

Reprezentacja wiedzy

ontologie, logiki deskrypcyjne

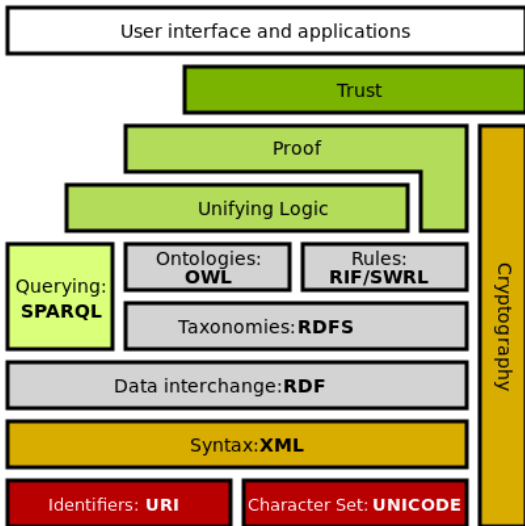
Agnieszka Ławrynowicz

28 listopada 2018

Plan wykładu

- 1 Powtórka: sieci semantyczne, RDF
- 2 Definicja ontologii
- 3 Logiki deskrypcyjne

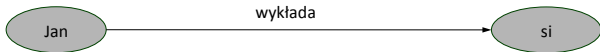
Semantyczny Internet - stos



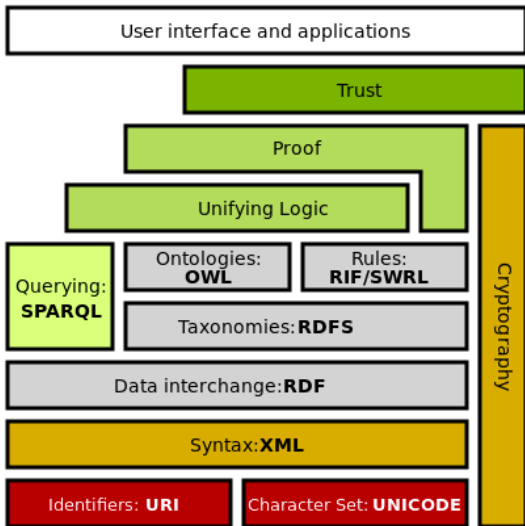
- sieci semantyczne w technologii WWW
- "trójkowy" model danych: podmiot, orzeczenie, dopełnienie

Example

```
<http://example.org/Jan> <http://example.org/wyklada> <http://dbpedia.org/ontology/si> .
```



Semantyczny Internet - stos

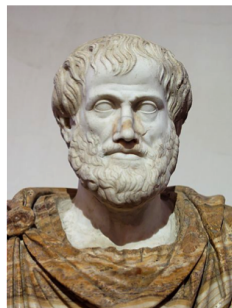


Ontologia w filozofii

Ontologia (metafizyka)

nauka o bycie

- co to jest byt? co charakteryzuje byt?
- jak dokonywać klasyfikacji bytów?



Arystoteles, Metafizyka, ks. IV

Ontologia w informatyce

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

“An *ontology* is a
formal specification
of a shared
conceptualization
of a domain of interest”

(Gruber 93, Studer 98)

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

Ontologia

formalna specyfikacja pojęć (najczęściej z określonej dziedziny)

Ontologia w informatyce

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

“An *ontology* is a
formal specification
of a *shared*
conceptualization
of a *domain of interest*”
(Gruber 93, Studer 98)

- ⇔ *maszynowa interpretacja*
- ⇔ *konsensus* grupy osób
- ⇔ abstrakcyjny model zjawisk, *pojęcia*
- ⇔ wiedza *dziedzinowa*

Ontologia

formalna specyfikacja pojęć (najczęściej z określonej dziedziny)

Ontologia w informatyce

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

“An *ontology* is a
formal specification
of a *shared*
conceptualization
of a *domain of interest*”

(Gruber 93, Studer 98)

