

Univerzita Karlova  
Filozofická fakulta  
Ústav informačních studií a knihovnictví  
obor informační věda

Disertační práce

PhDr. Helena Kučerová

**Ontologie organizace znalostí**  
**Ontology of knowledge organization**

Školitelka PhDr. Eva Bratková, Ph.D.

2018



Prohlašuji, že jsem disertační práci napsala samostatně s využitím pouze uvedených a řádně citovaných pramenů a literatury a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze, dne

## **Abstrakt**

Cílem výzkumu dokumentovaného v disertaci bylo navrhnout ontologii pro doménu organizace znalostí. Byly stanoveny tři požadavky specifikující oblasti užití ontologie: tvorba výroků pro znalostní bázi organizace znalostí, indexace zdrojů v doméně a propojení s věcně relevantními ontologiemi. Návrh byl zpracován metodou reengineeringu první verze ontologie, vytvořené v rámci řešení projektu NAKI Znalostní báze pro obor organizace informací a znalostí v letech 2013–2015. Druhá verze ontologie je založena na základní ontologii DOLCE, konkrétně na opětovném použití jejích vybraných tříd a predikátů včetně s nimi spojených axiomů. Reengineering obsahu první verze ontologie byl proveden metodou doménové a fasetové analýzy. Výsledkem jsou dva propojené moduly – axiomatizovaná ‚těžká‘ jádrová ontologie a ‚lehká‘ terminologická ontologie. Ontologie nabízí řešení kompatibility české a světové oborové terminologie, uplatňuje rozšířené pojetí domény přesahující hranice paměťových institucí a je dílčím příspěvkem k rozvoji teorie oboru organizace znalostí. Obsah práce je členěn do pěti kapitol. V první kapitole je podán přehled současného stavu teorie a praxe informatických ontologií, založený na literárním průzkumu. Druhá kapitola charakterizuje metody, využití k návrhu první i druhé verze ontologie organizace znalostí. Ve třetí kapitole je shrnutí průběhu řešení a hlavních výsledků projektu NAKI. Čtvrtá kapitola dokumentuje postup a výsledky doménové analýzy organizace znalostí. Pátá kapitola obsahuje dokumentaci návrhu druhé verze ontologie organizace znalostí.

## **Abstract**

The aim of the research documented in the thesis was to design an ontology for the knowledge organization domain. Three requirements specifying areas of ontology use have been identified: making of statements for a knowledge base for knowledge organization, indexing of resources in the domain and interlinking with content-relevant ontologies. The design was elaborated by the method of reengineering the first version of ontology, created in the framework of the NAKI research project Knowledge base for subject area of knowledge organization in 2013-2015. The second version of ontology is based on the DOLCE foundational ontology, specifically on the reusing of its selected classes and predicates, including the axioms associated with them. The reengineering of the content of the first version of ontology was done by the domain and facet analysis methodology. The outcome is divided in two interlinked modules – axiomatized heavy-weight core ontology and light-weight terminological ontology. The ontology offers a solution to the compatibility of Czech and world terminology, it applies an extended concept of domain beyond the scope of memory institutions, and it presents a partial contribution to the development of the theory of knowledge organization. The content of the thesis is divided into five chapters. The first chapter presents an overview of the current state of the theory and practice of information ontologies, based on a literary survey. The second chapter describes the methods used to design the first and second versions of the knowledge organization ontology. The third chapter summarizes the course of the solution and the main results of the NAKI project. The fourth chapter documents the progress and outcomes of the domain knowledge analysis. The fifth chapter contains the documentation of the second version of the knowledge organization ontology.

## **Klíčová slova**

organizace znalostí, systém organizace znalostí, ontologie, doménová ontologie, doménová analýza, DOLCE

## **Keywords**

knowledge organization, knowledge organization system, ontology, domain ontology, domain analysis, DOLCE

# Obsah

Úvod.....	8
<b>1 Současný stav teorie a praxe informatických ontologií .....</b>	<b>12</b>
1.1 Historický přehled dosavadního vývoje informatických ontologií .....	14
1.2 Ontologie a znalostní systém.....	17
1.3 Oblasti užití informatických ontologií.....	23
1.4 Typologie informatických ontologií .....	25
1.5 Teoretické a filozofické problémy informatických ontologií.....	29
1.5.1 Materialismus vs. idealismus .....	29
1.5.2 Různé přístupy k chápání/řešení reprezentace .....	33
1.5.3 Různé přístupy k chápání/řešení konceptualizace.....	36
1.5.4 Problém určení typů a vlastností ontologických entit.....	45
1.5.5 Předpoklad otevřeného světa .....	52
1.6 Dílčí závěry a diskuse .....	53
1.6.1 Problém sladění požadavků na jednoduchost a sémantickou sílu reprezentace .....	53
1.6.2 Možnosti vzájemné aplikovatelnosti principů konceptualizace v ‚přirozené‘ a v umělé inteligenci.....	54
1.6.3 Pracovní definice informatické ontologie .....	55
<b>2 Metody použité při návrhu ontologie organizace znalostí .....</b>	<b>58</b>
2.1 Metodiky ontologického inženýrství.....	58
2.1.1 Principy analýzy a metodika <i>OntoClean</i> .....	65
2.1.2 Metodika <i>UPON</i> .....	70
2.1.3 Návrhové vzory .....	73
2.2 Doménová analýza.....	74
2.3 Fasetová analýza .....	80
2.3.1 Fasetová analýza v informační vědě a v organizaci znalostí.....	81
2.3.2 Metodika fasetové analýzy Williama Dentona a Louise Spiteriové.....	83
2.3.3 Metoda <i>5W1H</i> .....	85
2.3.4 Teorie integrativních úrovní .....	85
2.4 Diskuse .....	87
2.4.1 Omezení systémového a analytického přístupu.....	87
2.4.2 Vztah metodik ontologického inženýrství a metod organizace znalostí.....	88
<b>3 První verze ontologie navržená pro prototyp znalostní báze organizace znalostí.....</b>	<b>90</b>
3.1 Cíle, harmonogram a řešitelé projektu.....	90
3.2 Architektura znalostní báze organizace znalostí .....	94
3.3 Ontologie pro znalostní bázi organizace znalostí .....	96
3.4 Implementace znalostní báze.....	99
3.4.1 Pracovní postup pro naplnění znalostní báze instancemi .....	100
3.4.2 Propojení / mapování / linkování ve znalostní bázi.....	104
3.4.3 Technologické řešení implementace znalostní báze .....	108
3.5 Shrnutí výsledků projektu.....	110
3.5.1 Výsledky projektu .....	110
3.5.2 Kritická analýza první verze ontologie.....	111
<b>4 Doménová analýza organizace znalostí.....</b>	<b>112</b>
4.1 Cíl organizace znalostí.....	114
4.2 Hlavní témata organizace znalostí .....	115
4.3 Jednotka organizace znalostí.....	131

4.4	Proces organizace znalostí .....	133
4.5	Realizační kontext procesu organizace znalostí.....	137
4.5.1	Aktéři – osobnosti, institucionální a komunikační základna .....	137
4.5.2	Typologie zdrojů v doméně organizace znalostí.....	140
4.5.3	Systém organizace znalostí .....	144
4.6	Kognitivní kontext organizace znalostí.....	151
4.6.1	Vědní disciplíny a aplikační oblasti .....	151
4.6.2	Terminologie organizace znalostí .....	155
4.7	Model domény organizace znalostí .....	157
4.7.1	Textový model domény organizace znalostí.....	158
4.7.2	Semiformalizovaný model domény organizace znalostí (výroky) .....	160
4.7.3	Vizualizace doménového modelu – pojmová mapa .....	162
<b>5</b>	<b>Návrh druhé verze ontologie organizace znalostí .....</b>	<b>163</b>
5.1	Specifikace požadavků .....	163
5.2	Architektura ontologie organizace znalostí .....	166
5.3	Pojmový model ontologie .....	168
5.3.1	Základní ontologie <i>DOLCE</i> .....	168
5.3.2	Jádrová ontologie organizace znalostí .....	172
5.4	Postup návrhu a implementace .....	176
5.5	Dílní závěry .....	176
5.5.1	Realizace případu užití 1.1 – opětovné použití zaznamenaných doménových znalostí .....	176
5.5.2	Realizace případu užití 1.2 – odvozování a objevování nových znalostí.....	178
5.5.3	Realizace případu užití 2.1 – vyjádření obsahu (indexace) zdrojů.....	178
5.5.4	Realizace případu užití 3.1 – sémantické mapování ontologických tříd a predikátů.....	180
<b>Závěr</b>	.....	<b>182</b>
Příloha 1	Přehled používaných jmenných prostorů .....	186
Příloha 2	Dokumentace výběru základní ontologie pomocí nástroje <i>ONSET</i> .....	188
Příloha 3	Seznam tříd jádrové ontologie organizace znalostí (páteřní taxonomie) .....	194
Příloha 4	Seznam predikátů ontologie organizace znalostí.....	198
Příloha 5	Ukázka implementace ontologie v editoru <i>Protégé</i> .....	204
Příloha 6	Ukázka formalizovaných výroků pro testování případu užití 1.1 .....	207
<b>Literatura</b>	.....	<b>210</b>

## Seznam obrázků

Obr. 1	Informační systém podle ISO/TR 9007 .....	18
Obr. 2	Znalostní systém s ontologií podle ISO/TR 9007 .....	20
Obr. 3	Typologie inforatických ontologií .....	25
Obr. 4	Typy ontologií podle Debory L. McGuinnessové.....	26
Obr. 5	Sémiotický trojúhelník.....	35
Obr. 6	Obecný rámec metodiky <i>UPON</i> .....	71
Obr. 7	Architektura znalostní báze organizace znalostí, založená na principech standardu ISO/TR 9007 .....	94
Obr. 8	Instance v modulu bibliografických entit.....	96
Obr. 9	Páteřní taxonomie (pojmová struktura) ontologie.....	97
Obr. 10	Pracovní postup pro naplnění znalostní báze instancemi .....	100
Obr. 11	Část grafu znázorňujícího formalizované výroky ve znalostní bázi.....	103
Obr. 12	Ukázka vizualizace formalizovaných výroků v prostředí <i>Drupal</i> .....	104
Obr. 13	Typy externího mapování a propojování obsahu znalostní báze .....	106
Obr. 14	Pracovní postup tvorby a zpřístupnění znalostní báze.....	109
Obr. 15	Typologie znalostí z hlediska jejich komunikace a organizace .....	132
Obr. 16	Pojmový model procesu organizace znalostí a procesu přístupu k znalostem.....	135
Obr. 17	Aktéři a kontext organizace znalostí.....	138
Obr. 18	Typologie systémů organizace znalostí .....	146
Obr. 19	Systém organizace znalostí jako model metadat .....	147
Obr. 20	Struktura systému organizace znalostí podle <i>SKOS</i> .....	149
Obr. 21	Pojmová mapa organizace znalostí .....	162
Obr. 22	Schéma propojení modulů ontologie.....	167
Obr. 23	Třídy nejvyšší úrovně převzaté z ontologie <i>DOLCE</i> .....	170
Obr. 24	Interpretace sémiotického trojúhelníku v modulu <i>LMM</i> .....	172
Obr. 25	Návrhový vzor pro popis zdrojů.....	179
Obr. 26	Obecný návrhový vzor pro označení, typování a časové údaje.....	180

## Seznam tabulek

Tab. 1	Terminologie sémiotiky .....	34
Tab. 2	Příklady základních kategorií .....	44
Tab. 3	Terminologie ontologických entit.....	46
Tab. 4	Fasetová typologie metodik ontologického inženýrství .....	60
Tab. 5	Činnosti v rámci pracovních postupů metodiky <i>UPON</i> .....	72
Tab. 6	Činnosti v rámci klíčových fází metodiky Williama Dentona .....	84
Tab. 7	Harmonogram řešení projektu <i>NAKI</i> a přehled jeho plnění.....	92
Tab. 8	Milníky a výstupy projektu <i>NAKI</i> a zodpovědní řešitelé.....	93
Tab. 9	Rozsah znalostní báze .....	95
Tab. 10	Kvantitativní přehled pojmové struktury ontologie .....	98
Tab. 11	Problémy organizace znalostí.....	118
Tab. 12	Kognitivní kontext procesu organizace znalostí.....	151
Tab. 13	Srovnání české a anglické terminologie organizace znalostí.....	157
Tab. 14	Fasetová klasifikace klíčových pojmů organizace znalostí.....	161
Tab. 15	Kvantitativní přehled tříd a predikátů v jádrové ontologii .....	172

## Seznam zkratek

AAA	anyone can say anything about anything
AP	application profile
ASIST	Association for Information Science and Technology
BFO	Basic formal ontology
CASE	computer aided software/system engineering
CIDOC	(ICOM's) International Committee for Documentation
CIDOC CRM	CIDOC Conceptual reference model
CIM	computation independent model
CRG	Classification Research Group
ČSN	česká technická norma
DCAT	Data catalog vocabulary
DCMI	Dublin Core Metadata Initiative
DDC	Dewey decimal classification
DDT	Deweyho desetinné třídění
DLP	DOLCE Lite-Plus
DUL	DOLCE+DnS Ultralite
ER	entity-relationship
FF UK	Filozofická fakulta Univerzity Karlovy
FIS VŠE	Fakulta informatiky a statistiky Vysoké školy ekonomické
FOAF	friend of a friend
FRBR	Functional requirements for bibliographic records
FRBR <sub>ER</sub>	Functional requirements for bibliographic records – entity-relationship
FRBR <sub>oo</sub>	Functional requirements for bibliographic records – object-oriented
FRSAD	Functional requirements for subject authority records
GO	Gene ontology
HTML	hypertext markup language
HTTP	hypertext transfer protocol
ICC	Information coding classification
ICP	International cataloguing principles
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IEKO	ISKO Encyclopedia of knowledge organization
IFLA	International Federation of Library Associations and Institutions
IFLA LRM	IFLA Library reference model
INTERPI	Interoperabilita v paměťových institucích
IOLite	Information objects ontology
ISKO	International Society for Knowledge Organization
ISNI	International standard name identifier
ISO	International Organization for Standardization
KO	knowledge organization
KOS	knowledge organization system
KOSO	Knowledge organization systems ontology
LMM	Lexical Meta-Model
MDA	model driven architecture
MDT	Mezinárodní desetinné třídění
NAKI	Program aplikovaného výzkumu a vývoje národní a kulturní identity
NKOS	Networked Knowledge Organization Systems/Services/Structures
NUNA	non-unique naming assumption
OCR	optical character recognition
ODPA	Ontology Design Patterns Association
OMG	Object Management Group



OMV	Ontology metadata vocabulary
OWA	open world assumption
OWL	Web ontology language
PIM	platform independent model
PMEST	personality, matter, energy, space, time
PSH	Polytematický strukturovaný heslář
PSM	platform specific model
RDA	Resource description and access
RDF	Resource description framework
RDFS	RDF schema
RUP	Rational unified process
SEO	Software enterprise ontology
SIG/CR	Special Interest Group for Classification Research
SKOS	Simple knowledge organization system
SOZ	system organizace znalostí
SPARQL	SPARQL protocol and RDF query language
SQL	Structured query language
TC	technical committee
TDKIV	Terminologická databáze knihovnictví a informační vědy
TR	technical report
UDC	Universal decimal classification
UDCC	UDC Consortium
UFO	Unified foundational ontology
UML	Unified modeling language
UNESCO	United Nations Educational, Scientific and Cultural Organization
UP	Unified process
UPON	Unified process for ontology
URI	Uniform resource identifier
URL	Uniform resource locator
VIAF	Virtual international authority file
W3C	World Wide Web Consortium
WWW	world wide web

## Typografické, jmenné a grafické konvence uplatněné v textu práce

V textu disertace jsou uplatněny následující typografické, jmenné a grafické konvence:

Doslovné citace jsou uváděny kurzívou a uzavřeny v uvozovkách. Cizojazyčné termíny jsou psány kurzívou (např. *aboutness*). Kurzívou jsou též psány názvy děl zmiňovaných v textu. Části textu uzavřené v apostrofech obsahují zdůrazňované pasáže nebo parafráze. Do apostrofů jsou rovněž uzavřeny texty, v nichž jsou slova používána v jiném než obvyklém smyslu (např. nezačíná ‚od nuly‘, kniha vs. označení ‚kniha‘).

Názvy děl, institucí a produktů jsou uváděny kurzívou. Názvy kategorií a tříd jsou v souladu s notací jazyka *UML* uváděny v jednotném čísle a s velkými počátečními písmeny, názvy atributů a vztahů s malými počátečními písmeny. Názvy vztahů jsou uzavřeny v špičatých závorkách.

Vzhledem k panující nejednotnosti terminologie organizace znalostí jsou používané termíny často doprovázeny v závorkách alternativními způsoby jejich vyjádření a cizojazyčnými ekvivalenty – např. ekvivalence (stejnost), jmenný prostor (angl. *namespace*).

Pojmové a procesní diagramy na obrázcích 1, 2, 3, 7, 10, 13, 14, 16, 23, 25, 26 jsou vyjádřeny v jazyce *UML*. Jejich grafická notace je založena na standardu *OMG Unified modeling language (OMG UML)*.<sup>1</sup> Používané grafické symboly jsou uvedeny v tabulce na následující straně. Třída, znázorněná v diagramu obdélníkem, zastupuje vždy množinu objektů se stejnými vlastnostmi a chováním. Obdélník je rozdělen na tři části: v první části je název třídy, v druhé části jsou znázorněny statické vlastnosti třídy neboli atributy a třetí část obsahuje seznam dynamických vlastností třídy neboli operací. Vztah asociace tříd je vyjádřen notací, znázorněnou na řádku 2 tabulky. Na řádku 3 je notace vztahu partitivní hierarchie. Šipka zakončená kosočtvercem míří od třídy – části k třídě – celku. Vztah generické hierarchie je vyjádřen symbolem generalizace, který je znázorněn na řádku 4. Šipka zakončená prázdným trojúhelníkem směřuje od dětské třídy k třídě rodičovské. Pro stavy i aktivity se používají ikony obdélníků se zaoblenými rohy, znázorněné v řádku 5 spolu s ikonami z Erikssonova-Penkerova aplikačního profilu. Vztahy posloupnosti aktivit a stavů se znázorňují orientovanými hranami. Notace vztahů posloupnosti je uvedena na řádku 6. Procesní diagramy kromě orientovaných hran též disponují symboly, jež umožňují znázornit alternativní možnosti průběhu modelovaných procesů (větvení) a jejich synchronizace (notace viz řádek 7 tabulky).

---

<sup>1</sup> Object Management Group. *OMG Unified modeling language (OMG UML)* [online]. OMG, 1997- [cit. 2018-01-30]. Dostupné z: <http://www.omg.org/spec/UML/>.

1	<table border="1"> <tr><td>Název třídy</td></tr> <tr><td>atribut</td></tr> <tr><td>operace ()</td></tr> </table>	Název třídy	atribut	operace ()
Název třídy				
atribut				
operace ()				
2				
3				
4				
5				
6				
7				

## Úvod

*Poznání věcí je dokonalé, pokud je úplné, správné a uspořádané.*<sup>2</sup>  
Jan Amos Komenský, 1637

Cílem výzkumu dokumentovaného v této práci bylo specifikovat pojmovou bázi domény organizace znalostí prostřednictvím návrhu doménové ontologie. Ontologii chápeme jako typ pojmového modelu, jehož účelem je kromě komunikace, opakovaného využití a organizace znalostí i popis skutečnosti pro počítačové zpracování a automatické odvozování znalostí ve vymezené oblasti, jíž je v našem případě organizace znalostí.

Termín organizace znalostí označuje proces, jehož historie je stejně stará jako historie lidského poznávání světa. Lze se s ním pod nejrůznějšími názvy setkat všude, kde se pracuje s informacemi a znalostmi. Nejvýznamnějšími oblastmi, jež se organizaci znalostí věnovaly v minulosti a jež dodnes významně ovlivňují její metody a techniky, jsou encyklopedistika, klasifikace věd a vědecké klasifikace, techniky duševní práce, knihovnické a bibliografické pořádací systémy, nejnověji též technologie zpracování a komunikace dat v počítačích a v počítačových sítích. Vzhledem k různorodosti názorů na vymezení této domény považujeme za vhodné vymezit vlastní pracovní pojetí organizace znalostí, které bude uplatněno v rámci této práce. Organizaci znalostí chápeme jako proces strukturování zaznamenaných znalostí, jenž k nim umožňuje přístup a jejich využití. Kontext tvoří informační zdroje, používané metody a nástroje a aktéři.

Jak proces organizace znalostí, tak jeho realizační kontext jsou součástí komunikace informací ve společnosti a jako takové patří mezi předměty zkoumané informační vědou. Zhruba od 70. let 20. století jsou navíc patrné snahy a aktivity směřující k ustavení samostatné vědní disciplíny zaměřené na zkoumání efektivních způsobů řešení problémů spojených s procesem organizace znalostí a s jeho kontextem. Na rozdíl od informační vědy, jež je považována za interdisciplinární teoretickou vědu, má obor organizace znalostí charakter inženýrské (aplikované) disciplíny. Co do předmětu zájmu, resp. aplikační oblasti, má organizace znalostí s informační vědou významný průnik, kterým je proces organizace znalostí. Liší se ovšem v používaných metodách: informační věda používá metody teoretického výzkumu, v organizaci znalostí se používají inženýrské postupy. Při řešení problémů organizace znalostí se v oboru organizace znalostí aplikují systematicky poznatky jiných vědních disciplín, zejména teorie poznání, logiky, lingvistiky, informační vědy, kognitivní vědy, systémové vědy, umělé inteligence. Formující se teoretickou platformu tvoří dále empirické principy a zobecněné příklady nejlepší praxe z oblastí paměťových a fondových institucí, podnikové informatiky, informatiky veřejné správy a vědecké informatiky.

Změna společenského kontextu v podobě nástupu znalostní ekonomiky, globalizace a multikulturalismu a především masové rozšíření nových informačních technologií (síťový model komunikace, elektronické dokumenty a propojená data) způsobily, že organizace znalostí se momentálně nachází ve fázi paradigmatu v krizi, resp. ve fázi hledání nového paradigmatu. Teoretická, pojmová a terminologická základna oboru, ustavená v 19.–20. století, se redefinuje a není jednotná. Nelze najít autoritativní zdroj, který by byl přijímán všemi odborníky, již se dané oblasti věnují. Tento fakt byl při volbě tématu vyhodnocen jako příležitost, jež umožní v rámci návrhu ontologie organizace znalostí zformulovat a nabídnout odborné veřejnosti nové originální poznatky a metodické a teoretické podněty pro konstrukci nového paradigmatu.

---

<sup>2</sup> KOMENSKÝ, Jan Amos. *Předehra pansofie: objasnění pansofických pokusů*. Praha: Academia, 2010, s. 79.

Záměrem výzkumu bylo reagovat na paradigmatické změny v doméně a v metodické rovině navázat na vědecký přístup Blahoslava Kováře, jenž se zasloužil o formování oboru na moderních základech v podmínkách tehdejšího Československa v 60.–80. letech 20. století.<sup>3</sup> Jeho vědecký přístup se snažíme aplikovat na současnou vývojovou etapu komunikace a organizace znalostí 21. století a zároveň rozšířit zkoumání problematiky za hranice vymezené působností tradičních paměťových institucí.

K výběru tématu mě vedl dlouholetý zájem o otázky pořádání informací a znalostí. Svůj zájem jsem měla možnost uplatnit v průběhu výuky předmětů *Organizace informací a znalostí* na *Vyšší odborné škole informačních služeb*, předmětů *Věcné pořádání informací a znalostí* a *Pořádání a vyhledávání informací* na *Ústavu informačních studií a knihovnictví FF UK* a předmětů *Organizace znalostí* a *Informační architektura* na *Katedře informačních a znalostních technologií FIS VŠE*. Významným impulsem a zdrojem praktických zkušeností s návrhem ontologie se pro mě stala účast v řešitelském týmu programového projektu *NAKI Znalostní báze pro obor organizace informací a znalostí*, jehož řešitelkou byla moje školitelka Eva Bratková. Ráda bych také touto prací vyjádřila vděčnost svému učiteli a vedoucímu diplomové práce Blahoslavu Kovárovi, jehož přednášky na tehdejší katedře knihovnictví a vědeckých informací *FF UK* ve mně vzbudily zájem o problematiku organizace znalostí a její hlubší studium.

Kromě osobního zájmu mě k volbě tématu podnítil i objektivně zjištěný stav oboru v tuzemském kontextu. V období 2010–2012 jsem v souvislosti s rozpracováním schváleného tématu disertace provedla orientační srovnávací průzkum stavu poznatkové základny oboru organizace znalostí v České republice a ve světě, zaměřený na následující oblasti: technologie, akreditované studijní obory, oborová terminologie a informační zdroje (vědecké monografie, učebnice, periodika, standardy). Na tento průzkum navázaly v prvním roce řešení projektu *NAKI* v roce 2013 rozhovory s doménovými experty a zpracování rozsáhlé bibliografické rešerše, jež plně potvrdily závěry průzkumu z roku 2012. Výsledkem průzkumu byla následující zjištění: Srovnatelné se světovým vývojem jsou v českém prostředí používaná technologická základna a akreditované vysokoškolské studijní obory a předměty. Problémy byly zjištěny v oblasti oborové terminologie a informačních zdrojů. Jako nejpalčivější se jeví nekompatibilita odborné terminologie, která brání mezinárodní komunikaci, a dále chybějící aktuální vědecké monografie a učebnice. Tato absence aktuálních publikací v mateřském jazyce je pocíťována českou odbornou veřejností jako citelný nedostatek. Potvrzují ji i zkušenosti z výuky výše uvedených předmětů, jejichž studenti často poukazovali na chybějící aktuální učebnice.

Průzkum potvrdil, že posledním zdrojem pokrývajícím komplexně problematiku organizace znalostí publikovaným v češtině jsou vědecké práce a učebnice Blahoslava Kováře z 80. let 20. století. Předností Kovářových textů jsou vědecké základy, na nichž jsou postaveny. Jeho zásluhou je rovněž ustavení původní české terminologie pro oblast věcného pořádání, dodnes používané v české knihovnické a informační praxi. Přes svoji kvalitu však už tyto publikace nejsou využitelné při studiu oboru, poněvadž nezachycují významné změny, k nimž došlo od 90. let 20. století v souvislosti se změnou společenského a technologického paradigmatu. Bibliografická rešerše dále prokázala, že přes množství vydávaných titulů jsou současné monografie vydávané v zahraničí vesměs pragmatičtěji orientovány. Zachycují sice aktuální stav vývoje disciplíny, chybí jim však až na nečetné výjimky hlubší teoretický základ. Navíc se v aplikační rovině

---

<sup>3</sup> KOVÁŘ, Blahoslav. *Problémy teorie procesu věcného pořádání informací a selekčních jazyků*. Praha: Univerzita Karlova, 1976. 165 [1] s. Acta Universitatis Carolinae. Philosophica et historica. Monographia, sv. 65/1976, ISSN 0567-8307.

zpravidla omezují na tradiční pojetí organizace znalostí v paměťových institucích (katalogizace, bibliografie, indexace dokumentů). Z tohoto důvodu a vzhledem k závislosti organizace znalostí na přirozeném jazyce a na specifickém kulturním kontextu nebyla varianta překladu zahraniční vědecké monografie do češtiny vyhodnocena jako optimální.

Již po ukončení orientačního průzkumu bylo proto po dohodě se školitelkou rozhodnuto rozšířit rozsah výzkumu a kromě ontologie vytvořit i znalostní bázi organizace znalostí. Zatímco pojmový model ve formě ontologie má přispět k řešení teoretických a terminologických problémů, znalostní báze má coby množina zaznamenaných a uspořádaných poznatků poskytnout zdrojový materiál pro publikační činnost a zejména pro zpracování původní vědecké monografie. Návrh a implementace prototypu znalostní báze byly realizovány za podpory *Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI)*. Tato ‚vedlejší větev‘ výzkumu způsobila prodloužení standardní doby doktorského studia, na druhou stranu přinesla mimo jiné zdrojový materiál pro testování úplnosti a vyjadřovací přesnosti (expresivity) ontologie.

Součástí úvodního průzkumu bylo i posouzení výsledků aktuálních výzkumných aktivit souvisejících s výzkumným záměrem – projektem *KOSO (Knowledge Organization Systems Ontology)*<sup>4</sup>, projektem *OMV (Ontology Metadata Vocabulary)*<sup>5</sup> a s českým projektem *INTERPI (Interoperabilita v paměťových institucích)*<sup>6</sup>. Bylo zjištěno, že každý z nich se týká problematiky stanovené výzkumným záměrem pouze okrajově. Ani jeden z uvedených výzkumných projektů neřeší problematiku v její úplnosti: *KOSO* zužuje obsah ontologie na systémy organizace znalostí, *OMV* se zaměřuje pouze na definování popisných metadat pro anotování ontologií, *INTERPI* se omezuje na organizaci znalostí v paměťových institucích a jeho znalostní báze má univerzální zaměření a postrádá doménově specifické detaily.

Východiskem pro formulaci dílčích cílů disertační práce je fakt, že inženýrské ontologie jsou produkty inženýrského návrhu a jsou vždy konstruovány pro určité případy použití a za účelem splnění konkrétních cílů. Stanovené dílčí cíle mají proto formu tří klíčových požadavků na ontologii organizace znalostí, specifikujících oblasti jejího užití: 1) Ontologie bude využitelná pro tvorbu výroků pro znalostní bázi organizace znalostí, 2) Ontologie bude využitelná pro indexaci zdrojů v doméně organizace znalostí a 3) Ontologie bude propojena s dalšími věcně relevantními ontologiemi. Každý z těchto požadavků je dále detailně konkretizován prostřednictvím případů použití.

Přínosy ontologie organizace znalostí jsou předpokládány ve třech oblastech. 1) Ontologie nabídne řešení kompatibility české a světové oborové terminologie. 2) Rozšířením pohledu na obor za hranice paměťových institucí umožní transdisciplinární komunikaci a výzkum. 3) Ontologie se stane dílčím příspěvkem k vymezení hranic a k rozvoji teorie oboru organizace znalostí a informační vědy.

---

<sup>4</sup> WELLER, Katrin. *KOSO: A reference ontology for reuse of existing knowledge organization systems* [online preprint]. [2008] [cit. 2018-01-30]. S. 31-39. Dostupné z: <http://www.phil-fak.uni-duesseldorf.de/fr/015-koso-a-reference-ontology-for-reuse-of-existing-knowledge-organization-systems/> nebo z [http://www.phil-fak.uni-duesseldorf.de/fileadmin/Redaktion/Institute/Informationswissenschaft/1215088990015\\_weller.pdf](http://www.phil-fak.uni-duesseldorf.de/fileadmin/Redaktion/Institute/Informationswissenschaft/1215088990015_weller.pdf).

<sup>5</sup> PALMA, Raúl, HARTMANN, Jens, HAASE, Peter. *OMV: Ontology Metadata Vocabulary for the Semantic web* [online]. Version 2.4.1. OMV Consortium, March 2009 [cit. 2018-01-30]. 76 s. OMV Report. Dostupné z: <http://sourceforge.net/projects/omv2/files/OMV%20Documentation/OMV-Reportv2.4.1.pdf>.

