

Polski system innowacji w obszarze informatyki - analiza bibliometryczna

dr Krzysztof Klincewicz, Wydział Zarządzania Uniwersytetu Warszawskiego

Referat na konferencję „Wiedza i innowacje w rozwoju gospodarki”, Kraków, 11.01.2007

Streszczenie

Referat analizuje polski system innowacji w obszarze informatyki, koncentrując się na działalności badawczej polskich specjalistów. W oparciu o bazę międzynarodowych artykułów z obszaru nauk technicznych INSPEC przeprowadzono analizy bibliometryczne, dotyczące: pozycji Polski w porównaniu z innymi krajami w badaniach informatycznych, aktywności poszczególnych polskich ośrodków akademickich i firm prywatnych oraz zainteresowań badawczych polskich informatyków w zestawieniu z globalnymi tendencjami w badaniach stosowanych. Analiza wskazuje na znaczący dorobek polskich informatyków, ale także na istotne słabości w poszczególnych obszarach badań stosowanych. Wyniki badań i rekomendacje mogą stanowić punkt wyjścia do budowy strategii technologicznych polskich przedsiębiorstw informatycznych, jak również przedmiot dyskusji w procesach tworzenia polityki naukowo-technologicznej.

1. Wstęp

Referat omawia specyfikę sektorowego systemu innowacji (SSI) związanego z rozwojem i wdrażaniem oprogramowania, w szczególności relacje pomiędzy instytucjami akademickimi a sektorem prywatnym. Polskie uczelnie podkreślają wyjątkowe kompetencje programistyczne studentów kierunków informatycznych, jednak warto zadać pytanie, czy umiejętności kodowania znajdują uzupełnienie w ambitnej tematyce badań akademickich i w rozwoju innowacyjnych technologii informatycznych. Krytyczne głosy dochodzące z polskich przedsiębiorstw informatycznych sugerują, że kształcimy ekspertów od konstruowania algorytmów i rozwiązywania abstrakcyjnych problemów, którzy na uczelniach mają jednak tylko ograniczony kontakt z najnowszymi technologiami czy popularną w krajach zachodnich problematyką badawczą. Niniejszy referat weryfikuje te opinie, prezentując dane empiryczne, dotyczące polskich zainteresowań badawczych, szczególnych kompetencji, jak również niedoborów wiedzy w obszarze nauk informatycznych. Polityka naukowo-technologiczna większości krajów przywiązuje szczególną wagę do transferu technologii pomiędzy instytucjami badawczymi a przemysłem, jednak wskazane jest zadanie bardziej

fundamentalnego pytania: czy ten wkład akademicki jest dostatecznie atrakcyjny dla firm informatycznych, czy faktycznie jest co transferować? Referat próbuje więc udzielić odpowiedzi na pytanie o możliwość zastosowania w praktyce wyników badań akademickich, stosując metody analityczne z obszaru bibliometrii i podejścia *tech mining*. Analiza koncentruje się na mierzalnych wynikach badań związanych z oprogramowaniem Polski i innych krajów: artykułów naukowych i wystąpień konferencyjnych, zgromadzonych w bazie INSPEC dla 5-letniego okresu 2000-2004. Badania prezentują najaktywniejsze naukowo polskie ośrodki informatyczne, pozycję rodzimych badaczy na tle innych krajów świata, jak również względne polskie przewagi i słabości w obszarze informatycznych badań stosowanych.

2. Sektorowe systemy innowacji

Sektory przemysłu są tradycyjnym obiektem analiz ekonomistów i badaczy zarządzania technologią (Malerba 2002; Carlsson et al. 2002). Sektorowy system innowacji i produkcji (SSI) jest definiowany jako zespół produktów i podmiotów, angażujących się w rynkowe i poza-rynkowe interakcje w celu tworzenia, produkcji i sprzedaży tych produktów (Malerba 2002: 250). Analizy SSI przywiązują szczególną wagę do powiązań instytucjonalnych (zwłaszcza relacji z organizacjami niekomercyjnymi, takimi jak publiczne instytuty badawcze, agendy rządowe lub grupy interesów) oraz do specyficznych procesów tworzenia wiedzy i uczenia się, które przebiegają w sieciach powiązań między organizacjami (Malerba 2002: 250). Metody badawcze związane z systemami innowacji zakładają dynamiczny charakter analizowanych obiektów, uwzględniający interakcje i sprzężenia zwrotne między aktorami, a także występowanie zmian technologicznych, ekonomicznych i społecznych (Carlsson et al. 2002: 237).

Istniejące od lat segmenty sektora mogą ulec zniszczeniu lub zmianom konfiguracji w wyniku zaburzających innowacji lub nowej generacji technologii (Dalum et al. 2002), w skrajnych przypadkach doprowadzając nawet do upadku całych lokalnych sektorów. Ten sam sektor może obejmować wiele nieporównywalnych technologii, przekształconych jednak w bezpośrednio konkurujące ze sobą produkty końcowe (Carlsson et al. 2002: 238) – istnieje więc wiele niewspółmiernych sposobów zaspokajania popytu rynkowego. Użytecznym sposobem strukturalizacji sektorowych systemów innowacji jest model sieci techno-ekonomicznej (Bell, Callon 1994), wprowadzający rozróżnienie między różnymi "biegunami" sieci: finansowym, rynkowym, regulacyjnym i naukowym. Niniejszy referat koncentruje się na biegunie naukowym, w którym najczęściej stosowanym miernikiem efektywności jest

liczba publikacji naukowych.