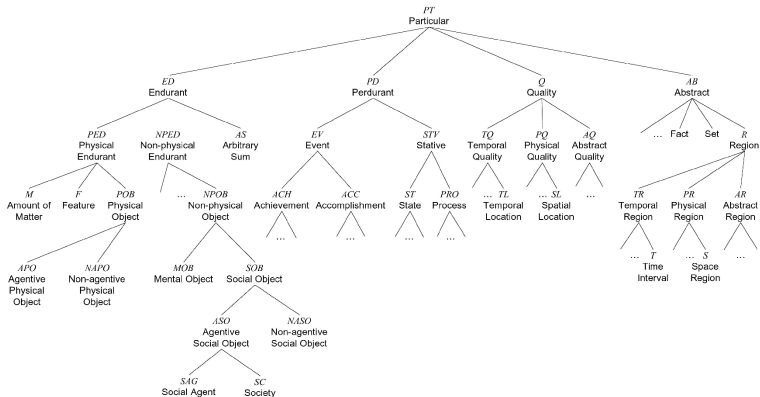
- ⇔ *maszynowa interpretacja*
- ⇔ *konsensus* grupy osób
- ⇔ abstrakcyjny model zjawisk, *pojęcia*
- ⇔ wiedza *dziedzinowa*

Ontologia

formalna specyfikacja pojęć (najczęściej z określonej dziedziny)

Przykłady ontologii: nadrzędne

DOLCE, BFO



Przykłady ontologii: złożone w dziedzinie biomedycyny

Obo Foundry (Open Biological and Biomedical Ontologies), w tym GO (Gene Ontology), Snomed CT, NCI, Galen

BioPortal Browse Search Mappings Recommender Annotator Resource Index Projects Sign In Help Feedback

Follow @bioontology 1,879 followers Lubio! Cite Us

Welcome to BioPortal, the world's most comprehensive repository of biomedical ontologies. Current Release: 4.29 (October 2016)

For help using BioPortal, click on this icon: ?

Search all ontologies

Enter concept, e.g. Melanoma Search

Advanced Search

Find an ontology

Enter ontology name, e.g. NCI Thesaurus Explore

Browse Ontologies >

Search resources

Enter a concept, e.g. Melanoma Search

Advanced Resource Search

Ontology Visits (October 2016)

| | |
|---|-------|
| Current Procedural Terminology (CPT) | 49961 |
| Medical Dictionary for Regulatory Activities (MEDDRA) | 18835 |
| RxNORM (RXNORM) | 14379 |
| Systematized Nomenclature of Medicine - Clinical Terms (SNOMEDCT) | 9949 |
| National Drug Data File (NDDF) | 6548 |
| More | |

Statistics

| | |
|----------------------------------|-----------------|
| Ontologies | 529 |
| Classes | 7,798,227 |
| Resources indexed | 48 |
| Indexed Records | 39,359,542 |
| Direct Annotations | 95,468,433,792 |
| Direct Plus Expanded Annotations | 144,789,582,932 |

Latest Notes

What is measured is carbon isotope composition (Plant Trait Ontology)
7 months ago by BrenOil
I would suggest that there should be subcategories for "Carbon Isotope Discrimination", as what I...

New Relationship Proposal: centiMorgan is by definition dimensionless (Units of Measurement Ontology)
11 months ago by grimaldi
centiMorgan is not a well defined length unit (Units of Measurement Ontology)
11 months ago by grimaldi
Dear curators, As one cannot give conversion factors to other length units, centiMorgan should ...

Grade definition (Common Terminology Criteria for Adverse Events)
about 1 year ago by hpotier
Hi, We would like for users to be able navigate down to the complete definition for each grade I...

Grade definition (Common Terminology Criteria for Adverse Events)
about 1 year ago by hpotier
Hi, We would like for users to be able navigate down to the complete definition for each grade I...

Latest Mappings

Alzheimer's disease (CRISP) <=> CDKS (ADO)
REST Mapping 10/24/2016 by SAM-PETERS

CDKS (ADO) <=> Alzheimer's disease (CRISP)
REST Mapping 10/24/2016 by SAM-PETERS

Malnutrition (MEDLINEPLUS) <=> Malnutrition (NCIT)
REST Mapping 09/29/2016 by Anagruz

Malnutrition (NCIT) <=> Malnutrition (MEDLINEPLUS)
REST Mapping 09/29/2016 by Anagruz

Cancer (LOINC) <=> Leukaemias (MEDDRA)
REST Mapping 07/11/2016 by mdoef

THE NATIONAL CENTER FOR BIOMEDICAL ONTOLOGY

Navigation icons: back, forward, search, etc.

