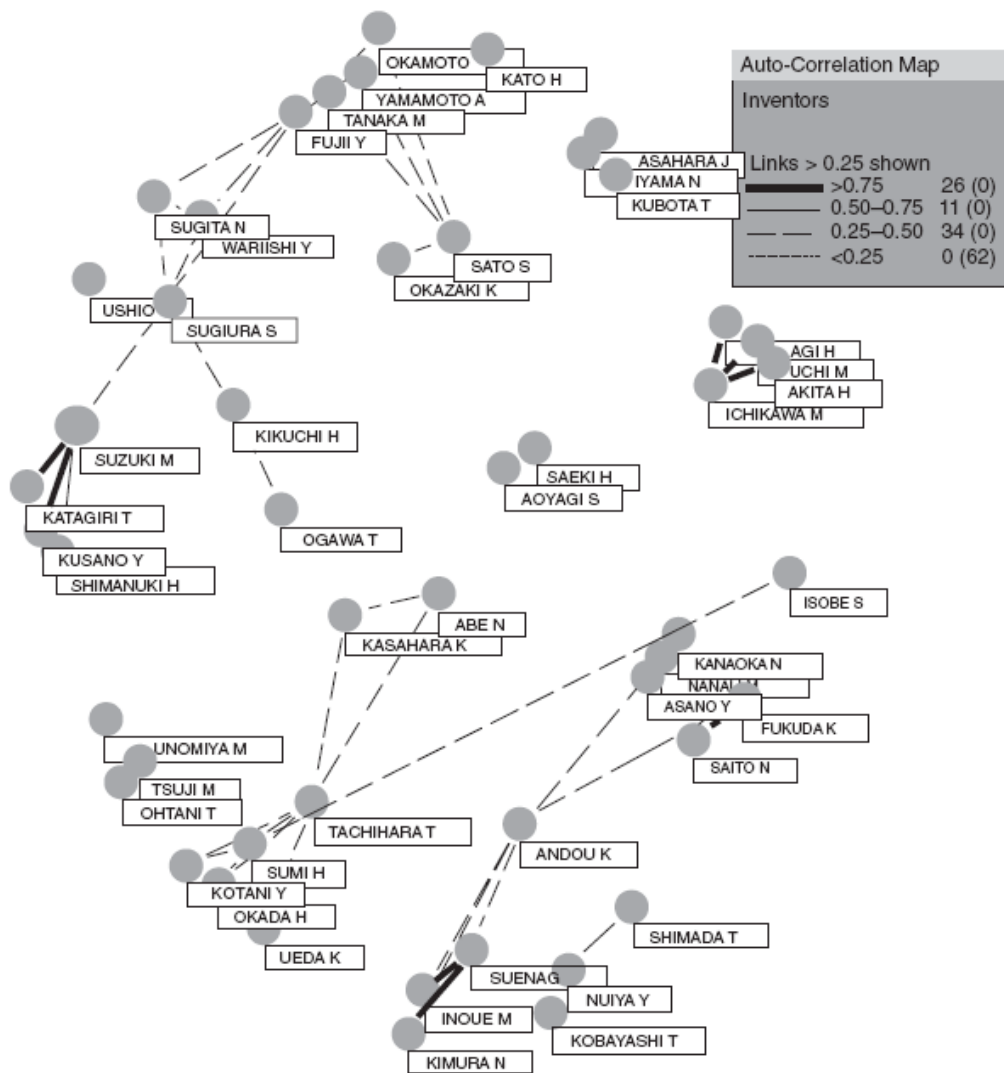
- Dodatkowy program *VP Reader* do oglądania plików stworzonych w pełnej wersji *VantagePoint*, posiadający uproszczony interfejs.
- Dobre wsparcie techniczne oraz pomoc *online*.

Podstawowa funkcjonalność:

- *Co-occurrence Matrix* – macierz pokazująca, ile rekordów w zbiorze danych zawiera te same dwa elementy. Pozwala to określić np. ile tekstów z konkretnego ośrodka badawczego zawiera dane słowa kluczowe. Funkcja ta pozwala także filtrować oraz wyróżniać wartości, pozostawiając wyłącznie wyniki powyżej określonego przez użytkownika progu, co ułatwia dostrzeżenie najmocniejszych powiązań (szczegółową dyskusję analiz krzyżowych zawiera rozdział książki, opisujący technikę *tech mining*).
- *Auto-Correlation Map* – prezentuje w formie graficznej relacje pomiędzy elementami na tej samej liście, np. wskazuje autorów, którzy często wspólnie publikują. Mapa ta pokazuje również siłę poszczególnych powiązań i częstotliwość występowania danych w zbiorze.
- *Cross-Correlation Map* – prezentuje w formie graficznej relacje pomiędzy elementami na różnych listach, np. może pokazać, jakie organizacje czy ośrodki publikują na ten

sam temat. Mapa ta pokazuje również siłę poszczególnych powiązań. Jest to narzędzie analogiczne do poprzedniego, ale umożliwia zestawienie dwóch różnych serii danych.

- *Factor Matrix* – macierz pozwalająca zidentyfikować ukryte zmienne - czynniki w zbiorze danych przy wykorzystaniu techniki *PCA*.
- *Factor Map* – pokazuje w formie graficznej powiązania pomiędzy grupami obiektów w oparciu o wyniki analizy czynnikowej. Mapa pokazuje również siłę poszczególnych powiązań. Jest to prosty i przejrzysty sposób graficznej prezentacji danych. Z mapy można przykładowo odczytać, które ośrodki najczęściej współpracują ze sobą przy określonej tematyce badań.



Rysunek 126: Program VantagePoint – przykładowy wynik analiz w postaci mapy auto-korelacji. Źródło: Porter, Cunningham (2005: 232).

- W celu przyspieszenia obróbki danych, pozwala wyodrębnić z dużego zbioru danych mniejsze podzbiory do analizy.

- Pozwala eksportować dane m.in. do formatu tekstowego oraz do arkusza *Microsoft Excel*.
- Posiada funkcję łączącą dane pochodzące z różnych źródeł (baz) w jednym zbiorze.
- Poprzez wprowadzenie predefiniowanych lub definiowanych samodzielnie tzw. *metatag*-ów ułatwia odnajdywanie i zaznaczanie wpisów o takiej samej wartości w różnych polach, np. nazwa tej samej firmy lub tego samego ośrodka w polu miejsca publikacji i właściciela patentu.
- Oczyszczanie danych może być przeprowadzone na kilka sposobów:
 - usuwanie powtarzających się rekordów,
 - łączenie powtarzających się rekordów,
 - usuwanie i łączenie powtarzających się rekordów, które nie są dokładnie identyczne,
 - usuwanie i łączenie powtarzających się rekordów, które zawierają określenia synonimiczne lub podobne (także według własnoręcznie zdefiniowanych przez użytkownika list synonimów).
- Podczas każdego oczyszczania listy *VantagePoint* tworzy nową jej wersję, co powoduje, że lista pierwotna pozostaje niezmieniona. Pozwala to na bezpieczne eksperymentowanie, bez obawy o modyfikację pierwotnego źródła danych.
- Oczyszczanie listy jest kontrolowane w czasie rzeczywistym i pozwala na ręczną modyfikację każdego pojedynczego kroku.
- *VantagePoint* dzięki obsłudze skryptów w języku *Visual Basic for Applications* pozwala na automatyzację powtarzalnych czynności, prowadzonych na zbiorach danych, np. tworzenie map, oczyszczanie danych czy łączenie list. Program zawiera dużą bibliotekę predefiniowanych skryptów, którą można dowolnie rozbudowywać.
- Możliwa jest też modyfikacja sposobów importu, obróbki i eksportu danych poprzez wykorzystanie predefiniowanych lub tworzonych przez użytkownika filtrów. Pozwalają one uprościć i przyspieszyć proces wyciągania danych z zewnętrznych baz (obecnie program obsługuje kilkadziesiąt baz).
- W najnowszych wersjach *VantagePoint* dodano także aplikację, wspomagającą graficzną prezentację wyników.

Wersje:

- Wersja pełna standardowa.
- *VP Reader* czyli program do odczytywania raportów i analiz utworzonych w pełnej wersji *VantagePoint*.
- *TechOASIS* – wersja *VantagePoint* przeznaczona wyłącznie do użytku dla instytucji rządowych w Stanach Zjednoczonych.
- *Derwent Analytics* – wersja *VantagePoint* przystosowana do analiz patentowych, pracuje z bazą patentów firmy *Thomson Reuters*.

Cena: 7.500 USD za jedno stanowisko - opłata jednorazowa, obejmująca 12-miesięczne wsparcie techniczne, dostęp do aktualizacji oraz bibliotek filtrów i skryptów. Przedłużenie umowy wsparcia technicznego na kolejny rok kosztuje 3.100 USD.

11.6. Opis programu HistCite

Producent: Thomson Reuters

Adres internetowy:

http://thomsonreuters.com/products_services/science/science_products/a-z/histcite/

Cechy charakterystyczne: program stworzony przez Eugene Garfielda, pioniera badań bibliometrycznych. Pozwala na przetwarzanie wyników wyszukiwań, pobranych z platformy *Web of Knowledge* (w tym bazy danych *Web of Science*) oraz tworzenie zestawień tabelarycznych i graficznych, odpowiadających prostym analizom bibliometrycznym, opisanym w rozdziale 8 książki.

Podstawowe cechy:

- Współpracuje z bazami dostępnymi na platformie *Web of Knowledge*, w tym bazą *Web of Science*.
- Program wymaga wykorzystania przeglądarki *Microsoft Internet Explorer*.
- Oferuje możliwości prostych analiz - zestawień liczby publikacji i cytowań.
- Generuje historiografy - wykresy, odzwierciedlające przemiany sieci cytowań w czasie.
- Pozwala na eksport danych do innych programów, w tym arkuszy kalkulacyjnych i narzędzi do analizy danych sieciowych.
- Analizy mogą być prowadzone na tekstach pisanych w języku naturalnym.

#	Institution	Recs	Percent	TLCS	TGCS
1	Garfield	330	66.1	105	3177
2	Inst Sci Informat	118	23.6	19	6530
3	The Scientist	14	2.8	19	139
4	Unknown	10	2.0	0	19
5	ISI	9	1.8	4	59
6	Russian Acad Sci	7	1.4	13	118
7	Thomson ISI	4	0.8	0	21
8	SCIENTIST	3	0.6	0	6
9	Thomson Sci	3	0.6	0	323
10	Washington State Univ	3	0.6	6	45
11	Drexel Univ	2	0.4	0	8
12	Inst Informat Sci	2	0.4	5	56
13	Univ Aquila	2	0.4	4	17
14	GARFIELD CO	1	0.2	0	0
15	HLTH SERV & HLTH EDUC	1	0.2	0	0
16	Inst Sci Informat Inc	1	0.2	0	0
17	IRCCS	1	0.2	1	1
18	Kitt Peak Natl Observ	1	0.2	1	3
19	Leiden Univ	1	0.2	0	10
20	Oncol Business Unit	1	0.2	0	0
21	RAVENSWOOD CITY SCH DIST	1	0.2	0	0
22	SIMMONS COLL	1	0.2	0	8

Rysunek 127: Program *HistCite* – przykładowy wynik analizy afiliacji autorów publikacji.
Źródło: opracowanie własne.

Grand Totals: LCS 153, LCSx 0, GCS 10341, SCS n/a, CR 6323, NA 551
Collection span: 1987 - 2011 (25 years)

Records: 499, Authors: 34, Journals: 55, Cited References: 4867, Words: 1920

#	Author	Recs	Percent	T LCS	T LCS/t	T LCSx	T GCS	T GCS/t	T LCR	T LCSb	T LCSe
1	Garfield E	497	99.6	153	8.80	0	10341	513.71	153	125	
2	WELLJAMSDOROF A	9	1.8	13	0.63	0	110	5.25	6	12	0
3	Pudovkin AI	7	1.4	13	1.39	0	118	12.68	10	9	0
4	Small H	4	0.8	0	0.00	0	12	0.55	0	0	0
5	Istomin VS	3	0.6	6	0.64	0	45	4.94	4	5	1
6	Marchionini G	2	0.4	0	0.00	0	8	0.53	0	0	0
7	Melino G	2	0.4	4	0.27	0	17	1.14	1	4	0
8	Abt HA	1	0.2	1	0.10	0	3	0.30	0	1	0
9	Cortex	1	0.2	0	0.00	0	0	0.00	0	0	0
10	Davis C	1	0.2	0	0.00	0	8	0.53	0	0	0
11	Engelbart D	1	0.2	0	0.00	0	8	0.53	0	0	0
12	Fagan J	1	0.2	0	0.00	0	0	0.00	0	0	0
13	GARFIELD EF	1	0.2	0	0.00	0	0	0.00	0	0	0
14	GARFIELD EP	1	0.2	0	0.00	0	0	0.00	0	0	0
15	Glanzel W	1	0.2	0	0.00	0	10	1.25	0	0	0
16	Griffith B	1	0.2	0	0.00	0	8	0.53	0	0	0
17	Grimwade A	1	0.2	0	0.00	0	1	0.10	0	0	0
18	Hargens LL	1	0.2	0	0.00	0	0	0.00	0	0	0
19	Harmon G	1	0.2	0	0.00	0	0	0.00	0	0	0
20	HICKS J	1	0.2	0	0.00	0	0	0.00	0	0	0
21	KNE T	1	0.2	0	0.00	0	0	0.00	0	0	0
22	Lievrouw LA	1	0.2	0	0.00	0	0	0.00	0	0	0
23	Lynch C	1	0.2	0	0.00	0	8	0.53	0	0	0

Rysunek 128: Program *HistCite* – przykładowy wynik analizy autorów publikacji. Źródło: opracowanie własne.

Grand Totals: LCS 153, LCSx 0, GCS 10341, SCS n/a, CR 6323, NA 551
Collection span: 1987 - 2011 (25 years)

Records: 499, Authors: 34, Journals: 55, Cited References: 4867, Words: 1920

Cited Reference List (4867) including 85 records, 9 on this page (Hide 9 records)

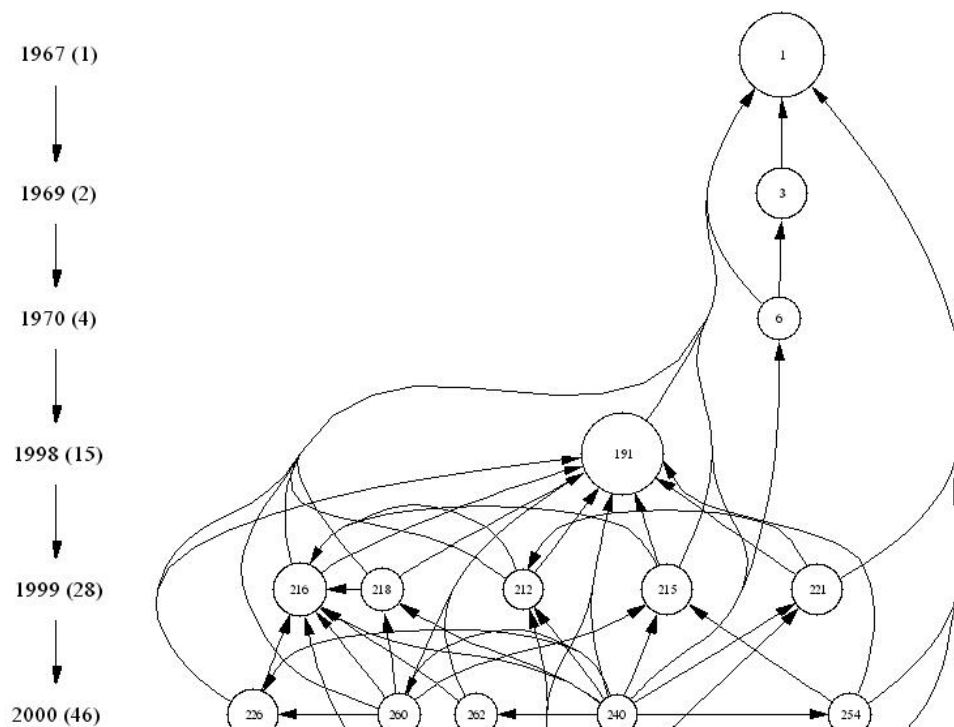
#	Author / Year / Journal	Recs	Percent
191	MESSING J, 1982, GENE, V19, P269	3	0.6
192	MESSING J, 1983, METHOD ENZYMOLOG, V101, P20	3	0.6
193	MICHELL RH, 1975, BIOCHIM BIOPHYS ACTA, V415, P81	3	0.6
194	MORAVCSI MJ, 1974, LEONARDO, V7, P255	3	0.6
195	MORAVCSIK MJ, 1975, CURRENT CONTENTS SOC, V5, P86	3	0.6
196	NISHIZUKA Y, 1984, NATURE, V308, P693	3	0.6
197	Ophof T, 1999, CARDIOVASC RES, V41, P1	3	0.6
198	PENDLEBURY D, 1989, SCIENTIST, V3, P19	3	0.6
199	PRICE DJD, 1965, SCIENCE, V149, P510	3	0.6
200	SHARPE WF, 1964, J FINANC, V19, P425	3	0.6
201	SINGLETON VL, 1985, CURRENT CONTENTS, V16, P18	3	0.6
202	SMALL H, 1994, SCIENTOMETRICS, V30, P229	3	0.6
203	SMALL HG, 1978, SOC STUD SCI, V8, P327	3	0.6
204	SOKOLOFF L, 1977, J NEUROCHEM, V28, P897	3	0.6
205	SPITZER RL, 1978, ARCH GEN PSYCHIAT, V35, P773	3	0.6
206	SWANSON DR, 1987, J AM SOC INFORM SCI, V38, P228	3	0.6
207	SWINBURNE JK, 1983, J INFORM SCI, V6, P75	3	0.6
208	THACKRAY A, 1986, CONT CLASSICS ENG AP	3	0.6
209	VIRGO JA, 1977, LIBR QUART, V47, P415	3	0.6
210	WEINBERG S, 1986, LESSON QUANTUM THEOR, P227	3	0.6
211	WIEDOEFT R, 1966, ASCAP BIOGRAPHICAL D, P789	3	0.6

Rysunek 129: Program *HistCite* – lista cytowań publikacji z odnośnikami do bazy *Web of Science*. Źródło: opracowanie własne.

Podstawowa funkcjonalność:

- Import danych, pobranych z bazy *Web of Science* lub innych baz dostępnych na platformie *Web of Knowledge* i zapisanych w pliku tekstowym.
- Tworzenie list (zestawień tabelarycznych), prezentujących wyniki wyszukiwań (liczba publikacji i liczba cytowań do nich) w podziale na:
 - autorów,
 - czasopisma,
 - lata publikacji,
 - typy dokumentów (artykuł, wystąpienie konferencyjne itd.),
 - języki publikacji,
 - instytucje, z których pochodzą autorzy,
 - kraje, z których pochodzą autorzy,
 - słowa, występujące w tytułach analizowanych publikacji,
 - cytowania analizowanych publikacji.
- Udostępnianie bezpośrednich odnośników do publikacji i ich cytowań na platformie *Web of Knowledge* (funkcjonalność nie działa poprawnie w odniesieniu do baz dostępnych na polskich uczelniach za pośrednictwem *Wirtualnej Biblioteki Nauki* ze względu na odmienne adresy odnośników do zasobów bazy).
- Generowanie wykresów, prezentujących rozkład cytowań w czasie czyli historyczny rozwój analizowanego obszaru badań (historiografów) - wykresy prezentują sieci cytowań poszczególnych publikacji, oznaczanych numerami (nie ma możliwości wygenerowania sieci zależności pomiędzy autorami, instytucjami lub krajami).
- Eksport danych w formacie *CSV*, pozwalającym na dalsze przetwarzanie danych w arkuszu kalkulacyjnym oraz programach do analizy sieci społecznych.
- Eksport danych w formacie *HTML*, pozwalającym na prezentowanie wyników analiz.
- Możliwość wygenerowania macierzy cytowań do dalszego przetwarzania w innych programach.

Wadą programu jest brak wbudowanych funkcji czyszczenia danych, ujednolicania autorów i afiliacji - jest jednak możliwa modyfikacja pliku tekstowego, importowanego do programu, przy wykorzystaniu zewnętrznych edytorów tekstu. Brak możliwości generowania złożonych zestawień macierzowych (np. zestawień autorów z poszczególnych instytucji, instytucji z odzwierciedleniem lat ich aktywności). Brak możliwości analizy słów kluczowych, pojęć występujących w streszczeniach artykułów ani kodów klasyfikacyjnych, opisujących publikacje - co wynika m.in. z ograniczeń bazy *Web of Science*, która nie oferuje większości tych danych. Ze względu na wymienione ograniczenia, *HistCite* nie pozwala na prowadzenie zaawansowanych analiz bibliometrycznych, które zostały opisane we wcześniejszych rozdziałach niniejszej książki. Jego możliwości ograniczone są do prostych analiz (generowania zestawień danych, opisywanych w rozdziale 8 książki). Zaletą jest jednak prostota obsługi, szybkość działania i integracja z platformą *Web of Knowledge*.



Rysunek 130: Program HistCite – przykładowy historiograf, prezentujący sieć cytowań w czasie.

Źródło: <http://www.garfield.library.upenn.edu/histcomp/guide.html>

Wersje:

- Darmowa, dostępna na stronie internetowej firmy.

Cena: Program darmowy.

11.7. Opis programu PatentNet

Producent: *Chen Shih Sheng*

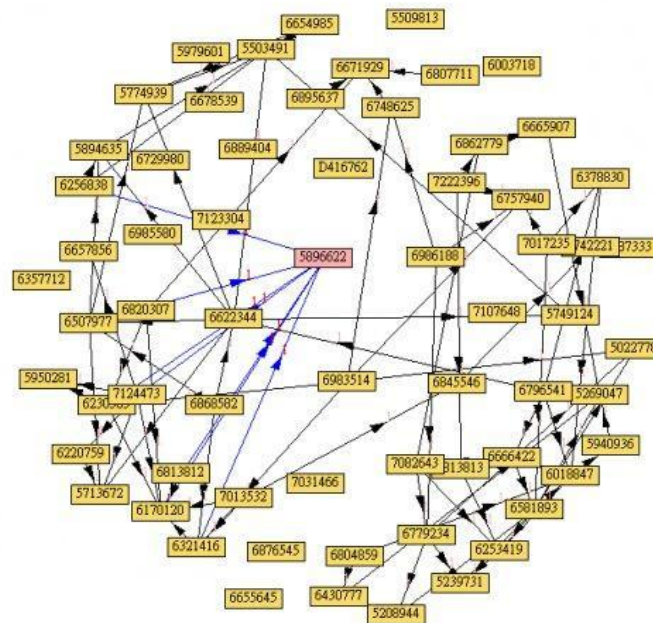
Adres internetowy: <http://sourceforge.net/projects/patentnet/>

Cechy charakterystyczne: oparta na środowisku *Java* aplikacja do pobierania i analizy opisów patentowych, oferująca możliwość szybkiego pobrania dokumentów z baz *esp@cenet* i *Google Patents* oraz stworzenia graficznych sieci zależności między patentami, wynalazcami i właścicielami patentów. Bezpłatny program został stworzony w 2008 roku i jest rozwijany przez jedną osobę.

Podstawowe cechy:

- Program wymaga instalacji na komputerze platformy *Java* i odpowiednich komponentów firmy *Microsoft*.
- Pozwala na pobranie danych o patentach i ich cytowaniach z baz Europejskiego Urzędu Patentowego i bazy *Google Patents* (zawierającej dane z *USPTO*).
- Generowanie graficznych sieci zależności pomiędzy patentami, wynalazcami i właścicielami patentów.

- Możliwość wyboru "poziomów zagnieżdżeń" w celu analizy cytowań cytowań, współwynałazców, firm-partnerów w projektach badawczo-rozwojowych.



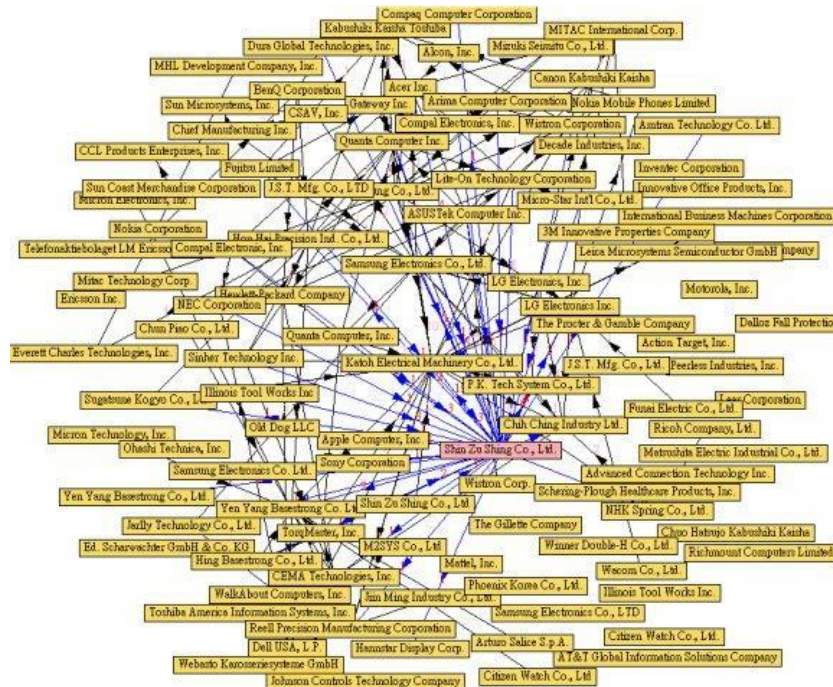
Rysunek 131: Program PatentNet – wykres sieci cytowań patentów. Źródło: strona internetowa PatentNet.

Podstawowa funkcjonalność:

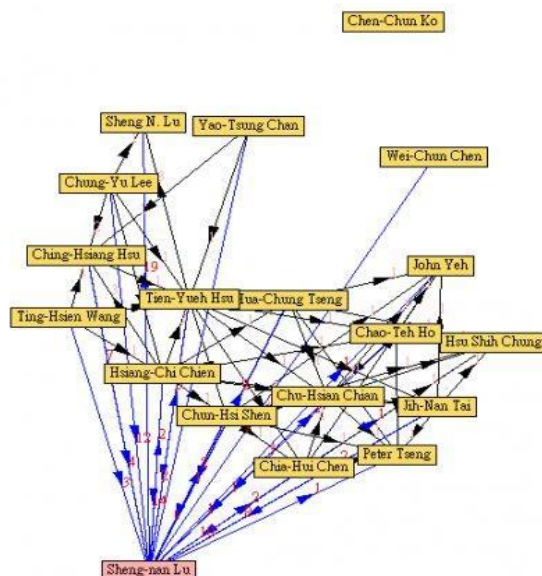
- Pobieranie danych z baz *esp@cenet* i *Google Patents* - polecenie wyszukiwania wpisywane jest do programu *PatentNet* i przekazywane do odpowiedniej bazy, a następnie rezultaty wyszukiwania są zapisywane na dysku lokalnym do celów dalszych analiz.
- Gromadzenie danych o patentach i ich cytowaniach z możliwością dalszego przetwarzania, także po odłączeniu od Internetu.
- Możliwość wyboru "poziomów zagnieżdżeń" w odniesieniu do wynalazców, właścicieli, cytowań wymienionych w zgłoszeniu patentowym i cytowań do zgłoszenia patentowego. Przykładowo, wybór wartości "1" w odniesieniu do wynalazców oznacza, że do zbioru danych włączone zostaną także wszystkie chronione patentami wynalazki, które zostały stworzone przez autorów wynalazków z podstawowego zbioru patentów, który uzyskano w wyniku zdefiniowanego wyszukiwania, a wybór wartości "2" oznacza, że dodatkowo uwzględnione zostaną również patenty osób, które były współwynałazcami patentów, stworzonych przez osoby, które były również twórcami podstawowych, objętych wyszukiwaniem wynalazków. Analogicznie, można w łatwy sposób zidentyfikować i objąć analizami wszystkie pozostałe patenty firmy-właściciela odnalezione w wyniku wyszukiwania wynalazku (poziom 1), jak również patenty uzyskane przez podmioty, z którymi ta firma współpracuje badawczo i jest współwłaścicielem niektórych patentów (poziom 2). W odniesieniu do analizy cytowań, możliwe jest odnalezienie nie tylko zgłoszeń

patentowych cytujących dany patent (poziom 1), ale też nowszych zgłoszeń, cytujących te wcześniejsze zgłoszenia, które z kolei odwołują się do interesującego nas patentu (poziom 2). Opisana funkcjonalność pozwala na odzwierciedlenie złożonych sieci zależności.

- Tworzenie zestawień słów kluczowych, występujących w patentach.
- Generowanie zestawień danych nagłówkowych patentów w formacie, który może być dalej przetwarzany w arkuszach kalkulacyjnych.
- Budowa graficznej sieci zależności pomiędzy patentami (prezentującej numery patentów i wzajemne cytowania) (ang. *patent number network*).
- Budowa graficznej sieci zależności pomiędzy wynalazcami (ang. *inventor network*).
- Budowa graficznej sieci zależności pomiędzy właścicielami patentów (ang. *assignee network*).
- Prosty, ale mało przejrzysty interfejs użytkownika i ograniczona dokumentacja techniczna.



Rysunek 132: Program PatentNet – wykres zależności pomiędzy organizacjami-właścicielami patentów, ujawnionych poprzez analizę sieci cytowań patentów. Źródło: strona internetowa PatentNet.



Rysunek 133: Program *PatentNet* – wykres zależności pomiędzy wynalazcami-twórcami patentów, ujawnionych poprzez analizę sieci cytowań patentów. Źródło: strona internetowa *PatentNet*.

Wersje:

- Darmowa, dostępna na stronie internetowej twórcy.

Cena: Program darmowy.

11.8. Porównanie zastosowań opisanych programów

W celu ułatwienia wyboru najbardziej właściwego rozwiązania, poniżej umieszczona została tabela pokazująca, na które pytania badawcze można uzyskać odpowiedź przy użyciu poszczególnych, opisanych wcześniej programów.

Pytania badawcze	<i>VantagePoint</i>	<i>Matheo</i>	<i>CiteSpace</i>	<i>HistCite</i>	<i>PatentNet</i>
Co?	+	+	+	+	+
Co dokładnie?	+	+	+	-	-
Kto?	+	+	+	+	+
Gdzie?	+	+	+	+	-
Kiedy?	+	+	+	+	-
Z kim?	+	+	+	+	+
Z jaką organizacją?	+	+	+	-	+
Ile?	+	+	+	+	+
Jak zmienia się w czasie?	+	+	+	+	-
Jak będzie się rozwijać?	-/+	-/+	-/+	-/+	-

Pytania badawcze	<i>VantagePoint</i>	<i>Matheo</i>	<i>CiteSpace</i>	<i>HistCite</i>	<i>PatentNet</i>
Jak popularne w porównaniu do?	+	+	+	+	+
Jakie są nisze?	+	+	+	-	-
Jakie są nowe popularne trendy?	-/+	+	-/+	-	-
Z czym związane?	+	+	+	-	-
W jaki sposób związane?	+	+	+	-	-
Co uzupełnia?	+	+	+	-	-
Jakich wariantów nie analizowano?	-/+	-/+	-/+	-	-
Kto uzupełnia?	+	+	+	-	-
Kto może być dostawcą?	+	+	+	-	-
Kto może być odbiorcą?	+	+	+	-	-
Kto jest konkurentem?	+	+	+	-	+
Kto może być konkurentem?	+	+	+	-	-
Kto może być pracownikiem?	+	+	+	+	+
Kto może być sukcesorem?	+	+	+	+	+
Jakie X ma strategie badawcze?	+	+	+	-	-
Jak ważne?	-/+	+	-/+	+	+
Jak łącznie ważne?	-/+	+	-/+	+	+
Kto jest ważny?	-/+	+	-/+	+	+
Do jakiej wiedzy nawiązuje?	-/+	+	-/+	+	+
Do których badań nawiązuje?	-/+	+	-/+	+	+

Tabela 55: Pytania badawcze w bibliometrii a możliwości analityczne oprogramowania. Źródło: opracowanie własne.

Tabela 55 wskazuje, że do wielu typowych zastosowań bibliometrycznych nadają się wszystkie opisane programy. Przy identyfikacji nowych tendencji badawczych oraz analizy cytowań, najlepiej sprawdzi się program *CiteSpace*, który został opracowany właśnie do tego typu zastosowań. Jednocześnie program ten może nie być najbardziej dogodnym narzędziem do niektórych podstawowych analiz – ma mało przyjazny dla użytkownika interfejs, który zniechęca praktyków, nieznających specjalistycznej terminologii, dlatego najlepiej stosować

go jako narzędzie uzupełniające. Tabela 55 obejmuje pola, oznaczone symbolami „-/+”, które dotyczą przypadków, gdy określone oprogramowanie może pomóc w udzieleniu odpowiedzi na określone pytanie pod warunkiem wykorzystania dodatkowego narzędzia (np. dalszej obróbki danych w programie do analiz statystycznych) lub filtra/skryptu/klasyfikacji (np. przetwarzanie danych o cytowaniach, analiza zgodna z podejściem *TRIZ*, identyfikacja nowych, narastających zjawisk w badaniach).