Naukowe aspekty technologii są istotne nie tylko dlatego, że pomagają rozwijać wiedzę i stymulować nowe odkrycia. Działania w obszarze badań podstawowych pozwalają poszerzać możliwości absorpcji wiedzy przez firmy, przygotowując je do lepszego zrozumienia i wykorzystania technologii (Cohen, Levinthal 1990: 128-129). Często trudno jest też dokonać jednoznacznego rozróżnienia między badaniami podstawowymi i stosowanymi. Wiodące firmy alokują budżety B&R w projekty, mające na celu analizę podstawowych zagadnień technologicznych, a nie tylko rozwój aplikacji. Nie powinno być zaskoczeniem, że wielu naukowców zatrudnionych przez firmy takie jak IBM lub Sony było zdobywcami Nagrody Nobla (Stephan 1996: 1209-1210). Z drugiej strony, uniwersytety aktywnie tworzą centra transferu technologii i firmy odpryskowe (*spin-off*), a czysto naukowa praca badaczy jest uzupełniana kompetencjami menedżerskimi, umożliwiającymi organizacjom przekształcanie specjalistycznej wiedzy w produkty i ich komercjalizację (Michelacci 2003: 208). Istnieją również przypadki, gdy znaczący uczestnicy rynku wycofują się lub ograniczają własny wkład w biegun naukowy - literatura wymienia przykłady firm celowo nie prowadzących własnych badań, lecz raczej uczących się szybkiego wykorzystywania wiedzy pochodzącej z publicznie dostępnych wyników badań innych podmiotów (Hage, Hollingsworth 2000: 973), jak również firm ograniczających własne badania do obszaru aplikacji, celowo utajniających ich zawartość (Lim 2004). Wyniki badań stają się dobrem publicznym tylko w wyniku kodyfikacji i publikacji (Stephan 1996: 1199-1200) - a publikacja w renomowanych magazynach lub prezentacja na konferencji nie jest zwykle priorytetem dla działów badawczo-rozwojowych dużych firm technologicznych. Równocześnie wiele istotnych inicjatyw badawczych nie jest ujętych w oficjalnych statystykach B&R, zwłaszcza gdy prowadzone są przez małe i średnie firmy (Jacobsson, Philipsson 1996: 246), które wykorzystują wyniki badań w swoich produktach, ale nie odzwierciedlają tych wydatków w sprawozdaniach finansowych, ani nie promują naukowych aspektów swoich odkryć.

Historyczne analizy rozwoju przemysłu high-tech w Dolinie Krzemowej wskazują na wkład relacji między instytucjami akademickimi i przemysłem w powstanie konkurencyjnego sektora i utrzymywanie innowacyjności (Castilla et al. 2000: 229-233). W historii Stanów Zjednoczonych, duże przedsiębiorstwa były często ważnymi klientami uniwersytetów, zlecając projekty badawcze i w ten sposób wpływając pośrednio na kierunki badań własnych uczelni (Kodama, Branscomb 1999: 6). Współczesne statystyki transferu technologii dla największych amerykańskich uniwersytetów wydają się imponujące i nieosiągalne dla innych krajów, jednak te doświadczenia są ważnym punktem odniesienia także dla polskiego sektora

oprogramowania.

3. Charakterystyka sektora oprogramowania komputerowego

Analiza sektora oprogramowania wiąże się z wyzwaniami metodologicznymi, wynikającymi z unikalnych cech oprogramowania, procesów jego rozwoju i dyfuzji. Wiele spośród typowych wskaźników, wykorzystywanych przez badaczy systemów innowacji (Carlsson et al. 2002: 242-244) jedynie w ograniczonym stopniu nadaje się do analiz SSI związanego z oprogramowaniem (Young 1996).

Patenty nie są użytecznym wskaźnikiem dla sektora oprogramowania, gdyż większość krajów nie oferuje ochrony patentowej programów komputerowych, a nawet w Stanach Zjednoczonych patentowaniu podlegają jedynie algorytmy i metody o dowiedzionym, radykalnie innowacyjnym charakterze - podczas gdy branże tworzące materialne, fizyczne technologie mogą łatwo chronić także mniej radykalne wynalazki i odmiany produktów. Branżę oprogramowania cechuje także znacząca przepaść między B (badaniami naukowymi) a R (rozwojem aplikacji): sprawny programista nie musi rozumieć ani bezpośrednio korzystać z osiągnięć badań naukowych.

Technologie software'owe są typowym przykładem dobra informacyjnego: kosztownego w produkcji, ale taniego w reprodukcji i ponownym wykorzystaniu (Shapiro, Varian 1999: 3). Innowacje produktowe i procesowe szybko rozpowszechniają się wśród specjalistów IT dzięki standardowym narzędziom programistycznym, systemom i metodykom, które 'opakowują' wiedzę, skracając czas potrzebny na uczenie się nowych technologii i redukując niezbędne inwestycje w infrastrukturę. Sektor cechują więc relatywnie niskie koszty wejścia, którym towarzyszy wysokie ryzyko niezamierzonego 'wycieku' wiedzy z firmy (*knowledge spillover*), szczególnie groźnego wobec braku efektywnych metod ochrony patentowej w większości krajów. Obok nadal powszechnego piractwa w dziedzinie oprogramowania, należy wspomnieć również zaskakująco duże zaangażowanie specjalistów z branży w nielegalną wymianę dokumentów w sieciach *peer-to-peer* i 'podziemnych' stronach internetowych. Programiści uzyskują w ten sposób szybki i bezpłatny dostęp do zaawansowanej literatury technicznej, wykorzystując zdobytą wiedzę dla swoich firm. Kilka lat temu rozwój złożonych aplikacji wymagał ukończenia szkoleń i zakupu kosztownej literatury - dzisiaj znający język angielski specjaliści mogą uzyskać tę samą wiedzę bezpłatnie (choć w nielegalny sposób), a skala zjawiska tej niepożądanego dyfuzji wiedzy jest nieporównywalna z doświadczeniami jakiegokolwiek innej branży.

Popularność offshoringu w rozwoju oprogramowania demonstruje z kolei znaczenie

"innowacji na zasadach outsourcingu": znaczące wynalazki nie są własnością ich rzeczywistych autorów, przykładowo indyjskich firm informatycznych, ale równocześnie firmy te mogą budować kompetencje i umiejętności technologiczne na koszt swoich klientów, aby później uzyskać niezależność. Usługi związane z oprogramowaniem nie wymagają znaczących początkowych inwestycji w B&R, a koszt zdobycia wiedzy może być częściowo pokryty przez zagranicznych partnerów-zleceniodawców, co radykalnie zmienia tradycyjną logikę zarządzania technologiami.

Nie należy wreszcie zapominać o wysoce oligopolistycznej strukturze sektora oprogramowania, w którym dominującą rolę odgrywają standardy technologiczne kontrolowane przez pojedyncze firmy, dążące do uzależnienia klientów (*lock-in*). Nie powinno więc zaskakiwać, że jedynie kilku znaczących graczy-firm software'owych pochodzi z krajów europejskich i azjatyckich i nie może skutecznie konkurować z kilkoma amerykańskimi gigantami, kontrolującymi ponad 50% światowego rynku oprogramowania komputerowego.

