

# Wprowadzenie do technologii semantycznych

Agnieszka Ławrynowicz  
Politechnika Poznańska

Poznań, rok akademicki 2014/2015

# Co to jest semantyka?

---

**Semantyka** (starogrecki: σημαντικός sēmantikós, z sema = „znak”) bada znaczenie symboli i zbiorów symboli wyrażonych w języku, kodzie, lub innej formie reprezentacji

# Wyszukiwanie informacji na temat zasobów wymagające znajomości wiedzy dziedzinowej

---

*Znajdź zdjęcia „zwierząt, które wykorzystują echolokację, ale nie są to ani nietoperze ani delfiny”*



## Podsumowanie

---

- **Wyszukiwanie informacji** w WWW bazujące na słowach kluczowych: problem z interpretacją wiedzy negatywnej
- Problem z granulacją informacji: strony a nie zasoby/encje
- Potrzeba danych na temat zasobów (np. typowe własności dla encji danego typu)
- Potrzeba **semantycznego wyszukiwania**

# Rezerwacja wycieczki z Edynburga do Poznania

---

Najpierw rezerwacja przelotu...

## ...duże linie lotnicze...

Lufthansa - Book your ticket and check-in online. We fly from 7 polish cities - Windows Internet Explorer

http://www.lufthansa.com/online/portal/lh/pl/homepage?l=en&cid=1000340&WT.srch=1

Google

Pluk Edycja Widok Ulubione Narzędzia Pomoc

Ulubione Google Sugerowane witryny Bezpłatna usługa pocztowa ... Galeria obiektów Web Slice

Lufthansa - Book your ticket and check-in online. We f...

Strona Bezpieczeństwo Narzędzia

Poland Polski Other countries Help & Contact Mobile version Search

There's no better way to fly.  
**Lufthansa**

Book & Plan Offers & Ideas My Bookings Information & Service Miles & More Login

Booking Check-in

Flight Hotel Car

Round-trip One-way

From → Cities Depart on  
Edinburgh, Edinburgh (E Su, 03.10.2010

To → Cities  
Poznan, Poznan (POZ).

Search by price (+/- 3 days)  
Search these dates

**lufthansa.com**  
Buenos Aires from  
**2199 PLN\***

Book now, fly summer 2011.

Fly with Lufthansa to Canada and USA at attractive price

→ Atlanta from 1959 PLN\*  
→ Boston from 1899

Bogota - Lufthansa's new destination. Starting from  
**2399 PLN\***

My bookings

Gotowe Internet 100%

...lub linie lotnicze kraju docelowego...

lot.com | Rezerwuj tanie bilety i podróżuj wygodnie z PLL LOT - Windows Internet Explorer

http://www.lot.com/web/lot/home

Plik Edycja Widok Ulubione Narzędzia Pomoc

Ulubione Google Sugerowane witryny Bezpłatna usługa pocztowa ... Galeria obiektów Web Slice

lot.com | Rezerwuj tanie bilety i podróżuj wygodnie z ...

Strona Bezpieczeństwo Narzędzia

**LOT** POLSKIE LINIE LOTNICZE

INFORMACJA I SERWIS MOJE REZERWACJE PROMOCJE MILES AND MORE O NAS

Polska / polski

Strona główna >

**ZAREZERWUJ LOT**

☒ W obie strony ☐ W jedną stronę

Miejsce wylotu:  Data wylotu:

Miejsce przylotu:  Data powrotu:

☒ Wyszukiwanie elastyczne (+/- 3 dni)

Rodzaj pasażerów

| Dorośli                        | Seniorzy                       | Dzieci                         | Niemowlęta                     |
|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| 12 lat                         | 60 lat                         | 2-11 lat                       | 2 lata                         |
| <input type="text" value="1"/> | <input type="text" value="0"/> | <input type="text" value="0"/> | <input type="text" value="0"/> |

Klasa podróży:

**od 1157 PLN w obie strony**

Otwieramy bezpośrednie połączenie **Warszawa - Kair**

**NARZĘDZIA**

- ZALOGUJ SIĘ
- ODPRAWA ONLINE
- ŚLEDZENIE LOTU
- ROZKŁAD LOTÓW
- ZNIŻKI DLA FIRM I AGENTÓW
- TWÓJ ASYSTENT

Gotowe

Internet 100%

...lub tanie linie...

Tanie linie lotnicze Ryanair - Tanie loty z Polski - Windows Internet Explorer

http://www.ryanair.com/pl

Plik Edycja Widok Ulubione Narzędzia Pomoc

Ulubione Google Sugerowane witryny Bezpłatna usługa pocztowa ... Galeria obiektów Web Slice

Tanie linie lotnicze Ryanair - Tanie loty z Polski

Polska (Polski)

**RYANAIR**

Połączenia lotnicze

STRONA GŁÓWNA WYNAJEM AUT UBEZPIECZENIE TURYSTYCZNE TANIE HOTELE DOJAZD DO LOTNISKA KARTA RYANAIR HOSTELE I B&BS CRUISE HOLIDAYS TANI ROAMING WYNAJEM NA WAKACJE KEMPING

Szukaj Zarządzaj Rezerwacją Informacje o lotach Pytania dotyczące podróży Miejsca docelowe Wiadomości O nas Zarezerwuj teraz Rozkłady lotów Opłaty Praca Obsługa klienta

**WYPRZEDAŻ**  
W JEDNĄ STRONĘ  
**21** OD PLN  
PODATKI I OPŁATY WLICZONE  
Zobacz warunki i postanowienia

**Loty** Odprawa online Hotele Hertz

☒ W obie strony ☐ W jedną stronę Czy mogę zmieniać daty swej podróży? ☒

Data odlotu: Edynburg (EDI) 02 Paź 2010

Data powrotu: Poznań (POZ) 02 Paź 2010

Liczba pasażerów: 1 Dorośli 0 Dzieci (poniżej 16 lat) 0 Niemowlęta (poniżej 2 lat)

☐ Przeczytałem/am i akceptuję Warunki korzystania ze strony internetowej Ryanair

**WYSZUKAJ TANIE LOTY**

**WIADOMOŚCI**

- Buy a Ryanair Approved Samsonite Cabin Bag

**Hostele**

**Wynajem aut**  
Save up to **25% OFF**

**Tanie Hotele**  
**BOOKING.COM**

**Transfer Prywatny**

Internet 100%

# Wycieczka z Edynburga do Poznania

---

Następnie szukamy noclegu...

...taniego...

semantic network eve is a ... x Hostels, Backpackers, Host... x

www.hosteltraveler.com/index.php?select\_searchregion=Poznan

drogeria natura pozn... Google Program Szwajcarski - ... polish swiss grants - S... BO Autobusy - BajkiOnlin... lastminute.com: hotel... Inne zakładki

# Hostel Traveler.com

Get Instant Hostel Rates & Availability

User ID [Login](#) [Forgot your password?](#)

[travelers](#) [work with us](#) [reviews](#) [my bookings](#)

Poznan [Back](#)

Arrival: 10 (Oct) 04 2010 Nights: 2

Guests: 2 Rooms/Units: 1 [Search](#)

Find Hotel or City [Quick Find](#)

Hostel Traveler IS  
Secure online hostel booking at worldwide youth hostels, backpackers, and budget hotels.

[My Hostel Travel](#)  
[Popular Hostels](#)

**We are the cheapest for hostels! We guarantee it!**  
While our friends at other hostel booking sites compete over who's cheaper, look no further. Hostel Traveler is the cheapest. Reserve with confidence with our Low Price Refund Guarantee.

[Cheaper than Hostelworld, Hostelbookers, Hostelsclub, and the rest](#)

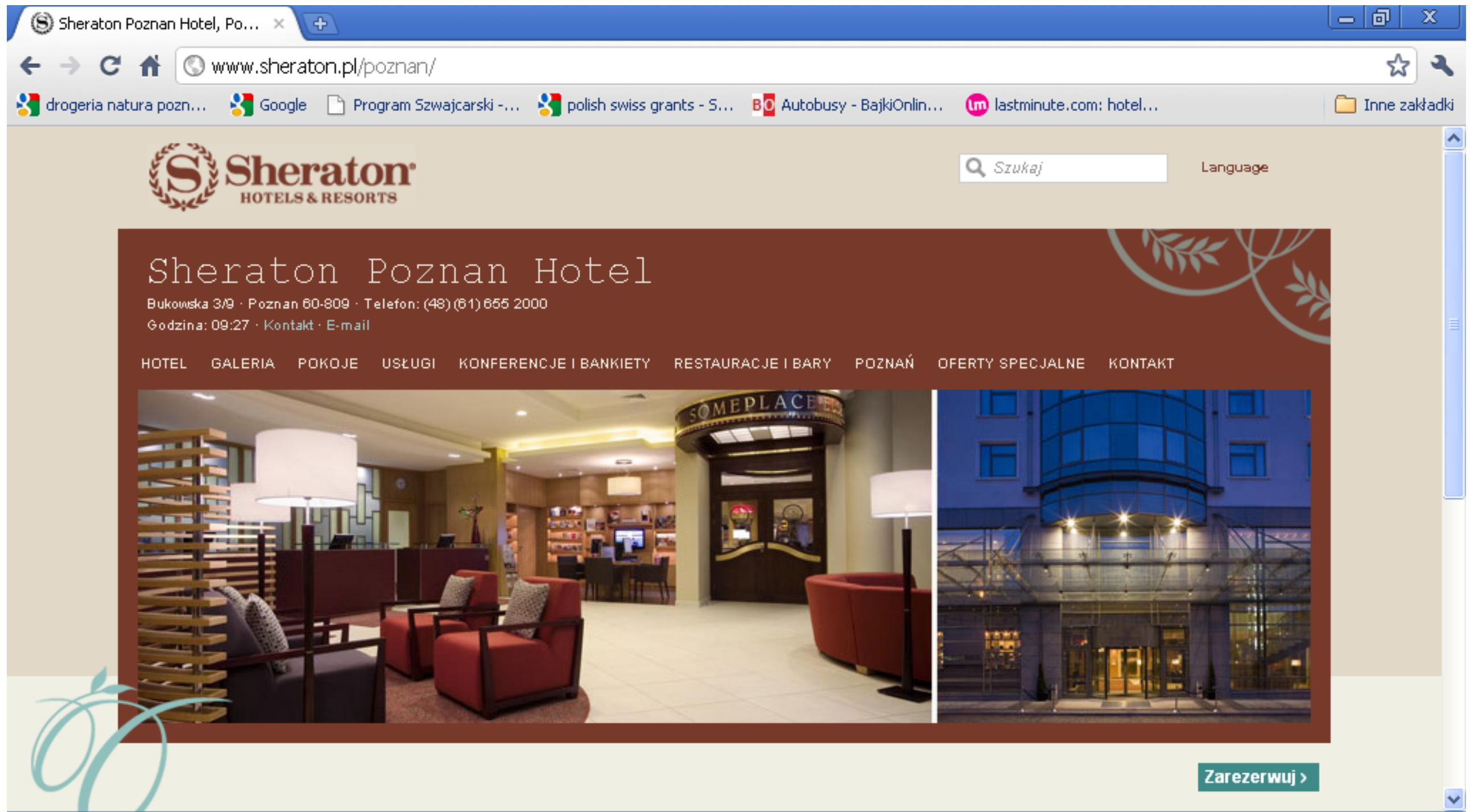
[Lowest prices guaranteed](#)  
[Find a lower price, and we refund double the difference](#)