Istotnym kryterium przy wyborze odpowiedniego narzędzia informatycznego powinno być wsparcie dla baz danych, które mają być wykorzystywane w pracach analitycznych. Najszerze wsparcie dla różnych formatów danych oferuje program *VantagePoint* dzięki rozbudowanej kolekcji *import filters*. *CiteSpace* jest dedykowane do obsługi bazy *Web of Science* (choć z pewnymi ograniczeniami może wspierać również bazy *MEDLINE* i *Scopus*), a *Matheo Analyzer* – do baz patentowych *USPTO* i *esp@cenet* (inne bazy można obsłużyć, definiując nowy format importu). Ograniczeniem programu *VantagePoint* jest z kolei brak standardowego interfejsu, pozwalającego na import danych z baz *USPTO* i *esp@cenet*, ale taka operacja jest możliwa w oparciu o odpowiednią konfigurację programu przez użytkownika. Dane patentowe pozwoli z kolei przetwarzać i wizualizować program *PatentNet*, nie oferujący jednak wielu dodatkowych funkcji, typowych dla narzędzi bibliometrycznych. *HistCite* jest wreszcie niszowym narzędziem, dedykowanym dla produktów firmy *Thomson Reuters* i pozwalającym na generowanie prostych zestawień i graficznych prezentacji cytowań publikacji w wybranych okresach historycznych.

Program *VantagePoint* jest najpopularniejszym narzędziem wspierającym *tech mining*, często opisywanym w specjalistycznych publikacjach i wykorzystywanym przez liczne instytucje rządowe i przedsiębiorstwa na całym świecie, współtworzonym m.in. przez osoby, których nazwiska były wielokrotnie wymieniane w niniejszej książce - Alana L. Portera (reprezentanta firmy *Search Technologies, Inc.*, profesora *Georgia Institute of Technology*) i Ronalda N. Kostoffa (reprezentanta pierwszego nabywcy programu – sektora obronności Stanów Zjednoczonych oraz posiadacza kluczowych patentów, opisujących algorytmy analityczne). *CiteSpace* zostało z kolei stworzone przez Chaomei Chena, wybitnego specjalistę z dziedziny wizualizacji danych i stanowi praktyczne wdrożenie wielu nowatorskich rozwiązań i algorytmów w zakresie analizy i wizualizacji danych bibliometrycznych. Europejski program *Matheo Analyzer* był inspirowany możliwościami programu *VantagePoint*, oferuje wiele zbliżonych funkcji, koncentrując się na ważnym komercyjnie zastosowaniu jakim jest analiza patentów.

11.9. Przegląd innych programów, wspierających analizy bibliometryczne

Pełne porównanie wszystkich dostępnych na rynku narzędzi informatycznych znacznie wykraczałoby poza ramy niniejszej publikacji. Znajomość większego wycinka rynku pozwala jednak na wybór oprogramowania w zależności do potrzeb osób prowadzących analizy. Jako uzupełnienie wcześniej zaprezentowanych informacji, przygotowano zostało tabelaryczne zestawienie podstawowych informacji na temat innych, dostępnych na rynku programów, ułatwiające dotarcie do informacji na temat konkretnych aplikacji. Dostępne na rynku programy zostały podzielone na dwie grupy: płatne i darmowe, zaprezentowane w formie dwóch odrębnych zestawień tabelarycznych.

Nazwa	Producent	Strona www	Zastosowania	Charakterystyka
<i>Aureka</i>	<i>Thomson Reuters</i>	http://ip.thomsonreuters.com/raining/aureka	wyszukiwanie informacji i przetwarzanie ich w celu poprawy efektywności funkcjonowania organizacji	obsługuje dane częściowo ustrukturalizowane, dobre funkcje wizualizacyjne, słabo obsługuje język naturalny, tańsza i uproszczona wersja interfejsu dla użytkowników niezaawansowanych, obsługuje patenty tylko z bazy USA
<i>Bizint</i>	<i>Bizint Solutions</i>	www.bizcharts.com	tworzenie i dystrybuowanie raportów z danych pochodzących z baz patentowych	zestaw kilku narzędzi do analiz różnych dziedzin (<i>Smart Charts Viewer, for Patents, for Drug Pipelines</i>), obsługuje różne bazy (<i>WPI, CA, PatBase, Questel, Orbit.com, Total Patent, STN, GenomeQuest</i>)
<i>ClearForest (ClearResearch, OpenCalais)</i>	<i>Thomson Reuters</i>	www.clearforest.com	wyszukiwanie i porządkowanie informacji pochodzących z ogólnodostępnych źródeł, przygotowywanie raportów przydatnych w zarządzaniu organizacją	możliwość analizy nieustrukturalizowanego tekstu, eksportowanie wyników w formie strukturalizowanej, skomplikowana obsługa i parametryzacja
<i>Cognos</i>	<i>IBM Cognos</i>	http://www-01.ibm.com/software/data/cognos	łączenie, eksploracja i wizualizacja danych pochodzących z różnych źródeł, dużo elementów ułatwiających podejmowanie decyzji zarządczych	łatwość użycia, mało możliwości analitycznych, konstrukcja modułowa
<i>Derwent Analytics (później TDA)</i>	<i>Thomson Reuters</i>	http://science.thomsonreuters.com/support/faq/analytics	narzędzie obsługujące bazy danych <i>Thomson Reuters</i> , wyszukiwanie, analiza i wizualizacja danych	wersja <i>VantagePoint</i> przystosowana do analiz patentowych, zmienny interfejs, mało uniwersalny, przystosowany do pracy z bazami <i>Thomson Reuters</i>
<i>Enterprise Miner/Text Miner</i>	<i>SAS Institute</i>	www.sas.com/text-analytics/text-miner/index.html	łączenie nieustrukturalizowanych danych tekstowych z danymi ilościowymi, analiz i wizualizacja wyników	możliwość analizy nieustrukturalizowanego tekstu, względnie prosty interfejs, dość złożona obsługa, systematycznie rozwijany od wielu lat
<i>Expertus</i>	<i>Splendor</i>	www.expertus.com.pl	analizy cytowań i dorobku naukowego, system wspomagania decyzji ewaluacyjnych dla jednostek naukowych	koncentracja na bibliografiach prac naukowych, mało uniwersalna, brak rozbudowanych funkcji graficznego prezentowania wyników i złożonych analiz

Nazwa	Producent	Strona www	Zastosowania	Charakterystyka
<i>Focust</i> (Wisdomain)	<i>Wisdomain</i>	www.wisdomain.com/overview.htm	analiza i wizualizacja danych pochodzących z baz patentowych, wyszukiwanie informacji wspomagających decyzje zarządcze	obsługuje dane częściowo strukturalizowane, ograniczone możliwości importu danych (współpraca z komercyjną bazą patentów)
<i>Goldfire Innovator</i>	<i>Invention Machine</i>	http://invention-machine.com	wspomaganie wynalazczości zgodnie z TRIZ	możliwość analizy nieustrukturalizowanego tekstu, zaawansowana analiza semantyczna tekstów naukowych i dokumentów patentowych w języku naturalnym, skomplikowana obsługa i parametryzacja
<i>Inxight Smart Discovery</i>	<i>Inxight Software</i>	www.inxight.com/products	wyszukiwanie i łączenie danych ogólnodostępnych, analiza tekstów w wielu językach narodowych (w tym polskim)	możliwość analizy nieustrukturalizowanego tekstu w języku naturalnym, wyszukiwanie równoległe w wielu bazach równocześnie, skomplikowana obsługa i parametryzacja
<i>Matheo Analyzer</i>	<i>IALE Tecnologia</i>	www.matheo-analyzer.com	opis w odrębnym podrozdziale	
<i>Omniviz</i>	<i>BioWisdom</i>	www.biowisdom.com/tag/omniviz	analizy publikacji i patentów dla wsparcia procesów badawczych	możliwość analizy nieustrukturalizowanego tekstu, dobra wizualizacja wyników, przeznaczone głównie dla sektora chemicznego i farmaceutycznego
<i>PatentInsight Pro</i>	<i>GridLogics</i>	www.patentinsightpro.com/product.html	analiza i wizualizacja danych patentowych oraz literatury naukowej na tematy techniczne	wyłącznie do analiz patentowych, dostępny z dodatkowymi narzędziami – do wizualizacji (VizMap) oraz CHEMPAT (analizy chemiczne)
<i>SciFinder Scholar</i>	<i>American Chemical Society</i>	www.cas.org/products/scifinder/index.html	analiza i wizualizacja danych z różnych źródeł w obszarze chemii	skoncentrowany na sektorze chemicznym, konstrukcja modułowa, powiązany z bazą CAS
<i>SPSS</i>	<i>IBM SPSS</i>	www.spss.com	uniwersalny pakiet do obróbki statycznej danych ilościowych i jakościowych	jeden z najbardziej rozbudowanych programów do analizy statystycznej, budowa modularna, wysoki koszt użytkowania, ograniczone możliwości eksportu danych do innych programów, eksploracja danych tekstowych to tylko jedna z wielu funkcji
<i>STN AnaVist</i>	<i>STN International</i>	http://www.stn-international.de/stn_home.html	analiza i wizualizacja danych patentowych i literatury naukowej	dużo opcji wizualizacyjnych, zarówno do literatury jak i patentów, działa dobrze tylko na ustrukturalizowanym tekście, lepiej przystosowane do analizy patentów, ograniczona liczba obsługiwanych baz danych, osobny program do przeglądania wyników analiz (<i>STN Viewer</i>)
<i>TEMIS Luxid</i>	<i>TEMIS</i>	www.temis.com	wyszukiwanie, analiza i	możliwość analizy nieustrukturalizowanego tekstu,

Nazwa	Producent	Strona www	Zastosowania	Charakterystyka
			wizualizacja danych tekstowych	dość proste możliwości wizualizacji wyników, kosztowny, budowa modułowa w zależności od odbiorcy (biznesowy/akademicki)
<i>UCINet</i>	<i>Analytic Technologies</i>	www.analytictech.com/ucinet/	analiza i wizualizacja danych sieciowych	możliwość przetwarzania uprzednio przygotowanych macierzy z danymi o cytowaniach lub innych współzależnościach, analizy i wizualizacji sieci zależności
<i>VantagePoint</i>	<i>Search Technologies</i>	www.thevantagepoint.com	opis w odrębnym podrozdziale	
<i>WizPatent Express Manager</i>	<i>WizSoft</i>	www.wizpatent.com/Default.aspx	wyszukiwanie i zapis na dysku danych patentowych	obsługa baz <i>USPTO</i> , <i>WIPO</i> i <i>EPO</i> , prosty interfejs, w droższej wersji programu – funkcje grupowania i opisu dokumentów
<i>WordStat</i>	<i>Provalis Research</i>	www.provalisresearch.com	analiza i wyszukiwanie danych ilościowych oraz jakościowych dla potrzeb naukowych i biznesowych	konstrukcja modułowa, kilka różnych programów współpracujących ze sobą: <i>WordStat</i> (eksploracja danych tekstowych), <i>QDAMiner</i> (analiza danych jakościowych) i <i>SimStat</i> (analiza statystyczna)

Tabela 56: Zestawienie najpopularniejszych płatnych programów, mogących wspierać analizy bibliometryczne. Źródło: opracowanie własne na podstawie stron internetowych producentów oprogramowania.

Nazwa	Producent	Strona www	Zastosowania	Charakterystyka
<i>Alibaba</i>	<i>Humboldt-Universität</i>	http://alibaba.informatik.hu-berlin.de/instructions.html	wyszukiwanie i wizualizacja danych medycznych (białka, choroby i leki), linkowanie bezpośrednio do tekstów źródłowych	obsługuje tylko wybrane bazy medyczne
<i>Arrowsmith</i>	<i>University of Illinois, Chicago</i>	http://arrowsmith.psych.uic.edu/arrowsmith_uic/index.html	wyszukiwanie powiązań pomiędzy artykułami z zakresu medycyny, analizuje odwołania w dwóch grupach artykułów równolegle	wyspecjalizowany w obsłudze baz medycznych, kilka współpracujących programów (<i>Author-ity</i> , <i>WETLAB</i> , <i>ADAM</i>), działa <i>online</i> , wspiera <i>Literature-Based Discovery</i>
<i>BibExcel</i>	<i>Umea University</i>	http://www8.umu.se/inforsk/Bibexcel	analiza danych bibliometrycznych, eksport do arkusza kalkulacyjnego w celu dalszej obróbki	program przetwarzający dane pobrane z bazy bibliograficznej na format programu <i>MS Excel</i>
<i>Bitola</i>	<i>University of Ljubljana</i>	http://ibmi.mf.uni-lj.si/bitola	wyszukiwanie powiązań pomiędzy różnymi badaniami i koncepcjami z zakresu medycyny	przeznaczony do analiz publikacji medycznych, obsługuje bazy <i>MEDLINE</i> , <i>Entrez Gene</i> , <i>OMIM</i> i <i>HUGO</i> , nacisk na badania nad genomem, działa <i>online</i> , wspiera <i>Literature-Based Discovery</i>
<i>CiteSpace</i>	<i>Chaomei Chen</i>	http://cluster.cis.drexel.edu/~cchen/citespace	opis w odrębnym podrozdziale	
<i>ExpertMapper</i>	<i>ExpertMapper</i>	www.expertmapper.com	tworzenie rankingu autorów i naukowców w zakresie medycyny	portal analizujący dane z bazy <i>MEDLINE</i>
<i>Free Patents Online</i>	<i>FreePatents Online.com</i>	www.freepatentsonline.com	baza patentów, oferująca możliwości analizy i wizualizacji danych	darmowa baza patentów, oparta na zasobach <i>USPTO</i> , pozwalająca na wizualizację danych patentowych na mapach <i>Google Maps</i> oraz dodatkowe analizy w oparciu o dane chemiczne; niestety zasoby patentowe nie wydają się kompletne i nie nadają się do poważnych zastosowań analitycznych
<i>HistCite</i>	<i>Thomson Reuters</i>	http://thomsonreuters.com/products_services/science_products/a-z/histcite/	opis w odrębnym podrozdziale	
<i>HTTrack Website Copier</i>	<i>Xavier Roche</i>	www.httrack.com	zapis na dysku plików <i>HTML</i> – wybranych elementów strony <i>www</i>	narzędzie przydatne przy ściąganiu na dysk lokalny dokumentów patentowych z bazy <i>USPTO</i>
<i>NetDraw</i>	<i>Analytic Technologies</i>	www.analytictech.com/netdraw/netdraw.htm	wizualizacja danych sieciowych	budowa graficznych sieci w oparciu o uprzednio przygotowane macierze z danymi o cytowaniach lub innych współzależnościach

Nazwa	Producent	Strona www	Zastosowania	Charakterystyka
<i>Pajek</i>	Vladimir Batagelj, Andrej Mrvar	http://vlado.fmf.uni-lj.si/pub/networks/pajek/	analiza i wizualizacja danych sieciowych	możliwość przetwarzania uprzednio przygotowanych macierzy z danymi o cytowaniach lub innych współzależnościach, analizy i wizualizacji sieci zależności
<i>PatentNet</i>	Chen Shih Sheng	http://sourceforge.net/projects/patentnet/	opis w odrębnym podrozdziale	
<i>Science Overlay Maps</i>	Georgia Institute of Technology / SPRU / University of Amsterdam	www.idr.gatech.edu	rzutowanie danych bibliometrycznych na mapę światowego dorobku badań	współpracuje z bazą <i>Web of Science</i> , pozwala na wizualizację interdyscyplinarnego charakteru badań prowadzonych przez daną instytucję lub w danym obszarze
<i>The Stanford Parser</i>	Stanford University	http://nlp.stanford.edu/software/lex-parser.shtml	statystyczna analiza tekstów	pozwała na przetwarzanie języka naturalnego (np. patentów) w celu wyodrębnienia z tekstów określonych struktur leksykalnych
<i>VOSviewer</i>	Leiden University	www.vosviewer.com	wizualizacja sieci powiązań na podstawie danych bibliometrycznych	współpracuje z bazą <i>Web of Science</i> , pozwala na analizy cytowań, współautorstwa i <i>co-words</i>
<i>WEKA</i>	University of Waikato	www.cs.waikato.ac.nz/ml/weka/	klasyfikacja, łączenie i wizualizacja danych tekstowych	napisany w języku <i>Java</i> , pozwala na przetwarzanie danych tekstowych, niezbędne podstawowe umiejętności programowania
	University of Amsterdam	http://www.leydesdorff.net/software.htm	zestaw narzędzi, wspierających analizy bibliometryczne	programy, pozwalające na: analizę <i>co-words</i> , przetwarzanie zbioru rekordów z baz <i>Web of Science</i> , <i>Scopus</i> i <i>Google Scholar</i> na format relacyjnej bazy danych (np. <i>Microsoft Access</i>), wizualną analizę cytowań i współautorstwa

Tabela 57: Zestawienie najpopularniejszych bezpłatnych programów, mogących wspierać analizy bibliometryczne. Źródło: opracowanie własne na podstawie stron internetowych producentów oprogramowania.

12. Rekomendacje dotyczące prowadzenia analiz bibliometrycznych

12.1. Wprowadzenie

W niniejszym rozdziale zaprezentowana zostanie propozycja **zakresów zadań**, które mogą być przedmiotem projektów analiz bibliometrycznych oraz rekomendacje dotyczące konkretnych **sposobów organizacji prac analitycznych**.

We wcześniejszych rozdziałach książki omówione zostały:

- teoretyczne uwarunkowania analiz bibliometrycznych,
- przykładowe doświadczenia zagranicznych instytucji publicznych i prywatnych,
- możliwe pytania badawcze, służące ukierunkowaniu prac analitycznych,
- specyficzne techniki analiz bibliometrycznych,
- najważniejsze bazy publikacji i patentów, które mogą dostarczać danych do analiz,
- aplikacje komputerowe, wspomagające analizy.

Po szczegółowej prezentacji możliwości prowadzenia analiz bibliometrycznych, źródeł danych oraz narzędzi, które mogą te analizy wspierać, nadszedł czas na sformułowanie propozycji dla przyszłych badaczy.

W pierwszej kolejności, należy dokonać **wyboru konkretnych pytań badawczych**, które mogą interesować decydentów z organizacji sektora B+R. Techniki bibliometryczne mogą pomóc w podejmowaniu decyzji strategicznych jako podstawa dla analizy SWOT - określania sił i słabości, szans i zagrożeń, stojących przed badaczami. To pozwoli z kolei podejmować decyzje o wyborze priorytetowych kierunków badawczych, alokacji środków na projekty badawczo-rozwojowe, jak również identyfikacji potencjalnych beneficjentów-realizatorów projektów i uczestników konsorcjów badawczych. Proponowane działania mogą wspierać wdrażanie koncepcji inteligentnej specjalizacji w polityce naukowej i innowacyjnej (Kardas 2010).

W oparciu o wcześniejszą dyskusję możliwych zastosowań i ograniczeń prezentowanych technik bibliometrycznych należy podkreślić, że omawiane w niniejszej pracy analizy **nie nadają się bezpośrednio do prowadzenia ewaluacji** indywidualnych badaczy oraz jednostek naukowych. Książka koncentruje się na zastosowaniach bibliometrii do obserwacji badań naukowych i rozwoju technologii, identyfikacji długookresowych trendów i krótkookresowych zjawisk oraz rozumieniu specyfiki poszczególnych obszarów badań, także poprzez określenie istotnych osób i instytucji, prowadzących badania w określonych obszarach. Istnieje również drugie, istotnie odmienne zastosowanie bibliometrii jako podstawy do ewaluacji - jednak opiera się na odrębnych założeniach koncepcyjnych, a jego poprawne metodologicznie zastosowanie wymaga stosowania innych narzędzi niż w przypadku podejścia eksploracyjnego, stanowiącego przedmiot niniejszej publikacji.

12.2. Wybór pytań badawczych

Tabela 58 zawiera zestawienie pytań badawczych, na które można udzielić odpowiedzi przy wykorzystaniu bibliometrii. Jej zawartość powiela elementy prezentowanej wcześniej Tabela 9, przypominając o potencjale analiz. Z perspektywy sprawnego prowadzenia analiz bibliometrycznych, warto szczególną uwagę zwrócić na zróżnicowane wymogi dotyczące znajomości analizowanej dyscypliny oraz stopnia trudności przetwarzania danych w przypadku prowadzenia analiz bibliometrycznych.

Odpowiedzi na niektóre z pytań badawczych mogą być udzielone bez dogłębnej znajomości analizowanego obszaru badań, co może okazać się bardzo ważne i korzystne dla organizacji, prowadzącej analizy bibliometryczne (symbol "+", Tabela 58). Takie analizy nie będą wymagały poszukiwań zewnętrznych ekspertów dziedzinowych i mogą być prowadzone przez osoby, które poznały podstawy technik bibliometrycznych oraz dysponują odpowiednimi narzędziami. Bardziej złożone analizy mogą wymagać doraźnych konsultacji z ekspertami w procesach identyfikacji podzbioru danych bibliometrycznych (definiowania odpowiednich wyszukiwań, wskazywania słów kluczowych i kodów klasyfikacyjnych) oraz analizy danych (np. w związku z niezbędnymi wyjaśnieniami, dotyczącymi różnic pomiędzy podobszarami badawczymi i szczegółowymi tematami poddanych analizom badań). Omawiane przypadki Tabela 58 oznacza symbolem "++" i można założyć, że takie analizy są bardziej czasochłonne ze względu na konieczność konsultacji z ekspertami, choć niekoniecznie wymagać będą dodatkowych nakładów finansowych, gdyż w większości przypadków niezbędne konsultacje z osobami zewnętrznymi mogą ograniczać się do relatywnie krótkich rozmów. Nieliczne pytania badawcze wymagają wreszcie bliskiej, intensywnej współpracy z ekspertami i dogłębnej znajomości analizowanego obszaru nauki i techniki - te przypadki oznaczono symbolem "+++", a związane z nimi pytania badawcze wydają się bardziej interesujące dla podmiotów prowadzących działalność badawczo-rozwojową, zainteresowanych dobrym zrozumieniem otoczenia rynkowego, niż dla reprezentantów instytucji rządowych.

Analogicznie można zinterpretować znaczenie kryterium stopnia trudności przetwarzania danych. Na wiele pytań badawczych można udzielić odpowiedzi w oparciu o kwerendy w bazach bibliometrycznych lub podstawową obróbkę danych bibliometrycznych w standardowym oprogramowaniu (scenariusze odpowiednio oznaczone jako "*" i "***"). Takie analizy mogą być przy niewielkich nakładach pracy przeprowadzone przez pracowników resortu nauki lub podmioty zewnętrzne, pod warunkiem posiadania dostępu do baz i odpowiedniego oprogramowania. Trudniejsze w realizacji wydają się warianty oznaczone symbolami "****" i "*****" - wymagają one eksportu przetworzonych danych bibliometrycznych z odpowiedniego narzędzia informatycznego do innych programów, pozwalających na dalsze analizy oraz wykorzystania zaawansowanych algorytmów analiz statystycznych lub lingwistycznych. Wydaje się, że stopień złożoności tych analiz, trudność opanowania technik oraz ograniczenia narzędzi informatycznych sprawiają, że te zaawansowane warianty nie przydadzą się bezpośrednio w typowej pracy specjalistów organizacji sektora B+R, choć mogą stanowić podstawę bardzo interesujących projektów badawczych. Gdyby pojawiło się ściśle określone zapotrzebowanie na omawiane analizy,

wstępnie przetworzony zbiór danych bibliometrycznych może być również przekazany odpowiedniemu wykonawcy zewnętrznemu, wyspecjalizowanego w określonym rodzaju analiz - przykładowo, budowa modeli prognostycznych lub analiza czynnikowa mogą być prowadzone przez analityków, którzy nie są specjalistami w zakresie bibliometrii, posiadają jednak odpowiednią wiedzę z zakresu technik ilościowych oraz niezbędne narzędzia informatyczne.

Pytanie analityczne	Wyjaśnienie	Znajomość analizowanej dyscypliny	Stopień trudności przetwarzania danych
Co?	Identyfikacja obszarów badań i technologii w zbiorze danych bibliometrycznych dla określonego kraju, regionu, ośrodka badawczego lub dla określonych lat	+	*
Co dokładnie?	Identyfikacja podobszarów badawczych w zbiorze danych bibliometrycznych, odzwierciedlających wybrany obszar badań i technologii	++	**
Kto?	Identyfikacja badaczy/wynalazców i zespołów badawczych, prowadzących badania na określony temat	+	*
Gdzie?	Identyfikacja afiliacji, w których prowadzone są badania na określony temat (poziom kraju, miasta, organizacji)	+	*
Kiedy?	Określenie momentu publikacji, zgłoszenia patentowego lub uzyskania patentu, jak również okresów, w których prowadzono badania na określony temat	+	*
Z kim?	Analiza współautorstwa publikacji i wynalazków (<i>co-authorship</i>) poprzez identyfikację osób, które wspólnie prowadziły badania	+	**
Z jaką organizacją?	Analiza współautorstwa publikacji i wynalazków (<i>co-authorship</i>) poprzez identyfikację organizacji, zatrudniających osoby, które wspólnie prowadziły badania	+	**
Ile?	Zliczanie publikacji, zgłoszeń patentowych lub udzielonych patentów dla wybranego obszaru lub podobszaru badań i technologii (możliwe jest także sporządzanie zestawień porównawczych liczby publikacji w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją?)	+	*
Ilu?	Zestawienie liczby autorów publikacji lub zgłoszeń patentowych w danym obszarze	+	**
Jak zmienia się w czasie?	Analiza szeregów czasowych, odzwierciedlających liczby publikacji lub patentów w wybranych obszarach badań i technologii (możliwe jest odzwierciedlenie zmian w czasie w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? ile? – czyli	+	*

Pytanie analityczne	Wyjaśnienie	Znajomość analizowanej dyscypliny	Stopień trudności przetwarzania danych
	prezentacja historycznego rozwoju obszaru badań)		
Jak będzie się rozwijać?	Prognoza krótkookresowa, oparta na ekstrapolacji szeregów czasowych (możliwe jest przygotowanie prognoz w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? pod warunkiem posiadania dostatecznie długich serii danych historycznych)	+	***
Jak popularne w porównaniu do?	Porównanie liczb publikacji, zgłoszeń patentowych lub patentów w kilku obszarach badawczych	+	*
Jakie są nisze?	Identyfikacja podobszarów badawczych, które nadal cieszą się niewielkim zainteresowaniem badaczy (w oparciu o liczbę publikacji, zgłoszeń patentowych lub patentów w danym podobszarze, porównaną z ich globalną liczbą dla szerokiego obszaru badań)	++	*
Jakie są nowe popularne tematy?	Identyfikacja nowych, cieszących się narastającą popularnością tematów badawczych (<i>hot topics</i>), w tym dynamicznie rozwijających się, nowych technologii (<i>emerging technologies</i>)	++	**
Z czym związane?	Analiza pokrewnych obszarów badań i tematów badawczych, które równocześnie opisywane są w publikacjach lub dokumentach patentowych – na poziomie współwystępowania specjalistycznych pojęć w streszczeniach i tytułach (<i>co-words</i>) lub w słowach kluczowych i kodach klasyfikacyjnych (<i>co-classification</i>)	++	**
W jaki sposób związane?	Analiza szczegółowych podejść badawczych, technik, środków i technologii, wykorzystywanych przez zidentyfikowane wcześniej obszary badań i tematy badawcze, pokrewne wobec podstawowego obszaru analiz	++	**
Co uzupełnia?	Identyfikacja komplementarnych obszarów wiedzy naukowej, komplementarnych umiejętności badawczych lub komplementarnych technologii, wykorzystywanych przy prowadzeniu badań w danym obszarze	++	**
Jakich wariantów nie analizowano?	Identyfikacja możliwych nowych kierunków badań i sposobów rozwiązania problemów naukowych i praktycznych w oparciu o analogie, techniki kreatywnego rozwiązywania problemów i ocenę potencjału ewolucyjnego	+++	****
Kto uzupełnia?	Identyfikacja osób i organizacji, posiadających komplementarne obszary wiedzy naukowej, komplementarne umiejętności badawcze lub komplementarne technologie, wykorzystywane przy prowadzeniu badań w danym obszarze	++	**

Pytanie analityczne	Wyjaśnienie	Znajomość analizowanej dyscypliny	Stopień trudności przetwarzania danych
Kto może być dostawcą?	Identyfikacja osób i organizacji, które mogą być podwykonawcami określonych zleceń badawczych ze względu na dotychczasowy dorobek naukowy	++	**
Kto może być odbiorcą?	Identyfikacja osób i organizacji, które mogą być zleceniodawcami projektów badawczo-rozwojowych lub nabywcami/licencjobiorcami ich rezultatów ze względu na dotychczasowe zainteresowania i potrzeby badawcze	++	**
Kto jest konkurentem?	Identyfikacja osób i organizacji, prowadzących badania w analogicznych, szczegółowych obszarach badawczych, przy wykorzystaniu zbliżonych metod (podmioty zidentyfikowane jako konkurenci mogą być też potencjalnymi partnerami w konsorcjach badawczych – zwłaszcza wtedy, gdy niezbędne jest połączenie zbliżonych zasobów dla zwiększenia skali prac badawczych)	++	**
Kto może być konkurentem?	Identyfikacja osób i organizacji, które mogą podjąć badania w analogicznych, szczegółowych obszarach badawczych, przy wykorzystaniu zbliżonych metod, a wskazuje na to ich profil badawczy, dotychczasowe zainteresowania i analizy statystyczne (zwłaszcza techniki analizy czynnikowej, pozwalające na identyfikację „ukrytych” podobieństw i zależności w dużych zbiorach danych)	+++	***
Kto może być pracownikiem?	Identyfikacja osób, które prowadzą lub mogą prowadzić badania w określonym obszarze i zgromadziły zadowalający dorobek naukowy	++	**
Kto może być sukcesorem?	Identyfikacja osób, które prowadziły badania wspólnie z określonym naukowcem (<i>co-authorship</i>) i mogą w przyszłości przejąć obowiązki w zakresie zarządzania projektami badawczymi	++	**
Jakie X ma strategie badawcze?	Analiza tematyki badawczej, mechanizmów współpracy autorskiej (<i>co-authorship</i>), kluczowych badaczy i zespołów badawczych, partnerów zewnętrznych i zmian w powyższych obszarach dla wybranej organizacji	++	**
Jak ważne?	Zestawienie liczby cytowań dla pojedynczych publikacji lub patentów	+	*
Jak łącznie ważne?	Zestawienie zagregowanej liczby cytowań dla grupy publikacji lub patentów (np. pochodzących od autorów z jednej organizacji, dotyczących jednego obszaru badawczego)	+	**
Kto jest ważny?	Porównania liczby cytowań w dorobku danych badaczy oraz obliczanie wskaźników dotyczących cytowań (np. indeksów h Hirscha)	+	*

Pytanie analityczne	Wyjaśnienie	Znajomość analizowanej dyscypliny	Stopień trudności przetwarzania danych
Do jakiej wiedzy nawiązuje?	Identyfikacja źródeł wiedzy, ujawnionych w cytowaniach, zawartych w publikacjach naukowych i zgłoszeniach patentowych (możliwe dalsze analizy zidentyfikowanego zbioru źródeł wiedzy w oparciu o pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? ile?)	++	**
Do czyich badań nawiązuje?	Identyfikacja naukowców i wynalazców, do prac których odwołują się cytowania w publikacjach naukowych i zgłoszeniach patentowych	++	**

Tabela 58. Zestawienie pytań analitycznych w projektach bibliometrycznych. Źródło: opracowanie własne (por. Tabela 9). Legenda:

Symbole użyte w kolumnie „Znajomość analizowanej dyscypliny”:

- + - odpowiedzi na pytania wymagają jedynie ogólnej orientacji w specyficie i terminologii analizowanej dyscypliny, mogą być wykonywane przez niespecjalistów pod warunkiem wstępnych konsultacji z ekspertami dziedzinowymi i/lub wstępnych studiów literaturowych,
- ++ - odpowiedzi na pytania wymagają wsparcia ze strony ekspertów dziedzinowych na etapie analizy i interpretacji danych, często także bezpośredniego zaangażowania ekspertów w prace analityczne,
- +++ - odpowiedzi na pytania mogą być udzielone tylko przez ekspertów dziedzinowych, dostatecznie dobrze rozumiejących specyfikę poddanego analizom obszaru badań naukowych i technologicznych.