4. Metoda badawcza

Prezentowane badania opierają się na metodzie *tech mining*, opisanej przez Portera i Cunninghama (2005), łączącej techniki badawcze bibliometrii i eksploracji tekstów (*text mining*) z metodami ilościowymi i socjometrią. Wykorzystywana metoda analizuje relacje między aktorami i technologiami w obrębie wybranego systemu innowacji w oparciu o dane wejściowe, pochodzące z bazy artykułów naukowych lub patentów. Takie podejście badawcze przypomina wcześniejsze studia innych branż, np. robotyki (Kumaresan, Miyazaki 1999), jak również porównywalne badania sektora oprogramowania w krajach azjatyckich (Klincewicz, Miyazaki 2005).

Badania opierają się na danych pobranych z bazy INSPEC, jednej z najbardziej reprezentatywnych kolekcji recenzowanych artykułów technicznych. INSPEC agreguje streszczenia artykułów z magazynów oraz konferencji związanych z informatyką. Z oczywistych względów nie obejmuje wszystkich istniejących magazynów naukowych, jednak wcześniejsze badania potwierdziły reprezentatywność bazy INSPEC (Porter et al. 2002).

Dane zgromadzono w drodze zapytań do bazy INSPEC, zawierających wybrany kraj jako miejsce pracy autora ("*Author Affiliation*") oraz kod klasyfikacji ("*Classification Code*") 'C' (odpowiadający informatyce oraz powiązanym z nią badaniom matematycznym, które można określić jako badania podstawowe). Zapytania sformułowano dla 5-letniego okresu (2000-2004) - przyczyną wyboru tej względnie krótkiej perspektywy są dramatyczne przemiany

globalnej branży informatycznej. Wyniki badań stosowanych sprzed 10 lat mogą dziś nie mieć już wartości praktycznej – a analizowany okres lat 2000-2004 charakteryzuje dominacja systemu Microsoft Windows jako platformy dla przedsiębiorstw oraz dynamiczny rozwój Internetu i aplikacji mobilnych, podczas gdy te tematy nie były dostatecznie interesujące dla badaczy przed rokiem 2000.

Uzyskane dane były czyszczone i wstępnie obrabiane przy wykorzystaniu programowalnych *regular expressions*, a następnie analizowane w arkuszu kalkulacyjnym *OpenOffice Calc* oraz programach do analizy sieci *NetDraw* i *KeyPlayer*.

5. Porównania międzynarodowe

Prezentację wyników badań rozpoczyna ogólne porównanie publikacji związanych z oprogramowaniem, pochodzących z różnych krajów świata. Tabela 1 zawiera informacje o liczbie artykułów dotyczących oprogramowania komputerowego i względnym znaczeniu tej tematyki w każdym z krajów (udziału artykułów w całości publikacji naukowych kraju, zarejestrowanych w bazie INSPEC). Polska zajmuje 16 miejsce w tym międzynarodowym rankingu publikacji, a ponad jedna czwarta polskich publikacji naukowych w bazie INSPEC jest powiązana z informatyką.

Tabela 1. Liczba publikacji związanych z informatyką z lat 2000-2004.

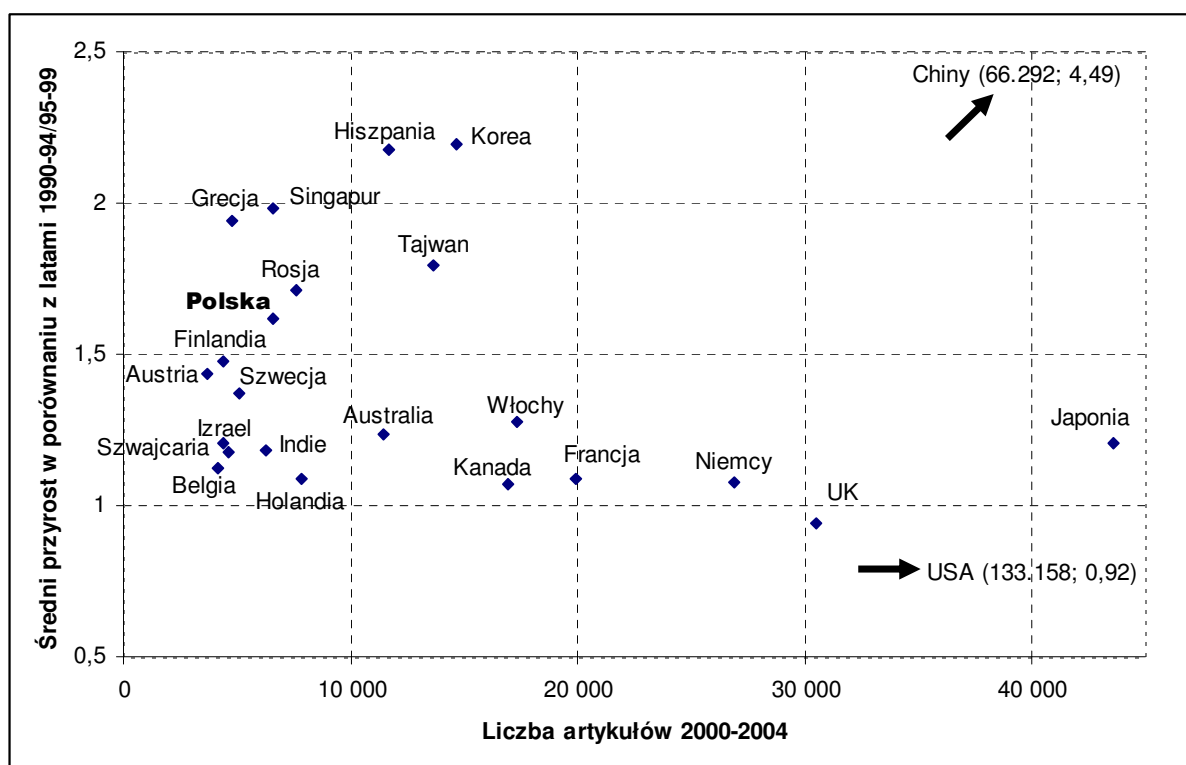
Kraj	Publikacje informatyczne (2000-2004)	% wszystkich publikacji (2000-2004)			
			17. Singapur	6,600	40.53%
			18. Szwecja	5,090	28.70%
1. USA	133,158	33.72%	19. Szwajcaria	4,619	24.91%
2. Japonia	43,616	24.96%	20. Izrael	4393	32.23%
3. Chiny	66,292	39.30%	21. Belgia	4,116	30.17%
4. Wielka Brytania	30,476	35.69%	22. Finlandia	4,372	39.45%
5. Niemcy	26,878	25.82%	23. Grecja	4,803	45.69%
6. Francja	19,945	27.91%	24. Austria	3,705	37.21%
7. Kanada	16944	40.36%	25. Dania	2,258	31.04%
8. Włochy	17,345	30.90%	26. Meksyk	3,030	27.20%
9. Australia	11434	44.24%	27. Portugalia	2,850	38.81%
10. Tajwan	13,678	40.39%	28. Węgry	1,833	33.72%
11. Korea	14,653	28.74%	29. Norwegia	1,701	38.98%
12. Hiszpania	11,715	36.19%	30. Rumunia	1,769	29.46%
13. Holandia	7,845	35.51%	31. Czechy	2,240	30.34%
14. Indie	6,277	19.23%	32. Słowenia	1,870	44.04%
15. Rosja	7,595	11.90%	33. Irlandia	1,625	40.32%
16. Polska	6,586	26.67%	34. Ukraina	1,546	10.45%
			35. Bułgaria	826	25.82%