Przykłady ontologii: "lekkie" w e-commerce

schema.org

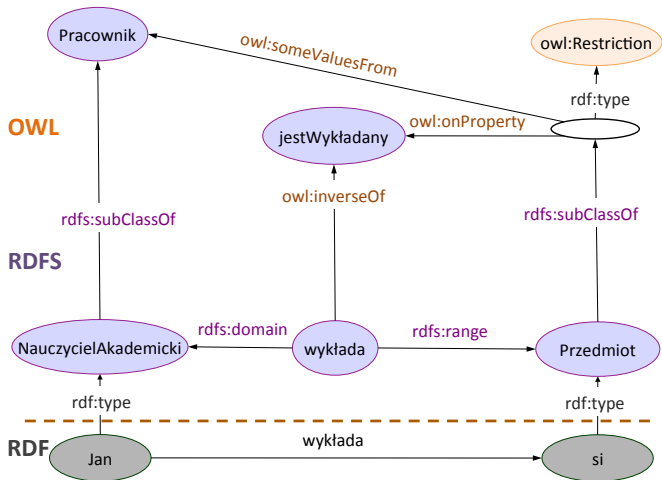
The screenshot shows the schema.org website interface. At the top, there is a red header with the 'schema.org' logo on the left, a search bar in the center, and navigation links for 'Home', 'Schemas', and 'Documentation' on the right. Below the header, the main content area is titled 'Person' in red. Underneath, it says 'Thing > Person' and provides a definition: 'A person (alive, dead, undead, or fictional)'. It also notes 'Usage: Over 1,000,000 domains' and a '[more...]' link. The core of the page is a table listing properties of the 'Person' class, with columns for 'Property', 'Expected Type', and 'Description'. The table lists various properties such as 'additionalName', 'address', 'affiliation', 'alumniOf', 'award', 'birthDate', 'birthPlace', 'brand', 'children', 'colleague', and 'contactPoint', each with its corresponding data type and a brief description.

| Property | Expected Type | Description |
|---|---|--|
| Properties from Person | | |
| additionalName | Text | An additional name for a Person, can be used for a middle name. |
| address | PostalAddress or Text | Physical address of the item. |
| affiliation | Organization | An organization that this person is affiliated with. For example, a school/university, a club, or a team. |
| alumniOf | EducationalOrganization or Organization | An organization that the person is an alumni of. Inverse property: alumni . |
| award | Text | An award won by or for this item. Supersedes awards . |
| birthDate | Date | Date of birth. |
| birthPlace | Place | The place where the person was born. |
| brand | Brand or Organization | The brand(s) associated with a product or service, or the brand(s) maintained by an organization or business person. |
| children | Person | A child of the person. |
| colleague | Person or URL | A colleague of the person. Supersedes colleagues . |
| contactPoint | ContactPoint | A contact point for a person or organization. Supersedes contactPoints . |

Ontologia w OWL

- **Encje**: klasy, własności, indywidua i wszelkie inne elementy modelowanej dziedziny
- **Wyrażenia**: złożone pojęcia
- **Aksjomaty**: twierdzenia, które są prawdziwe w modelowanej dziedzinie

Ontologie + dane = graf wiedzy



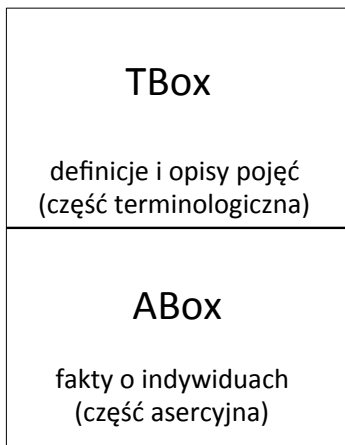
Logiczna interpretacja OWL

Logiki deskrypcyjne

rodzina języków będących podzbiorem rachunku predykatów do reprezentacji wiedzy szczególnie w postaci **terminologii (definicji pojęć)** i stanowiących fundament teoretyczny języka do reprezentacji ontologii OWL

- języki logik deskrypcyjnych są rozróżniane poprzez konstruktory (operatory logiczne) jakie oferują
- \mathcal{AL} ("attributive language"): minimalny język o praktycznym znaczeniu
- nazwy innych języków tworzy się na bazie zbiorów konstruktorów jakie dany język wspiera

