

Uczelniane Bazy Wiedzy w projekcie SYNAT

Doświadczenia Politechniki Warszawskiej

Jakub Koperwas, Łukasz Skonieczny, Marek Kozłowski,
Piotr Andruszkiewicz, Henryk Rybiński, Wacław Struk

Instytut Informatyki, Politechnika Warszawska

Agenda

1. Jak nas widzą ...
2. Koncepcja infrastruktury SYNAT – kilka słów
3. Repozytoria... czy raczej bazy wiedzy?
4. Funkcjonalność uczelnianej bazy wiedzy w systemie SYNAT
5. Metody sztucznej inteligencji w systemie Ω - Ψ^R
6. Doświadczenia Politechniki Warszawskiej

Jak nas widzą?

... np. w Webometrics

RANKING WEB OF REPOSITORIES

UniversitiesHospitalsBusiness SchoolsResearch Centers

SEARCH

HOME NORTH AMERICA LATIN AMERICA EUROPE ASIA AFRICA ARAB WORLD OCEANIA BY AREAS TOP INSTITUTIONALS TOP PORTALS

[Home](#) » Poland

About Us

- About Us
- Contact Us

About the Ranking

- Methodology
- Objectives

Resources

- Best Practices
- Links



www.mavir.net



Poland

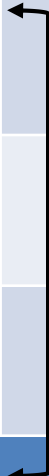
ranking	World Rank ▲	Instituto	Size	Visibility	Files Rich	scholar
1	278	Biblioteka Cyfrowa Politechniki Krakowskiej / Digital Library of Cracow University of Technology	456	523	495	176
2	396	Biblioteka Cyfrowa Uniwersytetu Wrocławskiego / Digital Library of Wrocław University	421	120	712	1015
3	400	Kujawsko Pomorska Digital Library University Nicolaus Copernicus	208	54	1095	1112
4	431	Biblioteki Cyfrowej Politechniki Warszawskiej / Warsaw University of Technology Digital Library	933	206	355	959
5	485	Academic Repository Adam Mickiewicz University in Poznan	803	749	583	354
6	658	Lublin University of Technology Digital Library	618	663	1170	587
7	688	Repozytorium Electrical Engineering Wrocław Polytechnics	639	701	881	840
8	951	Digital Library of Polish Institute of Anthropology	499	941	656	1117
9	1037	e-Biblioteka Uniwersytetu Warszawskiego / University of Warsaw e-bUW	215	255	1235	1300
10	1078	Biblioteka Cyfrowa Uniwersytetu Warmińsko-Mazurskiego / University of Warmia and Mazury Digital Library	1171	994	1119	1055

Jak nas widzą?

... Google Scholar

Jeżeli czegoś nie ma w internecie – nie istnieje

Zapytanie	Liczba odpowiedzi w Google scholar	Liczba odpowiedzi w Microsoft Academia
„rozprawa doktorska” „Politechnika Warszawska”	299	0
"Ph D thesis" "Politechnika Warszawska"	91	0
"Ph D thesis"	2.530.000	1237
"Ph D thesis" "Czech Technical University"	1320	0



Jak nas widzą? (nota bene)

... Google Scholar

WARSAW UNIVERSITY OF
TECHNOLOGY

Institute of Electrical Machines

[PDF] [WARSAW UNIVERSITY OF TECHNOLOGY](#)

[T Wichert - qucosa.de](#)

(Abstract) Information on the structure of considered drive system with reluctance motor, principle of operation and mathematical model of the motor, principle of the energy conversion in analysed system and elementary types of supply devices have been ...

[Related articles](#) [View as HTML](#) [All 2 versions](#) [Import into BibTeX](#)

[PDF] [from qucosa.de](#)



Ph.D. THESIS

Torsten Wichert, M.Sc.

**Design and Construction Modifications of
Switched Reluctance Machines**

Supervisor

Paweł Staszewski, Ph.D., D.Sc.

Hans Kuß, Prof. Dr.-Ing. habil.

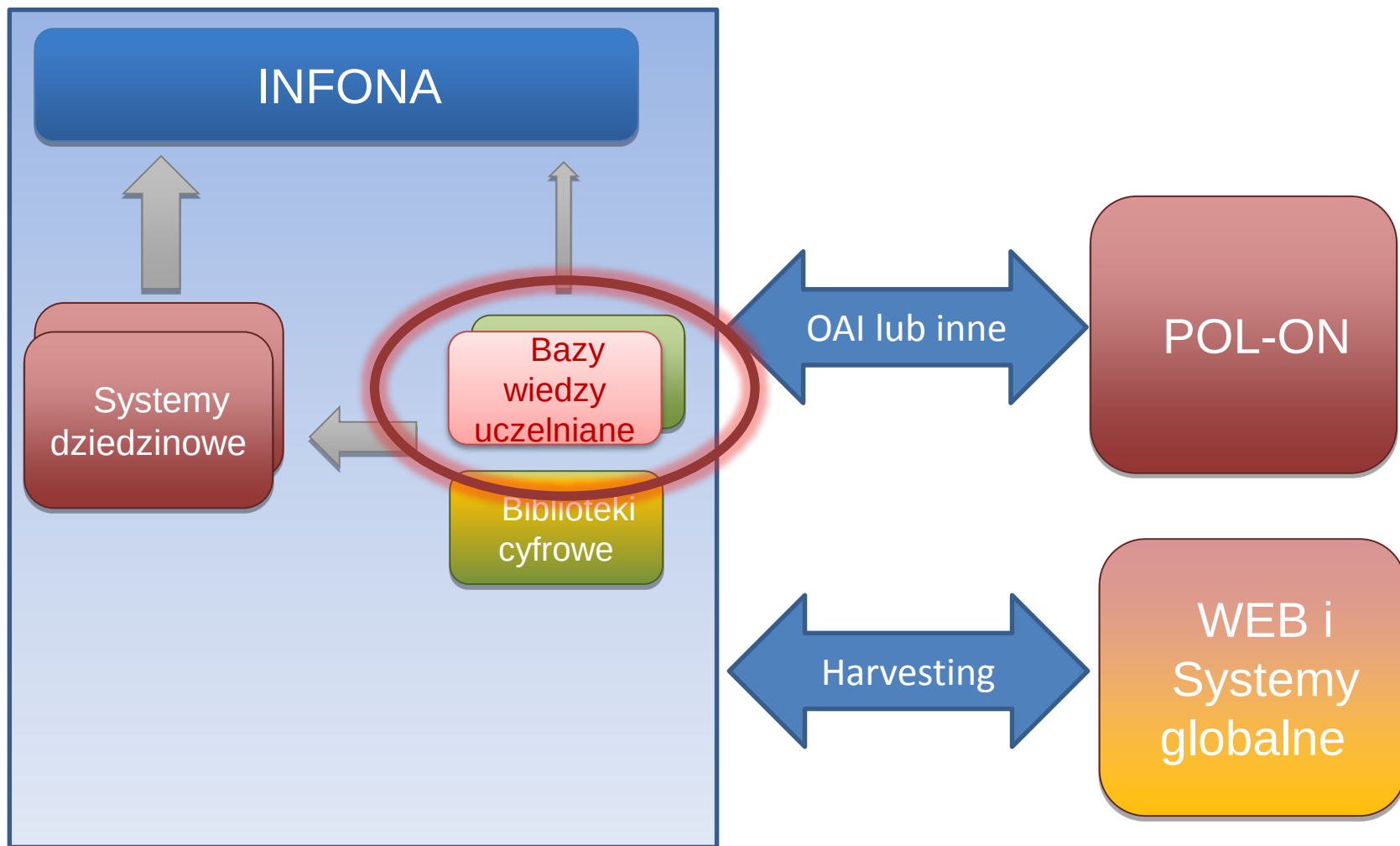
Jak nas widzą?

... Google Scholar

Zapytanie	Liczba odpowiedzi w Google scholar	Liczba odpowiedzi w Microsoft Academia
"Warsaw University of Technology"	30800/38700	14314
"Czech Technical University"	27300/34900	12695
"AGH University"	13600	10036
„Jagiellonian University”	36000	17900

Struktura projektu SYNAT

S
Y
N
A
T



Dlaczego Bazy Wiedzy
... a nie repozytoria

czy coś jest nie tak z repozytoriami?

Co jest nie tak z repozytoriami instytucjonalnymi

Cytat z artykułu

„Innkeeper at the Roach Motel”, autorstwa
Dorothea Salo:

„The institutional repository is like a roach
motel. Data goes in, but it doesn't come out”

... i dalej:

Co jest nie tak z repozytoriami instytucjonalnymi

Fragmenty z „Innkeeper at the Roach Motel”, by Dorothea Salo

„Institutional repositories have not fulfilled their early promise of increased access to the scholarly journal literature through faculty initiative (McDowell, 2007). ...

Academic libraries twist in the wind, enticed by possibilities but disillusioned by results, uncertain how to proceed. Thus far, at least in the United States, doubts about the viability of institutional repositories have been kept quiet or denied altogether....