<sup>6</sup> Dostupné z: <http://www.interpi.cz/> [cit. 2018-01-30].

Návrh ontologie probíhal ve dvou etapách. První etapa proběhla v letech 2013–2015 v rámci řešení grantového projektu *NAKI* a jejím výsledkem byla první verze ontologie. Návrh byl realizován řešitelským týmem grantového projektu *NAKI* za použití strukturované metodiky *UPON*. V roce 2014 byla první verze ontologie implementována do funkčního prototypu znalostní báze organizace znalostí. Druhou etapu v letech 2016–2017 jsem realizovala individuálně a v jejím průběhu byla navržena druhá verze ontologie s uplatněním metody reengineeringu první verze a opětovného použití základní ontologie *DOLCE*. Jádrem obsahového návrhu ontologie byla metoda doménové a fasetové analýzy.

Obsah práce je členěn do pěti kapitol. První dvě kapitoly mají přehledový charakter, následující tři kapitoly dokumentují provedený výzkum. V první kapitole je podán přehled současného stavu teorie a praxe informatických ontologií, založený na literárním průzkumu. Kapitola je zároveň shrnutím teoretických problémů, jež byly řešeny v průběhu návrhu ontologie. V závěru kapitoly je zformulována pracovní definice informatické ontologie, používaná pro účely této práce.

Druhá kapitola charakterizuje metody, využití k návrhu první i druhé verze ontologie organizace znalostí. Zaměřuje se především na použité metodiky ontologického inženýrství a na metody doménové a fasetové analýzy.

Ve třetí kapitole je shrnutí průběhu řešení a hlavních výsledků grantového projektu *NAKI Znalostní báze pro obor organizace informací a znalostí*. Její součástí je popis a zhodnocení první verze ontologie, jež je implementována v prototypu znalostní báze.

Čtvrtá kapitola dokumentuje postup a výsledky doménové analýzy organizace znalostí. Doménová analýza se zaměřila na stanovení cíle organizace znalostí, na specifikaci hlavních témat ve formě problémů a principů jejich řešení, na definování jednotky organizace znalostí a na analýzu procesu organizace znalostí a jeho realizačního a kognitivního kontextu. Výsledkem je model domény, který se stal zdrojem pro návrh druhé verze ontologie.

Pátá kapitola obsahuje dokumentaci návrhu druhé verze ontologie organizace znalostí. Zahrnuje i popis základní ontologie *DOLCE*, jež byla využita při návrhu a jejíž části jsou do ontologie organizace znalostí integrovány. Splnění požadavků na druhou verzi ontologie je dokumentováno prostřednictvím realizace jednotlivých případů užití.

Dílní poznatky a výsledky výzkumu předkládané v této práci byly již prezentovány na konferencích *ISKO 2014*, *Znalosti 2014* a v odborných časopisech *Ikaros*, *ProInflow* a *Knihovna*. Konkrétní reference jsou uvedeny na příslušných místech textu formou bibliografických odkazů. Výsledky doménové analýzy organizace znalostí byly podrobně rozpracovány v textu vědecké monografie *Organizace znalostí: klíčová témata*<sup>7</sup>.

---

<sup>7</sup> KUČEROVÁ, Helena. *Organizace znalostí: klíčová témata*. 1. vyd. Praha: Karolinum, 2017. 269 s. ISBN 978-80-246-3587-3 (brož). ISBN 978-80-246-3597-2 (pdf).

# Kapitola 1 Současný stav teorie a praxe informatických ontologií

*Je-li řešení založené na ontologiích odpovědí, jak zní otázka?*<sup>8</sup>  
Mike Uschold, 2005

Slovo ontologie pochází z řečtiny a znamená ‚přemýšlení o bytí‘, což je tradiční náplní filozofie. I když samotný termín ontologie se ve filozofii používá pro označení zkoumání otázek bytí až od 17. století<sup>9</sup>, zkoumání toho, co je (řeceno slovy Martina Heideggera „*jsoucí jakožto jsoucí*“<sup>10</sup>), bylo pod nejrůznějšími názvy (metafyzika, první filosofie) součástí filozofie od jejích počátků. Za zakladatele ontologického bádání je považován Parmenides z Eleje (510–450 př. n. l.), který jako první v evropské filozofické tradici zpochybnil smyslový a empirický způsob zkoumání a byl zastáncem racionálního a logického poznávání světa.<sup>11</sup>

Termín ontologie byl proto donedávna spojován především s filozofickou disciplínou zkoumající obecné principy bytí a s jejími výsledky. V poslední době se čím dál častěji objevuje i ve slovníku informatiků a informačních pracovníků včetně knihovníků, zejména v souvislosti s nástupem sémantického webu. Přes společné východisko, spočívající v zájmu o poznání skutečnosti, a obdobnou formu vyjádření v podobě pojmového / kategoriálního systému nabyt pojem ontologie v informatice svůj specifický význam. Specifika informatických ontologií<sup>12</sup> oproti ontologiím filozofickým vyplývají z odlišně stanoveného cíle. Filozofické ontologie jsou vytvářeny pro ‚čisté‘ poznání, informatické ontologie mají vždy specifický účel použití. Liší se i používané metody. Filozofické ontologie jsou výsledkem teoretického zkoumání v podobě abstraktních myšlenkových systémů, informatické ontologie jsou výsledkem inženýrského návrhu v podobě slovníků. Neusilují o řešení základních filozofických (ontologických) otázek, v tom jsou vždy závislé na některém filozofickém směru, tj. na některé z filozofických ontologií. John F. Sowa označuje filozofické ontologie (jež nazývá „vědecké ontologie“) za deskriptivní, informatické ontologie (jež nazývá „aplikované“ nebo „inženýrské“) mohou podle něj být deskriptivní i preskriptivní.<sup>13</sup>

Informatické ontologie jsou tedy inženýrské artefakty, které jsou navrhovány a implementovány se záměrem dosáhnout předem stanoveného cíle v nějaké konkrétní oblasti použití. Michael Grúninger a Jintae Lee zobecnili specifické oblasti využití ontologií do tří základních okruhů:

---

<sup>8</sup> USCHOLD, Michael. An ontology research pipeline. In: *Applied ontology*. 2005, **1**(1), s. 13. ISSN 1570-5838 (Print). ISSN 1875-8533 (Online).

<sup>9</sup> Viz např. hesla ‚bytí‘, ‚existence‘, ‚jsoucnost‘, ‚metafyzika‘, ‚ontický, ontologie, ontologický‘ v DUROZOL, Gérard, ROUSSEL, André. *Filozofický slovník*. Přeložili Jan Binder aj. 1. vyd. Praha: EWA, 1994, 352 s. ISBN 80-85764-07-5.

<sup>10</sup> HEIDEGGER, Martin. *Co je metafyzika?* Praha: Oikúmené 1993, s. 7.

<sup>11</sup> Z Parmenidova díla se ovšem zachovaly jen zlomky, jež se staly předmětem často velmi rozdílných interpretací.

<sup>12</sup> V některých anglicky psaných textech se používá pro odlišení typů ontologií místo přídavných jmen ‚filozofická‘ a ‚informatická‘ rozdílné psaní – filozofická disciplína se píše s počátečním velkým písmenem a je nepočitatelná, informatické ontologie jsou označeny počitatelným podstatným jménem s počátečním malým písmenem. Tento způsob odlišení používá např. Nicola Guarino v GUARINO, Nicola, OBERLE, Daniel, STAAB, Steffen. What is an ontology? In: STAAB, Steffen, STUDER, Rudi, ed. *Handbook on ontologies*. 2nd ed. Dordrecht: Springer, 2009, s. 1-17. doi:10.1007/978-3-540-92673-3. ISBN 978-3-540-70999-2 (Print). ISBN 978-3-540-92673-3 (Online).

<sup>13</sup> SOWA, John F. Signs and reality. In: *Applied ontology*. 2015, **10**(3-4), 273-284. doi:10.3233/AO-150159. ISSN 1570-5838 (Print). ISSN 1875-8533 (Online).



komunikace, počítačové odvozování, opakované využití a organizace znalostí<sup>14</sup> (blíže viz kapitola 1.3).

V dnes již značně rozsáhlé literatuře je možné se setkat s četnými pokusy o vytvoření definice informatické ontologie. Za reprezentativní lze považovat definici publikovanou v dokumentu specifikace požadavků na ontologický jazyk *OWL*, na jehož vytvoření se podílel Leo Obrst<sup>15</sup>. Definice vyjadřuje přístup konsorcia *W3C* k chápání obsahu pojmu ontologie:

*„Ontologie definuje termíny používané k popisu a reprezentaci oblasti znalostí. Ontologie používají lidé, databáze a aplikace, potřebující sdílet informace o určité doméně ... Ontologie zahrnují počítačem použitelné definice základních pojmů v doméně a vztahy mezi nimi ... Kódují znalosti v doméně a také znalosti o těchto doménách. Tím umožňují, aby tyto poznatky byly opětovně použitelné.“<sup>16</sup>*

Shrnutí pohledů na informatické ontologie z pozic oborů umělé inteligence a znalostního inženýrství nabízí monografie *Ontological engineering* (Ontologické inženýrství)<sup>17</sup>. V kapitole 1.2 autoři uvádějí 16 různých definic ontologie, přičemž konstatují, že jde pouze o výběr z velkého množství dalších definic, jež je možné najít v literatuře. V závěru vyvozují vlastní definici, v níž konstatují, že cílem ontologií je „zachytit konsenzuální znalosti obecným způsobem, aby mohly být znovu použity a sdíleny mezi softwarovými aplikacemi a skupinami lidí“<sup>18</sup>.

Obecně přijímaná a často diskutovaná definice informatické ontologie Thomase Grubera lapidárně vystihuje její gnozeologický a sémiotický základ: „*Ontologie je explicitní specifikace konceptualizace.*“<sup>19</sup> Detailnějšímu rozboru obsahu této definice se budeme věnovat v kapitole 1.6.

V této kapitole je podán přehled současného stavu problematiky informatických ontologií, založený na literárním průzkumu. Kapitola je členěna na šest částí. Část 1.1 obsahuje historický přehled dosavadního vývoje informatických ontologií. V části 1.2 je podán souhrn problematiky informatických ontologií tak, jak jsou chápány v oboru umělé inteligence a v znalostním inženýrství; je specifikován jejich vztah ke znalostním bázím a znalostním systémům. V části 1.3 je uveden přehled oblastí užití informatických ontologií a část 1.4 obsahuje jejich typologii. V části 1.5 jsou charakterizovány teoretické a filozofické problémy informatických ontologií, jež jsou zároveň shrnutím teoretických problémů, které byly řešeny v průběhu návrhu ontologie organizace znalostí. Část 1.6 shrnuje dílčí závěry této kapitoly. Jeden z dílčích závěrů tvoří

<sup>14</sup> GRÜNINGER, Michael, LEE, Jintae. Ontology applications and design. In: *Communications of the ACM*. February 2002, **45**(2), s. 40. doi:10.1145/503124.5031. ISSN 0001-0782 (Print). ISSN 1557-7317 (Online).

<sup>15</sup> Viz OBRST, Leo. Ontologies for semantically interoperable systems. In: *CIKM '03: Proceedings of the twelfth international conference on information and knowledge management, New Orleans, LA, USA - November 03-08, 2003*. New York: ACM, 2003, s. 366. doi:10.1145/956863.956932. ISBN 1-58113-723-0.

<sup>16</sup> HEFLIN, Jeff, ed. *OWL Web Ontology Language use cases and requirements* [online]. W3C Recommendation 10 February 2004. Cambridge (MA): World-Wide Web Consortium, © 2004 [cit. 2018-01-30]. Dostupné z: <https://www.w3.org/TR/webont-req/#onto-def>.

<sup>17</sup> GÓMEZ-PÉREZ, Asunción, FERNÁNDEZ-LÓPEZ, Mariano, CORCHO, Oscar. Theoretical foundations of ontologies. In: *Ontological engineering: with examples from the areas of knowledge management, e-commerce and the semantic web*. 1st ed. London: Springer, 2004, s. 1-45. doi:10.1007/b97353. ISBN 978-1-85233-551-9 (Print). ISBN 978-1-85233-840-4 (Online).

<sup>18</sup> Tamtéž, s. 8-9.

<sup>19</sup> GRUBER, Thomas R. A translation approach to portable ontology specification. In: *Knowledge acquisition*. June 1993, **5**(2), 199-220. doi:10.1006/knac.1993.1008. ISSN 1042-8143.

formulace vlastní pracovní definice informatické ontologie, jež bude používána v dalším textu práce.

## 1.1 Historický přehled dosavadního vývoje informatických ontologií

Vzhledem k bohaté existující literatuře na toto téma se omezíme pouze na stručné shrnutí. Podrobnější přehled historie informatických ontologií lze najít například v příručce *Handbook on ontologies*<sup>20</sup> nebo v přehledovém textu Zdeňka Zdráhal<sup>21</sup>.

Na úvod je vhodné zdůraznit, že přes poměrně krátkou historii má ontologické inženýrství k dispozici rozsáhlou poznatkovou bázi tvořenou výsledky tisíciletého filozofického zkoumání v podobě filozofických ontologií. Navíc mohou tvůrci informatických ontologií čerpat i z neméně dlouhé tradice organizace lidského poznání především v knihovnách, ať už byly příslušné činnosti nazývány třídění, klasifikace, indexace, katalogizace nebo pořádání. Tyto poznatky zobecňuje a teoreticky zkoumá informační věda. Následně jsou pak uplatněny v aplikační oblasti návrhu a implementace knihovních a informačních systémů, jejichž integrální součástí jsou i ‚knihovnické‘ ontologie, označované jako systémy organizace znalostí (v rámci typologie představené v kapitole 1.4 by bylo možné je zařadit do skupiny terminologických ontologií). Ty zůstávají zatím s výjimkou medicíny poněkud stranou pozornosti odborníků z oblasti informatiky. Jejich předností nebývá propracovaný systém vztahů a axiomů, ale bohatost lexikálního obsahu, který zase mnohdy chybí ontologiím vzniklým v laboratorních podmínkách. Na rozdíl od informatických ontologií, jež kladou důraz na opětovnou použitelnost znalostí, jsou ‚knihovnické‘ ontologie zaměřeny na umožnění přístupu ke znalostem. K této linii tradice se hlásí i konsorcium *OMG*, jež v rámci standardu pro ontologický metamodel nabízí následující vymezení rozsahu pojmu informatické ontologie:

*„Ontologie může co do expresivity sahat od taxonomie (znalosti strukturované hierarchicky nebo genericky stylem předek – potomek) k tezaurům (slova a synonyma), pojmovým modelům (se složitějším zachycením struktury znalostí) až k logické teorii (s velmi bohatými, složitými, konzistentními a smysluplnými znalostmi).“*<sup>22</sup>

Knihovníkům a informačním pracovníkům může takový výčet znít velmi povědomě a potvrdit jim, že informatické ontologie nepřicházejí odkudsi z neznáma, ale naopak že představují aktuální etapu v rámci kontinuálního historického vývoje nástrojů reprezentace a organizace znalostí, tj. indexů, katalogů, klasifikačních schémat, taxonomií, číselníků, souborů autorit, heslářů ad.

Lze tedy konstatovat, že při tvorbě informatických ontologií se po letech relativně izolovaného vývoje naskytá možnost synergie nahromaděných zkušeností filozofů, knihovníků a informatiků. Zkušenosti filozofů a knihovníků jsou nesporně rozsáhlejší a dlouhodobější. Informatici své

---

<sup>20</sup> STAAB, Steffen, STUDER, Rudi, ed. *Handbook on ontologies*. 2nd ed. Dordrecht: Springer, 2009. xix, 811 s. International Handbooks on Information Systems. doi:10.1007/978-3-540-92673-3. ISBN 978-3-540-70999-2 (Print). ISBN 978-3-540-92673-3 (Online).

<sup>21</sup> ZDRÁHAL, Zdeněk. Ontologie: od filosofie k umělé inteligenci. In: Vladimír MAŘÍK, Olga ŠTĚPÁNKOVÁ, Jiří LAŽANSKÝ a kol. *Umělá inteligence*. 6. 1. vyd. Praha: Academia, 2013, kap. 1, s. 21-84. ISBN 978-80-200-2276-9.

<sup>22</sup> Object Management Group. *Ontology definition metamodel (ODM)* [online]. Version 1.1. Needham (MA, USA): Object Management Group, September 2014 [cit. 2018-01-30], s. 31. OMG Document Number: formal/2014-09-02. Dostupné z: <http://www.omg.org/spec/ODM/>.

zkušenosti shromažďují sice nepoměrně kratší dobu cca od 40. let 20. století, nicméně s tou výhodou, že se ve své praxi mohou opřít o principy matematiky, logiky a systémové vědy a o nejnovější informační technologie.

Thomas Gruber (1959) v článku *Ontologie v Encyklopedii databázových systémů*<sup>23</sup> člení krátkou historii informatických ontologií na desetiletí. ‚Zakladatelská‘ osmdesátá léta 20. století byla podle něj v režii oboru umělá inteligence. Umělá inteligence chápala ontologii jednak jako teorii modelovaného světa, tj. svého druhu aplikovanou filozofii, a jednak jako komponentu znalostních systémů, umožňující formální reprezentaci znalostí. Devadesátá léta 20. století přidala chápání ontologie jako nástroje sémantické interoperability a přinesla již výše zmíněnou definici ontologie jako explicitní specifikace konceptualizace. Tato definice odrážela přesun pozornosti směrem k nástrojům vyjádření pojmového obsahu ontologie, tj. k nejrůznějším typům slovníků, jež byly počátkem 21. století doplněny návrhem specifických jazyků (*RDF*, *RDFS*, *OWL*). První desetiletí 21. století je pak jednoznačně ve znamení nástupu sémantického webu, přičemž sémantiku mají zajišťovat právě ontologie. Příznačně to vystihuje název přehledového příspěvku *Ontologie a WWW*, který přednesl Vojtěch Svátek na konferenci *Datakon 2002*.<sup>24</sup>

Jak bylo uvedeno, v 80. letech 20. století začali termín ontologie systematicky používat vědci z oboru umělá inteligence. V roce 1980 použil termín ontologie jeden ze zakladatelů oboru John McCarthy (1927–2011) jako název pro formalizaci kontextu, v němž probíhá nemonotónní odvozování.<sup>25</sup> Významným příspěvkem ke koncepci ontologií se stal obor reprezentace znalostí. Další z průkopníků umělé inteligence Allen Newell (1927–1992) definuje reprezentaci znalostí jako proces, který zajišťuje, aby znalost byla dostupná tomu, kdo ji má používat.<sup>26</sup> Zdeněk Zdráhal doplňuje, že proces reprezentace spočívá ve vyjádření znalosti nějakým přirozeným nebo umělým jazykem.<sup>27</sup> Toto široké pojetí se často zužuje na řešení problému reprezentace lidských znalostí strojům, typicky počítačovým programům. Výsledkem reprezentace je vyjádření znalostí a jejich vztahů, v případě vyjádření prostřednictvím pojmů nazývané konceptualizace (blíže viz kapitoly 1.5.2 a 1.5.3). Randall Davis se spoluautory pojmenovali následujících pět rolí reprezentace znalostí, jež se do značné míry kryjí se současným pohledem na oblasti užití informatických ontologií (viz kapitola 1.3). 1. Náhražka (surogát) věci samé, umožňující určit důsledky přemýšlením o ní, 2. množina ontologických závazků (k tomuto pojmu blíže v kapitole 1.5), 3. fragmentární teorie inteligentního usuzování s povolenými a nepovolenými možnostmi odvozování, 4. médium pro efektivní počítačové zpracování, 5. médium pro komunikaci.<sup>28</sup>

---

<sup>23</sup> GRUBER, Thomas R. Ontology. In: Ling LIU, M. Tamer ÖZSU, ed. *Encyclopedia of database systems*. London: Springer, 2009, s. 1963-1965. doi:10.1007/978-0-387-39940-9\_1318. ISBN 978-0-387-35544-3 (Print). ISBN 978-0-387-39940-9 (Online).

<sup>24</sup> SVÁTEK, Vojtěch. Ontologie a WWW. In: Dušan CHLAPEK, ed. *DATAKON 2002: sborník databázové konference: Brno, Česká republika, 19.-22. října 2002*. Brno: Masarykova univerzita, 2002, s. 27-55. ISBN 978-80-210-2958-3.

<sup>25</sup> MCCARTHY, John. Circumscription: a form of non-monotonic reasoning. In: *Artificial intelligence*. April 1980, **13**(1-2), s. 31. doi:10.1016/0004-3702(80)90011-9. ISSN 0004-3702.

<sup>26</sup> NEWELL, Allen. The knowledge level. In: *Artificial intelligence*. January 1982, **18**(1), s. 114. doi:10.1016/0004-3702(82)90012-1. ISSN 0004-3702.

<sup>27</sup> ZDRÁHAL, Zdeněk. Reprezentace znalostí. In: Vladimír MAŘÍK, Jiří LAŽANSKÝ, Olga ŠTĚPÁNKOVÁ a kol. *Umělá inteligence*. 1. vyd. Praha: Academia, 1993, kap. 4, s. 99-122. ISBN 978-80-200-0496-3.

<sup>28</sup> DAVIS, Randall, SHROBE, Howard, SZOLOVITS, Peter. What is a knowledge representation? In: *AI magazine*. Spring 1993, **14**(1), 17-33. ISSN 0738-4602.

Dalším myšlenkovým proudem, který přispěl k formování teoretické základny inforatických ontologií, je tvorba pojmových modelů informačních systémů. Už v 60. letech 20. století uvedl ontologickou problematiku do oblasti informatiky matematik a informatik George H. Mealy (1927–2010), autor koncepce tzv. Mealyho stroje – konečného automatu s výstupem. V článku *Jiný pohled na data*<sup>29</sup> z roku 1967 Mealy zformuloval náčrt obecné teorie dat použitelné pro tvorbu programovacích jazyků. Vychází z filozofického základu a koncipuje významné stavební prvky ontologického pohledu na data. Není bez zajímavosti, že Mealy vedl disertační práci Petera Pin-Shan Chena (1947), autora pojmového ER (z angl. *entity-relationship*, tj. entita-vztah) modelu<sup>30</sup>, který se spolu s relačním modelem dat Edgara F. Codd (1923–2003)<sup>31</sup> stal základem koncepce databázových informačních systémů v druhé polovině 20. století.

V roce 1982 Allen Newell k dvěma obecně akceptovaným úrovním reprezentace počítačových systémů, tj. k úrovni zařízení (komponenty, logické obvody...) a k úrovni symbolické (programy), doplnil úroveň znalostí<sup>32</sup>. Úroveň zařízení a symbolická úroveň obsahují znalosti o vnitřní struktuře systému, jež odrážejí jeho fyzickou stránku. Znalostní úroveň podle Newella pohlíží na systém jako na černou skříňku a zajímá se o jeho fungování vzhledem k okolí, o poskytované služby a řešené problémy. Na toto pojetí navázali v roce 1987 tvůrci technické zprávy ISO/TR 9007 *Pojmy a terminologie pro pojmové schéma a informační základnu*<sup>33</sup>, v níž je definována tříúrovňová architektura dat v informačním systému: pojmové schéma, vnější schéma a vnitřní schéma (podrobněji viz kapitola 1.2). Vnitřní a vnější schéma se zaměřují na znalosti o formě obsažených informací. Pojmové schéma zachycuje znalosti o obsahu informací v informačním systému, čímž odpovídá zhruba Newellově znalostní úrovni počítačových systémů.

Aktuální implementací principu znalostního/pojmového modelu v návrhu informačních systémů a aplikací je od roku 2001 *MDA – Model driven architecture* (Architektura řízená modelem)<sup>34</sup>. Tato specifikace konsorcia *OMG* definuje koncepci architektury softwarových aplikací, jež zahrnuje tři typy postupně konkretizovaných modelů, reprezentujících úrovně znalostí o informačních systémech a aplikacích. Na nejobecnější úrovni je model nezávislý na počítačovém zpracování (*CIM – Computation Independent Model*), zachycující znalosti o oblasti použití systému, například model podnikových procesů. Další úroveň je pojmový model nezávislý na platformě (*PIM – Platform Independent Model*), v němž jsou obsaženy znalosti o pojmech reprezentovaných systémem. Tento model odpovídá pojmovému schématu podle ISO/TR 9007.

---

<sup>29</sup> MEALY, George H. Another look at data. In: *Proceedings of the Fall Joint Computer Conference (AFIPS Fall '67)*, November 14-16, 1967. New York: ACM, 1967, 525-534. doi:10.1145/1465611.1465682.

<sup>30</sup> CHEN, Peter Pin-Shan. The entity-relationship model – toward a unified view of data. In: *ACM Transactions on Database Systems*. March 1976, **1**(1), 9-36. doi:10.1145/320434.320440. ISSN 0362-5915 (Print). ISSN 1557-4644 (Online).

<sup>31</sup> CODD, Edgar F. A relational model of data for large shared data banks. In: *Communications of the ACM*. June 1970, **13**(6), 377-387. doi:10.1145/362384.362685. ISSN 0001-0782 (Print). ISSN 1557-7317 (Online).

<sup>32</sup> NEWELL, Allen. The knowledge level. In: *Artificial intelligence*. January 1982, **18**(1), 87-127. doi:10.1016/0004-3702(82)90012-1. ISSN 0004-3702.

<sup>33</sup> ISO/TR 9007:1987. *Information processing systems - Concepts and terminology for the conceptual schema and the information base*. 1st ed. Geneva: International Organization for Standardization, 1987. 120 s.

<sup>34</sup> Object Management Group. *MDA guide. Revision 2.0* [online]. Needham (MA, USA): Object Management Group, June 2014 [cit. 2018-01-30]. 15 s. Document Number: ormsc/2014-06-01. Dostupné z: <http://www.omg.org/cgi-bin/doc?ormsc/14-06-01>.

Model pro specifickou platformu (*PSM – Platform Specific Model*) odpovídá vnitřnímu schématu informačního systému podle ISO/TR 9007.

Důležitý význam má i další specifikace konsorcia *OMG*, věnovaná metamodelu pro návrh ontologií: *ODM – Ontology Definition Metamodel* (Metamodel definice ontologie).<sup>35</sup> První verze standardu byla publikována v roce 2009 a v roce 2014 byla vydána aktualizovaná verze 1.1. Cílem je sloučit ‚softwarově‘ zaměřený koncept *MDA* s technologiemi sémantického webu a vytvořit zobecněný model struktury informatické ontologie. *ODM* zahrnuje i specifikace vzájemného mapování metamodelů v různých jazycích a jejich mapování do *UML* a *ER* modelů.

S nástupem sémantického webu a propojených dat v první a druhé dekádě 21. století přibyla nová oblast použití ontologií. Dosavadní představa ontologie jako pojmového modelu pro (navrhovaný) znalostní systém je doplněna o pojetí ontologie jako nástroje sémantické interoperability v spontánně vznikající a neustále se měnící síti propojených dat. V současné etapě vývoje ontologií zaměřené na aplikace sémantického webu hraje důležitou standardizační roli konsorcium *W3C*. Jeho nejvýznamnějším příspěvkem jsou jazyky pro návrh a zápis ontologie: *RDF (Resource description framework)*, jehož 1. verze byla publikována v roce 1999), *RDFS (RDF Schema)*, 1. verze publikována v roce 2014), *OWL (Web ontology language)*, 1. verze publikována v roce 2004) a *SKOS (Simple knowledge organization system)*, 1. verze publikována v roce 2009).

## 1.2 Ontologie a znalostní systém

Tradičním kontextem informatických ontologií je informační, resp. znalostní systém. Na nejobecnější úrovni lze dokonce konstatovat, že informační systémy a informatické ontologie vykazují shodné vlastnosti. Název napovídá, že obsah znalostního systému tvoří znalosti. Výchozím pojmem většiny definic informatické ontologie je rovněž znalost a možnosti jejího opětovného použití. Shoda panuje i v technologii: společnou bází informačních i znalostních systémů a informatických ontologií jsou informační technologie, techniky softwarového a znalostního inženýrství a umělé inteligence. V této kapitole se proto zaměříme na specifikaci vzájemného vztahu informačních systémů a ontologií. Využijeme k tomu pojmový model informačního systému definovaný v technické zprávě ISO/TR 9007 *Pojmy a terminologie pro pojmové schéma a informační základnu*<sup>36</sup>, která byla zmíněna v předchozí kapitole. Tento model budeme konkretizovat, aby vystihoval specifika znalostních systémů, a do takto konkretizovaného modelu začleníme pojem informatické ontologie.

Z hlediska funkce lze za informační systém považovat jakýkoli systém umožňující komunikaci a transformaci informací a/nebo znalostí – časově, prostorově i co do formy tak, aby byly lépe využity než v původním stavu. Jde tedy o systém, který přidává hodnotu ke zpracovávaným či komunikovaným informacím a/nebo znalostem za účelem odstranění informačních bariér. Typickými úlohami řešenými informačními systémy jsou kromě saturování potřeby informací a/nebo znalostí (pro poznání, pro rozhodování, pro realizaci určité činnosti) problémy složitosti,

---

<sup>35</sup> Object Management Group. *Ontology definition metamodel (ODM)* [online]. Version 1.1. Needham (MA, USA): Object Management Group, September 2014 [cit. 2018-01-30]. 348 s. OMG Document Number: formal/2014-09-02. Dostupné z: <http://www.omg.org/spec/ODM/>.

<sup>36</sup> Následující citace uvádíme podle českého překladu normy: ČSN ISO/TR 9007 (97 9702). *Systémy zpracování informací: Pojmy a terminologie pro pojmové schéma a informační základnu*. Praha: Český normalizační institut, 1996. 131 s.

opětovné použitelnosti, automatizace, komunikace a problémy bezpečnosti, spolehlivosti a minimalizace rizik.

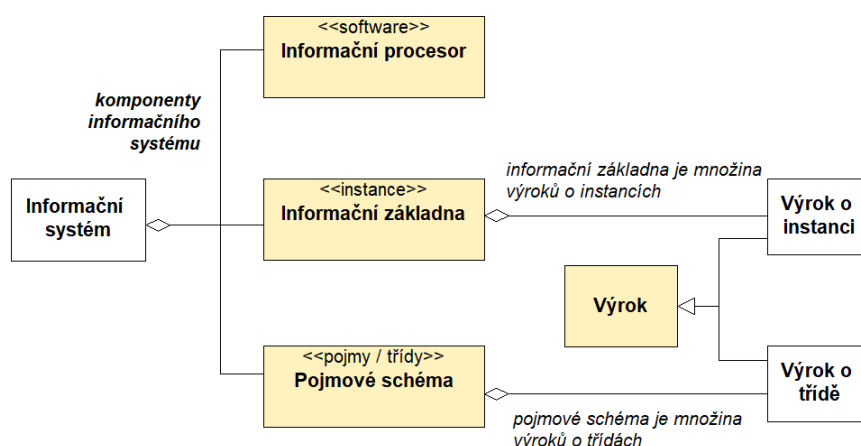
Informační systém představuje dialektickou jednotu statického a dynamického. Má jak svou stránku pasivní – je artefaktem modelujícím svět (modelem, reprezentací, odrazem reality) prostřednictvím informací a/nebo znalostí, tak stránku aktivní – je procesem působícím na svět, procesy, jež realizuje s informacemi, ovlivňují realitu. Yair Wand a Ron Weber nabízejí v souladu s tímto pojetím dva pohledy na informační systém:

„1. *Reprezentace systému reálného světa, tak jak jej někdo vnímá, prostřednictvím artefaktů, vytvořená k tomu, aby plnila funkce zpracování informací.*

2. *Mechanismus zachycující stavy systému reálného světa, který má modelovat.*“<sup>37</sup>

Gove Allen a Salvatore March navrhuji upřesnit formulaci druhého pohledu na „*mechanismus zpracovávající události*“<sup>38</sup>, tak aby byla explicitně vyjádřena dynamická stránka informačního systému.