Baza wiedzy w logikach deskrypcyjnych



TBox: schemat bazy wiedzy
ABox: dane

Budowanie pojęć za pomocą logiki deskrypcyjnej

Podstawowe części składowe: **pojęcia**, **własności**, **konstruktory**, **indywidua**

| | |
|----------|--|
| Przykład | <p>TBox</p> <p>Atomowe pojęcie: Pracownik, Przedmiot Własności: wykłada, jestWykładany Konstruktory: \sqcap, \exists Aksjomaty (definicje pojęcia): NauczycielAkademicki \equiv Pracownik \sqcap \existswykłada.Przedmiot Aksjomaty (deskrypcje): "każdy nauczyciel akademicki to pracownik": NauczycielAkademicki \sqsubseteq Pracownik "każdy przedmiot jest wykładany przez pracownika": Przedmiot \sqsubseteq \existsjestWykładany.Pracownik Aksjomaty (charakterystyki własności): własność odwrotna wykłada \sqsubseteq $\bar{\ }$jestWykładany</p> |
| | <p>ABox</p> <p>Asercja faktu: wykłada(Jan, si)</p> |

Konstruktory pojęć języka *ALC*

| Konstruktor | Składnia | Przykład |
|------------------------------|---------------|--|
| pojęcie (klasa) uniwersalne | \top | |
| pojęcie (klasa) puste | \perp | |
| iloczyn | $C \sqcap D$ | <i>Dziecko</i> \sqcap <i>Mężczyzna</i> |
| suma | $C \sqcup D$ | <i>Matka</i> \sqcup <i>Ojciec</i> |
| negacja | $\neg C$ | \neg <i>Mężczyzna</i> |
| kwantyfikator egzystencjalny | $\exists R.C$ | $\exists ma$ <i>Dziecko.Syn</i> |
| kwantyfikator ogólny | $\forall R.C$ | $\forall ma$ <i>Dziecko.Syn</i> |

Interpretacja

Interpretacja $\mathcal{I}=(\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ składa się z:

- $\Delta^{\mathcal{I}}$: **dziedzina interpretacji** (niepusty zbiór)
- $\cdot^{\mathcal{I}}$: **funkcja interpretacji**, przypisuje:
 - każdemu atomowemu pojęciu A zbiór $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$,
 - każdej atomowej własności R relację binarną $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$

Semantyka \mathcal{ALC}

| Konstruktor | Składnia | Semantyka |
|------------------------------|---------------|---|
| pojęcie uniwersalne | \top | $\Delta^{\mathcal{I}}$ |
| pojęcie puste | \perp | \emptyset |
| iloczyn | $C \sqcap D$ | $C^{\mathcal{I}} \cap D^{\mathcal{I}}$ |
| suma | $C \sqcup D$ | $C^{\mathcal{I}} \cup D^{\mathcal{I}}$ |
| negacja | $\neg C$ | $\Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$ |
| kwantyfikator egzystencjalny | $\exists R.C$ | $\{a \in \Delta^{\mathcal{I}} \mid \exists b. (a, b) \in R^{\mathcal{I}} \wedge b \in C^{\mathcal{I}}\}$ |
| kwantyfikator ogólny | $\forall R.C$ | $\{a \in \Delta^{\mathcal{I}} \mid \forall b. (a, b) \in R^{\mathcal{I}} \rightarrow b \in C^{\mathcal{I}}\}$ |

Pojęcia (klasy)

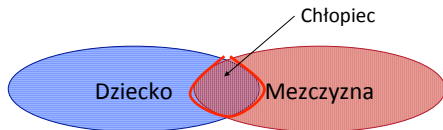
zbiory instancji

```
:NauczycielAkademicki rdf:type owl:Class
```

Iloczyn

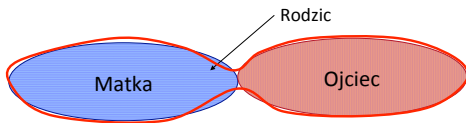
Dziecko \cap Mężczyzna

```
owl:intersectionOf ( :Dziecko  
                    :Mężczyzna  
                    )
```



Matka \sqcup Ojciec

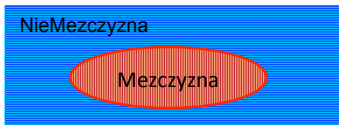
```
owl:unionOf ( :Matka  
              :Ojciec  
            )
```



Negacja (dopełnienie)

→ Męczyzna

`owl:complementOf :Męczyzna`



Uwaga na założenie otwartego świata (omówione dalej)

Kwantyfikator egzystencjalny

"każdy nauczyciel akademicki musi wykładać co najmniej jeden przedmiot"

$\text{NauczycielAkademicki} \sqsubseteq \exists \text{wykłada.Przedmiot}$

```
:NauczycielAkademicki rdf:type owl:Class ;  
    rdfs:subClassOf [ rdf:type owl:Restriction ;  
                    owl:onProperty :wykłada ;  
                    owl:someValuesFrom :Przedmiot  
                    ] .
```

Kwantyfikator ogólny

"asystenci prowadzą tylko laboratoria"

Asystent \sqsubseteq prowadzi.Laboratorium

```
:Asystent rdf:type owl:Class ;  
  rdfs:subClassOf [ rdf:type owl:Restriction ;  
                    owl:onProperty :prowadzi ;  
                    owl:allValuesFrom :Laboratorium  
                  ] .
```

Kwantyfikator ogólny c.d.