Symbole użyte w kolumnie „Stopień trudności przetwarzania danych”:

- * - odpowiedzi na pytania nie wymagają posiadania specjalistycznego oprogramowania, mogą być udzielone na podstawie kwerend (poleceń wyszukiwania) w bibliometrycznych bazach danych,
- ** - odpowiedzi na pytania wymagają dodatkowego przetwarzania podzbiorów danych, pobranych z bibliometrycznych baz danych, przy wykorzystaniu specjalistycznego oprogramowania,
- *** - odpowiedzi na pytania wymagają statystycznej obróbki podzbiorów danych, pobranych z bibliometrycznych baz danych, przy wykorzystaniu specjalistycznego oprogramowania w oparciu o wybór odpowiednich algorytmów,
- **** - odpowiedzi na pytania wymagają jakościowej analizy podzbiorów danych, pobranych z bibliometrycznych baz danych w oparciu o wiedzę ekspercką z analizowanej dyscypliny oraz sprawność w wykorzystywaniu zaawansowanych technik analitycznych odwołujących się do *LBD*, *KBMA* i/lub *TRIZ*.

Tabela 59 przedstawia zestawienie pytań badawczych w oparciu o dwa wyodrębnione wcześniej kryteria. Ujawnia ona zaskakująco szeroki zakres analiz, które mogą być prowadzone przez osoby, nie posiadające dogłębnej wiedzy dziedzinowej ani zaawansowanych narzędzi technicznych, a jedynie przeszkolonych w zakresie podstaw analiz bibliometrycznych. Warto podkreślić, że omawiane analizy pozostają w zasięgu możliwości realizacyjnych pracowników wielu organizacji z sektora B+R. W wielu przypadkach niezbędne okazuje się prowadzenie wielu analiz w formie szybkiej reakcji na trudne do przewidzenia zdarzenia oraz doraźne zapotrzebowanie na odpowiednie dane jako podstawy decyzji. Wydaje się, że samodzielna gotowość do realizacji przynajmniej części spośród opisanych analiz przez pracowników instytucji naukowych i przedsiębiorstw będzie wskazana, pozwalając na oszczędność kosztów, zwiększoną elastyczność i łatwość działania w oparciu o wiedzę, jak również budowę i utrzymanie niezbędnych kompetencji merytorycznych.

		Wymóg znajomość analizowanej dyscypliny		
		+	++	+++
Stopień trudności przetwarzania danych	*	<ul style="list-style-type: none"> • Co? • Kto? • Gdzie? • Kiedy? • Ile? • Jak zmienia się w czasie? • Jak popularne w porównaniu do? • Jak ważne? • Kto jest ważny? 	<ul style="list-style-type: none"> • Jakże są nisze? 	-
	**	<ul style="list-style-type: none"> • Z kim? • Z jaką organizacją? • Ilu? • Jak łącznie ważne? 	<ul style="list-style-type: none"> • Co dokładnie? • Jakie są nowe popularne tematy? • Z czym związane? • W jaki sposób związane? • Co uzupełnia? • Kto uzupełnia? • Kto może być dostawcą? • Kto może być odbiorcą? • Kto jest konkurentem? • Kto może być pracownikiem? • Kto może być sukcesorem? • Jak X ma strategie badawcze? • Do jakiej wiedzy nawiązuje? • Do czyich badań nawiązuje? 	-
	***	<ul style="list-style-type: none"> • Jak będzie się rozwijać? 	-	<ul style="list-style-type: none"> • Kto może być konkurentem?
	****	-	-	<ul style="list-style-type: none"> • Jakich wariantów nie analizowano?

Tabela 59: Zestawienie pytań badawczych w projektach bibliometrycznych w podziale na stopień trudności przetwarzania danych i wymóg znajomości analizowanej dyscypliny. Źródło: opracowanie własne.

12.3. Przydatność technik analitycznych

Kolejnym krokiem przy definiowaniu zakresu przyszłych analiz bibliometrycznych jest wybór przydatnych technik analitycznych. Poszczególne techniki zostały zaprezentowane we wcześniejszych rozdziałach książki. Warto ocenić ich przydatność do udzielania odpowiedzi na konkretne pytania badawcze, co prezentuje Tabela 60.

W oparciu o analizę tego zestawienia danych można sformułować propozycję wykorzystania techniki analitycznej *tech mining* jako podstawy do prowadzenia szczególnie wielu analiz.

Prezentowany wcześniej przegląd literatury wskazywał na analogiczne doświadczenia instytucji rządowych i przedsiębiorstw z innych krajów oraz popularność *tech mining*-u jako podstawy do decyzji w obszarze zarządzania działalnością badawczo-rozwojową oraz tworzenia polityki innowacyjnej.

W podejściu *tech mining* zawierają się już elementy mniej zaawansowanych technik analitycznych - analiz liczebności, *co-words*, *co-classification* i *co-authorship*. Jest ono zbliżone do technik tomografii bazodanowej, a programy takie jak *VantagePoint* pozwalają na prowadzenie również takich analiz, jednak współcześnie wydaje się, że ograniczanie się do analizy złożonych przypadków współwystępowania słów w tytułach i streszczeniach publikacji naukowych i patentów nie jest wystarczające do kompleksowej obserwacji badań naukowych i rozwoju technologii. Analogicznie, niektóre możliwości analizy cytowań są już udostępniane w ramach analiz *tech mining*, a pozostałe korzyści związane z cytowaniami wydają się bardziej adekwatne do zastosowań ewaluacyjnych, którymi nie zajmuje się niniejsza publikacja.

Obiecujący potencjał badawczy posiadają techniki, oparte na *Literature-Based Discovery*, *TRIZ* i *keyword-based morphological analysis*. Należy jednak zauważyć, że ich podstawowym zastosowaniem jest stymulowanie kreatywności badaczy i wynalazców w celu odkrywania nowych możliwości technicznych. Techniki te pozwalają przede wszystkim udzielić odpowiedzi na pytanie "Jakich wariantów nie analizowano?" czyli służyć mogą przede wszystkim zespołom badawczym i przedsiębiorcom, poszukującym praktycznych rozwiązań naukowych i technologicznych, a nie decydom w instytucjach administracji publicznej. Zastosowanie *TRIZ* pozwala dodatkowo ocenić potencjał ewolucyjny technologii oraz przygotować prognozy dalszego rozwoju technologii, przydatne przykładowo w projektach typu *foresight*, jednak nie jest to klasyczny scenariusz analiz bibliometrycznych, oparty na przetwarzaniu danych ilościowych, a jakościowa analiza zawartości dokumentów patentowych. Dodatkowo, jak wskazano w poprzednim podrozdziale, odpowiedź na pytanie "Jakich wariantów nie analizowano?" wymaga wiedzy eksperckiej w zakresie analizowanej dyscypliny oraz wykorzystania kosztownych narzędzi do przetwarzania danych, co ogranicza możliwość realizacji tego typu analiz w ramach typowych projektów bibliometrycznych.

Interesujące mogą z kolei okazać się prace analityczne, dotyczące analizy nowych, popularnych zjawisk i tendencji w nauce i technologii (ang. *hot topics*) oraz mapowanie sieci zależności w oparciu o dane bibliometryczne. Te zadania mogą być realizowane przy wykorzystaniu tych samych narzędzi informatycznych, m.in. opisanego wcześniej, bezpłatnego programu *CiteSpace*, dedykowanego do wizualizacji sieci oraz identyfikacji nowych tendencji (ang. *burst effect*). Innym przydatnym narzędziem jest również opisany wcześniej program, wspólnie stworzony przez badaczy z Wielkiej Brytanii, Holandii i Stanów Zjednoczonych i bezpłatnie udostępniony na stronie internetowej, pozwalający na "rzutowanie" na światową mapę wiedzy (zależności pomiędzy specjalistycznymi tematami badań) własnego zestawu danych, przykładowo danych o publikacjach autorstwa pracowników wybranej uczelni lub wszystkich badaczy w danym kraju, jak również zbioru światowych publikacji w danym interdyscyplinarnym obszarze badań. Oba narzędzia pozwolą uzupełnić analizy typu *tech mining* w wielu projektach analitycznych.

Pytanie analityczne	Analiza liczebności	Analiza <i>co-words</i>	Analiza <i>co-classification</i>	Analiza <i>co-authorship</i>	Wskaźniki cytowań	Analiza <i>co-citations</i>	Analiza <i>tech mining</i>	Tomografia bazodanowa	LBD / KMBA / TRIZ	Analiza <i>hot topics</i>	Mapowanie sieci
Co?	X	X	X				X	X			
Co dokładnie?	X	X	X				X	X			
Kto?	X			X			X				
Gdzie?	X			X			X				
Kiedy?	X						X				
Z kim?				X			X				X
Z jaką organizacją?				X			X				X
Ile?	X						X	X			
Ilu?	X			X			X				
Jak zmienia się w czasie?	X						X	X			
Jak będzie się rozwijać?	X						X				
Jak popularne w porównaniu do?	X						X	X			X
Jakie są nisze?	X	X	X				X	X			X
Jakie są nowe popularne tematy?	X	X	X				X	X		X	
Z czym związane?		X	X				X	X			X
W jaki sposób związane?		X	X				X	X			X
Co uzupełnia?		X	X				X	X			X
Jakich									X		

Pytanie analityczne	Analiza liczebności	Analiza <i>co-words</i>	Analiza <i>co-classification</i>	Analiza <i>co-authorship</i>	Wskaźniki cytowań	Analiza <i>co-citations</i>	Analiza <i>tech mining</i>	Tomografia bazodanowa	<i>LBD / KMBA / TRIZ</i>	Analiza <i>hot topics</i>	Mapowanie sieci
wariantó w nie analizowano?											
Kto uzupełnia?							X				X
Kto może być dostawcą?	X						X				
Kto może być odbiorcą?	X						X				
Kto jest konkurentem?	X						X				
Kto może być konkurentem?							X				
Kto może być pracownikiem?	X						X				
Kto może być sukcesorem?				X			X				
Jakie X ma strategię badawcze?							X				
Jak ważne?					X						
Jak łącznie ważne?					X	X					
Kto jest ważny?					X	X					
Do jakiej wiedzy nawią--						X					X

Pytanie analityczne	Analiza liczebności	Analiza <i>co-words</i>	Analiza <i>co-classification</i>	Analiza <i>co-authorship</i>	Wskaźniki cytowań	Analiza <i>co-citations</i>	Analiza <i>tech mining</i>	Tomografia bazodanowa	<i>LBD / KMBA / TRIZ</i>	Analiza <i>hot topics</i>	Mapowanie sieci
zuje?											
Do których badań nawiązuje?						X					X

Tabela 60: Przydatność poszczególnych technik analiz bibliometrycznych w udzielaniu odpowiedzi na pytania analityczne. Źródło: opracowanie własne.

12.4. Przydatność baz bibliograficznych i oprogramowania

We wcześniejszym rozdziale książki szczegółowo opisane zostały cztery najpopularniejsze bazy streszczeń publikacji i cytowań, które mogą być wykorzystywane w projektach badawczych: *Web of Science*, *Scopus*, *INSPEC* i *MEDLINE*. Nie jest możliwe bezpośrednio porównanie tych czterech baz i wskazanie najlepszej, najbardziej dostosowanej do potrzeb analiz - w zależności od specyfiki badań, przydatna okazać się może inna baza. Przykładowo, analiza cytowań jest możliwa przede wszystkim w oparciu o bazy *Web of Science* i *Scopus*. Szczegółowy indeks tematyczny, pozwalający na łatwą identyfikację publikacji dotyczących specjalistycznych tematów w naukach ścisłych i technicznych zawiera *INSPEC*, podczas gdy baza *MEDLINE* indeksuje w analogiczny sposób artykuły medyczne.

Niektóre bazy posiadają też istotne ograniczenia - np. *Web of Science* nie klasyfikuje indywidualnych artykułów, przypisując całe pisma do poszczególnych obszarów badań (co uniemożliwia analizę interdyscyplinarnych obszarów badawczych), *INSPEC* przez wiele lat podawał tylko afiliację pierwszego autora publikacji (eliminując możliwość analizy zależności pomiędzy instytucjami B+R).

Dobór bazy powinien być więc uzależniony od:

- analizowanego obszaru badań,
- pytania badawczego.

Bazy *Web of Science* i *Scopus* są porównywalne w przypadku wielu zastosowań. *Web of Science* jest produktem, którego wcześniejsze wersje były dostępne od kilkadziesiąt lat, a wiele specjalistycznych narzędzi informatycznych bezpośrednio odwołuje się do danych z tej bazy. Nowsza baza *Scopus* wydaje się jednak bardziej interesująca dla celów analiz bibliometrycznych - obejmuje więcej czasopism i innych publikacji, wzbogaca opis wielu publikacji o słowa kluczowe oraz udostępnia dodatkowe narzędzia wspierające wyszukiwanie. Jej ograniczeniem jest gromadzenie informacji na temat cytowań publikacji od roku 1996, choć nie wydaje się to istotnym problemem w przypadku projektów analitycznych w obszarach zarządzania technologiami i polityki naukowej, polegających zwykle na obserwacji relatywnie niedawnych tendencji badawczych. Warto też pamiętać, że niektóre

aplikacje komputerowe, wykorzystywane do analiz bibliometrycznych, współpracują jedynie z bazą *Web of Science* - przykładowo, jest ona niezbędna do tworzenia map wiedzy przy wykorzystaniu bezpłatnego pakietu, udostępnionego przez Rafolsa, Leydesdorffa i Portera. Chociaż do wielu zastosowań warto zarekomendować bazę *Scopus*, konkurencyjna baza *Web of Science* nadal pozostaje potrzebna.

Pytanie analityczne	Wyjaśnienie	<i>Web of Science</i>	<i>Scopus</i>	<i>INSPEC</i>	<i>MEDLINE</i>
Co?	Identyfikacja obszarów badań i technologii w zbiorze danych bibliometrycznych dla określonego kraju, regionu, ośrodka badawczego lub dla określonych lat	++	++	++	++
Co dokładnie?	Identyfikacja podobszarów badawczych w zbiorze danych bibliometrycznych, odzwierciedlających wybrany obszar badań i technologii	+	+	++	++
Kto?	Identyfikacja badaczy/wynalazców i zespołów badawczych, prowadzących badania na określony temat	++	++	++	++
Gdzie?	Identyfikacja afiliacji, w których prowadzone są badania na określony temat (poziom kraju, miasta, organizacji)	++	++	+	+
Kiedy?	Określenie momentu publikacji, zgłoszenia patentowego lub uzyskania patentu, jak również okresów, w których prowadzono badania na określony temat	++	++	++	++
Z kim?	Analiza współautorstwa publikacji i wynalazków (<i>co-authorship</i>) poprzez identyfikację osób, które wspólnie prowadziły badania	++	++	++	++
Z jaką organizacją?	Analiza współautorstwa publikacji i wynalazków (<i>co-authorship</i>) poprzez identyfikację organizacji, zatrudniających osoby, które wspólnie prowadziły badania	++	++	+	+
Ile?	Zliczanie publikacji, zgłoszeń patentowych lub	++	++	++	++

Pytanie analityczne	Wyjaśnienie	Web of Science	Scopus	INSPEC	MEDLINE
	udzielonych patentów dla wybranego obszaru lub podobzaru badań i technologii (możliwe jest także sporządzanie zestawień porównawczych liczby publikacji w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją?)				
Ilu?	Zestawienie liczby autorów publikacji lub zgłoszeń patentowych w danym obszarze	++	++	++	++
Jak zmienia się w czasie?	Analiza szeregów czasowych, odzwierciedlających liczby publikacji lub patentów w wybranych obszarach badań i technologii (możliwe jest odzwierciedlenie zmian w czasie w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? ile? – czyli prezentacja historycznego rozwoju obszaru badań)	++	++	++	++
Jak będzie się rozwijać?	Prognoza krótkookresowa, oparta na ekstrapolacji szeregów czasowych (możliwe jest przygotowanie prognoz w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? pod warunkiem posiadania dostatecznie długich serii danych historycznych)	++	++	++	++
Jak popularne w porównaniu do?	Porównanie liczb publikacji, zgłoszeń patentowych lub patentów w kilku obszarach badawczych	+	++	++	++
Jakie są nisze?	Identyfikacja podobzarów badawczych, które nadal cieszą się niewielkim zainteresowaniem badaczy (w oparciu o liczbę publikacji, zgłoszeń	+	++	++	++

Pytanie analityczne	Wyjaśnienie	Web of Science	Scopus	INSPEC	MEDLINE
	patentowych lub patentów w danym podobszarze, porównaną z ich globalną liczbą dla szerokiego obszaru badań)				
Jakie są nowe popularne tematy?	Identyfikacja nowych, cieszących się narastającą popularnością tematów badawczych (<i>hot topics</i>), w tym dynamicznie rozwijających się, nowych technologii (<i>emerging technologies</i>)	+	++	++	++
Z czym związane?	Analiza pokrewnych obszarów badań i tematów badawczych, które równocześnie opisywane są w publikacjach lub dokumentach patentowych – na poziomie współwystępowania specjalistycznych pojęć w streszczeniach i tytułach (<i>co-words</i>) lub w słowach kluczowych i kodach klasyfikacyjnych (<i>co-classification</i>)	++	++	++	++
W jaki sposób związane?	Analiza szczegółowych podejść badawczych, technik, środków i technologii, wykorzystywanych przez zidentyfikowane wcześniej obszary badań i tematy badawcze, pokrewne wobec podstawowego obszaru analiz	+	++	++	++
Co uzupełnia?	Identyfikacja komplementarnych obszarów wiedzy naukowej, komplementarnych umiejętności badawczych lub komplementarnych technologii, wykorzystywanych przy prowadzeniu badań w danych obszarze	+	++	++	++
Jakich wariantów nie analizowano?	Identyfikacja możliwych nowych kierunków badań i sposobów rozwiązania problemów naukowych i praktycznych w oparciu o	+	+	+	+

Pytanie analityczne	Wyjaśnienie	Web of Science	Scopus	INSPEC	MEDLINE
	analogie, techniki kreatywnego rozwiązywania problemów i ocenę potencjału ewolucyjnego				
Kto uzupełnia?	Identyfikacja osób i organizacji, posiadających komplementarne obszary wiedzy naukowej, komplementarne umiejętności badawcze lub komplementarne technologie, wykorzystywane przy prowadzeniu badań w danym obszarze	+	++	++	++
Kto może być dostawcą?	Identyfikacja osób i organizacji, które mogą być podwykonawcami określonych zleceń badawczych ze względu na dotychczasowy dorobek naukowy	++	++	++	++
Kto może być odbiorcą?	Identyfikacja osób i organizacji, które mogą być zleceniodawcami projektów badawczo-rozwojowych lub nabywcami/licencjobiorcami ich rezultatów ze względu na dotychczasowe zainteresowania i potrzeby badawcze	++	++	++	++
Kto jest konkurentem?	Identyfikacja osób i organizacji, prowadzących badania w analogicznych, szczegółowych obszarach badawczych, przy wykorzystaniu zbliżonych metod (podmioty zidentyfikowane jako konkurenci mogą być też potencjalnymi partnerami w konsorcjach badawczych – zwłaszcza wtedy, gdy niezbędne jest połączenie zbliżonych zasobów dla zwiększenia skali prac badawczych)	++	++	++	++
Kto może być konkurentem?	Identyfikacja osób i organizacji, które mogą podjąć badania w analogicznych, szczegółowych obszarach badawczych, przy	+	+	++	++

Pytanie analityczne	Wyjaśnienie	Web of Science	Scopus	INSPEC	MEDLINE
	wykorzystaniu zbliżonych metod, a wskazuje na to ich profil badawczy, dotychczasowe zainteresowania i analizy statystyczne (zwłaszcza techniki analizy czynnikowej, pozwalające na identyfikację „ukrytych” podobieństw i zależności w dużych zbiorach danych)				
Kto może być pracownikiem?	Identyfikacja osób, które prowadzą lub mogą prowadzić badania w określonym obszarze i zgromadziły zadowalający dorobek naukowy	++	++	++	++
Kto może być sukcesorem?	Identyfikacja osób, które prowadziły badania wspólnie z określonym naukowcem (<i>co-authorship</i>) i mogą w przyszłości przejąć obowiązki w zakresie zarządzania projektami badawczymi	++	++	++	++
Jakie X ma strategie badawcze?	Analiza tematyki badawczej, mechanizmów współpracy autorskiej (<i>co-authorship</i>), kluczowych badaczy i zespołów badawczych, partnerów zewnętrznych i zmian w powyższych obszarach dla wybranej organizacji	++	++	++	++
Jak ważne?	Zestawienie liczby cytowań dla pojedynczych publikacji lub patentów	++	++	-	-
Jak łącznie ważne?	Zestawienie zagregowanej liczby cytowań dla grupy publikacji lub patentów (np. pochodzących od autorów z jednej organizacji, dotyczących jednego obszaru badawczego)	++	++	-	-
Kto jest ważny?	Porównania liczby cytowań w dorobku danych badaczy oraz obliczanie wskaźników dotyczących cytowań (np. indeksów h Hirscha)	++	++	-	-
Do jakiej wiedzy nawiązuje?	Identyfikacja źródeł wiedzy, ujawnionych w cytowaniach, zawartych w publikacjach	++	++	-	-

Pytanie analityczne	Wyjaśnienie	Web of Science	Scopus	INSPEC	MEDLINE
	naukowych i zgłoszeniach patentowych (możliwe dalsze analizy zidentyfikowanego zbioru źródeł wiedzy w oparciu o pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? ile?)				
Do czyich badań nawiązuje?	Identyfikacja naukowców i wynalazców, do prac których odwołują się cytowania w publikacjach naukowych i zgłoszeniach patentowych	++	++	-	-

Tabela 61: Przydatność baz bibliograficznych do analiz, mających na celu udzielenie odpowiedzi na konkretne pytania analityczne. Źródło: opracowanie własne. Legenda:

- ++ - baza oferuje pełne możliwości odpowiedzi na to pytanie,
- + - baza oferuje ograniczone możliwości uzyskania odpowiedzi na to pytanie (brak części istotnych danych, np. systematycznej kategoryzacji treści lub informacji o afiliacjach współautorów tekstu),
- -- baza nie oferuje możliwości uzyskania odpowiedzi na to pytanie.

Analogiczne rozważania dotyczą baz patentowych. Najwygodniejsze okazują się bazy komercyjne, jednak ich koszty zwykle przekraczają możliwości finansowe badaczy. Interesującym wyjątkiem jest oferta *Patent Integration*, która kosztuje tylko 80 USD miesięcznie (48 USD w przypadku instytucji akademickich). Nie stanowi problemu wykorzystanie dla potrzeb analiz bibliometrycznych baz bezpłatnych pod warunkiem odpowiedniego przetwarzania uzyskanych danych. Ograniczeniem baz *USPTO* i *Google Patents* jest brak mechanizmów zapisywania danych patentowych na dysku lokalnym lub innych form eksportu danych - istnieją jednak programy komputerowe, które wspierają ten proces. Baza *esp@cenet* pozwala na ograniczony eksport danych. Prowadzone na dużą skalę projekty bibliometryczne opierają się często na bazach komercyjnych, które w pełni wspierają eksport i dalsze przetwarzanie rekordów danych, jednak ich wykorzystanie bywa kosztowne.

Ze względu na zróżnicowany zasięg geograficzny, bazy *USPTO* i *esp@cenet* nie mogą być traktowane jako alternatywy - pierwsza dotyczy patentów chronionych na terenie Stanów Zjednoczonych, druga - na terenie Unii Europejskiej. Wskazane jest prowadzenie równoległych analiz dla obu obszarów geograficznych. W przypadku ograniczeń czasowych i technicznych, rekomendowana jest koncentracja na danych patentowych ze Stanów Zjednoczonych jako najbardziej innowacyjnego rynku technologicznego, na którym najważniejsze wynalazki do opatentowania zgłaszają także firmy europejskiej (bazy *USPTO* lub *Google Patents*). Z punktu widzenia obserwacji badań naukowych i rozwoju technologii, nie wydają się przydatne analizy oparte o bazę patentów polskich, udzielonych przez Urząd Patentowy RP - znaczące komercyjnie i najbardziej innowacyjne wynalazki są później przez polskich wynalazców zgłaszane także do *EPO* i znajdują się w bazie *esp@cenet*, pozwalającej na porównania międzynarodowej i zawierającej bogatszy przegląd osiągnięć wynalazczych.

Pytanie analityczne	Wyjaśnienie	USPTO	Esp@ce-net	Google Patents	Patent Integration
Co?	Identyfikacja obszarów badań i technologii w zbiorze danych bibliometrycznych dla określonego kraju, regionu, ośrodka badawczego lub dla określonych lat	++	++	++	++
Co dokładnie?	Identyfikacja podobszarów badawczych w zbiorze danych bibliometrycznych, odzwierciedlających wybrany obszar badań i technologii	++	++	++	++
Kto?	Identyfikacja badaczy/wynalazców i zespołów badawczych, prowadzących badania na określony temat	++	++	++	++
Gdzie?	Identyfikacja afiliacji, w których prowadzone są badania na określony temat (poziom kraju, miasta, organizacji)	++	++	++	++
Kiedy?	Określenie momentu publikacji, zgłoszenia patentowego lub uzyskania patentu, jak również okresów, w których prowadzono badania na określony temat	++	++	++	++
Z kim?	Analiza współautorstwa publikacji i wynalazków (<i>co-authorship</i>) poprzez identyfikację osób, które wspólnie prowadziły badania	++	++	++	++
Z jaką organizacją?	Analiza współautorstwa publikacji i wynalazków (<i>co-authorship</i>) poprzez identyfikację organizacji, zatrudniających osoby, które wspólnie prowadziły badania	++	++	++	++
Ile?	Zliczanie publikacji, zgłoszeń patentowych lub udzielonych patentów dla wybranego obszaru lub podobszaru badań i technologii (możliwe jest także sporządzanie zestawień porównawczych liczby publikacji w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją?)	++	++	++	++
Ilu?	Zestawienie liczby autorów publikacji lub zgłoszeń	++	++	++	++

Pytanie analityczne	Wyjaśnienie	USPTO	Esp@ce-net	Google Patents	Patent Integration
	patentowych w danym obszarze				
Jak zmienia się w czasie?	Analiza szeregów czasowych, odzwierciedlających liczby publikacji lub patentów w wybranych obszarach badań i technologii (możliwe jest odzwierciedlenie zmian w czasie w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? ile? – czyli prezentacja historycznego rozwoju obszaru badań)	++	++	++	++
Jak będzie się rozwijać?	Prognoza krótkookresowa, oparta na ekstrapolacji szeregów czasowych (możliwe jest przygotowanie prognoz w powiązaniu z odpowiedziami na pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? pod warunkiem posiadania dostatecznie długich serii danych historycznych)	++	++	++	++
Jak popularne w porównaniu do?	Porównanie liczb publikacji, zgłoszeń patentowych lub patentów w kilku obszarach badawczych	++	++	++	++
Jakie są nisze?	Identyfikacja podobszarów badawczych, które nadal cieszą się niewielkim zainteresowaniem badaczy (w oparciu o liczbę publikacji, zgłoszeń patentowych lub patentów w danym podobszarze, porównaną z ich globalną liczbą dla szerokiego obszaru badań)	++	++	++	++
Jakie są nowe popularne tematy?	Identyfikacja nowych, cieszących się narastającą popularnością tematów badawczych (<i>hot topics</i>), w tym dynamicznie rozwijających się, nowych technologii (<i>emerging technologies</i>)	++	++	++	++
Z czym związane?	Analiza pokrewnych obszarów badań i tematów badawczych,	++	++	++	++

Pytanie analityczne	Wyjaśnienie	USPTO	Esp@ce-net	Google Patents	Patent Integration
	które równocześnie opisywane są w publikacjach lub dokumentach patentowych – na poziomie współwystępowania specjalistycznych pojęć w streszczeniach i tytułach (<i>co-words</i>) lub w słowach kluczowych i kodach klasyfikacyjnych (<i>co-classification</i>)				
W jaki sposób związane?	Analiza szczegółowych podejść badawczych, technik, środków i technologii, wykorzystywanych przez zidentyfikowane wcześniej obszary badań i tematy badawcze, pokrewne wobec podstawowego obszaru analiz	++	++	++	++
Co uzupełnia?	Identyfikacja komplementarnych obszarów wiedzy naukowej, komplementarnych umiejętności badawczych lub komplementarnych technologii, wykorzystywanych przy prowadzeniu badań w danych obszarze	++	++	++	++
Jakich wariantów nie analizowano?	Identyfikacja możliwych nowych kierunków badań i sposobów rozwiązania problemów naukowych i praktycznych w oparciu o analogie, techniki kreatywnego rozwiązywania problemów i ocenę potencjału ewolucyjnego	+	+	+	+
Kto uzupełnia?	Identyfikacja osób i organizacji, posiadających komplementarne obszary wiedzy naukowej, komplementarne umiejętności badawcze lub komplementarne technologie, wykorzystywane przy prowadzeniu badań w danym obszarze	++	++	++	++
Kto może być dostawcą?	Identyfikacja osób i organizacji, które mogą być podwykonawcami określonych zleceń badawczych ze względu na dotychczasowy dorobek naukowy	++	++	++	++
Kto może	Identyfikacja osób i	++	++	++	++

Pytanie analityczne	Wyjaśnienie	USPTO	Esp@ce-net	Google Patents	Patent Integration
być odbiorcą?	organizacji, które mogą być zleceniodawcami projektów badawczo-rozwojowych lub nabywcami/licencjobiorcami ich rezultatów ze względu na dotychczasowe zainteresowania i potrzeby badawcze				
Kto jest konkurentem?	Identyfikacja osób i organizacji, prowadzących badania w analogicznych, szczegółowych obszarach badawczych, przy wykorzystaniu zbliżonych metod (podmioty zidentyfikowane jako konkurenci mogą być też potencjalnymi partnerami w konsorcjach badawczych – zwłaszcza wtedy, gdy niezbędne jest połączenie zbliżonych zasobów dla zwiększenia skali prac badawczych)	++	++	++	++
Kto może być konkurentem?	Identyfikacja osób i organizacji, które mogą podjąć badania w analogicznych, szczegółowych obszarach badawczych, przy wykorzystaniu zbliżonych metod, a wskazuje na to ich profil badawczy, dotychczasowe zainteresowania i analizy statystyczne (zwłaszcza techniki analizy czynnikowej, pozwalające na identyfikację „ukrytych” podobieństw i zależności w dużych zbiorach danych)	++	++	++	++
Kto może być pracownikiem?	Identyfikacja osób, które prowadzą lub mogą prowadzić badania w określonym obszarze i zgromadziły zadowalający dorobek naukowy	++	++	++	++
Kto może być sukcesorem?	Identyfikacja osób, które prowadziły badania wspólnie z określonym naukowcem (<i>co-authorship</i>) i mogą w przyszłości przejąć obowiązki	++	++	++	++

Pytanie analityczne	Wyjaśnienie	USPTO	Esp@ce-net	Google Patents	Patent Integration
	w zakresie zarządzania projektami badawczymi				
Jakie X ma strategie badawcze?	Analiza tematyki badawczej, mechanizmów współpracy autorskiej (<i>co-authorship</i>), kluczowych badaczy i zespołów badawczych, partnerów zewnętrznych i zmian w powyższych obszarach dla wybranej organizacji	++	++	++	++
Jak ważne?	Zestawienie liczby cytowań dla pojedynczych publikacji lub patentów	++	+	++	
Jak łącznie ważne?	Zestawienie zagregowanej liczby cytowań dla grupy publikacji lub patentów (np. pochodzących od autorów z jednej organizacji, dotyczących jednego obszaru badawczego)	++	+	++	
Kto jest ważny?	Porównania liczby cytowań w dorobku danych badaczy oraz obliczanie wskaźników dotyczących cytowań (np. indeksów h Hirscha)	++	+	++	
Do jakiej wiedzy nawiązuje?	Identyfikacja źródeł wiedzy, ujawnionych w cytowaniach, zawartych w publikacjach naukowych i zgłoszeniach patentowych (możliwe dalsze analizy zidentyfikowanego zbioru źródeł wiedzy w oparciu o pytania: co? co dokładnie? kto? gdzie? kiedy? z kim? z jaką organizacją? ile?)	++		++	
Do czyich badań nawiązuje?	Identyfikacja naukowców i wynalazców, do prac których odwołują się cytowania w publikacjach naukowych i zgłoszeniach patentowych	++		++	

Tabela 62: Przydatność baz patentowych do analiz, mających na celu udzielenie odpowiedzi na konkretne pytania badawcze. Źródło: opracowanie własne. Legenda:

- ++ - baza oferuje pełne możliwości odpowiedzi na to pytanie,
- + - baza oferuje ograniczone możliwości uzyskania odpowiedzi na to pytanie (brak części istotnych danych, np. systematycznej kategoryzacji treści lub informacji o afiliacjach współautorów tekstu),
- - - baza nie oferuje możliwości uzyskania odpowiedzi na to pytanie.