36. Chorwacja	1,083	42.67%	42. Łotwa	233	24.63%
37. Słowacja	865	30.51%	43. Estonia	249	25.15%
38. Malezja	1,292	50.12%	44. Luksemburg	37	53.62%
39. Tajlandia	1,172	46.81%	45. Wietnam	77	29.06%
40. Argentyna	789	17.01%	46. Brazylia	8	34.78%
41. Litwa	393	22.50%			

Źródło: badania własne w oparciu o bazę INSPEC

Porównanie osiągnięć badawczych z lat 2000-2004 z dwoma poprzednimi pięcioletnimi okresami pozwala na identyfikację tendencji rozwojowych w tematyce badawczej poszczególnych krajów. W Stanach Zjednoczonych i Wielkiej Brytanii informatyka od lat stanowiła istotny obszar badań (w obu krajach odnotowano nieznaczny spadek liczby artykułów w analizowanym okresie), podczas gdy rozwijające się gospodarki Chin, Malezji, Tajlandii czy Wietnamu odnotowały ponad czterokrotny przyrost, a niektóre kraje doceniające znaczenie informatyki (w tym Korea, Irlandia, Czechy, Chorwacja i Estonia) – wzrost ponad dwukrotny. Polskę podobnie jak większość krajów europejskich charakteryzuje umiarkowany wzrost (liczba publikacji w latach 2000-2004 stanowiła 162% średniej wartości dla okresów 1990-1994 i 1995-1999). Wykres 1 prezentuje graficznie omawiane zależności.

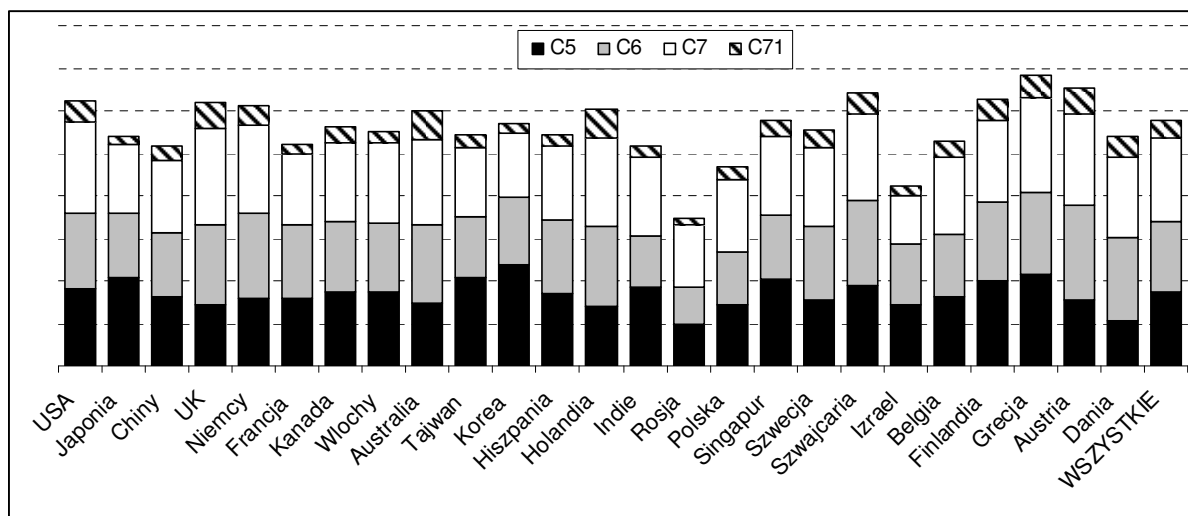
Wykres 1. Liczba artykułów informatycznych z lat 2000-2004 a średni przyrost ich liczby w stosunku do okresów 1990-94 i 1995-99.



Źródło: badania własne w oparciu o bazę INSPEC

Baza INSPEC oferuje własny system klasyfikacji artykułów w oparciu o hierarchię zagadnień badawczych – przez odwołanie się do alfanumerycznego kodu można więc zidentyfikować na wysokim poziomie ogólności temat artykułu. Wykres 2 prezentuje aktywność badawczą poszczególnych krajów w obszarach badań stosowanych dotyczących: sprzętu komputerowego i oprogramowania wspierającego jego wykorzystanie (kod C5), programowania (kod C6), aplikacji użytkowych (kod C7) oraz podzbioru aplikacji użytkowych, dedykowanych dla biznesu (kod C71). Ze względu na możliwość przypisania do tego samego artykułu kilku kodów w bazie INSPEC, wartości nie powinny być sumowane. Udziały poszczególnych obszarów tematycznych w polskich badaniach pozostają na poziomach niższych niż dla większości krajów zachodnich. 28,33% polskich publikacji dotyczyło sprzętu komputerowego (C5), 25,24% - programowania (C6), 34,16% - aplikacji (C7) i 5,72% - aplikacji biznesowych (C71). Dla porównania, niemieckie wyniki to: 31,36% dla C5, 40,49% dla C6, 41,15% dla C7 i 9,19% dla C71.

Wykres 2. Udziały poszczególnych obszarów badawczych w publikacjach informatycznych analizowanych krajów.