Uwaga! Może to dotyczyć także asystentów, którzy nie prowadzą żadnych zajęć! Wynika to ze znaczenia kwantyfikatora ogólnego w logice pierwszego rzędu.

"wszystkie moje worki z pieniędzmi leżą na tym stole"



Aksjomaty

- zawieranie się (subsumcja)
- równoważność
- rozłączność

Deskrypcje (opisy) versus definicje

deskrypcje (warunki konieczne)

Asystent $\sqsubseteq \forall$ prowadzi.Laboratorium

```
:Asystent rdf:type owl:Class ;  
  rdfs:subClassOf [ rdf:type owl:Restriction ;  
    owl:onProperty :prowadzi ;  
    owl:allValuesFrom :Laboratorium  
  ] .
```

definicje (warunki konieczne i wystarczające)

Chłopiec \equiv Dziecko \sqcap Mężczyzna

```
:Chłopiec rdf:type owl:Class ;  
  owl:equivalentClass [ rdf:type owl:Class ;  
    owl:intersectionOf ( :Dziecko  
      :Mężczyzna  
    ) ] .
```

Rozłączność

Dopóki nie są wprowadzone jawnie **ograniczenia rozłącznościowe**, klasy mogą mieć część wspólną

```
:NauczycielAkademicki owl:disjointWith :Przedmiot
```

Wnioskowanie

- **Subsumcja:** Czy C jest podpojęciem pojęcia D ?
- **Spełnialność:** Czy (złożone) pojęcie C nie jest sprzeczne?
- **Spójność:** Czy ABox \mathcal{A} jest niesprzeczny?
- **Populacja/sprawdzenie przynależności do klasy:** Czy a jest instancją pojęcia C mając dany ABox \mathcal{A} ?

Algorytm tablic semantycznych dla języka *ALC*

Cel: na wejściu mając pojęcie w języku *ALC* zdecyduj czy jest ono spełnialne

- metoda tablic semantycznych
- próba konstrukcji modelu
- strukturalne rozłożenie pojęcia na części
- wnioskowanie o nowych ograniczeniach na modelu
- drzewo przeszukiwania, reguły

Algorytm tablic semantycznych dla języka

ALC

| Reguła | Warunek | Akcja |
|------------------|---|--|
| reguła \sqcap | \mathcal{L} zawiera $x : (C_1 \sqcap C_2)$, ale nie oba $x : C_1$ i $x : C_2$ | $\mathcal{L}' := \mathcal{L} \cup \{x : C_1, x : C_2\}$ |
| reguła \sqcup | \mathcal{L} zawiera $x : (C_1 \sqcup C_2)$, ale ani $x : C_1$ ani $x : C_2$ | $\mathcal{L}' := \mathcal{L} \cup \{x : C_1\}$, $\mathcal{L}'' := \mathcal{L} \cup \{x : C_2\}$ |
| reguła \exists | \mathcal{L} zawiera $x : (\exists R.C)$, ale nie ma indywiduum z , takiego, że $z : C$ i $(x, z) : R$ są w \mathcal{L} | $\mathcal{L}' := \mathcal{L} \cup \{y : C, (x, y) : R\}$, gdzie y jest indywiduum nie występującym w \mathcal{L} |
| reguła \forall | \mathcal{L} zawiera $x : (\forall R.C)$ i $(x, y) : R$, ale nie $y : C$ | $\mathcal{L}' := \mathcal{L} \cup \{C(y)\}$ |

Algorytm tablic semantycznych dla języka *ALC*

Drzewo wyvodu:

- **Sprzeczne:** drzewo przeszukiwania zawiera sprzeczność jeśli posiada węzeł z \perp lub $\{A, \neg A\}$
- **Domknięte:** drzewo przeszukiwania jest domknięte jeśli nie można już zastosować żadnych reguł

Algorytm jest inicjalizowany wstawieniem danego pojęcia do korzenia drzewa

Daje odpowiedź "tak" jeśli uzyskamy domknięte i nie sprzeczne drzewo przeszukiwania