Technická zpráva ISO/TR 9007 definuje jak statické, tak dynamické komponenty informačního systému. Klíčovými statickými komponentami informačního systému jsou informační základna (angl. *information base*) a pojmové schéma<sup>39</sup> (angl. *conceptual schema*) a dynamickou složku představuje informační procesor (angl. *information processor*). Obrázek 1 znázorňuje obecnou strukturu informačního systému ve formě diagramu tříd v jazyce UML.



Obr. 1 Informační systém podle ISO/TR 9007

Informační procesor je v ISO/TR 9007 specifikován jako „*mechanismus, který podle zadaného příkazu vykonává požadovanou činnost v pojmovém schématu a informační základně*“<sup>40</sup>.

V převážné většině typů informačních systémů plní tuto funkci software.

<sup>37</sup> WAND, Yair, WEBER, Ron. Toward a theory of the deep structure of information systems. In: Janice I. DeGROSS, Maryam ALAVI, Hans J. OPPELLAND, ed. *Proceedings of the Eleventh International Conference on Information Systems, December 16-19, 1990, Copenhagen, Denmark*. Baltimore: ACM Press, © 1990, s. 62.

<sup>38</sup> ALLEN, Gove N., MARCH, Salvatore T. A critical assessment of the Bunge-Wand-Weber ontology for conceptual modeling [online]. In: *16th Annual workshop on information technologies and systems (WITS)*, Milwaukee, WI, Dec. 9-10 2006 [cit. 2018-01-30]. Dostupné z: <http://ssrn.com/abstract=951803>.

<sup>39</sup> Poznámka: V dalším textu práce používáme termín pojmový model, v této pasáži je ponechána původní terminologie použitá v českém překladu normy.

<sup>40</sup> ČSN ISO/TR 9007. *Systémy zpracování informací: Pojmy a terminologie pro pojmové schéma a informační základnu*. Praha: Český normalizační institut, 1996, s. 49.

Na obrázku 1 je znázorněno, že základním stavebním prvkem statických komponent informačního systému jsou výroky (věty), jež tvoří základ jak pro informační základnu, tak pro pojmové schéma.

Informační základnu tvoří podle ISO/TR 9007 „soubor vět konzistentních jak navzájem, tak i s pojmovým schématem, vyjadřující výroky odlišné od nutných výroků, které platí pro určitý svět entit“<sup>41</sup>. V konkrétních typech informačních systémů jsou pro informační základnu používány specifické termíny – informační základna transakčních databázových systémů se označuje jako databáze<sup>42</sup> (angl. *data base*), informační základnu znalostních systémů tvoří znalostní báze (angl. *knowledge base*). V současné době se výroky v informační základně často označují jako instance.

Norma ISO/TR 9007 věnuje pozornost především pojmovému schématu. V souvislosti s ním zavádí významný pojem ‚univerzum diskurzu‘, tj. obor rozpravy či úvah<sup>43</sup>, jímž se jednoznačně vymezuje sémantický prostor v rámci nějaké úlohy (například při pojmovém modelování) či zkoumání v konkrétních vědních disciplínách. Norma vymezuje univerzum diskurzu následovně: „*Jsou to všechny entity, o které se zajímáme, které byly, jsou nebo mohou být.*“<sup>44</sup>. Pojmové schéma je v normě definováno jako „*ucelený soubor vět vyjadřujících nutné výroky, které platí pro univerzum diskursu*“<sup>45</sup>. Na rozdíl od instancí v informační základně jsou jednotky pojmového schématu obvykle označovány jako třídy. Cílem pojmového schématu je jednoznačně definovat význam dat v informační základně pro různé uživatele. Význam jednoznačného definování se v normě zdůvodňuje následovně:

*„Nejdůležitější charakteristikou prostředí datových základen je, že společná data jsou sdílena mezi mnoha uživateli jednoho systému. Sdílením společných dat zakládají tito uživatelé prostřednictvím systému mezi sebou dialog. Je zřejmé, že pro užitečnost a spolehlivost této komunikace musí existovat jednotné chápání informací prezentovaných daty. Protože se může stát, že dva uživatelé se nikdy nesetkají, musí být společné chápání řízeno něčím, co stojí mimo ně. Toto společné chápání musí být zaznamenáno a v pravidlech k založení dialogu musí být ustanovena potřebná, předem stanovená gramatická pravidla.“*<sup>46</sup>

Na obrázku 2 je základní schéma informačního systému doplněno o entity významné pro znalostní systém. Znalostní systém je znázorněn pomocí symbolu generalizace jako specifický typ informačního systému, což vyjadřuje, že jeho charakteristika je odvozena z obecných vlastností informačního systému, týkajících se jeho obsahu, funkce a klíčových komponent. Tyto ‚zdeděné‘ obecné vlastnosti jsou pak v rámci znalostního systému specifikovány a doplněny o další důležité charakteristiky.

---

<sup>41</sup> Tamtéž, s. 50.

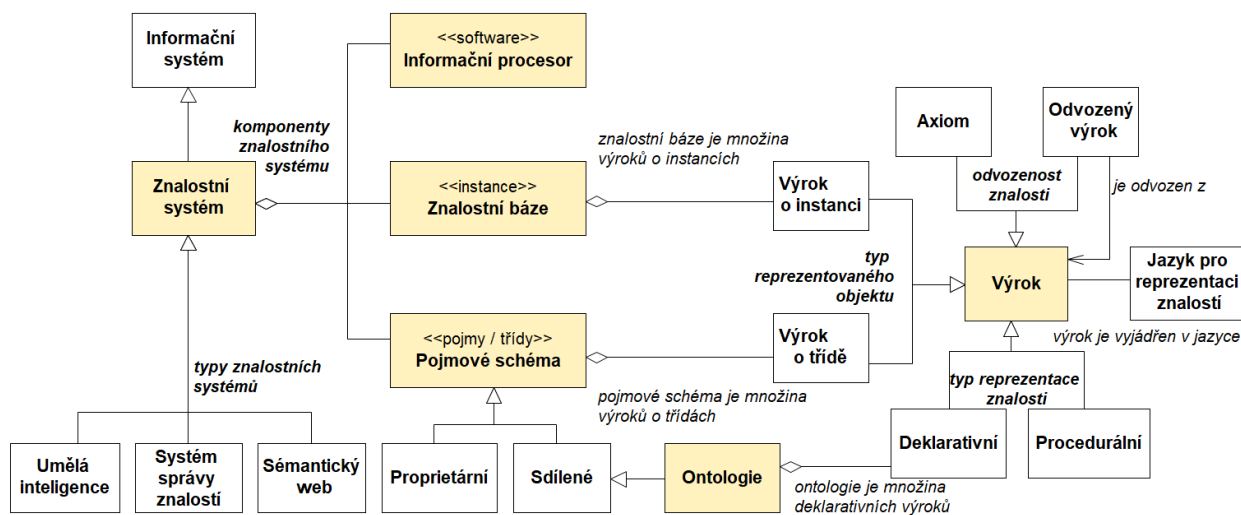
<sup>42</sup> Poznámka: V českém překladu normy se používá termín ‚datová základna‘.

<sup>43</sup> „*Diskurz je projev racionálního myšlení v podobě souvislé řady úsudků, jež se opírají o dílčí operace a pojmy.*“ DUROZOI, Gérard, ROUSSEL, André. *Filozofický slovník*. Přeložili Jan Binder aj. 1. vyd. Praha: EWA, 1994, s. 61. ISBN 80-85764-07-5.

<sup>44</sup> ČSN ISO/TR 9007. *Systémy zpracování informací: Pojmy a terminologie pro pojmové schéma a informační základnu*. Praha: Český normalizační institut, 1996, s. 16.

<sup>45</sup> Tamtéž, s. 50.

<sup>46</sup> Tamtéž, s. 8.



Obr. 2 Znalostní systém s ontologií podle ISO/TR 9007

K chápání obsahu a funkce znalostního systému se zpravidla jinak přistupuje v oblasti umělé inteligence, a jinak v pragmatičtější orientované oblasti znalostního managementu. Vyjdeme z typologie Joerga Evermanna, který rozlišuje dvě kategorie znalostních systémů<sup>47</sup>: systémy založené na znalostech a systémy správy znalostí.

Systémy založené na znalostech (angl. *knowledge-based systems*) používají explicitní znalosti k automatickému usuzování a k odvozování nových znalostí. Obsažené znalosti jsou reprezentované formálními jazyky založenými na logice. Jejich uživatelem je software, například inteligentní agent. Tato kategorie je ve schématu na obrázku 2 označena v souladu s jejími teoretickými principy jako ‚umělá inteligence‘.

Systémy správy znalostí (angl. *knowledge management systems*) spravují explicitní znalosti pro účely jejich sdílení mezi lidskými aktéry. Obvykle se vymezují v kontextu nějaké organizace. Organizace jsou potom považovány za místo zpracování a uchování znalostí a vytváření podmínek pro vznik nových znalostí (inovace). Tyto systémy obsahují zpravidla znalosti reprezentované přirozeným jazykem.

Třetí kategorií znalostních systémů znázorněnou na obrázku 2 je sémantický web, do jehož prostředí se aktuálně přesouvají mnohé znalostní systémy z původních oblastí použití. Znalosti jsou obsaženy v distribuovaných webových zdrojích a jejich typickou vlastností je diverzita (blíže viz v kapitole 1.5.5). Webové zdroje mohou obsahovat jak rigorózně formalizované znalosti, tak znalosti vyjádřené neformálně.

Specifika znalostních systémů v rámci množiny informačních systémů se týkají všech komponent popsaných výše: informační základny, pojmového schématu i informačního procesoru.

Informační procesor znalostního systému umožňuje realizovat takové operace s daty, při nichž dochází ke změně a/nebo k tvorbě nového obsahu či rozhodovací procedury díky netriviálnímu odvozování nových znalostí ze znalostí zachycených ve znalostní bázi. To je umožněno tím, že na rozdíl od transakčních informačních systémů je softwaru, který realizuje operace s daty,

<sup>47</sup> EVERMANN, Joerg. Towards a cognitive foundation for knowledge representation. In: *Information systems journal*. April 2005, 15(2), 147-178. doi:10.1111/j.1365-2575.2005.00193.x. ISSN 1350-1917 (Print). ISSN 1365-2575 (Online).



srozumitelná nejen syntax, ale díky ontologii i sémantika znalostní báze.

Stejně jako v informačním systému je i ve znalostním systému základním stavebním prvkem výrok coby konkrétní forma znalosti zaznamenané ve znalostním systému. Aby byl obsah pojmového schématu a informační základny komunikovatelný, je zapotřebí jej vyjádřit nějakým jazykem. Tuto komponentu znalostního systému znázorňuje na obrázku 2 třída ‚Jazyk pro reprezentaci znalostí‘.

Na obrázku 2 je znázorněna typologie výroků ve znalostním systému. K již uvedené typologii výroků v informačním systému, jež člení výroky podle typu reprezentovaného objektu na výroky o instancích a na výroky o třídách, jsou doplněny další dvě typologie. Podle odvozenosti znalosti se výroky člení na axiomatické a odvozené. Z hlediska obsahu, přesněji řečeno podle způsobu reprezentace znalostí, se výroky ve znalostním systému obvykle člení na deklarativní a procedurální. ‚Intelligence‘ deklarativních výroků je obsažena v datech a jejich strukturách, ‚intelligence‘ výroků procedurálních je obsažena v procesech/programech, pravidlech, heuristikách. K těmto dvěma typům intelligence by bylo možné ještě připojit další typ, který přináší koncept propojených dat v sémantickém webu – ‚intelligence‘ je obsažena v sémantickém propojení dat, tj. v jejich vztazích.

Členění výroků na procedurální a deklarativní je odvozeno z termínů používaných pro odlišení dvou typů znalostí – procedurálních a deklarativních. Procedurální (operační) znalost umožňuje úspěšnou realizaci nějaké akce nebo činnosti, deklarativní (konceptuální) znalost má formu tvrzení či výroku o poznané entitě, u něž má smysl určovat, zda je pravdivý. Jejich specifika se projevují jak v modelování, tak ve způsobu reprezentace a v mechanismech odvozování a usuzování. Modely procedurálních znalostí se zaměřují na metody a algoritmy řešení úloh či problémů ve formě pravidel využívaných k diagnostice a rozhodování komplikovaných problémů, typickou aktuální aplikací jsou webové služby. Modely deklarativních znalostí směřují k umožnění přístupu ke znalostem (vyhledávání a objevování) a ke generování nových znalostí. Jejich převažující formou jsou informatické ontologie. Typickou aktuální aplikací jsou znalostní báze. Zatímco pro rané stadium výzkumu a aplikací v oblasti umělé intelligence byla typická snaha o modelování, reprezentaci a odvozování procedurálních znalostí, ontologie obrátily pozornost odborníků na znalosti deklarativní.

Uvedené členění navozuje dojem, že je vždy možné prohlásit znalost buď za deklarativní, nebo za procedurální, ve skutečnosti se však běžně vyskytují znalosti, vykazující rysy obou skupin. Tento názor zastává Allen Newell, který upozorňuje na obtíže s vyjádřením a zachycením znalosti a zároveň na její dialektickou povahu:

*„Znalost není reprezentovatelná strukturou na úrovni symbolů. Vyžaduje jak struktury, tak procesy. Znalost zůstává navždy abstraktní a nikdy ji nelze mít skutečně k dispozici (pod kontrolou)”<sup>48</sup>.*

Příkladem snahy o propojení deklarativních a procedurálních znalostí je současné objektově orientované paradigma softwarových aplikací. V diagramech tříd, které se používají k jejich modelování, se v jedné třídě současně objevují jak deklarativní znalosti v podobě atributů, tak procedurální znalosti v podobě metod. V abstraktním pohledu von Neumannovy architektury počítačů dokonce deklarativní a procedurální znalosti splývají – jak data, tak programy jsou uloženy v paměti ve stejném formátu.

---

<sup>48</sup> NEWELL, Allen. The knowledge level. In: *Artificial intelligence*. January 1982, **18**(1), s. 125. doi:10.1016/0004-3702(82)90012-1. ISSN 0004-3702.

Kromě sdílení různými uživateli připadá v úvahu i sdílení pojmového schématu více systémy. Z tohoto pohledu se pojmová schémata znalostních systémů člení na proprietární nebo sdílená. Proprietární pojmové schéma je vytvořeno speciálně pro jeden konkrétní systém, sdílené schéma je použitelné pro více systémů. Nejvýznamnějším typem sdílených pojmových schémat jsou inforatické ontologie.<sup>49</sup> Jak bylo uvedeno, obvykle mají formu množiny deklarativních výroků o třídách objektů v dané oblasti zájmu. Díky své deklarativní povaze vykazují vyšší míru nezávislosti na programech, jež s nimi mohou pracovat, a tím i vyšší míru opětovné použitelnosti.

Informační základnu znalostního systému představuje znalostní báze, tvořená množinou výroků reprezentujících tvrzení o univerzu diskurzu, vyjádřených v jazyce pro reprezentaci znalostí. Na rozdíl od informačních základen transakčních informačních systémů se znalostní báze vnímá jako těsněji spjatá s pojmovým schématem, v literatuře se lze setkat i s tím, že termínem znalostní báze je označován komplex pojmového schématu a instancí (tj. dat ve znalostní bázi). Existují i ontologie, v nichž jsou kromě výroků o třídách obsaženy i výroky o instancích.

Znalostní báze je klíčovou komponentou všech tří výše uvedených typů znalostních systémů, často ovšem s odlišnými funkcemi. V systémech umělé inteligence se báze znalostí využívá jako podmínka, východisko pro procedury či aktivity – například pro plánování, rozhodování, diagnostiku, rozpoznávání ad. Systémy správy znalostí typicky chápou bázi znalostí jako výsledný produkt – zdroj poznání umožňující opětovné použití a vyhledávání a jako podmínku, východisko pro komunikaci. Znalostní bázi sémantického webu tvoří distribuované webové zdroje opatřené sémantickými metadaty. Na rozdíl od proprietárních systémů umělé inteligence a systémů správy znalostí disponuje sémantický web standardizovaným jazykem i softwarem. Jazykem pro reprezentaci znalostí v sémantickém webu je *RDF*<sup>50</sup>. ‚Věty‘ v *RDF* tvoří výroky ve formě oznamovacích vět v jednotné struktuře trojic, formalizovaných jako výroky predikátové logiky: subjekt (*URI*<sup>51</sup> pojmu A) – predikát (*URI* vztahu) – objekt (*URI* pojmu B, příp. literál, tj. hodnota vztahující se k pojmu B, vyjádřená řetězcem znaků; je-li například pojmem B tvůrce, může literál obsahovat jeho jméno). Analogií k textovým odstavcům, jež tvoří množina vět, jsou *RDF* grafy, tvořené množinou trojic *RDF*. Vztahy v sémantickém webu jsou implementovány jako linky *RDF*, jež mají rovněž formát trojic *RDF*. Na rozdíl od *HTML* linku (značka <a> s atributem ‚href‘), který neobsahuje žádnou sémantiku, je *RDF* link ‚typovaný‘, jeho sémantika je určena sémantikou použitého predikátu. Funkci informačního procesoru plní protokol *HTTP*.

---

<sup>49</sup> Vzájemný vztah pojmových modelů a inforatických ontologií je rozpracován v: WEBER, Ron. Conceptual modeling and ontology: possibilities and pitfalls. In: *Journal of database management*. July-September 2003, 14(3), 1-20. doi:10.4018/jdm.2003070101. ISSN 1063-8016.

<sup>50</sup> *Resource description framework (RDF)* [online]. W3C RDF Working Group, last modified 2017-09-26 [cit. 2018-01-30]. Dostupné z: <http://www.w3.org/RDF/>.

<sup>51</sup> BERNERS-LEE, Tim, FIELDING, Roy T., MASINTER, Larry. *Uniform resource identifier (URI): generic syntax* [online]. RFC 3986. IETF, 2005 [cit. 2018-01-30]. Dostupné z: <https://www.ietf.org/rfc/rfc3986.txt>.

### 1.3 Oblasti užití inforatických ontologií

V úvodu této kapitoly bylo uvedeno, že za hlavní oblasti využití inforatických ontologií lze na obecné úrovni považovat komunikaci znalostí, počítačové odvozování znalostí a opakované využití a organizaci znalostí. V následujícím souhrnu doplníme k této trojici i oblasti užití, formulované Zdeňkem Zdráhalem<sup>52</sup>, a v teoretické rovině se je pokusíme propojit s odpovídajícími sémiotickými pojmy, tak jak jsou interpretovány v kapitole 1.5.2.

#### Využití ontologií pro komunikaci znalostí

*„Klíčovým problémem komunikace je porozumění a pochopení smyslu sdělení.“*<sup>53</sup> Převědeme-li tato slova Zdeňka Zdráhalu do terminologie, použité při interpretaci sémiotického trojúhelníku v kapitole 1.5.2, lze tuto úlohu formulovat slovy ‚od označení k pojmu‘, což je hrana trojúhelníku označující vztah ‚význam‘. Ontologie plní v rámci této úlohy funkci slovníku s explicitně definovanými významy zahrnutých pojmů.

V případě využití ontologií pro komunikaci znalostí je vhodné zdůraznit, že jde o komunikaci v obou jejích významech – tj. sdělování i sdílení. Sdělování znalostí spočívá v jejich přenosu v prostoru nebo v čase. Sdílení znalostí je založeno na existenci nějaké komunikační platformy, jež umožňuje spojení a společnou činnost. Rozpor mezi heterogenním sémantickým prostředím počítačových sítí a potřebou společné platformy pro komunikaci (interoperabilitu) pomáhají řešit právě ontologie ve funkci sémantického rozhraní. Protože uživatelé ontologie mohou být nejen lidé, ale i technická zařízení, především počítačové programy, má ontologie za úkol zprostředkovat smysluplnou komunikaci jak mezi lidmi (člověk–člověk), tak mezi lidmi a počítačovými systémy (člověk–stroj) a dokonce i mezi počítačovými systémy navzájem (např. webové služby, inteligentní agenti, tj. komunikace stroj–stroj). Tedy všude tam, kde je třeba se domluvit s ostatními členy odborné komunity a zejména tam, kde se počítá s inteligentní softwarovou podporou takovéto komunikace, využijeme ontologie.

#### Využití ontologií pro opakované použití a organizaci znalostí

Ve schématu sémiotického trojúhelníku na obrázku 5 v kapitole 1.5.2 je tato funkce ontologie vyznačena interakcí agenta s pojmy a s jejich označeními. V této oblasti použití funguje ontologie jako opakovaně použitelný pojmový model (viz kapitola 1.2). Opakovatelné využití neboli znovupoužitelnost znalostí zachycených v ontologii se projevuje několika způsoby. Jednak tím, že ontologický pojmový model lze použít pro více různých zdrojů (v tom se liší například od jednoúčelových databázových schémat) a navíc jej může použít více různých uživatelů pro více různých účelů. Další možností je propojení s jinými ontologiemi. Ve všech případech je podmínkou jednotné chápání sémantiky znalostí těmi, kdo je vytvořili a zaznamenali a těmi, kdo je hodlají využít nebo zpracovat.

Tento způsob aplikace ontologií v praxi je nejnepřítěžnějším případem využití tzv. lehkých terminologických ontologií (viz kapitola 1.4). Do značné míry se překrývá s tím, jak mají svůj cíl definován tradiční systémy organizace znalostí (o nich blíže v kapitole 4.5.3). Zjednodušeně řečeno, využívají se k popisu obsahu a struktury zaznamenaných organizovaných znalostí.

---

<sup>52</sup> ZDRÁHAL, Zdeňk. Ontologie: od filosofie k umělé inteligenci. In: Vladimír MAŘÍK, Olga ŠTĚPÁNKOVÁ, Jiří LAŽANSKÝ a kol. *Umělá inteligence*. 6. 1. vyd. Praha: Academia, 2013, s. 21-28. ISBN 978-80-200-2276-9.

<sup>53</sup> Tamtéž, s. 26.

## Využití ontologií k popisu skutečnosti pro její zpracování / podporu počítačovými systémy

Při úvahách o úloze nástrojů pro zpracování znalostí lze obrazně konstatovat, že zatímco cílem slovníků je nastolit pořádek mezi slovy a systém organizace znalostí se o totéž snaží u pojmů, vlastním zájmem ontologie je realita, již se snaží reprezentovat natolik explicitně, aby to bylo srozumitelné i počítačovým programům.<sup>54</sup> Zdeněk Zdráhal to formuluje následovně: „*Jednou ze základních součástí většiny informačních systémů je explicitní či implicitní popis (model, zobrazení) části reality, ke které se informační systém určitým způsobem vztahuje.*“<sup>55</sup> Tento případ užití chápe ontologii jako účelovou reprezentaci či model znalostí o skutečnosti, určenou k využití počítačovými systémy (tj. informačními systémy, softwarovými aplikacemi a službami). Koresponduje s tím, jak byla popsána úloha ontologií ve znalostních systémech (viz kapitola 1.2).

Sémiotický trojúhelník zobrazuje tuto úlohu trajektorií ve směru od věcí k pojmům a k jejich označení. V tomto směru ontologie podporují myšlenku sémantického webu, jehož cíl charakterizoval Tim Berners Lee lapidárně jako přechod od webu dokumentů k webu dat (angl. *web of data*)<sup>56</sup>. Tato vize předpokládá, že počítačové programy, které byly dosud využívány především k rychlé mechanické manipulaci se soubory dat (typicky k uložení a přenosu z jednoho místa na druhé), budou díky ontologiím schopny takových operací, které vyžadují porozumět obsahu dat – kategorizace, klasifikace, určení relevance ad.

### Využití ontologií pro automatické odvozování znalostí

Podmínkou úspěšné realizace této úlohy je opět jednotné chápání sémantiky pojmů kognitivními agenty (tj. lidmi i stroji). Využití ontologie pro automatické usuzování, odvozování a objevování nových znalostí umožňuje explicitní zachycení netriviálních vztahů mezi zahrnutými pojmy a pravidel vymezujících jejich validitu. Při odvozování se pak používá jazyk založený na některém z logických kalkulů (predikátová logika, deskripční logika ad.) Touto ambicí se ontologie přibližují k oboru umělé inteligence, v jejímž rámci ostatně byl termín ‚uveden do oběhu‘ (viz kapitola 1.1).

---

<sup>54</sup> Je vhodné upozornit, že ‚realita‘ v tomto pojetí zahrnuje nejen to, co aktuálně a prokazatelně existuje, ale i představy a myšlenkové konstrukce.

<sup>55</sup> ZDRÁHAL, Zdeněk. Ontologie: od filosofie k umělé inteligenci. In: Vladimír MAŘÍK, Olga ŠTĚPÁNKOVÁ, Jiří LAŽANSKÝ a kol. *Umělá inteligence*. 6. 1. vyd. Praha: Academia, 2013, s. 21. ISBN 978-80-200-2276-9.

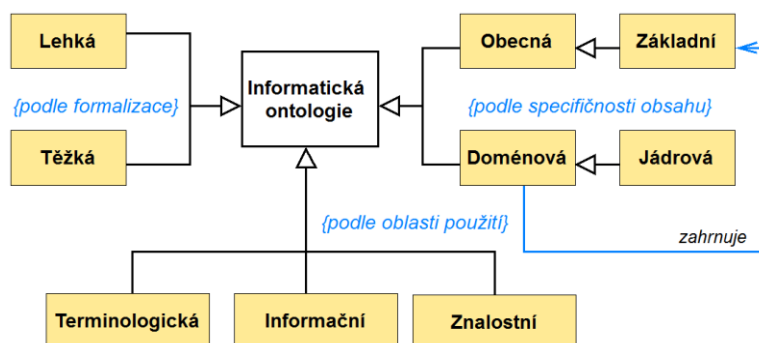
<sup>56</sup> ‚Data‘ jsou v tomto případě metaforou pro informace a znalosti.

## 1.4 Typologie inforatických ontologií

Stejně jako nepanuje jednotné mínění o vymezení rozsahu a obsahu pojmu ontologie, existují i různé přístupy k jejich typologii. Kupříkladu typologie zpracovaná účastníky konference *Ontology Summit 2016*<sup>57</sup> zahrnuje následující typy:

- vrchní/vrcholové ontologie, ontologie vyšší úrovně (*upper ontologies*)
- referenční ontologie (*reference ontologies*)
- aplikační a lokální ontologie (*application and local ontologies*)
- překlenovací („přemostující“) ontologie (*bridge ontologies*)
- obecné metadatové šablony a metadatová schémata (*common metadata templates and metadata schema*)

V následujícím stručném přehledu se zaměříme na tři kritéria členění, jež vedou ke vzniku tří rozdílných typologií: podle formalizace, podle oblasti použití a podle specifčnosti obsahu. Tyto tři typologie jsou schematicky znázorněny na obrázku 3 ve formě diagramu tříd UML. Je ovšem třeba upozornit, že obecně je pocítována absence přesných dělicích hranic mezi jednotlivými typy a převažuje pojetí typologie inforatických ontologií jako plynulého spektra.



Obr. 3 Typologie inforatických ontologií

### Typologie podle míry formalizace a typu konceptualizace

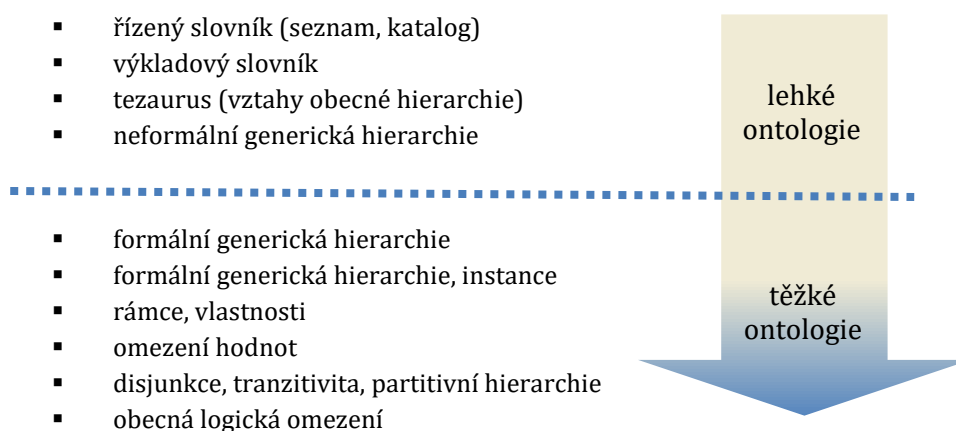
Charakter plynulých přechodů je patrný již v jedné z prvních typologií zpracované Michaellem Uscholdem a Michaellem Grüningerem<sup>58</sup>, kteří člení ontologie podle míry formálnosti jazyka, v němž jsou jimi reprezentované znalosti vyjádřeny. Vysoce neformální ontologie jsou vyjádřené volně v přirozeném jazyce, semi-neformální ontologie jsou vyjádřené v omezené a strukturované formě přirozeného jazyka, semi-formální ontologie jsou vyjádřené v umělém, formálně definovaném jazyce a rigorózně formální ontologie obsahují přesně (striktně) definované termíny s formální sémantikou, teorémy a důkazy (např. vlastnosti bezespornosti nebo úplnosti), plně vyjádřitelné programovacím jazykem.

Typologie Debory L. McGuinnessové je další z často citovaných typologií ve formě spektra. Řadí inforatické ontologie podle vyjadřovací (sémantické) síly a zároveň dle stupně formalizace.

<sup>57</sup> *Ontology Summit 2016 communique: Ontologies within semantic interoperability ecosystems* [online]. Donna Fritzsche, Michael Gruninger, ed. 2016 [cit. 2018-01-30]. Dostupné z: <http://ontologforum.org/index.php/OntologySummit2016/Communique>.

<sup>58</sup> USCHOLD, Michael, GRÜNINGER, Michael. *Ontologies: principles, methods and applications* [online]. Edinburgh: February 1996 [cit. 2018-01-30], s. 6. Technical Report, AIAI-TR-191. University of Edinburgh, School of Informatics, Artificial Intelligence Applications Institute. Dostupné z: <http://www.aiai.ed.ac.uk/publications/documents/1996/96-ker-intro-ontologies.pdf>.