Do przetwarzania danych, pobranych z baz bibliograficznych, niezbędne będą odpowiednie pakiety oprogramowania. Zaprezentowana wcześniej analiza wskazała na trzy najbardziej rozbudowane aplikacje - *Matheo Analysis*, *CiteSpace* i *VantagePoint*. Porównanie ich

funkcjonalności wskazało na przewagę programu *VantagePoint*, oferującego najbardziej rozbudowane możliwości, a jednocześnie opartego na doświadczeniach dużej liczby użytkowników w instytucjach administracji publicznej i firmach przemysłowych. *VantagePoint* oferuje użytkownikom bibliotekę nakładek, pozwalających na import i pracę z danymi, pochodzącymi z kilkudziesięciu baz (ang. *import filters*), co istotnie wzbogaca wartość programu. Obok komercyjnego *VantagePoint*-a, wskazane jest również wykorzystanie bezpłatnej aplikacji *CiteSpace*, która oferuje dodatkowe funkcje, związane z zaawansowaną wizualizacją danych bibliometrycznych oraz identyfikacją nowych tendencji badawczych i technicznych (ang. *emerging technologies*). *CiteSpace* współpracuje tylko z niektórymi bazami bibliometrycznymi, przede wszystkim z bazą *Web of Science*.

Tabela 63 zawiera informacje na temat możliwości współpracy poszczególnych pakietów oprogramowania z opisywanymi wcześniej bazami.

	<i>VantagePoint</i>	<i>Matheo</i>	<i>CiteSpace</i>
<i>Web of Science</i>	++	+	++
<i>Scopus</i>	++	+	+
<i>INSPEC</i>	++	+	-
<i>MEDLINE</i>	++	+	+
<i>USPTO</i>	+	++	-
<i>esp@cenet</i>	+	++	-
<i>Google Patents</i>	+	-	-
<i>Patent Integration</i>	+	-	-

Tabela 63: Możliwości wykorzystania pakietów oprogramowania do analizy danych, pobranych z poszczególnych baz bibliograficznych. Źródło: opracowanie własne. Legenda:

- ++ - oprogramowanie oferuje standardowe funkcje, pozwalające na przetwarzanie danych z określonej bazy,
- +- oprogramowanie oferuje możliwość przetwarzania danych z określonej bazy, wymaga jednak odpowiedniej konfiguracji i dostosowań,
- -- oprogramowanie nie pozwala na przetwarzanie danych z określonej bazy.

Do specyficznych zastosowań przydatne mogą też okazać się dodatkowe programy narzędziowe, wspierające poszczególne kroki w analizach bibliometrycznych:

- *HistCite* - do wizualizacji sieci cytowań publikacji w oparciu o bazę *Web of Science*,
- *HTTrack Website Copier* – pozwalające na pobieranie dokumentów patentowych z bazy *USPTO*,
- *PatentNet* - do wizualizacji sieci cytowań patentów w oparciu o bazy *Google Patents* i *esp@cenet*,
- programy narzędziowe rozwinięte przez prof. Leydesdorffa - w tym aplikacja pozwalająca na konwersję rekordów bazy *Scopus* do formatu bazy *Web of Science*.

Powyższe aplikacje są udostępniane nieodpłatnie i uzupełniają możliwości dwóch wybranych wcześniej programów.

Analizy bibliometryczne mogłyby również oprzeć się na wykorzystaniu dedykowanego oprogramowania, napisanego specjalnie dla specyficznych potrzeb badaczy - byłoby to jednak rozwiązanie kosztowne i nieefektywne z poniższych powodów:

- Rozbudowane aplikacje bibliometryczne opierają się doświadczeniach setek projektów analitycznych, a ich kolejne wersje były odpowiednio doskonalone, podczas gdy pisany od nowa program nie może skorzystać z owych najlepszych praktyk.
- Koszt pojedynczej licencji programu *VantagePoint* (7.500 USD) jest niewspółmiernie niższy od kosztów napisania nowego programu, dostawca programu oferuje też dodatkowe bezpłatne licencje programu *VantagePoint Viewer* do wykorzystania przez innych pracowników organizacji-nabywcy w celu podglądu analiz – a koszty rozwoju komercyjnie dostępnego oprogramowania są rozliczane pomiędzy wszystkimi jego użytkownikami.
- Ewentualna rozbudowa oprogramowania i uzupełnienie go o nowe funkcje będzie wymagała dodatkowych inwestycji i uzależniała badacza od jednego dostawcy usług programistycznych.
- Ze względu na regularne zmiany struktury rekordów baz bibliograficznych, koszty utrzymania systemu byłyby wysokie ze względu na konieczność jego modyfikacji.
- W Polsce brakuje potencjalnych dostawców oprogramowania, którzy posiadaliby dobrą orientację w zagadnieniach bibliometrycznych, a stworzenie oprogramowania wymagałoby zdobycia tej wiedzy i umiejętności, jak również zapoznania się z już dostępnymi rozwiązaniami.
- Pisany na zamówienie program nie będzie prawdopodobnie oferował wielu zaawansowanych funkcji, których rozwój wymagał znaczących inwestycji, wieloletnich prób i zdobycia wiedzy specjalistycznej - np. heurystyczne metody czyszczenia danych bibliometrycznych, techniki analizy czynnikowej lub wizualizacji sieci zależności.

Potencjalną zaletą wykorzystania rozwiązania pisanego na zamówienie byłaby własność kodu przez zamawiającego, co zwykle pozwala na udostępnianie programu także innym podmiotom i zespołom badawczym. Opis możliwej architektury systemu, służącego do analizy danych bibliometrycznych, można znaleźć w pracach Yu i in. (2008) oraz Malliga (2010). obie publikacje zawierają też przegląd wcześniejszej literatury, propozycje struktury relacyjnej bazy danych, wykorzystywanej do przetwarzania rekordów pobranych z zewnętrznych baz publikacji i patentów, a także przykładowe zapytania w języku *SQL*, pozwalające na generowanie podstawowych zestawień danych.

Należy jednak podkreślić, że przedstawione opisy nie oferują rozwiązania szczególnie trudnych zadań, które są niezbędne przy prowadzeniu analiz bibliometrycznych: czyszczenia danych, identyfikacji duplikatów i synonimów wśród nazw zawartych w poszczególnych listach, jak również stosowania zaawansowanych metod ilościowego przetwarzania danych. Twórcy dostępnego oprogramowania potrzebowali znaczących nakładów pracy i środków finansowych na rozwój tych elementów swoich aplikacji - przykładowo, w raporcie podsumowującym jeden z etapów rozwoju programu *VantagePoint*, przygotowanym dla

zewnętrznej instytucji finansującej prace, Frey, Minsk i Porter (2005) szczegółowo opisują wyzwania, związane z czyszczeniem danych w oparciu o metody heurystyczne i przetwarzanie języka naturalnego. Chociaż jest możliwe technicznie powtórzenie prac, podejmowanych wcześniej przez innych dostawców oprogramowania, takie działania nie wydaje się racjonalne ekonomicznie w typowych projektach analiz bibliometrycznych.

13. Podstawowe analizy bibliometryczne – przykład analizy badań nad grafenem

W 2010 roku rosyjscy naukowcy Andriej Gejm i Konstantin Nowosiołow zdobyli Nagrodę Nobla w dziedzinie fizyki za **badania nad materiałem półprzewodnikowym grafenem**, który wielu technologów uważa za możliwego następcę krzemu w przemyśle elektronicznym. 6 kwietnia 2011 r. dziennik „*Rzeczpospolita*” opisał odkrycie naukowców z Wydziału Fizyki Uniwersytetu Warszawskiego i Instytutu Technologii Materiałów Elektronicznych – metodę wytwarzania grafenu, pozwalającą na jego komercyjne wykorzystanie. Artykuł był inspirowany publikacją w internetowym wydaniu pisma naukowego „*Nano Letters*”, opisującą możliwości wykorzystania epitaksji do wytwarzania warstw grafenu na podłożu z węgla krzemu (*SiC*).

Grafen jest relatywnie mało rozpoznany materiałem półprzewodnikowym, a osoby prowadzące nad nim badania nie są powszechnie znane. Publikacja prasy popularnej może być **zaskoczeniem dla niektórych reprezentantów sektora B+R**, a **bibliometria pozwoli na szybkie zdobycie orientacji** w omawianym obszarze badawczym.

Poniżej zaprezentowane zostaną wyniki **podstawowych analiz**, przeprowadzonych wyłącznie w oparciu o bazę bibliometryczną, bez wykorzystania dodatkowych narzędzi informatycznych i **trwających łącznie około 30 minut**. Niniejszy przykład wskazuje na wymierne korzyści, jakie można osiągnąć za sprawą wykorzystania bibliometrii w codziennej pracy. Poniższe zestawienia są wynikami najprostszych, szybkich analiz, bez odwoływania się do dodatkowych narzędzi informatycznych. Dodatkowo oferuje wykorzystanie opisanych w książce narzędzi takich jak *VantagePoint* czy *CiteSpace* i prowadzenie analiz typu *tech mining*.

W omawianym przypadku, do wyszukiwania publikacji naukowych, dotyczących grafenu, wykorzystana została baza *Scopus*. Sformułowane zapytanie brzmiało “*TITLE-ABS-KEY(graphene)*” (wyszukiwanie dokumentów, w których tytule, streszczeniu lub zestawie słów kluczowych występuje słowo ‘grafen’). Wyszukiwanie dotyczyło dokumentów ze wszystkich dostępnych lat i zostało przeprowadzone dnia 6 kwietnia 2011 r. Łączna liczba odnalezionych dokumentów wynosiła 10.112, w tym 8.175 artykułów i 1.253 wystąpień konferencyjnych. Dla porównania, baza *Web of Science* zawiera 11.455 publikacji, dotyczących grafenu – jednak nie umożliwia rozbudowanych analiz słów kluczowych. W bazach patentowych znalazła się znacząca liczba dokumentów, dotyczących grafenu – 768 patentów udzielonych przez *USPTO*, 902 przez *EPO* oraz 213 zgłoszenia do *USPTO*.

Z analiz wynika, że na przestrzeni lat popularność badań nad grafenem dynamicznie wzrastała – liczba publikacji była szczególnie duża w ostatnich 5 latach (por. Tabela 64).

Rok	Liczba publikacji
1985	1
1986	1
1987	2
1988	12
1989	9
1990	3
1991	7
1992	11
1993	15
1994	16
1995	29
1996	36
1997	40
1998	51
1999	57
2000	59
2001	80
2002	95
2003	118
2004	161
2005	214
2006	330
2007	745
2008	1.332
2009	2.097
2010	3.634
2011 (do 6.04.2011)	957

Tabela 64: Liczba publikacji naukowych dotyczących grafenu w kolejnych latach. Źródło: opracowanie własne na podstawie danych pobranych z bazy *Scopus*.

Interesujące może być zapoznanie się z nazwiskami najbardziej produktywnych badaczy- autorów publikacji, dotyczących grafenu (Tabela 65).

Autor	Liczba publikacji
Geim, A.K.	72
Novoselov, K.S.	69
Guinea, F.	57
Peres, N.M.R.	53
Castro Neto, A.H.	49
Katsnelson, M.I.	49
De Heer, W.A.	43
Ruoff, R.S.	42
Das Sarma, S.	38
Mullen, K.	37
Avouris, P.	36
Otsuji, T.	35
Peeters, F.M.	35
Loh, K.P.	34
Ando, T.	33
First, P.N.	31
Ensslin, K.	31
Ruoff, R.S.	30

Autor	Liczba publikacji
Berger, C.	30
Endo, M.	28
Tour, J.M.	27
Gaskill, D.K.	27
Liang, G.	27
Katsnelson, M.I.	27
Dragoman, D.	27
Cohen, M.L.	26
Seyller, T.	26
Lin, M.F.	26
Koshino, M.	25
MacDonald, A.H.	24
Tedesco, J.L.	24
Bostwick, A.	24
Kim, P.	24
Ryzhii, V.	24
Louie, S.G.	23
Morpurgo, A.F.	23
Dresselhaus, M.S.	23
Molitor, F.	22
Lau, C.N.	22
Ihn, T.	22

Tabela 65: Liczba publikacji naukowych dotyczących grafenu autorstwa poszczególnych naukowców. Źródło: opracowanie własne na podstawie danych pobranych z bazy *Scopus*.

Dla naukowców, poszukujących publikacji na omawiany temat lub zastanawiających się nad odpowiednim medium dla upowszechnienia wyników swoich badań, ważne będzie zdobycie wiedzy o czasopismach naukowych, najczęściej publikujących artykuły dotyczące grafenu (Tabela 66).

Czasopismo	Liczba publikacji
<i>Physical Review B Condensed Matter and Materials Physics</i>	1.451
<i>Applied Physics Letters</i>	539
<i>Physical Review Letters</i>	448
<i>Carbon</i>	393
<i>Nano Letters</i>	335
<i>Journal of Physical Chemistry C</i>	253
<i>ACS Nano</i>	214
<i>Journal of Physics Condensed Matter</i>	161
<i>Nanotechnology</i>	159
<i>Journal of Applied Physics</i>	149
<i>Journal of Materials Chemistry</i>	132
<i>Physica E Low Dimensional Systems and Nanostructures</i>	111
<i>Journal of the American Chemical Society</i>	111
<i>Chemical Physics Letters</i>	107
<i>Advanced Materials</i>	101
<i>Solid State Communications</i>	101
<i>New Journal of Physics</i>	100
<i>Materials Research Society Symposium Proceedings</i>	89
<i>Chemical Communications</i>	84
<i>Physica Status Solidi B Basic Research</i>	80
<i>Journal of Chemical Physics</i>	77
<i>Ecs Transactions</i>	76
<i>Nature Nanotechnology</i>	76
<i>Langmuir</i>	68
<i>Small</i>	68
<i>Journal of Physics Conference Series</i>	68
<i>Science</i>	66
<i>Chemistry of Materials</i>	64
<i>Physics Letters Section A General Atomic and Solid State Physics</i>	63
<i>Journal of the Physical Society of Japan</i>	57
<i>Epl</i>	57
<i>Electrochimica Acta</i>	55
<i>European Physical Journal B</i>	52
<i>AIP Conference Proceedings</i>	51
<i>Nature</i>	50
<i>Electrochemistry Communications</i>	50
<i>Proceedings of SPIE the International Society for Optical Engineering</i>	50
<i>Journal of Physical Chemistry B</i>	49
<i>Physica B Condensed Matter</i>	48
<i>Nano Research</i>	47

Tabela 66: Liczba publikacji dotyczących grafenu publikowana w poszczególnych czasopismach naukowych. Źródło: opracowanie własne na podstawie danych pobranych z bazy Scopus.

Decydenci zajmujący się polityką naukową, kierownicy jednostek naukowych oraz indywidualni badawcze zainteresują się z pewnością zestawieniem najbardziej produktywnych organizacji, które publikują wyniki badań nad grafenem (Tabela 67).

Instytucja badawcza	Liczba publikacji
<i>National University of Singapore</i>	205
<i>Massachusetts Institute of Technology</i>	160
<i>Tsinghua University</i>	158
<i>University of Texas at Austin</i>	152
<i>Instituto de Ciencia de Materiales de Madrid</i>	151
<i>Japan Science and Technology Agency</i>	150
<i>Lawrence Berkeley National Laboratory</i>	148
<i>Nanyang Technological University</i>	145
<i>UC Berkeley</i>	143
<i>Tohoku University</i>	131
<i>Columbia University in the City of New York</i>	120
<i>Georgia Institute of Technology</i>	115
<i>Russian Academy of Sciences</i>	111
<i>National Institute of Advanced Industrial Science and Technology</i>	111
<i>University of Manchester</i>	107
<i>Tokyo Institute of Technology</i>	106
<i>University of Tokyo</i>	102
<i>Rice University</i>	99
<i>Radboud University Nijmegen</i>	98
<i>Pennsylvania State University</i>	96
<i>University of California, Riverside</i>	95
<i>Purdue University</i>	92
<i>University of Maryland</i>	92
<i>Universiteit Antwerpen</i>	91
<i>Northwestern University</i>	88
<i>CNRS Centre National de la Recherche Scientifique</i>	87
<i>Boston University</i>	85
<i>Nanjing University</i>	83
<i>IBM Thomas J. Watson Research Center</i>	82
<i>University of Cambridge</i>	79
<i>Korea Advanced Institute of Science & Technology</i>	79
<i>Institute of Physics Chinese Academy of Sciences</i>	76
<i>Chinese Academy of Sciences</i>	75
<i>Max Planck Institute for Solid State Research</i>	74
<i>National Institute for Materials Science Tsukuba</i>	73
<i>University of Tsukuba</i>	71
<i>Universite Paris-Sud XI</i>	71
<i>Fudan University</i>	70
<i>Naval Research Laboratory</i>	70
<i>Harvard University</i>	68
<i>Lancaster University</i>	68
<i>Pohang University of Science and Technology</i>	68
<i>Consiglio Nazionale delle Ricerche</i>	67
<i>University of California, Los Angeles</i>	67
<i>Sungkyunkwan University</i>	66
<i>Universidade do Minho</i>	66
<i>Indian Institute of Science</i>	65
<i>Cornell University</i>	61
<i>Osaka University</i>	61
<i>Universität Erlangen-Nürnberg</i>	60
<i>University of Pennsylvania</i>	60
<i>University of Illinois at Urbana-Champaign</i>	58

Instytucja badawcza	Liczba publikacji
<i>Oak Ridge National Laboratory</i>	58
<i>Nagoya University</i>	57
<i>University of Oxford</i>	57
<i>Kyoto University</i>	56
<i>University of Wollongong</i>	56
<i>Peking University</i>	56
<i>Jawaharlal Nehru Centre for Advanced Scientific Research</i>	56
<i>Princeton University</i>	54
<i>National Cheng Kung University</i>	54
<i>University of Florida</i>	52
<i>Fritz Haber Institute of the Max Planck Society</i>	51
<i>Hokkaido University</i>	51
<i>Rensselaer Polytechnic Institute</i>	51
<i>North Carolina State University</i>	50
<i>Jilin University</i>	50
<i>Linköpings Universitet</i>	50
<i>Los Alamos National Laboratory</i>	49
<i>Bogolyubov Institute for Theoretical Physics Nasu</i>	49
<i>Seoul National University</i>	49
<i>Stanford University</i>	48
<i>Shinshu University</i>	48
<i>Nankai University</i>	47
<i>Leiden University</i>	47
<i>Ecole Polytechnique Federale de Lausanne</i>	46
<i>Trinity College Dublin</i>	45
<i>Graduate University of Chinese Academy of Sciences</i>	45
<i>Technische Universität Dresden</i>	44
<i>National Tsing Hua University</i>	43
<i>Eidgenössische Technische Hochschule Zurich</i>	43
<i>Hefei National Laboratory for Physical Sciences at Microscale</i>	43
<i>Max Planck Institute for Polymer Research</i>	42
<i>Universität Regensburg</i>	42
<i>Changchun Institute of Applied Chemistry Chinese Academy of Sciences</i>	41
<i>Brown University</i>	40
<i>Universidad Autónoma de Madrid</i>	40
<i>Universite Pierre et Marie Curie</i>	39
<i>Arizona State University</i>	39
<i>Universite Joseph Fourier</i>	39
<i>University of Aizu</i>	39
<i>Delft University of Technology</i>	38
<i>Consejo Superior de Investigaciones Científicas</i>	38
<i>Universite Henri Poincare</i>	38
<i>East China University of Science and Technology</i>	38
<i>Universidade Federal de Minas Gerais</i>	38
<i>College of Chemistry and Molecular Engineering, Peking University</i>	38
<i>Zhejiang University</i>	37
<i>Kyushu University</i>	37
<i>University of Notre Dame</i>	36
<i>National Institute of Standards and Technology</i>	36
<i>Brookhaven National Laboratory</i>	35

Instytucja badawcza	Liczba publikacji
<i>National High Magnetic Field Laboratory</i>	35
<i>University of Science and Technology of China</i>	35
<i>Nanjing University of Aeronautics and Astronautics</i>	35
<i>University of Utah</i>	35
<i>XiangTan University</i>	35
<i>Rutgers, The State University of New Jersey</i>	35
<i>NEC Corporation</i>	35
<i>Universidade Federal do Ceara</i>	34
<i>Institute of Metal Research Chinese Academy of Sciences</i>	34
<i>University of California, San Diego</i>	34
<i>Southeast University</i>	33
<i>Leibniz Institut fur Festkorper und Werkstofforschung Dresden</i>	33
<i>Universita degli Studi di Catania</i>	33
<i>CEA Grenoble</i>	33
<i>Imperial College London</i>	33
<i>Universität Augsburg</i>	33
<i>Institute for Materials Research, Tohoku University</i>	33
<i>Universite Montpellier 2 Sciences et Techniques</i>	33
<i>Sharif University of Technology</i>	33
<i>Korea University</i>	32
<i>Pacific Northwest National Laboratory</i>	32
<i>Shanghai Jiaotong University</i>	32
<i>Facultes Universitaires Notre-Dame de la Paix</i>	31
<i>Sandia National Laboratories, New Mexico</i>	31
<i>Uniwersytet Warszawski</i>	31
<i>University of Exeter</i>	31
<i>Argonne National Laboratory</i>	31
<i>Universita di Pisa</i>	31
<i>Laboratorium fur Festkorperphysik</i>	31
<i>Max-Planck-Institute for the Physics of Complex Systems</i>	31
<i>Indiana University</i>	30
<i>Universidad del Pais Vasco - Euskal Herriko Unibertsitatea</i>	30
<i>University of Queensland</i>	30
<i>Institute of High Performance Computing, Singapore</i>	29
<i>University of Illinois</i>	29
<i>University of Texas at Dallas</i>	29
<i>Meijo University</i>	29
<i>California Institute of Technology</i>	29
<i>Michigan State University</i>	29
<i>Landau Institute for Theoretical Physics</i>	29
<i>Aalto University School of Science and Technology</i>	29
<i>Universitatea din Bucuresti</i>	28
<i>Research Institute for Technical Physics and Materials Science Hungarian Academy of Sciences</i>	28
<i>Université Simon Fraser</i>	28
<i>Tianjin University</i>	28
<i>Nippon Telegraph & Telephone</i>	28
<i>Universidade Estadual de Campinas</i>	28
<i>National Institute for Research and Development in Microtechnologies Romania</i>	27
<i>University of Houston</i>	27
<i>University of Central Florida</i>	27

Instytucja badawcza	Liczba publikacji
<i>Institute of Solid State Physics Chinese Academy of Sciences</i>	27
<i>Hunan University</i>	27
<i>Universite Catholique de Louvain</i>	26
<i>Instituto Potosino de Investigacion Cientifica y Tecnologica</i>	26
<i>Technische Universität Berlin</i>	26
<i>Universität Hamburg</i>	26
<i>University of Minnesota Twin Cities</i>	26
<i>Beijing National Laboratory for Molecular Sciences</i>	26

Tabela 67: Najbardziej aktywne instytucje, publikujące wyniki badań nad grafenem. Źródło: opracowanie własne na podstawie danych pobranych z bazy Scopus.

Tabela 67 wskazuje, że Uniwersytet Warszawski jest najaktywniejszą w Polsce instytucją publikującą badania nad grafenem. Jego łączna liczba publikacji wynosi 31 artykułów z lat 2006-2011. Ta informacja może być przydatna dla resortu nauki, poszukującego potencjalnego gracza na scenie innowacji, który mógłby aktywnie reprezentować Polskę i uczestniczyć w międzynarodowych projektach badawczych. Lepsze zrozumienie specyfiki polskich badań wymaga jednak sprawdzenia, jacy naukowcy prowadzą omawiane badania.

Tabela 68 prezentuje autorów, publikujących artykuły dotyczące grafenu, które opatrzone afiliacją Uniwersytetu Warszawskiego (nie wszyscy są pracownikami UW, niektórzy tylko współpracują z uczelnią przy realizacji projektów badawczych).

Autor	Liczba publikacji
Baranowski, J.M.	12
Tworzydło, J.	12
Strupinski, W.	10
Beenakker, C.W.J.	11
Borysiuk, J.	8
Grodecki, K.	7
Bozek, R.	7
Wysmolek, A.	7
Stepniewski, R.	5
Akhmerov, A.R.	4
Drabinska, A.	4
Potemski, M.	4
Faugeras, C.	3
De Heer, W.A.	3
Kosciewicz, K.	3
Rycerz, A.	3
Berger, C.	3
Kossacki, P.	2
Stepniewski, R.	2
Korona, K.P.	2
Sprinkle, M.	2
Pelc, M.	2
Marciak-Kozłowska, J.	2
Groth, C.W.	2
Snyman, I.	2
Amado, M.	2
Bardarson, J.H.	2
Plochocka, P.	1

Autor	Liczba publikacji
Wysmollek, A.	2
Recher, P.	1
Schleede, J.	1
Schubert, G.	1
Sepkhanov, R.A.	1
Soucy, G.	1
Witowski, A.M.	1
Vollhardt, D.	1
Tymicki, E.	1

Tabela 68: Autorzy publikacji na temat grafenu, które opatrzone afiliacją Uniwersytetu Warszawskiego. Źródło: opracowanie własne na podstawie danych pobranych z bazy Scopus.

Publikacje Uniwersytetu Warszawskiego, dotyczące grafenu, opisane są w bazie Scopus przy pomocy następujących słów kluczowych: *4H-SiC substrate, Absorption, A-carbon, Acetylene, AFM, Agglomeration, Angle of incidences, Annealing time, Atomic force microscopy, Atomic physics, Average numbers, Carbon layers, CVD, Dirac fermions, Electrons, Epitaxial graphene, Epitaxy, Fermions, Graphene growth, Graphene layers, Graphite, Heat transfer, High-resolution TEM, Off-axis, Olefins, Phonon excitations, Raman scattering, Scanning electron microscopy, Scanning tunneling microscopy, Scattering, Semiconducting silicon compounds, Shot noise, SiC substrates, Silicon carbide, Substrates, TEM, Transmission electron microscopy.*

Naukowcy z Uniwersytetu Warszawskiego w badaniach nad grafenem współpracowali z reprezentantami następujących ośrodków badawczych: *CNRS Centre National de la Recherche Scientifique; Cornell University; Ecole des Mines de Paris; Ernst-Moritz-Arndt-Universität Greifswald; Georgia Institute of Technology; Institut Néel; Institute of Physics Academy of Sciences of the Czech Republic; Instytut Fizyki Polskiej Akademii Nauk; Instytut Technologii Elektronowej; Instytut Technologii Materiałów Elektronicznych w Warszawie; Lawrence Berkeley National Laboratory; LCMi Laboratoire des Champs Magnétiques Intenses; Leiden University; LNCMI-CNRS; Ludwig-Maximilians-Universität München; Politechnika Warszawska; UC Berkeley; Universidad Complutense de Madrid; Universität Augsburg; Universität Erlangen-Nürnberg; Universität Konstanz; Université de Sherbrooke; Université Joseph Fourier; Université Montpellier 2 Sciences et Techniques; Univerzita Karlova v Praze; Uniwersytet Jagielloński w Krakowie; Uniwersytet Marii Curie-Skłodowskiej w Lublinie; Whiteshell Lab.*

Najczęściej cytowane polskie artykuły dotyczące grafenu to:

- Tworzydło, J., Trauzettel, B., Titov, M., Rycerz, A., Beenakker, C.W.J. (2006) Subpoissonian shot noise in graphen. *Physical Review Letters*, 96, 24 (170 cytowań w bazie Scopus),
- Rycerz, A., Tworzydło, J., Beenakker, C.W.J. (2007) Valley filter and valley valve in grapheme. *Nature Physics*, 3, 3 (167 cytowań w bazie Scopus).

Ich współautorem jest dr hab. Jakub Tworzydło, pracownik Instytutu Fizyki Teoretycznej Wydziału Fizyki Uniwersytetu Warszawskiego, a zarazem autor 40 publikacji zebranych w bazie Scopus. Publikuje od 1994 roku, zdobył łącznie 737 cytowań (jego indeks h wynosi 13) i współpracował z 46 autorami. Wcześniej publikował też pod afiliacją Uniwersytetu w Lejdzie w Holandii.

Współautorzy publikacji dra hab. Jakuba Tworzydło reprezentują: *Chalmers Tekniska Högskola; Chung Yuan Christian University; CNRS Centre National de la Recherche Scientifique; Delft University of Technology; Institut für Halbleiterphysik und Optik; Institute of Biomedical Sciences Academia Sinica Taiwan; Institute of Physics Polish Academy of Sciences; Institut-Lorentz for Theoretical Physics, Leiden; Instytut Technologii Elektronowej; Instytutu Technologii Materiałów Elektronicznych w Warszawie; Johannes Kepler Universität Linz; Julius-Maximilians-Universität Würzburg; Laboratoire de Physique Statistique UMR 8550; Lancaster University; Landau Institute for Theoretical Physics; Lawrence Berkeley National Laboratory; Leiden University; Los Alamos National Laboratory; Max Planck Gesellschaft zur Förderung der Wissenschaften; Max-Planck-Institute for the Physics of Complex Systems; National Taiwan University; National Taiwan University Hospital; National Tsing Hua University; Polish Academy of Sciences; Politechnika Wroclawska; Stanford University; Technische Universität Eindhoven; Tohoku University; Tufts University; UC Berkeley; Universität Konstanz; Universität Regensburg; Universität Wien; Université de Geneve; Université Montpellier 2 Sciences et Techniques; Université Pierre et Marie Curie; University at Buffalo State University of New York; University of Missouri-Kansas City; Uniwersytet Jagielloński w Krakowie; Uniwersytet Warszawski.*

Publikacje dra hab. Jakuba Tworzydło dotyczyły następujących zagadnień (słów kluczowych): *Acoustic wave velocity, Acoustic waves, Algorithms, Alloys, Amplification, Amplifying disordered waveguides, Andersons, Angle of incidences, Approximation theory, Asymmetric field dependent spin splitting, A-thermal, Band edge, Band gaps, Band insulators, Band structure, Bosonic creation operators, Bound state, Calculations, Chaos theory, Classical dynamics, Coherent light, Computer simulation, Concentration dependence, Conductance, Critical angles, Crystal impurities, Crystals, Decay modes, Defects, Degree of freedoms, Degrees of freedom mechanics, Diffusion, Diluted magnetic semiconductor, Dirac fermions, Dirac Hamiltonian, Disordered systems, Effective medium theories, Eigenvalues and eigenfunctions, Electric conductivity, Energy gap, Exchange and super-exchange, Graphene, High energy physics, Impurities in semiconductors, Light absorption and reflection, Magnetic field effects, Magnetic semiconductors, Magnetism, Mathematical analysis, Matrix algebra, Metal surfaces, Model, Molecular dynamics, Molecular field approximation, Molecular physics, Multimode theory, Nanostructured materials, N-doped, Negative magnetoresistance, Noise, Nonperturbative theory, Numerical methods, Off diagonal scattering term, Olefins, Open quantum systems, Optical communication, Optical properties, Optical waveguides, Oscillations, Parameter estimation, Parameter scaling, Perturbation techniques, Phase space methods, Phase transitions, Phosphorus, Photodetectors, Photons, Physicochemical model, Poisson equation, Polarized magnetorelectance, Polymers, Probability, Propagating modes, Pseudospin, Pyrolysis, Quantized conductance, Quantum dot, Quantum mechanics, Quantum optics, Quantum theory, Quantum well, Quasibound states, Random impurity scattering, Random matrix theory, Random potentials, Random processes, Rayleigh distribution, Reflected beams, Refractive index, Resistors, Scale-dependent velocity, Scaling functions, Scattering matrix, Scattering theory, Semiconducting cadmium compounds, Semiconducting gallium arsenide, Semiconductor doping, Semiconductor quantum dots, Sensitivity analysis, Shot noise, Signal to noise ratio, Silicon carbide, Spin dynamics, Spin splitting, Spin-orbit couplings, Spin-rotation symmetry, Spurious signal noise, Statistical methods, Stepwise increase, Stroboscopy, Sub lattices, Superconducting materials, Thermal decompositions, Time-reversal, Total internal reflections, Total reflections, Trajectories, Transport data, Transport experiments, Transport properties, Transverse resistivity, Two-dimensional superconductors, Universal conductance fluctuations (UCF), Virtual crystal approximation, Von Neumann entropy, Vortex flow, Wave equations, Wave packets, Wave superconductors, Wavelet transforms, Weak disorder, Weak localization (WL), Zero field concentration dependence.*

Twórcą polskiej metody wytwarzania grafenu jest z kolei prof. dr hab. Jacek M. Baranowski, reprezentujący dwie instytucje: Instytut Fizyki Eksperymentalnej Wydziału Fizyki Uniwersytetu Warszawskiego oraz Instytut Technologii Materiałów Elektronicznych w Warszawie. Publikuje od 1967 roku, na przestrzeni lat napisał 172 zgromadzone w bazie Scopus publikacje ze 150 współautorami i otrzymał łącznie 1420 cytowania (indeks h = 19).