**Popular cities:** [Dubrovnik](#) [New York](#) [Hong Kong](#) [Florence](#)  
[Rio de Janeiro](#) [Salinas](#) [San Francisco](#) [New York City](#) [Brandon](#) [Honolulu](#)

## Europe

**#1 Search for Hostels**

# lub luksusowego...



# Przykładowy problem 2

## ....może skorzystać z wyspecjalizowanej witryny...

Create your trip from Edinb... x

www.expedia.com/pub/agent.dll?qscr=cmsh&subm=1&CMBTX\_0\_rgnm=poznan&dsct=&dnam=poznan&dlvl=&rtmn=&rtmx=&dcty=

drogeria natura pozn... Google Program Szwajcarski -... polish swiss grants - S... Autobusy - BajkiOnlin... lastminute.com: hotel... Inne zakładki

Welcome - Already a member? [ Sign In ] My Itineraries My Account Customer Support Feedback Search Expedia Go

Expedia

Home Vacation Packages Hotels Cars Flights **NO FEES** Cruises Activities DEALS & OFFERS Maps Business Travel Rewards

Start search over

Change your search

Departing: 11/2/2010

Returning: 11/12/2010

Star Rating: Show all

Name contains: Go

Change Travelers

2 Adults  
1 Room  
[Change travelers](#)

### Create your trip from Edinburgh to Poznan

Maps: [Hotel map view](#) Hotel amenities: [Narrow your search](#)

Not what you're looking for? [Choose a different destination](#)

View trips: 1 - 10 Previous | Next

Sort by: ☒ Expedia Picks ☐ Price ☐ Hotel Name ☐ City ☐ Star Rating ☐ Traveler Opinion

#### Flight + Hotel Trip

**Sheraton Poznan Hotel**

Centrally located in Poznan, this hotel is near Adam Mickiewicz University. Another nearby point of interest is National Museum. [More lodging info](#)

**Hotel promotion - Stay 3 Nights and Save 30% on Your Stay!**

**Insiders' Select**  
2010 WINNER

Book online or call **1-800-222-0892**

★★★★★ Poznan, Poland Check in: 11/03/10  
Classic Room-Promotion Check out: 11/12/10

Total Price: \$2071

**\$1036** per person  
includes: Flight + Hotel, Taxes & Fees\*\*

Select and Continue

Get ThankYou® Points

1-800-222-0892  
Same great rates plus expert advice.



...lub innej...

Poznan Search Results - Tri... x

www.tripadvisor.com/HACSearch?geo=&q=poznan&inDay=8&inMonth=10/2010&checkIn=10/8/2010&outDay=10&outMonth=10/21

drogeria natura pozn... Google Program Szwajcarski -... polish swiss grants - S... Autobusy - BajkiOnlin... lastminute.com: hotel... Inne zakładki

Sign in with Facebook | Sign in | Register Now! |

**tripadvisor** World's most trusted travel advice™

Home Hotels Flights Restaurants Vacation Rentals **NEW** Trip Ideas Write a Review

City, hotel name, etc. Search Over 35 million trusted traveler reviews & opinions

Poznan

- Poznan Tourism
- Poznan Hotels
- Vacation Rentals
- Flights to Poznan
- Poznan Deals
- More ▾

**Refine search**

30 of 57

Price per night  
\$0 - \$350+

advertisement

**Best Western**

NOW EARN DOUBLE OR TRIPLE REWARDS  
AAA MEMBERS EARN EVEN MORE  
BOOK NOW  
Limited Time - Stay Between 9/12 & 11/21

Home → Europe → Poland → Central Poland → Poznan → Poznan Hotels

## Poznan, Central Poland Hotels

Hotels (30) B&Bs / Inns (8) Specialty Lodging (2) Vacation Rentals (2)

**Find available hotels in Poznan** Clear dates

Check-in 10/8/2010 Check-out 10/10/2010 Adults 2 Search Change city

Search by address or point of interest

Other TripAdvisor sites: Cruise Critic SeatGuru See all sites

2010's Best Vacation Rentals »

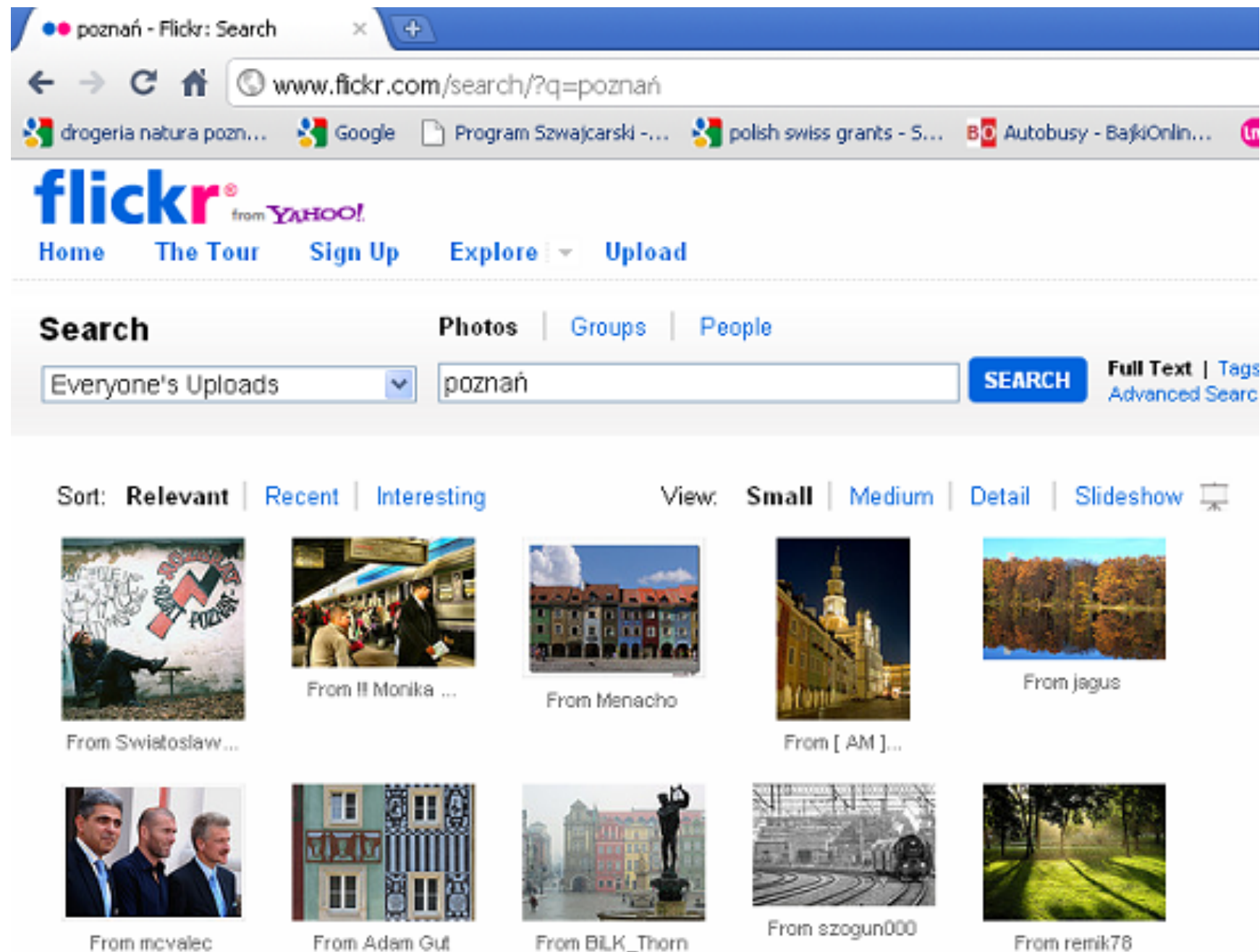
4 października 2010

# Wycieczka z Edynburga do Poznania

---

Być może chcielibyśmy się dowiedzieć czegoś więcej na temat Poznania, poszukać zdjęć...