Jednotlivé typy znázorněné na obrázku 4 jsou označeny prostřednictvím pojmenování nástrojů, jež jsou pro daný typ ontologií charakteristické.<sup>59</sup>



Obr. 4 Typy ontologií podle Debory L. McGuinnessové

Autorka navíc člení toto plynulé spektrum na dvě skupiny, jež se obvykle označují jako tzv. lehké a těžké (angl. *light-weight, heavy-weight*) ontologie.<sup>60</sup>

Lehké ontologie obsahují pojmy, vztahy mezi pojmy a definice, jež popisují pojmy. Zahrnují tedy nástroje, pro něž se v oblasti knihovnictví používá termín ‚systém organizace znalostí‘ (podrobněji viz kapitola 4.5.3).

Za informatické ontologie v pravém slova smyslu se zpravidla považují až systémy s formální generickou hierarchií. Rovněž podle normy ISO 25964, jež ontologie uvažuje v kontextu systémů organizace znalostí, je účelné za ontologie považovat pouze těžké ontologie.<sup>61</sup> Přínos těžkých ontologií oproti tradičním systémům organizace znalostí spočívá v bohatší sémantice pojmů i jejich vztahů, vyšším stupni formalizace a v těsnějším sepětí s informačními a komunikačními technologiemi.

Pojmy těžkých ontologií jsou stejně jako v lehkých ontologiích určeny označením, definicemi a navíc vymezením svých vlastností. Obsahují více omezení, jež vyjasňují význam zahrnutých pojmů. Dalším přínosem je bohatší sémantika vztahů – za těžké ontologie se považují systémy schopné vyjádřit komplexnější vztahy mezi pojmy, než jsou obecné vztahy ekvivalence, hierarchie a asociace, obvyklé například v knihovnických klasifikacích a v tezaurech, případně v databázových schématech. Sémantika těžkých ontologií je nejen explicitně vyjádřena, ale i formalizována prostřednictvím logických axiomů, omezení nebo pravidel. Jednoznačně a

<sup>59</sup> MCGUINNESS, Deborah L. Ontologies come of age. In: FENSEL, Dieter, HENDLER, James A., LIEBERMAN, Henry, WAHLSTER, Wolfgang, ed. *Spinning the semantic Web: bringing the World Wide Web to its full potential*. 1st ed. Cambridge (Mass.): MIT Press, 2003, s. 171-196. ISBN 978-0-26256212-6 (brož.). ISBN 978-0-262-06232-9 (váz.).

<sup>60</sup> GÓMEZ-PÉREZ, Asunción, FERNÁNDEZ-LÓPEZ, Mariano, CORCHO, Oscar. *Ontological engineering: with examples from the areas of knowledge management, e-commerce and the semantic web*. 1st ed. London: Springer, 2004, s. 8. doi:10.1007/b97353. ISBN 978-1-85233-551-9 (Print). ISBN 978-1-85233-840-4 (Online).

<sup>61</sup> ISO 25964-2:2013. *Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies*. 1st ed. Geneva: International Organization for Standardization, 2013-03-04, s. 72-73.

formalizovaně vymezená sémantika pojmů umožňuje automatické odvozování znalostí. Odvozovat je možné nejen na deklarativní (pojmové) úrovni, ale i na úrovni procesů v plánovacích a analytických úlohách. Významnou součástí těžké ontologie je obvykle software, který umožňuje její vytvoření, editaci i využití.

Lze konstatovat, že rozdíl mezi lehkými a těžkými ontologiemi se projevuje i v možnostech jejich využití (viz kapitola 1.3). Všechny oblasti užití jsou zpravidla schopny pokrýt pouze těžké ontologie, oblasti užití lehkých ontologií se obvykle omezují na první dva uvedené okruhy, tj. na podporu komunikace a opakovaného využití a organizace znalostí.

### Typologie ontologií podle oblasti použití

Zajímavou typologii, již nazývá členění „podle historických paradigmat“, představuje Vojtěch Svátek<sup>62</sup> v podobě členění na terminologické, informační a znalostní ontologie. Tato typologie odráží existenci různých odborných komunit, jež ontologie vytvářejí a používají, a jimi aplikovaných vědních disciplín.

Terminologické či lexikální ontologie<sup>63</sup> využívají obory terminologie, jazykověda, informační věda a knihovnictví. Jejich jádrem je formalizovaný slovník, tj. zahrnují řízené slovníky a systémy organizace znalostí (viz kapitola 4.5.3). V komunitě paměťových a fondových institucí se pro ně aktuálně používá označení slovníky metadat (angl. *value vocabularies*).<sup>64</sup> Obvykle spadají do kategorie lehkých ontologií. Příkladem je ontologie *WordNet* nebo tezaurus *AGROVOC*. John F. Sowa je definuje jako ontologie, jejichž kategorie nemusejí být plně specifikovány axiomy a definicemi a terminologickou ontologií staví do protikladu k tzv. axiomatizované ontologii, což je podle něj „*ontologie, jejíž kategorie jsou rozlišeny pomocí axiomů a definic uvedených v jazyce logiky nebo v nějakém počítačovém jazyce, jenž může být automaticky do logického jazyka přeložen.*“ Sowa zároveň konstatuje, že rozdíl mezi axiomatizovanou a terminologickou ontologií je spíše záležitostí míry axiomatizace a možností automatického odvozování a definování než rozdíl v typu.<sup>65</sup>

Teoretickým základem informačních ontologií je informatika. Zahrnují datové modely a metadatová schémata (tj. modely struktury metadat, angl. *metadata element sets*). Typickým zástupcem tohoto typu ontologií jsou *Abstraktní model DCMI*<sup>66</sup> nebo ontologický metamodel

---

<sup>62</sup> SVÁTEK, Vojtěch. Ontologie a WWW. In: Dušan CHLAPEK, ed. *DATAKON 2002: sborník databázové konference: Brno, Česká republika, 19.-22. října 2002*. Brno: Masarykova univerzita, 2002, s. 29-30. ISBN 978-80-210-2958-3.

<sup>63</sup> Viz např. LACASTA, Javier, NOGUERAS-ISO, Javier, ZARAZAGA-SORIA, Francisco Javier. *Terminological ontologies: design, management and practical applications*. New York: Springer, 2010. xvii, 197 s. Semantic Web and Beyond. Computing for Human Experience, vol. 9, ISSN 1559-7474. doi:10.1007/978-1-4419-6981-1. ISBN 978-1-4419-6980-4 (Print). ISBN 1-4419-6980-2 (Print). ISBN 978-1-4419-6981-1 (Online).

<sup>64</sup> ISAAC, Antoine, WAITES, William, YOUNG, Jeff, ZENG, Marcia. *Library Linked Data Incubator Group: Datasets, value vocabularies, and metadata element sets* [online]. W3C Incubator Group report 25 October 2011. W3C, 2011 [cit. 2018-01-30]. Dostupné z: <http://www.w3.org/2005/Incubator/lld/XGR-lld-vocabdataset-20111025/>.

<sup>65</sup> SOWA, John F. *Knowledge representation: logical, philosophical, and computational foundations*. Pacific Grove: Brooks/Cole, © 2000. s. 493-497. Computer Science Series. ISBN 0-534-94965-7. ISBN 978-0-534-94965-5.

<sup>66</sup> POWELL, Andy, NILSSON, Mikael, NAEVE, Ambjörn, JOHNSTON, Pete, BAKER, Thomas. *DCMI abstract model*. DCMI, 2007-06-04 [cit. 2018-01-30]. Dostupné z: <http://dublincore.org/documents/abstract-model/>.

ODM<sup>67</sup>.

Znalostní ontologie tvoří formalizované logické teorie (např. základní ontologie *KR Ontology* navržená J. F. Sowou). Obvykle spadají do kategorie těžkých ontologií, jejich teoretickým základem je logika a umělá inteligence.

### Typologie podle povahy zahrnutých entit / podle specifčnosti

Tato typologie se zaměřuje na povahu univerza diskurzu stanoveného pro danou ontologii. Podle toho, jak je vymezena množina jimi reprezentovaných entit, se ontologie člení na obecné (též univerzální, angl. *generic*) a doménové (angl. *domain*).

Obecné ontologie jsou všeobecně zaměřené, doménově nezávislé a co do rozsahu neohraničené. Příkladem obecné ontologie je *DBpedia Ontology*<sup>68</sup>, vytvořená z označení položek infoboxů v článcích Wikipedie. Obdobně ontologie *Schema.org* zahrnuje témata používaná k označení obsahu webových stránek, jež byla shromážděna provozovateli internetových vyhledávačů. *YSO – General Finnish ontology*<sup>69</sup> je příkladem obecné ontologie vzniklé v prostředí paměťových a kulturních institucí.

Specifickým typem obecných ontologií jsou základní (angl. *foundational*) / vrchní (angl. *upper, top level*) ontologie, nazývané též ontologie vyšší úrovně. Stefano Borgo a Claudio Masolo charakteristiku základních ontologií shrnují následovně:

„1. mají velký rozsah, 2. jsou ve vysoké míře opětovně použitelné v různých modelovacích scénářích, 3. mají dobré filozofické a konceptuální základy, a 4. jsou sémanticky transparentní a (proto) bohatě axiomatizované.“<sup>70</sup>

Účelem základních ontologií je sloužit jako výchozí bod pro tvorbu nových doménových ontologií. Obsahují vícenásobně použitelné entity, obvykle základní kategorie na nejvyšší úrovni obecnosti (viz kapitola 1.5.3) a někdy i návrhové vzory (viz kapitola 2.1.3). Vztah základní ontologie a doménové ontologie vytvořené s její pomocí je hierarchický (inkluzivní), tj. doménová ontologie zahrnuje, typicky na nejvyšší hierarchické úrovni, pojmy převzaté ze základní ontologie. Příklady základních ontologií jsou *BFO – Basic formal ontology*<sup>71</sup> a *DOLCE* (viz kapitola 5.3.1).

Doménové ontologie se omezují na konkrétní oblast. Příkladem doménově specifické ontologie je genová ontologie *GO – Gene ontology*<sup>72</sup>. Doménu lze vymežit nejen tematicky, ale i problémově – tak by bylo možné rozlišovat mezi ontologiemi orientovanými na deskriptivní a na procedurální znalosti (viz např. ontologie webových služeb *OWL-S*<sup>73</sup>). Zvláštním typem doménové ontologie je tzv. jádrová (angl. *core*) ontologie, obsahující centrální pojmy domény.

<sup>67</sup> Object Management Group. *Ontology definition metamodel (ODM)* [online]. Version 1.1. Needham (MA, USA): Object Management Group, September 2014 [cit. 2018-01-30]. 348 s. OMG Document Number: formal/2014-09-02. Dostupné z: <http://www.omg.org/spec/ODM/>.

<sup>68</sup> Dostupné z: <http://wiki.dbpedia.org/services-resources/ontology> [cit. 2018-01-30].

<sup>69</sup> Dostupné z: <http://finto.fi/yso/> [cit. 2018-01-30].

<sup>70</sup> BORGIO, Stefano, MASOLO, Claudio. Foundational choices in DOLCE. In: STAAB, Steffen, STUDER, Rudi, ed. *Handbook on ontologies*. 2nd ed. Dordrecht: Springer, 2009, s. 361. doi:10.1007/978-3-540-92673-3\_16. ISBN 978-3-540-70999-2 (Print). ISBN 978-3-540-92673-3 (Online).

<sup>71</sup> Dostupné z: <http://ifomis.uni-saarland.de/bfo/> [cit. 2018-01-30].

<sup>72</sup> Dostupné z: <http://geneontology.org/> [cit. 2018-01-30].

<sup>73</sup> Dostupné z: <https://www.w3.org/Submission/OWL-S/> [cit. 2018-01-30].



## 1.5 Teoretické a filozofické problémy informatických ontologií

*Informace je informace, není to ani hmota, ani energie. Žádný materialismus, který toto nepřipouští, nemůže přetrvat dnešek.*<sup>74</sup>  
Norbert Wiener, 1948

Přes to, co bylo uvedeno o pragmatickém zaměření a relativně samostatném vývoji informatických ontologií, nelze konstrukci informatické ontologie postavit na jiných než na filozofických základech. Jedním z názorných příkladů je metodika *OntoClean* (viz kapitola 2.1.1), která explicitně vychází z filozofických principů.

Nejvýznamnější filozofické problémy informatických ontologií lze zhruba rozdělit do těchto okruhů:

- materialismus vs. idealismus
- různé způsoby reprezentace a z toho vyplývající různé sémantické systémy
- různé způsoby konceptualizace a z toho vyplývající různé kategoriální systémy
- různé způsoby určení typů a vlastností ontologických entit
- předpoklad otevřeného světa a z toho vyplývající logické a sémantické problémy

Je vhodné poznamenat, že toto členění je ryze pracovní a v žádném případě není disjunktní. Ve skutečnosti se okruhy navzájem prolínají a jsou propojeny četnými významnými souvislostmi.

Ke každému z výše uvedených problémů lze zaujmout různé postoje a dospět tak k různým způsobům jejich řešení. Přijatý či zvolený způsob řešení těchto problémů pak představuje nejvyšší úroveň tzv. ontologických závazků pro danou ontologii. Termín ontologické závazky (angl. *ontological commitment*) se používá i ve filozofii, kam jej zavedl Willard Van Orman Quine<sup>75</sup>. V kontextu informatických ontologií se tímto termínem označují předpoklady, pravidla a omezení uplatňovaná při konstrukci ontologie. V negativním slova smyslu se též označují jako *cognitive bias* (kognitivní zkreslení). V užším slova smyslu představují ontologické závazky dohodu uživatelů ontologie o významu používaných znaků (slov, termínů) pro vyjádření pojmů a jejich vztahů. Plní obdobnou úlohu jako integritní omezení a tzv. *business rules* (obchodní pravidla, aplikační logika) v databázových systémech.

### 1.5.1 Materialismus vs. idealismus

V počátcích návrhu jakékoli informatické ontologie je nutnost přiklonit se k některé z variant řešení základní ontologické otázky – co je, tj. rozhodnout, kterému fenoménu univerza diskurzu bude přisuzována ‚reálná‘ existence. Klasické filozofické dilema materialismu a idealismu znovu ožívá při úvahách o způsobu existence pojmů (kategorií), jež jsou základní stavební jednotkou informatických ontologií. Problémy materialismu a idealismu jsou spojeny i s rozhodnutím o podstatě znalosti, informace a digitálních informačních zdrojů.

#### Podstata znalosti a informace

V kapitole 1.2 bylo konstatováno, že významným kontextem informatických ontologií jsou informační a znalostní systémy. To vede k otázce po definici znalosti a informace. Oba pojmy jsou používány v mnoha oblastech i vědních oborech, v jejichž rámci se uplatňují často velmi

---

<sup>74</sup> WIENER, Norbert. *Cybernetics: or control and communication in the animal and the machine*. New York: John Wiley, 1948, s. 132.

<sup>75</sup> Viz PEREGRIN, Jaroslav. *Kapitoly z analytické filosofie*. 2. vyd. Praha: Filosofia, 2014, s. 196. ISBN 978-80-7007-420-6.

rozdílné přístupy k vymezení jejich významu. Z toho vyplývají i problémy při určování jejich vzájemného vztahu. Obecně lze rozlišit dvě rozdílné dimenze chápání a zkoumání znalosti a informace: filozofickou rovinu a rovinu speciálních věd.

Filozofie pohlíží na znalost jako na vědění, či dokonce moudrost. Na tuto dimenzi se zaměřuje teorie poznání (gnozeologie, epistemologie, noetika), psychologie a kognitivní vědy. Klíčovou otázkou je: ‚Co je to znalost / informace?‘ Není překvapením, že filozofie neposkytuje na tuto otázku jednoznačnou odpověď. I často citovaná klasická Platónova definice znalosti z dialogu *Theaitétos* – „*správné mínění spojené s vědomím rozdílu*“<sup>76</sup> je jím samým vzápětí vyvrácena<sup>77</sup> a vybízí k dalšímu zkoumání. Je zajímavé, že v rámci filozofického zkoumání se při definování znalosti vůbec neoperuje s pojmem informace a platí to i opačně. Na informaci se ve filozofii obvykle pohlíží jako na vlastnost hmoty a zkoumání z našeho pohledu blízkých jevů informace a znalost se realizuje odděleně. Ojedinělou výjimkou mezi filozofickými koncepcemi je komplexní filozofie informace Luciana Floridiho, zahrnující i její sémantické a sociální aspekty<sup>78</sup>.

Klíčovou otázkou speciálních věd je: ‚Jak zpracovat / komunikovat / využívat znalost / informaci?‘ Znalost je chápána jako použitelná informace a informace jako sdělitelná znalost. Významnými obory zaměřujícími se na tuto dimenzi jsou informační věda (znalosti a informace ve společnosti), umělá inteligence (znalosti a informace v počítači) a informační / znalostní management (znalosti a informace lidí v podniku). Teorie i praxe inforatických ontologií pracuje s pojetím informací a znalostí založeným na této dimenzi.

Zatímco abstraktní povaha znalosti je všeobecně přijímána, lze konstatovat, že v rámci speciálních věd zkoumajících informaci jsou k dispozici velmi rozdílné názory včetně navzájem neslučitelných koncepcí informace jako hmotné, fyzické věci a informace jako abstraktního, nehmotného, duševního fenoménu. V oblasti systémové vědy a kybernetiky se dokonce lze setkat s názory, že informace představuje nový ontologický fenomén, jak to například vyjadřuje výrok Norberta Wienera v mottu kapitoly 1.5. Obdobně jeden ze zakladatelů systémové vědy Kenneth Boulding prohlašuje, že „*pojmem informace může sloužit k vývoji obecných pojmů strukturovanosti a abstraktní míry organizace, jež nám jako takové poskytují třetí základní dimenzi kromě hmoty a energie*“<sup>79</sup>.

Významným příspěvkem k chápání informace v rámci informační vědy je studie Michaela K. Bucklanda *Information as thing* (Informace jako věc).<sup>80</sup> Jeho pohled, založený na rozboru definic slova informace v *Oxfordském slovníku angličtiny*, nahlíží na informaci ze tří různých úhlů jako na proces, znalost (jež je výsledkem procesu) a věc. Buckland doporučuje věnovat se především zkoumání informace jako věci, jež považuje pro účely informační vědy za nejpřínosnější.

---

<sup>76</sup> PLATÓN. *Theaitétos*. Z řec. orig. (1933) přel. František Novotný. 4., oprav. vyd. Praha: OIKOYMENH, 2007, s. 103 (210). Platónovy dialogy, sv. 11. ISBN 978-80-7298-294-3.

<sup>77</sup> „*Tedy vědění asi není ani vnímání, Theaitéte, ani pravdivé mínění ani výměr připojující se k pravdivému mínění.*“ Tamtéž, s. 103 (210b).

<sup>78</sup> FLORIDI, Luciano. *The philosophy of information*. Oxford: Oxford University Press, 2013. 405 s. ISBN 978-0-19-923239-0.

<sup>79</sup> BOULDING, Kenneth E. General systems theory – the skeleton of science. In: *Management science*. April 1956, 2(3), s. 202.

<sup>80</sup> BUCKLAND, Michael K. Information as thing. In: *Journal of the American Society for Information Science*. June 1991, 42(5), 351-360. ISSN 0002-8231.

Na opačném pólu spektra názorů na podstatu informace je přístup Bertrama C. Brookese<sup>81</sup>, který považuje informace a znalosti za abstraktní fenomény. Jejich vzájemný vztah vyjadřuje následující rovnici:  $K[S] + \Delta I = K[S + \Delta S]$ . Písmeno K v rovnici označuje znalost (angl. *knowledge*), písmeno I informaci a písmeno S strukturu.  $K[S]$  je struktura znalostí,  $\Delta I$  označuje přírůstek informace,  $K[S + \Delta S]$  je struktura znalostí modifikovaná novou informací. Brookes prohlásil tuto rovnici za základní rovnici informační vědy a vyzval k interpretacím a dalšímu rozvíjení této myšlenky. V roce 2008 ale David Bawden s politováním konstatoval, že se tak dosud nestalo a praktické aplikace této rovnice stále chybí.<sup>82</sup>

### Digitální informační zdroje

Inspirující pohledy na ontologickou povahu digitálních informačních zdrojů nabízejí Nicholas Negroponte a David Weinberger. Negroponte odvozuje specifika digitálních zdrojů z odlišnosti jejich konstrukčních prvků. Základním stavebním prvkem analogových zdrojů jsou podle něj atomy, jež jsou hmotné, kdežto digitální zdroje jsou komponovány z abstraktních bitů.<sup>83</sup> Weinberger není v názoru na abstraktní povahu bitů tak kategorický jako Negroponte, nicméně i on upozorňuje na to, že bity a atomy jsou odlišné typy hmoty, jejichž specifika určují i odlišné principy jejich komunikace.<sup>84</sup> Díky svým specifickým vlastnostem se tak digitální objekty v mnoha ohledech spíše podobají objektům nemateriálním, abstraktním.

Problém ‚materiality‘ digitálních informačních zdrojů souvisí s výše uvedenými problémy určení podstaty informace, má ovšem i své vlastní aspekty, týkající se procesů zaznamenání, zpracování/organizování a vnímání/využívání digitálních zdrojů a jejich odlišností oproti zdrojům analogovým.

Zaznamenání digitálních zdrojů nepřináší z technického hlediska výrazná specifika. Je zřejmé, že jejich forma či médium (nosič), použité pro záznam nebo přenos či vysílání, jsou rovněž materiální (fyzické) objekty, stejně jako je to u analogových zdrojů. Není však sporu o tom, že uživatelsky nejvýznamnější složkou informačních zdrojů je jejich abstraktní obsah. A právě v odlišném způsobu záznamu obsahu tkví rozdíl analogových a digitálních zdrojů. Zatímco záznam obsahu analogového zdroje má materiální/fyzickou podstatu (např. molekuly tiskařské černi na papíře), základní stavební jednotkou digitálního záznamu obsahu je bit, tj. abstraktní jednotka množství informace.

Proces zpracování/organizování digitálních zdrojů je odlišný tím, že na rozdíl od atomů jsou bity přímo zpracovatelné počítačem. Dovedeno do důsledků, je-li obsah digitálně zakódován, je ‚srozumitelný‘ počítači. Důrazněji než dříve přichází ke slovu požadavek organizovat skutečně ‚čistý‘ obsah a ne pouze fyzické předměty, na nichž je obsah zaznamenán. Samozřejmě příprava dat umožňujících ‚zhmotnění‘ obsahu je složitější než příprava dat pro zpracování na syntaktické úrovni například v databázových systémech.

---

<sup>81</sup> BROOKES, Bertram C. Foundations of information science: Part I. Philosophical aspects. In: *Journal of information science*. June 1980, 2(3-4), 131. doi:10.1177/016555158000200302. ISSN 0165-5515.

<sup>82</sup> BAWDEN, David. Smoother pebbles and the shoulders of giants: the developing foundations of information science. In: *Journal of information science*. 2008, 34(4), 415-426. doi:10.1177/0165551508089717. ISSN 0165-5515.

<sup>83</sup> NEGROPONTE, Nicholas. Bity a atomy. In: *Digitální svět*. 1. vyd. Praha: Management Press, 2001, s. 15-17. ISBN 80-7261-046-5.

<sup>84</sup> WEINBERGER, David. *Everything is miscellaneous: the power of the new digital disorder*. 1st ed. New York: Times Books, 2007, s. 4-5. ISBN 978-0-8050-8043-8.

Další specifikum představuje samotný vztah obsahu a formy a nutnost jeho reformulace pro digitální zdroje. Digitální technologie stírají rozdíly mezi obsahem a formou, protože bity poskytují jednotnou platformu pro jejich reprezentaci. Navíc na rozdíl od analogových dokumentů, u nichž se rozlišování obsahu a formy řešilo lidskou inteligencí, u digitálních objektů se spoléhá na počítačové aplikace. Zvýrazňuje se tedy potřeba jasně odlišit, co je obsah a co je forma. Analýza obsahu i formy digitálních zdrojů musí být formalizovaná a její výsledky explicitní. Přitom otázka, kde leží hranice mezi (abstraktním) obsahem a (materiální) formou, zdaleka není triviální. To dokládají problémy s aplikací abstraktního modelu *IFLA LRM* na konkrétní bibliografické entity: dílo je definováno jako čistý obsah, vyjádření jako abstraktní forma obsahu, provedení jako konkrétní obsah formy, jednotka jako čistá forma. Rozhodnout, zda to, co popisujeme, je forma obsahu nebo obsah formy, je ovšem velmi obtížné.<sup>85</sup>

Na druhou stranu se dá obrazně konstatovat, že bity umožňují ‚materializovat‘, tj. jednotně reprezentovat hmotné i nehmotné entity včetně abstraktních fenoménů, například množiny. Až v prostředí relačních databází a s pomocí jazyka *SQL* lze doslova ‚vidět‘ množiny a výsledky logických operací s nimi. Tentýž efekt, tentokrát dokonce na sémantické úrovni, je patrný v případech ontologií a jejich jazyků *RDF* a *OWL*, popisujících obsah dokumentů. Pojmy, doposud abstraktní a nesdílitelné, dostávají na webu díky identifikátorům *URI* své konkrétní zhmotnění.<sup>86</sup> Bezeztrátová duplikace a virtualizace způsobuje, že se vztah originál–kopie blíží ekvivalenci a vede k přehodnocení funkce popisu a odkazů, mimo jiné tím stírá rozdíly mezi katalogem a bibliografií.

Co se týče procesu vnímání a využívání analogových a digitálních zdrojů, lze konstatovat jeho symetricky opačné vlastnosti. Zatímco obsah analogových zdrojů je přímo vnímatelný lidskými smysly, u digitálních zdrojů je k vnímání obsahu člověkem zapotřebí nějaká technická pomůcka, pro niž je třeba explicitně definovat struktury obsahu i formy dokumentů. Jak už bylo uvedeno, obsah digitálních zdrojů je přímo využitelný počítačem, zatímco pro využití analogových zdrojů počítačem je zapotřebí definovat dodatečné procedury (například skenování, *OCR* apod.).

### **Problém existence kategorií**

Za základní jednotku obsahu jsou tradičně považovány pojmy neboli kategorie. Jednou ze základních filozofických otázek je rozhodnutí, zda kategorie existují objektivně, a priori, či zda jsou subjektivními konstrukcemi. Ve středověku se pokusy o řešení této otázky označovaly jako spor o univerzálie. Termínem univerzálie byly ve středověké filozofii označovány obecné pojmy (obecniny) či kategorie, jejichž protikladem jsou tzv. partikulárie (jednotliviny). Kromě otázky objektivní existence univerzálií je další významnou otázkou diskutovanou rovněž už ve středověké filozofii vztah pojmů a věcí. Krajními póly řešení této otázky jsou realismus a nominalismus. Realismus tvrdí, že univerzálie, eventuálně univerzální charakteristiky jednotlivých věcí, objektivně existují (tj. jsou to věci). Pro realisty univerzálie existují ve věcech (*universalia sunt in re*) nebo dokonce před věcmi (*universalia sunt ante rem*). Nominalismus považuje univerzálie za slova, jež pouze odkazují na věci (*universalia sunt nomina post rem*). Někdy se ještě vyčleňuje tzv. konceptualismus, jehož zastánci rovněž tvrdí, že univerzálie existují po věcech (*universalia sunt post rem*), ovšem nikoli jako slova, ale pouze jako myšlenky o věcech

---

<sup>85</sup> Podrobně byla tato problematika diskutována v článku: BRATKOVÁ, Eva, KUČEROVÁ, Helena. K otázkám metadatového popisu systémů organizace znalostí. In: *Knihovna*. 2015, 26(1), 5-36. ISSN 1801-3252 (Print). ISSN 1802-8772 (Online).

<sup>86</sup> Například pojem ‚ontologie‘ v Polytematickém strukturovaném hesláři má *URI* <http://psh.techlib.cz/skos/PSH2611>.

(tj. pojmy). Středověký spor o univerzálie dnes ožívá v oblasti ontologického inženýrství v souvislosti s digitálními zdroji, jejichž prostřednictvím jsou vyjadřovány ontologické pojmy.

Problematiku realismu, konceptualismu a nominalismu v již zmíněném článku *Jiný pohled na data*<sup>87</sup> řeší i George H. Mealy. Ten v úvodu článku upozorňuje, že jeho teorie je založena na „četných starých myšlenkách“ (míní tím především myšlenky filozofické a sémiotické), jež ovšem nejsou v informatice běžně uplatňovány. Připomíná, že na filozofické úrovni se obvykle uvažují tři oblasti zpracování dat: reálný svět, myšlenky lidí o světě a symboly zaznamenané na nějakém záznamovém médiu, přičemž myšlenky (ve středověké terminologii univerzálie) se považují za modely reality a symboly za modely myšlenek. Konstatuje, že na to, která z těchto oblastí „skutečně existuje“ v ontologickém smyslu, existují různé filozofické názory. Mealy se přiklání k nominalistické teorii; považuje data za reálně existující symboly, jež nazývá „fragmenty teorie o reálném světě“.

### 1.5.2 Různé přístupy k chápání/řešení reprezentace

Termínem reprezentace se v sémiotice označuje takový vztah mezi znakem a realitou, kdy skutečnost je reprezentována (zastupována) znakem. Z tohoto pohledu lze reprezentaci znalostí, již zmíněnou v kapitole 1.1, považovat za specifický typ reprezentace. Různé způsoby chápání a řešení vztahu mezi znakem a realitou vedly k ustavení různých sémiotických systémů, jež si často protiřečí. Příčinou jsou rozdílná filozofická východiska jednotlivých autorů, resp. jejich ontologické a gnozeologické chápání světa (materialismus, idealismus, pragmatismus, pozitivismus, fenomenologie...). Názorným příkladem jsou vzájemně neslučitelné sémiotické koncepce Ferdinanda de Saussura (1857–1913) a Charlese Sanderse Peirce (1839–1914).

Pro znázornění funkčních vztahů mezi znakem a realitou na nejvyšší úrovni abstrakce je v různých modifikacích<sup>88</sup> tradičně používán tzv. sémiotický trojúhelník zvaný též trojúhelník reference. Ukazuje využití znaku k vyjádření smyslu a k označení reality. Zřejmě nejčastěji interpretovanou variantou je sémiotický trojúhelník publikovaný v roce 1923 Charlesem K. Ogdenem a Ivorem A. Richardsem<sup>89</sup>, i když s obdobnou triádou se lze setkat už v dříve publikovaných pracích, kupříkladu u Gottloba Fregeho (1848–1925)<sup>90</sup> a Charlese S. Peirce<sup>91</sup>. Praktickou využitelnost tohoto modelu v každodenní praxi dokládá i jeho využití ve studii *IFLA* věnované funkčním požadavkům na věcné autority (*FRSAD*)<sup>92</sup>: entity Dílo – Téma – Název (*Work – Thema – Nomen*) jsou situovány na odpovídajících vrcholech sémiotického trojúhelníku.

---

<sup>87</sup> MEALY, George H. Another look at data. In: *Proceedings of the Fall Joint Computer Conference (AFIPS Fall '67), November 14-16, 1967*. New York: ACM, 1967, 525-534. doi:10.1145/1465611.1465682.

<sup>88</sup> Podrobněji viz např. ČERMÁK, František. *Jazyk a jazykověda: přehled a slovníky*. 2. dotisk 3. dopl. vyd. Praha: Karolinum, 2001, 2004, 2007. Kapitola 1.52, Struktura znaku a jeho vztahy, s. 24-28. ISBN 987-80-246-0154-0.

