

Znalost a sémantika



ISKM89 Organizace dat - sémantický web | podzim 2023
Zuzana Nevěřilová | Centrum zpracování přirozeného jazyka

Data - Informace - Znalost - Moudrost - (a dál)

Data = symboly, signály

Data existují, nemají žádný účel.

Nicméně, počítačová data vznikají sběrem (capture)

Capta

Informace = data, která mají smysl

Informace nemusí být užitečné

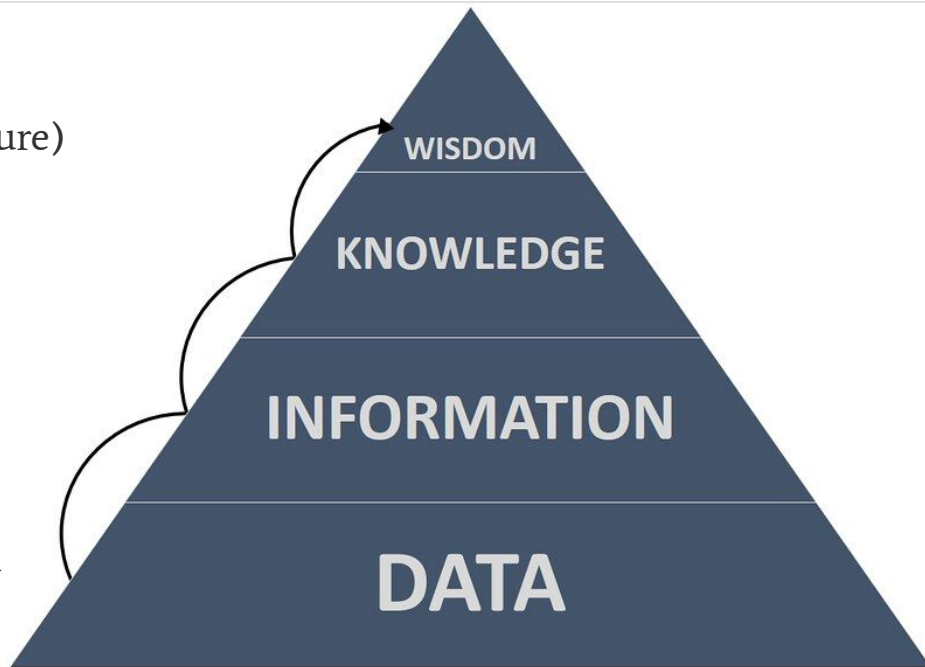
Znalosti = kolekce užitečných informací

Znalosti obsahují záměr

Moudrost = porozumění znalostem

Moudrost umožňuje použít znalosti v budoucnu

https://cs.wikipedia.org/wiki/DIKW_pyramida



<https://www.flickr.com/photos/ryan2point0/34764532445>

Znalost

Znalosti = kolekce užitečných informací
Znalosti obsahují záměr

Pravdy (Truths) a přesvědčení (Beliefs)

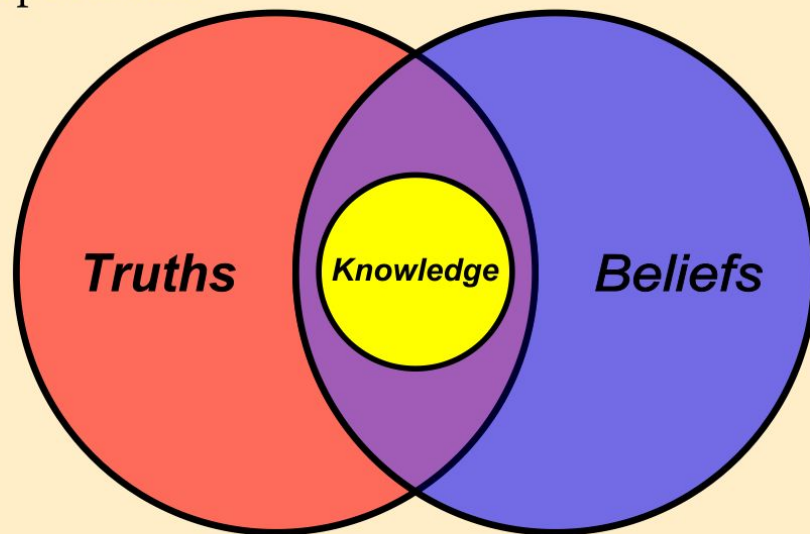
důvodné pravdivé přesvědčení - **Justified True Belief (JTB)**

A subject S knows that a proposition P is true if and only if:

1. P is true, and
2. S believes that P is true, and
3. S is justified in believing that P is true

https://en.wikipedia.org/wiki/Gettier_problem

Propositions



<https://www.flickr.com/photos/ryan2point0/34764532445>

Sdílená znalost - když “mluvíme stejným jazykem”

Jazyk: syntax, sémantika (sdílené znaky, sdílené struktury, sdílené koncepty)

Koncepty: taxonomie (hierarchie konceptů)

Asociace a relace: thesaurus (jak lze alternativně něco nazvat)

Ontologie: pravidla o relacích (které jsou povolené a “dávají smysl”)