As long as academic libraries and repository managers remain silent about the current deplorable situation, no one can rectify it; libraries can only wait and hope”

Co jest nie tak z repozytoriami instytucjonalnymi

Excerpts from „Innkeeper at the Roach Motel”, by
Dorothea Salo

Faculty and Self-Archiving:

Institutional repository? Forgive me, but—that sounds
vaguely obscene. (Graduate student in psychology)

What? No! I'd never want those [preprints] on the web!
They're not authoritative! I'd never use them, either!
(Senior professor of engineering)

[Engineering faculty] don't even know the library exists.
They never go there; they download all they need. The
library doesn't even register with them.
(Engineering IT manager)

Co jest nie tak z repozytoriami instytucjonalnymi

*“The institutional-repository software platforms, plagued by innovation-hostile architectures and an ideology-driven **rather than user-centered understanding of the problem domain**, have been slow to align development with needs. Interested faculty, librarians, administrators, and developers must reframe their approaches to institutional repositories if they are to recover from their current neglect”*

Z drugiej strony

Naukowcy budują swoje profile w takich systemach jak

1. ResearchGate
2. Academia.edu
3. Microsoft Academia
4. Scholar

... a nawet w komercyjnych systemach

1. Web of Science
2. Scopus

A może bazy wiedzy wokół repozytorium

Co należałoby zmienić:

1. Nie ograniczać się do repozytorium, raczej budować wielofunkcjonalną bazę wiedzy wokół repozytorium,
 - a) integrując pełną wiedzę o działalności naukowej naukowców, zespołów, jednostek uczelni i uczelni jako całości i
 - b) Zaspokajając różne potrzeby bardzo różnych aktorów (w tym uczelni)
2. Funkcjonalność bazy wiedzy ma być zorientowana na naukowca:
 - a) Profil naukowca powinien prezentować różne aspekty jego aktywności, i integrować pełną wiedzę o naukowcu w jednym miejscu
 - b) Wiedza nie powinna być ograniczona do tego co jest wprowadzone ręcznie

A może bazy wiedzy wokół repozytorium

Co należałoby zmienić:

Zasada podstawowa:

Minimalizować obowiązki, maksymalizować zyski naukowca z bazy wiedzy

A może bazy wiedzy wokół repozytorium

Co należałoby zmienić:

1. Nie ograniczać się do repozytorium, raczej budować wielofunkcjonalną bazę wiedzy wokół repozytorium, integrując pełną wiedzę o działalności naukowej naukowców, zespołów, jednostek uczelni i uczelni jako całości

Akwizycja i Archiwizacja

- pracowników i jednostek
- dorobku naukowego i prac dyplomowych
- uwzględnienie pełnych tekstów dokumentów

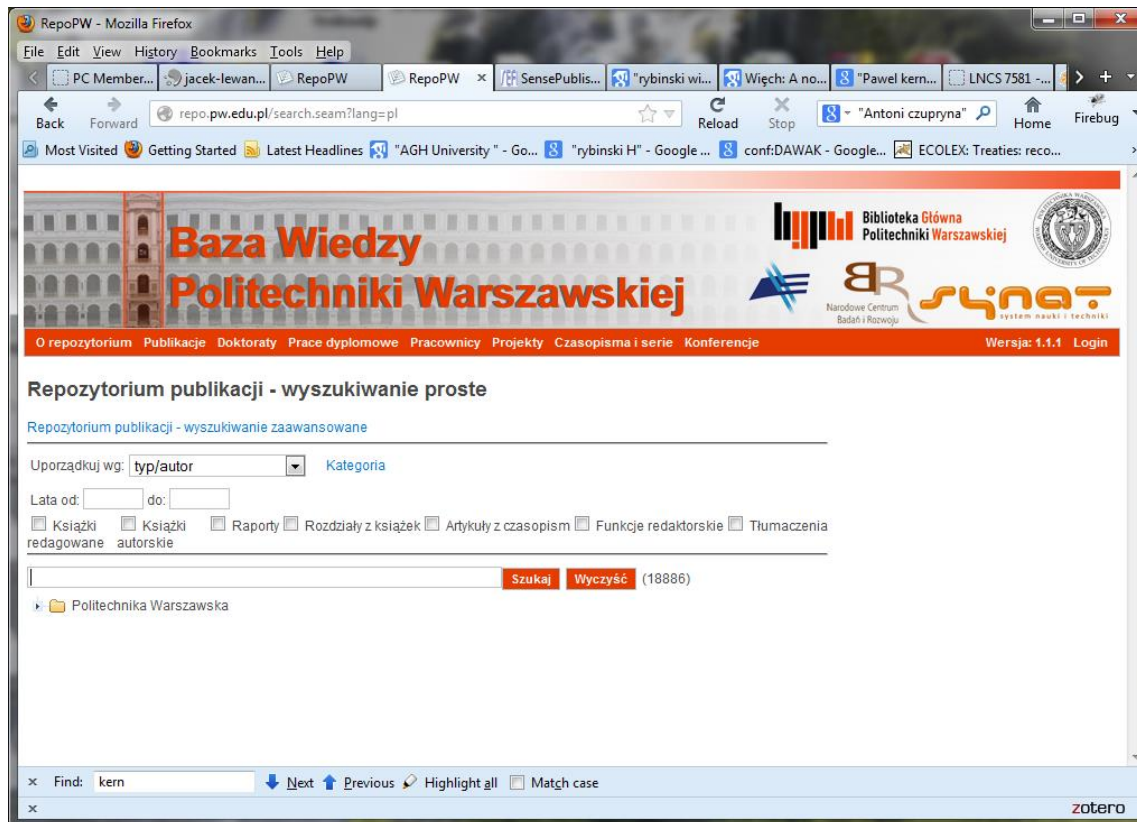
Prezentacja i Udostępnianie

- Wymiana informacji z systemami krajowymi (INFONA, Pol-On, BNP)
- Ekspozycja dorobku naukowego uczelni na świat

Analityka i Sprawozdawczość

- Ocena pracowników
- Raporty uczelniane
- Raporty ministerialne

Baza Wiedzy PW - doświadczenia

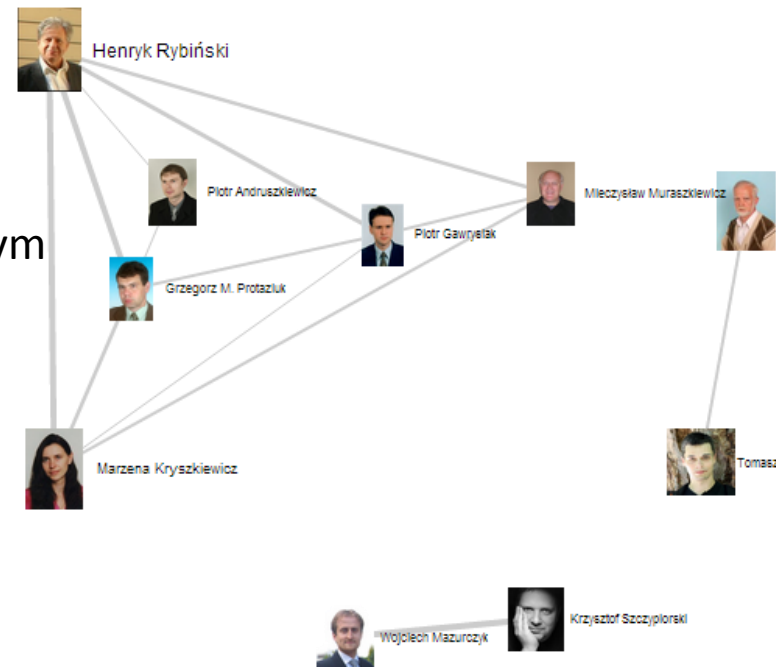


Zrealizowaliśmy:
oprogramowanie
integrujące rozmaite
funkcje
i
wdrożyliśmy Bazę
Wiedzy Politechniki
Warszawskiej



Uczelniana Baza Wiedzy

- Relacje i powiązania pomiędzy obiektami: *author, paper, book, project, journal, conference*
- Automatycznie generowane strony naukowców, instytutów, wydziałów (wkrótce także strony czasopism, konferencji)
 - statystyki dot. działalności naukowej
 - Chmura słów charakteryzująca badania
 - Graf współpracy
- Elementy sieci społecznościowej
 - utrzymywanie aktualnego profilu
 - udostępnianie naukowego dorobku innym
 - Wyszukiwanie naukowców w danej dziedzinie



Czym jest Omega-PSIR?

- Repozytorium
- Bazą wiedzy (kto jest kto, poszukiwanie potencjału, ekspertów, gdzie publikować, gdzie są jakie konferencje)
- Systemem oceny pracowników
- Systemem promocji nauki realizowanej na uczelni
- Platformą komunikacji z biznesem (promocja technologii)
- Systemem sprawozdawczości (raporty ministerialne)



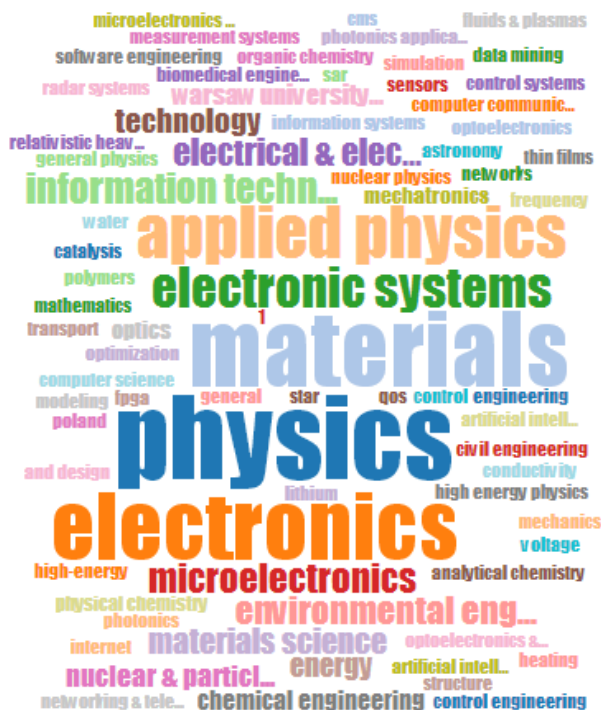
Czym jest Omega-PSIR?