<sup>89</sup> OGDEN, Charles Kay, RICHARDS, Ivor Armstrong. *The meaning of meaning: a study of the influence of language upon thought and of the science of symbolism*. New York: Harcourt Brace & Company, 1923. 363 s.

<sup>90</sup> FREGE, Gottlob. Über Sinn und Bedeutung. In: *Zeitschrift für Philosophie und philosophische Kritik*. 1892, NF 100, s. 25-50.

<sup>91</sup> PEIRCE, Charles Sanders. *Collected papers of Charles Sanders Peirce*. Cambridge, Harvard University Press, 1931-1958. 8 sv.

<sup>92</sup> IFLA. IFLA Working Group on the Functional Requirements for Subject Authority Records. *Functional requirements for subject authority data (FRSAD): a conceptual model*. Marcia Lei ZENG, Maja ŽUMER, Athena SALABA, ed. Berlin: De Gruyter Saur, © 2011. vi, 74 s. IFLA series on bibliographic control, vol. 43, ISSN 1868-8438. ISBN 978-3-11-025323-8 (Print). ISBN 978-3-11-026378-7 (Online).

Klíčové problémy reprezentace spočívají v tom, jakou interpretaci sémiotického trojúhelníku zvolit a jak řešit případy vícenásobných vztahů reprezentace mezi pojmem a jeho označením.

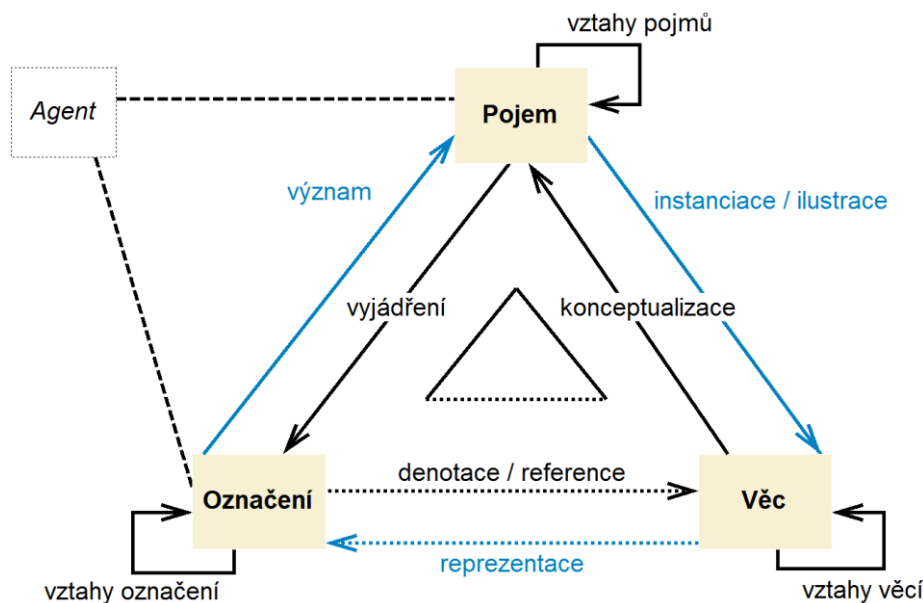
I když je triadický princip sémiotického trojúhelníku s kořeny v Aristotelově filozofii všeobecně přijímán, názory zastánců různých směrů v lingvistice a sémiotice na vymezení klíčových komponent sémiotického trojúhelníku a jejich vztahů nejsou jednotné. Různorodost přístupů se projevuje i v rozdílných terminologických systémech. V tabulce 1 jsou pro názornost uvedeny termíny, jež používáme v této práci, spolu s nejfrekventovanějšími významově podobnými termíny používanými v různých sémiotických systémech.

Věc		Pojem		Označení	
denotát (třída)		deznát		deskriptor	
entita		interpretans ( <i>Peirce</i> )		forma	
individuum		myšlenka		jazyk	
instance		označované ( <i>Saussure</i> )		lexikální jednotka, lexém	
objekt ( <i>Peirce</i> )		poznatek		označující ( <i>Saussure</i> )	
objektivní realita		reference		representamen ( <i>Peirce</i> )	
referent ( <i>individuum</i> )		referendum		symbol	
skutečnost		smysl		termín	
		třída		vehikulum	
		význam		výraz	
		znalost		znak	
Vztah					
Věc ↓ Pojem	Pojem ↓ Věc	Označení ↓ Pojem	Pojem ↓ Označení	Označení ↓ Věc	Věc ↓ Označení
konceptualizace	exemplifikace instanciace ilustrace mínit něco něčím	deznace konotace označení referenční sémantika signifikace symbolizace význam znamenání	relační sémantika specifikace vyjádření	denotace ( <i>třída</i> ) reference ( <i>individuum</i> )	reprezentace zastupování

Tab. 1 Terminologie sémiotiky

Sémiotický trojúhelník byl původně koncipován jako nástroj pochopení znaku. Informatické ontologie ho ovšem využívají jako kognitivní nástroj, umožňující pochopení vztahu reprezentace a konceptualizace reality.<sup>93</sup> Na obrázku 5 představujeme vlastní pracovní interpretaci sémiotického trojúhelníku, tak jak je aplikována v této práci.

<sup>93</sup> Viz například SMITH, Barry. Beyond concepts: ontology as reality representation. In: Achille C. VARZI, Laure VIEU, ed. *Formal ontology in information systems: proceedings of the Third International Conference (FOIS 2004)*: [International Conference on Formal Ontology and Information Systems, Turin, November 4-6, 2004]. Amsterdam: IOS Press, © 2004, s. 73-84. *Frontiers in artificial intelligence and applications*, vol. 114, ISSN 0922-6389. ISBN 978-158603-468-9.



Obr. 5 Sémiotický trojúhelník

V našem znázornění sémiotického trojúhelníku na obrázku 5 je cokoli, co je součástí reality, zastoupeno pracovním termínem ‚Věc‘. Bylo by možné zvolit i jiné označení ‚všeho‘, například v informační vědě i v databázovém světě populární termín entita, případně objekt. Pro termín ‚Věc‘ jsme se rozhodli z důvodu jeho běžného používání ve filozofických textech a rovněž v informatických ontologiích. Například na vrcholu hierarchie slovníku ontologického jazyka *OWL* je ‚*owl:Thing*‘, tj. věc; referenční model *IFLA LRM* používá jako vrcholovou entitu ‚*Res*‘, tj. latinsky věc. Slovo věc tedy nepoužíváme ve smyslu ‚materiální objekt‘, zastupuje pro naše účely i abstraktní entity a události, včetně informací a znalostí. V zájmu souladu s pojetím uplatněným v terminologických systémech a v systémech organizace znalostí jsou zbývající dva vrcholy trojúhelníka obsazeny entitami ‚Pojem‘ a ‚Označení‘.

Vztahy entit na vrcholech sémiotického trojúhelníku jsou znázorněny šipkami. Jak je z obrázku 5 zřejmé, jsou zastoupeny jak vztahy mezi různými typy entit, tak i vzájemné vztahy mezi entitami stejného typu. První skupinu tvoří binární asymetrické obousměrné vztahy věcí a pojmů, pojmů a označení a označení a věcí. Ve směru Věc → Pojem → Označení naznačeném šipkami na obrázku 5 sémiotický trojúhelník ukazuje cestu od skutečnosti k myšlení (pochopení) a od myšlenek k jejich vyjádření. V obráceném směru Označení → Pojem → Věc lze sledovat, jak jazyk zastupuje pojmy a jak pojmy zastupují realitu. Do druhé skupiny vztahů patří rekurzivní vzájemné vztahy pojmů, označení a věcí.

Sémantický vztah reprezentace mezi znakem a realitou, tj. ‚Označení <označuje> Věc‘ je na obrázku 5 znázorněn prostřednictvím tečkované čáry jako vztah nepřímý, zprostředkovaný konceptualizací a vyjádřením pojmů pomocí nějaké znakové soustavy. Konceptualizace (blíže viz kapitola 1.5.3) je pojmový (myšlenkový) odraz (zpracování) představy o významu skutečnosti. Pojem je určený intenzí a extenzí. Intenze je smysl, význam pojmu, nejčastěji stanovený prostřednictvím množiny charakteristik nebo definicí. Extenze je množina věcí, které pojem zahrnuje. Pojem jako takový je implicitní (nesdělitelný, tacitní), pro jeho komunikaci je nutné explicitní vyjádření pojmu označením. Vyjádření označuje situaci, kdy pojem je zastoupen znakem (tj. označen), nejčastěji ve formě nějakého jazykového výrazu. Význam je pojmová interpretace obsahu znaku. Instanciacie/ilustrace vyjadřuje uskutečnění významu pojmu nebo jeho doložení příkladem. Denotace/reference je vztah, v němž znak označuje skutečnost,

vypovídá (referuje) o realitě.

Ke třem faktorům na vrcholech sémiotického trojúhelníku se ještě někdy přidává čtvrtý – subjekt nebo agent, tj. ten, kdo pojmy tvoří ve své mysli či kognitivním systému (tím vyjadřujeme možnost, že v roli agenta může vystupovat nejen člověk, ale i stroj) a vyjadřuje je nějakými znaky.<sup>94</sup> Agent vnáší do sémiotického trojúhelníku subjektivní dimenzi, tj. jedinečný přístup každého individua k chápání významu, a dimenzi pragmatickou, tj. porovnání významu s tím, k čemu ho chce agent využít.<sup>95</sup> Pragmatický aspekt vztahu agenta k jím nahlížené realitě vyjadřuje vztah relevance mezi jeho kognitivní potřebou a mezi významem poznávané věci, tedy její použitelnost pro agenta. Vztah interakce mezi agentem a označením a agentem a pojmem znázorňuje na obrázku 5 přerušovaná čára.

Nejjednodušší implementací sémiotického trojúhelníku je jednoznačná lineární posloupnost vrcholů trojúhelníku se vztahy typu 1 : 1, v níž jedné věci odpovídá právě jeden význam (pojem), jenž je vyjádřen právě jedním označením.<sup>96</sup> Ve skutečnosti se může jednat až o vztahy typu N : M, tj. více-více (více v tomto smyslu zastupuje interval  $0-\infty$ , tedy je možné, že existují i znaky bez významu a znaky nebo pojmy bez denotátů). V přirozeném jazyce jsou nejčastější případy komplikovanější reprezentace, založené na vícenásobném vztahu mezi pojmem a označením: synonymie (více označení pro jeden pojem), polysémie a homonymie (jedno označení pro více pojmů), složená témata (složený pojem – celek, který je rozložitelný na více dílčích pojmů, například ‚biomedicína‘ nebo ‚biomedicínská informatika‘) a slovní spojení (složené označení jednoho pojmu, například ‚umělá inteligence‘).

Aby bylo možné tyto komplikované vztahy zpracovávat v počítačovém prostředí, je nutné jejich explicitní zjednoznačnění. Obvyklým nástrojem explicitního zjednoznačnění je formalizovaný či přímo umělý jazyk. V současné praxi jsou nejobvyklejšími nástroji tohoto typu řízené slovníky, notace a identifikátory. Další možností, jež ovšem už vyžaduje určitý podíl inteligence, je odvození jednoznačného významu z kolokace, tj. z kontextu dané entity (například jsou-li v textu v blízkosti slova ‚zámek‘ slova ‚šlechta‘ a ‚sídlo‘, lze předpokládat, že zámek označuje budovu, a nikoli součást dveří).

### 1.5.3 Různé přístupy k chápání/řešení konceptualizace

V této části se budeme věnovat procesu, který je v sémiotickém trojúhelníku na obrázku 5 v předchozí kapitole vyznačen jako vztah mezi věcí a pojmem. Konceptualizace neboli tvorba kognitivního pojmového odrazu skutečnosti je nutně spojena s abstrakcí, tj. s výběrem entit a charakteristik, jež jsou pro daný účel považovány za podstatné. Po tomto kroku následuje seskupování ekvivalentních entit do kategorií (tříd). V průběhu procesu konceptualizace je zapotřebí řešit následující problémy: volba kritérií ekvivalence, existence různých

---

<sup>94</sup> Například Ladislav Tondl používá pro subjekt označení „pozorovatel“, případně „uživatel jazyka“. TONDL, Ladislav. *Problémy sémantiky*. Praha: Karolinum, 2006, s. 30-31. ISBN 978-80-246-1075-7.

<sup>95</sup> Podrobněji viz např. HUDÁKOVÁ, Miriam. Nevyhnutelnost subjektivity při pořádání informací a znalostí. In: *Ikaros* [online]. 2006, **10**(3) [cit. 2018-01-30]. urn:nbn:cz-ik3204. ISSN 1212-5075. Dostupné z: <http://www.ikaros.cz/node/3204>.

<sup>96</sup> Poznámka: V opačném směru by však bylo přesnější znázornit vztah pojem–věc jako vícenásobný, protože jeden pojem může zpravidla vystihnout smysl více denotátů (entit, přesněji řečeno jejich instancí či individuí – objektů). Tento vztah se označuje jako abstrakce, zobecnění. Umožňuje označit více konkrétních případů (například všechny pondělky) jedním pojmem. Takovéto chápání pojmu je blízké způsobu, jak jsou definovány kategorie – pojem je intenzí kategorie, zatímco jednotlivé entity zařazené do kategorie jsou její extenzí.



kategoriálních systémů, konceptualizace obsahu zdrojů (*aboutness*).

Rozhodnutí, které věci jsou či nejsou ekvivalentní, není triviální. Nabízí se více možností od teoreticky racionálně podložených po zcela pragmatické a subjektivní. Situaci komplikuje navíc skutečnost, že na určení ekvivalence má vliv kontext, v jehož rámci je uvažována: kupříkladu v určitém kontextu se ruka a noha mohou jevit jako ekvivalentní, protože obě jsou končetiny, v ortopedické praxi naopak vyniknou jejich rozdíly.

Lze rozlišovat mezi přístupem ke konceptualizaci v lidské mysli, jež je součástí subjektivního poznání, a mezi konceptualizací spadající do sféry objektivních zaznamenaných znalostí, relativně na subjektu nezávislých. V této části vycházíme při charakteristice konceptualizace v lidské mysli z přehledu přístupů kognitivní vědy, jež zpracoval George Lakoff<sup>97</sup>. Shrnutí principů „umělé“ konceptualizace, resp. především jejich podmnožiny založené na společných attributech, je zpracováno v souladu s jejich prezentací v mezinárodních normách ISO 704, ISO 1087-1 a ISO 25964 (charakteristiky těchto norem jsou podány v kapitole 4.5.2).

### **Konceptualizace jako součást lidského myšlení**

V rámci kognitivní vědy se konceptualizace/kategorizace zkoumá jako přirozená součást a nástroj lidského myšlení. Kategorie je považována za kognitivní a lingvistický model umožňující využít existující znalosti v nových situacích a rozpoznat instance nebo události (hovoří se o kognitivní ekonomii kategorizace). George Lakoff ve svém historickém přehledu zkoumání kategorizace člení filozofický přístup ke zkoumání poznání a myšlení na dva hlavní proudy – „tradiční“ a „nový“. První, tzv. tradiční či abstraktní proud, který Lakoff nazývá objektivismus, tvrdí, že myšlení je abstraktní a nezávislé na těle. Představitelé tohoto proudu chápou konceptualizaci v lidské mysli jako objektivně probíhající proces, jehož pravidla jsou v zásadě totožná s tím, co bude v následující části představeno jako systémový/obsahový přístup. Druhý, tzv. interakcionistický proud, tvrdí, že myšlení je závislé na těle a že je imaginativní. Sám Lakoff je zastáncem druhého směru a svůj přístup nazývá zkušenostním realismem (*experientialismem*).<sup>98</sup>

*„Podle tradičního přístupu je rozum abstraktní a nezávislý na těle. Podle nového přístupu má rozum tělesnou podstatu. Tradiční přístupy popisují strukturu rozumu jako soubor doslovných tvrzení, především propozic, které jsou objektivně buď pravdivé, nebo nepravdivé. Nový přístup považuje imaginativní aspekty rozumu – jako jsou metafora, metonymie a mentální obraznost – spíše za centrální než periferní nebo bezvýznamné přívazky k aspektům doslovným.“<sup>99</sup>*

Z Lakoffova přehledu výsledků empirických výzkumů lidského myšlení, prováděných v 20. století v rámci psychologie a nově se formující kognitivní vědy, se jako potenciálně užitečné jeví zejména rodinné podobnosti, centralita a prototypy, odstupňované členství v kategorii, bázevová kategorizace a používání metonymie.

Princip rodinných podobností takto pojmenoval Ludwig Wittgenstein, Lakoff ho interpretuje jako vzájemný vztah příbuznosti členů kategorie, daný sérií řetězců rodinných podobností, aniž by existovala množina jejich společných vlastností. Princip centrality a prototypů spočívá v tvrzení, že některé členy jsou lepšími reprezentanty kategorie než jiné (například vlaštovka

---

<sup>97</sup> LAKOFF, George. *Ženy, oheň a nebezpečné věci: co kategorie vypovídají o naší mysli*. 1. vyd. Praha: Triáda, 2006. 655 s. ISBN 978-80-86138-78-7.

<sup>98</sup> Tamtéž, s. 11-15.

<sup>99</sup> Tamtéž, s. 11.

bude zřejmě považována za lepšího reprezentanta kategorie ‚ptáci‘ než pštros). Odstupňované členství v kategorii škáluje příslušnost k třídě ekvivalence do více stupňů. Při bázev kategorizaci se uplatňuje tzv. postup *middle-out*, od středu ven: východiskem (bází) nejsou ani členy s nejvyšší mírou konkrétnosti, ani členy na nejvyšší úrovni abstrakce, nýbrž nejdůležitější zástupci, od nichž se postupuje jak směrem vzhůru, tak směrem dolů. S používáním metonymie způsobem *pars pro toto* (pojmenováním celku jeho částí) se lze setkat například při označování všech kopírek jako xerox, což byla původně obchodní značka pouze jednoho z typů kopírovacích strojů.

### **Konceptualizace jako objektivní proces ontologického inženýrství**

V rámci ‚umělé‘ konceptualizace lze rozlišit tři typy, založené na různých způsobech určování ekvivalence: systémový/obsahový přístup, formální přístup a ‚nesystémový‘, subjektivní přístup.<sup>100</sup>

Přístup ke konceptualizaci, který pracovní nazýváme ‚systémový‘, hraje klíčovou roli v ontologickém inženýrství. Systémový přístup se stal především díky svému využití v oblasti informačních technologií natolik všudypřítomným, až se stal takřka ‚neviditelným‘ pozadím či základnou specifických metod. Jeho integrace do teoretických i praktických aktivit souvisejících s životním cyklem informací a znalostí je podmínkou i stimulem jejich dramatického rozvoje v posledním půlstoletí. S vědomím toho, že jde o obecně známé a přijímané skutečnosti, nejprve stručně připomeneme základní pojmy a principy systémového přístupu, systémové vědy a systémové analýzy.

Připomeňme, že je třeba rozlišovat mezi systémovou analýzou a systémovým přístupem. Systémový přístup je způsobem myšlení, jež chápe předmět svého zájmu jako komplexní celek s vnitřními i vnějšími vztahy. Nemá své vlastní specifické metody ani formální aparát. Obojí je záležitostí až systémové vědy (teorie systémů) jako standardního vědního oboru se specifickým předmětem zkoumání (tj. systémy) a specifickými metodami, jejichž jádrem je systémová analýza.

Ústředním pojmem systémové analýzy je systém. Autoři významné české monografie *Systémová analýza a syntéza* Jaroslav Habr a Jaromír Vepřek zformulovali definici, jež výstižně zachycuje všechny podstatné rysy systému:

*„Za systém považujeme složitý reálný nebo abstraktní objekt, v němž rozlišujeme části, vztahy mezi nimi, vlastnosti. Vůči okolí vystupuje systém jako celek. Části systému jsou ve vzájemné interakci a interagují i se systémem jako celkem. Označujeme je jako prvky systému a vztahy mezi nimi nazýváme vazbami systému. Pro to, abychom považovali složitý reálný objekt za systém, je rozhodující náš přístup k tomuto objektu, způsob jeho pojetí, způsob práce s ním, nikoli jeho věcná povaha.“<sup>101</sup>*

Tento arbitrární charakter systému podtrhuje jeden ze zakladatelů systémové vědy Brian Gaines výrokem „systém je to, co je za systém považováno.“<sup>102</sup>

---

<sup>100</sup> Poznámka: Tyto názvy byly zformulovány jako ryze pracovní pouze pro účely tohoto textu, nejsou tedy součástí odborné terminologie.

<sup>101</sup> HABR, Jaroslav, VEPŘEK, Jaromír. *Systémová analýza a syntéza: zdokonalování a projektování systémů*. 2. přeprac. vyd. Praha: SNTL, 1986, s. 25.

<sup>102</sup> GAINES, Brian R. General systems research: quo vadis? In: *General systems: yearbook of the Society for General Systems Research*. 1979, vol. 24, s. 1. Dostupné z: <http://pages.cpsc.ucalgary.ca/~gaines/reports/> [cit. 2018-01-30]

System je chápán jako jednota struktury a funkce, jež má definován svůj cíl. Strukturou se míní víceméně stálý způsob uspořádání prvků systému, tj. množina prvků a jejich vzájemných vztahů. Prvek systému „lze chápat jako primitivní pojem, nebo jej pro naše účely můžeme charakterizovat jako při dané rozlišovací úrovni dále nedělitelnou část celku“<sup>103</sup>. Funkce systému představují procesy, jež jsou opět spjaty vzájemnými vztahy. Základními systémovými procesy jsou proces zpracování, který transformuje vstupy na výstupy, a proces řízení, jenž upravuje chování systému s ohledem na definované cíle. Slovo ‚jednota‘ z definice systému upozorňuje na to, že prvky ani procesy nemohou v systému existovat izolovaně, ale vzájemně na sobě závisejí. Aby mohl proběhnout proces, potřebuje prvky, na nichž se uskuteční. Prvky systému se zase dynamicky mění v čase a tuto dynamiku realizují právě procesy. Jakkoli jsou ovšem v realitě všechny zmíněné komponenty pevně propojeny vzájemnými vztahy a fyzicky neoddělitelné, díky abstraktní povaze systémové analýzy, jež je analýzou myšlenkovou, si můžeme dovolit uvažovat jednotlivé entity odděleně.

Systémový přístup ke konceptualizaci je založený na ekvivalenci obsahu, tj. na intenzionálním definování. V lingvistice se pro tento typ ekvivalence tradičně používá označení synonymie. Podle systémových komponent zastoupených v kritériu ekvivalence lze rozlišit několik typů kritérií obsahové ekvivalence: struktura, funkce, vztahy, prvky, vlastnosti, hodnoty.

Ve všech případech systémového přístupu je ekvivalence určena zevnitř, obsahově, analýzou struktury (tím, co jsme zjistili, ‚co to je‘, angl. *isness*). Tento přístup předpokládá intelektuální aktivitu, vyžadující obvykle lidskou inteligenci. Aby bylo možné zařadit určitý objekt do třídy ekvivalence, je nutné ho důkladně analyzovat. Intelligence je zapotřebí jednak pro zjištění charakteristik objektů, jichž může být nekonečný počet, ale zejména pro výběr těch vlastností, které se projeví jako užitečné pro daný účel v daném kontextu.

Nejvýznamnějším typem konceptualizace, používaným tradičně v praxi ontologického inženýrství, je konceptualizace založená na stejných vlastnostech (atributech), případně vyžadující nejen stejné atributy, ale i jejich hodnoty. Tento postup, jakkoli vzdálený ‚přirozenému‘ lidskému uvažování, se stal de facto standardem v návrhu struktur informačních systémů a softwarových aplikací.

Formální ekvivalence se určuje zvenčí, podle ‚povrchu‘, formy (tím, co vidíme, ‚jak to vypadá‘). V lingvistice se pro tento typ ekvivalence tradičně používá označení homonymie. Na rozdíl od systémového přístupu, který zkoumá vnitřní vlastnosti, se zájem soustředí na metadata přidělená zkoumaným objektům, tj. na vnější vlastnosti. Kritériem ekvivalence je zpravidla stejné označení (názvy, metadata) prvků, funkcí, vlastností, vztahů. Tento postup lze automatizovat a svěřit tak konceptualizaci počítači.

Třetím způsobem je subjektivní, arbitrární určování ekvivalence, přizpůsobené individuálním potřebám toho, kdo konceptualizaci realizuje.<sup>104</sup> Tento typ ekvivalence je obvykle vyjadřován extenzionálně, tj. vyjmenováním (enumerací) ekvivalentních prvků. Do této skupiny patří i tzv. kanonicky určená ekvivalence a ekvivalence zdůvodněná různými pragmatickými důvody, kdy jsou ekvivalentní věci sdružovány na základě společného cíle či účelu. Zcela arbitrárním způsobem je určení kategorií založené na náhodě, například stanovení skupin soutěžících týmů

---

<sup>103</sup> HABR, Jaroslav, VEPŘEK, Jaromír. *Systémová analýza a syntéza: zdokonalování a projektování systémů*. 2. přeprac. vyd. Praha: SNTL, 1986, s. 26-27.

<sup>104</sup> Poznámka: Jako subjektivní lze samozřejmě označit i kritéria ekvivalence pro účely kategorizace v lidské mysli, již zkoumá kognitivní věda.

ve sportovních soutěžích losováním.

### **Konceptualizace obsahu informačních zdrojů (indexace, *aboutness*)<sup>105</sup>**

Obecný princip ekvivalence, platící pro konceptualizaci věcí, je platný samozřejmě i pro konceptualizaci jejich specifických případů, jimiž jsou zaznamenané znalosti. Zde však nastává problém. Z hlediska organizace se totiž informační zdroje liší od ostatních věcí tím, že při jejich organizaci se málokdy používá kritérium ekvivalence založené na jejich podstatných vnitřních vlastnostech. Spíše než o to, „co jsou“ informační zdroje (ontologická rovina – angl. *isness*, např. dokument, je‘ rukopis, dokument, je‘ počítačový soubor ve formátu MP3), se zajímáme o to, „o čem jsou“ (gnozeologická rovina – angl. *aboutness* např. dokument, je o‘ rukopisech, dokument, je o‘ formátu MP3). To má své důsledky při určování typu sémantických vztahů. Kupříkladu o konkrétním medvědu v pražské zoo lze prohlásit, že představuje extenzi pojmu medvěd. Řečeno termíny sémiotického trojúhelníku (viz kapitola 1.5.3), medvěd v zoo je instancí či spíše ilustrací třídy ‚medvěd‘. Jiná situace ovšem nastane v případě knihy o témže medvědu. Mezi pojmem medvěd a knihou o medvědu bezesporu nepanuje vztah extenze, kniha o medvědu v pražské zoo není instancí medvěda. Vzájemný vztah díla a pojmu, jenž vyjadřuje jeho obsah, je vztahem asociace. Proces asociování zdrojů a témat se obvykle označuje termínem indexace.

Řešení problému *aboutness* se tradičně věnují odborníci v oblasti paměťových a fondových institucí a nově se přidávají i zástupci komunity sémantického webu. Metadatové schéma *Dublin Core* nabízí pro vyjádření *aboutness* prvek popisu ‚dc:subject‘. Toto jednoduché řešení chápe *aboutness* jako vlastnost popisovaného zdroje a z hlediska filozofické perspektivy se kloní k realismu (viz kapitola 1.5.1). Konsorcium *W3C* vydalo v souvislosti se sémantickým anotováním zdrojů a s využitím třídy v roli hodnoty vlastnosti doporučení s návrhem pěti variant řešení, jak reprezentovat obsah zdrojů s využitím vlastnosti ‚dc:subject‘ v jazycích *RDFS* a *OWL*.<sup>106</sup>

Postupný vývoj řešení problému *aboutness* lze sledovat od 90. let 20. století na půdě mezinárodní organizace *IFLA* v dokumentaci iniciativy zaměřené na návrh pojmového modelu bibliografického univerza: *FRBR<sub>ER</sub>*<sup>107</sup>, *FRSAD*<sup>108</sup>, *IFLA LRM*<sup>109</sup>. Příspěvkem k tomuto tématu je článek Jonathana Furnera z roku 2012 věnovaný ontologii předmětu díla.<sup>110</sup> Článek teoreticky

---

<sup>105</sup> Podrobně je tato problematika diskutována v článku: KUČEROVÁ, Helena. Co analyzujeme při obsahové analýze dokumentů? K pojmu *aboutness* v organizaci znalostí. In: *Knihovna: knihovnická revue*. 2014, 25(1), 36-54. ISSN 1801-3252 (Print). ISSN 1802-8772 (Online).

<sup>106</sup> NOY, Natalya F., ed. *Representing classes as property values on the semantic web* [online]. W3C Working group note 05 April 2005. W3C, 2005 [cit. 2018-01-30]. Dostupné z: <https://www.w3.org/TR/swbp-classes-as-values/>.

<sup>107</sup> IFLA. *Funkční požadavky na bibliografické záznamy: závěrečná zpráva*. Praha: Národní knihovna ČR, 2002. 116 s. ISBN 80-7050-400-5.

<sup>108</sup> IFLA. *Functional requirements for subject authority data (FRSAD): a conceptual model*. Marcia Lei Zeng, Maja Žumer, Athena Salaba, ed. IFLA Working Group on the Functional Requirements for Subject Authority Records (FRSAR), 2011. vi, 74 s. Dostupné z: <http://www.ifla.org/node/5849> [cit. 2018-01-30].

<sup>109</sup> IFLA. *IFLA library reference model: a conceptual model for bibliographic information* [online]. Pat Riva, Patrick LeBoeuf, Maja Žumer, ed. Hague: International Federation of Library Associations and Institutions, rev. August 2017 as amended and corrected through December 2017 [cit. 2018-01-30]. 101 s. Dostupné z: <https://www.ifla.org/publications/node/11412>.

<sup>110</sup> FURNER, Jonathan. *FRSAD and the ontology of subjects of works*. In: *Cataloging & classification quarterly*. 2012, 50(5-7), 494-516. ISSN 0163-9374. doi:10.1080/01639374.2012.681252.

zdůvodňuje přístup pracovní skupiny *FRSAD* k tvorbě pojmového modelu věcných autorit. Furner rozlišuje *aboutness* a tzv. *subjecthood* (podle *Oxfordského slovníku angličtiny* „stav nebo podmínka bytí předmětem“<sup>111</sup>) a zdůrazňuje, že toto rozlišení není čistě teoretické, ale má vliv například na návrh knihovních katalogů. *Aboutness* odpovídá na otázku, jaká je logická povaha vztahu mezi výrokem/dokumentem a věcí, o níž tvrdíme, že výrok/dokument je. *Subjecthood* odpovídá na otázku, jaká je ontologická povaha věci, o níž tvrdíme, že výrok/dokument je. *Aboutness* je podle Furnera vztah k předmětu díla, *subjecthood* je předmět sám. Furner charakterizuje různá pojetí *aboutness* a vymezuje dva typy, pro jejichž označení používá filozofické termíny realismus a nominalismus. Realismus podle něj považuje *aboutness* za objektivně existující atribut díla, nominalismus tvrdí, že *aboutness* je vztah mezi dílem a tématem. Furner konstatuje, že jakkoli se v odborné literatuře dají jen obtížně nalézt obhajoby realistického pohledu, praxe bibliografických systémů klasifikace a indexace dokumentů je právě na tomto přístupu založena. V zájmu toho, aby se i v praxi prosadil nominalistický model, navrhuje uvažovat o předmětech analogicky jako o dílech v modelu *FRBR*, tj. jako o produktu tvůrčí aktivity. „*Stejně jako jsou díla realizována prostřednictvím vyjádření, ztělesněna v provedení a ilustrována jednotkami, je rozumné předpokládat, že předměty (obsahy) jsou realizovány prostřednictvím pojmů, jež jsou ztělesněny v termínech (term-type) a ilustrovány konkrétním výskytem termínů (term-tokens)?*“<sup>112</sup> Použitý tázací způsob naznačuje, že spíše než o definitivní závěr šlo autorovi o výzvu k diskusi.