Źródło: badania własne w oparciu o bazę INSPEC

Powyzsze analizy wskazują na znaczącą międzynarodową pozycję Polski w badaniach informatycznych – ujawniają jednak także koncentrację na badaniach podstawowych a nie stosowanych, jak również słabszą dynamikę rozwoju aktywności badawczej w porównaniu z krajami aktywnie inwestującymi w rozwój technologii informacyjnych i komunikacyjnych.

6. Polskie ośrodki badań informatycznych

Kolejny etap analizy dotyczy specyfiki polskich badań informatycznych – ich dystrybucji geograficznej, udziału placówek nie-akademickich, w tym szczególnie firm prywatnych w

tworzeniu wiedzy, oraz identyfikacji wiodących ośrodków i badaczy. W oparciu o artykuły pobrane z bazy INSPEC, których autorzy deklarowali polskie afiliacje, przygotowano rankingi, wskazujące na znaczące centra rozwoju wiedzy informatycznej w Polsce.

Wykres 3. Najważniejsze centra tworzenia wiedzy informatycznej w Polsce (w nawiasach liczba artykułów informatycznych).



Źródło: badania własne w oparciu o bazę INSPEC

Wykres 3 wskazuje najważniejsze miasta, z których pochodziły analizowane artykuły. Zdecydowany prym w skali kraju wiedzie Warszawa (1847 artykuły), inne bardzo aktywne ośrodki to: Wrocław (763), Aglomeracja Katowicka (722 – większość publikacji z Gliwic), Kraków (686), Poznań (587) i Trójmiasto (513). Graficzna prezentacja dowodzi jednocześnie istnienia ogromnych obszarów „informatycznych pustyń” w naszym kraju, pokrywających się z najuboższymi polskimi regionami.

Tabela 2 prezentuje ranking najbardziej aktywnych w badaniach informatycznych instytucji w Polsce – nie powinno być zaskoczeniem, że najwyższe pozycje zajmują uczelnie wyższe i oddziały Polskiej Akademii Nauk. Zaskakiwać może jednak relatywnie niewielkie znaczenie rozpoznawanych na świecie uczelni, uważanych za kuźnie kadr informatycznych, w tym Uniwersytetu Warszawskiego i Uniwersytetu Jagiellońskiego. Działalność dydaktyczna i doskonalenie procesów kształcenia programistów niekoniecznie idą więc w parze z aktywnym tworzeniem nowej wiedzy. Oba uniwersytety posiadają też znacznie mniejsze

zasoby kadrowe niż uczelnie politechniczne, które badania powiązane z informatyką i programowaniem prowadzą także na innych kierunkach studiów technicznych.

Tabela 2. Polskie instytucje najaktywniejsze w badaniach informatycznych.

Miejsce	Instytucja	Liczba artykułów
1	Politechnika Warszawska	874
2	Politechnika Wroclawska	608
3	Politechnika Śląska (Gliwice)	596
4	Akademia Górniczo-Hutnicza (Kraków)	523
5	Politechnika Poznańska	439
	<i>Polska Akademia Nauk (cała Polska)</i>	432
6	Polska Akademia Nauk (Warszawa)	336
7	Politechnika Gdańska	315
8	Politechnika Łódzka	283
9	Uniwersytet Warszawski	242
10	Uniwersytet w Zielonej Górze	115
11	Politechnika Szczecińska	114
12	Politechnika Białostocka	113
13	Uniwersytet Wrocławski	110
14	Politechnika Rzeszowska	105
15	Politechnika w Zielonej Górze	99
16	Akademia Morska (Gdynia)	96
17	Uniwersytet Jagielloński (Kraków)	96
18	Politechnika Częstochowska	92
19	Wojskowa Akademia Techniczna (Warszawa)	75
20	Uniwersytet Adama Mickiewicza (Poznań)	62

Źródło: badania własne w oparciu o bazę INSPEC

Spośród 6.586 polskich artykułów informatycznych, tylko 339 powstało w placówkach nieakademickich, w tym instytutach przemysłowych i firmach prywatnych (5,15%), a jedynie 95 w przedsiębiorstwach (1,44%), głównie w polskich oddziałach firm międzynarodowych. Tak niskie zaangażowanie sektora prywatnego w tworzenie wiedzy informatycznej może świadczyć o postrzeganiu przez firmy informatyki jako działalności odtwórczej, rzemiosła sprowadzającego się do wdrażania rozwiniętych przez innych systemów – a nie wykorzystywania tej specjalistycznej dziedziny wiedzy czy wręcz „algorytmicznej sztuki” do rozwoju własnych innowacji.

Rozczarowująco niskie jest zaangażowanie lokalnych firm informatycznych w tworzenie wiedzy naukowej - spośród znaczących informatycznych spółek w Polsce, jedynie Computerland i InfoVIDE znalazły się wśród autorów analizowanych publikacji. To sugeruje

wczesny etap rozwoju sektora informatycznego, koncentrującego się na zaspokajaniu nadal znaczącej luki w popycie wewnętrznym, korzystającego z przewag kosztowych w porównaniu z zachodnimi konkurentami i nie muszącego jeszcze wypracowywać unikalnych rozwiązań i wyróżników na rynkach międzynarodowych. W krajach zachodnich firmy informatyczne wykorzystują badania naukowe jako dodatkowy sposób poznawania nowych technologii, wypracowywania innowacyjnych rozwiązań, a także budowania wizerunku nowoczesnej, zaawansowanej technologicznie firmy.

Tabela 3. Najaktywniejsze badawczo polskie organizacje nie-akademickie.

Instytucja	Liczba artykułów
Przemysłowy Instytut Elektroniki (Warszawa)	25
Poznańskie Centrum Superkomputerowo-Sieciowe (Poznań)	19
Instytut Fizyki Jądrowej (Kraków)	18
Instytut Energii Atomowej (Warszawa i Otwock-Świerk)	15
Przemysłowy Instytut Telekomunikacji (Warszawa)	14
Instytut Technologii Elektronowej (Warszawa)	14
ABB (Kraków, Gdańsk i Łódź)	12
Przemysłowy Instytut Automatyki i Pomiarów (Warszawa)	10
Instytut Techniki i Aparatury Medycznej (Zabrze)	10
Siemens (Warszawa)	9

Źródło: badania własne w oparciu o bazę INSPEC

Tabela 4. Najaktywniejsze badawczo przedsiębiorstwa.