Mluvíme stejným jazykem

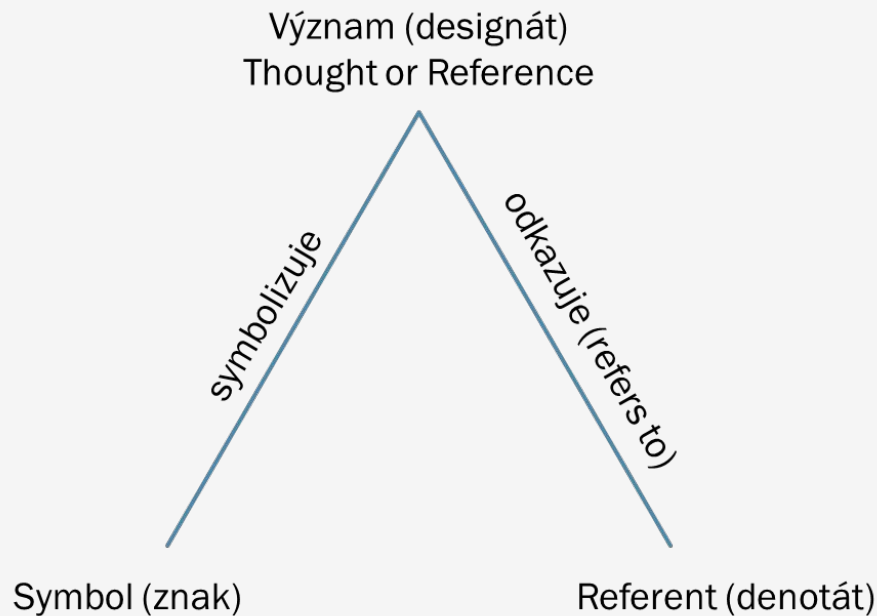
Když o něčem mluvíme, rozumíme tytéž koncepty?

Odgen-Richardsův trojúhelník

Slova nejsou koncepty

Slova jsou nálepky

Koncepty nejsou objekty v našem světě.



Mluvíme stejným jazykem

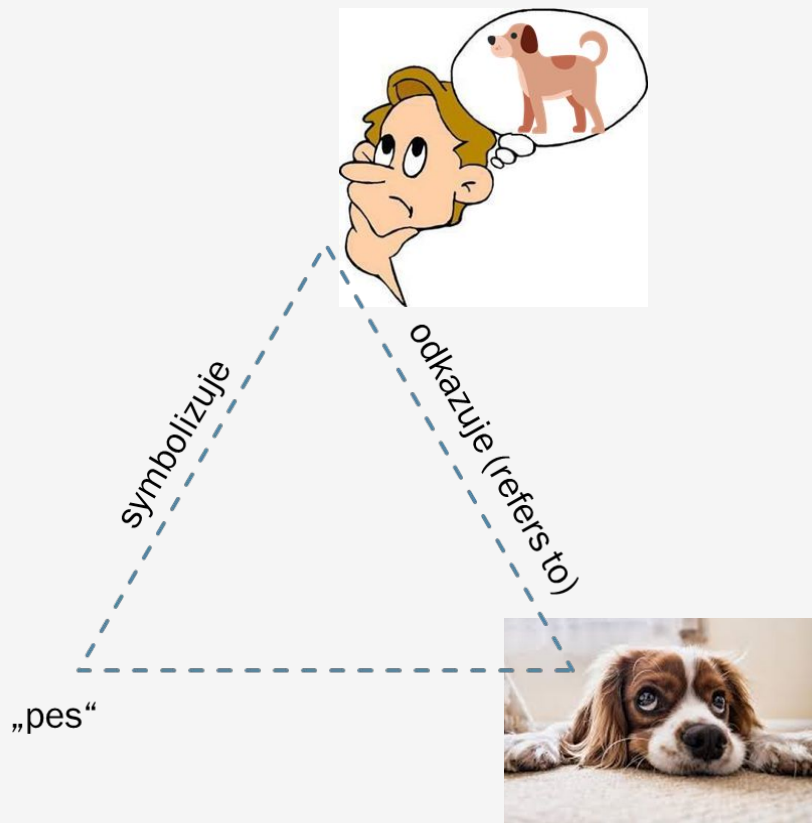
Když o něčem mluvíme, rozumíme tytéž koncepty?

Odgen-Richardsův trojúhelník

Slova nejsou koncepty

Slova jsou nálepky

Koncepty nejsou objekty v našem světě.



Mluvíme stejným jazykem?

Když o něčem mluvíme, rozumíme tytéž koncepty?

Víceznačnost (jeden znak ukazuje na více konceptů)

Parafráze (jeden koncept můžeme vyjádřit více způsoby)

-tel

Koruna

Kolej

Praštil se sluchátkem.

I saw a man with a telescope.

Šeřík. Bez.

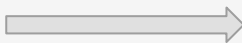
Mnohem více, než si dovedeme představit

Mluvíme stejným jazykem?

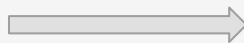
Přirozený jazyk:

- morfologie
- syntax
- lexikum
- sémantika
- pragmatika

} gramatika



struktura



interpretace v kontextu

kontext = “okolí” komunikace
záměr



Formální popis faktů - abychom rozuměli totéž

Logika (FOPL, deskripční logika)

Logika popisuje relace (typy relací)

Logika umožňuje:

- ověřování faktů (pravdivosti tvrzení)
- odvozování nových faktů (inference podle pravidel)

Inferenční stroj (Inference Engine)

Ontologie

An ontology is a formal, explicit specification of a shared conceptualization.

Studer et al., 1998