### **Různé kategoriální systémy (systémy základních kategorií)**

Významnou roli ve filozofii, speciálních disciplínách i v infromatických ontologiích hrají tzv. základní kategorie, tj. kategorie nejvyšší úrovně, zahrnující ve svém souhrnu veškeré univerzum poznání. Vymezení takových kategorií patří k základním filozofickým otázkám.

V průběhu historie se o vymezení základních kategorií v rámci filozofických ontologií snažili mnozí myslitelé. Na evropskou a obecně západní filozofickou tradici významně zapůsobil Aristotelův ontologický systém v interpretaci Porfyria z Tyru. Aristotelův systém zahrnuje deset kategorií: podstata (substance, věc – jediná samostatná kategorie, všechny další kategorie jsou považovány za tzv. případy, akcidenty), kvantita, kvalita, vztah, místo, čas, poloha, mít (vlastnictví, stav), činnost, trpnost.<sup>113</sup> Tento kategoriální systém spolu s přesvědčením, že příslušnost do kategorie je založena na společných vlastnostech, byl bezmála po dvě tisíciletí považován za standard a stal se základem četných vědních klasifikací a systémů organizace znalostí.

Jedním z příkladů aplikace aristotelského systému je šestice kategorií označovaná jako *5W1H*. Zkratka je vytvořena z prvních písmen anglických tázacích zájmen *who* – kdo, *what* – co, *where* – kde, *when* – kdy, *why* – proč, *how* – jak (blíže viz kapitola 2.3.3).

O vymezení základních kategorií se snažili i odborníci v oblasti informační vědy, kteří se zaměřují především na aplikace v oblasti systémů organizace znalostí. Například dle Anthony C.

---

<sup>111</sup> „*The state or condition of being a subject.*“ *subjecthood*, n. In: *OED: Oxford English Dictionary* [online]. Oxford: Oxford University Press, © 2014 [cit. 2018-01-30]. Dostupné komerčně z: <http://www.oed.com/view/Entry/192696>.

<sup>112</sup> FURNER, Jonathan. *FRSAD and the ontology of subjects of works*. In: *Cataloging & classification quarterly*. 2012, **50**(5-7), s. 513. ISSN 0163-9374. doi:10.1080/01639374.2012.681252.

<sup>113</sup> ARISTOTELÉS. *Organon. I, Kategorie*. Přeložil A. Kříž. [Komentář A. Kříž a K. Berka. Úvodní studii napsal Karel Berka.] 1. vyd. Praha: Nakladatelství Československé akademie věd, 1958. 75 s.

Fosketta existuje pět základních kategorií pojmů: entity, aktivity, abstraktní pojmy, vlastnosti, heterogenní pojmy.<sup>114</sup> Obdobný přístup je uplatněn v mezinárodní normě pro tezaury ISO 25964-1, která v částech 5.1.2 a 5.1.3 nabízí orientační výběr z typických základních kategorií: věci a jejich fyzické části, materiály, činnosti nebo procesy, události nebo výskyty, vlastnosti osob, věcí, materiálů nebo akcí, vědní disciplíny nebo obory, jednotky měření, typy osob a organizací. Individuální entity označené vlastními jmény lze podle normy rovněž seskupovat do kategorií – nabízí se kategorie pro místa, specifické objekty, topografické úkazy, jednotlivce, organizace, společnosti.

Na systému základních kategorií jsou založeny i všechny typy infromatických ontologií – terminologické, informační i znalostní (viz kapitola 1.4).

Typickým zástupcem terminologických ontologií je *Mezinárodní desetinné třídění (MDT)*, jež je jedním z nejvýznamnějších a nejrozšířenějších univerzálních systémů organizace znalostí. Strukturu klasifikačního schématu *MDT* vyjadřují tzv. tabulky, z nichž každá obsahuje vlastní hierarchický seznam znaků, kombinovatelných se znaky z ostatních tabulek. Základní schéma představuje deset vrcholových tříd hlavních tabulek<sup>115</sup>, postupně členěných podle narůstající specifičnosti na další úrovně s uplatněním desetinného principu. Tyto základní kategorie slouží k vyjádření tématu a jsou založeny na koncepci klasifikace věd Francise Bacona. V aktuální verzi *MDT* mají následující označení:<sup>116</sup>

0 Věda a poznání. Organizace. Věda o počítačích. Informace. Dokumentace. Knihovnictví. Instituce. Publikace

1 Filozofie. Psychologie

2 Náboženství. Teologie

3 Společenské vědy

5 Matematika a přírodní vědy

6 Aplikované vědy. Lékařství. Technika

7 Umění. Rekreace. Zábava. Sport

8 Jazyk. Lingvistika. Literatura

9 Geografie. Biografické studie. Dějiny

Tyto hlavní třídy jsou doplněny tzv. pomocnými tabulkami, jež v kombinaci s hlavními znaky umožňují uplatnit další hlediska, například formu či jazyk organizovaného zdroje. Třetí sadou jsou zvláštní pomocné znaky, navržené podle specifik konkrétního předmětu, které lze kombinovat s hlavními znaky určité třídy.

Velký vliv na vývoj moderního přístupu k organizaci znalostí v knihovnách mělo pojetí kategorií jako základního principu metody fasetové analýzy (viz kapitola 2.3). Prvotním impulsem byla pětice kategorií stanovená Shiyali R. Ranganathanem jako základ jeho *Dvojtečkového třídění*: ‚osobnost‘ (angl. *personality*, zkratka P) – podstata, kategorie příbuzná s Aristotelovou kategorií substance, hmota/látka/materiál (angl. *matter*, zkratka M), energie (angl. *energy*, zkratka E) – aktivita, prostor (angl. *space*, zkratka S) a čas (angl. *time*, zkratka T). Tuto pětici kategorií rozpracovala britská *Skupina pro výzkum klasifikace (CRG)* do systému 13 kategorií: věc, druh, část, vlastnost, materiál, proces, operace, pasivní prvek činnosti, produkt, vedlejší produkt,

---

<sup>114</sup> FOSKETT, Anthony Charles. *The subject approach to information*. 5th ed. London: Library Association Publishing, 1996, s. 77.

<sup>115</sup> Poznámka: Třída 4 je momentálně neobsazena.

<sup>116</sup> Zdroj: UDC Consortium. *České MDT Online* [online]. Den Haag: UDC Consortium, 2018 [cit. 2018-01-30]. Dostupné z: <http://cz.udc-hub.com/cs/login.php>. Dostupné také z: <http://www.udc-hub.com/>. Přístup volný po registraci.

agent, prostor, čas.

Zástupci skupiny informačních ontologií relevantní pro doménu organizace znalostí jsou referenční pojmové modely *IFLA LRM*<sup>117</sup> a *CIDOC CRM*<sup>118</sup>.

Referenční pojmový model *IFLA LRM* je důležitým příspěvkem k tvorbě základní ontologie pro oblast knihoven. Přes svoje zakotvení ve sféře paměťových a fondových institucí a tím dané omezení na tzv. bibliografické univerzum představuje zároveň model *IFLA LRM* v současnosti nejobecněji přijímaný explicitně specifikovaný model (ontologii) organizace znalostí. V rámci modelu je definováno 11 kategorií propojených hierarchickými a asociativními vztahy. Na nejvyšší úrovni hierarchie se nachází entita věc (*res*, kód E1), zahrnující všechny entity v rámci univerza diskurzu, které v modelu *IFLA LRM* tvoří bibliografické univerzum. Podkategoriemi entity věc jsou dílo (angl. *work*, kód E2), vyjádření (angl. *expression*, kód E3), provedení (angl. *manifestation*, kód E4), jednotka (angl. *item*, kód E5), agent (angl. *agent*, kód E6) s podkategoriemi osoba (angl. *person*, kód E7) a kolektivní agent (angl. *collective agent*, kód E8), jméno (*nomen*, kód E9), místo (angl. *place*, kód E10), časové rozpětí (angl. *time-span*, kód E11).

Ontologie *CIDOC CRM* byla vytvořena mezinárodní muzeologickou organizací jako doménová referenční ontologie pro výměnu informací o kulturním dědictví, má ovšem rovněž rysy základní ontologie, protože obsahuje i obecné pojmy pro vyjádření časoprostorových charakteristik a abstraktních pojmů. Ontologie je tedy použitelná i v jiných oblastech. Díky vysoké míře formalizace je možné ji řadit mezi těžké ontologie. Zahrnuje 82 tříd a 262 predikátů, jež pokrývají i pojmy nejvyšší úrovně. Na vrcholu hierarchie je třída ‚E1 CRM entita‘, jež zahrnuje pět vrcholových kategorií: dočasnou entitu, trvalou položku, časové rozpětí, místo a rozměr.

Jako příklady kategoriálních systémů formalizovaných znalostních ontologií v našem stručném přehledu uvedeme vrcholové kategorie základních ontologií *DOLCE* a *BFO (Basic Formal Ontology)*<sup>119</sup> Barryho Smithe a spoluautorů.

Základní ontologie *DOLCE* má na nejvyšší úrovni čtyři základní kategorie: objekt (endurant), událost (perdurant), kvalita (vlastnost) a abstraktní entita. Obdobně základní ontologie *BFO* obsahuje dvě nejobecnější kategorie – statické závislé a nezávislé kontinuanty a dynamické výskyty, jež se člení na procesy, hranice procesů a dočasné a časoprostorové entity.

V následující tabulce jsou pro srovnání uvedeny některé výše zmíněné seznamy základních kategorií. V některých případech je dvojitou čarou naznačeno vnitřní členění daného systému. Je patrné, že četné kategorie se (třeba i s pozměněnými názvy) vyskytují ve většině systémů. V plné míře to platí pro známé filozofické kategorie prostor a čas. Zároveň je však zřetelně vidět, že každý z představených systémů je založen na poněkud jiných základech a nelze je tedy jednoduše propojit přímými odkazy z jedné soustavy do druhé.

<sup>117</sup> *IFLA Library Reference Model (LRM)* [online]. Pat Riva, Patrick LeBoeuf, Maja Žumer, ed. Hague: International Federation of Library Associations and Institutions, rev. August 2017 [cit. 2018-01-30]. 101 s. Dostupné z: <https://www.ifla.org/publications/node/11412>.

<sup>118</sup> ICOM/CIDOC Special Interest Group. *Definition of the CIDOC Conceptual Reference Model* [online]. Produced by the ICOM/CIDOC Documentation Standards Group, continued by the CIDOC CRM Special Interest Group. Christian Emil ORE, Martin DOERR, Patrick Le BOEUF, Stephen STEAD, ed. Version 6.2. © 2003, version from May 2015 [cit. 2018-01-30]. 206 s. Dostupné z: <http://www.cidoc-crm.org/Version/version-6.2>.

<sup>119</sup> ARP, Robert, SMITH, Barry, SPEAR, Andrew D. *Building ontologies with Basic Formal Ontology*. Cambridge, Mass.: MIT Press, 2015. 248 s. ISBN 978-0-262-52781-1 (brož.). ISBN: 978-0-262-32957-3 (eBook). doi:<https://doi.org/10.7551/mitpress/9780262527811.001.0001>.

Aristoteles	5W1H	CRG	MDT	Ranganathan	IFLA LRM	CIDOC CRM	DOLCE	BFO
podstata	kdo	věc	Hlavní tabulky (0-9) – ‚téma‘	osobnost (P)	dílo	dočasná entita	abstraktní entita	kontinuant
kvantita	co	druh	Tabulky všeobecných pomocných znaků	hmota (M)	vyjádření	trvalá položka	objekt (endurant)	▪ závislý
kvalita	kde	část	...jazyka	energie (E)	provedení	časové rozpětí	událost (perdurant)	▪ nezávislý
vztah	proč	vlastnost	...formy	prostor (S)	jednotka	místo	kvalita	výskyt
místo	jak	materiál	...místa	čas (T)	agent	rozměr		▪ hranice procesu
čas	kdy	proces	... plemen, etnických skupin a národů		▪ osoba			▪ proces
poloha		operace	...času		▪ kolektivní agent			▪ časoprostorový
mít (vlastnictví, stav)		pasivní prvek	...vlastností		jméno			▪ dočasný
činnost		produkt	...materiálů		místo			
trpnost		vedlejší produkt	...vztahů, procesů a operací (postupů)		časové rozpětí			
		agent	...osob a osobních charakteristik					
		prostor	Zvláštní pomocné znaky					
		čas						

Tab. 2 Příklady základních kategorií



### 1.5.4 Problém určení typů a vlastností ontologických entit

Problém rozlišení typů a určení vlastností ontologických entit je živě diskutovaným tématem. Podrobně se mu věnují např. Vojtěch Svátek a Miroslav Vacura<sup>120</sup>. Teoretické volby v pozadí návrhu základních ontologií popisují Stefano Borgo a Claudio Masolo<sup>121</sup>, Natalya Noyová a Deborah McGuinnessová se zaměřují na doménové ontologie<sup>122</sup>. Přes poměrně rozsáhlou literaturu na toto téma však stále platí konstatování Giancarla Guizzardiho z roku 2005, že v praxi pojmového modelování a ontologického inženýrství se tyto závažné problémy obvykle řeší ad hoc a intuitivním způsobem<sup>123</sup>, závazná pravidla a metodiky zatím nebyly zformulovány.

Klíčovými entitami každé ontologie jsou pojmy, jež popisují vlastnosti věcí v rámci univerza diskurzu. Tradičně se pro ontologické pojmy používá označení třída a věci se označují jako instance. Vlastnosti se obvykle ještě rozlišují podle toho, zda jsou přisuzovány pouze jedné věci (například jméno osobě), či zda jsou sdíleny více věcmi (například vlastnost autorství sdílená osobou a knihou). V prvním případě jsou často označovány jako atributy a v druhém případě se v běžném jazyce mluví o vztazích. Zatímco třídy mají definované abstraktní vlastnosti, instance tříd mají vlastnosti „naplněné“ hodnotami (někdy jsou hodnoty označovány jako instance vlastností). Hned v úvodu je možné konstatovat, že s každou uvedenou entitou jsou spojené významné problémy. Způsob jejich řešení přitom ovlivňuje všechny oblasti užití ontologie (viz kapitola 1.3).

Problémy lze zhruba rozdělit do tří okruhů, jež se týkají určení entit, vztahů entit a vlastností entit. Potíže ovšem začínají už při označování ontologických entit, pro něž se v praxi volí podle konkrétní aplikační oblasti různé názvy. Kupříkladu v ontologickém jazyce *OWL* se používá následující názvosloví: Pojmy se označují jako třídy (*class*), věci jako individua (*individual*). Vlastnosti (*property*) se člení na datové (*data property*) a objektové (*object property*). Objektové vlastnosti jsou binární relace individuí, datové vlastnosti jsou relace mezi individuem a nějakou hodnotou vyjádřenou literálem. V modelovacím jazyce *UML* se pro tytéž entity používají termíny klasifikátor (*classifier* – nejznámějším klasifikátorem je třída v diagramu tříd), objekt (*object*), vlastnosti (*properties*) se člení na atributy (*attributes*), operace (*operations*) a vztahy (*relationships*). Vztahy se člení na asociativní (*association*), generalizaci/specializaci (*generalization/specialization*) agregaci (*aggregation*) kompozici (*composition*) a závislost (*dependency*).

Shrnutí nejčastěji se vyskytujících označení ontologických entit obsahuje tabulka 3. Jak je

---

<sup>120</sup> SVÁTEK, Vojtěch, VACURA, Miroslav. Ontologické inženýrství na sémantickém webu. In: Vladimír MAŘÍK, Olga ŠTĚPÁNKOVÁ, Jiří LAŽANSKÝ a kol. *Umělá inteligence*. 6. 1. vyd. Praha: Academia, 2013, kap. 3, s. 150-152. ISBN 978-80-200-2276-9.

<sup>121</sup> BORGIO, Stefano, MASOLO, Claudio. Foundational choices in DOLCE. In: STAAB, Steffen, STUDER, Rudi, ed. *Handbook on ontologies*. 2nd ed. Dordrecht: Springer, 2009, s. 361-381. doi:10.1007/978-3-540-92673-3\_16. ISBN 978-3-540-70999-2 (Print). ISBN 978-3-540-92673-3 (Online).

<sup>122</sup> NOY, Natalya F., MCGUINNESS, Deborah L. *Ontology development 101: a guide to creating your first ontology* [online]. March 2001 [cit. 2018-01-30]. 25 s. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880. Dostupné z: [http://protege.stanford.edu/publications/ontology\\_development/ontology101.html](http://protege.stanford.edu/publications/ontology_development/ontology101.html).

<sup>123</sup> GUIZZARDI, Giancarlo. *Ontological foundations for structural conceptual models*. Enschede: CTIT, Telematica Instituut, 2005, s. 96. Telematica Instituut Fundamental Research Series, no. 015 (TI/FRS/015). CTIT PhD Thesis Series, No. 05-74, ISSN 1381-3617. ISBN 90-75176-81-3. Dostupné z: <https://research.utwente.nl/en/publications/ontological-foundations-for-structural-conceptual-models> [cit. 2018-01-30].

z obsahu tabulky zřejmé, termíny ‚pojem‘ a ‚věc‘ jsou používány v oblasti ontologického inženýrství v jiném významu než v sémiotice (viz tabulka 1).

Pojem	Věc	Vlastnost
doména	člen	atribut
druh	entita	funkce
entitní typ <sup>124</sup>	exemplář / exemplifikace	charakteristika
faseta	ilustrace	predikát
kategorie	(logické) individuum	proměnná
klasifikátor <sup>125</sup>	instance	relace
klastr <sup>126</sup>	jednotlivina/partikulárie	role
množina	objekt <sup>130</sup>	rys
námět <sup>127</sup>	prvek kategorie/množiny	slot (položka)
obecnina/univerzálie	případ	vztah
pojmové pole <sup>128</sup>	subjekt / objekt (v <i>RDF</i> )	
rod	token	
synset <sup>129</sup>	událost	
taxon	výskyt	
třída		
typ		

Tab. 3 Terminologie ontologických entit

### Určení ontologických entit

Definování ontologických entit spočívá v rozhodnutí, co z univerza diskurzu bude považováno za třídu, co za instanci, a co bude vlastností či vztahem. Problém je v tom, že pojmy třída, instance, vlastnost a vztah jsou přes existenci odlišných definic ve skutečnosti relativní – jejich užití často závisí na kontextu, v němž se použijí. Knihu můžeme chápat jako třídu s instancemi v podobě konkrétních knižních publikací, je ale představitelná i jako kniha – instance, např. v případě uvažování o konkrétních instancích obecně vymezené třídy dokument. Knihu a jejího autora si lze představit jako dvě třídy spojené vztahem autorství, ale i jako třídu a její vlastnost (autor jako vlastnost knihy nebo kniha jako vlastnost autora). Výpůjčku knihy mohou představovat dvě třídy spojené asociací (kniha a její výpůjčka), dvojice třída – vlastnost (výpůjčka jako vlastnost knihy nebo kniha jako vlastnost výpůjčky), nebo dvojice třída – vztah (výpůjčka jako vztah mezi knihou a nějakou další třídou – např. čtenář). Dokonce i zdánlivě jasné rozlišení vlastnosti a její hodnoty není vždy snadné. Jedním z problémů jsou vlastnosti s předpokládanými booleovskými hodnotami (např. ‚ano‘ – ‚ne‘), kupříkladu ‚pilný‘ nebo ‚dospělý‘. Specifické jsou i vlastnosti, jejichž hodnoty jsou naplňovány instancemi nějaké třídy (to je případ tzv. autoritních dat používaných v paměťových institucích).

Všechny konstatované možnosti vedou k situaci, kdy je možné v různých ontologiích najít entity se stejnou sémantikou, ale odlišného typu. Konkrétním příkladem jsou entity, používané

<sup>124</sup> Používá se v omezené doméně relačních databází, někdy též zkráceně označováno jako entita.

<sup>125</sup> Obecné označení pro skupinu prvků v jazyce *UML* se společnými vlastnostmi, angl. *classifier*.

<sup>126</sup> Např. množina autoritních záznamů označujících stejnou entitu ve virtuální databázi *VIAF*.

<sup>127</sup> Tzv. topik v systému *Topic Maps*, angl. *topic*.

<sup>128</sup> Skupina pojmů v tezaurech, jež mají stejný nadřazený pojem.

<sup>129</sup> Soubor synonymních termínů v lexikální ontologii *WordNet*.

<sup>130</sup> Používá se v omezené doméně objektově orientovaného programování – zde jako protiklad ke třídě.

několika významnými ontologiemi k vyjádření obsahu informačních zdrojů (tzv. *aboutness* – viz kapitola 1.5.3). Metadatové schéma *Dublin Core* definuje obsah jako datovou (unární) vlastnost prostřednictvím prvku ‚dc:subject‘. V ontologii *CIDOC CRM* je zavedena objektová (binární) vlastnost ‚je o (je předmětem)‘. Model *FRBR<sub>ER</sub>* používá k vyjádření obsahu vztah ‚má jako předmět‘. V modelu *FRSAD* je pro určení obsahu díla definována samostatná třída ‚Thema‘ a v modelu *IFLA LRM* je opět zaveden vztah ‚má jako předmět / je předmětem‘. Mapování těchto entit je sice možné, ale odvozování a usuzování nad takovou heterogenní množinou povede pravděpodobně ke ztrátě informace nebo dokonce k chybným výsledkům.

Důležitou úlohu při definování ontologických entit hraje rozlišení třídy a její instance. Z tohoto rozhodnutí vyplývají pro dané entity významné důsledky. V ontologickém inženýrství se zpravidla do ontologie zařazují pouze třídy, instance tvoří náplň znalostní báze (viz kapitola 1.2). Zatímco třídy mohou mít instance, instance už další instance mít nemohou, tj. mají pouze intenzi, nikoli extenzi. Rozlišení třídy a instance je důležité pro určení vztahu mezi těmito entitami: vztah třídy a instance se označuje jako instanční hierarchie, hierarchický vztah tříd může být buď generický nebo partitivní (o těchto vztazích bude pojednáno v následující části). Stejně jako není jednoduché rozhodnout, které věci jsou ekvivalentní, není ani explicitní určení třídy a instance triviální. I v tomto případě se znovu ‚vrací do hry‘ témata propracovaná scholastickou filozofií ve sporu o univerzálie (viz kapitola 1.5.1). Zvolený přístup k řešení se někdy odráží i v použité terminologii. Nominalistický přístup, který předpokládá jako prvotní existenci věcí, jež jsou až dodatečně kategorizovány a seskupeny do tříd, označuje vztah třídy a instance termíny exemplifikace nebo ilustrace. Realistický přístup používá pro tentýž vztah, který ovšem chápe opačně v čase (neboli existence třídy předchází vytvoření instance), termín instanciace. Termín instanciace tedy označuje uskutečnění třídy, exemplifikace a ilustrace jsou chápány jako doložení obecné třídy konkrétním příkladem.

Určitou pomůckou pro rozlišení třídy a instance se mohou stát specifické typy jejich označení. Pro slovní označení instance se používá vlastní jméno (*proprium*), například pro označení hlavního města ČR se používá jméno Praha (nebo Prag, Prague, Praga...). Pro instance označené vlastním jménem se aktuálně používá název pojmenované entity a vyvíjejí se programy pro jejich automatické rozpoznávání v textech (*named entity recognition* – *NER*). V realitě však existují i instance bez označení, takže tato pomůcka není vždy aplikovatelná. Třídy se označují obecnými jmény (*apelativy*), v odborném jazyce pak jsou nejčastěji používanými typy obecných jmen termíny. Obvyklým způsobem označení třídy jsou jména či jiné symboly, jejichž extenze zahrnuje všechny instance – například jméno ‚pryskyřníkovité‘ označuje skupinu rostlin, do níž patří orsej, blatouch a sasanka. Situaci ovšem komplikuje to, že v praxi jsou časté i případy metonymie, kdy se použije jméno či symbol typického reprezentanta obsažených objektů, tj. instance – například obrázek knihy zastupující všechny druhy dokumentů v knihovně. Při organizaci webových sídel se používají i označení kategorií prezentovaných dokumentů objasňující jejich účel (úlohu, kterou mají pomoci řešit, nebo otázku, na kterou mají dát odpověď), případně pojmenovávající typického uživatele organizovaných zdrojů či služeb.

### **Určení vlastností ontologických entit**

Jak bylo uvedeno, termínem ontologické entity označujeme třídy a jejich instance, vlastnosti a případně jejich hodnoty. V této části si všimneme charakteristik tříd a vlastností.

Protože volba charakteristik tříd výrazně ovlivňuje charakter celé ontologie, jsou někdy v literatuře tyto charakteristiky vztaženy na ontologii jako celek a tvoří tak další fasety v typologii ontologií (viz kapitola 1.4). Tímto způsobem pracuje s charakteristikami tříd i nástroj

ONSET<sup>131</sup>, navržený pro výběr základní ontologie vhodné pro opětovné použití při konstrukci doménové ontologie. Nejčastěji uváděnými charakteristikami tříd jsou způsob konceptualizace, existence reprezentovaných entit, řešení problému prostoru a řešení problému času.

Podle způsobu konceptualizace skutečnosti lze třídy označovat buď jako deskriptivní nebo jako revizionistické. Deskriptivní třída vychází z běžného chápání skutečnosti za použití přirozeného jazyka; deskriptivní princip v systémech organizace znalostí uplatňuje *literary warrant*. Revizionistická třída je založena na vědeckých principech, jež se mohou lišit od běžné empirie. Příkladem revizionistické ontologie je Mendělejevova periodická tabulka prvků, založená na periodickém zákoně, jež dokonce antcipovala existenci prvků v době jejího vzniku neznámých. Toto členění je obvyklé i v oblasti terminologické práce. M. Teresa Cabréová to formuluje následovně: „*Musíme rozlišovat mezi popisnou prací, která se omezuje na shromažďování pojmů v rámci oboru, a prací s preskriptivním (normativním) účelem, jejímž cílem je ovlivnit používání jazyka poukázáním na termíny, jejichž použití je doporučeno.*“<sup>132</sup>

Podle způsobu existence reprezentovaných entit se rozlišují třídy aktualistické a posibilistické. Aktualistická třída popisuje pouze to, co aktuálně existuje, posibilistická třída umožňuje popsat i věci zatím neexistující. Obdobnou dvojici tvoří vlastnosti prezentistická – eternální, u nichž se porovnává, zda instance tříd existují v aktuálním čase.

Způsob řešení problému prostoru přináší členění tříd na multiplikativní a redukcionistické. Zatímco redukcionistická třída umožňuje, aby na jednom místě v daném čase byla pouze jedna entita, multiplikativní třídy připouštějí tzv. kolokalizaci entit v jednom časoprostoru. V multiplikativní ontologii je tedy možné, aby v jednom časoprostoru bylo více různých entit. Tím se přibližuje koncepci fasetového členění (viz kapitola 2.3).

Podle toho, jak řeší časovou dimenzi, se třídy člení na 3D (třírozměrné) a 4D (čtyřrozměrné). Na nejobecnější úrovni pohledu jde o rozdíl mezi prvky/objekty a procesy/událostmi, které se liší způsobem trvání v čase. Instance 3D třídy je v každém časovém okamžiku přítomna celá (zachovává si svou identitu), instance 4D třídy je v každém časovém okamžiku přítomná jen v jedné své části. Kupříkladu na oběd lze nahlížet jako na objekt – množinu potravin na talíři, nebo jako na proces – oběd jako událost trvající v čase. Vztahy celek–část v 3D a 4D třídách jsou odlišného typu: zatímco části objektu (např. maso a brambory jako suroviny, z nichž se uvaří oběd) mohou existovat společně v jednom čase, části procesu tvoří časové úseky (např. aperitiv, předkrm, hlavní jídlo, dezert), jež nemohou existovat společně v jednom časovém okamžiku či úseku. V informatických ontologiích se pro dvojice tříd 3D a 4D často používá označení *endurant* – *perdurant* (angl. též *continuant* – *occurent*), které vychází z filozofické tradice.<sup>133</sup>

---

<sup>131</sup> KHAN, Zubeida, KEET, C. Maria. ONSET: automated foundational ontology selection and explanation. In: Annette ten Teije et al., ed. *Knowledge Engineering and Knowledge Management: 18th International Conference, EKAW 2012, Galway City, Ireland, October 8-12, 2012*. Proceedings. Berlin, Heidelberg: Springer-Verlag, 2012, s. 237-251. doi:10.1007/978-3-642-33876-2\_22. Lecture notes in computer science, no. 7603, ISSN 0302-9743 (Print). ISSN 1611-3349 (Online). ISBN 978-3-642-33875-5 (Print). ISBN 978-3-642-33876-2 (Online).

<sup>132</sup> CABRÉ, M. Teresa. *Terminology: theory, methods and applications*. Edited by Juan C. SAGER. Translated by Janet Ann DeCESARIS. Amsterdam; Philadelphia: John Benjamins Publishing Company, © 1999, s. 132. Terminology and lexicography research and practice, ISSN 1388-8455, vol. 1. ISBN 978-90-272-1633-5 (váz.). ISBN 978-90-272-1634-2 (brož.). ISBN 978-90-272-9865-2 (Online).

<sup>133</sup> Podrobnější teoretický komentář k této problematice viz např. JOHANSSON, Ingvar. Qualities, quantities, and the *endurant-perdurant* distinction in top-level ontologies [online]. In: *CEUR Workshop Proceedings*. 2015 [cit. 2018-01-30], vol. 130. ISSN 1613-0073. Dostupné z: <http://ceur-ws.org/Vol-130/>.

Zatímco charakteristiky tříd jsou obvykle využívány k určení typu ontologie, charakteristiky vlastností jsou v praxi často využívány pro hodnocení ontologie. Vlastnosti lze přisoudit nejen třídě, ale i vztahu a dokonce i samotné vlastnosti (jako příklady lze uvést definiční obor a obor hodnot vlastnosti v jazyce *OWL* nebo datový typ a kardinalitu atributu v databázové tabulce). Na typologii „vlastností vlastností“ je přímo založena metodika analýzy ontologií *OntoClean* (blíže o této metodice v kapitole 2.1.2). V následujícím stručném přehledu se zaměříme na tyto charakteristiky vlastností: existenční podstata, trvalost, účel v informačním systému, význam (důležitost), počet tříd, role entit, reflexivita a tranzitivita.