Politechnika Warszawska

[Strona domowa](#)

Profesorowie	506
Adiunkci	1125
Pozostali pracownicy	1102
Doktoranci	417
Publikacje	34264
Wypromowane rozprawy doktorskie	3025
Udział w projektach	776
Wypromowane prace inżynierskie	1692
Wypromowane prace magisterskie	1665
Patenty	179
Rozprawy doktorskie	2927



[Powrót](#)

Czym jest Omega-PSIR?

Wyniki dla zapytania: 'data mining'

Liczba wszystkich pozycji: 659



Jednostki Uczelni



Instytut Informatyki



Wydział Matematyki i Nauk
Informacyjnych



Zakład Zastosowań Informatyki
i Metod Numerycznych

Ludzie ➡

Pierwsze 4 z 61 wyników

Wyszukaj eksperta



Czym jest Omega-PSIR?

Dziedzina: data mining

Eksperti wg. kryterium

Dorobek naukowy

Algorytm ocenia całość dorobku naukowego



prof. dr hab. inż. Henryk Rybiński
Profesor zwyczajny
Instytut Informatyki
Wydział Elektroniki i Technik Informacyjnych
email : H.Rybinski@ii.pw.edu.pl
Dopasowanie: [\[szczegóły\]](#)

	w dziedzinie	wszystkie
Publikacje	36	93
Rozprawy doktorskie	7	14
Udział w projektach	7	31
Wypromowane prace dyplomowe		2



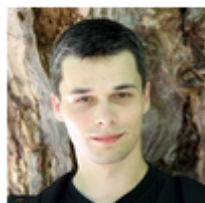
prof. dr hab. inż. Marzena Kryszkiewicz
Profesor nadzwyczajny
Instytut Informatyki
Wydział Elektroniki i Technik Informacyjnych
email : M.Kryszkiewicz@ii.pw.edu.pl
Dopasowanie: [\[szczegóły\]](#)

	w dziedzinie	wszystkie
Publikacje	47	83
Rozprawy doktorskie	5	6
Udział w projektach	6	9
Wypromowane prace dyplomowe	2	6



prof. dr hab. inż. Krzysztof Walczak
Profesor nadzwyczajny
Instytut Informatyki
Wydział Elektroniki i Technik Informacyjnych
email : K.Walczak@ii.pw.edu.pl
Dopasowanie: [\[szczegóły\]](#)

	w dziedzinie	wszystkie
Publikacje	17	39
Rozprawy doktorskie	4	4
Udział w projektach	2	2
Wypromowane prace dyplomowe	3	6



dr inż. Tomasz Gambin
Adiunkt
Instytut Informatyki
Wydział Elektroniki i Technik Informacyjnych
email : T.Gambin@ii.pw.edu.pl
Dopasowanie: [\[szczegóły\]](#)

	w dziedzinie	wszystkie
Publikacje	7	23
Rozprawy doktorskie	1	1
Udział w projektach	1	1

Czym jest Omega-PSIR?



Profesor zwyczajny

Wydział Elektroniki i Technik Informacyjnych

Telefon: +48 22 234-7432, fax +48 22 234 6091

Konsultacje: poniedziałek 14.00-16.00

Raport dorobku

Publikacje	91
Rozprawy doktorskie	13
Udział w projektach	30
Wypromowane prace dyplomowe	2
Aktywność zawodowa	9

h-index*:11

Tytuł magistra inżyniera uzyskał w 1970 roku, stopień naukowy doktora nauk technicznych uzyskał w roku 1974, stopień naukowy doktora habilitowanego nauk technicznych uzyskał w 1988 roku na Wydziale Elektroniki Politechniki Warszawskiej. Tytuł naukowy profesora nauk technicznych uzyskał w 2001 roku nadany przez Prezydenta Rzeczypospolitej Polskiej. Pracuje w Instytucie Informatyki od 1989 roku. Od 1994 do 2008 roku kierownik Zakładu Systemów Informatycznych. Kierownik specjalności Inżynieria Systemów Informatycznych do 2008 roku. Od 2008 roku Dyrektor Instytutu Informatyki. Specjalność naukowa: informatyka, bazy danych i systemy informacyjne. Członek Komitetu Informatyki PAN od 2011 roku. Przewodniczący Rectorskiej Komisji ds. Opracowania Strategii Informatyzacji Politechniki Warszawskiej (2013-2020).



Czym jest Omega-PSIR?

[Profil](#)
[Publikacje](#)
[Doktoraty](#)
[Projekty](#)
[Prace dyplomowe](#)
[Aktywność](#)
[Cytowania](#)
[Statystyki](#)
[Współpraca](#)
[Edytuj](#)



prof. dr hab. inż. Henryk Rybiński

Profesor zwyczajny

Instytut Informatyki

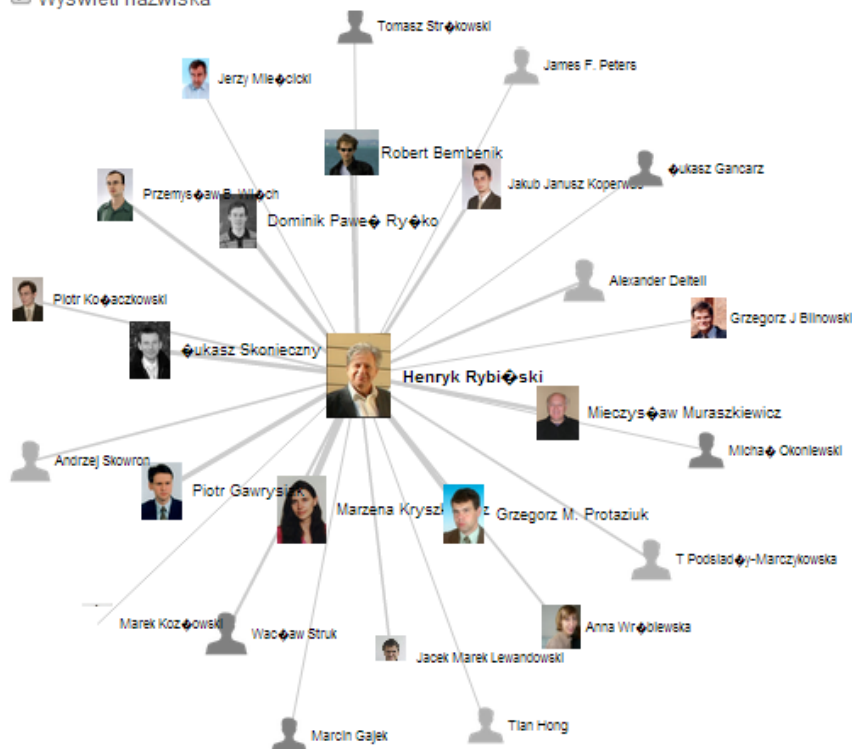
Wydział Elektroniki i Technik Informacyjnych

email : H.Rybinski@ii.pw.edu.pl

Telefon: +48 22 234-7432, fax +48 22 234 6091

Pokój: 204

Konsultacje: poniedziałek 14.00-16.00

☒ Wyświetl nazwiska

Wskazówki:

- Wybierając na chmurze tagów dziedzinę ograniczysz graf współpracowników do wybranej dziedziny
- Podwójne kliknięcie na współpracownika spowoduje przejście do jego profilu
- Po najechaniu myszką na połączenie między współpracownikami zostanie wyświetlona liczba ich wspólnych publikacji

Czym jest Omega-PSIR?

Profil Publikacje Doktoraty Projekty Prace dyplomowe Aktywność Cytowania Statystyki Współpraca Edytuj



prof. dr hab. inż. Henryk Rybiński

Profesor zwyczajny

Instytut Informatyki

Wydział Elektroniki i Technik Informacyjnych

email : H.Rybinski@ii.pw.edu.pl

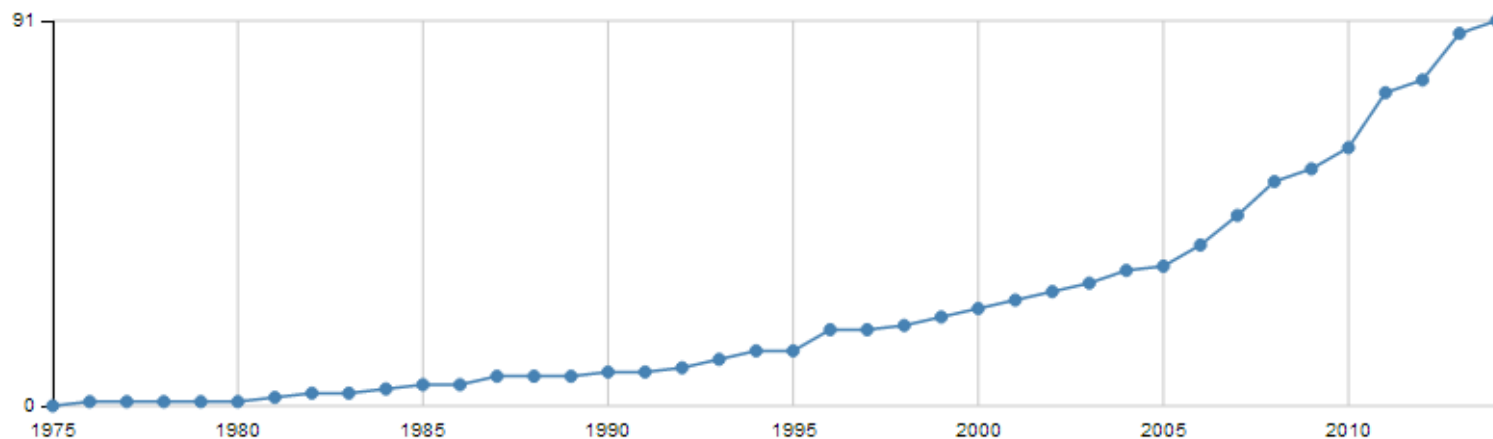
Telefon: +48 22 234-7432, fax +48 22 234 6091

Pokój: 204

Konsultacje: poniedziałek 14.00-16.00

knowledge base association rules
text mining
text analysis rough sets learning algorithm
proper nouns default logic decision rules
data mining
information systems synonyms
frequent itemsets rough set theory knowledge discovery
repository information retrieval
pattern recognition distributed computing

● Publikacje ● Punktacja



● Skumulowany ● Roczny

Czym jest Omega-PSIR?

