

Organizace znalostí VIII.

PhDr. Jiří Stodola, PhD.

28. 11. 2020

Index

- Lat. – někdo či něco, co ukazuje cestu
- Označení něčeho
- Rejstřík
- Označuje polohu záznamů souboru
-
- Peirce – indexální znak

Index a indexování

- “the process of analyzing the informational content of records of knowledge and expressing the informational content in the language of the indexing system” (Borko, Bernier)
- “[t]he act of describing or identifying a document [4] in terms of its subject content” (ISO 5963:1985)

Složky procesu indexování

- Určit věcný obsah jednotky
- Identifikovat rozličné předměty, jejich aspekty a vzájemné vztahy
- Reprezentovat je prostřednictvím výrazů jazyka předmětového hesláře

Index

- Vede k lokalizaci dokumentů
- Obsahuje přístupové body k textům
- Nejběžněji se používá abecední řazení, ale existují i indexy sestavené jinak
-
- Je buď samostatným dokumentem nebo jeho částí

Index

- Derivace termínů ze zdrojového dokumentu nebo přidělení termínů z již existujícího indexu
- Uspořádání termínů
- Zachycení vztahů mezi termíny
-
-

Podle jejich zdrojových dokumentů

- Podle indexovaného materiálu
 - Knižní, časopisové, databázové, obrázkové, hudební, multimediální, netextové
 -
- Podle přístupových bodů
 - Autorské, názvové, deskriptorové a klíčvoslovové, analytické
 -
- Analytické
- Kompletní
-
- Podle pokrytí zdrojů
 - Kumulativní, selektivní - komplexní
 -
 -

Podle indexů samotných

- Podle způsobů organizace
 - Abecední, systematické

-

Podle užitých symbolů

- Textové, obrazové

-

- Podle způsobu syntaxe

- Prekoordinované, postkoordinované

-

- Podle lokátorů

- Konkrétní (na čísla stran), relativní (na koncepty v klasifikačním systému)

-

-

Podle indexů samotných

- Podle pravděpodobnosti
 - Nepravděpodobnostní – termín přidělen dokumentu
 - Pravděpodobnostní – vyjádřena pravděpodobnost, zda bude relevantní
-
- Podle informací
 - S abstrakty, nahé
 -
- Podle fyzické formy
 - Tištěné, elektronické
-

Podle procesu indexování

- Kdo indexuje
 - Ruční, automatické
 -
- Podle způsobu získání termínů
 - Odvozené neboli extrahované, přiřazené
 -
- Podle zdroje přiřazených termínů
 - Volně tvořené, založené na nadpisech, založené na tezaurech (postkoordinované), založených na systematických klasifikacích předmětů
 -
- Podle teoretických předpokladů
 - Založené na dokumentech, založené na informačních potřebách
 -
 -

Teoretické předpoklady

- Racionalismus – izolovat základní kategorie, z nich poskládat libovolný předmět, logické dělení, apriorismus
- Kognitivní paradigma – vnitřní pravidla
- Empirismus – podobnost, statistika
- Historismus – nutnost znalost domény
- Pragmatismus - účelovost

Jednotky indexování

- Slova – KWIC a KWOC
- Concepty – tezaurus
- Předměty

Jazyk a myšlení

- Překlad z jazyka do jazyka
- Ideální plán – verbální vyjádření – vyjádření prostřednictvím notace
- Konceptuální analýza - překlad

Subjektivita indexátora

- Rekonstruktivní hermeneutika – 4 kánony
-
- Autor jako indexátor
- Indexátor trénovaný v LIS
- Lidové indexování / tagování
- Odborník v určité oblasti
- Odborník v určité oblasti trénovaný v LIS

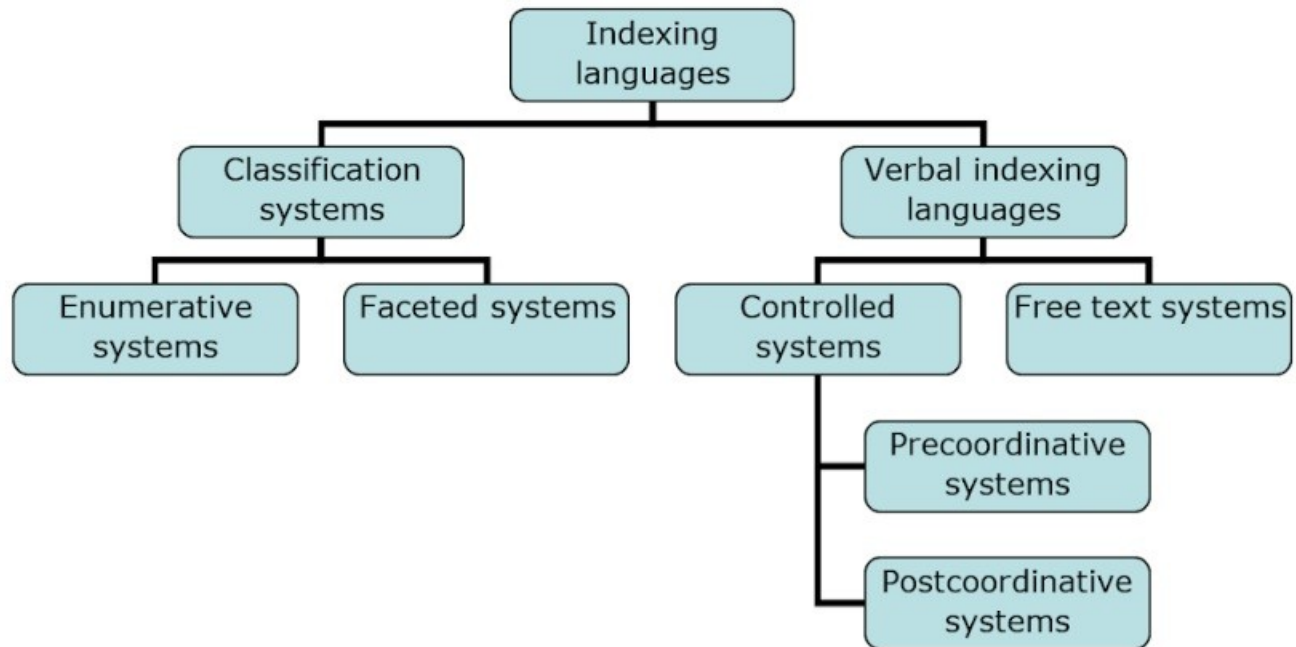
Automatické algoritmické indexování

- Chybí porozumění – Searle a Čínský pokoj
- Etické problémy předsudky

Selekční jazyk

- '*Indexing language*' (*documentary language*) as used in this book = any language (broadly defined) for the representation and/or for the arrangement of retrieval objects and/or their substitutes with the objective of making the items retrievable.

Dělení selekčních jazyků



Vlastnosti selekčních jazyků

- Hloubka – do jaké míry je téma v indexu zastoupeno
- Vyčerpání – počet termínů reprezentujících dokument v indexu
- Specifičnost – shoda mezi indexujícím termínem a indexovaným pojmem
- Granularita – odkazuje na hloubku indexování; jak velké jsou jednotky zahrnuté do indexování
-
-
- Selekční jazyky jsou metajazyky vzhledem k textům a dalším dokumentům, které jsou indexovány