Wcześniej prof. Baranowski publikował też pod wieloma innymi afiliacjami: *Aalto University School of Science and Technology, Helsinki University of Technology, Espoo, Finland; Berkeley Lab, Materials Sciences Division, Berkeley, United States; Institute of Plasma Physics, Department of Pulse Plasma Systems, Prague, Czech Republic; Instytut Problemów Jądrowych im. Andrzeja Sołtana; Instytut Problemów Jądrowych, Warszawa; Instytut Technologii Elektronowej, Warszawa; Instytut Wysokich Ciśnień Polskiej Akademii Nauk, Warszawa; Karl-Franzens-Universität Graz, Graz, Austria; Laboratoire de Physique Statistique UMR 8550, Paris, France; Linköpings Universitet, Linköping, Sweden; Max Planck Gesellschaft zur Förderung der Wissenschaften, München, Germany; Office of Naval Research, Arlington, United States; Stanford University, Stanford Electronics Laboratories, Palo Alto; Universidad de Buenos Aires, Buenos Aires, Argentina; Universität Bremen, Bremen, Germany; Universite Pierre et Marie Curie, Laboratory of Physics and Solids, Paris, France; University of California, Berkeley, Department of Materials Science and Mineral Engineering, Berkeley, United States; University of St Andrews, School of Physical Sciences, St Andrews, United Kingdom.*

Współautorzy publikacji naukowych profesora Baranowskiego reprezentują następujące instytucje: *Aalto University School of Science and Technology; Australian National University; CEA Grenoble; CNRS Centre National de la Recherche Scientifique; Conseil national de recherches Canada; CRHEA Centre de Recherche sur l'Hétéro-Epitaxie et ses Applications; Ecole des Mines de Paris; FOM Institute for Plasma Physics - Rijnhuizen; Helmholtz-Zentrum Berlin für Materialien und Energie GmbH; Humboldt-Universität zu Berlin; Institute for Nuclear Problems; Institute of Physics Academy of Sciences of the Czech Republic; Institute of Plasma Physics, Academy of Sciences of the Czech Republic; Institute of Plasma Physics, Kharkov; Institute of Precision Mechanics; Instituto de Fisica del Plasma, Buenos Aires; Instytut Chemii Fizycznej Polskiej Akademii Nauk; Instytut Chemii i Techniki Jądrowej, Warszawa; Instytut Fizyki Polskiej Akademii Nauk; Instytut Problemów Jądrowych; Instytut Technologii Elektronowej; Instytutu Technologii Materiałów Elektronicznych w Warszawie; Jet Propulsion Laboratory, California Institute of Technology; Kansas State University at Manhattan; Karl-Franzens-Universität Graz; Kharkiv National University; Kirtland Air Force Base; Laboratoire de Physique Statistique UMR 8550; Lawrence Berkeley National Laboratory; LCMI Laboratoire des Champs Magnétiques Intenses; Linköpings universitet; Macquarie University; Max Planck Institute for Solid State Research; Meijo University; Ministerstwo Środowiska; MPI-FKF/CNRS; Office of Naval Research; Pn Lebedev Physics Institute, Russian Academy of Sciences; Politechnika Warszawska; Samsung Group; Stanford University; UC Berkeley; UC MERCED; United Kingdom Ministry of Defence; Universidad de Buenos Aires; Universidad Nacional del Centro de la Provincia de Buenos Aires; Universität Bonn; Universität Bremen; Universite Joseph Fourier; Universite Montpellier 2 Sciences et Techniques; Universite Pierre et Marie Curie; University of Nottingham; University of St Andrews; University of Technology Sydney; Univerzita Karlova v Praze; Uniwersytet w Białymstoku; Uniwersytet Warszawski; Van der Waals-Zeeman Institute for Experimental Physics; VTT Technical Research Centre of Finland; Wright State University; Wright-Patterson AFB.*

Słowa kluczowe opisujące publikacje Baranowskiego to: *Adsorption, Acceptors, AlGaN layers, Aluminum compounds, Aluminum gallium nitride, Aluminum nitride, Anisotropy, Annealing, Atomic force microscopy, Cadmium selenide, Calibration, Capillary pinch, Carbon layers, Carrier concentration, Cathodoluminescence, Crystal defects, Crystal growth, Crystal lattices, Defects, Deuterium, Diffusion, Dislocations crystals, Donors, Electric conductivity, Electric Discharges, Electric discharges, Electric field effects, Electric potential, Electric properties, Electric resistance, Electrodes, Electron emission, Electron energy levels, Electron gas, Electron microscopy, Electron scattering, Electron transitions, Electronic structure, Elementary particle, Energy gap, Epitaxial graphene, Epitaxial growth, Etching, Excitons, Film growth, Gallium compounds, Gallium nitride, Gallium phosphide, Gas Puffed Discharges, Graphene growth, Graphene layers, Graphite, Ground state, Hall effect, Hamiltonians, Heterojunctions, High pressure effects in solids, High temperature effects, High-resolution TEM, Homoepitaxial growth, Hydrogen, Infrared radiation, Interfaces materials, Ion beams, Ion implantation, Ion-beams, Kelvin force microscopy (KFM), Light absorption, Light emission, Low temperature effects, Low temperature operations, Luminescence, Luminescence of solids, Magnetic field effects, Magnetic resonance, Mass spectrometry, Metallorganic chemical vapor deposition, Metallorganic vapor phase epitaxy, Molecular beam epitaxy, Molecular vibrations, Neutron yield, Nitrides, Nitrogen, Nuclear track detectors, Off-axis, Optical microscopy, Optical properties, Optical variables measurement, Oscillations, Oxygen, Phonons, Photoluminescence, Photon energy, Plasma applications, Plasma focus, Plasma theory, Positron annihilation, Quenching, Radiation detectors, Rutherford backscattering spectroscopy, Sapphire, Scanning electron microscopy, Scanning tunneling microscopy, Schottky barrier diodes, Secondary ion mass spectrometry, Semiconducting aluminum compounds, Semiconducting films, Semiconducting gallium arsenide, Semiconducting gallium compounds, Semiconducting indium compounds, Semiconducting indium gallium arsenide, Semiconducting silicon compounds, Semiconducting zinc compounds, Semiconducting-cadmium compounds, Semiconductor doping, Semiconductor growth, Semiconductor quantum dots, Semiconductor quantum wells, SiC substrates, Silanes, Silicon carbide, Single crystals, Spectroscopic analysis, Spectroscopy - absorption, Spectrum analysis, Strain, Substrates, Superlattices, TEM, Thermal effects, Transmission electron microscopy, Uniaxial stress, Vapor phase epitaxy, X ray diffraction analysis.*

Profesor Baranowski uczestniczył w przeszłości w rozwoju innych, ważnych naukowo i komercyjnie technologii półprzewodnikowych. Był wskazany jako współtwórca wynalazków, zgłoszonych do *EPO* i *USPTO* przez Instytut Badań Wysokociśnieniowych Polskiej Akademii Nauk, dotyczących podłoży z azotku galu, wykorzystywanych później do produkcji niebieskiego lasera półprzewodnikowego. Współpracował też jako współtwórca innego wynalazku, zgłoszonego do UPRP przez Uniwersytet Warszawski, z Robertem Dwilińskim, późniejszym założycielem firmy *Ammono*, ważnego gracza na globalnym rynku optoelektronicznym.

Powyższe informacje można zdobyć i opracować w oparciu o bazę *Scopus* w przeciągu około 30 minut, bez wykorzystania dodatkowych narzędzi informatycznych oraz bez potrzeby indywidualnych kontaktów z naukowcami lub kierownictwem jednostek naukowych. Zastosowanie specjalistycznego oprogramowania pozwoli dodatkowo na przygotowanie zestawień i pogłębionych analiz, opisanych w niniejszej książce. Omówiony przykład pokazuje potencjał, tkwiący w bibliometrii jako metodzie wspierającej obserwację badań naukowych i rozwoju technologii.

14. Analiza bibliometryczna aktywności publikacyjnej polskich uczelni wyższych, 2000-2009

Uczelnie wyższe w Polsce rywalizują o środki na badania i dążą do uzyskiwania najwyższych pozycji w rankingach, procedurach ewaluacyjnych i akredytacyjnych. Bibliometria pozwala na zgromadzenie cennych danych porównawczych na temat poszczególnych instytucji.

Poniżej zaprezentowane zostanie porównanie aktywności publikacyjnej 24 najbardziej aktywnych uczelni wyższych w Polsce w latach 2000-2009. Dane wykorzystane do analiz zostały pobrane w czerwcu 2011 r. z bazy *Scopus*. W momencie gromadzenia danych, baza *Scopus* wymieniała aż 566 afiliacji w Polsce, jednak wiele spośród nich było duplikatami (z nieznacznie zmienioną nazwą). Tylko 25 polskich instytucji opublikowało ponad 5.000 artykułów i wystąpień konferencyjnych – pozostałe instytucje z listy były znacznie mniej produktywne. Wśród 25 organizacji znalazła się również Polska Akademia Nauk, która ze względu na specyfikę i skalę prowadzonej działalności nie nadawała się do bezpośrednich porównań z uczelniami wyższymi.

Dla pozostałych 24 uczelni przeprowadzono wyszukiwania w bazie, gromadząc podzbiór danych, który następnie został poddany analizom bibliometrycznym. Warto wspomnieć, że *Scopus* ułatwia poszukiwania publikacji pochodzących z konkretnej instytucji dzięki oznaczaniu instytucji unikalnymi numerami ID, co pozwala uniknąć niektórych błędów wynikających ze zróżnicowanego nazewnictwa lub podawania nazw tych samych instytucji czasami po angielsku a czasami po polsku. Przykładowo, publikacje stworzone przez badaczy z Politechniki Warszawskiej zidentyfikowano dzięki poniższemu poleceniu wyszukiwania:

```
(AF-ID("Politechnika Warszawska" 60003675)) AND (LIMIT-TO(PUBYEAR, 2000) OR  
(LIMIT-TO(PUBYEAR, 2001) OR (LIMIT-TO(PUBYEAR, 2002) OR (LIMIT-  
TO(PUBYEAR, 2003) OR (LIMIT-TO(PUBYEAR, 2004) OR (LIMIT-TO(PUBYEAR,  
2005) OR (LIMIT-TO(PUBYEAR, 2006) OR (LIMIT-TO(PUBYEAR, 2007) OR (LIMIT-  
TO(PUBYEAR, 2008) OR (LIMIT-TO(PUBYEAR, 2009)
```

Dalsze procedury czyszczenia danych ujawniły konieczność usunięcia niektórych rekordów, błędnie przypisanych w bazie do analizowanych instytucji. Do właściwych analiz wykorzystano tylko dane dotyczące opublikowanych artykułów naukowych, eliminując wystąpienia konferencyjne, książki i inne rodzaje publikacji. Wskaźniki budowano w oparciu o metodę *whole counting* (por. podrozdział 8.9), czyli każdej instytucji, na której pracowali współautorzy analizowanego artykułu, przyporządkowano jedną publikację. Zgromadzono łączne liczby artykułów z lat 2000-2009, jak również sumaryczne zestawienia liczb cytowań bezpośrednich wszystkich artykułów autorstwa pracowników danej uczelni, opublikowanych w latach 2000-2009 (cytowania obejmowały okres od roku 2000 do momentu prowadzenia analiz czyli do czerwca 2011 r.). Zestawienie wyników prezentuje Tabela 69.

Uczelnia	Liczba artykułów	Liczba cytowań	Liczba cytowań / liczba artykułów
Uniwersytet Warszawski	9.735	104.706	10,76
Uniwersytet Jagielloński, Kraków	12.722	122.374	9,62
Politechnika Warszawska	6.078	41.882	6,89
Politechnika Wrocławska	5.791	35.069	6,06
Uniwersytet im. Adama Mickiewicza, Poznań	6.465	42.625	6,59
Uniwersytet Wrocłowski	4.873	46.295	9,50
Warszawski Uniwersytet Medyczny	4.995	39.192	7,85
Uniwersytet Medyczny w Łodzi	5.048	33.999	6,74
Politechnika Łódzka	4.117	21.549	5,23
Śląski Uniwersytet Medyczny	4.429	28.354	6,40
Uniwersytet Śląski	4.350	28.436	6,54
Uniwersytet Medyczny w Poznaniu	3.938	22.935	5,82
Uniwersytet im. Marii Curie-Skłodowskiej, Lublin	3.417	21.870	6,40
Akademia Górniczo-Hutnicza, Kraków	5.845	30.126	5,15
Gdański Uniwersytet Medyczny	3.785	34.889	9,22
Uniwersytet Medyczny w Lublinie	3.844	21.438	5,58
Politechnika Gdańska	3.182	22.961	7,22
Uniwersytet Łódzki	3.232	24.433	7,56
Akademia Medyczna, Wrocław	4.873	21.824	4,48
Uniwersytet im. Mikołaja Kopernika, Toruń	5.582	40.086	7,18
Uniwersytet Medyczny w Białymstoku	3.671	21.742	5,92
Politechnika Śląska	2.891	15.248	5,27
Uniwersytet Gdański	3.217	27.150	8,44
Politechnika Poznańska	2.348	19.601	8,35
Łącznie dla 24 uczelni	118.428	868.784	7,34

Tabela 69: Liczba artykułów i cytowań 24 najbardziej produktywnych uczelni w Polsce z lat 2000-2009. Źródło danych: baza Scopus.

Zaprezentowane dane ujawniają istotne różnice między poszczególnymi uczelniami. Dwie spośród nich – Uniwersytet Warszawski i Uniwersytet Jagielloński – prowadzą pod względem liczby publikacji i cytowań, pozostawiając w tyle kolejne 22 instytucje. Uniwersytet Jagielloński odpowiada za 10,74% publikacji wszystkich 24 uczelni i 14,08% wszystkich cytowań w podzbiorze danych, podczas gdy Uniwersytet Warszawski stworzył 8,22% wszystkich publikacji podzbioru i otrzymał 12,05% spośród wszystkich objętych analizą cytowań. Pozostałe uczelnie osiągały istotnie niższe wyniki, a każda odpowiadała tylko 2-5,1% publikacji zbioru i 1,8-5,3% cytowań.

Warto pamiętać, że Uniwersytet Jagielloński obejmuje *Collegium Medicum* (w bazie *Scopus* wyodrębnione jako "*Collegium Medicum Uniwersytetu Jagiellońskiego*" i opatrzone identyfikatorem 60021323), podczas gdy Uniwersytet Warszawski nie posiada własnego wydziału medycznego. W czerwcu 2011 roku, ponad 21% wszystkich publikacji Uniwersytetu Jagiellońskiego pochodziło od autorów, pracujących w strukturach Collegium Medicum, dlatego bezpośrednie porównanie dwóch największych polskich uczelni wydaje się trudne. Chociaż całkowita liczba publikacji Uniwersytetu Jagiellońskiego przewyższa odpowiedni wskaźnik dla Uniwersytetu Warszawskiego, średnia liczba cytowań przypadających na jedną publikację jest wyższa dla artykułów autorstwa pracowników Uniwersytetu Warszawskiego, co może świadczyć o większym znaczeniu prac warszawskich badaczy dla światowych środowisk naukowych.

Na kolejnych pozycjach pod względem liczby publikacji są: Uniwersytet im. Adama Mickiewicza w Poznaniu, Politechnika Warszawska, Akademia Górniczo-Hutnicza w Krakowie, Politechnika Wrocławska i Uniwersytet im. Mikołaja Kopernika w Toruniu. Te uczelnie generowały znaczące liczby publikacji, jednak nie otrzymywały proporcjonalnie wysokiej liczby cytowań – co oznacza, że badacze z wymienionych uczelni są bardzo produktywni, ale ich publikacje nie mają podobnie wysokiego wpływu na dyscypliny badawcze (choć indywidualne publikacje mogą być oczywiście wpływowe i ważne naukowo). Niektóre uczelnie otrzymują z kolei relatywnie dużo cytowań przy mniejszej liczbie publikacji i mogą być postrzegane jako szczególnie dobrze skoncentrowane na badaniach, które wywierają istotny wpływ na światową naukę. Wysokie poziomy wskaźnika cytowań do publikacji charakteryzują: Uniwersytet Wrocławski, Gdański Uniwersytet Medyczny, Uniwersytet Gdański i Politechnikę Poznańską. Również Warszawski Uniwersytet Medyczny i Uniwersytet Łódzki utrzymują się powyżej krajowej średniej wartości 7,34 cytowań na jeden opublikowany artykuł naukowy z lat 2000-2009.

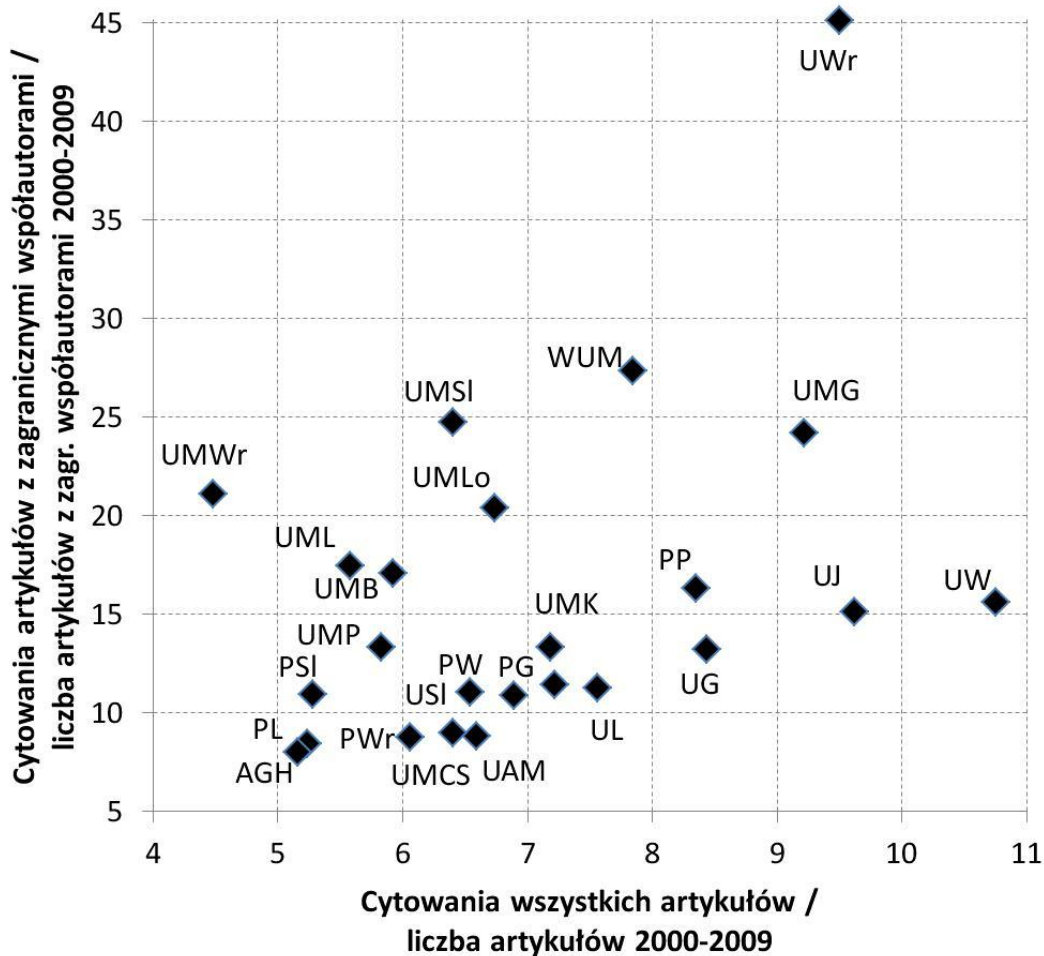
Interesujące może być porównanie stopnia umiędzynarodowienia badań, prowadzonych przez poszczególne uczelnie. Szkoły wyższe w Polsce są aktywnymi uczestnikami międzynarodowych projektów badawczych, a publikacje ze współautorami, reprezentującymi zagraniczne instytucje, dobrze dokumentują przypadki ponadnarodowej współpracy badawczej, także współpracy *ad-hoc*, która może polegać na konsultacjach, wymianie danych empirycznych, aparatury lub porad metodologicznych, czy wizytach studyjnych i stypendialnych. W dalszych analizach mierzono więc stopień umiędzynarodowienia uczelni jako udział w ogóle artykułów tych publikacji, które zostały stworzone wspólnie z reprezentantami ośrodków badawczych z innych krajów. Proponowany wskaźnik odzwierciedla wymierne rezultaty prac badawczych – dlatego wydaje się lepszy od stosowanych czasami miar odnoszących się do „wejść” procesu badawczego takich jak: liczba wspólnych projektów badawczych, liczba wizytujących badaczy, nakłady na projekty międzynarodowe.

Uczelnie	Liczba artykułów ze współautorami z zagranicznych instytucji	Udział artykułów ze współautorami z zagranicznych instytucji we wszystkich artykułach uczelni	Liczba cytowań artykułów z zagranicznymi współautorami / liczba artykułów z zagranicznymi współautorami
Uniwersytet Warszawski	4.332	44,50%	15,59
Uniwersytet Jagielloński, Kraków	4.964	39,02%	15,12
Politechnika Warszawska	1.939	31,90%	10,90
Politechnika Wrocławska	1.739	30,03%	8,76
Uniwersytet im. Adama Mickiewicza, Poznań	2.320	35,89%	8,81
Uniwersytet Wrocławski	516	10,59%	45,11
Warszawski Uniwersytet Medyczny	789	15,80%	27,33
Uniwersytet Medyczny w Łodzi	574	11,37%	20,38
Politechnika Łódzka	1.049	25,48%	8,42
Śląski Uniwersytet Medyczny	564	12,73%	24,76
Uniwersytet Śląski	1.435	32,99%	11,02
Uniwersytet Medyczny w Poznaniu	710	18,03%	13,35
Uniwersytet im. Marii Curie-Skłodowskiej, Lublin	962	28,15%	9,00
Akademia Górniczo-Hutnicza, Kraków	2.152	36,82%	7,98
Gdański Uniwersytet Medyczny	902	23,83%	24,23
Uniwersytet Medyczny w Lublinie	426	11,08%	17,47
Politechnika Gdańska	889	27,94%	11,44
Uniwersytet Łódzki	1.002	31,00%	11,28
Akademia Medyczna, Wrocław	516	10,59%	21,10
Uniwersytet im. Mikołaja Kopernika, Toruń	1.750	31,35%	13,30
Uniwersytet Medyczny w Białymstoku	369	10,05%	17,05
Politechnika Śląska	623	21,55%	10,92
Uniwersytet Gdański	1.194	37,12%	13,20
Politechnika Poznańska	586	24,96%	16,34
Łącznie dla 24 uczelni	32.302	27,28%	14,03

Tabela 70: Artykuły z 24 uczelni w Polsce, pisane ze współautorami, reprezentującymi zagranicznej instytucje, 2000-2009. Źródło danych: baza Scopus.

Tabela 70 przedstawia stopień umiędzynarodowienia poszczególnych uczelni. W rankingu przoduje Uniwersytet Warszawski, na którym aż 44,50% publikacji z lat 2000-2009 powstało we współpracy z zagranicznymi badaczami. Drugie miejsce zajmuje Uniwersytet Jagielloński (39,02%). Do uczelni z umiędzynarodowionymi badaniami należą też: Uniwersytet Gdański (37,12%), Akademia Górniczo-Hutnicza w Krakowie (36,82%), Uniwersytet im. Adama Mickiewicza w Poznaniu (35,89%), Uniwersytet Śląski (32,99%), Politechnika Warszawska (31,90%), Uniwersytet im. Mikołaja Kopernika w Toruniu (31,55%) i Uniwersytet Łódzki (31,00%).

Najmniej zorientowane międzynarodowo uczelnie wśród analizowanej grupy 24 szkół wyższych to: Uniwersytet Medyczny w Białymstoku (10,05%), Uniwersytet Wrocławski (10,59%), Uniwersytet Medyczny w Lublinie (11,08%) i Uniwersytet Medyczny w Łodzi (11,37%). Uzyskane wyniki są zastanawiające, gdyż badania w obszarze medycyny należą zwykle do najbardziej umiędzynarodowionych i prowadzonych przez największe zespoły badawcze. Ciekawa jest również konstatacja dotycząca Uniwersytetu Wrocławskiego, który należy do najbardziej produktywnych i najczęściej cytowanych uczelni w kraju, a jednak znalazł się w grupie uczelni z najmniej umiędzynarodowionymi badaniami naukowymi. Jedynie 10,59% artykułów badacze z Uniwersytetu Wrocławskiego napisali z zagranicznymi współautorami – ale ten podzbiór artykułów był szczególnie często cytowany, otrzymując rekordowy wynik średnio 45,11 cytowań przypadających na każdy artykuł.



Rysunek 134: Wskaźniki cytowań dla wszystkich artykułów 24 polskich uczelni i dla artykułów z zagranicznymi współautorami, 2000-2009. Źródło danych: baza *Scopus*.

Warto zastanowić się, czy współpraca międzynarodowa może wpływać na podniesienie międzynarodowej rangi prowadzonych badań i publikowanych artykułów. W celu weryfikacji tej zależności, porównano wskaźniki, zbudowane jako iloraz liczby cytowań uzyskanych przez zbiór publikacji i liczebności tego zbioru dla wszystkich artykułów pochodzących z danej uczelni oraz dla artykułów z danej uczelni, które powstały w ramach współpracy międzynarodowej. Graficzne zestawienie tych wskaźników oferuje Rysunek 134.

Okazało się, że współpraca międzynarodowa była szczególnie korzystna dla niektórych szkół wyższych, znacząco podnosząc cytawalność artykułów. Najlepiej korzystają z międzynarodowej współpracy badawczej: Uniwersytet Wrocławski (wskaźnik 45,11 dla artykułów z cudzoziemcami wobec 9,50 dla wszystkich artykułów), Warszawski Uniwersytet Medyczny (27,33 i 7,85), Śląski Uniwersytet Medyczny (24,76 i 6,40), Uniwersytet Medyczny w Gdańsku (24,23 i 9,22), Akademia Medyczna we Wrocławiu (21,10 i 4,48) oraz Uniwersytet Medyczny w Łodzi (20,38 i 6,74). Nawet uczelnie o relatywnie niewielkim stopniu umiędzynarodowienia badań naukowych skorzystały z pozytywnego wpływu współpracy ponadnarodowej – przykładowo, Uniwersytet Medyczny w Lublinie otrzymywał średnio 11,90 cytowań do publikacji „międzynarodowych”, a Uniwersytet Medyczny w Białymstoku – 11,13 cytowań. Największe uniwersytety w kraju nie doświadczały tak

dramatycznego wzrostu liczby cytowań przypadających na publikację w przypadku porównywania artykułów tworzonych z cudzoziemcami i całego dorobku publikacyjnego uczelni, co może wynikać z wysokiej rangi naukowej badań, które są prowadzone również we własnym zakresie przez polskie zespoły badawcze.

Niezależnie od tego, każda z 24 szkół wyższych odczuwała w analizowanym okresie lat 2000-2009 pozytywny wpływ współpracy międzynarodowej na wskaźniki cytowań swoich publikacji, a przeprowadzenie odpowiednich testów statystycznych pozwala na zweryfikowanie istnienia statystycznie istotnych różnic między wynikami dla wszystkich publikacji i publikacji umiędzynarodowionych.

15. Analiza bibliometryczna badań nad laserami półprzewodnikowymi, 1988-2005

Rozwój lasery półprzewodnikowych, w szczególności laserów światła niebieskiego, wydawał się być obiecującym obszarem współpracy nauki i biznesu w Polsce. Aktywność firmy TopGaN, spółki odpryskowej Polskiej Akademii Nauk, stwarzała nadzieje na zdobycie przez Polaków znaczącego udziału w globalnym rynku optoelektronicznym. Bibliometria pozwala na lepsze zrozumienie uwarunkowań prac badawczo-rozwojowych w analizowanym obszarze, a poniższy przykład opiera się na analizach, zaprezentowanych w książce Klincewicza (2010).

Analizy opierają się na danych pobranych wiosną 2007 roku z bazy *INSPEC* i obejmujących okres od pierwszego wystąpienia publikacji na analizowany temat do końca roku 2005. Przy wyszukiwaniu odpowiednich publikacji skorzystano z taksonomii bazy *INSPEC* czyli tzw. *classification codes*, wybierając podzbiór publikacji dotyczących laserów półprzewodnikowych.

Liczba publikacji i wystąpień konferencyjnych, dotyczących laserów półprzewodnikowych, systematycznie przyrastała na przestrzeni lat, co prezentuje Tabela 71.

Rok	Liczba publikacji
1988	25
1989	665
1990	1.547
1991	1.617
1992	1.624
1993	1.861
1994	2.238
1995	2.447
1996	2.991
1997	2.596
1998	3.036
1999	2.705
2000	2.911
2001	2.494
2002	2.811
2003	2.975
2004	3.087
2005	2.709
Łącznie	40.339

Tabela 71: Liczba publikacji związanych z laserami półprzewodnikowymi na przestrzeni lat 1988–2006. Źródło danych: baza *INSPEC*.

Wzrost nie przejawiał nasycenia ani spadku dynamiki i mógł zostać opisany funkcją regresji, która pozwala na prognozowanie dalszego, ilościowego rozwoju badań:

$$\ln(C) = 1,3548 \times \ln(Y) + 2,9435$$

gdzie:

C – skumulowana liczba artykułów,

Y – kolejny numer roku (seria 17 punktów danych, rok 1989 = 1, rok 2005 = 17).

Uzyskana funkcja była dobrze dopasowana do danych empirycznych ze współczynnikiem determinacji R^2 na poziomie 0,92.

Kraj	Liczba publikacji
Stany Zjednoczone	11.330
Japonia	6.660
Niemcy	3.249
Wielka Brytania	2.814
Chiny	2.433
Francja	1.785
Rosja	1.325
Korea Południowa	1.191
Kanada	861
Włochy	732
Tajwan	728
Szwajcaria	658
Holandia	485
Szwecja	442
Australia	374
Hiszpania	373
Polska	363
Belgia	330
Dania	332
Irlandia	266
Singapur	251
Finlandia	242
Izrael	229
Białoruś	161
Indie	153
Austria	149
Grecja	121
Ukraina	115
Brazylia	109
Portugalia	98
Turcja	89
Meksyk	55
Czechy	55

Tabela 72: Liczba publikacji związanych z laserami półprzewodnikowymi z poszczególnych krajów na przestrzeni lat 1988–2005. Źródło danych: baza *INSPEC*.

Interesujące było zróżnicowanie regionalne badań, co odzwierciedla Tabela 72. Dominowały Stany Zjednoczone (11.300 publikacji), Japonia (6.600 publikacji) i Niemcy (3.249 publikacji), a Polska była na 17 miejscu rankingu międzynarodowego z jedynie 363 publikacjami. Aktywność poszczególnych krajów można porównywać nie tylko pod względem całkowitej liczby publikacji, ale też dynamiki przyrostu liczby publikacji na przestrzeni lat (por. Tabela 73). Pod względem średniego rocznego przyrostu przodowały Tajwan (średnio 160,61% rocznie), Holandia (158,60%) i Korea Południowa (154,80%), a w tym rankingu Polska znalazła się na 7 miejscu (125,42%), co wskazywało na nadrobienie zaległości przez polskich badaczy.

Kraj	Średni roczny przyrost liczby publikacji 1989–2005
Tajwan	160,61%
Holandia	158,60%
Korea Południowa	154,80%
Australia	145,21%
Niemcy	128,51%
Chiny	126,49%
Polska	125,42%
Kanada	124,14%
Szwajcaria	122,62%
Hiszpania	120,81%
Włochy	119,99%
Francja	118,37%
Szwecja	117,97%
Wielka Brytania	115,72%
Stany Zjednoczone	110,44%
Japonia	109,68%
Rosja	-88,84%

Tabela 73: Średni roczny przyrost liczby publikacji związanych z laserami półprzewodnikowymi z poszczególnych krajów na przestrzeni lat 1989–2005. Źródło danych: baza INSPEC.

Obliczenie indeksów *RTA* wskazuje na istnienie lub brak względnych przewag konkurencyjnych poszczególnych krajów (por. podrozdział 9.7). Wartość *RTA* powyżej 1,0 oznacza konkurencyjność obszaru badań lub technologii. Kalkulacje oparte na dorobku publikacyjnym w obszarze badań nad laserami półprzewodnikowymi prezentuje Tabela 74. Konkurencyjną pozycję mają m.in. Niemcy (1,31), Wielka Brytania (1,23), Japonia (1,73), Szwajcaria (1,51), Korea Południowa (1,50), Francja (1,05), Tajwan (1,20) i Stany Zjednoczone (1,11). Polska osiąga *RTA* na poziomie 0,73, co świadczy o braku przewagi konkurencyjnej. Gdyby jednak przeanalizować zmiany indeksu *RTA* na przestrzeni lat (Tabela 75) okaże się, że polska konkurencyjność systematycznie wzrastała, od poziomu indeksu 0,47 (okres 1989-1994), poprzez 0,62 (1995-2000) do 0,94 (2001-2005), podczas gdy wartości *RTA* dla krajów takich jak Stany Zjednoczone czy Japonia w tym samym okresie stopniowo zmniejszały się.