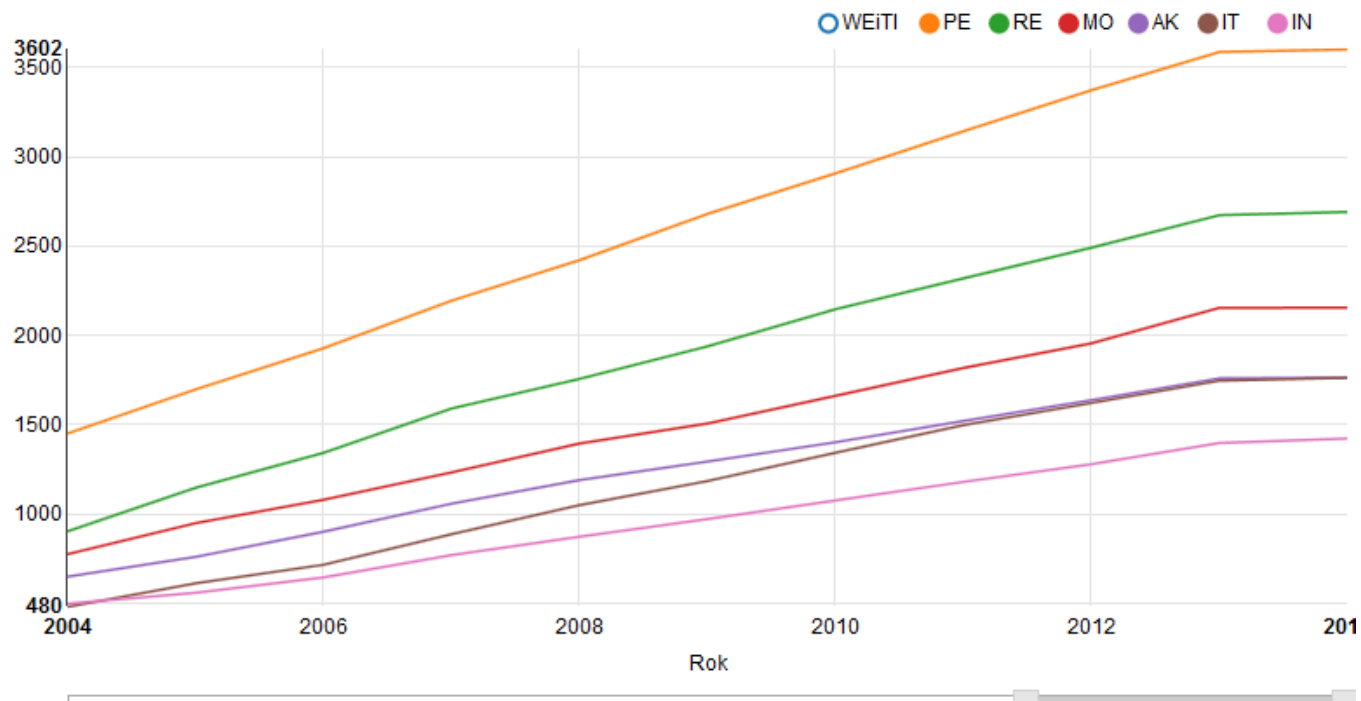
Profil Statystyki Współpraca pracowników Współpraca wewnętrzna Współpraca zewnętrzna



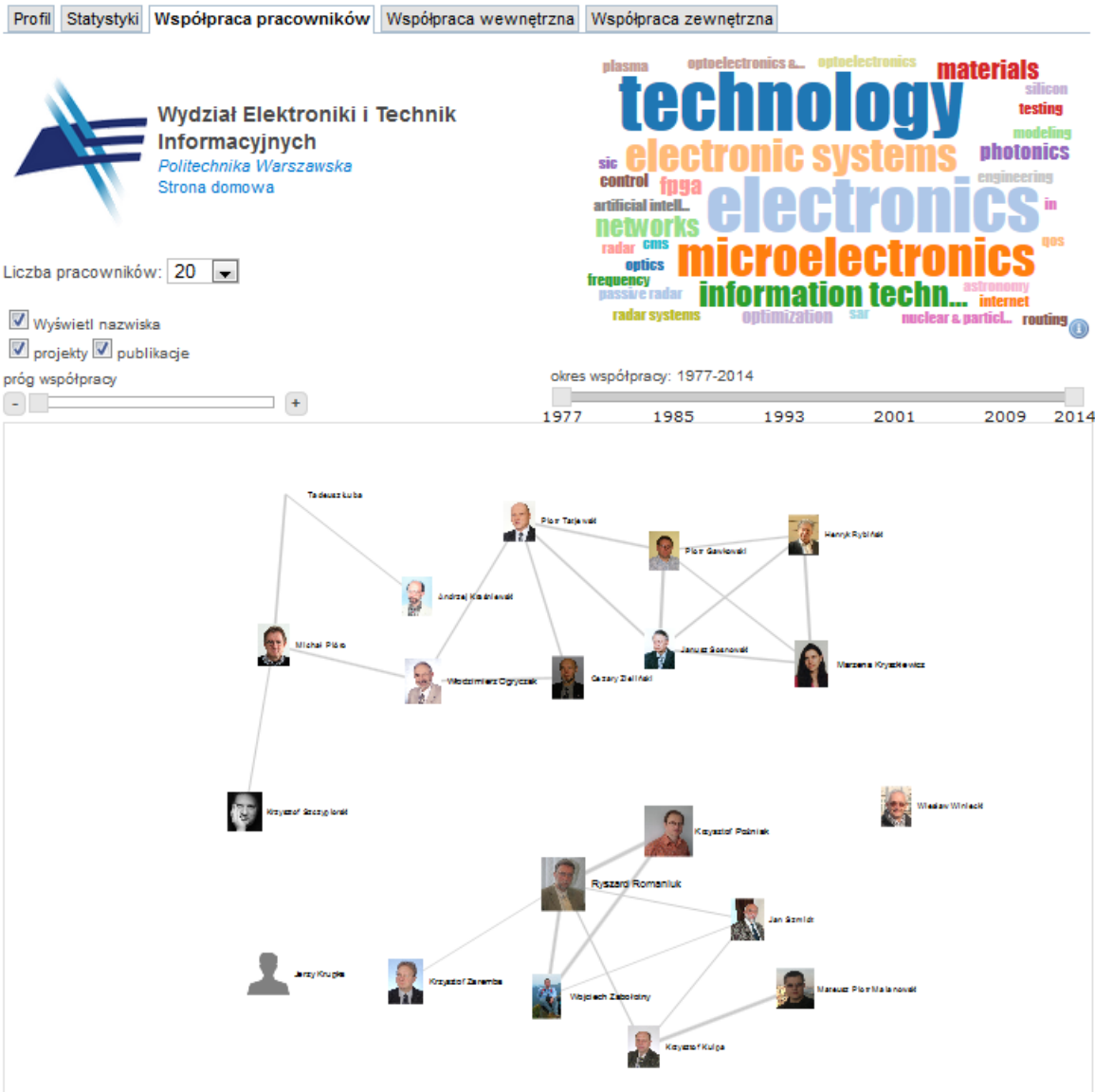
Wydział Elektroniki i Technik
Informacyjnych
Politechnika Warszawska
Strona domowa

information techn... networks astronom
technology fpga
control warsaw university... physics
applied physics optimization

☒ Publikacje ☐ Punkcja



Czym jest Omega-PSIR?



Czym jest Omega-PSIR?