## ....na flickr....



## ....przez Google...

poznań - Szukaj w Google

www.google.pl/images?hl=pl&source=imghp&biw=1024&bih=481&q=poznań&gbv=2&aq=f&aqi=g10&aql=&oq=&gs\_rfai=

drogeria natura pozn... Google Program Szwajcarski -... polish swiss grants - S... Autobusy - BajkiOnlin... lastminute.com: hotel... Inne zakładki

Sieć Grafika Wideo Mapy Wiadomości Książki Gmail więcej ▼ Ustawienia wyszukiwania | Zaloguj

**Google** poznań Szukaj Filtrowanie umiarkowane ▼

Okolo 2,250,000 wyników (0,21 s) Szukanie zaawansowane

Hasła pokrewne: [lech poznań](#) [poznań koziolki](#) [poznań ratusz](#) [poznań herb](#) [poznań mapa](#)

**Wszystko**  
Mapy  
Wiadomości  
Filmy  
**Grafika**  
Więcej

**Dowolny rozmiar**  
Duże  
Średnie  
Ikony  
Większe niż...  
Dokładnie...

**Dowolny typ**  
Twarz  
Zdjęcie

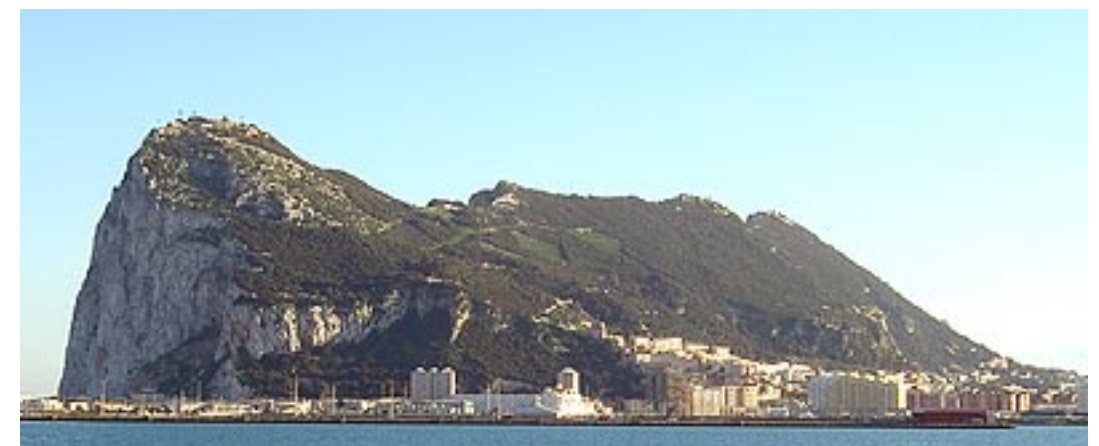


www.google.pl/imgres?imgurl=http://i.wp.pl/a/f/jpeg/22405/poznan-450.jpeg&imgrefurl=http://polskajestfajna.wp.pl/gid,11202923,t,,title,Poznan-za-pol-ceny,galeria.html&usg=\_\_HZ1ePHvhoZqpZwTllu...

# Zlecanie złożonych zadań agentom software'owym



*Zarezerwuj mi wczasy  
na następny weekend,  
gdzieś w miarę blisko w  
łagodnym klimacie z  
urzędowym angielskim*



# Podsumowanie

---

- Zapoznanie się z dużą liczbą stron o różnym stylu, celu, być może też języku...
- Integracja danych częściowo „w myślach” żeby osiągnąć założony cel
- Czasami długi i żmudny proces!
- Na chwilę obecną:
  - maszyny dokonują prezentacji (łatwe)
  - ludzie dokonują integracji i interpretacji danych (trudne)

# Podsumowanie

---

- Każda ze stron tylko „wierzchołkiem góry lodowej”
  - prawdziwe **dane** ukryte w bazach danych, plikach XML, arkuszach kalkulacyjnych, ...
  - dostęp jedynie do tego o czym zdecydowali twórcy strony
- Wyspecjalizowane strony (Expedia, TripAdvisor) idą o krok dalej:
  - gromadzą i agregują dane z różnych źródeł (zazwyczaj za zgodą właścicieli)
  - jednak kontrolują ich prezentację

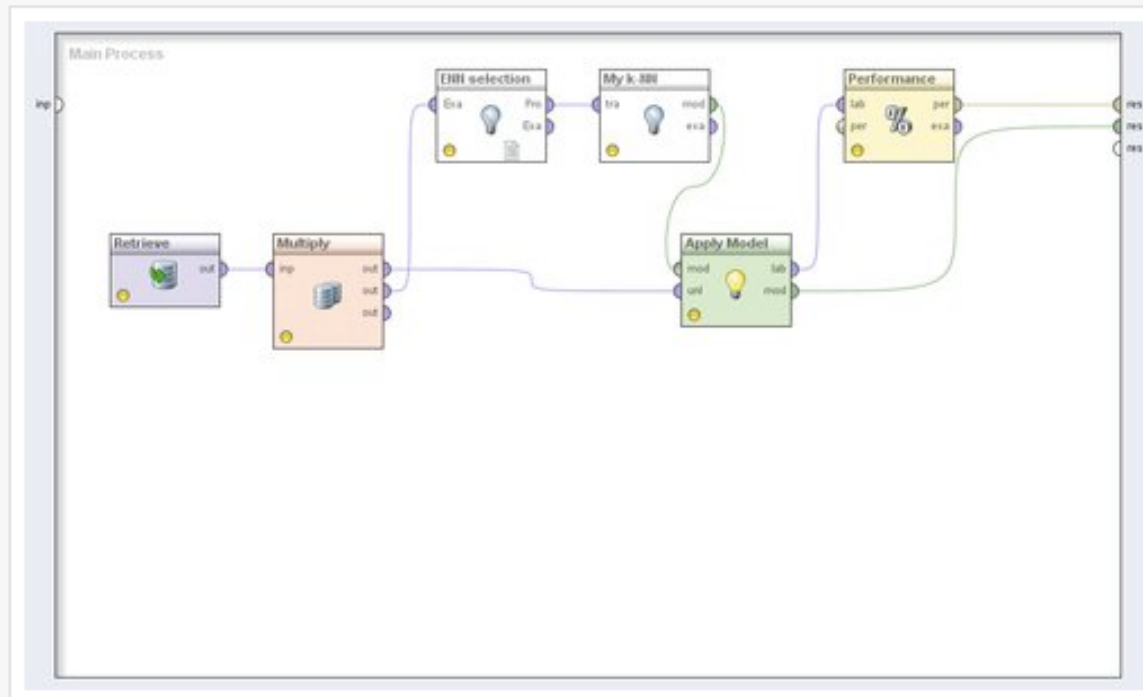
# Kompozycja usług sieciowych

**Title:** Instance Selection and Prototype Based Rules Example 1

**Type:** RapidMiner

## Preview

(Click on the image to get the full size)



[Download Scalable Diagram \(SVG\)](#)

## Description

Example of Instance selection (Requires Instance Selection and Prototype Based Rules plugin):

Before training kNN we do instance selection based on ENN algorithm. The ENN algorithm can be replaced by any other instance selection algorithms from Prules/Selection/\*

Instance selection algorithms from PRules plugin works as View on original ExampleSet



**Marcin Blachnik**

## License

All versions of this Workflow are licensed under:



## Credits (0)

(People/Groups)  
None

## Attributions (0)

(Workflows/Files)  
None

## Tags (1)

☐ Original Uploader tags

**prules plugin**

[Log in](#) to add Tags

## Shared with Groups (0)

None

## Featured In Packs (0)

None

[Log in](#) to add to one of your Packs

## Ratings (0)

Current:

# Podsumowanie

---

- Które usługi mogę połączyć ze sobą w przepływ i w jakiej kolejności?
- Które usługi mogą być użyte zamiennie z daną usługą (realizują podobną funkcję)?
- Które usługi będą działać optymalnie dla konkretnych danych wejściowych i zadania?

# Question answering

“ What does a velociraptor look like ”

This might answer your question:

Input interpretation

Velociraptor Image

Result



# Jeopardy!

---

- *Jeopardy!* to amerykański quiz show (odpowiednik polskiego *Va Banque!*)
- 1964 – do dzisiaj
- format odpowiedź-i-pytanie
- Przykład:
  - Kategoria: Nauka ogólnie
  - Wskazówka: *W zderzeniu z elektronami, fosfor wydziela energię elektromagnetyczną w tej formie*
  - Odpowiedź: Czym jest światło?



# Podsumowanie

---

## Odpowiadanie na pytania

- Dla ludzi wyzwaniem jest **znajomość odpowiedzi**
- Wymaga posiadania **wiedzy** z wielu **dziedzin** i **łączenia ze sobą faktów** aby **wyciągnąć wnioski**
- Dla maszyn wyzwaniem jest także **zrozumienie pytania**

# Technologie semantyczne

## Przegląd tematów TSiSS 2014/2015

---

- Metadane
- Semantyczna Sieć
  - sieci semantyczne
  - ontologie
  - powiązane dane
- Dane strukturalne w HTMLu
- Semantyczne systemy Wiki, crowdsourcing wiedzy  
Bazy i Grafy wiedzy
- Question answering
- Systemy dialogowe, wirtualni personalni asystenci

# Metadane

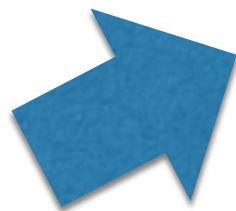
---

**Metadane** = dane, które opisują inne dane

# Biblioteki

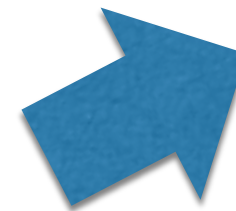


Biblioteka Aleksandryjska  
III w. p.n.e.



Biblioteka Kongresu  
1800 r.