Kraj	Indeks RTA
Irlandia	3,27
Białoruś	2,28
Dania	2,01
Niemcy	1,31
Wielka Brytania	1,23
Japonia	1,73
Szwajcaria	1,51
Korea Południowa	1,50
Francja	1,05
Holandia	0,90
Tajwan	1,20
Szwecja	1,17
Finlandia	1,13
Belgia	1,12
Stany Zjednoczone	1,11
Singapur	1,05
Rosja	0,84
Kanada	0,80
Chiny	0,79
Polska	0,73
Australia	0,63
Hiszpania	0,61
Włochy	0,61

Tabela 74: Indeksy RTA w obszarze badań nad laserami półprzewodnikowymi dla poszczególnych krajów. Źródło danych: baza *INSPEC*.

Kraj	Indeks RTA 1989–1994	Indeks RTA 1995–2000	Indeks RTA 2001–2005
Stany Zjednoczone	1,20	1,15	1,03
Japonia	2,25	1,66	1,44
Niemcy	0,80	1,23	1,63
Wielka Brytania	1,11	1,11	1,52
Chiny	0,87	0,93	0,70
Francja	1,04	1,02	1,08
Korea Południowa	1,03	1,44	1,48
Rosja	0,70	0,84	0,98
Tajwan	0,96	0,91	1,41
Szwajcaria	1,08	1,36	1,98
Polska	0,47	0,62	0,94

Tabela 75: Zmiany indeksów RTA w obszarze badań nad laserami półprzewodnikowymi dla poszczególnych krajów na przestrzeni lat. Źródło danych: baza *INSPEC*.

Baza *INSPEC* oferuje możliwości bardziej szczegółowych analiz metadanych, dotyczących publikacji. Przykładowo, możliwe jest wyodrębnienie publikacji o charakterze badań podstawowych i stosowanych (w oparciu o przypisanie ich do odpowiedniej grupy przez redaktorów bazy *INSPEC*), przy czym publikacja może być jednocześnie przypisana do obu

rodzajów badań (por. Tabela 76). W Stanach Zjednoczonych, jedynie 28,56% publikacji dotyczyło badań podstawowych, a 87,32% publikacji reprezentowało badania stosowane. W Japonii ta relacja była jeszcze bardziej jednoznaczna, gdyż tylko 20,51% tekstów wchodziło w obszar badań podstawowych, a aż 92,62% z nich miało charakter stosowany. Spośród polskich publikacji dotyczących laserów półprzewodnikowych, aż 52,45% było z kolei poświęconych badaniom podstawowym, co wskazuje na ograniczony potencjał komercjalizacji wyników tych badań. 68,73% polskich publikacji można było określić jako reprezentujące badania stosowane.

Kraj	%badania podstawowe	%badania stosowane
Stany Zjednoczone	28,56%	87,32%
Japonia	20,51%	92,62%
Niemcy	25,39%	89,57%
Wielka Brytania	37,97%	80,12%
Chiny	40,23%	79,85%
Francja	28,73%	87,31%
Korea Południowa	28,67%	88,06%
Rosja	38,37%	79,74%
Kanada	45,65%	74,11%
Tajwan	23,30%	91,94%
Włochy	45,25%	79,16%
Szwajcaria	25,92%	90,13%
Holandia	43,98%	79,12%
Szwecja	27,43%	87,39%
Hiszpania	68,45%	57,00%
Australia	43,81%	79,90%
Polska	52,45%	68,73%

Tabela 76: Publikacje związane z laserami półprzewodnikowymi, reprezentujące badania podstawowe i stosowane w poszczególnych krajach na przestrzeni lat 1989–2005 (grupy nie są rozłączne, udziały nie sumują się do 100%). Źródło danych: baza *INSPEC*.

Indeksy bazy *INSPEC* pozwalają również na analizy szczegółowych zagadnień i tematów badawczych, których dotyczą publikacje. Autorowi udało się przykładowo zidentyfikować kody klasyfikacyjne, odpowiadające różnym rodzajom materiałów półprzewodnikowych, wykorzystywanych do wytwarzania laserów. Pomogła w tym analiza indeksu związków chemicznych (ang. *Chemical Index*) i studia literaturowe, dzięki którym podzielono związki na grupy, odpowiadające trzem głównym rodzajom podłoży półprzewodnikowych – związkom odpowiednich grup układu okresowego (tzw. związki II-VI, związki III-V i związki IV-VI). Identyfikacja odpowiednich związków i ich przypisanie do grup jest czasochłonne i wymaga specjalistycznej wiedzy, nie byłoby też możliwe bez wykorzystania dostępnego w bazie *INSPEC* indeksu chemicznego. W celu weryfikacji poprawności klasyfikacji, po dokonaniu podziału zweryfikowano jego rozłączność, obliczając współczynniki korelacji współwystępowania poszczególnych rodzajów związków w analizowanych publikacjach, a wszystkie obliczane współczynniki były na bardzo niskim poziomie (ok. 0,01). Trzy grupy związków wyznaczały kierunki badań nad laserami, a ostatecznie tylko związki III-V sprawdziły się w praktyce wytwórczej. Porównanie

aktywności badawczej prezentuje Tabela 77. Niektóre kraje dużo uwagi poświęcały „ślepych zaułkom” badań, koncentrując uwagę na związkach IV-VI (Stany Zjednoczone: 5,25% publikacji, Japonia: 5,86%, Rosja: 5,06%, Tajwan: 5,16%, Szwajcaria: 5,74%, Szwecja: 6,19%) lub związkach II-VI (Stany Zjednoczone: 4,27%, Japonia: 5,37%, Niemcy: 7,27%, Rosja: 7,85%). Polsce udało się skoncentrować na najbardziej obiecującym kierunku badań, bo aż 51,42% publikacji dotyczyło związków III-V, przy mniejszym nacisku na związki II-VI (3,10%) i związki IV-VI (4,65%).

Kraj	% publikacji dotyczących związków II-VI	% publikacji dotyczących związków III-V	% publikacji dotyczących związków IV-VI
Stany Zjednoczone	4,27%	44,24%	5,25%
Japonia	5,37%	48,63%	5,86%
Niemcy	7,27%	49,39%	4,53%
Wielka Brytania	1,71%	38,62%	3,46%
Chiny	2,39%	30,12%	2,16%
Francja	4,52%	38,84%	3,34%
Korea Południowa	2,36%	41,75%	3,73%
Rosja	7,85%	51,28%	5,06%
Kanada	0,67%	39,29%	2,68%
Tajwan	2,02%	42,19%	5,16%
Włochy	4,49%	24,27%	3,03%
Szwajcaria	3,09%	44,92%	5,74%
Holandia	0,80%	28,51%	4,82%
Szwecja	1,99%	50,22%	6,19%
Hiszpania	0,76%	16,79%	0,00%
Australia	2,06%	19,33%	1,80%
Polska	3,10%	51,42%	4,65%

Tabela 77: Publikacje związane z laserami półprzewodnikowymi dotyczące poszczególnych rodzajów związków chemicznych dla wybranych krajów na przestrzeni lat 1989–2005. Źródło danych: baza INSPEC.

Analizy zawartości indeksów, słów kluczowych, tytułów i streszczeń publikacji pozwoliły na wyodrębnienie szczegółowych technologii produkcyjnych, wykorzystywanych do wytwarzania laserów. Do najpopularniejszych należały *MBE* (ang. *molecular beam epitaxy*) i *MOCVD* (ang. *metal-organic vapour phase epitaxy*). Przegląd listy dziesiątków tysięcy słów kluczowych i pojęć, występujących w tytułach i streszczeniach artykułów pomógł zidentyfikować terminy najlepiej opisujące *MBE* i *MOCVD*. Współczynnik korelacji, odzwierciedlający prawdopodobieństwo jednoczesnego przypisania publikacji do obu obszarów technologicznych, był na poziomie 0,04, co świadczyło o rozłączności podziału. Tylko nieliczne publikacje dotyczyły omawianych metod – *MBE* było dyskutowane w 3,6% publikacji, a *MOCVD* – w 4,24% publikacji. Możliwa była jednak identyfikacja znaczących różnic regionalnych, gdy badacze z danego kraju szczególnie interesowali się jedną z technologii. Przykładowo, naukowcy z Japonii i Korei Południowej stawiali na *MOCVD* (odpowiednio 7,44% i 6,08% publikacji), podczas gdy Niemcy i Polacy przewodzili pod względem koncentracji na *MBE* (odpowiednio 6,32% i 6,20% publikacji).

Kraj	%publikacji dotyczących MOCVD	%publikacji dotyczących MBE
Stany Zjednoczone	3,52%	3,38%
Japonia	7,44%	3,55%
Niemcy	4,77%	6,32%
Wielka Brytania	2,81%	2,67%
Chiny	4,48%	3,78%
Francja	3,82%	3,98%
Korea Południowa	6,08%	1,67%
Rosja	3,85%	4,14%
Kanada	1,23%	3,13%
Tajwan	4,53%	3,40%
Włochy	2,24%	1,58%
Szwajcaria	1,91%	2,65%
Holandia	2,01%	0,60%
Szwecja	4,65%	2,88%
Hiszpania	0,51%	2,29%
Australia	2,58%	2,58%
Polska	3,10%	6,20%

Tabela 78: Publikacje dotyczące poszczególnych rodzajów technologii produkcyjnych dla wybranych krajów. Źródło danych: baza *INSPEC*.

Warto wreszcie postawić pytanie, jakiego rodzaju instytucje angażowały się w prowadzenie badań nad laserami półprzewodnikowymi. W Stanach Zjednoczonych 63,95% publikacji pochodziło z instytucji publicznych, a 36,05% - od firm prywatnych. Zbliżone proporcje dotyczyły Niemiec (82,07% - 17,93%), Wielkiej Brytanii (75,96% - 24,04%), Francji (68,27% - 31,73%) i Korei Południowej (80,65% - 19,35%). W przodującej pod względem sprzedaży laserów Japonii dominowały z kolei badania prywatne, odpowiadające aż za 61,90% publikacji tego kraju, przy jedynie 38,10% publikacji z uczelni i publicznych instytutów badawczych. Polskie badania nad laserami półprzewodnikowymi były zdominowane przez podmioty sektora publicznego (97,39% publikacji) z bardzo ograniczoną aktywnością prywatną (2,61%), co wskazuje na dodatkowe bariery komercjalizacji wyników badań.

Kraj	% uczelnie i publiczne instytuty badawcze	% firmy prywatne
Stany Zjednoczone	63,95%	36,05%
Japonia	38,10%	61,90%
Niemcy	82,07%	17,93%
Wielka Brytania	75,96%	24,04%
Chiny	99,05%	0,95%
Francja	68,27%	31,73%
Korea Południowa	80,65%	19,35%
Rosja	95,68%	4,33%
Kanada	79,42%	20,57%
Tajwan	97,67%	2,33%
Włochy	90,99%	9,01%
Szwajcaria	76,68%	23,32%
Holandia	57,69%	42,31%
Szwecja	83,01%	16,99%
Hiszpania	98,64%	1,36%
Australia	92,27%	7,73%
Polska	97,39%	2,61%

Tabela 79: Odsetek publikacji dotyczących laserów półprzewodnikowych, autorstwa pracowników instytucji publicznych i prywatnych na przestrzeni lat 1989–2005. Źródło danych: baza *INSPEC*.

W sposób analogiczny do analizy krajów, przeprowadzono też analizy aktywności badawczej i publikacyjnej najważniejszych firm i uczelni, odzwierciedlając:

- kraje, w których dana firma prowadziła badania nad laserami półprzewodnikowymi,
- okresy największej aktywności i ewentualne momenty rezygnacji z dalszych badań,
- koncentrację na związkach II-VI, III-V i IV-VI,
- koncentrację na technologiach *MBE* i *MOCVD*,
- najważniejszych badaczy, pracujących dla poszczególnych firm.

Wyniki analizy publikacji zestawiono z przeprowadzoną w analogiczny sposób analizą patentów, opartą o zasoby *USPTO*. Szczegóły projektu badawczego opisane są w książce Klincewicz (2010).

Literatura

Alterovitz, G., Ramoni, M. (red.) (2010) Knowledge-based bioinformatics. From analysis to interpretation. Chichester: John Wiley & Sons

Altszuller, H. (1972) Algorytm wynalazku. Warszawa: Wiedza Powszechna

Archambault, E., Campbell, D., Gingras, Y., Lariviere, V. (2009) Comparing bibliometric statistics obtained from the Web of Science and Scopus. *Journal of the American Society for Information Science and Technology*, 60, 7, s. 1320-1326

Arman, H., Foden, J. (2010) Combining methods in the technology intelligence process: application in an aerospace manufacturing firm. *R&D Management*, 40, 2, s. 181-194

Banville, D.L. (2009) Chemical information mining: Facilitating Literature-Based Discovery. Boca Raton, FL: CRC Press

Bar-Ilan, J., Levene, M., Lin, A. (2007) Some measures for comparing citation databases. *Journal of Informetrics*, 1, s. 26-34

Bell G., Callon, M. (1994) Techno-economic networks and science and technology policy. *OECD STI Review*, 14, s. 59-118

Bengisu, M., Nekhili, R. (2006) Forecasting emerging technologies with the aid of science and technology databases. *Technological Forecasting & Social Change*, 73, s. 835-844

Börner, K., Chen, Ch., Boyack, K.W. (2003) Visualizing knowledge domains. *Annual Review of Information Science & Technology*, 37, s. 179-255

Boyack, K.W., Klavans, R. (2010) Co-citation analysis, bibliographic coupling, and direct citation: which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology*, 61, 12, s. 2389-2404

Boyack, K.W., Klavans, R. (2010) Co-citation analysis, bibliographic coupling, and direct citation: which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology*, 61, 12, s. 2389-2404

Boyack, K.W., Wylie, B.N., Davidson, G.S. (2002) Domain visualization using VxInsight for science and technology management. *Journal of the American Society for Information Science and Technology*, 53, 9, s. 764-774

Brenner, M. (2005) Technology intelligence at air products. Leveraging analysis and collection techniques. *Competitive Intelligence Magazine*, 8, 3, s. 6-19

Breschi, S., Lissoni, F. (2004) Knowledge networks from patent data. Methodological issues and research targets. w: Moed, H.F. i in. (red.), *Handbook*

of Quantitative Science and Technology Research, Amsterdam: Kluwer Academic Publishers, s. 613-643

Brusoni, S., Geuna, A. (2004) Specialisation and integration. Combining patents and publications data to map the 'structure' of specialised knowledge. w: Moed, H.F. i in. (red.), Handbook of Quantitative Science and Technology Research, Amsterdam: Kluwer Academic Publishers, s. 733-758

Calcagno, M. (2008) An investigation into analysing patents by chemical structure using Thomson's Derwent World Patent Index codes. World Patent Information, 30, s. 188-198

Callan, A., Hockema, S., Eysenbach, G. (2010) Contextual cocitation: augmenting cocitation analysis and its applications. Journal of the American Society for Information Science and Technology, 61, 6, s. 1130-1143

Callon, M., Courtial, J.P., Laville, F. (1991) Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. Scientometrics, 22, 1, s. 155-205

CAS (2011a). CAS Databases. www.cas.org/expertise/cascontent/index.html

CAS (2011b). CAS Databases. Registry. www.cas.org/expertise/cascontent/registry/index.html

CAS (2011c). CAS Databases. Registry. CAS Registry and CASRNs. www.cas.org/expertise/cascontent/registry/regsys.html

CAS (2011d). SciFinder. How To Guides. www.cas.org/support/scifi/htguides.html

Cavallucci, D., Rousselot, F., Zanni, C. (2009) Linking contradictions and laws of engineering system evolution within the TRIZ framework. Creativity and Innovation Management, 18, 2, s. 71-80

Chen, Ch. (2006a) CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. Journal of the American Society for Information Science and Technology, 57, 3, s. 359-377

Chen, Ch. (2006b) Information visualization. Beyond the horizon. London: Springer

Chen, Ch. (2009) Visualizing and analyzing scientific literature with CiteSpace. Workshop presentation, part 4. http://cluster.cis.drexel.edu/~cchen/citespace/doc/tutorial/4_working_with_citespace.pdf

Chen, K., Guan, J. (2011) A bibliometric investigation of research performance in emerging nanobiopharmaceuticals. Journal of Informetrics, 5, s. 233-247

- Chen, Ch., Song, I.-Y., Yuan, X., Zhang, J. (2008) The thematic and citation landscape of 'Data and Knowledge Engineering' (1985-2007). *Data & Knowledge Engineering*, 67, s. 234-259
- Chen, Ch., Zhang, J., Vogeley, M.S. (2010) Making sense of the evolution of a scientific domain: a visual study of the Sloan Digital Sky Survey research. *Scientometrics*, 83, s. 669-688
- Christensen, C.M. (2000) *The innovator's dilemma*. New York: Harper Business
- Clarke, D.W. Sr. (2000) Strategically evolving the future: Directed evolution and technological systems development. *Technological Forecasting & Social Change*, 64, s. 133-153
- Claverie, J.-M., Notredame, C. (2007) *Bioinformatics for dummies*. Hoboken, NJ: Wiley Publishing
- Coates, V., Farooque, M., Klavans, R., Lapid, K., Linstone, H.A., Pistorius, C., Porter, A.L. (2001) On the future of technological forecasting. *Technological Forecasting & Social Change*, 67, s. 1-17
- Cockburn, I.M., Henderson, R.M. (1998) Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. *The Journal of Industrial Economics*, 46, 2, s. 157-182
- Cohen, W.D., Levinthal, D.A. (1990) Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35, s. 128-152
- Cohen, W.M., Nelson, R.R., Walsh, J.P. (2002) Links and impacts: The influence of public research on industrial R&D. *Management Science*, 48, 1, s. 1-23
- Coleman, A. (2007) Assessing the value of a journal beyond the impact factor. *Journal of the American Society for Information Science and Technology*, 58, 8, s. 1148-1161
- Cong, H., Tong, L.H. (2005) Similarity between TRIZ principles. *The TRIZ Journal*, 9, s. 1-5
- Cong, H., Tong, L.H. (2007) Patent classification inspires new innovation. *Innovation*, 7, 3, s. 62-64
- Cong, H., Tong, L.H. (2008) Grouping of TRIZ inventive principles to facilitate automatic patent classification. *Expert System with Applications*, 34, s. 788-795
- Coombs, R., Saviotti, P., Walsh, V. (1987) *Economics and technological change*. Totowa, NJ: Rowman & Littlefield
- Courtial, J.P., Callon, M., Sigogneau, A. (1993) The use of patent titles for identifying the topics of invention and forecasting trends. *Scientometrics*, 26, 2, s. 231-242
- Courtial, J.P. (1994) A cword analysis of scientometrics. *Scientometrics*, 31, 3, s. 251-260

D'Angelo, C.A., Giuffrida, C., Abramo, G. (2011) A heuristic approach to author name disambiguation in bibliometric databases for large-scale research assessments. *Journal of the American Society for Information Science and Technology*, 62, 2, s. 257-269

Delanghe, H., Sloan, B., Muldur, U. (2010) European research policy and bibliometric indicators, 1990-2005. *Scientometrics*, w druku, www.springerlink.com/content/a27r064003g21564/

Devoino, I.G., Koshevoy, O.E., Litvin, S.S., Tsourikov, V., Skuratovich, A. (2001) Computer based system for imaging and analyzing a process system and indicating values of specific design changes. United States Patent No. 6.202.043-B1

Ding, Y., Yan, E., Frazho, A., Caverlee, J. (2009) PageRank for ranking authors in co-citation networks. *Journal of the American Society for Information Science and Technology*, 60, 11, s. 2229-2243

EC (2011) ICT - Information and Communication Technologies. Work Programme 2011-12. A Theme for research and development under the specific programme "Cooperation" implementing the Seventh Framework Programme (2007-2013) of the European Community for research, technological development and demonstration activities. ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/ict-wp-2011-12_en.pdf

Eck, N.J. van, Waltman, L. (2009) How to normalize cooccurrence data? An analysis of some well-known similarity measures. *Journal of the American Society for Information Science and Technology*, 60, 8, s. 1635-1651

Eck, N.J. van, Waltman, L., Dekker, R., Berg, J. van den (2010) A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. *Journal of the American Society for Information Science and Technology*, 61, 12, s. 2405-2416

EI (2011a). Compendex & Ei Backfile. www.ei.org/compendex

EI (2011b). Compendex. www.ei.org/userfiles/ProductDescriptions/Compendex_LoRes.pdf

Espacenet (2011a). Wyszukaj za pomocą klasyfikacji europejskiej. http://v3.espacenet.com/eclsrch?locale=pl_PL&classification=ecla

Espacenet PL (2011a). Wyszukiwanie zaawansowane. http://pl.espacenet.com/advancedSearch?locale=pl_PL

Espacenet PL (2011b). Ograniczenia. http://pl.espacenet.com/help?topic=limitations&locale=pl_PL&method=handleHelpTopic

Espacenet PL (2011c). Dane bibliograficzne.
http://pl.espacenet.com/help?locale=pl_PL&method=handleHelpTopic&topic=bibliographic

Espacenet PL (2011d). Oryginał dokumentu.
http://pl.espacenet.com/help?locale=pl_PL&method=handleHelpTopic&topic=original

Eurostat (2010) Europa in figures. Eurostat yearbook 2010. Luxembourg: Publication Office of the European Union,
http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-CD-10-220/EN/KS-CD-10-220-EN.pdf

Fattori, M., Pedrazzi, G., Turra, R. (2003) Text mining applied to patent mapping: a practical business case. *World Patent Information*, 25, s. 335-342

Fischer, G., Lalyre, N. (2006) Analysis and visualisation with host-based software - The features of STN® AnaVist™. *World Patent Information*, 28, s. 312-318

Frey, P.R., Minsk, B.S., Porter, A.L. (2005) An integrated suite of text and data mining tools – phase II. Scientific and Technical Report STI-TR-948-02-F

Ganiz, M.C., Pottenger, W.M., Janneck, Ch.D. (2005) Recent advances in Literature Based Discovery. Lehigh University, Technical Report LU-CSE-05-027

Garcia-Perez, M.A. (2010) Accuracy and completeness of publication and citation records in the Web of Science, PsycINFO, and Google Scholar: A case study for the computation of h indices in psychology. *Journal of the American Society for Information Science and Technology*, 61, 10, s. 2070-2085

Garfield, E. (1955) Citation indexes for science. A new dimension in documentation through association of ideas. *Science*, 122, s. 108-111

Garfield, E. (1972) Citation analysis as a tool in journal evaluation. *Science*, 178, s. 471-479

Garfield, E. (2005) The agony and the ecstasy. The history and meaning of the journal impact factor. International Congress on Peer Review And Biomedical Publication, Chicago. <http://garfield.library.upenn.edu/papers/jifchicago2005.pdf>

Glaser, M., Miecznik, B. (2009) TRIZ for reverse inventing in market research: A case study from WITTENSTEIN AG, identifying new areas of application of a core technology. *Creativity and Innovation Management*, 18, 2, s. 90-100

Google Patents (2011a). About Google Patent Search.
www.google.com/googlepatents/about.html#adv

Google Patents (2011b). Google Patent Search Help.
www.google.com/googlepatents/help.html

Google Scholar (2011a). About Google Scholar.
<http://scholar.google.pl/intl/pl/scholar/about.html>

Grupp, H., Linstone, H.A. (1999) National technology foresight activities around the globe. Resurrection and new paradigms. *Technological Forecasting & Social Change*, 60, s. 85-94

Guo, Y., Huang, L., Porter, A.L. (2010) The research profiling method applied to nano-enhanced, thin-film solar cells. *R&D Management*, 40, 2, s. 195-208

Héraud, J.-A., Cuhls, K. (1999) Current foresight activities in France, Spain, and Italy. *Technological Forecasting & Social Change*, 60, s. 55-70

Hicks, D. (1987) Limitations of co-citation analysis as a tool for science policy. *Social Studies of Science*, 17, 2, s. 295-316

Hicks, D., Kroll, P., Narin, F., Thomas, P., Ruegg, R., Tomizawa, H., Saitoh, Y., Kobayashi, S. (2002) Quantitative methods of research evaluation used by the U.S. Federal Government. National Institute of Science and Technology Policy, NISTEP Study Material No. 86, Tokyo, Japan,
www.nistep.go.jp/achiev/ftx/eng/mat086e/pdf/mat086e.pdf

Hinze, S., Schmoch, U. (2004) Opening the black box. Analytical approaches and their impact on the outcome of statistical patent analyses. w: Moed, H.F. i in. (red.), *Handbook of Quantitative Science and Technology Research*, Amsterdam: Kluwer Academic Publishers, s. 215-235

Hirsch, J.E. (2005) An index to quantify an individual's scientific research output. *PNAS*, 102, 46, s. 16569-16572

Holtmannspötter, D., Rijkers-Defrasne, S., Ploetz, Ch., Thaller-Honold, S., Zweck, A. (2010) *Technologieprognosen. Internationaler vergleich 2010*. Düsseldorf: Zukünftige Technologien Consulting, Bundesministerium für Bildung und Forschung,
www.bmbf.de/pubRD/technologieprognosen_int_vergleich_ZTC_bd_58.pdf

Hsieh, Ch. (2011) Explicitly searching for useful inventions: dynamic relatedness and the costs of connecting versus synthesizing. *Scientometrics*, 86, s. 381-404

IET (2011a). INSPEC direct data. Which product?
www.theiet.org/publishing/inspec/products/range/data.cfm

IET (2011b). INSPEC Archive - Science Abstracts journals from 1898 – 1968.
www.theiet.org/publishing/inspec/products/range/archive.cfm

IET (2011c). INSPEC pay-as-you-go searching.
www.theiet.org/publishing/inspec/products/range/pay-go.cfm

IET (2011d). INSPEC print products.
www.theiet.org/publishing/inspec/products/range/print.cfm

IET (2011e). INSPEC Search Aids.
www.theiet.org/publishing/inspec/products/range/aids.cfm

IET (2011f) INSPEC, The Institute of Engineering and Technology.
www.theiet.org/publishing/inspec/

InspecDirect (2011). Content, coverage and indexing.
<http://inspecdirect.theiet.org/contentcoverage/inspec/index.cfm>

Intellogist (2011a). Compare: Patent Search System – Esp@cenet, Google Patent Search.

www.intellogist.com/wiki/Compare:Patent_Search_System?n=Esp%40cenet/Google+Patent+Search

Intellogist (2011b). Compare: Patent Search System – Google Patent Search.
www.intellogist.com/wiki/Compare:Patent_Search_System?n=Google+Patent+Search

Intellogist (2011c). Compare: Patent Search System – Delphion.
www.intellogist.com/wiki/Compare:Patent_Search_System?n=Delphion

Intellogist (2011d). Compare: Patent Search System – MicroPatent PatentWeb.
www.intellogist.com/wiki/Compare:Patent_Search_System?n=MicroPatent+PatentWeb

Intellogist (2011e). Compare: Patent Search System – PatBase.
www.intellogist.com/wiki/Compare:Patent_Search_System?n=PatBase

Intellogist (2011f). Compare: Patent Search System – QPat.
www.intellogist.com/wiki/Compare:Patent_Search_System?n=QPat

Intellogist (2011g). Compare: Patent Search System – WIPS Global.
www.intellogist.com/wiki/Compare:Patent_Search_System?n=WIPS+Global

Intellogist (2011h). Compare: Patent Search System – TotalPatent.
www.intellogist.com/wiki/Compare:Patent_Search_System?n=TotalPatent

Intellogist (2011i). Compare: Patent Search System – Thomson Innovation.
www.intellogist.com/wiki/Compare:Patent_Search_System?n=Thomson+Innovation

Intellogist (2011j). Compare: Patent Search System – SureChem.
www.intellogist.com/wiki/Compare:Patent_Search_System?n=SureChem

Intellogist (2011k). Compare: Patent Search System – JP-NETe.
www.intellogist.com/wiki/Compare:Patent_Search_System?n=JP-NETe

Jahangirian, M., Eldabi, T., Garg, L., Jun, G.T., Naseer, A., Patel, B., Stergioulas, L., Young, T. (2011) A rapid review method for extremely large corpora of literature: Applications to the domains of modelling, simulation, and management. *International Journal of Information Management*, 31, 3, s. 234-243

- Järvenpää, H.M., Mäkinen, S.J., Seppänen, M. (2011) Patent and publishing activity sequence over a technology's life cycle. *Technological Forecasting & Social Change*, 78, s. 283-293
- Jones, E., Mann, D., Harrison, D., Stanton, N.A. (2001) An eco-innovation case study of domestic dishwashing through the application of TRIZ tools. *Creativity and Innovation Management*, 10, 1, s. 3-14
- Kardas, M. (2010) Inteligentna specjalizacja - (nowa) koncepcja polityki innowacyjnej. Referat na konferencję "Instytucjonalne aspekty rozwoju sektora B+R w Polsce. Od gospodarki imitacyjnej do innowacyjnej", Uniwersytet w Białymstoku, Goniądz, 28-30 września 2010 r.
- Klavans, R., Boyack, K.W. (2006) Quantitative evaluation of large maps of science. *Scientometrics*, 68, 3, s. 475-499
- Klincewicz, K. (2005) *Strategic alliances in the hightech industry*. Berlin: Logos
- Klincewicz, K. (2008) *Polska innowacyjność. Analiza bibliometryczna*. Warszawa: Wydawnictwo Wydziału Zarządzania Uniwersytetu Warszawskiego
- Klincewicz, K. (2010) *Zarządzanie technologiami. Przypadek niebieskiego lasera*. Warszawa: Wydawnictwo Wydziału Zarządzania Uniwersytetu Warszawskiego
- Klincewicz, K., Miyazaki, K. (2011) Sectoral systems of innovation in Asia. The case of software research activities. *International Journal of Technology Management*, 53, 2-4, s. 161-189
- Kodama, F. (1992) Technology fusion and the new R&D. *Harvard Business Review*, 70, 4, s. 70-78
- Kopcsa, A., Schiebel, E. (1998) Science and technology mapping: A new iteration model for representing multidimensional relationships. *Journal of the American Society for Information Science*, 49, 1, s. 7-17
- Kostoff, R.N., Miles, D.L., Eberhart, H.J. (1995) System and method for database tomography. United States Patent No. 5.440.481
- Kostoff, R.N., Toothman, D.R., Eberhart, H.J., Humenik, J.A. (2001) Text mining using database tomography and bibliometrics: A review. *Technological Forecasting & Social Change*, 68, s. 223-253
- Kostoff, R.N. (2002) Citation analysis of research performer quality. *Scientometrics*, 53, 1, s. 49-71
- Kostoff, R.N. (2003) Science and technology text mining. Origins of database tomography and multi-word phrase clustering. Office of Naval Research, Technical Report, www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA416268
- Kostoff, R.N., Geisler, E. (2004) Science and technology text mining strategic management and implementation in government organizations. Office of Naval

Research, Technical Report, www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA421060

Kostoff, R.N. (2005) Method for data and text mining and literature-based discovery. United States Patent No. 6.886.010

Kostoff, R.N., Briggs, M.B., Rushenberg, R.L., Bowles, Ch.A., Pecht, M. (2006) The structure and infrastructure of Chinese science and technology. Office of Naval Research, Technical Report, www.fas.org/irp/world/china/docs/science.pdf

Kostoff, R.N. (2008a) Literature-related discovery (LRD): Introduction and background. *Technological Forecasting & Social Change*, 75, s. 165-185

Kostoff, R.N., Block, J.A., Solka, J.L., Briggs, M.B., Rushenberg, R.L., Stump, J.A., Johnson, D., Lyons, T.J., Wyatt, J.R. (2008b) Literature-related discovery (LRD): Lessons learned, and future research directions. *Technological Forecasting & Social Change*, 75, s. 276-299

Kostoff, R.N., Briggs, M.B., Solka, J.L., Rushenberg, R.L. (2008c) Literature-related discovery (LRD): Methodology. *Technological Forecasting & Social Change*, 75, s. 186-202

Kozłowski, J. (2010) Statystyka nauki, techniki i innowacji w krajach UE i OECD. Stan i problemy rozwoju. Departament Strategii Ministerstwa Nauki i Szkolnictwa Wyższego, www.nauka.gov.pl/fileadmin/user_upload/Nauka/Polityka_naukowa_panstwa/Analizy_raporty_statystyki/20100830_Statystyka_nauki_tehniki_i_innowacji_w_krajach_UE_i_OECD.pdf