Naukowcy/Pracownicy/Studenti



Bibliotekarze

Interesariusze/Beneficjenci



Administracja
(sekretariaty, biura projektów)



Kierownictwo: Rektorat,
dziekanaty



Uczelnia i jej jednostki: wydziały, instytuty

Techniczna charakterystyka

Home Publikacje Doktoraty Prace dyplomowe Pracownicy Projekty Czasopisma i serie Konferencje

Funkcje główne

Książki
Artykuły i rozdziały
Raporty
Tłumaczenia
Funkcje redaktorskie
Prace inżynierskie
Prace magisterskie
Rozprawy doktorskie
Projekty
Patenty

Funkcje Pomocnicze

Afilacje
Autorzy i pracownicy
Autorzy instytucjonalni
Konferencje (wydarzenia)
Konferencje (tytuły)
Czasopisma i serie
Lista serii książkowych (do usunięcia)
Języki
Tabele walidacyjne
Typy projektów
Import

System wprowadzania i modyfikowania danych w Repozytorium PW.
Jesteś zalogowany jako jkoperwa
Wybierz

Życzy

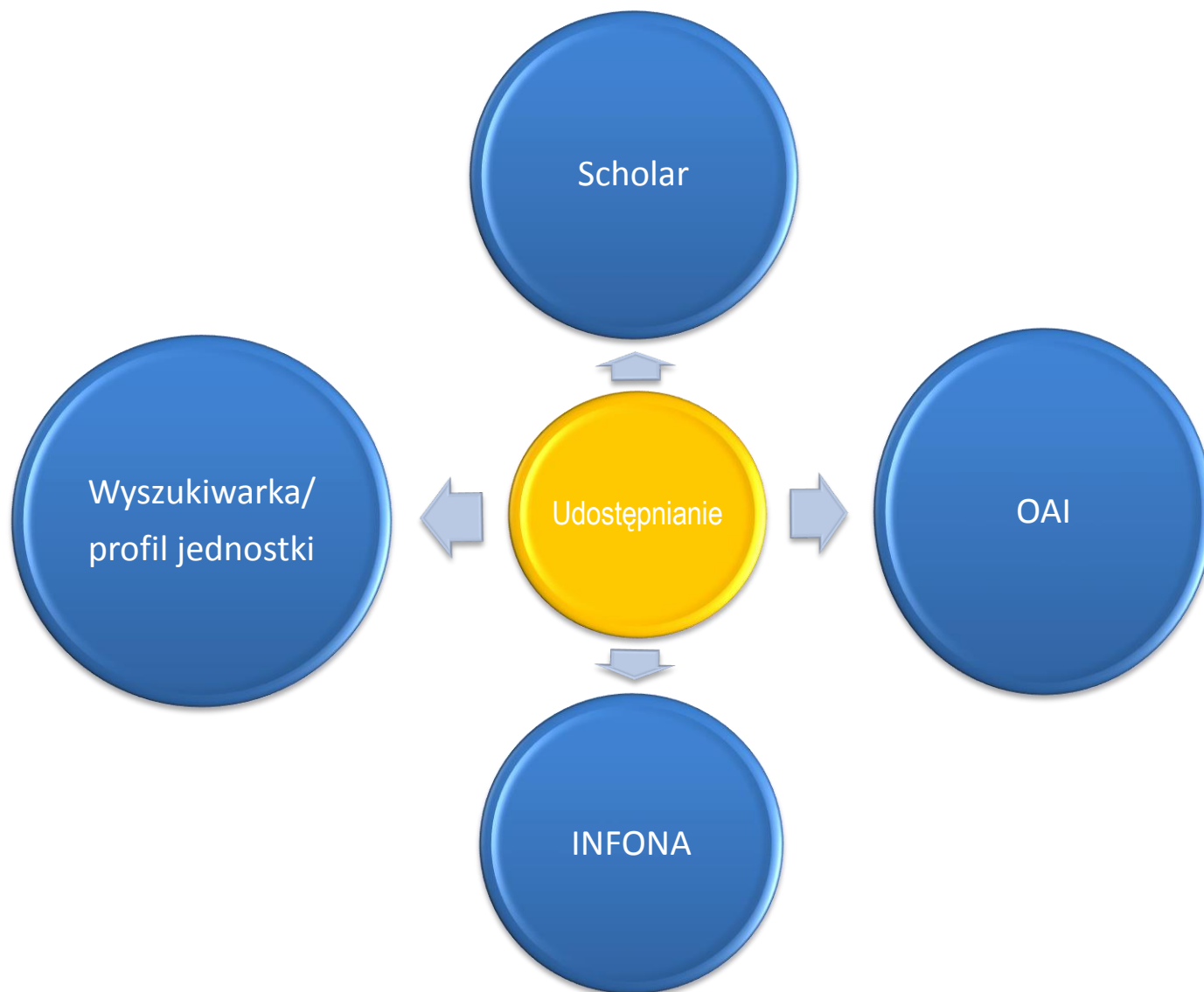
Dane

- Powiązania pomiędzy obiektami (struktura grafowa)
- Przechowywanie historyczne danych
- Prezentacja powiązań pomiędzy różnymi obiektami
- Kontrola uprawnień dostępu na poziomie wyszukiwania/prezentowania/wprowadzania danych
- Wyszukiwanie pełnotekstowe
- Wersjonowanie
- Rozszerzalne o dowolny typ zasobu, którego potrzeba przechowywania może pojawić się w przyszłości
- Definiowanie ekranów, procedur, itp. poza oprogramowaniem

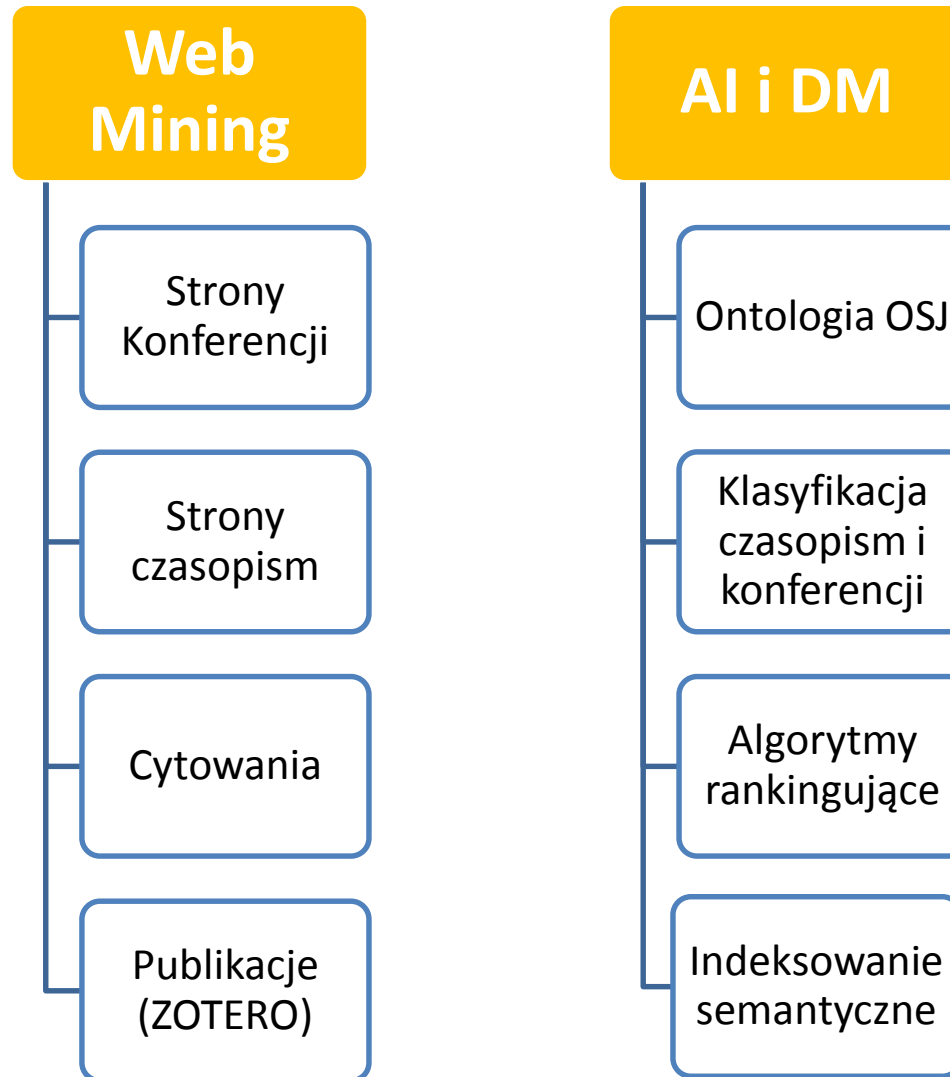
Czym jest Omega-PSIR?



Czym jest Omega-PSIR?