Instytucja	Liczba artykułów
ABB (Kraków, Gdańsk i Łódź)	12
Siemens (Warszawa)	9
Alcatel (Warszawa)	5
InfoVIDE (Warszawa)	4
PKP (Warszawa)	4
Lucent Technologies (Warszawa)	4
Computerland (Wrocław i Warszawa)	3
Polska Telefonii Cyfrowa (Warszawa)	3
ADB (Zielona Góra)	3
Ponetex-Logistics (Rakoniewice)	2
Aldec-ADT (Gliwice)	2
Telekomunikacja Polska (Warszawa)	2
Rodan Systems (Warszawa)	2
APATOR Control (Toruń)	2
BRE Bank (Warszawa)	2

Źródło: badania własne w oparciu o bazę INSPEC

Tabela 5. Najaktywniejsi polscy badacze w dziedzinie informatyki.

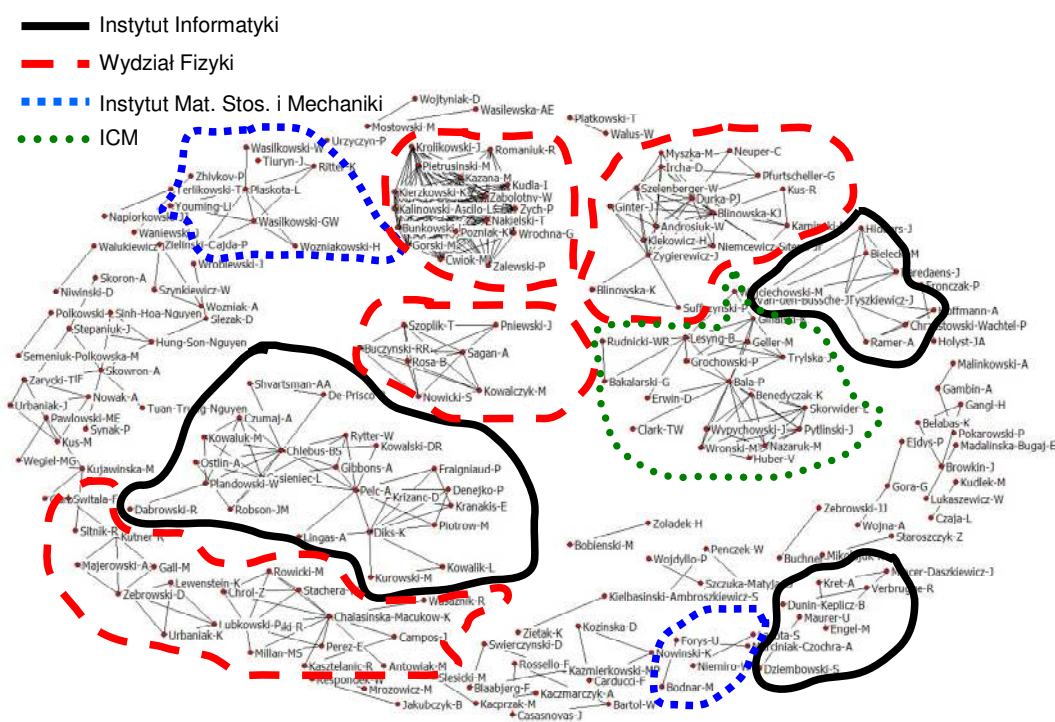
	Autor	Liczba artykułów	Uczelnia macierzysta
1	prof. dr hab. Tadeusz Kaczorek	48	Politechnika Warszawska
2	prof. dr hab. Andrzej Napieralski	47	Politechnika Łódzka
3	dr hab. Bogdan Smołka	42	Politechnika Śląska
4	prof. dr hab. Ryszard Tadeusiewicz	37	AGH Kraków
5	prof. dr hab. Andrzej Świerniak	31	Politechnika Śląska
6	prof. dr hab. Krzysztof Gałkowski	30	Uniwersytet w Zielonej Górze
7	dr Marian Bubak	30	AGH Kraków
8	prof. Dr hab. Janusz Kacprzyk	29	PAN Warszawa
9	prof. dr hab. Zdzisław Bubnicki	28	PAN / Politechnika Wrocławska
10	prof. dr hab. Marian P. Kaźmierkowski	27	Politechnika Warszawska

Źródło: badania własne w oparciu o bazę INSPEC

Zgromadzone dane umożliwiły opracowanie rankingu najaktywniejszych polskich badaczy w obszarze informatyki (tabela 5), a także wykorzystanie socjometrii i metod sieciowych do analizy współpracy między badaczami. Wykres 4 przedstawia przykładową sieć zależności współautorów publikacji informatycznych z Uniwersytetu Warszawskiego, zbudowaną w oparciu o mechanizmy skalowania wielowymiarowego – zakreślone liniami skupienia to grupy bezpośrednio współpracujących badaczy, wspólnie prowadzących badania i publikujących artykuły związane z informatyką.

Do dalszej analizy danych wykorzystano aplikację *KeyPlayer*, która stosuje algorytm poszukiwania węzłów, których eliminacja spowoduje rozkład sieci. W ten sposób zidentyfikowano jednostki niezbędne dla poprawnej współpracy w ramach istniejących struktur. Dla Uniwersytetu Warszawskiego i badań informatycznych są to: profesorowie Piotr Bała (Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego), Katarzyna Chałasińska-Macukow (Wydział Fizyki), Bogdan Stanisław Chlebus (Instytut Informatyki), Andrzej Pelc (Instytut Informatyki) oraz dr Jarosław Żygierewicz (Wydział Fizyki). W przypadku publikacji naukowych, niekoniecznie są to najbardziej aktywni badacze – ale przede wszystkim osoby najaktywniej angażujące się w pracę zespołową.

Wykres 4. Sieć powiązań autorów publikacji informatycznych z Uniwersytetu Warszawskiego.