Tezaurus  
Library of Congress  
Subject Headings (LCSH)



**FEDERACJA BIBLIOTEK CYFROWYCH**

Wyszukiwanie Przeglądanie Biblioteki cyfrowe

**Zestawienie polskich bibliotek cyfrowych**

| Lp. | Logo | Nazwa                                                 | Stan     |    |        |       |     |   |
|-----|------|-------------------------------------------------------|----------|----|--------|-------|-----|---|
| 1.  |      | Akademicka Biblioteka Cyfrowa KRAKÓW [wiecej...]      | Dostępna | 2  | 11 882 | 21    | 11% | ✓ |
| 2.  |      | Archiwum Muzyki Węgierskiej [wiecej...]               | Dostępna | 1  | 3 190  | ×     | 0%  | ✓ |
| 3.  |      | Armarium, dominikańska biblioteka cyfrowa [wiecej...] | Dostępna | 3  | 872    | 6     | 94% | ✓ |
| 4.  |      | Batycka Biblioteka Cyfrowa [wiecej...]                | Dostępna | 27 | 40 357 | 3 761 | 99% | ✓ |
| 5.  |      | Baza Rozpraw Doktorskich Uniwersytetu Przyrodniczo-   | Dostępna | 1  | ×      | ×     | ×   | × |

Biblioteki cyfrowe

Dublin Core Schema

# WWW

---

- **Tim Berners-Lee**, fizyk pracujący w CERN w Genewie, stworzył ideę sieci WWW w 1989, obecnie szef W3C
- Dążył do uzyskania szybkiego, elektronicznego dostępu „online” do dokumentów laboratoriów fizyki wysokich energii



**Wynik:** fundamentalna zmiana w sposobie komunikacji międzyludzkiej

# Sieć Semantyczna – Wizja

---

- Oryginalna wizja sieci wg Sir Tim'a BL o wiele bardziej ambitna niż istniejąca obecnie sieć WWW:

“I have a dream for the Web [in which computers] become capable of analyzing all the data on the Web — the content, links, and transactions between people and computers. A [Semantic Web](#), which should make this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines. The intelligent agents people have touted for ages will finally materialize.”

(Berners-Lee, 1999)

# Scientific American, 2001

---



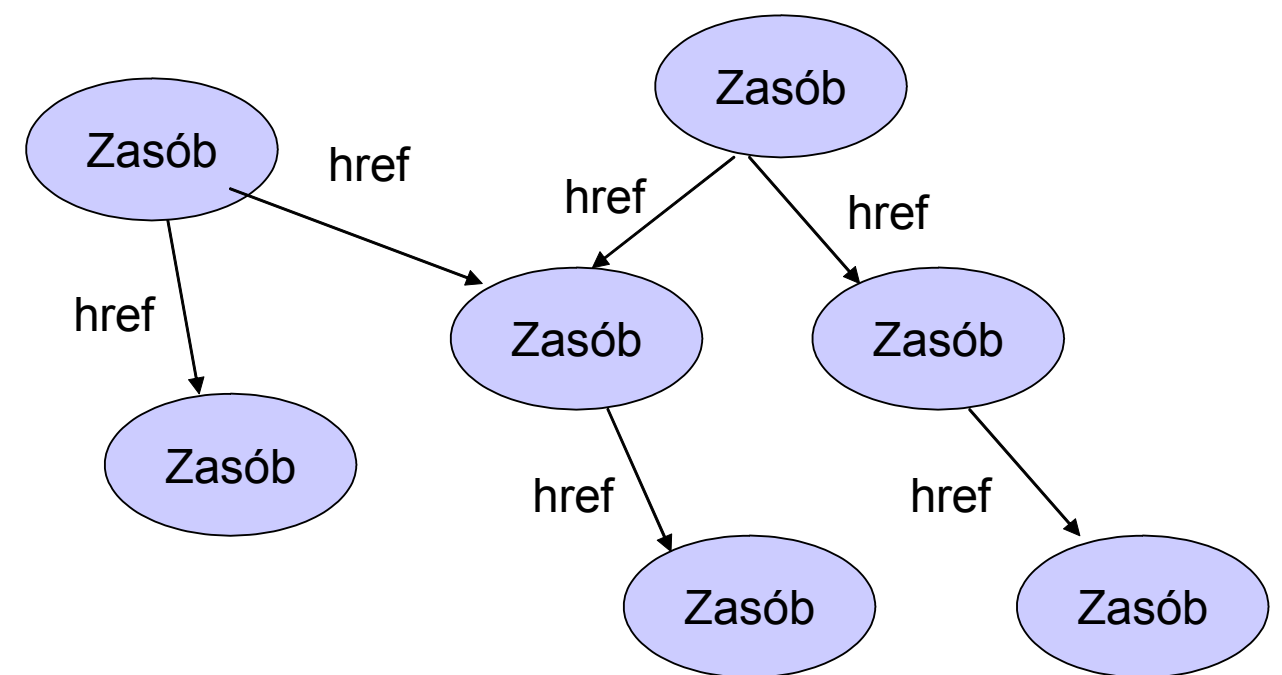
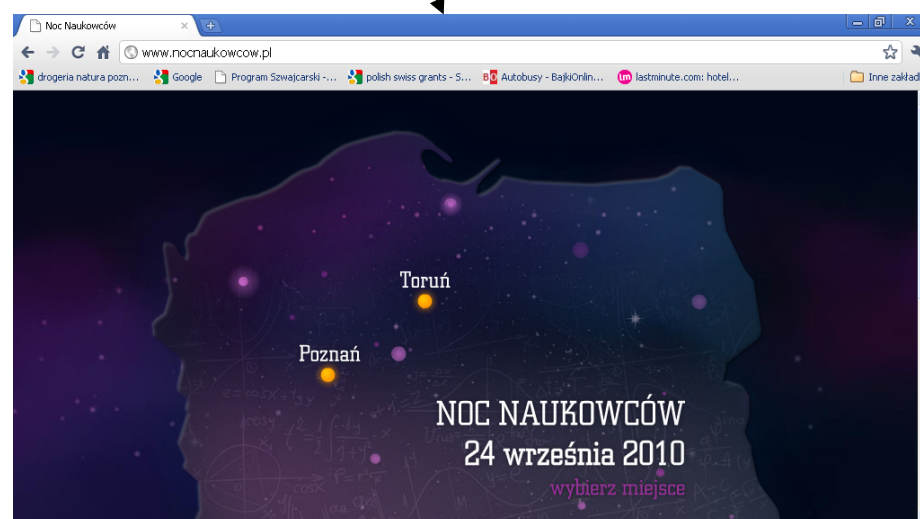
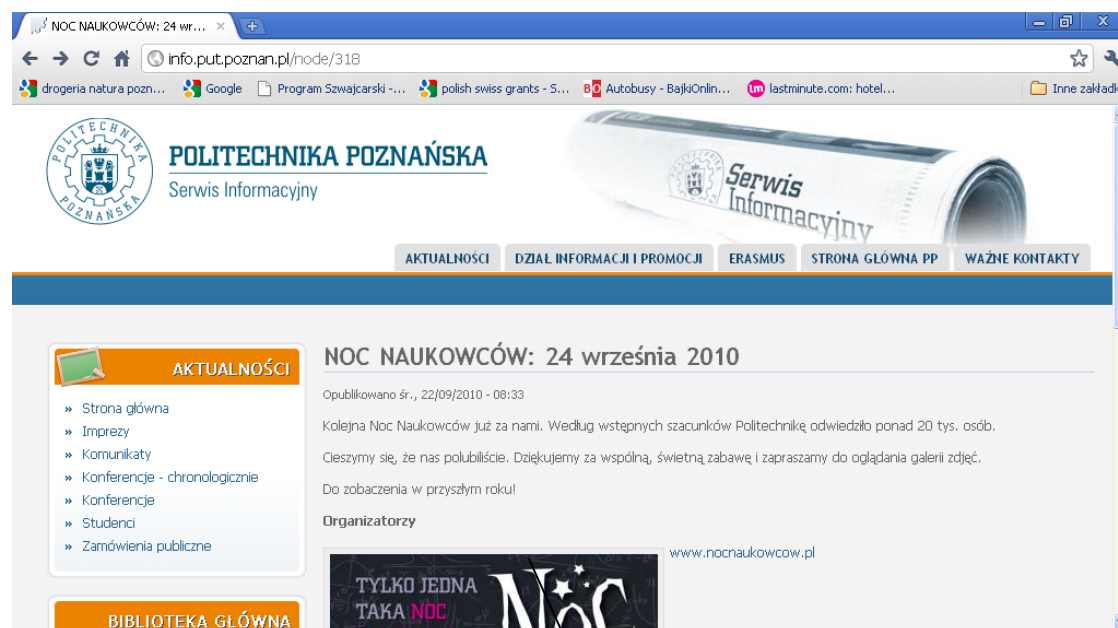
## THE SEMANTIC WEB

A new form of Web content  
that is meaningful to computers  
will unleash a revolution of new abilities

by  
TIM BERNERS-LEE,  
JAMES HENDLER and  
ORA LASSILA

PHOTO CREDIT HERE

# Klasyczna sieć WWW



# „Sieć danych”

---

- czasami potrzeba dostępu do oryginalnych danych i dowolnego ich łączenia!
- wykorzystywać dane w Sieci podobnie jak dokumenty:
  - łączyć dane (niezależnie od ich prezentacji)
  - prezentować, eksplorować
  - wykorzystywać programy agentowe, które mogą częściowo *interpretować* dane
- rozszerzyć aktualną Sieć do Sieci Danych (ang. Web of Data) i umożliwić aplikacjom bezpośrednio wykorzystanie danych

# Klasyczna Sieć WWW - HTML

```
<HTML>
<HEAD>
  <TITLE>BIURO PODROZY</TITLE>
</HEAD>
<BODY bgcolor="ccfff">
<CENTER>
  <H1>Biuro podrozy</H1>
</CENTER>
<BR></BR>
<TABLE BORDER="1" align="center">
  <TR bgcolor="#ffff00">
    <TD>Francja</TD>
    <TD>Nicea</TD>
    <TD>Hilton</TD>
    <TD>***</TD>
    <TD>2500</TD>
    <TD>0001</TD>
  </TR>
```

...