Podle existenční podstaty se vlastnosti člení na vnitřní (vlastní, přirozené, angl. *intrinsic*) a vnější (zevní, cizí, přidělené, angl. *extrinsic*). Za vnitřní vlastnosti se považují ty, s nimiž objekt vznikl (například krevní skupina osoby), vnější vlastnosti jsou získané v průběhu jeho existence (například jméno, rodinný stav osoby). Při identifikaci ontologických entit se samozřejmě dává přednost vnitřním vlastnostem, které charakterizují entitu po celou dobu její existence.

Podle trvalosti vlastnosti se rozlišují statické (neměnné v čase) a dynamické vlastnosti (proměnlivé v čase – existencí či hodnotou). Například osoba může během svého života změnit jméno (resp. hodnotu atributu „jméno“), může získat i ztratit řidičské oprávnění, nemůže ovšem přestat mít krevní skupinu.

Podle jejich účelu v informačním systému lze vlastnosti rozdělit na identifikační, popisné a vyhledávací. Identifikační vlastnosti věc odlišují od všech ostatních. Popisné vlastnosti tvoří maximální (úplná) množina atributů, které věc reprezentují (zastupují) nebo doplňují. Vyhledávací vlastnosti, podle nichž lze věc vyhledat, jsou často označovány jako přístupové body.

Význam (důležitost) vlastnosti se vztahuje k existenci entity – posuzuje se, zda je vlastnost pro existenci entity nezbytná. Aristotelská tradice rozlišuje vlastnosti podle tohoto kritéria na podstatné a nepodstatné. Podstatná vlastnost představuje souhrn obecných stálých určení dané věci. Podstatné vlastnosti spadají pod kategorii podstaty (blíže k Aristotelovu kategoriálnímu systému viz kapitola 1.5.3), lze je vyjádřit substantivně. Nepodstatné vlastnosti (akcidenty) spadají pod kategorii kvality a lze je vyjádřit pouze adjektivně. Tyto vlastnosti mohou či nemusí být přítomné, aniž by se tím změnila povaha věci, již náležejí (např. jablko nepřestane být jablkem, pokud nebude mít stanovenou prodejní cenu). Obdobnou typologii vlastností podává norma ISO 704 pro terminologickou práci, jež uvádí v části 5.4.5 tři typy vlastností objektů: nutné (angl. *necessary*), postačující (angl. *sufficient*) a podstatné (angl. *essential*). Vlastnosti nutné musí mít každý objekt/instance dané třídy. Vlastnosti postačující samy o sobě stačí k přiřazení objektu/instance k dané třídě. Vlastnosti podstatné splňují jednak podmínky nutnosti, jednak jsou postačující.<sup>134</sup>

V normě ISO 704 je rovněž uplatněno kritérium počtu tříd, jež mají danou vlastnost. V části 5.4.4 jsou rozlišeny sdílené (angl. *shared*) vlastnosti, jež mají třídy patřící do stejné skupiny, od vlastností vymezujících (angl. *delimiting*), jež odlišují třídy navzájem.<sup>135</sup>

Kritérium role entit se týká sdílených vlastností, jež plní funkci vztahu mezi entitami. Jsou-li role entit sdílejících danou vlastnost stejné, prohlásíme vlastnost za symetrickou. Mají-li entity ve

---

<sup>134</sup> ISO 704:2009. *Terminology work – Principles and methods*. 3rd ed. Geneva: International Organization for Standardization, 2009, s. 7.

<sup>135</sup> Tamtéž, s. 6-7.

vztahu různé role, jde o vlastnost asymetrickou.

Pro označení počtu hodnot vlastnosti se používají termíny násobnost nebo kardinalita. Počet větší než 1 se obvykle nevyjadřuje číselně, ale pomocí zobecnujícího symbolu, např. N, M, \*, ∞. Jazyk UML rozlišuje násobnosti: právě 1 (1..1, 1), 0 nebo 1 (0..1), 0 až více (0..\*, \*), 1 až více (1..\*). Obvyklé jsou způsoby vyjadřování násobnosti vztahů: jedna–jedna (1 : 1, angl. *one-to-one*), jedna–více (jeden k mnoha, 1 : N, angl. *one-to-many*), více–jedna (N : 1, angl. *many-to-one*), více–více (mnohý k mnoha, N : M, angl. *many-to-many*). Dalším kvantitativním parametrem je tzv. povinnost výskytu vlastnosti či členství (účasti) ve vztahu.

Reflexivní/ireflexivní vlastnost (z lat. *reflexio* – obrácení zpět) určuje vztah entity s ní samou. Například vztahy totožnosti, ekvivalence a podobnosti jsou reflexivní, vztah generické hierarchie je ireflexivní (druh nemůže být svým rodem).

Tranzitivitu neboli přenositelnost vlastnosti lze vyjádřit formulí: když platí  $A \rightarrow B \rightarrow C$ , platí, že  $A \rightarrow C$ . Dalším příkladem tranzitivity je přenos sémantiky vztahu tříd na konkrétní vztah jejich instancí.

### Určení vztahů ontologických entit

Samostatný problém představuje rozlišení vztahů ontologických entit. Opět platí, že způsob určení typu vztahu zásadně ovlivňuje případy užití ontologických entit, jež jsou daným vztahem spojeny. Jedním z významných případů užití je tranzitivita (přenositost) vztahu, která umožňuje odvozování. Obecně jsou všechny typy vztahů považovány za tranzitivní. Při jejich vzájemné kombinaci a řetězení už ale toto všeobecné pravidlo nelze uplatnit.<sup>136</sup> Tranzitivita zpravidla spolehlivě funguje pouze v řetězcích vztahů stejného typu. Navíc v rozdílných typech vztahů je tranzitivita uplatněna na různé aspekty – přenos samotného vztahu nebo přenos vztahu i vlastností entit, jež jsou ve vztahu.

Na nejobecnější úrovni se uvažují vztahy ekvivalence, asociace a hierarchie. V rámci hierarchie lze ještě rozlišovat mezi hierarchií generickou a partitivní a mezi mono- a polyhierarchií (nemožnost či možnost, aby jedna entita měla více nadřazených entit). Ekvivalence a asociace jsou symetrické, hierarchie je asymetrická. Tranzitivita ekvivalence se projevuje jak ve vztazích, tak ve vlastnostech. Je-li A ekvivalentní s B a B je ekvivalentní s C, platí, že A je ekvivalentní s C, tj. vlastnosti A se dají vztáhnout i na C. Rovněž v generické hierarchii se uplatňuje jak tranzitivita vztahu, tak tranzitivita vlastností. Tranzitivita asociace a partitivní hierarchie se uplatňuje pouze ve vztazích. Je-li A asociováno s B nebo je-li částí B a B je asociováno s C nebo je částí C, platí, že A je asociováno s C nebo je částí C, nelze ovšem usuzovat, že by měly podobné vlastnosti.

I zdánlivě jednoznačný vztah ekvivalence začne být problematickým, jakmile se připomene, že v sémantickém kontextu nejde o totožnost, ale o shodu významných vlastností, tedy o určitou míru podobnosti. Pro formální vyjádření ekvivalence ontologických entit je tedy místo znaménka = vhodnější použít symbol ≈. Problém pak spočívá v tom, jakou toleranci pro obsahovou ekvivalenci zvolit a kde stanovit hranice mezi ekvivalencí a asociací. V jazyce OWL platí, že entity s vlastností ‚owl:sameAs‘ (instance), ‚owl:equivalentClass‘ (třída) a ‚owl:equivalentProperty‘ (vlastnost) jsou považovány za ekvivalentní pro účely odvozování (a jsou tedy například zahrnuty do výsledků vyhledávání nebo do automatického mapování). Jazyk datového slovníku SKOS oproti tomu nabízí bohatší škálu typů ekvivalence. Vlastnosti

---

<sup>136</sup> Viz například detailní specifikaci podmínek pro tranzitivitu vztahů v *Genové ontologii* na <http://www.geneontology.org/page/ontology-relations> [cit. 2018-01-30].

,skos:closeMatch', ,skos:exactMatch', ,skos:broadMatch', ,skos:narrowMatch', ,skos:relatedMatch' jsou určeny pro mapování mezi různými systémy organizace znalostí a propojené pojmy jsou považovány za ekvivalentní pro účely vyhledávání. Obdobně norma ISO 25964-1 stanoví stupně mezijazykové a pojmové ekvivalence k rozlišení míry podobnosti zúčastněných pojmů: ekvivalence přesná, nepřesná, částečná, širší/užší, nonekvivalence. Exaktní pravidla či kritéria pro určení jednotlivých typů ekvivalence ovšem zatím chybí.

Jak bylo uvedeno, hierarchické vztahy lze členit na dva základní typy – generické a partitivní. Mnohé lehké ontologie však často rezignují na striktně formální specifikaci hierarchie a používají volně definovanou ,obecnou' či ,hlediskovou' hierarchii<sup>137</sup>, již nelze členit na partitivní a generickou ani ji korektně použít k automatickému usuzování a odvozování. Kupříkladu z toho, že třída *MDT 677.4 Chemická vlákna* je podtřídou třídy *677 Textilní průmysl*, nelze odvodit, že dokument zařazený do třídy 677 skutečně obsahově pokrývá (zahrnuje) i téma chemických vláken. Jediným korektním tvrzením v tomto typu hierarchického vztahu je, že třída na vyšší úrovni má širší extenzi a užší intenzi než její podtřída.

V generické hierarchii jde obecně o inkluzivní vztah množiny a její podmnožiny, třída zahrnuje podtřidu (tj. její extenzi i intenzi). Pragmatickou pomůckou pro rozpoznání generické hierarchie je test validity ,všichni – někteří', doporučený například normou ISO 25964-1<sup>138</sup>. Její uplatnění na dvojici tříd ,Dopravní prostředek' a ,Vlak' by vypadalo následovně: všechny vlaky jsou dopravní prostředky, některé dopravní prostředky jsou vlaky. Zdánlivě obdobná dvojice tříd ,Dopravní prostředek' a ,Velbloud' už tímto testem neprojde: všichni velbloudi nejsou dopravní prostředky, i když některé dopravní prostředky jsou velbloudi. Tranzitivita vlastností v generické hierarchii se projevuje tím, že jednou definovaná vlastnost v nadřazené úrovni se stává vlastností prvků ve všech podřazených úrovních. Jinými slovy – pokud je třída ,Vlak' podtřídou třídy ,Dopravní prostředek', pak každá instance vlaku je též instancí dopravního prostředku. Tato tranzitivita se někdy označuje jako dědičnost a platí i pro vztahy – veškeré vztahy nadřazených tříd (například asociace s jinými třídami) platí i pro jejich podtřídy.

Klíčový význam generické hierarchie v konstrukci infromatických ontologií vyjadřuje označení ,páteřní taxonomie' (angl. *backbone taxonomy*). Zahrnuje strukturu, jejímiž prvky jsou ontologické třídy a jejich generické vztahy.

Vztah generické hierarchie umožňuje členit třídy na podtřídy prakticky neomezeně co do granularity a z toho důvodu je někdy obtížné jej odlišit od vztahu třída – instance. Pro označení obecné třídy se obvykle používá označení rod, podtřída se označuje jako druh. Protože druh je třída, může mít kromě zděděných vlastností i své specifické vlastnosti a může mít instance, zatímco instance nikoli. Situaci ještě komplikuje existence tzv. metatříd či abstraktních tříd, jež nemají své vlastní instance, ale jejich instancemi (tj. nikoli druhem) jsou třídy. Kdy máme před sebou třídu–druh, třídu–instanci a kdy instanci, je často obtížné rozhodnout. V některých případech lze pro určení druhu použít pomůcku ,je typ'. O dvojici Město (třída) – Hlavní město lze prohlásit, že hlavní město je typ města, tedy třída–druh. U dvojice Hlavní město (třída) –

---

<sup>137</sup> Viz např. pojetí hlediskové hierarchie Elaine Svenoniové, která ji chápe jako syntagmatický ad hoc vztah, závislý na specifickém úhlu pohledu: SVENONIUS, Elaine. Subject languages: referential and relational semantics. In: *The intellectual foundation of information organization*. Cambridge (Mass): MIT Press, 2000, chapter 9, s. 162-168. ISBN 978-0-262-19433-4.

<sup>138</sup> ISO 25964-1:2011. *Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval*. 1st ed. Geneva: International Organization for Standardization, 2011-08-08, s. 59.

Praha však už nelze tvrdit, že Praha je typem hlavního města, ta je jeho instancí.

Partitivní vztah, který vyjadřuje vztah mezi celkem a jeho částí, má z hlediska odvozování stejné vlastnosti jako vztah asociace. Hranice mezi partitivní hierarchií a asociací a mezi vztahem třídy a její vlastnosti opět nejsou zcela zřetelné. Problematika partitivních vztahů je natolik specifická, že se jí dokonce věnuje samostatná vědní disciplína na pomezí filozofie a logiky – mereologie. Rozlišují se četné specifické typy partitivních vztahů, například vztahy prostorové, konstitutivní nebo vztahy členství či účasti v procesu. Lze konstatovat, že tranzitivita partitivního vztahu platí pouze u entit (a vztahů) stejného typu. Pokud bychom zkusili odvodit z trojice tříd Ruka <je část> Jan Novák <je člen> FC Kolín, že ruka je členem FC Kolín, dospěli bychom bezesporu k chybnému výsledku. Norma ISO 25964-1 dále doporučuje, aby partitivní vztahy byly definovány jako monohierarchické, část má patřit vždy jen do jednoho celku. Takovému požadavku může vyhovět poměrně malý okruh entit, které norma vyjmenovává v části 10.2.3.1: tělesné systémy a orgány, geografické lokality, disciplíny nebo oblasti rozpravy, hierarchické společenské struktury.

Za asociční vztah se považuje jakýkoli vztah, spojující třídy, jež se liší svými vlastnostmi (tj. nejsou ekvivalentní), a nelze jej označit za hierarchický. Vzhledem k významu standardizace vztahů pro komunikaci v počítačovém prostředí není divu, že mnozí autoři volají po prohloubení specifikace a větší formalizaci asociativních vztahů. Elaine Svenoniová však upozorňuje, že problém může mít dvě stránky:

*„Rozčlenění obecných vztahů asociace na skupiny specifitějších vztahů se zdá být nevyhnutelným v obecném vývoji předmětových jazyků směrem k specifčnosti a formalismu. Nicméně existuje i možnost, že skutečně užitečné asociativní vztahy mohou být ve skutečnosti ty, jež jsou lexikálně neurčitelné, ty, jež se nehodí k formalizaci.“<sup>139</sup>*

### 1.5.5 Předpoklad otevřeného světa

Sémantický web představuje v současné vývojové etapě bezesporu nejrozsáhlejší oblast aplikace informatických ontologií. V porovnání s „tradičními“ znalostními systémy vykazuje sémantický web významná specifika, jež lapidárně vyjadřují anglické zkratky *OWA*, *AAA* a *NUNA*.<sup>140</sup>

Princip *OWA* (z angl. *open world assumption*) označuje předpoklad otevřeného světa v tom smyslu, že výroky nezahrnuté v sémantickém webu nejsou považovány za nepravdivé, věci nezastoupené na webu nejsou považovány za neexistující. Univerzum poznání, reprezentované prostřednictvím webových zdrojů, je chápáno jako dynamicky se měnící. Mění se nejen suma dostupných znalostí (nové zdroje se objevují, stávající zdroje mizí), ale i vztahy mezi nimi. Žádné tvrzení se nepovažuje za definitivní.

Princip *AAA* (z angl. *anyone can say anything about anything* – kdokoli může říci cokoli o čemkoli) označuje možnost různých způsobů konceptualizace, diskutovanou v kapitole 1.5.3. Konstatuje, že jedna věc může být chápána a vyjádřena různými způsoby, tj. jedna věc může mít více významů. O stejném předmětu mohou být k dispozici různá tvrzení. Tato tvrzení nemusejí

<sup>139</sup> SVENONIUS, Elaine. Subject languages: referential and relational semantics. In: *The intellectual foundation of information organization*. Cambridge (Mass): MIT Press, 2000, chapter 9, s. 162. ISBN 978-0-262-19433-4.

<sup>140</sup> ALLEMANG, Dean, HENDLER, James. *Semantic web for the working ontologist: effective modeling in RDFS and OWL*. 2nd ed. Waltham (MA) Amsterdam: Morgan Kaufmann/Elsevier, © 2011, s. 13. ISBN 978-0-12-385966-2 (Online). ISBN 978-0-12-385965-5 (Print).



být navzájem konzistentní, ale nemusejí se ani vylučovat.

Princip *NUNA* (z angl. *non-unique naming assumption* – nepravdivost předpokladu jedinečných jmen) vychází z předpokladu nejednotného pojmenování a připouští, že jedna věc může mít více označení. Tento problém možnosti různého vyjádření stejného pojmu (synonymie) byl již zmíněn v kapitole 1.5.2.

Z výše uvedeného vyplývá, že ontologie pro sémantický web musí být flexibilní, tj. připravená na změny a schopná zpracovávat neurčitě (neúplně) a často i nekonzistentní informace. Princip otevřeného světa a s ním související fenomény *AAA* a *NUNA* na jedné straně komplikují možnosti zpracování dat, někteří autoři však míní, že se právě proto mohou stát impulsem pro rozvoj filozofické ontologie.<sup>141</sup>

## 1.6 Dílčí závěry a diskuse

### 1.6.1 Problém sladění požadavků na jednoduchost a sémantickou sílu reprezentace

Při tvorbě inforatické ontologie se předpokládá, že její konstrukce bude založena na sémantické analýze vymezené domény. Problém vyjádření sémantiky je těsně spjat s volbou způsobu její reprezentace. Obecně lze zaznamenat dvě tendence: 1. reprezentovat co nejvěrněji často nejednoznačný obsah a mnohdy velmi komplikované a spletité vztahy a 2. vytvořit nad nevyslovitelným, implicitním obsahem jednoduchou abstraktní strukturu, jež bude moci být konsenzuálně přijata a umožní základní orientaci (s tím, že kognitivní procesy a struktury v celé jejich komplexnosti a individualitě nelze algoritmizovat a obsah nelze jednoznačně vyjádřit).

Klíčovou otázkou je množství a míra specifičnosti popisovaných vztahů. Zatímco Zdeněk Zdráhal konstatuje, že pro účely zpracování znalostí počítačovými aplikacemi „*jsme ochotni obětovat přesnost popisu za zjednodušení vztahů mezi pojmy*“<sup>142</sup>, následující úvaha Rebeccy Greenové vystihuje komplikovanost problematiky a svou výzvu k jejímu řešení doprovází jasně zformulovanými argumenty.

*„Vztahy podílející se na organizaci znalostí jsou jak početné, tak složité ... jejich ohromné množství a složitost působí proti jejich konzistentnímu používání informačními profesionály, což platí ještě více pro koncové uživatele ... Někdo pak přirozeně vznáší otázku, zda hlubší a/nebo širší důraz na vztahy by byl vůbec pro organizaci znalostí prospěšný, nebo by pouze překážel ... Ale skutečné důvody pro důraz na vztahy pocházejí z kombinace pokračující exploze informací a jejich záznamu v nepřehledném množství forem a z narůstajícího podílu automatizovaných nástrojů na jejich správě. Intuitivní porozumění, jež vnášejí lidé do vztahů, není sdíleno výpočetními zařízeními. Přitom jsou vyjádření a manipulace vztahů pravděpodobně naší největší nadějí na prostoupení*

---

<sup>141</sup> LIVET, Pierre. Web ontologies as renewal of classical philosophical ontology. In: *Metaphilosophy*. July 2012, 43(4), 396-404. ISSN 0026-1068.

<sup>142</sup> ZDRÁHAL, Zdeněk. Ontologie: od filosofie k umělé inteligenci. In: In: Vladimír MAŘÍK, Olga ŠTĚPÁNKOVÁ, Jiří LAŽANSKÝ a kol. *Umělá inteligence*. 6. 1. vyd. Praha: Academia, 2013, kap. 1, s. 56. ISBN 978-80-200-2276-9.

našich vyhledávacích systémů vyšší kvalitou.“<sup>143</sup>

Hledání rovnováhy mezi sémantickou silou (vypovídací schopností) a jednoduchostí použití systému je obtížné. Explicitní reprezentace složitých struktur v jejich komplexitě se zdá být správným krokem směrem k poznání skutečnosti v ontologickém smyslu. Vzhledem k tomu, že v současné etapě vývoje informačních technologií vzrůstá důležitost zprostředkování významu počítačovým programům, je vhodné zdůraznit, že to je dobré pro stroje. Ty si nedokážou poradit s realitou, tak jak je, a potřebují její reprezentaci ve formátu, který je jim dostupný. Otázkou je, co tato komplexní reprezentace skutečnosti přináší pro člověka v gnozeologickém smyslu. Dovedeno do extrému, vedle nekonečně složité reality, kterou svými smysly nedokáže pochopit, se mu nabídne její explicitní specifikace v poměru 1 : 1, která klade stejný odpor lidskému poznání. Zdá se, že toto dilema je spojené s následujícím problémem konceptualizace v ‚přirozené‘ a v umělé inteligenci.

### 1.6.2 Možnosti vzájemné aplikovatelnosti principů konceptualizace v ‚přirozené‘ a v umělé inteligenci

Názory odborníků na aplikovatelnost principů fungování lidské mysli zjištěných kognitivní vědou v oblasti ontologického inženýrství nejsou jednotné. Je otevřenou otázkou, zda je vhodné usilovat o podobnost principů konceptualizace v externích pamětech s principy fungování lidské mysli, jak to naznačuje např. Vannevar Bush v článku *As we may think* (Jak asi myslíme).<sup>144</sup> Jeho návrh osobního informačního systému *memex* vycházel z psychologické teorie asocianismu, která se soustředí při výkladu psychických jevů na jejich fyzickou podstatu v podobě základních prvků vědomí (například počítků a představ), spojovaných asociacemi do komplexnějších celků. Alternativou je pojetí rozšiřování možností lidské mysli technologií, která ji nekopíruje, ale využívá vlastní efektivní techniky. Guillaume Apollinaire to v předmluvě k dramatu *Prsy Tirésiovoy* obrazně vyjadřuje přirovnáním, že „když chtěl člověk napodobiti chůzi, vynalezl kolo, které se naprosto nepodobá noze“.<sup>145</sup> V současné době má obor umělá inteligence už k dispozici technologická řešení jak pro alternativu napodobující lidské myšlení (např. neuronové sítě, ontologie *WordNet* založená na psycholingvistické teorii lexikální paměti), tak pro druhý typ založený na matematických a statistických metodách a logickém odvozování.

Výsledky zkoumání psychických procesů určování ekvivalence, dosažené v průběhu 20. století, jsou bezesporu inspirativní i pro obor ontologického inženýrství. Jejich problémem je ovšem praktická nemožnost formulování exaktních pravidel a tím méně algoritmizace. Problém teorií kognitivní vědy spočívá v jejich (ne)přenositelnosti ze světa lidského myšlení do světa počítačových a dalších technologií. Je nesporné, že mezi konceptualizací v mysli a konceptualizací v externích pamětech lze pozorovat četné analogie. Jak kognitivní věda, tak obor ontologického inženýrství mají totožný předmět zájmu, tj. znalosti. Přesto je možné zřetelně vymezit rozdíly mezi nimi – kognitivní věda je deskriptivní, ontologické inženýrství je preskriptivní. Ontologické inženýrství se soustředí výhradně na principy záměrné, vědomé

<sup>143</sup> GREEN, Rebecca. Relationships in the organization of knowledge: an overview. In: Carol A. Bean, Rebecca Green, ed. *Relationships in the organization of knowledge*. Dordrecht: Kluwer Academic Publishers, 2011, s. 13-14.

<sup>144</sup> BUSH, Vannevar. As we may think. In: *Atlantic monthly*. July 1945, 176(7), 101-108. ISSN 0160-6506.

<sup>145</sup> Citováno dle: APOLLINAIRE, Guillaume. *Prsy Tirésiovoy: nadrealistické drama o dvou jednáních s prologem*. Praha: Odeon, 1926, s. 10.

konceptualizace, otázky nevědomí (např. intuice) přenechává právě kognitivním a psychologickým vědám. Rozdíly jsou patrné i v metodách. Při zkoumání konceptualizace v mysli je možné se zaměřit na objektivně působící principy, jež mají své kořeny v biologicky determinovaných funkcích mozku. V případě konceptualizace zaznamenaných znalostí jde o výsledek vědomé aktivity směřující k řešení problému, tj. o inženýrskou disciplínu, jejíž „pravidla hry“ lze mnohem více ovlivnit a formovat.

Nespornou výhodou „umělé“ konceptualizace založené na společných vlastnostech je, že umožní počítačovou podporu a vyhledávání, počítačové programy budou „rozumět“ takto organizovaným objektům. Lze přesně stanovit její pravidla, postup je reprodukovatelný a tudíž splňuje požadavky kladené na vědeckou metodu. Problémem je kromě vědecky prokázaného rozporu s tím, jak lidé ve skutečnosti konceptualizují a v důsledcích i chápou realitu, vysoká míra abstrakce. Faktorů, od nichž se v „systémovém“ pojetí konceptualizace odhlíží, je poměrně hodně. Patří mezi ně již zmíněný kontext, subjektivní potřeba informace/znalosti, čas (minulý, budoucí), a v neposlední řadě i typ entit (např. přírodní objekty versus abstraktní entity). Ukazuje se, že některé typy entit jsou prostřednictvím svých vlastností lépe určitelné než jiné. Kupříkladu barvy lze objektivně rozlišovat podle specifické vlnové délky a frekvence záření. Takový atribut už ale postrádají vůně. Například vůně vína se obvykle nepopisují prostřednictvím specifických charakteristik vína, ale vyjadřují se pomocí analogie k vůním jiných potravin či rostlin (černý rybíz, jahody, kokos, kopřiva, tráva apod.).

### 1.6.3 Pracovní definice informatické ontologie

Většina definic informatické ontologie je založena na výčtu typů entit, jež jsou považovány za ontologie. Prvním pokusem o vymezení ontologie klasickým aristotelským způsobem (odkaz na nadřazený pojem, uvedení specifických rozdílů) je definice Thomase Grubera z roku 1993<sup>146</sup>, doplněná Willemem Nico Borstem v roce 1997<sup>147</sup> do následující formulace definice ontologie:

*„Ontologie je formální explicitní specifikace sdílené konceptualizace skutečností“.*

Tato značně kompaktní formulace si zaslouží podrobnější komentář, v němž se ji pokusíme propojit s teoretickým konstruktem sémiotického trojúhelníku a ukázat, jak odkazuje na všechny tři jeho vrcholy.

Lze konstatovat, že pojem „explicitní specifikace“ z definice ontologie odpovídá straně sémiotického trojúhelníku označené na obrázku 5 v kapitole 1.5.2 jako „vyjádření“. Slovo „formální“ znamená, že specifikace pojmu je zakódovaná v nějakém formálním jazyce a tedy srozumitelná i počítačovým programům. Zbývá doplnit komentář k významu adjektiva „sdílená“ u slova konceptualizace. Jak bylo uvedeno v kapitole 1.5.3, konceptualizace spočívá v tvorbě kognitivního odrazu skutečnosti prostřednictvím pojmů. „Sdílená konceptualizace“ znamená, že konceptualizace je využívána ke komunikaci (komunikace v pojetí sdílení a opakovatelné použitelnosti významu), a že se tedy nejedná o individuální pojmovou reprezentaci skutečnosti,

---

<sup>146</sup> „An ontology is an explicit specification of a conceptualization.“ GRUBER, Thomas R. A translation approach to portable ontology specification. In: *Knowledge acquisition*. June 1993, 5(2), 199. doi:10.1006/knac.1993.1008. ISSN 1042-8143.

<sup>147</sup> „Ontologies are defined as formal specification of a shared conceptualization.“ BORST, Willem Nico. *Construction of engineering ontologies for knowledge sharing and reuse*. Enschede: Centre for Telematics and Information technology, 1997, s. 12. CTIT Ph. D-thesis series, no. 97-14, ISSN 1381-3617. Dissertation (Ph.D.). Proefschrift Universiteit Twente, Enschede. ISBN 978-90-365-0988-6.

nýbrž o výsledek dohody o významu vyjadřovaných pojmů v nějaké širší komunitě, například v rámci vědního oboru. Potřeba sdílené konceptualizace ve formě pojmového schématu byla konstatována již v normě ISO/TR 9007 (viz kapitola 1.2). Příčinou nutnosti dohody o významu pojmů je nejen již zmíněný vztah typu více–více mezi pojmy a znaky, který se projevuje jako synonymie a homonymie, ale i obecně známý fenomén vágnosti a různá míra ustálenosti (stability) významu zejména v přirozeném jazyce. Sám termín ontologie se svými různými významy proměňujícími se v čase je toho konkrétním příkladem. Dohoda uživatelů ontologie o významu používaných znaků (slov, termínů) je proto jedním z nejvýznamnějších typů ontologických závazků.

Jako stavební bloky pro vlastní pracovní definici ontologie pro účely tohoto textu použijeme následující pojmy, jež představují pojmovou základnu systémové analýzy (viz kapitola 1.5.3)<sup>148</sup>: třída, prvek, atribut a hodnota, doplněné o specifické ontologické pojmy axiom a ontologický závazek. Význam použitých pojmů vymežíme pomocí těchto výroků:

#### **Třída**

Výsledek konceptualizace. Obsahem třídy jsou vlastnosti zahrnutých objektů, rozsahem třídy je množina objektů se stejnými vlastnostmi, tj. instancí.

Poznámka: Takto definovaná třída je ekvivalentní s pojmem.

#### **Instance**

Objekt, jenž má vlastnosti třídy, naplněné konkrétními hodnotami (daty).

#### **Konceptualizace**

Pojmová reprezentace skutečnosti.

#### **Vyjádření**

Zastoupení (specifikace) pojmu znakem (tj. jeho označení).

#### **Pojmový model**

Model vymezené domény vytvořený pomocí pojmů a jejich vztahů.

#### **Vztah**

Typ vlastnosti, sdílený více entitami (tj. třídami nebo instancemi).

#### **Axiom**

Výrok, tvrzení o třídě, které se nedokazuje a jehož účelem je upřesnit její význam (sémantiku).<sup>149</sup>

#### **Ontologické závazky**

Předpoklady, pravidla a omezení uplatňovaná při konceptualizaci a konstrukci ontologie.

S pomocí výše uvedených pojmů a jejich definic zkonstruujeme klasickou aristotelskou definici informatické ontologie prostřednictvím nejbližšího nadřazeného pojmu a určením druhových specifik.

**Nejbližším nadřazeným pojmem** pro informatickou ontologii je pojmový model.

**Účelem** informatické ontologie je komunikace, opakované využití a organizace znalostí, popis skutečnosti pro počítačové zpracování a automatické odvozování znalostí.

**Obsah (složení):** Jádrem každé informatické ontologie jsou třídy a jejich vlastnosti.

**Specifické vlastnosti** informatické ontologie jsou vyjádřeny následujícími výroky:

Informatická ontologie je inženýrský artefakt navržený s předem stanoveným cílem.

---

<sup>148</sup> Poznámka: Dalším používaným způsobům označení těchto pojmů jsme se věnovali v kapitole 1.5.2.