Źródło: badania własne w oparciu o bazę INSPEC i program NetDraw

7. Informatyczne badania stosowane w Polsce i na świecie

Podsumowujące prezentowane badania analizy dotyczą zróżnicowania tematycznego prowadzonych badań, w szczególności badań stosowanych, których wyniki są najbardziej podatne na transfer technologii i ewentualną komercjalizację. Porównano aktywność badawczą w poszczególnych krajach dla wybranych obszarów badawczych, wyodrębnionych w oparciu o specjalistyczną literaturę. Do każdego obszaru przypisano listę słów kluczowych (hasła i ich odmiany, dodatkowo poddanych lematyzacji), a następnie wyszukano w bazie INSPEC dokumenty, zawierające w tytule lub streszczeniu wskazane terminy, w rozbiciu na poszczególne analizowane kraje. Dla lepszego zrozumienia pozycji Polski w poszczególnych obszarach badawczych oraz identyfikacji polskich przewag i słabości, obliczono wskaźniki RTA (ujawniona przewaga technologiczna, *revealed technological advantage*). Analiza RTA nawiązuje do teorii przewag komparatywnych w handlu międzynarodowym, postulującej koncentrację poszczególnych krajów-partnerów wymiany na obszarach, stanowiących ich względne przewagi w porównaniu z globalną produkcją. W przypadku technologii, analizie podlegają publikowane artykuły techniczne lub zdobyte patenty. Najczęściej stosowana formuła obliczania RTA to:

$$RTA = \frac{(P_{ij} / \sum_j P_{ij})}{(\sum_j P_{ij} / \sum_i \sum_j P_{ij})} , \text{ gdzie: } P_{ij} = \text{liczba publikacji w obszarze } i \text{ napisana w kraju } j$$

Wartość RTA powyżej 1.0 oznacza istnienie przewagi komparatywnej – z kolei wartości indeksu poniżej jedności to słabości analizowanego podmiotu. Tabela 6 prezentuje wartości indeksu RTA dla Polski i wybranych krajów, ujawniając raczej słabości niż przewagi naszego kraju. Względne przewagi polskiego dorobku badawczego widoczne są jedynie w obszarach projektowania wspieranego komputerowo (CAD) oraz szeroko rozumianych technologii sieciowych (zagadnieniu bardziej związanym z infrastrukturą fizyczną niż programowaniem). Alarmujące są z kolei polskie słabości w obszarach związanych z najnowszymi technologiami, zwłaszcza przeznaczonymi dla przedsiębiorstw i szczególnie podatnymi na transfer i komercjalizację, w tym: badaniach sieci Internet i powiązanych z nią technologii portalowych i handlu elektronicznego; bezpieczeństwa danych i sieci komputerowych; *embedded software* tworzonego do sterowania specjalistycznymi maszynami i urządzeniami.

Tabela 6. Indeks RTA dla wybranych obszarów badań stosowanych w informatyce.

	Polska	USA	Japonia	Niemcy	Chiny	Indie	Rosja
Projekty informatyczne	0.84	1.10	0.79	1.03	0.87	1.01	0.61
Integracja	0.57	1.11	0.66	1.57	0.91	0.79	0.45
Rozwój aplikacji	0.84	1.09	0.87	1.12	0.87	1.00	0.82
Open source	0.57	1.30	0.55	1.17	0.64	1.02	0.30
Embedded software	0.37	1.16	0.79	1.21	0.76	1.08	0.18
Sieci komputerowe	1.01	0.98	1.04	0.86	1.12	1.25	0.51
Internet	0.45	1.22	0.81	0.89	0.64	0.81	0.24
E-business i portale	0.24	1.10	0.44	1.11	0.67	1.02	0.14
Web 2.0	0.24	1.18	0.79	1.27	0.55	0.59	0.32
Technologie mobilne	0.57	1.00	1.26	1.21	0.68	0.77	0.18
Bezpieczeństwo	0.49	1.07	0.73	0.83	1.00	0.98	0.31
Kontrola dostępu i szyfrowanie	0.42	0.86	0.97	0.88	1.02	1.10	0.23
Ochrona systemów	0.31	1.58	0.58	0.25	1.33	0.75	0.49
Multimedia	0.69	1.07	1.22	1.06	0.88	0.87	0.59
Video i obraz	0.68	1.07	1.31	1.04	0.97	0.82	0.58
Dźwięk i mowa	0.64	0.99	1.58	1.20	0.68	1.13	0.35
Gry i wirtualna rzeczywistość	0.73	1.02	1.33	1.27	0.80	0.52	0.93
Kompresja	0.99	1.16	0.85	0.70	1.02	1.48	0.49
Systemy dla przedsiębiorstw	0.65	0.92	0.40	1.51	1.12	1.06	0.29
Eksploatacja	0.75	1.02	1.10	0.93	0.97	0.99	0.54

danych							
Projektowanie wspierane komputerowo	1.04	0.96	0.82	0.89	1.53	1.14	0.77

Źródło: badania własne w oparciu o bazę INSPEC

8. Wnioski

Referat przedstawił analizę polskiego dorobku badawczego w obszarze informatyki. Polska zajmuje 16 miejsce na świecie pod względem ilości publikacji badawczych tworzonych w latach 2000-2004. Ponad 26% wszystkich polskich publikacji w bazie INSPEC jest związana z zagadnieniami informatycznymi – co oznacza docenianie przez rodzimych badaczy tej tematyki, choć jest jednocześnie słabszym wynikiem niż w innych aktywnych badawczo krajach (z wyjątkiem Rosji i Ukrainy). Dynamika wzrostu liczby publikacji informatycznych w Polsce jest duża, w latach 2000-2004 stworzono 1,62 razy więcej publikacji niż średnio w dwóch wcześniejszych pięcioleciach - wynik ten jest jednak niższy niż w niektórych zorientowanych na informatykę gospodarkach takich jak Singapur, Korea czy Chiny. Analiza klasyfikacji artykułów w bazie INSPEC wykazała niższy niż dla innych rozwiniętych gospodarek udział badań stosowanych – relatywnie niewiele polskich artykułów dotyczy problemów programistycznych, a szczególnie słabo rozwinięte są badania nad aplikacjami biznesowymi, co może ograniczać potencjał transferu wyników badań technologii informatycznych do sektora prywatnego.

Dystrybucja geograficzna artykułów w Polsce wskazuje na ogromną rolę dużych ośrodków badawczych takich jak Warszawa, Wrocław, Aglomeracja Katowicka, Kraków, Poznań i Trójmiasto, przy jednoczesnym występowaniu „pustyń informatycznych” w regionach peryferyjnych. Ranking aktywności badawczej polskich instytucji wykazał duże znaczenie uczelni politechnicznych oraz dużo skromniejszych wymiernych efektach badań wiodących uniwersytetów. Słabo zaangażowane w badania informatyczne są działające w Polsce przedsiębiorstwa; jedynie dwie znane polskie firmy informatyczne przygotowały w analizowanym okresie publikacje badawcze zindeksowane w bazie INSPEC. Badania podsumowała kalkulacja indeksu RTA, określająca słabości Polski w badaniach stosowanych, w szczególności związanych z Internetem i bezpieczeństwem danych. Względne polskie przewagi wiążą się z badaniami projektowania wspieranego komputerowo i infrastrukturalnych aspektów sieci komputerowych.