- **HTML** - forma nie sprawiająca problemów w odbiorze **ludziom**
- **maszyny** (agenty) - problem z odróżnieniem np. nazwy miejscowości od nazwy hotelu
- treść generowana z baz danych często prezentowana bez zachowania oryginalnej informacji o jej strukturze

# Klasyczna Sieć WWW - XML

---

- XML – krok w stronę maszynowej interpretacji

```
<biuro_podrozy>
  <oferta numer="0001">
    <miejsce_pobytu>
      <kraj> Francja </kraj>
      <miejscowosc> Nicea </miejscowosc>
      <nazwa_hotelu klasa="***"> Hilton </nazwa_hotelu>
    </miejsce_pobytu>
    <termin>
      ...
    </termin>
    <cena>2500</cena>
  </oferta>
```



# Potrzeba reprezentacji semantyki

---

- XML *sam w sobie* nie zapewnia:
  - specyficznego dla danej dziedziny **słownika**
  - **podstawowych elementów składniowych**
- Wykorzystanie **ontologii** do określenia znaczenia
  - ontologie dostarczają **słownik pojęć**
  - nowe pojęcia mogą być konstruowane poprzez **łączenie** istniejących (“Pojęciowe Lego”)
  - znaczenie (**semantyka**) pojęć jest formalnie określona
  - możliwe określenie relacji między pojęciami z **różnych ontologii**

# Sieć Semantyczna – podsumowanie

---

- dane w sieci, powiązane w taki sposób aby uzyskać **spójną, logiczną sieć danych**
- rozszerzenie obecnej sieci, w którym informacje mają dobrze zdefiniowane **znaczenie (semantykę)**, co ułatwić ma **współpracę** pomiędzy komputerami i ludźmi
- zbiór technologii do realizacji „**Sieci Danych**”

# Jak zrealizować Sieć Semantyczną?

Wizja 1

---

- rozszerzenie aktualnej sieci WWW przez adnotacje obecnych zasobów znacznikami opisującymi znaczenie (semantykę), przez klasyfikację i indeksację zasobów
- wykonane ręcznie lub pół-automatycznie
- zastosowania: personalizacji, wyszukiwanie i przeglądanie w sieci WWW

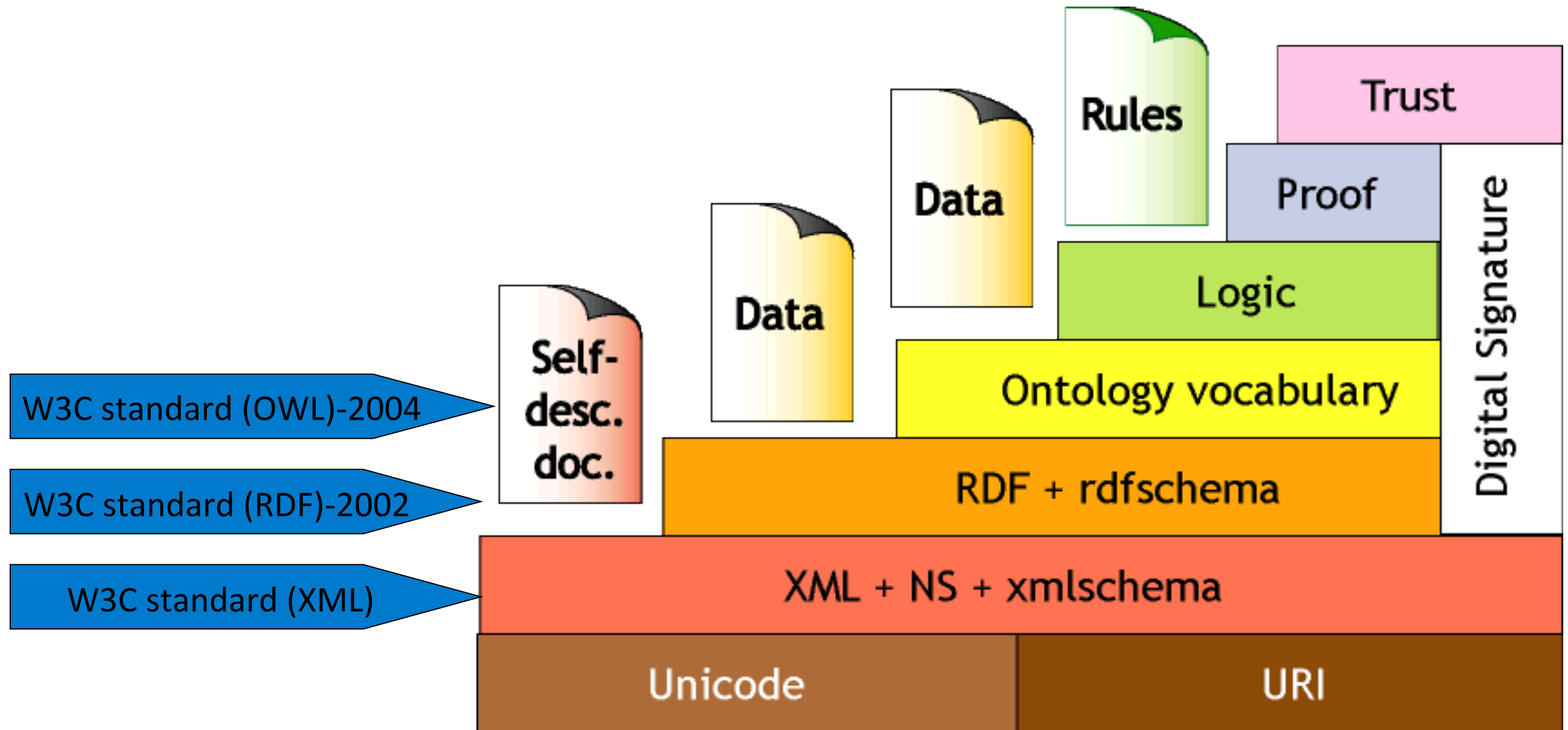
# Jak zrealizować Sieć Semantyczną?

Wizja 2 – Sieć danych („Web of data”)

---

- bliższa oryginalnej wizji Tim’a Berners’a Lee
- ekspozycja do sieci WWW zbiorów danych z baz danych, integracja tych zbiorów
- semantyka schematów baz danych powinna zostać wyeksponowana w sieci WWW razem z danymi
- umożliwienie maszynowej interpretacji, integracji i nieprzewidzianych przypadków użycia
- Sieć Semantyczna = wielka, globalna baza danych

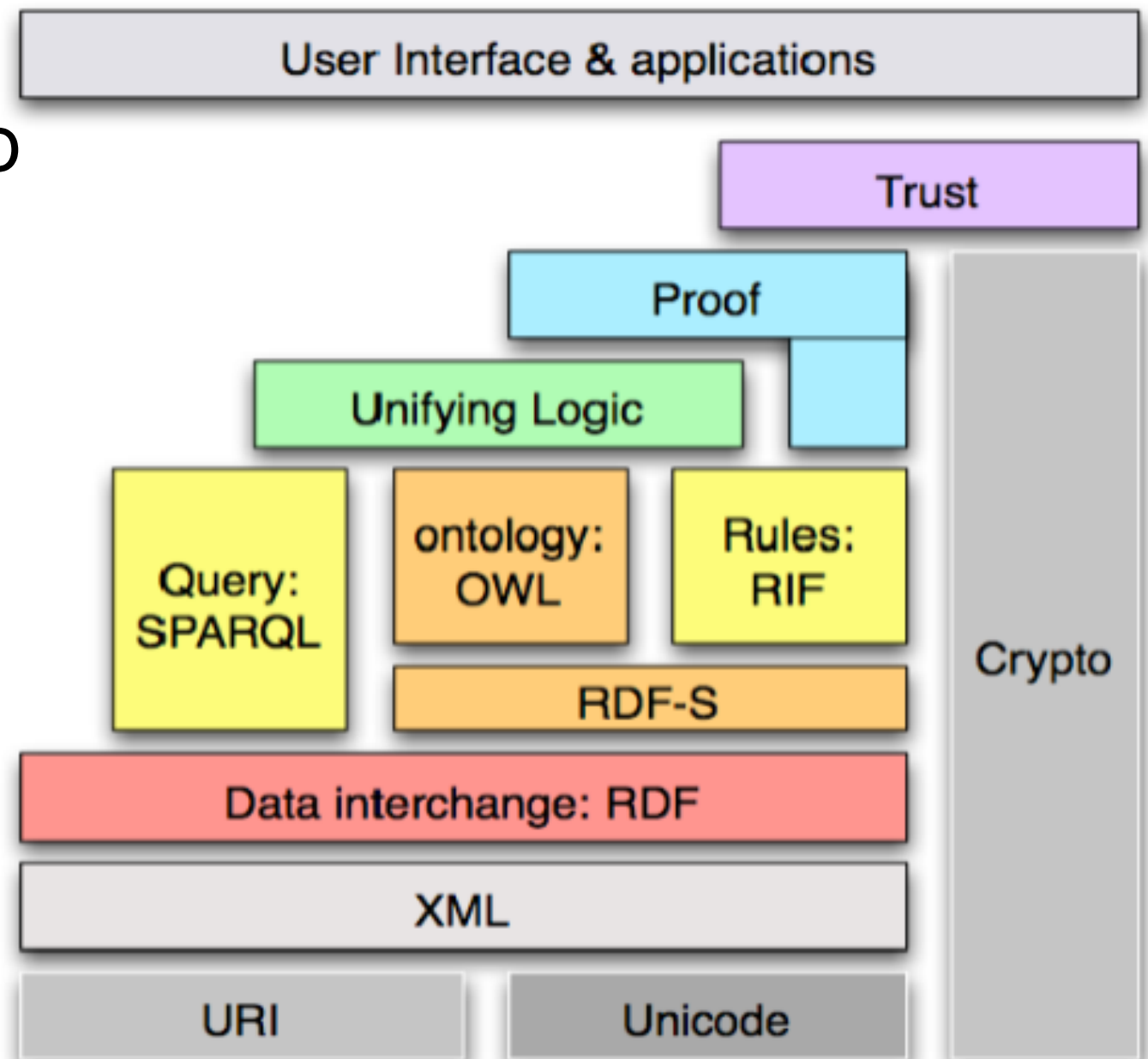
# Warstwy Sieci Semantycznej



*Tim Berners-Lee*

# Bardziej aktualna wersja

- **SPARQL**: język zapytań do RDF, “SQL dla Sieci Semantycznej”, standard W3C 2008
- **RIF** (lub **SWRL**): reguły