Kuwahara, T. (1999) Technology forecasting activities in Japan. *Technological Forecasting & Social Change*, 60, s. 5-14

Leeuwen, T. van (2004) Descriptive versus evaluative bibliometrics. Monitoring and assessing of national R&D systems. w: Moed, H.F. i in. (red.), *Handbook of Quantitative Science and Technology Research*, Amsterdam: Kluwer Academic Publishers, s. 373-388

Lehvo, A., Nuutinen, A. (2006) Finnish science in international comparison. A bibliometric analysis. Publications of the Academy of Finland No. 16/06, www.aka.fi/Tiedostot/Tiedostot/Julkaisut/15_06%20Finnish%20Science%20in%20International%20Comparison.pdf

Lewison, G. (2004) Citations to papers from other documents. Evaluation of the practical effects of biomedical research. w: Moed, H.F. i in. (red.), *Handbook of Quantitative Science and Technology Research*, Amsterdam: Kluwer Academic Publishers, s. 457-472

Leydesdorff, L. (2008) On the normalization and visualization of author co-citation data: Salton's cosine versus the Jaccard index. *Journal of the American Society for Information Science and Technology*, 59, 1, s. 77-85

- Leydesdorff, L. (2008) Patent classifications as indicator of intellectual organization. *Journal of the American Society for Information Science and Technology*, 59, 10, s. 1582-1597
- Leydesdorff, L. (2009) How are new citation-based journal indicators adding to the bibliometric toolbox? *Journal of the American Society for Information Science and Technology*, 60, 7, s. 1327-1336
- Leydesdorff, L., Opthof, T. (2010) Scopus's Source Normalized Impact per Paper (SNIP) versus a Journal Impact Factor based on fractional counting of citations. *Journal of the American Society for Information Science and Technology*, 61, 11, s. 2365-2369
- Leydesdorff, L., Opthof, T. (2011) Scopus' SNIP Indicator: Reply to Moed. *Journal of the American Society for Information Science and Technology*, 62, 1, s. 214-215
- Leydesdorff, L., Rafols, I. (2009) A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 60, 2, s. 348-362
- Librarians ACM (2011). The Digital Library. <http://librarians.acm.org/digital-library>
- Lichtenthaler, Eckhard (2004) Technology intelligence processes in leading European and North American multinationals. *R & D Management*, 34, 2, s. 121-135
- MacRoberts, M.H., MacRoberts, B.R. (1989) Problems of citation analysis: A critical review. *Journal of the American Society for Information Science and Technology*, 40, 5, s. 342-349
- MacRoberts, M.H., MacRoberts, B.R. (2010) Problems of citation analysis: A study of uncited and seldom-cited influences. *Journal of the American Society for Information Science and Technology*, 61, 1, s. 1-13
- Mallig, N. (2010) A relational database for bibliometric analysis. *Journal of Informetrics*, 4, s. 564-580
- Mann, D. (2002) Manufacturing technology evolution trends. *Integrated Manufacturing Systems*, 13, 2, s. 86-90
- Mann, D.L. (2003) Better technology forecasting using systematic innovation methods. *Technological Forecasting & Social Change*, 70, s. 779-795
- Meho, L.I., Yang, K. (2007) Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology*, 58, 13, s. 2105-2125
- Merton, R. K. 1968. The Matthew effect in science. *Science*, 159, s. 56-63

Meyer, M. (2000) What is special about patent citations? Differences between scientific and patent citations. *Scientometrics*, 49, 1, s. 93-123

Micropat (2011). PatentWeb. First Level Analysis. www.micropat.com/static/first_level.htm#alerts

Mićić, Pero (2010) Future markets-radar: A case study of applied strategic foresight. *Technology Forecasting & Social Change*, 77, s. 1499-1505

Miranda Santo, M. de, Coelho, G.M., Santos, D.M. dos, Filho, L.F. (2006) Text mining as a valuable tool in foresight exercises: A study on nanotechnology. *Technological Forecasting & Social Change*, 73, s. 1013-1027

Mishra, S., Deshmukh, S.G., Vrat, P. (2002) Matching of technological forecasting technique to a technology. *Technological Forecasting & Social Change*, 69, s. 1-27

Mizrachi, Y. (2010) Don't predict the future - direct it! Comments of the intellectual history, the logical and applicative visibility, and the underlying assumptions of directed evolution. *World Futures*, 66, s. 26-52

MNiSW (2008) Krajowy Program Badań Naukowych i Prac Rozwojowych. Ministerstwo Nauki i Szkolnictwa Wyższego, Załącznik do Komunikatu nr 22 z dnia 30 października 2008 r., www.bip.nauka.gov.pl/_gALLERY/54/32/5432/Krajowy_Program_Badan_Naukowych_i_Prac_Rozwojowych.pdf

MNiSW (2010) Informacja w sprawie ujednoczonego wykazu ustalonych kategorii jednostek naukowych. Ministerstwo Nauki i Szkolnictwa Wyższego, 15 grudnia 2010 r., www.nauka.gov.pl/fileadmin/user_upload/Finansowanie/finansowanie_nauki/Dzialalnosc_statutowa/20101217_wykaz_kategorii.pdf

Moed, H.F. (2005) Citation analysis in research evaluation. Dordrecht: Springer

Moed, H.F. (2010) Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, 4, s. 265-277

Moehrle, M.G. (2005) How combinations of TRIZ tools are used in companies - results of a cluster analysis. *R&D Management*, 35, 3, s. 285-296

Moehrle, M.G., Lessing, H. (2004) Profiling technological competencies of companies: A case study based on the Theory of Inventive Problem Solving. *Creativity and Innovation Management*, 13, 4, s. 231-238

Mogoutov, A., Cambrosio, A., Keating, P., Mustar, Ph. (2008) Biomedical innovation at the laboratory, clinical and commercial interface: A new method for mapping research projects, publications and patents in the field of microarrays. *Journal of Informetrics*, 2, s. 341-353

Mogoutov, A., Cambrosio, A., Keating, P., Mustar, Ph. (2008) Biomedical innovation at the laboratory, clinical and commercial interface: A new method for

mapping research projects, publications and patents in the field of microarrays. *Journal of Informetrics*, 2, s. 341-353

Mortara, Letizia, Thomson, Ruth, Moore, Chris, Armara, Kalliopi, Kerr, Clive, Phaal, Robert, Probert, David (2010) Developing a technology intelligence strategy at Kodak European Research: scan & target. *Research – Technology Management*, July – August 2010, s. 27-38

Myszka, W. (2010) O pewnych zastosowaniach naukometrii. w: Myszka, W. (red.) Wycinki do prezentacji 'Szybki wywiad technologiczny' - wersja 8. Politechnika Wroclawska, http://kreatywny.wmech.pwr.wroc.pl/wp-uploads/2010/03/qtip_rap.pdf

National Science Board (2010) Science and Engineering Indicators 2010. Arlington, VA: National Science Foundation, www.nsf.gov/statistics/seind10/pdf/seind10.pdf

Nesta, N., Patel, P. (2004) National patterns of technology accumulation: Use of patent statistics. w: Moed, H.F. i in. (red.), *Handbook of Quantitative Science and Technology Research*, Amsterdam: Kluwer Academic Publishers, s. 531-551

NLM (2011a). Fact Sheet. Medical Subject Headings (MeSH®). www.nlm.nih.gov/pubs/factsheets/mesh.html

NLM (2011b). Fact Sheet. MEDLINE®. www.nlm.nih.gov/pubs/factsheets/medline.html

NLM (2011c). Fact Sheet. Loansome Doc®. www.nlm.nih.gov/pubs/factsheets/loansome_doc.html

NLM (2011d). Copyright Information & Downloading National Library of Medicine Data. www.nlm.nih.gov/databases/download.html

NLM (2011e). Fact Sheet. MedlinePlus®. www.nlm.nih.gov/pubs/factsheets/medlineplus.html

NLM (2011f). OLD MEDLINE Data. www.nlm.nih.gov/databases/databases_oldmedline.html

NLM (2011g). Fact Sheet. Authorship in MEDLINE®. www.nlm.nih.gov/pubs/factsheets/authorship.html

NLM (2011h). MEDLINE®/PubMed® Data Element (Field) Descriptions. www.nlm.nih.gov/bsd/mms/medlineelements.html

Nowak, P. (2006) *Bibliometria. Webometria. Podstawy. Wybrane zastosowania*, Poznań: Wydawnictwo Naukowe Uniwersytetu Adama Mickiewicza

OECD (2002) *Frascati manual. Proposed standard practice for surveys on research and experimental development*. Paris: OECD

- OECD (1994) The measurement of scientific and technological activities. Using patent data as science and technology indicators. Patent manual. Paris: OECD, www.oecd.org/dataoecd/33/62/2095942.pdf
- Ohniwa, R.L., Hibino, A., Takeyasu, K. (2010) Trends in research foci in life science fields over the last 30 years monitored by emerging topics. *Scientometrics*, 85, s. 111-127
- Okubo, Y. (1997) Bibliometric indicators and analysis of research systems: Methods and examples. OECD STI Working Papers, OECD/GD(97)41, www.oecd-ilibrary.org/science-and-technology/bibliometric-indicators-and-analysis-of-research-systems_208277770603
- Orloff, M. (2006) Inventive thinking through TRIZ. A practical guide. Berlin-Heidelberg-New York: Springer
- OST (2008) Indicateurs de sciences et de technologies. Édition 2008. Rapport de l'Observatoire des Sciences et des Techniques. Paris: Oeconomica, www.obs-ost.fr/fileadmin/medias/tx_ostdocuments/Partie0_Debut.indd.pdf
- Patbase (2011). An introduction to Patbase. www.patbase.com/pboverview.pdf
- Persson, O. (2010) Identifying research themes with weighted direct citation links. *Journal of Informetrics*, 4, s. 415-422
- Podolny, J.E., Stuart, T.E. (1995) A role-based ecology of technological change. *American Journal of Sociology*, 100, 5, s. 1224–1260
- Portal ACM (2011). ACM Computing Classification System (CCS). portal.acm.org/ccs.cfm?CFID=11168962&CFTOKEN=59839880
- Porter, A.L., Roper, A.T., Mason, T.W., Rossini, F.A., Banks, J. (1991) *Forecasting and management of technology*. New York: John Wiley & Sons
- Porter, A.L., Detampel, M.J. (1995) Technology opportunities analysis. *Technological Forecasting & Social Change*, 49, s. 237-255
- Porter, A.L., Ashton, W.B., Clar, G., Coates, J.F., Cuhls, K., Cunningham, S.W., Ducatel, K., Duin, P. van der, Georghiou, L., Gordon, Th., Linstone, H., Marchau, V., Massari, G., Miles, I., Moguee, M., Salo, A., Scapolo, F., Smits, R. (2004) Technology futures analysis: Toward integration of the field and new methods. *Technological Forecasting & Social Change*, 71, s. 287-303
- Porter, A.L., Cunningham, S.W. (2005) *Tech mining. Exploiting new technologies for competitive advantage*. New Jersey: Wiley-Interscience
- Porter, A.L., Rafols, I. (2009) Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics*, 81, 3, s. 719-745
- Porter, A.L., Newman, N.C. (2011) Mining external R&D. *Technovation*, w druku,

www.thevantagepoint.com/data/documents/Mining%20External%20R&D%20-%20preprint%20Technovation.pdf

Postma, Th.J.B.M., Liebl, F. (2005) How to improve scenario analysis as a strategic management tool? *Technological Forecasting Social Change*, 72, s. 161-173

Price, D.J. de S. (1965) Networks of scientific papers. *Science*, 149, s. 510-515

Rafols, I. Porter, A.L., Leydesdorff, L. (2010) Science overlay maps: a new tool for research policy and library management. *Journal of the American Society for Information Science & Technology*, 61, 9, s. 1871-1887

Reger, G. (2001) Technology foresight in companies: From an indicator to a network and process perspective. *Technology Analysis & Strategic Management*, 13, 4, s. 533-553

Ruegg, R., Jordan, G. (red.) (2007) Overview of evaluation methods for R&D programs. A directory of evaluation methods relevant to technology development programs. U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, http://www1.eere.energy.gov/ba/pba/pdfs/evaluation_methods_r_and_d.pdf

Schiling, M.A. (1998) Technological lockout: An integrative model of the economic and strategic factors driving technology success and failure. *Academy of Management Review*, 23, 2, s. 267-284

Schloegl, Ch., Gorraiz, J. (2011) Global usage versus global citation metrics: The case of pharmacology journals. *Journal of the American Society for Information Science and Technology*, 62, 1, s. 161-170

Schmoch, U. (2007) Leistungsfähigkeit und Strukturen der Wissenschaft im internationalen Vergleich 2006. Analysen im Rahmen der jährlichen Berichterstattung zur Technologischen Leistungsfähigkeit Deutschlands. Studien zum deutschen Innovationssystem, Nr. 11-2007, Fraunhofer-Institut für Systemtechnik und Innovationsforschung, www.bmbf.de/pubRD/sdi-11-07.pdf

SCIRUS (2011a). Advanced search. www.scirus.com/srsapp/advanced/index.jsp?saInstance=2&savalue=all;0&ffl=2&ffvalue=all;0&dsInstance=2&jnlvalue=web;nom;jnl;0&dtvalue=all;0&fdtSelected=before%201900&tdtSelected=2012&q1=&q2=&t1=all&t2=all&cn1=all&cn2=all&co=AND&js=1#sa

SCIRUS (2011b). Scirus help. www.scirus.com/html/help/index.htm

Sciverse (2011a). What does it cover? www.info.sciverse.com/scopus/scopus-in-detail/facts

SciVerse (2010) Scopus User Guide www.info.sciverse.com/UserFiles/2508%20SciVerse%20Scopus%20User%20Guide%20%28LR%29.pdf

- Shibata, N., Kajikawa, Y., Takeda, Y., Matsushima, K. (2009) Comparative study on methods of detecting research fronts using different types of citation. *Journal of the American Society for Information Science and Technology*, 60, 3, s. 571-580
- Slocum, M.S., Lundberg, C.O. (2001) Technology forecasting: From emotional to empirical. *Creativity and Innovation Management*, 10, 2, s. 139-152
- Smalheiser, N.R. (2005) The Arrowsmith project: 2005 status report. w: Hoffmann, A., Motoda, H., Scheffer, T. (red.), *Proceedings of the 8th International Conference on Discovery Science - DS 2005*, LNAI 3735, s. 26-43, http://arrowsmith.psych.uic.edu/arrowsmith_uic/tutorial/smalheiser_statusreport_2005.pdf
- Small, H. (1973) Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24, 4, s. 265-269
- Small, H. (1999) Visualizing science by citation mapping. *Journal of the American Society for Information Science*, 50, 9, s. 799-913
- Small, H. (2006) Tracking and predicting growth areas in science. *Scientometrics*, 68, 3, s. 595-610
- Stegmann, J., Grohmann, G. (2003) Hypothesis generation guided by co-word clustering. *Scientometrics*, 56, 1, s.111-135
- Swanson, D.R., Smalheiser, N.R. (1996) Undiscovered public knowledge: a ten-year update. *KDD-96 Proceedings*, American Association for Artificial Intelligence, www.aaai.org/Papers/KDD/1996/KDD96-051.pdf
- Thissen, W. (2004) Technology futures analysis: Toward integration of the field and new methods. *Technology Forecasting & Social Change*, 71, s. 287-303
- Thom, Nico, Rohrbeck, René (2009) Technology foresight in the ICT sector. Exploration of new business opportunities. 2nd ISPIM Innovation Symposium: New York City, USA. <http://ssrn.com/abstract=1522883>
- Tijssen, R.J.W. (2004) Measuring and evaluating science-technology connections and interactions. Towards international statistics. w: Moed, H.F. i in. (red.), *Handbook of Quantitative Science and Technology Research*, Amsterdam: Kluwer Academic Publishers, s. 695-715
- Todhunter, J. (2009) Method for problem formulation and for obtaining solutions from a database. United States Patent No. 7.536.368-B2
- Todhunter, J., Bialiauski, Y., Brykman-Weiser, L.V. (2010) System and method for knowledge research. United States Patent Application No. US 2010/0235340-A1
- Tong, L.H., Cong, H., Lixiang, S. (2006) Automatic classification of patent documents for TRIZ users. *World Patent Information*, 28, s. 6-13

- Tseng, Y.-H., Lin, Y.I., Lee, Y.-Y., Hung, W.-Ch., Lee, Ch.-H. (2009) A comparison of methods for detecting hot topics. *Scientometrics*, 81, 1, s. 73-90
- Tsourikov, V.W., Batchilo, L.S., Sovpel, I.V. (2000) Document semantic analysis/selection with knowledge creativity capability utilizing Subject-Action-Object (SAO) structures. United States Patent No. 6.167.370
- Tushman, Ph., Anderson, M.L. (1986) Technological discontinuities and organizational environments. *Administrative Science Quarterly*, 31, s. 439-465
- Tushman, Ph., Anderson, M.L. (1990) Technological discontinuities and dominant designs: A cyclical model of technological change. *Administrative Science Quarterly*, 35, s. 604–633
- Upham, S.P., Small, H. (2010) Emerging research fronts in science and technology: patterns of new knowledge development. *Scientometrics*, 83, s. 15-38
- Urazajew [Уразаев], W. (2006) TRIZ w elektronice [ТРИЗ в Электронике]. Moskwa: Technosfera [Техносфера]
- USPTO (2011a). USPTO Patent Full – Text and Image Database. Boolean Search. <http://patft.uspto.gov/netahtml/PTO/search-bool.html>
- USPTO (2011b). USPTO Patent Full – Text and Image Database. Manual Search. <http://patft.uspto.gov/netahtml/PTO/search-adv.htm>
- USPTO (2011c). USPTO Patent Full – Text and Image Database. Help on the Advanced Search Page. www.uspto.gov/patft/help/helpadv.htm
- USPTO (2011d). USPTO Patent Full – Text and Image Database. Stopwords. www.uspto.gov/patft/help/stopword.htm
- USPTO (2011e). Check the Filing Status of Your Patent Application. <http://patents.uspto.gov/patents/process/status/index.jsp>
- Wang, M.-Y., Chang, D.-S., Kao, Ch.-H. (2010) Identifying technology trends for R&D planning. *R&D Management*, 40, 5, s. 491-508
- Watts, R.J., Porter, A.L. (1997) Innovation forecasting. *Technological Forecasting & Social Change*, 56, s. 25-47
- Weeber, M., Kors, J.A., Mons, B. (2005) Online tools to support literature-based discovery in the life sciences. *Briefings in Bioinformatics*, 6, 3, s. 277-286
- Whittaker, J. (1989) Creativity and conformity in science: Titles, keywords and co-word analysis. *Social Studies of Science*, 19, s. 473-496
- WoK (2011a). Web of Knowledge. What it is. <http://wokinfo.com/about/whatitis/>
- WoK (2011b). Web of Knowledge. FAQ. <http://wokinfo.com/about/faq/>
- WoK (2011c). Web of Knowledge. Multidisciplinary Resources. http://wokinfo.com/products_tools/multidisciplinary/

- WoK (2011d). Web of Knowledge. Discipline – Specific Resources. http://wokinfo.com/products_tools/specialized/
- WoK (2011e). Web of Science. Skrócona instrukcja obsługi. http://science.thomsonreuters.com/m/pdfs/wos_qrc_po.pdf
- WoK (2011f). Citation Mapping. http://isiwebofknowledge.com/products_tools/multidisciplinary/webofscience/citm_ap/
- WoK (2011g). Web of Science. http://wokinfo.com/products_tools/multidisciplinary/webofscience/
- Wyk, R.J. van (1997) Strategic technology scanning. *Technological Forecasting & Social Change*, 55, s. 21-38
- Yang, Y.Y., Akers, L., Klose, T., Yang, C.B. (2008) Text mining and visualization tools - Impressions of emerging capabilities. *World Patent Information*, 30, s. 280-293
- Yoon, J., Choi, S., Kim, K. (2011) Invention property-function network analysis of patents: a case of silicon-based thin film solar cells. *Scientometrics*, 86, s. 687-703
- Yoon, B., Lee, S., Lee, G. (2010) Development and application of a keyword-based knowledge map for effective R&D planning. *Scientometrics*, 85, s. 803-820
- Yoon, B., Park, Y. (2005) A systematic approach for identifying technology opportunities: Keyword-based morphology analysis. *Technological Forecasting & Social Change*, 72, s. 145-160
- Yoon, B., Phaal, R., Probert, D. (2008) Morphology analysis for technology roadmapping: application of text mining. *R&D Management*, 38, 1, s. 51-68
- Yu, H., Davis, M., Wilson, C.S., Cole, F.T.H. (2008) Object-relational data modelling for informetric database. *Journal of Informetrics*, 2, s. 240-251
- Zhou P., Leydesdorff, L. (2006) The emergence of China as a leading nation in science. *Research Policy*, 35, s. 83-104

Wykaz rysunków

Rysunek 1: Proces komercjalizacji pomysłów. Źródło: Klincewicz (2010: 27).....	24
Rysunek 2: Model rozwoju technologii w oparciu o krzywą S-kształtną. Źródło: Klincewicz (2010: 41).....	29
Rysunek 3: Doskonalenie systemów transportowych poprzez substytucję kolejnych technologii. Źródło: Orloff (2006: 200)..	30
Rysunek 4: Sieć techno-ekonomiczna jako podstawa analiz bibliometrycznych. Źródło: Bell i Callon (1994: 75).	38
Rysunek 5: Sieć zależności pomiędzy badaniami medycznymi a poprawą zdrowia publicznego. Źródło: Lewison (2004: 459).....	38
Rysunek 6: Ocena ważności poszczególnych publicznie dostępnych źródeł wiedzy dla firm z wybranych przemysłów w oparciu o badania ankietowe. Źródło: opracowanie własne na podstawie Cohen, Nelson, Walsh (2002: 11).....	50
Rysunek 7: Streszczenie artykułu w bazie <i>Web of Science</i> . Źródło: <i>Web of Science</i>	52
Rysunek 8: Streszczenie artykułu w bazie <i>Scopus</i> . Źródło: <i>Scopus</i>	53
Rysunek 9: Streszczenie artykułu w bazie <i>MEDLINE</i> . Źródło: <i>MEDLINE</i>	53
Rysunek 10: Streszczenie artykułu chemicznego w bazie <i>INSPEC</i> . Źródło: <i>INSPEC</i>	54
Rysunek 11: Streszczenie artykułu astronomicznego w bazie <i>INSPEC</i> . Źródło: <i>INSPEC</i>	54
Rysunek 12: Opis patentu w bazie <i>USPTO</i> . Źródło: <i>USPTO</i>	55
Rysunek 13: Opis patentu w bazie <i>Google Patents</i> . Źródło: <i>Google Patents</i>	56
Rysunek 14: Opis patentu w bazie <i>UPRP</i> . Źródło: <i>UPRP</i>	56
Rysunek 15: Opis patentu w bazie <i>EPO esp@cenet</i> . Źródło: <i>Esp@cenet</i>	57
Rysunek 16: Problemy z identyfikacją patentów w oparciu o kody i słowa kluczowe. Źródło: Hinze, Schmoch (2004: 231).....	71
Rysunek 17: Liczba publikacji dotyczących robotyki zidentyfikowanych w bazie <i>INSPEC</i> w oparciu o wybrane słowa kluczowe. Źródło: opracowanie własne w oparciu o badania Gity Ghiasi Hafezi (2011), stan na 27 lutego 2011.	73
Rysunek 18: Liczba publikacji naukowych w naukach ścisłych i technicznych a liczba patentów, udzielonych wynalazcom z poszczególnych krajów dla lat 1990-2007 w oparciu o bazy <i>INSPEC</i> i <i>USPTO</i> . Źródło: Klincewicz (2008: 57).....	79
Rysunek 19: Średnie roczne całkowite wydatki gospodarek na B+R (GERD) w latach 1989-2006 a liczba publikacji naukowych w naukach ścisłych i technicznych w latach 1990-2007 w oparciu o bazy <i>INSPEC</i> i <i>OECD-Stat</i> . Źródło: Klincewicz (2008: 60).	79
Rysunek 20: Średnia liczba badaczy w sektorze B+R w latach 1989-2006 a liczba publikacji naukowych w naukach ścisłych i technicznych w latach 1990-2007 w oparciu o bazy <i>INSPEC</i> i <i>OECD-Stat</i> . Źródło: Klincewicz (2008: 64).	80

Rysunek 21: Przykład graficznej prezentacji wyników analizy <i>co-words</i> . Źródło: Leydesdorff (2008: 1595).	81
Rysunek 22: Przykład analizy współautorstwa: sieć autorów publikacji informatycznych z Uniwersytetu Warszawskiego. Źródło: Klincewicz (2008: 183).	84
Rysunek 23: Lista publikacji cytujących wybrany artykuł naukowy. Źródło: <i>Web of Science</i>	88
Rysunek 24: Graficzna forma prezentacji publikacji cytujących wybrany artykuł naukowy (mapa cytowań). Źródło: <i>Web of Science</i>	88
Rysunek 25: Sieć cytowań patentów z <i>USPTO</i> , dotyczących laserów półprzewodnikowych opartych na azotku galu. Źródło: Klincewicz (2010: 148).	90
Rysunek 26: Sieć cytowań patentów pomiędzy grupami organizacji australijskich. Źródło: Hicks (2002: 141).	91
Rysunek 27: Warianty analizy cytowań. Źródło: Boyack, Klavans (2010: 2390).	92
Rysunek 28: Wskaźniki dotyczące liczby pobrań artykułu z bazy jako sposób na określanie popularności publikacji. Źródło: <i>ACM</i>	94
Rysunek 29: Przykładowa operacja oczyszczania danych bibliometrycznych – krok 1. Źródło: Frey, Minsk, Porter (2005: 18).	100
Rysunek 30: Przykładowa operacja oczyszczania danych bibliometrycznych – krok 2. Źródło: Frey, Minsk, Porter (2005: 18).	100
Rysunek 31: Przykładowa operacja oczyszczania danych bibliometrycznych – krok 3. Źródło: Frey, Minsk, Porter (2005: 20).	101
Rysunek 32: Czyszczenie listy afiliacji autorów publikacji naukowych w programie <i>VantagePoint</i> . Źródło: Myszkka (2010: 18).	101
Rysunek 33: Przykład macierzy autokorelacji dla podobszarów badań medycznych w programie <i>VantagePoint</i> . Źródło: materiały marketingowe firmy Search Technology, Inc..	106
Rysunek 34: Przykład zaawansowanych analiz w programie <i>VantagePoint</i> . Źródło: Frey, Minsk, Porter (2005: 36).	107
Rysunek 35: Graficzna forma prezentacji wyników analizy <i>PCA</i> w programie <i>VantagePoint</i> . Źródło: Frey, Minsk, Porter (2005: 51).	109
Rysunek 36: Wyszukiwanie literatury dotyczącej podstawowego analizowanego zagadnienia („A”) zgodnie z podejściem <i>Literature-Based Discovery</i> w programie <i>Arrowsmith</i> . Źródło: strona internetowa programu <i>Arrowsmith</i>	116
Rysunek 37: Lista zagadnień („B”) występujących wspólnie z dwoma wskazanymi terminami („A” i „C”) zgodnie z podejściem <i>Literature-Based Discovery</i> w programie <i>Arrowsmith</i> . Źródło: strona internetowa programu <i>Arrowsmith</i>	116

Rysunek 38: Zestawienie literatury dotyczącej potencjalnego odkrycia („B”) zgodnie z podejściem <i>Literature-Based Discovery</i> w programie <i>Arrowsmith</i> . Źródło: strona internetowa programu <i>Arrowsmith</i>	117
Rysunek 39: Fragment „tablicy sprzeczności” TRIZ. Źródło: opracowanie własne w oparciu o: Altszuller (1972).	119
Rysunek 40: Algorytm komputerowego przetwarzania tekstów naukowych, wykorzystywany przez program firmy <i>Invention Machine</i> . Źródło: Tsourikov i in. (2000).	122
Rysunek 41: Przykładowe możliwości poszukiwania rozwiązania problemu wynalazczego w oprogramowaniu firmy <i>Invention Machine</i> . Źródło: Devoino i in. (2001).	123
Rysunek 42: Możliwości przeszukiwania zasobów wiedzy naukowej w programie <i>Invention Machine Goldfire</i> . Źródło: strona internetowa firmy <i>Invention Machine</i>	123
Rysunek 43: Przykład analizy potencjału ewolucyjnego technologii zgodnie z podejściem <i>TRIZ</i> . Źródło: Mann (2003: 788).	126
Rysunek 44: Przykładowa mapa zależności pomiędzy pojęciami, występującymi w publikacjach naukowych, stworzona w programie <i>CiteSpace</i> . Źródło: strona internetowa programu <i>CiteSpace</i>	131
Rysunek 45: Wizualizacja sieci cytowań w przekroju czasowym przy wykorzystaniu programu <i>CiteSpace</i> . Źródło: Chen, Guan (2011: 245).	132
Rysunek 46: Mapa obszarów badawczych, które są mało konkurencyjne i niedostatecznie rozwinięte w Stanach Zjednoczonych. Źródło: MapOfScience.com	134
Rysunek 47: Dorobek badawczy Uniwersytetu w Amsterdamie zaprezentowany na globalnej mapie wiedzy. Źródło: www.idr.gatech.edu/detail.php?tab=2&id=3	135
Rysunek 48: Badania nad kinezyną zaprezentowane na globalnej mapie wiedzy. Źródło: www.idr.gatech.edu/detail.php?show=dyn&d=1	136
Rysunek 49: Przykładowa lista popularnych zagadnień badawczych „ <i>Top Topics</i> ”, oferowana przez serwis <i>ScienceWatch</i> . Źródło: http://sciencewatch.com/dr/tt/2011/11-febtt	137
Rysunek 50: Przykładowa mapa światowej nauki, oferowana przez serwis <i>ScienceWatch</i> . Źródło: http://sciencewatch.com/dr/rfm/mos/10maymosGLOBAL	138
Rysunek 51: Mapa przewag technologicznych krajów lub instytucji. Źródło: Nesta, Patel (2004: 244).	143
Rysunek 52: Zestawienie RTA i liczby publikacji dla badań nad <i>business intelligence</i> . Źródło: Klineciewicz, Miyazaki (2011: 181).	143
Rysunek 53: Interfejs użytkownika w bazie <i>Web of Science</i> - wyszukiwanie. Źródło: <i>Web of Science</i>	153
Rysunek 54: Interfejs użytkownika w bazie <i>Web of Science</i> – opcje eksportowania wyników wyszukiwania. Źródło: <i>Web of Science</i>	156

Rysunek 55: Interfejs użytkownika w bazie <i>Web of Science</i> –ogólne kategorie nauk i obszary tematyczne. Źródło: <i>Web of Science</i>	156
Rysunek 56: Interfejs użytkownika w bazie <i>Web of Science</i> – narzędzie analizy wyników. Źródło: <i>Web of Science</i>	158
Rysunek 57: Interfejs użytkownika w bazie <i>Scopus</i> – kryteria wyszukiwania. Źródło: <i>Scopus</i>	161
Rysunek 58: Interfejs użytkownika w bazie <i>Scopus</i> – opcje eksportowania wyników wyszukiwania. Źródło: <i>Scopus</i>	162
Rysunek 59: Interfejs użytkownika w bazie <i>Scopus</i> – ogólne kategorie nauk i obszary tematyczne. Źródło: <i>Scopus</i>	163
Rysunek 60: Interfejs użytkownika w bazie <i>Scopus</i> – analizator czasopism. Źródło: <i>Scopus</i>	164
Rysunek 61: Przykładowy przegląd cytowań w bazie <i>Scopus</i> – wersja tabelaryczna. Źródło: <i>Scopus</i>	165
Rysunek 62: Przykładowy przegląd cytowań w bazie <i>Scopus</i> – wersja graficzna. Źródło: <i>Scopus</i>	165
Rysunek 63: Opcje generowania listy cytowań w bazie <i>Scopus</i> . Źródło: <i>Scopus</i>	166
Rysunek 64: Przegląd opcji wyszukiwania w bazie <i>INSPEC</i> . Źródło: <i>INSPEC</i>	167
Rysunek 65: Opcja wydruku wyników wyszukiwania w bazie <i>INSPEC</i> . Źródło: <i>INSPEC</i>	169
Rysunek 66: Opcja przesłania e-mailem wyników wyszukiwania w bazie <i>INSPEC</i> . Źródło: <i>INSPEC</i>	170
Rysunek 67: Opcja przesłania eksportu i zapisu wyników wyszukiwania w bazie <i>INSPEC</i> . Źródło: <i>INSPEC</i>	170
Rysunek 68: Wyszukiwanie z pomocą klasyfikacji <i>MeSH</i> w bazie <i>MEDLINE</i> . Źródło: <i>MEDLINE</i>	172
Rysunek 69: Sugerowane przez tezaurs powiązane terminy w bazie <i>MEDLINE</i> . Źródło: <i>MEDLINE</i>	172
Rysunek 70: Interfejs użytkownika w bazie <i>MEDLINE</i> – kryteria wyszukiwania. Źródło: <i>MEDLINE</i>	174
Rysunek 71: Interfejs użytkownika w bazie <i>MEDLINE</i> – opcje eksportowania wyników wyszukiwania. Źródło: <i>MEDLINE</i>	175
Rysunek 72: Formularz wyszukiwania w bazie <i>EPO</i> . Źródło: <i>Esp@cenet</i>	183
Rysunek 73: Przykład widoku dokumentu w bazie <i>EPO</i> . Źródło: <i>Esp@cenet</i>	184
Rysunek 74: Opcja eksportu dokumentu w bazie <i>EPO</i> . Źródło: <i>Esp@cenet</i>	185
Rysunek 75: Opcja eksportu listy wyszukiwania w bazie <i>EPO</i> . Źródło: <i>Esp@cenet</i>	185