Należy pamiętać o ograniczeniach interpretacji prezentowanych badań. Wątpliwość może budzić milczące założenie o tym, że ilość publikacji odpowiada jakości osiągnięć krajów,

ośrodków lub pojedynczych badaczy – zestawienia liczby artykułów bez uwzględniania ich znaczenia (np. cytowalności i rankingu ważności w oparciu o tzw. *impact factor*) mogą prowadzić do nieuzasadnionych wniosków. Na poziomie zagregowanej analizy tysięcy artykułów dla całego kraju koncentracja na ilości jest już jednak typowa dla podejścia bibliometrycznego. Kolejnym problemem mogą być ewentualne błędy klasyfikacji bazy INSPEC – przykładowo, niektóre artykuły kategorii ‘C’ mogą nie dotyczyć informatyki i być niewłaściwie do niej przypisane. Przy tysiącach analizowanych artykułów, ewentualne jednostkowe pomyłki twórców INSPEC-a nie powinny jednak wpłynąć znacząco na interpretacje danych. Baza INSPEC wymienia tylko jedną afiliację autorów każdego artykułu. Możliwe jest więc pominięcie wielu artykułów, w których powstawanie byli zaangażowani pojedynczy polscy badacze i wieloosobowe zespoły z naszego kraju, a których pierwszy autor pracuje za granicą. Analogiczny błąd może być popełniony w przypadkach przygotowywania artykułów wspólnie przez badaczy akademickich i pracowników firm – możliwe jest pominięcie wymiernych osiągnięć badawczych sektora przedsiębiorstw, które współpracują z uczelniami. Ostatnia wreszcie wątpliwość dotyczy wykorzystania wyszukiwania pełnotekstowego do analizy tematyki badań stosowanych. Takie podejście jest ryzykowne, gdyż niektóre odnalezione terminy mogą występować w innych niż oczekiwane kontekstach, jednak przy analizach dużych zbiorów tekstowych, zastosowanie techniki *text mining* wydaje się jednak być uzasadnione ekonomicznie.

Wyniki zaprezentowanych badań oferują cenne wskazówki dla twórców polityki naukowo-technologicznej, w szczególności dotyczące niedocenianych przez lokalnych badaczy obszarów informatyki i lokalizacji „pustyń informatycznych”. Wnioski mogą być także przydatne dla samych badaczy, sugerując im możliwe oryginalne tematy badawcze i sposoby na wyróżnienie się w skali krajowej i międzynarodowej. Z przedstawionych analiz skorzystać mogą wreszcie firmy informatyczne, poszukujące możliwości współpracy badawczej, kompetentnych i znających najnowsze technologie pracowników lub obszarów rozwoju nowych produktów, które mogą znaleźć wsparcie w polskim dorobku naukowym.

Bibliografia

- Bell, G., M. Callon. 1994. Techno-Economic Networks and Science and Technology Policy. *OECD STI Review*, 14, s. 59-118
- Carlsson, B., Jacobsson, S., Holmen, M., A. Rickne. 2002. Innovation systems: analytical and methodological issues. *Research Policy*, 31, s. 233-245.
- Castilla, E.J., Hwang, H., Granovetter, E., M. Granovetter. 2000. Social Networks in Silicon Valley. [w:] Lee, Ch-M., Miller, W.F., Hancock, M.G., H.S. Rowen. *The Silicon Valley Edge*:

- A Habitat for Innovation and Entrepreneurship*. Stanford University Press, Stanford, s. 218-247.
- Cohen, W.M., D.A. Levinthal. 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35, s. 128-152.
- Dalum, B., Pedersen, Ch.O.R., G. Villumsen. 2002. Technological Life Cycles: Regional Clusters Facing Disruption. Danish Research Unit for Industrial Dynamics. *DRUID Working Paper* 2002-10.
- Hage, J., J.R. Hollingsworth. 2000. A Strategy for the Analysis of Idea Innovation Networks and Institutions. *Organization Studies*, 21, 5, s. 971-1004.
- Jacobsson, S., J. Philipson. 1996. Sweden's technological profile. What can R&D and patents tell and what do they fail to tell us? *Technovation*, 26, 5, s. 245-253.
- Klincewicz, K., K. Miyazaki. 2005. *Software Systems of Innovation in Asia. Empirical Analysis of Industry and Academia Research Activities*. STEPI International Symposium, Seoul.
- Kodama, F., L.M. Branscomb. 1999. University Research as an Engine for Growth: How Realistic Is the Vision? [w:] Branscomb, L.M., Kodama, F., R. Florida (red.). *Industrializing Knowledge. University-Industry Linkages in Japan and the United States*. MIT Press, Cambridge, Massachusetts, s. 3-19.
- Kumaresan, N., Miyazaki, K. 1999. An Integrated network approach to systems of innovation - the case of robotics in Japan. *Research Policy*, 28, s. 563-585
- Lim, K. 2004. The relationship between research and innovation in the semiconductor and pharmaceutical industries (1981-1997). *Research Policy*, 33, 2, s. 287-321
- Malerba, F. 2002. Sectoral systems of innovation and production. *Research Policy*, 31, s. 247-264
- Michelacci, C. 2003. Low Returns in R&D due to the Lack of Entrepreneurial Skills. *The Economic Journal of the Royal Economic Society*, 113, s.207-225.
- Porter, A.L., Rossner, J.D., Jin, X.-Y., N.C. Newman (2002) Measuring national 'emerging technology' capabilities. *Science and Public Policy*, 29, 3.
- Porter, A.L., S.W. Cunningham. 2005. *Tech Mining. Exploiting New Technologies for Competitive Advantage*. New Jersey: Wiley-Interscience.
- Shapiro, C., H.R., Varian. 1999. *Information Rules. A Strategic Guide to the Network Economy*. Harvard Business School Press, Boston, Massachusetts.
- Stephan, P.E. 1996. The Economics of Science. *Journal of Economic Literature*, 34, s. 1199-1235.
- Young, A. 1996. Measuring R&D in Services. *OECD STI Working Paper*, OCDE/GD(96)132.