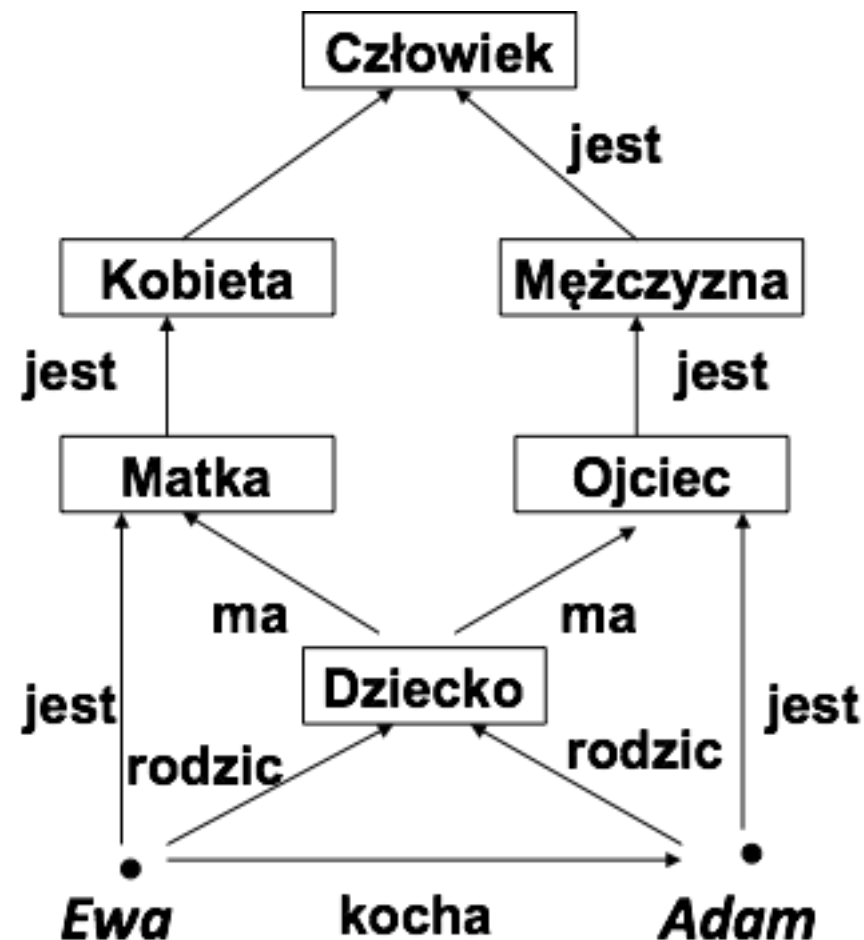
# Sieci semantyczne

---

**Sieci semantyczne** = graficzna notacja do reprezentacji wiedzy reprezentowanej jako zbiór węzłów (pojęć) połączonych etykietowanymi łukami, które reprezentują relacje między węzłami

- w ramach sztucznej inteligencji - początki w latach 60tych XX wieku
- dużo wcześniej wykorzystywane w filozofii, psychologii, lingwistyce

# Sieci semantyczne - przykład



*Ewa* jest *matką*

*Ewa* kocha *Adama*

Każde *dziecko* ma *ojca*

Każda *matka* jest *kobietą*

Każda *kobieta* jest *człowiekiem*

...

Dzisiejsze laboratoria: [RDF \(Resource Description Framework\)](#)

# Ontologia z punktu widzenia filozofa

---

- **Ontologia (metafizyka)** = *nauka o bycie* (Arystoteles, *Metafizyka*, ks. IV)
- *co to jest byt? co charakteryzuje byt?*
- *jak dokonywać klasyfikacji bytów?*



# Ontologia z punktu widzenia informatyka

- “*engineering artefact [...]*” (Guarino 98)

“*An **ontology** is a formal specification of a shared conceptualization of a domain of interest*”  
(Gruber 93)

- ⇒ maszynowa interpretacja
- ⇒ grupa osób, konsensus
- ⇒ abstrakcyjny model zjawisk, pojęcia
- ⇒ wiedza dziedzinowa

- ontologia = **formalna specyfikacja pojęć z danej dziedziny**

## „Powiązane Dane” (ang. Linked Data): definicja

---

*“The Semantic Web isn't just about putting data on the web. It is about making links, so that a person or machine can explore the web of data. With linked data, when you have some of it, you can find other, related, data.”* (Tim Berners-Lee)

**Powiązane Dane** – wykorzystanie technologii Sieci Semantycznej do publikowania ustrukturalizowanych danych w Sieci i do ustanawiania powiązań między źródłami danych

# Projekt „Linking Open Data” (Otwarte Powiązane Dane)

---



- Projekt społecznościowy ze wsparciem W3C
- Cel: Pomoc w wystartowaniu Sieci Semantycznej poprzez publikowanie zbiorów danych z wykorzystaniem RDF.
- Spełnia zasady „połączonych danych” (Linked Data principles)
- Główna idea: wziąć istniejące (otwarte) zbiory danych i uczynić je dostępnymi w Sieci w formacie RDF
- Raz opublikowane w RDF, połączyć je linkami z innymi zbiorami danych
- Przykładowy link RDF: <http://dbpedia.org/resource/Berlin> [Identyfikator Berlina w DBPedia] owl:sameAs <http://sws.geonames.org/2950159> [Identyfikator Berlina w Geonames].



# RDFa

**RDFa** = „RDF w atrybutach”, rekomendacja W3C dodająca rozszerzenia na poziomie atrybutów do HTML, XHTML, i różnych typów dokumentów XML w celu osadzania bogatych metadanych w dokumentach Web



```
<div
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:gr="http://purl.org/goodrelations/v1#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:v="http://rdf.data-vocabulary.org/#"
  typeof="gr:Offering" about="#product_data">
  <span property="gr:name">Hepp Research Personal SCSI Controller Card</span>
  <span rel="foaf:depiction">
    
  </span>
  <span rel="v:hasReview">
    <span typeof="v:Review-aggregate" about="#review_data">
      Average rating: <span property="v:average" datatype="xsd:string">4.9</span>,
      based on <span property="v:count" datatype="xsd:string">99</span> reviews
    </span>
  </span>
  <span rel="gr:hasPriceSpecification">
```

# Mikrodane

**Mikrodane** = specyfikacja HTML grupy WHATWG używana do zagnieżdżania metadanych w ramach istniejących treści na stronach internetowych



## Thing > Intangible > Offer

An offer to transfer some rights to an item or to provide a service—for example, an offer to sell tickets to an event, to rent the DVD of a movie, to stream a TV show over the internet, to repair a motorcycle, or to loan a book.

For GTIN-related fields, see [Check Digit calculator](#) and [validation guide](#) from [GS1](#).

| Property                                     | Expected Type                     | Description                                                                                                                                                  |
|----------------------------------------------|-----------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Properties from <a href="#">Offer</a></b> |                                   |                                                                                                                                                              |
| <a href="#">acceptedPaymentMethod</a>        | <a href="#">PaymentMethod</a>     | The payment method(s) accepted by seller for this offer.                                                                                                     |
| <a href="#">addOn</a>                        | <a href="#">Offer</a>             | An additional offer that can only be obtained in combination with the first base offer (e.g. supplements and extensions that are available for a surcharge). |
| <a href="#">advanceBookingRequirement</a>    | <a href="#">QuantitativeValue</a> | The amount of time that is required between accepting the offer and the actual usage of the resource or service.                                             |
| <a href="#">aggregateRating</a>              | <a href="#">AggregateRating</a>   | The overall rating, based on a collection of reviews or ratings, of the item.                                                                                |
| <a href="#">availability</a>                 | <a href="#">ItemAvailability</a>  | The availability of this item—for example In stock, Out of stock, Pre-order, etc.                                                                            |
| <a href="#">availabilityEnds</a>             | <a href="#">DateTime</a>          | The end of the availability of the product or service included in the offer.                                                                                 |
| <a href="#">availabilityStarts</a>           | <a href="#">DateTime</a>          | The beginning of the availability of the product or service included in the offer.                                                                           |
| <a href="#">availableAtOrFrom</a>            | <a href="#">Place</a>             | The place(s) from which the offer can be obtained (e.g. store locations).                                                                                    |
| <a href="#">availableDeliveryMethod</a>      | <a href="#">DeliveryMethod</a>    | The delivery method(s) available for this offer.                                                                                                             |

# Wikipedia

[illegible]

# Semantyczne Wiki



nawigacja

[Strona główna](#)[Hokej](#)[Fotogalerie](#)[Losuj stronę](#)[Ostatnie zmiany](#)[Portal użytkowników](#)[Kontakt](#)[Bibliografia](#)

Narzędzia

[Linkujące](#)[Zmiany w linkowanych](#)[Strony specjalne](#)[Wersja do druku](#)[Link do tej wersji](#)[Informacje o tej stronie](#)[Przeglądaj właściwości](#)

Znajdź nas

[Założ nowe konto](#) [Zaloguj się](#)Artykuł [Dyskusja](#)Czytaj [Wyświetl historię](#)[Przejdź](#)[Szukaj](#)

## WikiPasy.pl

Encyklopedia Klubu Sportowego Cracovia

Piłka nożna - [Hokej](#)

Dołącz do nas! Na przepisanie czekają relacje z 549 meczów.