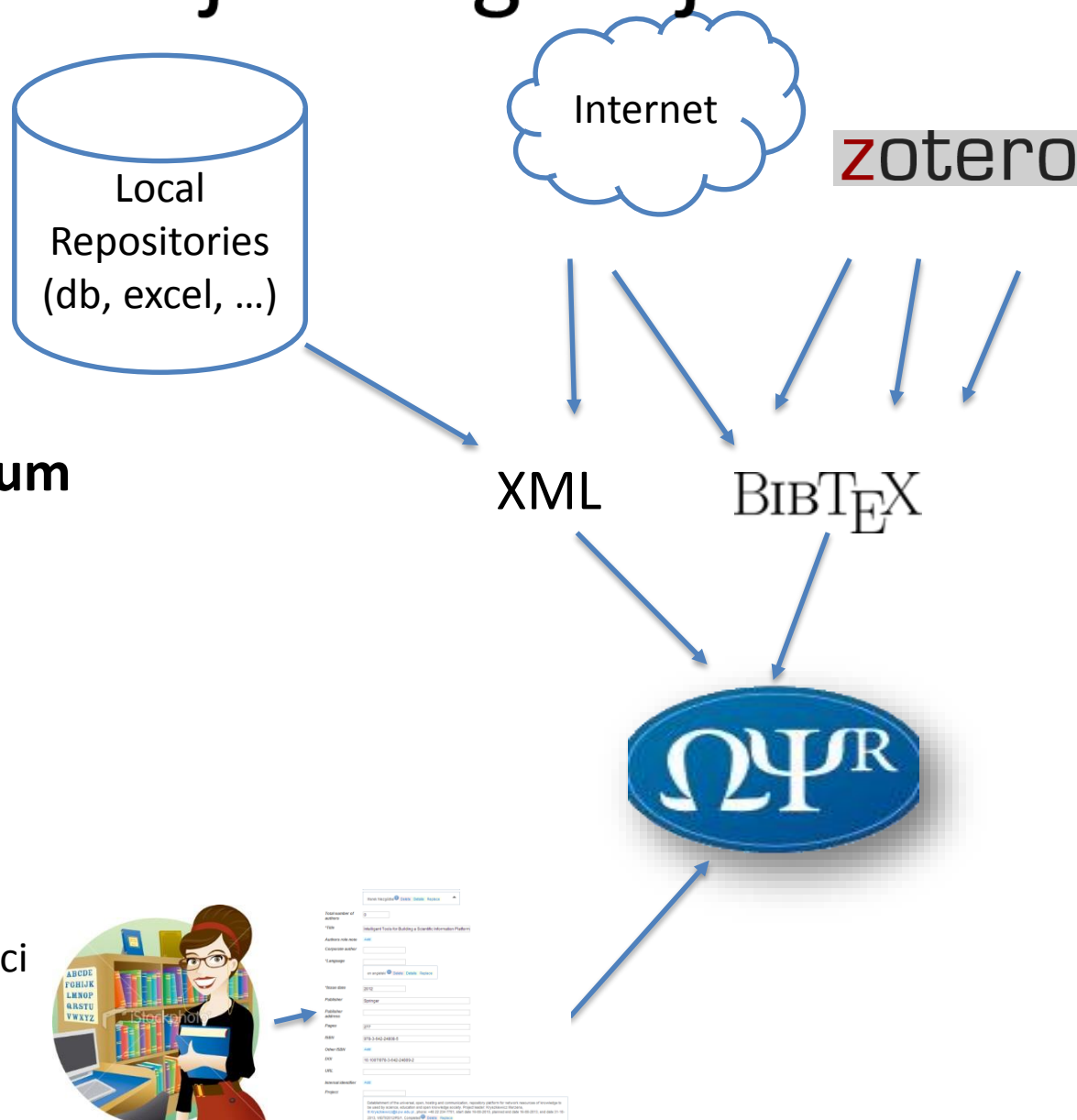


Metody Sztucznej Inteligencji



Metody Sztucznej Inteligencji

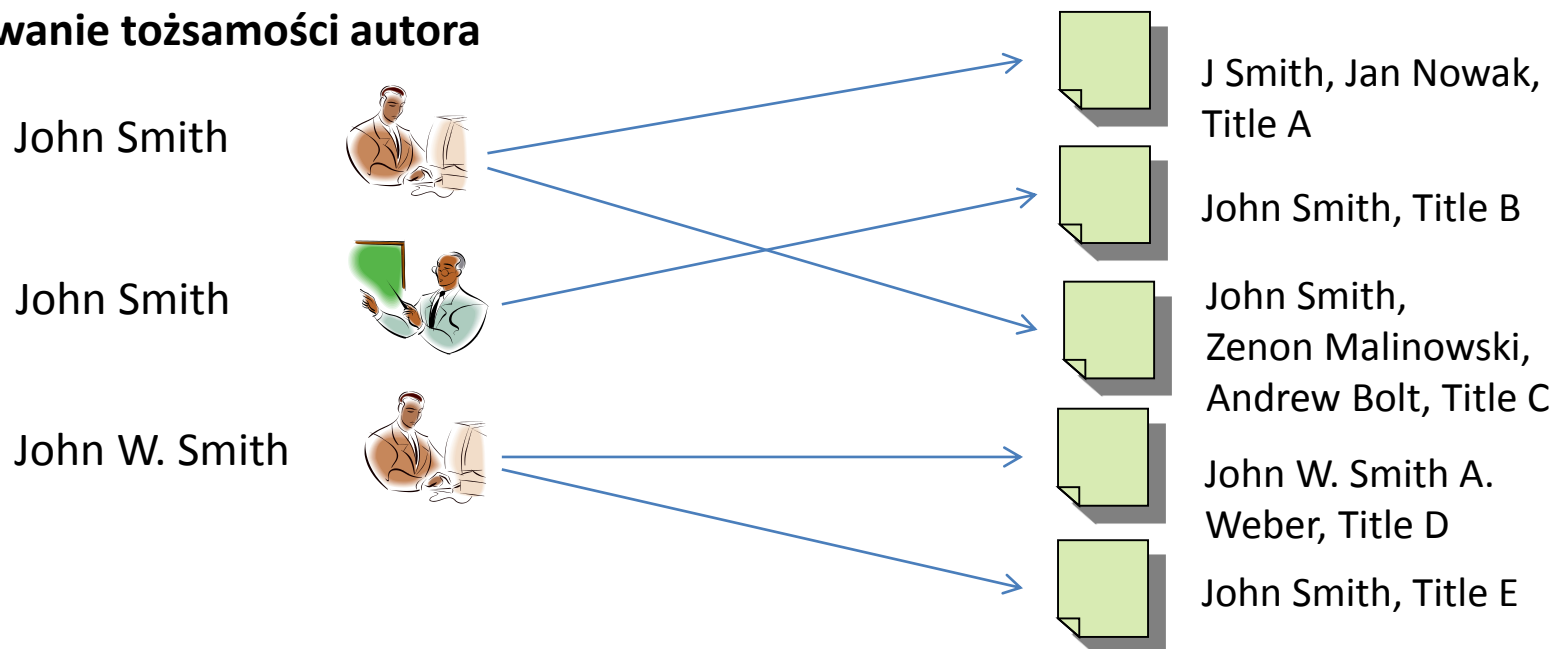
Narzędzia AI do wypełniania repozytorium



- Wykrywanie duplikatów
- Scalanie
- Rozstrzyganie niejednoznaczności
- Integrowanie

Metody Sztucznej Inteligencji

Wykrywanie tożsamości autora



Metody:

Tworzenie klastrów publikacji wg:

- Treści dokumentu (wykorzystanie tagów semantycznych)
- Miary podobieństwa (współautorstwo, cytowania, podobieństwo tytułów i abstraktów)

Metody Sztucznej Inteligencji

Przetwarzanie semantyczne

Cele:

- Wzbogacenie danych przez dodanie semantycznie znaczących terminów
 - Poprawa wyszukiwania
 - Odkrywanie obszarów badawczych (dla autorów, zespołów, jednostek)
-
1. Tagowanie publikacji kategoriami OSI
 2. Ekstrakcja i indukcja słów
 3. Indeksowanie bazujące na znaczeniach



Metody Sztucznej Inteligencji

Tagowanie za pomocą OSJ (Ontology of Scientific Journals, utworzona przez Science-Metrix, Canada, 2011)

15000 czasopism, hierarchiczna klasyfikacja (3 poziomy, kilkanaście języków, BG PW dodała polski)

- Gotowa do wykorzystania dla większości znanych czasopism (JCR) – i wykorzystywana
- Dla publikacji nie objętych OSJ zaimplementowaliśmy (na podstawie OSJ) klasyfikator dla czasopism spoza ontologii, materiałów konferencyjnych i rozdziałów z książek

Wynik

tree of bayesian classifier, 85% accuracy

Metody Sztucznej Inteligencji

Ekstrakcja i indukowanie słów

IEEE.org | IEEE Xplore Digital Library | IEEE Standards | IEEE Spectrum | More Sites Cart (0)

IEEE Xplore®
DIGITAL LIBRARY

OSJ: Applied Technologies: ICT

Access provided by:
POLITECHNIKI WARSZAWSKIEJ
[» Sign Out](#)

Download Citations

Email

Print

Request Permissions

Save to Project

0

Word meaning disambiguation has always been an important problem in many computer science tasks, such as information retrieval and extraction. One of the problems faced in automatic word sense discovery, is the number of different senses a word can have. Often, senses are dominated by some other, more frequent ones. Discovering such dominated meanings can significantly improve quality of many text-related algorithms. In particular, Web search quality can be leveraged. In the paper, we present a novel approach for discovering word senses. The method is based on concise representations of frequent patterns. The method attempts to discover not only word senses that are dominating, but also senses that are dominated and under represented in the repository.