Algorytm tablic semantycznych dla języka \mathcal{ALC} - przykład

Mamy dane na wejściu pojęcie $C_0 = (A \sqcap \neg A) \sqcup B$

Zaczynamy od inicjalizacji korzenia drzewa $x_0 := \mathcal{L}(x_0) = \{C_0\}$

$$\mathcal{L}_0 = \{x : (A \sqcap \neg A) \sqcup B\}$$

reguła \sqcup

reguła \sqcup

$$\mathcal{L}_1 = \mathcal{L}_0 \cup \{x : A \sqcap \neg A\} \quad \mathcal{L}_1^* = \mathcal{L}_0 \cup \{x : B\}$$

reguła \sqcap

$$\mathcal{L}_2 = \mathcal{L}_1 \cup \{x : A, x : \neg A\} \text{ sprzeczność}$$

Do \mathcal{L}_1^* nie można już zaaplikować więcej reguł i nie zawiera sprzeczności, więc pojęcie C_0 jest spełnialne

Model \mathcal{I} : $\Delta^{\mathcal{I}} = \{x\}$, $B^{\mathcal{I}} = \{x\}$, $A^{\mathcal{I}} = \emptyset$

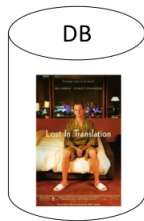
"Świat zamknięty" kontra "świat otwarty"

- **Zamknięty świat** (programowanie w logice, bazy danych)
 - kompletna wiedza o indywidualach
 - brak informacji jest informacją negatywną (negation-as-failure)
- **Otwarty świat** (logika deskrypcyjna, Sieć Semantyczna)
 - niekompletna wiedza o indywidualach
 - negacja faktu musi być jawnie podana (monotonic negation)

"Świat zamknięty" kontra "świat otwarty" - przykład

Założmy, że w bazie mamy następujące dane:

OscarMovie(lostInTranslation)
Director(sofiaCoppola)
creates(sofiaCoppola, lostInTranslation)



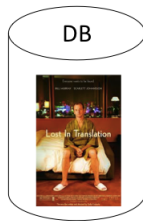
Czy wszystkie filmy Sofii Coppoli są filmami oskarowymi?

TAK - zamknięty świat

"Świat zamknięty" kontra "świat otwarty" - przykład

Założmy, że w bazie mamy następujące dane:

OscarMovie(lostInTranslation)
Director(sofiaCoppola)
creates(sofiaCoppola, lostInTranslation)



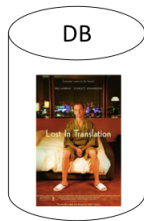
Czy wszystkie filmy Sofii Coppoli są filmami oskarowymi?

TAK - zamknięty świat

"Świat zamknięty" kontra "świat otwarty" - przykład

Założmy, że w bazie mamy następujące dane:

**OscarMovie(lostInTranslation)
Director(sofiaCoppola)
creates(sofiaCoppola, lostInTranslation)**



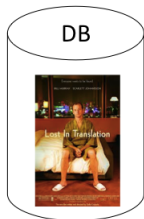
Czy wszystkie filmy Sofii Coppoli są filmami oskarowymi?

TAK - zamknięty świat

"Świat zamknięty" kontra "świat otwarty" – przykład

Założmy, że w bazie mamy następujące dane::

OscarMovie(lostInTranslation)
Director(sofiaCoppola)
creates(sofiaCoppola, lostInTranslation)



Czy wszystkie filmy Sofii Coppoli są filmami oskarowymi?

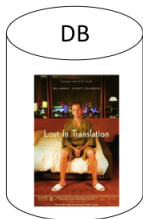
TAK - zamknięty świat **NIE WIEM** - otwarty świat

Różne wnioski!

"Świat zamknięty" kontra "świat otwarty" – przykład

Założmy, że w bazie mamy następujące dane::

OscarMovie(lostInTranslation)
Director(sofiaCoppola)
creates(sofiaCoppola, lostInTranslation)



Czy wszystkie filmy Sofii Coppoli są filmami oskarowymi?

TAK - zamknięty świat NIE WIEM - otwarty świat

Różne wnioski!

Brak założenia o unikalności nazw

JFK

John F. Kennedy

John Fitzgerald Kennedy

wszystkie (różne) nazwy mogą oznaczać ten sam obiekt

Dziękuję za uwagę!