Aktualności

[\[edytuj\]](#)

- **09.12.2013** - W ostatnim meczu na własnym stadionie w tym roku, Pasy pokonały [Zagłębie Lubin](#) 2:0.
- **04.12.2013** - Trzecie w sezonie ale dopiero drugie w obecności widzów zwycięstwo na "Ziemii Świętej". Cracovia pokonała [Lechię Gdańsk](#) 2:1.
- **01.12.2013** - Koniec wyjazdowej serii spotkań bez porażki. W meczu 18.kolejki ekstraklasy Cracovia przegrała w Chorzowie z [Ruchem](#) 2:3
- **25.11.2013** - Cracovia przegrała z [poznajskim Lechem](#) 1:6. Jest to [najwyższa porażka](#) przed własną publicznością w historii występów Cracovii w pierwszej klasie rozgrywkowej.
- **19.11.2013** - Młodzieżowa Reprezentacja Polski pokonała na [Stadionie Cracovii](#) rówieśników z Grecji 3:1. Zobacz [fotogalerię](#) z meczu.
- **15.11.2013** - W sparingowym spotkaniu Pasy pewnie wygrały ze słowackim [MFK Ružomberok](#) 3:0.
- **09.11.2013** - Cracovia zremisowała z [Piastem](#) w Gliwicach 1:1.
- **02.11.2013** - [Korona Kielce](#) przerwała zwycięską passę Pasów i zwyciężyła przy ul. Kałuży 2:1.
- **28.10.2013** - Czwarte zwycięstwo z rzędu! Cracovia wygrała w wyjazdowym spotkaniu z [Górnikiem Zabrze](#) 1:0!
- **25.10.2013** - Po kolejnym wyjazdowym spotkaniu Pasy wracają do Krakowa z trzema punktami! [Widzew Łódź](#) - Cracovia 1:3.
- **20.10.2013** - Na własnym stadionie przy pustych trybunach Cracovia wygrała 4:2 z [bielskim Podbeskidziem](#).
- **12.10.2013** - [Stadion Cracovii](#) ponownie gościł Młodzieżową Reprezentację Polski. W swoim czwartym meczu Eliminacji Mistrzostw Europy Polska pokonała Szwecję 2:0. Zobacz [fotogalerię](#) z meczu.
- **11.10.2013** - W przerwie ligowej, pozbawiona czołowych zawodników Cracovia przegrała z [MFK Ružomberok](#) 2:0.
- **05.10.2013** - W wyjazdowym spotkaniu Pasy pokonały [Jagiellonię Białystok](#) 2:1.

### Terminarz (runda jesienna)

 15.12, n, 18:00, [Legia Warszawa](#) (w)

Mamy już  
**9883**  
artykułów.

### Sezony:

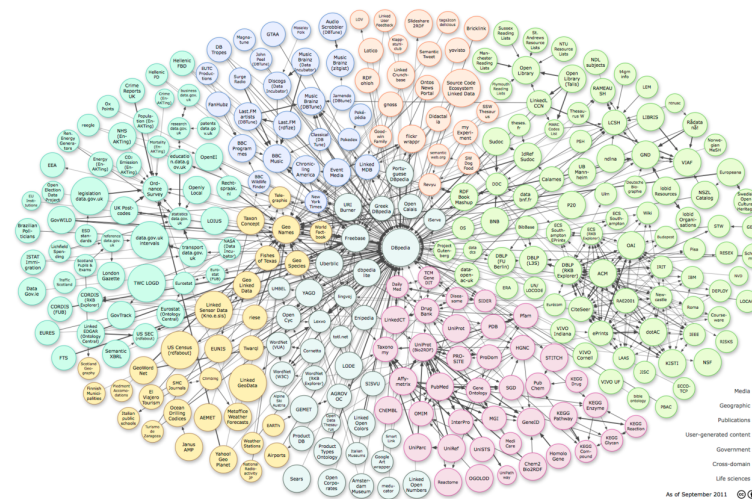
2013/14  
2012/13  
2011/12  
2010/11  
2009/10  
2008/09  
2007/08  
2006/07  
2005/06  
2004/05  
2003/04  
2002/03  
2001/02  
2000/01  
1999/00  
1998/99  
1997/98  
1996/97  
1995/96  
1994/95  
1993/94  
1992/93  
1991/92  
1990/91  
1989/90  
1988/89  
1987/88  
1986/87  
1985/86  
1984/85  
1983/84  
1982/83  
1981/82  
1980/81

# Baza wiedzy – pragmatyczna definicja

---

**Baza wiedzy**  $\approx$  kolekcja encji, klas i faktów o postaci „subject-predicate-object” (atrybuty, relacje), istotnych ogólnie lub w danej dziedzinie, która jest kompleksowa, zorganizowana semantycznie i zdatna do odczytu maszynowego

# Grafy wiedzy



Facebook Entity Graph



Microsoft Bing Satori



# Google: Graf Wiedzy

---

- **semantyczne wyszukiwanie**
- maj 2012: **baza wiedzy** wykorzystywana przez Google do **rozszerzenia wyników wyszukiwania**
  - wiele źródeł wiedzy: *CIA World Factbook*, *Freebase*, *Wikipedia*
  - sieć semantyczna zawiera ponad 570 mln obiektów i ponad 18 mld faktów
- maj 2013: polska wersja językowa; **zadawanie pytań** raczej niż wyszukiwanie, **informacje i powiązania** między nimi raczej niż zestaw linków
  - system poszukujący nie fraz kluczowych, lecz "bytów" stojących za wpisanymi w wyszukiwarce słowami

## Dwa główne sposoby działania Grafu Wiedzy

---

- dopasowywanie odpowiedzi do kontekstu; w przypadku dwuznacznych haseł prezentacja różnych wersji odpowiedzi
- podsumowania najbardziej istotnych informacji: -  
biogramy, wyróżnione najważniejsze elementy,  
powiązania między kluczowymi hasłami, odnośniki  
do kolejnych informacji

# Graf Wiedzy: przykład



[Internet](#) [Grafika](#) [Mapy](#) [Więcej ▾](#) [Narzędzia wyszukiwania](#)

Okolo 824,000 wyników (0,22 s)

[Maria Skłodowska-Curie – Wikipedia, wolna encyklopedia](#)  
[pl.wikipedia.org/wiki/Maria\\_Skłodowska-Curie ▾](#)  
**Maria** Salomea Skłodowska-Curie (ur. 7 listopada 1867 w Warszawie, zm. 4 lipca 1934 w Paryżu) – uczona polsko-francuska, fizyczka, chemiczka, dwukrotna ...  
[Życiorys](#) - [Ciekawostki](#) - [Wyróżnienia i nagrody](#) - [Upamiętnienie](#)

[Marie Curie - Wikipedia, the free encyclopedia](#)  
[en.wikipedia.org/wiki/Marie\\_Curie ▾](#) [Tłumaczenie strony](#)  
**Marie Curie** (7 November 1867 – 4 July 1934), née Maria Salomea Skłodowska was a Polish physicist and chemist, working mainly in France, who is famous for ...

[Uniwersytet Marii Curie-Skłodowskiej w Lublinie](#)  
[www.umcs.lublin.pl/ ▾](#)  
Serwis Uniwersytetu Marii Curie - Skłodowskiej; informacje o studiach pierwszego i drugiego stopnia, jednolitych magisterskich, międzywydziałowych, ...

[Marie Curie Actions - Research Fellowship Programme - European ...](#)  
[ec.europa.eu ▸ ... ▸ Research & Innovation ▾](#) [Tłumaczenie strony](#)  
Marie Curie Actions are European research grants for scientists in Europe and abroad.

[Maria Curie-Skłodowska \(1867-1834\) - Historia Kobiet](#)  
[historia\\_kobiet.w.interia.pl/teksty/sklodowska.html ▾](#)  
**Maria Curie-Skłodowska** (1867-1834) wybitna fizyczka i chemiczka polska, żyjąca i pracująca we Francji, pierwsza kobieta będąca profesorem Sorbony, ...

[Maria Skłodowska-Curie – Wikicytaty](#)  
[pl.wikiquote.org/wiki/Maria\\_Skłodowska-Curie ▾](#)  
**Maria** Salomea Skłodowska-Curie (1867–1934) – polsko-francuska fizyczka i chemiczka, z wykształcenia także matematyczka, dwukrotna laureatka Nagrody ...

[Marie Curie - Biography - Nobelprize.org](#)  
[www.nobelprize.org ▸ ... ▸ Nobel Prize in Physics ▾](#) [Tłumaczenie strony](#)  
Short profile from the foundation that awards the Nobel Prize.