Published in:
Data Mining Workshops, 2008. ICDMW '08. IEEE International Conference on

Terminy wyekstrahowane i zaindukowane:

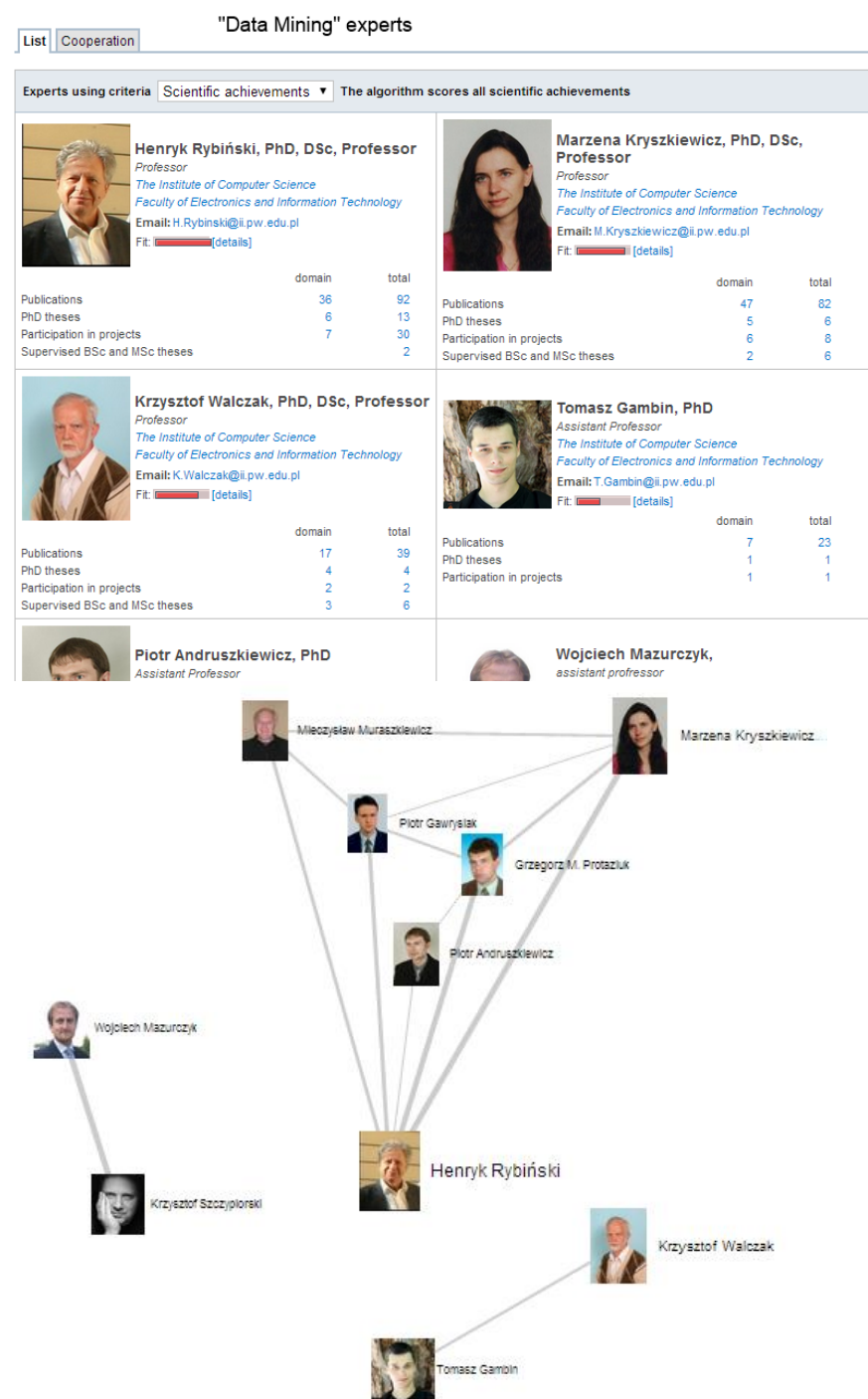
information, information retrieval, web search engine,
computer science, word sense, World Wide Web

Ekstrakcja i indukcja termów

- Wikipedia jako zasób semantyczny
- Metoda inspirowana przez prace Milne'go, Medelyana i Wittena
- Nasz wkład do metody: tagowanie z dokładaniem kontekstu znaczeń wykrywanych przez algorytm SnS

Poszukiwanie ekspertów i ranking

- Wyszukiwanie pełnotekstowe publikacji (np. „data mining”)
- Ranking wg relewancji względem zapytania (w oparciu o podobieństwo wektorów z podbiciem atrybutu „tagowanie semantyczne”)
- Ocena publikacji wg „research impact” (punkty ministerialne, IF, cytowania – możliwość zdefiniowania kryterium)
- Ocena za rolę (autor, współautor, promotor, recenzent, ...)
- Uporządkowanie autorów wg. tej oceny



Wdrożenie na Politechnice Warszawskiej

- Trzeci rok eksploatacji
- Ustanowione zasady gromadzenia i udostępniania
- Ponad 35000 publikacji (roczny przyrost w 2013 > 4000),
- Seria szkoleń organizowanych przez BG PW
- Wdrożone procedury typu „crowd sourcing” wspomagające utrzymywanie wysokiego poziomu jakości danych
- W trakcie wdrażania jest harvesting stron czasopism i stron konferencyjnych

Zainteresowanie systemem

- Politechnika Łódzka
- WAT
- Politechnika Poznańska
- Politechnika Wrocławska
- Uniwersytet Ekonomiczny we Wrocławiu
- Uniwersytet Ekonomiczny w Katowicach
- SGH
- Uniwersytet Gdański
- Uniwersytet UMK Toruń
- Akademia Pedagogiki Specjalnej Warszawa
- Uniwersytet Przyrodniczy w Lublinie
- UJ
- Instytut Kardiologii w Aninie
- Instytut Technologii Eksploatacji – Państwowy Instytut Badawczy Radom

Uniwersytet Ekonomiczny we Wrocławiu

← → ↻

Aplikacje RepoPW jmeter Read Later CEP jbos monitoring wnioski bieszczady

[O repozytorium](#) [Szukaj](#) [Publikacje](#) [Doktoraty](#) [Prace dyplomowe](#) [Pracownicy](#) [Aktywność zawodowa](#) [Projekty](#) [Czasopisma i serie](#) [Konferencje](#)

Repozytorium publikacji - wyszukiwanie proste

[Repozytorium publikacji - wyszukiwanie zaawansowane](#)

Uporządkuj wg:

[Kategoria](#)

Lata od: do:

☐ Książki redagowane ☐ Książki autorskie ☐ Rozdziały z książek ☐ Artykuły z czasopism ☐ Raporty ☐ Tłumaczenia ☐ Patenty

(1555)

Wyszukiwanie wykonywane jest w trybie pełno-tekstowym, tzn. zadane terminy mogą wystąpić w dowolnym polu opisu lub w dokumentach. Precyzyjne wyszukiwanie ze wskazaniem pola (np. wg autora) należy wykonywać z wykorzystaniem formularza [wyszukiwanie zaawansowane](#).

☒ Uniwersytet Ekonomiczny we Wrocławiu

- ☐ Biblioteka Główna
- ☐ Biuro Promocji
- ☐ Centrum Jakości Kształcenia
- ☐ Dział Obsługi Projektów Rozwojowych
- ☒ Wydział Ekonomii, Zarządzania i Turystyki
- ☒ Wydział Inżynieryjno-Ekonomiczny
- ☒ Wydział Nauk Ekonomicznych
- ☒ Wydział Zarządzania, Informatyki i Finansów

Adaptowalność systemu



Repozytorium publikacji - wyszukiwanie proste

[Repozytorium publikacji - wyszukiwanie zaawansowane](#)

Uporządkuj wg:

Lata od: do:

☐ Książki ☐ Książki ☐ Raporty ☐ Rozdziały z książek ☐ Artykuły z czasopism ☐ Funkcje redaktorskie ☐ Tłumaczenia
redagowane autorskie

Szukaj **Wyczyść** (15334)

- Politechnika Warszawska
 - Biblioteka Główna PW
 - Kolegium Nauk Ekonomicznych i Społecznych w Płocku
 - Studium Języków Obcych
 - Studium Wychowania Fizycznego i Sportu
 - Wydział Administracji i Nauk Społecznych
 - Wydział Architektury
 - Wydział Budownictwa, Mechaniki i Petrochemii w Płocku
 - Wydział Chemiczny
 - Wydział Elektroniki i Technik Informatycznych
 - Instytut Automatyki i Informatyki Stosowanej
 - Instytut Informatyki
 - Instytut Mikro i Optoelektroniki
 - Instytut Radioelektroniki
 - Instytut Systemów Elektronicznych
 - Instytut Telekomunikacji

Wygląd

Reguły dostępu

Raporty

...