Rysunek 76: Przykład wyszukiwania wg klasyfikacji <i>ECLA</i> w bazie <i>EPO</i> . Źródło: <i>Esp@cenet</i>	186
Rysunek 77: Kryteria wyszukiwania w bazie <i>USPTO</i> . Źródło: <i>USPTO</i>	190
Rysunek 78: Opcja przeglądania rysunków w bazie <i>USPTO</i> . Źródło: <i>USPTO</i>	190
Rysunek 79: Opcja odpłatnego pobrania dokumentacji w bazie <i>USPTO</i> . Źródło: <i>USPTO</i> . .	191
Rysunek 80: Opcje cen dokumentacji w bazie <i>USPTO</i> . Źródło: <i>USPTO</i>	191
Rysunek 81: Formularz wyszukiwania zaawansowanego w bazie <i>USPTO</i> . Źródło: <i>USPTO</i>	192
Rysunek 82: Przykład okna pomocy w bazie <i>USPTO</i> . Źródło: <i>USPTO</i>	192
Rysunek 83: Przykład wyszukiwania w bazie <i>USPTO</i> . Źródło: <i>USPTO</i>	193
Rysunek 84: Opcja przeglądania pełnej dokumentacji w bazie <i>USPTO</i> . Źródło: <i>USPTO</i>	193
Rysunek 85: Formularz wyszukiwania podstawowego w bazie <i>Google Patents</i> . Źródło: <i>Google Patents</i>	196
Rysunek 86: Formularz wyszukiwania zaawansowanego w bazie <i>Google Patents</i> . Źródło: <i>Google Patents</i>	197
Rysunek 87: Przykład listy wyników wyszukiwania w bazie <i>Google Patents</i> . Źródło: <i>Google Patents</i>	198
Rysunek 88: Przykład graficznej listy wyników w bazie <i>Google Patents</i> . Źródło: <i>Google Patents</i>	198
Rysunek 89: Widok przykładowego dokumentu w bazie <i>Google Patents</i> . Źródło: <i>Google Patents</i>	199
Rysunek 90: Mapowanie cytowań w bazie <i>Delphion</i> . Źródło: strona internetowa <i>Delphion</i>	205
Rysunek 91: Sporządzanie wykresów w bazie <i>Delphion</i> . Źródło: strona internetowa <i>Delphion</i>	205
Rysunek 92: <i>PatentLab II</i> w bazie <i>Delphion</i> . Źródło: strona internetowa <i>Delphion</i>	206
Rysunek 93: Przykład analizy słów kluczowych w bazie <i>Delphion</i> . Źródło: strona internetowa <i>Delphion</i>	206
Rysunek 94: Przykład drzewa korporacyjnego w bazie <i>Delphion</i> . Źródło: strona internetowa <i>Delphion</i>	207
Rysunek 95: Przykład wykresów w bazie <i>Micro Patent</i> . Źródło: strona internetowa <i>Micropat</i>	208
Rysunek 96: Przykład analizy cytowań w bazie <i>PatBase</i> . Źródło: materiały marketingowe <i>Patbase</i>	209
Rysunek 97: Przykład użycie translatora w bazie <i>PatBase</i> . Źródło: materiały marketingowe <i>Patbase</i>	209

Rysunek 98: Przykład analizy graficznej danych w bazie <i>PatBase</i> . Źródło: materiały marketingowe <i>Patbase</i> .	210
Rysunek 99: Przykład wykresów geograficznych w bazie <i>QPat</i> . Źródło: strona internetowa <i>Questel</i> .	211
Rysunek 100: Przykład relacyjnych wykresów sieciowych i słupkowych w bazie <i>QPat</i> . Źródło: strona internetowa <i>Questel</i> .	212
Rysunek 101: Przykład wykresu organizacji-właścicieli patentów i ich klasyfikacji w bazie <i>QPat</i> . Źródło: strona internetowa <i>Questel</i> .	213
Rysunek 102: Przykład wykresu cytowań organizacji-właścicieli patentów w bazie <i>QPat</i> . Źródło: strona internetowa <i>Questel</i> .	213
Rysunek 103: Przykład krzyżowego wykresu relacyjnego w bazie <i>QPat</i> . Źródło: strona internetowa <i>Questel</i> .	214
Rysunek 104: Przykład analitycznego wykresu cytowań w bazie <i>WIPS</i> . Źródło: strona internetowa <i>WIPS</i> .	215
Rysunek 105: Przykład analitycznego wykresu relacyjnego w bazie <i>WIPS</i> . Źródło: strona internetowa <i>WIPS</i> .	215
Rysunek 106: Przykładowy widok ekranu <i>PM Manager</i> w bazie <i>WIPS</i> . Źródło: strona internetowa <i>WIPS</i> .	216
Rysunek 107: Przykład mapy tematycznej w bazie <i>Thomson Innovation</i> . Źródło: strona internetowa <i>Thomson Innovation</i> .	217
Rysunek 108: Widok formularza wyszukiwania struktur chemicznych w bazie <i>SureChem</i> . Źródło: strona internetowa <i>SureChem</i> .	218
Rysunek 109: Przykład wizualizacji i grupowania patentów w bazie <i>Patent Integration</i> . Źródło: strona internetowa <i>Patent Integration</i> .	220
Rysunek 110: Przykład wykresu prezentującego liczby patentów w bazie <i>Patent Integration</i> . Źródło: strona internetowa <i>Patent Integration</i> .	220
Rysunek 111: Program <i>Matheo Analyzer</i> – okna opcji importu danych. Źródło: strona internetowa <i>Matheo Analyzer</i> .	227
Rysunek 112: Program <i>Matheo Analyzer</i> – okno prezentujące wyniki pobierania danych z programu <i>Matheo Patent</i> . Źródło: strona internetowa <i>Matheo Analyzer</i> .	228
Rysunek 113: Program <i>Matheo Analyzer</i> – przykładowa macierz czynnikowa. Źródło: strona internetowa <i>Matheo Analyzer</i> .	229
Rysunek 114: Program <i>Matheo Analyzer</i> – przykładowa mapa prezentująca powiązania pomiędzy organizacjami-właścicielami patentów. Źródło: strona internetowa <i>Matheo Analyzer</i> .	229

Rysunek 115: Program <i>Matheo Analyzer</i> – przykładowa mapa prezentująca powiązania pomiędzy danymi w obrębie dwóch zmiennych. Źródło: strona internetowa <i>Matheo Analyzer</i>	230
Rysunek 116: Program <i>Matheo Analyzer</i> – przykładowy klaster danych. Źródło: strona internetowa <i>Matheo Analyzer</i>	230
Rysunek 117: Program <i>CiteSpace</i> – mapa wygenerowana przy wykorzystaniu serwisu <i>Google Earth</i> . Źródło: Chen (2009).	232
Rysunek 118: Program <i>CiteSpace</i> – przykładowe okna obrazujące równocześnie różne poziomy tej samej analizy. Źródło: strona internetowa <i>CiteSpace</i>	232
Rysunek 119: Program <i>CiteSpace</i> – schemat prezentujący kluczowych autorów w analizowanym obszarze badań. Źródło: Chen (2009).	233
Rysunek 120: Program <i>CiteSpace</i> – schemat prezentujący analizę zmian publikacji w czasie. Źródło: strona internetowa <i>CiteSpace</i>	233
Rysunek 121: Program <i>CiteSpace</i> – schemat prezentujący zestawienie kilku serii danych. Źródło: strona internetowa <i>CiteSpace</i>	234
Rysunek 122: Program <i>CiteSpace</i> – schemat prezentujący wyniki analizy z użyciem krajów jako węzłów sieci. Źródło: Chen (2009).	234
Rysunek 123: Program <i>CiteSpace</i> – schemat prezentujący wyniki analizy z użyciem ośrodków badawczych jako węzłów sieci. Źródło: Chen (2009).	235
Rysunek 124: Program <i>CiteSpace</i> – schemat prezentujący wyniki analizy z użyciem nazwisk badaczy jako węzłów sieci. Źródło: Chen (2009).	235
Rysunek 125: Program <i>VantagePoint</i> – interfejs użytkownika. Źródło: plik pomocy programu <i>VantagePoint</i>	237
Rysunek 126: Program <i>VantagePoint</i> – przykładowy wynik analiz w postaci mapy auto-korelacji. Źródło: Porter, Cunningham (2005: 232).	238
Rysunek 127: Program <i>HistCite</i> – przykładowy wynik analizy afiliacji autorów publikacji. Źródło: opracowanie własne.	240
Rysunek 128: Program <i>HistCite</i> – przykładowy wynik analizy autorów publikacji. Źródło: opracowanie własne.	241
Rysunek 129: Program <i>HistCite</i> – lista cytowań publikacji z odnośnikami do bazy <i>Web of Science</i> . Źródło: opracowanie własne.	241
Rysunek 130: Program <i>HistCite</i> – przykładowy historiograf, prezentujący sieć cytowań w czasie. Źródło: http://www.garfield.library.upenn.edu/histcomp/guide.html	243
Rysunek 131: Program <i>PatentNet</i> – wykres sieci cytowań patentów. Źródło: strona internetowa <i>PatentNet</i>	244

- Rysunek 132: Program *PatentNet* – wykres zależności pomiędzy organizacjami-właścicielami patentów, ujawnionych poprzez analizę sieci cytowań patentów. Źródło: strona internetowa *PatentNet*. 245
- Rysunek 133: Program *PatentNet* – wykres zależności pomiędzy wynalazcami-twórcami patentów, ujawnionych poprzez analizę sieci cytowań patentów. Źródło: strona internetowa *PatentNet*. 246
- Rysunek 134: Wskaźniki cytowań dla wszystkich artykułów 24 polskich uczelni i dla artykułów z zagranicznymi współautorami, 2000-2009. Źródło danych: baza *Scopus*. 296

Wykaz tabel

Tabela 1: Zależność pomiędzy badaniami podstawowymi, stosowanymi i pracami rozwojowymi a rodzajami wyników badań. Źródło: opracowanie własne.....	27
Tabela 2: Porównanie mechanizmów transferu wiedzy technicznej. Źródło: Tijssen (2004: 701).....	37
Tabela 3: Metody ewaluacji wykorzystywane przez instytucje publiczne w Stanach Zjednoczonych w okresie do roku 2002. Źródło: Ruegg, Jordan (2007: 4).....	43
Tabela 4: Zestawienie analiz badań naukowych i rozwoju technologii, prowadzonych na zlecenie instytucji publicznych w wybranych krajach w latach 2004-2010. Źródło: Holtmannspötter i in. (2010: 173).	46
Tabela 5: Intensywność wykorzystywania metod w zakresie technologii intelligence w wybranych branżach. Źródło: Lichtenthaler (2004: 130).....	47
Tabela 6: Intensywność wykorzystywania źródeł informacji w procesach technologii intelligence w wybranych branżach. Źródło: Lichtenthaler (2004: 130).	49
Tabela 7: Liczba zgłoszeń patentowych do <i>EPO</i> (2006) i patentów udzielonych przez <i>USPTO</i> (2006) podmiotom z poszczególnych krajów. Źródło: Eurostat (2010: 608).	58
Tabela 8: Artykuły naukowe opublikowane w roku 2007, zgromadzone w bazie <i>Web of Science</i> . Źródło: National Science Board (2010: 5-31).....	59
Tabela 9. Zestawienie pytań analitycznych w projektach bibliometrycznych. Źródło: opracowanie własne.	67
Tabela 10: Przykład fragmentu wyników analizy współklasyfikacji. Źródło: opracowanie własne.....	82
Tabela 11: Przykład fragmentu analizy współautorstwa. Źródło: opracowanie własne.	84
Tabela 12: Możliwe scenariusze analiz krzyżowych przy wykorzystaniu technik <i>tech mining</i> . Źródło: opracowanie własne.	104
Tabela 13: Wskaźniki RTA obliczone w oparciu o patenty w poszczególnych obszarach technologicznych dla wybranych krajów OECD dla lat 1991-2000. Źródło: Nesta, Patel (2004: 539).	142
Tabela 14: Podsumowanie informacji na temat bazy <i>CAS</i> . Źródło: opracowanie własne.	146
Tabela 15: Podsumowanie informacji na temat bazy <i>Compendex</i> . Źródło: opracowanie własne.....	147
Tabela 16: Podsumowanie informacji na temat bazy <i>ACM DL</i> . Źródło: opracowanie własne.	148
Tabela 17: Podsumowanie informacji na temat wyszukiwarki <i>Scirus</i> . Źródło: opracowanie własne.....	149

Tabela 18: Podsumowanie informacji na temat wyszukiwarki <i>Google Scholar</i> . Źródło: opracowanie własne.	150
Tabela 19: Podsumowanie informacji na temat bazy <i>Web of Science</i> . Źródło: opracowanie własne.	151
Tabela 20: Podsumowanie informacji na temat bazy <i>Scopus</i> . Źródło: opracowanie własne.	151
Tabela 21: Podsumowanie informacji na temat bazy <i>INSPEC</i> . Źródło: opracowanie własne.	152
Tabela 22: Podsumowanie informacji na temat bazy <i>MEDLINE</i> . Źródło: opracowanie własne.	152
Tabela 23: Kryteria wyszukiwania w bazie <i>Web of Science</i> . Źródło: opracowanie własne.	154
Tabela 24: Opcje eksportu wyników w bazie <i>Web of Science</i> . Źródło: opracowanie własne.	155
Tabela 25: Przykładowe wyniki wyszukiwania w <i>Web of Science</i> dla kilku haseł z wybranych obszarów nauki (w polu: temat) dla okresu 2005 – 2010. Źródło: opracowanie własne.	157
Tabela 26: Kryteria wyszukiwania w bazie <i>Scopus</i> . Źródło: opracowanie własne.	161
Tabela 27: Opcje eksportu wyników w bazie <i>Scopus</i> . Źródło: opracowanie własne.	162
Tabela 28: Przykładowe wyniki wyszukiwania w bazie <i>Scopus</i> dla kilku haseł z wybranych obszarów nauki (w polu: tytuł artykułu, abstrakt, słowa kluczowe) dla okresu 2005 – 2010. Źródło: opracowanie własne.	163
Tabela 29: Kryteria wyszukiwania w bazie <i>INSPEC</i> . Źródło: opracowanie własne.	168
Tabela 30: Opcje eksportu wyników w bazie <i>INSPEC</i> . Źródło: opracowanie własne.	170
Tabela 31: Przykładowe wyniki wyszukiwania w bazie <i>INSPEC</i> dla kilku haseł z wybranych obszarów nauki (w polu: „ <i>terms anywhere</i> ”) dla okresu 2005 – 2010. Źródło: opracowanie własne.	171
Tabela 32: Kryteria wyszukiwania w bazie <i>MEDLINE</i> . Źródło: opracowanie własne.	173
Tabela 33: Opcje eksportu wyników w bazie <i>MEDLINE</i> . Źródło: opracowanie własne.	174
Tabela 34: Przykładowe wyniki wyszukiwania w bazie <i>MEDLINE</i> dla kilku haseł z wybranych obszarów nauki (w polu: „ <i>all fields</i> ”) dla okresu 2005 – 2010. Źródło: opracowanie własne.	176
Tabela 35: Kryteria wyszukiwania – zestawienie zbiorcze. Źródło: opracowanie własne.	179
Tabela 36: Opcje eksportu wyników w bazach – tabela zbiorcza. Źródło: opracowanie własne.	180
Tabela 37: Przykładowe wyniki wyszukiwania w bazach dla kilku haseł z wybranych obszarów nauki (w najszerszej określonej opcji wyszukiwania) dla okresu 2005 – 2010. Źródło: opracowanie własne.	181
Tabela 38: Kryteria wyszukiwania w bazie <i>EPO</i> . Źródło: opracowanie własne.	182

Tabela 39: Przykładowe wyniki wyszukiwania w bazie <i>EPO</i> dla kilku haseł z wybranych obszarów nauki (w polu: słowa kluczowe w tytule i abstrakcie) w roku 2010. Źródło: opracowanie własne.	187
Tabela 40: Podsumowanie bazy <i>EPO</i> . Źródło: opracowanie własne, przy wykorzystaniu danych z: Intellogist (2011a).	188
Tabela 41: Kryteria wyszukiwania w bazie <i>USPTO</i> . Źródło: opracowanie własne.	189
Tabela 42: Charakterystyka bazy patentów <i>USPTO</i> . Źródło: opracowanie własne.	190
Tabela 43: Przykładowe wyniki wyszukiwania w bazie <i>USPTO</i> dla kilku haseł z wybranych obszarów nauki (w polu: „ <i>all fields</i> ”) w roku 2010. Źródło: opracowanie własne.	194
Tabela 44: Podsumowanie bazy <i>USPTO</i> . Źródło: opracowanie własne.	195
Tabela 45: Kryteria wyszukiwania w bazie <i>Google Patents</i> . Źródło: opracowanie własne. .	196
Tabela 46: Charakterystyka bazy patentów <i>Google Patents</i> . Źródło: opracowanie własne. .	197
Tabela 47: Przykładowe wyniki wyszukiwania w bazie <i>Google Patents</i> dla kilku haseł z wybranych obszarów nauki (w polu: „ <i>exact phrase</i> ”) w 2010 roku. Źródło: opracowanie własne.	200
Tabela 48: Podsumowanie bazy <i>Google Patents</i> . Źródło: opracowanie własne przy wykorzystaniu Intellogist (2011b).	201
Tabela 49: Kryteria wyszukiwania – zestawienie zbiorcze. Źródło: opracowanie własne. ...	203
Tabela 50: Przykładowe wyniki wyszukiwania w bazach dla kilku haseł z wybranych obszarów nauki (w najszerszym polu wyszukiwania) w 2010. Źródło: opracowanie własne.	203
Tabela 51: Najważniejsze cechy oprogramowania wzbogaconego o moduł do eksploracji danych tekstowych. Źródło: opracowanie własne na podstawie (Porter, Cunningham 2005) oraz materiałów producentów.	224
Tabela 52: Najważniejsze cechy oprogramowania zintegrowanego z bibliometrycznymi bazami danych. Źródło: opracowanie własne na podstawie (Porter, Cunningham 2005) oraz materiałów producentów.	224
Tabela 53: Najważniejsze cechy oprogramowania o charakterze uniwersalnym. Źródło: opracowanie własne na podstawie (Porter, Cunningham 2005) oraz materiałów producentów.	225
Tabela 54: Najważniejsze cechy oprogramowania stworzonego do wizualizacji wyników. Źródło: opracowanie własne na podstawie (Porter, Cunningham 2005) oraz materiałów producentów.	225
Tabela 55: Pytania badawcze w bibliometrii a możliwości analityczne oprogramowania. Źródło: opracowanie własne.	247

Tabela 56: Zestawienie najpopularniejszych płatnych programów, mogących wspierać analizy bibliometryczne. Źródło: opracowanie własne na podstawie stron internetowych producentów oprogramowania.....	251
Tabela 57: Zestawienie najpopularniejszych bezpłatnych programów, mogących wspierać analizy bibliometryczne. Źródło: opracowanie własne na podstawie stron internetowych producentów oprogramowania.....	253
Tabela 58. Zestawienie pytań analitycznych w projektach bibliometrycznych. Źródło: opracowanie własne	259
Tabela 59: Zestawienie pytań badawczych w projektach bibliometrycznych w podziale na stopień trudności przetwarzania danych i wymóg znajomości analizowanej dyscypliny. Źródło: opracowanie własne.	260
Tabela 60: Przydatność poszczególnych technik analiz bibliometrycznych w udzielaniu odpowiedzi na pytania analityczne. Źródło: opracowanie własne.	264
Tabela 61: Przydatność baz bibliograficznych do analiz, mających na celu udzielenie odpowiedzi na konkretne pytania analityczne. Źródło: opracowanie własne.	270
Tabela 62: Przydatność baz patentowych do analiz, mających na celu udzielenie odpowiedzi na konkretne pytania badawcze. Źródło: opracowanie własne.	275
Tabela 63: Możliwości wykorzystania pakietów oprogramowania do analizy danych, pobranych z poszczególnych baz bibliograficznych. Źródło: opracowanie własne.	276
Tabela 64: Liczba publikacji naukowych dotyczących grafenu w kolejnych latach. Źródło: opracowanie własne na podstawie danych pobranych z bazy <i>Scopus</i>	280
Tabela 65: Liczba publikacji naukowych dotyczących grafenu autorstwa poszczególnych naukowców. Źródło: opracowanie własne na podstawie danych pobranych z bazy <i>Scopus</i>	281
Tabela 66: Liczba publikacji dotyczących grafenu publikowana w poszczególnych czasopismach naukowych. Źródło: opracowanie własne na podstawie danych pobranych z bazy <i>Scopus</i>	282
Tabela 67: Najbardziej aktywne instytucje, publikujące wyniki badań nad grafenem. Źródło: opracowanie własne na podstawie danych pobranych z bazy <i>Scopus</i>	286
Tabela 68: Autorzy publikacji na temat grafenu, które opatrzone afiliacją Uniwersytetu Warszawskiego. Źródło: opracowanie własne na podstawie danych pobranych z bazy <i>Scopus</i>	287
Tabela 69: Liczba artykułów i cytowań 24 najbardziej produktywnych uczelni w Polsce z lat 2000-2009. Źródło danych: baza <i>Scopus</i>	292
Tabela 70: Artykuły z 24 uczelni w Polsce, pisane ze współautorami, reprezentującymi zagranicznej instytucje, 2000-2009. Źródło danych: baza <i>Scopus</i>	294
Tabela 71: Liczba publikacji związanych z laserami półprzewodnikowymi na przestrzeni lat 1988–200o. Źródło danych: baza <i>INSPEC</i>	298

Tabela 72: Liczba publikacji związanych z laserami półprzewodnikowymi z poszczególnych krajów na przestrzeni lat 1988–2005. Źródło danych: baza <i>INSPEC</i>	299
Tabela 73: Średni roczny przyrost liczby publikacji związanych z laserami półprzewodnikowymi z poszczególnych krajów na przestrzeni lat 1989–2005. Źródło danych: baza <i>INSPEC</i>	300
Tabela 74: Indeksy RTA w obszarze badań nad laserami półprzewodnikowymi dla poszczególnych krajów. Źródło danych: baza <i>INSPEC</i>	301
Tabela 75: Zmiany indeksów RTA w obszarze badań nad laserami półprzewodnikowymi dla poszczególnych krajów na przestrzeni lat. Źródło danych: baza <i>INSPEC</i>	301
Tabela 76: Publikacje związane z laserami półprzewodnikowymi, reprezentujące badania podstawowe i stosowane w poszczególnych krajach na przestrzeni lat 1989–2005 (grupy nie są rozłączne, udziały nie sumują się do 100%). Źródło danych: baza <i>INSPEC</i>	302
Tabela 77: Publikacje związane z laserami półprzewodnikowymi dotyczące poszczególnych rodzajów związków chemicznych dla wybranych krajów na przestrzeni lat 1989–2005. Źródło danych: baza <i>INSPEC</i>	303
Tabela 78: Publikacje dotyczące poszczególnych rodzajów technologii produkcyjnych dla wybranych krajów. Źródło danych: baza <i>INSPEC</i>	304
Tabela 79: Odsetek publikacji dotyczących laserów półprzewodnikowych, autorstwa pracowników instytucji publicznych i prywatnych na przestrzeni lat 1989–2005. Źródło danych: baza <i>INSPEC</i>	305

Informacja o autorach

Książka została opracowana przez zespół pracowników Zakładu Teorii i Metod Organizacji (TiMO) Wydziału Zarządzania Uniwersytetu Warszawskiego, kierowany przez Krzysztofa Klincewicza.

Wydział Zarządzania UW posiada bogate doświadczenia badawcze i dydaktyczne w obszarze zarządzania i jest najdłużej działającą w Polsce placówką, kształcąca kadry w zakresie zarządzania (nieprzerwanie od 1972 roku). Zajmuje wiodące miejsca ogólnopolskich rankingów edukacji menedżerskiej, posiada 1 kategorię badawczą w ocenie parametrycznej MNiSW, akredytację PKA z wyróżnieniem oraz akredytację międzynarodową AMBA. W ostatnich latach, tematyka zarządzania innowacjami i technologiami, w tym komercjalizacji wyników badań naukowych, odgrywa szczególnie ważną rolę w dydaktyce i badaniach Wydziału Zarządzania UW. W tym zakresie Wydział współpracował w projektach badawczych m.in. z zagranicznymi uczelniami: *London School of Economics and Political Science* (Wielka Brytania), *Tokyo Institute of Technology* (Japonia), *Kyungpook National University* (Korea Południowa) oraz instytucjami polskimi: Ministerstwem Spraw Zagranicznych, Ministerstwem Środowiska oraz Urzędem Patentowym RP.

prof. UW dr hab. Krzysztof Klincewicz - kierownik Zakładu Teorii i Metod Organizacji Wydziału Zarządzania Uniwersytetu Warszawskiego. Pełnomocnik Rektora do spraw Ochrony Własności Intelektualnej Uniwersytetu Warszawskiego. Zajmuje się zarządzaniem technologiami, w szczególności tematyką międzynarodowego transferu technologii, uwarunkowań skutecznej komercjalizacji wyników badań, mechanizmów współpracy partnerskiej na rynkach technologicznych oraz wykorzystania technik bibliometrycznych do analiz technologii i ich dostawców. Prowadził badania dotyczące komercjalizacji technologii informatycznych, telekomunikacyjnych, środowiskowych, optoelektronicznych, nanotechnologii i biotechnologii.

Stypendysta *Collegium Invisibile*, rządu japońskiego i Fundacji na rzecz Nauki Polskiej, posiadacz tytułu zawodowego *chartered marketer* brytyjskiego *The Chartered Institute of Marketing*. W latach 1997-2004 był zatrudniony na stanowiskach doradczych i menedżerskich w firmach zaawansowanych technologii w Polsce, Finlandii i Wielkiej Brytanii. W latach 2004-2006 pracował w *Graduate School of Innovations Management, Tokyo Institute of Technology* w Japonii.

Autor licznych artykułów naukowych i wystąpień konferencyjnych, ekspertyz zleconych m.in. przez Ministerstwo Spraw Zagranicznych, Ministerstwo Nauki i Szkolnictwa Wyższego i Urząd Patentowy RP oraz pięciu książek z dziedziny zarządzania innowacjami i technologiami, w tym: "*Strategic alliances in the hightech industry*" (Berlin 2005), "*Management fashions. Turning bestselling ideas into objects and institutions*" (New Brunswick 2006), "*Polska innowacyjność. Analiza bibliometryczna*" (Warszawa 2008), "*Zarządzanie technologiami. Przypadek niebieskiego lasera*" (Warszawa 2010) i „*Dyfuzja innowacji. Jak odnieść sukces w komercjalizacji nowych produktów i usług*” (Warszawa 2011).

Od 2009 r. kieruje zespołem ekspertów w projekcie Akceleratora Zielonych Technologii *GreenEvo*, prowadzonym przez Ministerstwo Środowiska przy merytorycznym wsparciu

Wydziału Zarządzania Uniwersytetu Warszawskiego i mającym na celu międzynarodowy transfer polskich rozwiązań technologicznych, sprzyjających ochronie środowiska. Od 2011 roku reprezentuje kraje Unii Europejskiej jako członek *Technology Executive Committee* Organizacji Narodów Zjednoczonych.

dr inż. Marcin Żemigala – absolwent Wydziału Organizacji i Zarządzania Politechniki Łódzkiej na kierunku Zarządzanie i Marketing, doktor nauk ekonomicznych w zakresie nauk o zarządzaniu, adiunkt w Zakładzie Teorii i Metod Organizacji Wydziału Zarządzania Uniwersytetu Warszawskiego. Prowadził wykłady między innymi we Francji (*IAE - Institut d'Administration des Entreprises, Université Montpellier 2*), Belgii (*Haute Ecole Libre Mosane, Liege*) i Portugalii (*Instituto Politécnico de Bragança, Escola Superior de Tecnologia e Gestão*).

Do jego zainteresowań naukowych i badawczych należą nowoczesne koncepcje zarządzania, takie jak: społeczna odpowiedzialność biznesu, zarządzanie jakością, zarządzanie wiedzą oraz zarządzanie ekologiczne. Jest autorem lub współautorem ponad 20 artykułów naukowych oraz kilku książek: „*Spoleczna odpowiedzialność przedsiębiorstwa; budowanie zdrowej, efektywnej organizacji*” (Warszawa 2007), „*Jakość w systemie zarządzania przedsiębiorstwem*” (Warszawa 2008), „*Menedżer nowych czasów, najlepsze metody i techniki zarządzania*” (Warszawa 2010), „*Życie i praca współczesnego menedżera, aspekty teoretyczno–empiryczne*” (Warszawa 2010). Reprezentuje Uniwersytet Warszawski jako ekspert w Komitecie Technicznym KT nr 305 ds. społecznej odpowiedzialności przy Polskim Komitecie Normalizacyjnym.

dr Michał Mijał – adiunkt w Zakładzie Teorii i Metod Organizacji Wydziału Zarządzania Uniwersytetu Warszawskiego, doktor nauk ekonomicznych w zakresie nauk o zarządzaniu z bogatym pozauniwersyteckim doświadczeniem zawodowym: tłumacz, analityk biznesowy, kierownik oraz koordynator w różnych firmach i organizacjach, a także trener biznesu. Kierował największym w Polsce działem sprzedaży artykułów multimedialnych w sklepach wielkopowierzchniowych, dostosowywał do warunków polskiego rynku nowoczesne narzędzie informatyczne wspierające modelowanie procesów biznesowych, które jest obecnie stosowane zarówno w praktyce, jak i procesach kształcenia menedżerów, kierował także przez dwa lata wydziałową sekcją międzynarodowej wymiany pracowniczej i studenckiej. Stypendysta rządu Austrii i Szwajcarii, interesuje się zawodowo społecznymi i psychologicznymi aspektami zarządzania, migracjami w zarządzaniu oraz zastosowaniem gier symulacyjnych w procesie dydaktycznym. Z sukcesem wdrożył nowe metody nauczania na przedmiotach kursowych studiów I i II stopnia na Wydziale Zarządzania, wykładał też za granicą: we Włoszech, w Hiszpanii, w Wielkiej Brytanii i w Turcji. Obecnie współtworzy kilka dużych projektów współfinansowanych ze środków Unii Europejskiej i pełni funkcję Koordynatora Krajowego w Centrum Studiów Antymonopolowych i Regulacyjnych Wydziału Zarządzania Uniwersytetu Warszawskiego.

Książka prezentuje możliwości wykorzystania bibliometrii (ilościowych analiz publikacji, w tym artykułów naukowych i dokumentów patentowych) w zarządzaniu technologiami i badaniami naukowymi. Odpowiada na potrzeby osób zarządzających podmiotami w sektorze B+R oraz formułujących politykę naukową lub politykę innowacyjną.

Autorzy pokazują, w jaki sposób wykorzystać techniki bibliometryczne do monitorowania i skanowania technologii, prowadzenia wywiadu technologicznego, ocen istniejących technologii oraz prognozowania ich dalszego rozwoju. W oparciu o opis doświadczeń zagranicznych instytucji administracji publicznej oraz przedsiębiorstw, prezentują najpopularniejsze techniki analityczne, wykorzystywane do celów analiz bazy danych oraz narzędzia informatyczne. Książka zawiera również rekomendacje, dotyczące prowadzenia projektów bibliometrycznych przez polskie instytucje sektora B+R oraz praktyczne przykłady możliwych analiz.

Publikacja powstała w ramach projektu współfinansowanego ze środków UE w ramach projektu systemowego Ministerstwa Nauki i Szkolnictwa Wyższego "Wsparcie systemu zarządzania badaniami naukowymi oraz ich wynikami" (Poddziałanie 1.1.3 Programu Operacyjnego Innowacyjna Gospodarka).