[Ludzie \(People\)](#)

[Więcej obrazów](#)

## Maria Skłodowska-Curie

Maria Salomea Skłodowska-Curie – uczona polsko-francuska, fizyczka, chemiczka, dwukrotna noblistka. Większość życia spędziła we Francji, tam też studiowała, a następnie rozwinęła swoją karierę naukową. Prekursora nowej gałęzi chemii – radiochemii. [Wikipedia](#)

**Data i miejsce urodzenia:** 7 listopada 1867, [Warszawa](#)

**Data i miejsce śmierci:** 4 lipca 1934, [Paryż, Francja](#)

**Odkrycia:** [Polon](#), [Rad](#)

**Wykształcenie:** [Sorbona](#) (1903), [Sorbona](#) (1894), [Sorbona](#) (1891–1893)

**Dzieci:** [Irène Joliot-Curie](#), [Ève Curie](#)

**Nagrody:** Nagroda Nobla w dziedzinie fizyki, Nagroda Nobla w dziedzinie chemii

### Podobne wyszukiwania

[Pierre Curie](#)[Henri Becquerel](#)[Albert Einstein](#)[Ernest Rutherford](#)[Irène Joliot-Curie](#)

# IBM Watson

---

- Watson – system komputerowy stworzony przez IBM do odpowiadania na pytania zadawane w języku naturalnym
- Watson wystąpił w Jeopardy! w trzydniowej rozgrywce (2011) ...
- <http://www.youtube.com/watch?v=Puhs2LuO3Zc>

# IBM Watson

---

- ...
- przeciwnikami IBM Watsona byli:
  - Brad Rutter – do tej pory wygrał najwięcej pieniędzy,
  - Ken Jennings – był najdłużej niepokonanym mistrzem
- IBM Watson zajął pierwsze miejsce

# Problem automatycznego i niezależnego od dziedziny odpowiadania na pytania (QA)

---

- Mając dane
  - „treściwe” pytania w języku naturalnym dot. szerokiej dziedziny wiedzy
- Dostarcz (w czasie  $< 3s$ ):
  - precyzyjnych odpowiedzi: określ czego dotyczy pytanie & daj precyzyjną odpowiedź
  - dokładnie wyliczoną pewność odpowiedzi
  - strawne wyjaśnienia co do poprawności odpowiedzi

# IBM Watson – „trójząb”\*

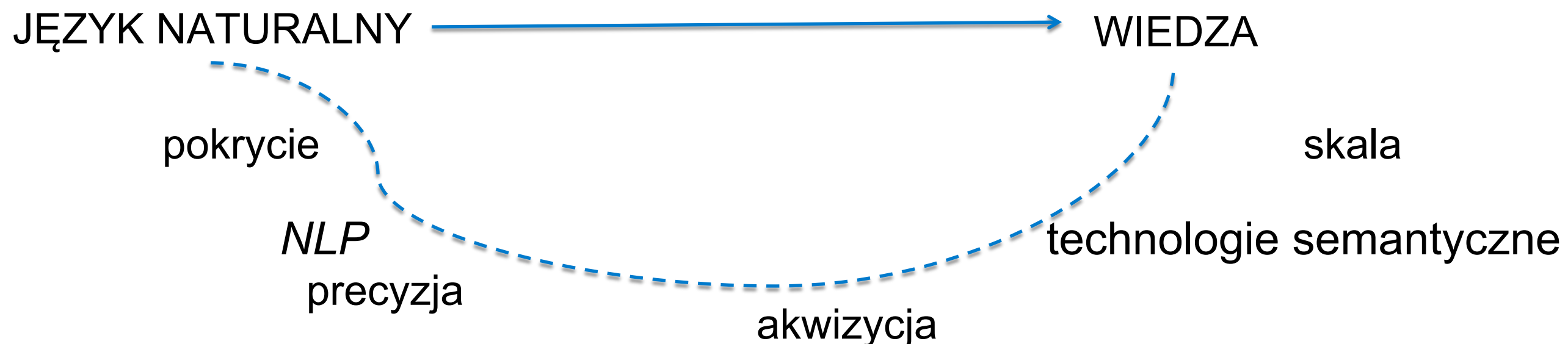
---

- Nowy paradygmat oprogramowania
  - coraz więcej zadań obliczeniowych wymaga rozwiązań niedokładnych, które łączą wiele metod w nieprzewidziany sposób
- Wiedza nie jest celem (o tym za chwilę)
- Inteligencja maszynowa *nie* jest inteligencją ludzką
- Różnica jest najbardziej znaczna w przypadku pomyłek

# IBM Watson – wiedza nie jest celem

- **Klasyczne podejście QA**

- Od zarania SI zakładano, że odpowiadanie na pytania będzie działać na bazie procesu, który całkowicie przekłada język naturalny na jednoznaczną (logiczną) reprezentację; proces wnioskowania będzie działać na tej reprezentacji aby wyprodukować odpowiedzi.



# IBM Watson – wiedza nie jest celem

- **Klasyczne podejście QA**

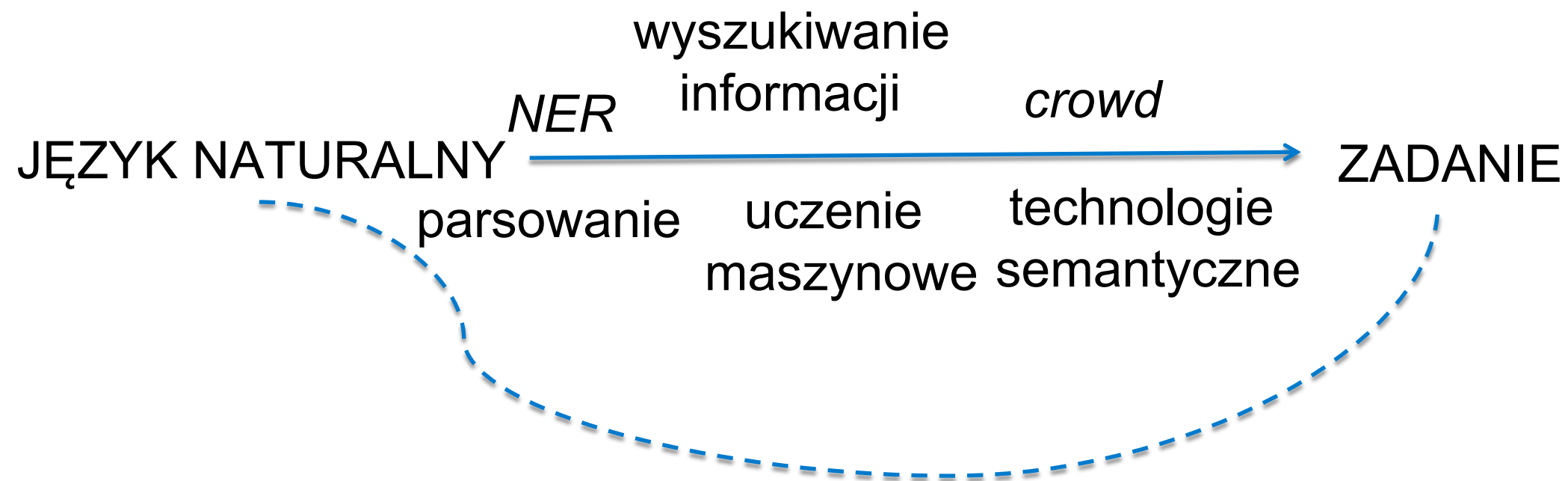
- Od zarania SI zakładano, że odpowiadanie na pytania będzie działać na bazie procesu, który całkowicie przekłada język naturalny na jednoznaczną (logiczną) reprezentację; proces wnioskowania będzie działać na tej reprezentacji aby wyprodukować odpowiedzi.



# IBM Watson – wiedza nie jest celem

- **DeepQA (Watson)**

- generuje i ocenia wiele hipotez wykorzystując kolekcję metod z dziedziny przetwarzania języka naturalnego, uczenia maszynowego, reprezentacji wiedzy i wnioskowania;
- gromadzą one i wąż dowody pochodzące ze źródeł danych niestukturalnych i strukturalnych (np. otwartych powiązanych danych) aby ustalić odpowiedź o najwyższej pewności na podstawie odpowiedzi wielu (setek) metod



# IBM Watson – jak to działa

---

- The science behind an answer

<http://www.youtube.com/watch?v=DywO4zksfXw>

## Systemy dialogowe

**System dialogowy** = system komputerowy, przeznaczony do rozmowy z człowiekiem w spójny sposób

### Wirtualny doradca

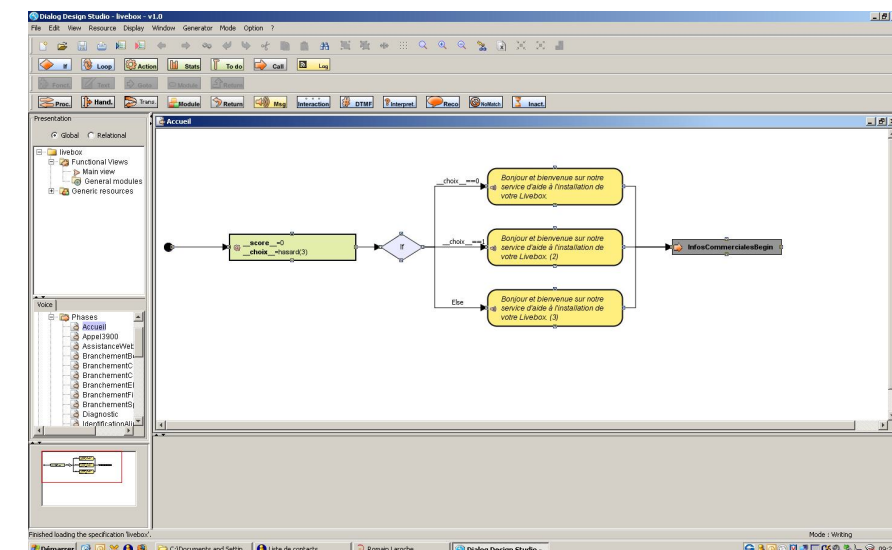
#### W samochodzie



<https://www.youtube.com/watch?v=fC57kONNzDk>



### Telekomunikacja, call centers



# Dziękuję za uwagę

---



"Now! ... That should clear up  
a few things around here!"