Warunki udostępniania

- Prace zrealizowane w ramach SYNAT (stan obecny) udostępnione NIEODPŁATNIE (w tym rozległa dokumentacja użytkownika)
- Deklaracja dalszego rozwoju systemu ze strony II PW, oraz niezbędne aktualizacje (reguły punktacji MNiSW), uwzględnienie potrzeb uczelni w mapie drogowej kolejnych wersji
- Możliwość uzyskania wsparcia podczas wdrożenia i eksploatacji systemu

Najbliższe spotkanie dla zainteresowanych uczelni

16 września 2014
w godzinach 10:00-15:00
w Centralnej Auli, Gmach WEiTl,
Warszawa, Nowowiejska 15/19

Dziękuję za uwagę



<http://omega-psir.ii.pw.edu.pl>

<http://repo.pw.edu.pl>

<http://wizzar.ii.pw.edu.pl/RepoPW/>

information extraction
from unstructured data (web pages)

AI Tools

Find relevant information on homepages

- Researcher: position, affiliation, address, email, phone...
- Journal: name, issn, impact factor
- Conference: name, important dates, location

Methods:

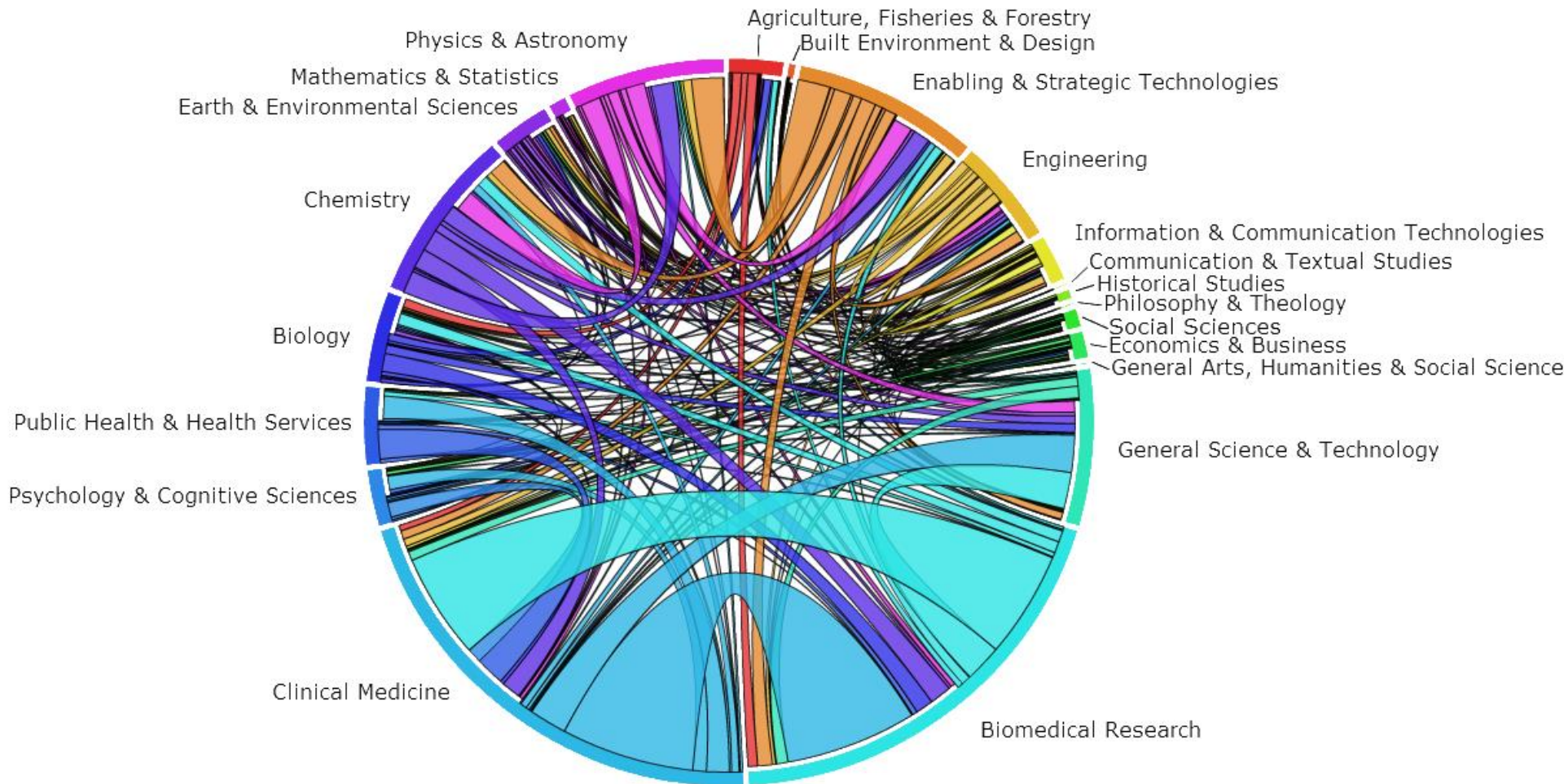
Regular expressions, generalized sequence patterns, SVM, HMM, conditional random fields

As corporate [\[position\]](#) **vice_president** [\[/position\]](#) of the [\[affiliation\]](#) **Natural Interactive Services** [\[/affiliation\]](#) Division NISD at Microsoft Corp. Kai-Fu Lee was responsible for the development of the technologies and services for making the user interface simpler and more natural. NISD includes technologies and products for speech natural language advanced search and help and authoring and learning technologies. The mission of NISD is to make these technologies usable and useful for Microsoft's customers. Lee joined Microsoft in 1998 and was the [\[position\]](#) **founder** [\[/position\]](#) of [\[affiliation\]](#) **Microsoft Research Asia** [\[/affiliation\]](#) which has since become one of the [\[position\]](#) **best** [\[/position\]](#) research laboratories in the world with a prolific publication and product transfer record. Before coming to Microsoft he was the [\[position\]](#) **president** [\[/position\]](#) of Cosmo Software the [\[affiliation\]](#) **Silicon Graphics Inc** [\[/affiliation\]](#). SGI multimedia software business unit. Before that he was [\[position\]](#) **vice_president** [\[/position\]](#) and general manager of Silicon Graphics Web products division responsible for several product lines and the company's corporate Web strategy. Before joining SGI Lee spent six years at Apple most recently as vice_president of the company's interactive media group which developed QuickTime QuickDraw D QuickTime VR and PlainTalk speech technologies. Prior to his position at Apple he was an [\[position\]](#) **assistant professor** [\[/position\]](#) at [\[affiliation\]](#) **Carnegie Mellon University** [\[/affiliation\]](#) where he developed the world's first speaker speaker-independent independent continuous speech-recognition system. While at Carnegie Mellon [\[affiliation\]](#) **Lee** [\[/affiliation\]](#) also developed the world-champion champion computer program that plays the game Othello and that defeated the human world champion in 1988. Lee holds a doctorate in computer science [\[affiliation\]](#) **from Carnegie Mellon University** [\[/affiliation\]](#) a bachelor's in computer science with highest honors [\[affiliation\]](#) **from Columbia University** [\[/affiliation\]](#). Lee is a Fellow of the IEEE.

Sense indexing

- Co to robi i jak działa?

OSJ – level 2



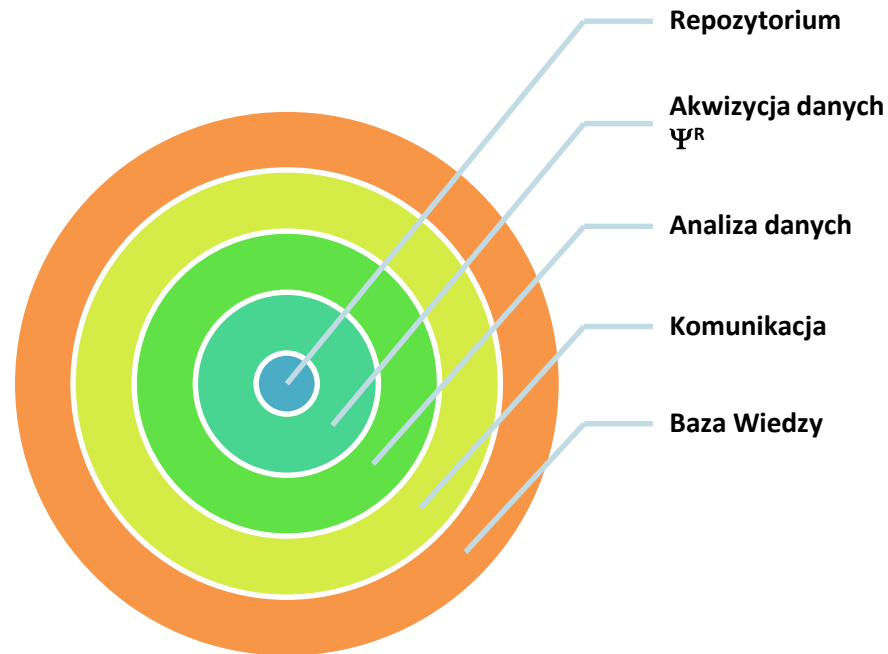
Institutional Repository

- Storing scientific information (meta data with full documents) – papers, theses, reports, patents, projects
- Exposing and sharing (with clean GUI, REST API and OAI)
- Advanced querying
- Scoring scientific achievements (with government rules)
- Reporting
 - for the national systems (POL-on), evaluation of research units,
 - complete bibliographies,
 - Annual Reports for the institutes, faculties, etc.,
- communication with other national systems (INFONA)
- Key technology: Apache Jackrabbit
(free implementation of **Content Repository API for Java - JCR**)

A może bazy wiedzy wokół repozytorium

Co należałoby zmienić:

1. Nie kończyć na repozytorium, raczej budować wielofunkcyjną bazę wiedzy wokół repozytorium



Czy tylko repozytorium?

Poza funkcjami repozytoryjnymi (obejmującymi publikacje, patenty, bazy doktoratów, dyplomów, projekty):

1. Generator strony domowej pracownika
2. Poszukiwanie eksperta (mamy wszystko w jednym miejscu),
3. Mapa obszarów badań zakładu, instytut, katedry, wydziału uczelni (j.w.)
4. Działalność naukowca (funkcje redaktorskie, aktywność w ciałach naukowych/komisjach, itp.)
5. Czasopisma (z informacją o IF, call for SI)
6. Konferencje (z informacją o CFP)

Czy tylko repozytorium?

... ale także funkcjonalność sprawozdawania

1. dla władz wydziałów i uczelni
2. do ministerstwa (POL-ON),
3. budowania raportów instytutowych, dziekańskich, tworzenia bibliografii pracowników i tematycznych

... ale także komunikacja z systemami

1. INFONA
2. Bibliografia Nauki Polskiej