<sup>149</sup> Poznámka: Takovými axiomy jsou v naší ‚ontologii informatické ontologie‘ pojmy ‚data‘, ‚pojmem‘, ‚objekt‘, ‚množina‘, ‚vlastnost‘.

Informatická ontologie je opakovaně použitelná.<sup>150</sup>

Informatická ontologie je sdílena nějakým rozsáhlejším okruhem uživatelů.<sup>151</sup>

Výběr tříd, jejich atributů a způsob strukturování informatické ontologie je založen na implicitně či explicitně stanovených ontologických závazcích.

Význam tříd v informatické ontologii je vymezen pomocí axiomů.

### **Pracovní definice:**

Informatická ontologie je typ pojmového modelu vymezené domény. Jejím účelem je komunikace, opakované využití a organizace znalostí, popis skutečnosti pro počítačové zpracování a automatické odvozování znalostí. Má formu inženýrského artefaktu, navrženého s předem stanoveným cílem. Je opakovaně použitelná a sdílená nějakým rozsáhlejším okruhem uživatelů. Obsah informatické ontologie tvoří třídy a jejich vlastnosti. Význam tříd je vymezen pomocí axiomů. Výběr tříd, jejich atributů a způsob strukturování ontologie je založen na implicitně či explicitně stanovených ontologických závazcích.

Poznámka: Protože v dalším textu práce už nebudeme uvažovat problematiku filozofických ontologií, budeme nadále v zájmu stručnosti používat pro označení informatických ontologií termín ‚ontologie‘.

---

<sup>150</sup> Tj. jedna ontologie může být využita pro více účelů.

<sup>151</sup> Tj. ontologie je výsledkem nějakého společenského konsenzu.

## Kapitola 2 Metody použité při návrhu ontologie organizace znalostí

*Za souborného nazírání uvádět pod jednu způsobu věci porůznu rozptýlené, aby vymezením každého pojmu činil jasným, o čem by pokaždé chtěl poučovat... Býti schopen zase naopak rozdělovat podle pojmových druhů na přirozené členy a snažit se nezломit přitom žádnou část po způsobu špatného kuchaře.*

Platón (427–347 př. n. l.), *Faidros*, 265 d, e<sup>152</sup>

V této kapitole jsou specifikovány metody, jež byly použity při návrhu ontologie organizace znalostí. První verze ontologie, na jejímž zpracování jsem se podílela v rámci řešení grantového projektu NAKI *Znalostní báze pro obor organizace informací a znalostí*, byla navržena podle metodiky *UPON*. Její fasetová struktura byla vytvořena podle metodiky Williama Dentona a Louise Spiteriové s uplatněním metody *5W1H*. Druhou verzi ontologie jsem navrhla samostatně za použití metod doménové analýzy, fasetové analýzy a opětovného použití základní ontologie. Při konstrukci první verze ontologie byl uplatněn postup ‚od nuly‘ směrem *bottom-up*, tj. od shromáždění termínů přes konceptualizaci k strukturování a fasetizaci. Druhá verze ontologie je výsledkem reengineeringu první verze a byla vytvořena způsobem *top-down*: konceptualizací výroků získaných doménovou analýzou organizace znalostí byly získány ‚nejvyšší‘ pojmy, které byly integrovány se základní ontologií *DOLCE*. Tato jádrová ontologie byla následně propojena s terminologickou ontologií, jež je založena na upravené první verzi ontologie.

Obsah kapitoly je členěn na čtyři části. První část shrnuje charakteristiky metod ontologického inženýrství a její podkapitoly jsou věnovány detailnějšímu popisu metodik aplikovaných v návrhu ontologie: *UPON*, *OntoClean* a návrhové vzory. V druhé části je popsána metoda doménové analýzy. Protože tato metoda je v českém prostředí zatím málo diskutovaná<sup>153</sup>, je její charakteristika podána poněkud obšírněji. Třetí část kapitoly je věnována fasetové analýze. Podrobněji jsou charakterizovány metody fasetové analýzy, jež byly využity při návrhu ontologie organizace znalostí: metodika zpracovaná Williamem Dentonem a Louise Spiteriovou a metoda *5W1H*. Závěrečná diskuse ve čtvrté části je zaměřena na omezení systémového a analytického přístupu se zřetelem na jejich aplikaci v klíčovém procesu ontologického inženýrství, jímž je konceptualizace, a na vzájemný vztah metodik ontologického inženýrství a organizace znalostí.

### 2.1 Metodiky ontologického inženýrství<sup>154</sup>

Již bylo konstatováno, že informatické ontologie na rozdíl od ontologií filozofických nejsou výsledkem vědeckého zkoumání, ale jsou to navrhované a konstruované artefakty. Přístup k návrhu informatických ontologií prodělal analogický vývoj jako v případě počítačových programů a znalostních systémů, jen s cca třicetiletým časovým posunem: od původního pojetí návrhu ontologie jako umění či vědy, dostupných jen několika výjimečně disponovaným jedincům, k inženýrství. Analogicky k softwarovému a znalostnímu inženýrství se zformovala nová inženýrská disciplína – ontologické inženýrství. Ontologické inženýrství zahrnuje soubor aktivit, jež se týkají procesu vývoje, životního cyklu a metod tvorby (konstrukce) ontologií a

<sup>152</sup> Citováno dle: Platón. *Faidros*. Z řeckého originálu přeložil František Novotný. 4. oprav. vyd. Praha: Oikoymenth, 1993, s. 58.

<sup>153</sup> Poznámka: Obecné konstatování z úvodního průzkumu o nedostatku odborných textů z oboru organizace znalostí v češtině lze vztáhnout i na metodiku doménové analýzy.

<sup>154</sup> Text této kapitoly je zkrácenou a aktualizovanou verzí článku: KUČEROVÁ, Helena. Metodiky ontologického inženýrství. In: *Ikaros* [online]. 2011, 15(5) [cit. 2018-01-30]. urn:nbn:cz:ik-13630. ISSN 1212-5075. Dostupné z: <http://www.ikaros.cz/node/13630>.

pomůcek (například jazyků), jež tyto aktivity podporují. Kromě obvyklých obecných inženýrských procesů (analýza, návrh, testování, implementace, správa) se ontologické inženýrství zaměřuje i na specifické oblasti mapování, integrace, sdílení a opětovného použití ontologií. Analogicky ke CASE systémům v softwarovém inženýrství už jsou k dispozici i specializované softwarové nástroje podporující návrh ontologií (*Protégé, Ontopia* ad.).

V ontologickém inženýrství se uvažují čtyři typy aktérů. Kromě ontologického/znalostního inženýra se zpravidla vývoje účastní ještě doménový specialista / odborník a ve stále větší míře i koncový uživatel – osoba. Dalším uvažovaným subjektem je počítačová aplikace (agent) v roli uživatele ontologie.

Stejně jako v softwarovém inženýrství se aktivity ontologického inženýrství zahajují specifikací požadavků a jejich konkretizací formou případů užití. Podle terminologické normy ISO/IEC 2382 je požadavek (angl. *requirement*) „základní podmínka, kterou musí systém splňovat“.<sup>155</sup> Případem užití rozumíme zaznamenaný požadavek, určující, co má systém vykonávat či jak se má chovat vzhledem ke konkrétnímu aktérovi (např. k uživateli). Případy užití se často specifikují formou scénářů pro používání systému; každý scénář obsahuje sekvenci (posloupnost) událostí, které v jeho rámci probíhají. Specifickým typem případu užití v ontologickém inženýrství je zodpovídání tzv. kompetenčních otázek (angl. *competency questions*). Pojem kompetenční otázky zavedli do ontologického inženýrství v rámci své metodiky *TOVE* Michael Grüninger a Mark S. Fox.<sup>156</sup> Definují je jako otázky, na něž má být ontologie schopna poskytovat odpovědi, a doporučují, aby každý objekt v ontologii byl spojen s nějakou otázkou, k jejímuž zodpovězení má přispět. Aldo Gangemi s Valentinou Presuttiovou dále upřesňují, že kompetenční otázka je „typický dotaz, který by odborník mohl chtít položit za určitým účelem znalostní bázi ze své cílové domény. Správná doménová ontologie by měla specifikovat všechny a zároveň pouze ty konceptualizace, potřebné k zodpovězení kompetenčních otázek formulovaných odborníky nebo získaných od nich.“<sup>157</sup>

První krok směrem ke standardizaci tvorby ontologií učinil v roce 1995 Thomas Gruber. Zformuloval obecná kritéria, jež doporučil sledovat během návrhu ontologií, a jejich použitelnost doložil dvěma případovými studiemi.<sup>158</sup>

1. Jasnost, srozumitelnost: Každý pojem reprezentovaný termínem má mít objektivní (jednoznačnou, přesnou a kontextově nezávislou) definici. Preferují se formální definice tvořené logickými axiomy. Definice mají být úplné, tj. mají splňovat jak logické podmínky nutnosti, tak podmínky postačitelnosti.

2. Soudržnost: Povoleno je jen takové odvozování, které je v souladu s obsahem definic termínů.

---

<sup>155</sup> ISO/IEC 2382:2015. *Information technology – Vocabulary* [online]. 1st ed. Geneva: International Organization for Standardization, 2015-05 [cit. 2018-01-30]. Dostupné z: <https://www.iso.org/obp/ui/#iso:std:iso-iec:2382:ed-1:v1:en:term:2122664>.

<sup>156</sup> GRÜNINGER, Michael, FOX, Mark S. Methodology for the design and evaluation of ontologies. In: *Workshop on basic ontological issues in knowledge sharing, held in conjunction with IJCAI-95* [online], Montreal, 1995, s. 1-10 [cit. 2018-01-30]. Dostupné z: <http://www.eil.utoronto.ca/enterprise-modelling/entmethod/index.html>.

<sup>157</sup> GANGEMI, Aldo, PRESUTTI, Valentina. Ontology design patterns. In: STAAB, Steffen, STUDER, Rudi, ed. *Handbook on ontologies*. 2nd ed. Dordrecht: Springer, 2009, s. 230-231. [https://doi.org/10.1007/978-3-540-92673-3\\_10](https://doi.org/10.1007/978-3-540-92673-3_10). ISBN 978-3-540-70999-2 (Print). ISBN 978-3-540-92673-3 (Online).

<sup>158</sup> GRUBER, Thomas R. Toward principles for the design of ontologies used for knowledge sharing. In: *International journal of human-computer studies*. November/December 1995, **43**(5-6), 907-928. doi:10.1006/ijhc.1995.1081. ISSN 1071-5819.

3. Rozšiřitelnost: Rozšiřování a specializace ontologie má umožnit monotónní usuzování, tj. definování nových termínů nemá vyžadovat revizi existujících definic.

4. Minimální závislost na znakovém (kódovém) systému: Konceptualizace, tj. pojmová reprezentace skutečnosti, má být provedena na úrovni znalostí a nemá být omezována zvoleným typem notace nebo kódování.

5. Minimální ontologické závazky: Termíny v ontologii mají mít svůj význam co nejméně omezen, pouze v míře nezbytně nutné pro komunikaci.

Tato kritéria patří k základním principům ontologického inženýrství. Nezapřou své kořeny v pravidlech klasické formální logiky. Dalším rysem, který zase odkazuje na principy organizace znalostí v knihovnách, je orientace na flexibilitu a jednoduchost systému. Požadavek minimálních ontologických závazků je analogický požadavku Seymoura Lubetzkyho z roku 1953 na zjednodušení katalogizačních pravidel, formulovanému slavnou otázkou: „*Jsou všechna tato pravidla nezbytná?*“<sup>159</sup>

Stejně jako ve vědeckém výzkumu, i v inženýrském návrhu mají klíčovou roli metodiky.

V přístupu k typologii metodik ontologického inženýrství lze uplatnit poměrně značné množství pohledů a jejich typologie je tak vhodnou oblastí pro aplikaci fasetové analýzy (viz kapitola 2.3). Jednotlivé fasety a jejich obsah jsou uvedeny v přehledu v tabulce 4 a v následujícím komentáři.

Využití softwaru	<ul style="list-style-type: none"> <li>▪ intelektuální</li> <li>▪ poloautomatizovaná</li> <li>▪ automatizovaná</li> </ul>
Postup řešení	<ul style="list-style-type: none"> <li>▪ top-down</li> <li>▪ middle-out</li> <li>▪ bottom-up</li> </ul>
Životní cyklus	<ul style="list-style-type: none"> <li>▪ vodopádový</li> <li>▪ iterativní</li> </ul>
Směr vývoje	<ul style="list-style-type: none"> <li>▪ dopředný</li> <li>▪ zpětný</li> </ul>
Účel metodiky	<ul style="list-style-type: none"> <li>▪ plánovací</li> <li>▪ konstrukční</li> </ul>
Pravidla	<ul style="list-style-type: none"> <li>▪ rigorózní</li> <li>▪ agilní</li> </ul>
Formát	<ul style="list-style-type: none"> <li>▪ vědecký zákon</li> <li>▪ odborný text</li> <li>▪ legislativní norma</li> <li>▪ technický standard</li> </ul>
Zaměření	<ul style="list-style-type: none"> <li>▪ metodiky pro systémy vytvářené ‚na zelené louce‘</li> <li>▪ metodiky opětovného použití</li> <li>▪ metodiky reengineeringu</li> <li>▪ metodiky mapování</li> <li>▪ metodiky vývoje</li> <li>▪ metodiky učení</li> <li>▪ metodiky hodnocení</li> </ul>
Zdroj	<ul style="list-style-type: none"> <li>▪ umělá inteligence, znalostní inženýrství</li> <li>▪ softwarové inženýrství</li> <li>▪ knihovní a informační věda</li> </ul>

Tab. 4 Fasetová typologie metodik ontologického inženýrství

<sup>159</sup> LUBETZKY, Seymour. *Cataloging rules and principles: a critique of the A.L.A. rules for entry or entry and a proposed design for their revision*. Washington: Library of Congress, 1953, s. 1.



## Faseta Využití softwaru

Metodiky návrhu a tvorby ontologií lze rozdělit na intelektuální (manuální) a automatizované, tj. realizované s použitím softwarových aplikací, jež jsou často vybavené nástroji umělé inteligence a umožňují například zpracování přirozeného jazyka nebo formální konceptuální analýzu. V úvahu připadají ještě metodiky poloautomatizované, kombinující oba postupy.

## Faseta Postup řešení

Dalším přístupem k členění metodik je postup uplatněný při řešení problémů. Rozlišuje se mezi postupem *top-down* (shora dolů) a *bottom-up* (zdola nahoru). Norma ISO/IEC 2382 definuje tyto metody v souvislosti s metodami analýzy nebo syntézy následovně: *top-down* začíná na nejvyšší úrovni abstrakce a směřuje k nejméně abstraktním úrovním, *bottom-up* začíná u nejnižší úrovně abstrakce a směřuje k nejabstraktnější úrovni.<sup>160</sup> V současných projektech se lze setkat i s postupem *middle-out* (od středu ven), který vychází od nejdůležitějších prvků a postupuje oběma směry k nadřazeným a podřazeným úrovním. V případě návrhu ontologií se tímto způsobem specifikuje postup při konceptualizaci. Postup *top-down* je deduktivní – celá doména se rozčlení na několik základních kategorií a ty se postupně naplňují konkrétními pojmy, resp. termíny. Při postupu *bottom-up* se uplatňuje indukce – nejprve se shromáždí všechny termíny a ty se pak seskupují do kategorií. Postup *middle-out* vychází z pojmů považovaných za nejvýznamnější, ty se indukci zobecňují do vrcholových kategorií a dedukcí specializují na konkrétní termíny zachycující detaily dané domény.

## Faseta Životní cyklus

Z hlediska životního cyklu uplatněného při návrhu je možné se setkat s různými kategorizacemi, mezi odborníky nepanuje úplná shoda ani ve významu jednotlivých termínů. Pro naše účely vyjdeme z přístupu Martina Fowlera<sup>161</sup>, který rozlišuje dva základní typy: metodiky uplatňující vodopádový model<sup>162</sup> (případně jeho varianty – fontánový model nebo V-model), a metodiky iterativní (a jejich varianty označované jako spirálový model<sup>163</sup>, přírůstkový neboli inkrementální model, evoluční model, prototypový model). Vodopádový model člení projekt podle činností (typicky na specifikaci požadavků, analýzu, návrh, testování a implementaci), jež za sebou následují v pevně stanoveném pořadí. Návraty k předchozím fázím se připouštějí, je však snaha je minimalizovat. Iterativní model je založen na plánovaném opakování procesů. Vychází se z hypotézy, že opakovaným řešením (verzováním, prototypováním) lze dospět ke kvalitnějšímu výsledku. Zpravidla se před zahájením iterací navrhne základní architektura systému, a poté se přírůstkově řeší jednotlivé části (komponenty, funkcionality). Výstupem každé iterace má být nový přírůstek rozsahu a kvality projektu v podobě použitelné části navrhovaného systému. V průběhu dalších iterací se nejen řeší další přírůstek, ale znovu se

---

<sup>160</sup> ISO/IEC 2382:2015. *Information technology – Vocabulary* [online]. 1st ed. Geneva: International Organization for Standardization, 2015-05 [cit. 2018-01-30]. Dostupné z: <https://www.iso.org/obp/ui/#iso:std:iso-iec:2382:ed-1:v1:en>.

<sup>161</sup> FOWLER, Martin. *Destilované UML*. 1. vyd. Praha: Grada, 2009. Kapitola 2, Proces vývoje, s. 37-49. ISBN 978-80-247-2062-3.

<sup>162</sup> ROYCE, Winston W. Managing the development of large software systems. In: *Proceedings of IEEE WESCON*. August 1970, 26, s. 1-9.

<sup>163</sup> BOEHM, Barry. A spiral model of software development and enhancement. In: *IEEE Computer*. May 1988, 21(5), 61-72. ISSN 0018-9162. Dostupné z: <http://csse.usc.edu/csse/TECHRPTS/1988/usccse88-500/usccse88-500.pdf> [cit. 2018-01-30].

revidují už dokončené části.

### **Faseta Směr vývoje**

Nejobvyklejším způsobem inženýrského návrhu je postup směrem *top-down* od modelu k implementaci – tzv. dopředné inženýrství (angl. *forward engineering*). Pokud se z hotové funkční aplikace ve směru *bottom-up* dodatečně odvozuje abstraktní pojmový model, hovoří se o tzv. zpětném inženýrství (angl. *reverse engineering*). Oba směry vývoje mohou být v současné době podporovány softwarovými prostředky a do jisté míry automatizovány.

### **Faseta Účel metodiky**

Podle zaměření na konkrétní aktivity v průběhu projektu se metodiky člení na plánovací a konstrukční. Plánovací metodiky obsahují opakovatelné procesy a postupy. Obsahem konstrukční metodiky jsou opětovně použitelné modely či návrhové vzory (viz kapitola 2.1.3), případně pravidla pro jejich konstruování.

### **Faseta Pravidla**

V souvislosti s množstvím a závazností pravidel obsažených v metodice se mluví o metodikách rigorózních (striktních), kladoucích důraz na dobře definované procesy vývoje, a o agilních (angl. *agile, light, lightweight*) metodikách, jež za nejdůležitější faktor úspěchu projektu považují kvalitu zúčastněných osob a jejich efektivní spolupráci. Agilní metodiky se snaží minimalizovat regulaci procesů ve prospěch kreativních ad hoc řešení a pružných reakcí na změny v průběhu projektu.

### **Faseta Formát**

Kromě obsahu je možné uvažovat i formáty, v nichž jsou metodiky zaznamenány. Ty mohou mít rozpětí od vědeckého zákona přes odborné texty až po pravidla a standardy na různé úrovni závaznosti. V úvahu připadají jak dobrovolné technické normy – například *ISO* nebo *ČSN*, tak legislativní normy – např. zákony jednotlivých zemí či direktivy *EU*. Podle rozsahu působnosti je možné se setkat s pravidly a standardy na mezinárodní, národní, odvětvové úrovni (např. vydávané organizacemi *W3C, OMG, IFLA, CIDOC, IEEE*), s podnikovými předpisy a s tzv. *best practices* (příklady dobré praxe).

### **Faseta Zaměření**

Metodiky ontologického inženýrství se už stačily typologicky rozvrstvit podle toho, které specifické úlohy při vývoji systému plní. Kromě kompletního vývoje celého systému ‚od nuly‘, ‚na zelené louce‘, ‚tabula rasa‘ (angl. *from scratch*) jsou k dispozici i četné další typy metodik. Vojtěch Svátek a Miroslav Vacura navrhuje v rámci této fasety kategorizaci metodik podle klíčových technologií do tří skupin: strukturované metodiky, metodiky použití základních ontologií a metodiky založené na automatické extrakci z textů (ty lze považovat za podmnožinu kategorie automatizovaných metodik z fasety Využití softwaru).<sup>164</sup>

Příkladem strukturované metodiky, určené pro vývoj ‚na zelené louce‘, je metodika *UPON*, která bude detailně představena v kapitole 2.1.2.

Metodiky opětovného použití nabízejí postupy, jak začlenit do nové ontologie již existující

---

<sup>164</sup> SVÁTEK, Vojtěch, VACURA, Miroslav. Ontologické inženýrství. In: Lubomír POPELÍNSKÝ, Ondřej VÝBORNÝ, ed. *DATAKON 2007: sborník databázové konference: Brno, Česká republika, 20.-23. října 2007*. Brno: Masarykova univerzita, 2007, s. 74-80. ISBN 978-80-7355-076-9.

ontologii či její část. Příkladem takové metodiky jsou *Pravida pro aplikační profily Dublin Core*.<sup>165</sup> Jako podpůrný nástroj pro uplatnění metodiky použití základní ontologie slouží *ONSET* – nástroj výběru základní ontologie pro návrh doménové ontologie.<sup>166</sup> Nástroj používá metodu vícekritériálního rozhodování. Kritéria jsou odvozena z detailního popisu základních ontologií a jsou seskupena do pěti kategorií. Uživatel má možnost kategoriím přiřadit váhy od 0 do 5, tyto váhy se poté přenášejí na všechna kritéria zahrnutá v kategorii. Jsou-li některá dílčí kritéria vyhodnocena jako nerelevantní, lze je z analýzy vynechat.

Dalším typem jsou metodiky specializované na úlohy reengineeringu, redesignu a refaktoringu<sup>167</sup>, jež se orientují na zásadní přepracování existujícího systému. Lze konstatovat, že zatím nebylo zformulováno mnoho specifických metodik reengineeringu ontologií, k dispozici jsou ovšem četné případové studie. Realizátoři obvykle používají osvědčené metody reengineeringu ze softwarového inženýrství. To je i případ rozšiřujícího modulu metodiky *Methontology* pro její použití při reengineeringu. Samotná metodika *Methontology*, kterou v roce 1997 vytvořila skupina *Ontology Engineering Group z Polytechnické univerzity v Madridu*, je založena na standardu IEEE 1074 pro návrh životního cyklu softwarových procesů.<sup>168</sup> Autorky publikovaného rozšíření pro úlohy reengineeringu uvádějí, že adaptovaly metody reengineeringu softwaru E. J. Chikofskyho.<sup>169</sup> Za cíl procesu reengineeringu označují transformaci pojmového modelu a v rámci tohoto procesu identifikovaly tři obecné aktivity: reverzní (zpětné) inženýrství, restrukturování a dopředné inženýrství. Návrh je doplněn případovou studií reengineeringu ontologie standardních jednotek a ontologie chemických prvků, jež byly založeny na tomto postupu. Dalším příkladem případové studie je popis přepracování tezauru *AGROVOC* do formátu ontologie.<sup>170</sup> Autoři se soustředili na specifikaci vztahů a přímo doporučují formulovat pravidla pro ni až „za běhu“ (*rules as you go*), tj. odvozovat je z rozpoznávaných vzorů při studiu stávajících vztahů v tezauru. Specifickým postupem

---

<sup>165</sup> COYLE, Karen, BAKER, Thomas. *Guidelines for Dublin Core Application Profiles* [online]. Dublin (Ohio): DCMI, 2009-05-18 [cit. 2018-01-30]. Dostupné z: <http://dublincore.org/documents/profile-guidelines/>.

<sup>166</sup> KHAN, Zubeida, KEET, C. Maria. ONSET: automated foundational ontology selection and explanation. In: Annette ten Teije et al., ed. *Knowledge Engineering and Knowledge Management: 18th International Conference, EKAW 2012, Galway City, Ireland, October 8-12, 2012*. Proceedings. Berlin, Heidelberg: Springer-Verlag, 2012, s. 237-251. doi:10.1007/978-3-642-33876-2\_22. Lecture notes in computer science, no. 7603, ISSN 0302-9743 (Print). ISSN 1611-3349 (Online). ISBN 978-3-642-33875-5 (Print). ISBN 978-3-642-33876-2 (Online).

<sup>167</sup> FOWLER, Martin. *Refactoring: zlepšení existujícího kódu*. 1. vyd. Praha: Grada, 2003. 394 s. ISBN 80-247-0299-1.

<sup>168</sup> GÓMEZ-PÉREZ, Asunción, FERNÁNDEZ-LÓPEZ, Mariano, CORCHO, Oscar. METHONTOLOGY. In: *Ontological engineering: with examples from the areas of knowledge management, e-commerce and the Semantic Web*. 1st ed. London: Springer, 2004, kapitola 3.3.5, s. 125-142. doi:10.1007/b97353. ISBN 978-1-85233-551-9 (Print). ISBN 978-1-85233-840-4 (Online).

<sup>169</sup> GÓMEZ-PÉREZ, Asunción, ROYAS-AMAYA, Dolores. Ontological reengineering for reuse. In: Dieter Fensel, Rudi Studer, ed. *Knowledge Acquisition, Modeling and Management: 11th European Workshop, EKAW'99 Dagstuhl Castle, Germany, May 26-29, 1999 Proceedings*. Berlin, Heidelberg: Springer-Verlag, 1999, s. 139-156. doi: [https://doi.org/10.1007/3-540-48775-1\\_9](https://doi.org/10.1007/3-540-48775-1_9). Lecture notes in computer science, no. 1621, ISSN 0302-9743 (Print). ISSN 1611-3349 (Online). ISBN 978-3-540-66044-6 (Print). ISBN 978-3-540-48775-3 (Online).

<sup>170</sup> SOERGEL, Dagobert, LAUSER, Boris, LIANG, Anita, FISSEHA, Frehiwot, KEIZER, Johannes, KATZ, Stephen. Reengineering thesauri for new applications: the AGROVOC example. In: *Journal of digital information* [online]. 2004, 4(4) [cit. 2018-01-30]. ISSN 1368-7506. Dostupné z: <http://journals.tdl.org/jodi/article/view/112/111>.

ontologického inženýrství je reengineering s využitím základní ontologie. V případové studii reengineeringu doménové ontologie *SEO (Software enterprise ontology)* popisují Monalessa Perini Barcellos a Ricardo de Almeida Falbo využití základní ontologie *UFO* k tomuto účelu.<sup>171</sup>

Metodiky mapování se zaměřují na procesy vedoucí k sémantickému propojení vzájemně nezávislých ontologií. Metodiky vývoje či učení existujících systémů se věnují postupnému doplňování ontologie o další pojmy a vztahy, často realizovanému s podporou specializovaného softwaru. Specifickou kategorií představují také metodiky pro hodnocení ontologií, jejichž zástupcem je metodika *OntoClean*, která bude detailně představena v kapitole 2.1.1.

### Faseta Zdroj

Metodiky návrhu informatických ontologií zpravidla využívají osvědčené metody příbuzných oborů, nejčastěji umělé inteligence a znalostního inženýrství, softwarového inženýrství a knihovní a informační vědy.

Umělá inteligence a znalostní inženýrství se zaměřují na návrh znalostních a expertních systémů, případně znalostníchází. Jejich klíčovými metodikami jsou logická pravidla (angl. *rules*) a rámce (angl. *frames*)<sup>172</sup>. Logická pravidla reprezentují procedurální znalosti pomocí produkčních pravidel typu předpoklad – závěr (angl. *if-then*). Rámce jsou systémem reprezentace objektů v bázi znalostí prostřednictvím položek (angl. *slots*), sloužících k popisu jejich vlastností. Značný význam mají vyspělé metody získávání znalostí (angl. *knowledge acquisition*)<sup>173</sup>, jež pomáhají v etapě, která je obecně považována při návrhu ontologie za kritickou. Patří mezi ně techniky získávání (elicitace) znalostí od expertů (interview, *brainstorming*, třídění kartiček, repertoárové tabulky) a automatizované nebo poloautomatizované získávání znalostí z dat a z textů (např. *data mining*, *text mining*). Další metodou ověřenou v této oblasti a uplatnitelnou při návrhu ontologií je formální konceptuální analýza (angl. *formal concept analysis*)<sup>174</sup> – metodika automatické derivace ontologií z datových objektů a jejich vlastností.

Softwarové inženýrství je disciplína zaměřená na návrh softwaru, případně informačního systému (systémové inženýrství, systémová integrace). Typickým produktem je softwarová aplikace nebo databáze. Soustředí se zpravidla na unikátní aplikace, navrhované na míru specifickým konkrétním podmínkám a uživatelským potřebám (tzv. *usability* – individuální

---

<sup>171</sup> PERINI BARCELLOS, Monalessa, DE ALMEIDA FALBO, Ricardo. Using a foundational ontology for reengineering a Software enterprise ontology. In: Heuser C. A., Pernul G., ed. *Advances in conceptual modeling - challenging perspectives*. ER 2009. Berlin, Heidelberg: Springer, 2009, s. 179-188. doi: [https://doi.org/10.1007/978-3-642-04947-7\\_22](https://doi.org/10.1007/978-3-642-04947-7_22). Lecture Notes in Computer Science, no. 5833, ISSN 0302-9743 (Print). ISSN 1611-3349 (Online). ISBN 978-3-642-04946-0 (Print). ISBN 978-3-642-04947-7 (Online).

<sup>172</sup> MINSKY, Marvin. *A framework for representing knowledge* [online]. Cambridge: MIT Media Laboratory, 1974 [cit. 2018-01-30]. MIT-AI Laboratory Memo 306, June, 1974. Dostupné z: <http://web.media.mit.edu/~minsky/papers/Frames/frames.html>.

<sup>173</sup> BERKA, Petr, MAŘÍK, Vladimír, SVÁTEK, Vojtěch. Proces získávání znalostí. In MAŘÍK, Vladimír, ŠTĚPÁNKOVÁ, Olga a LAŽANSKÝ, Jiří. *Umělá inteligence 2*. 1. vyd. Praha: Academia, 1997, s. 107-124. ISBN 80-200-0504-8.

<sup>174</sup> PRISS, Uta. Formal concept analysis in information science. In: *Annual review of information science and technology*. Vol. 40. Blaise CRONIN, ed. Medford (N.J.): Information Today for American Society for Information Science & Technology (ASIST), © 2006, chapter 13, s. 521-543. Annual review of information science and technology, 44, ISSN 0066-4200 (Print). ISSN 1550-8382 (Online). doi:10.1002/aris.1440400120. ISBN 978-1-57387-242-3.

použitelnost). Cestou směrem k opětovné použitelnosti osvědčených řešení jsou návrhové vzory (o nich podrobněji v kapitole 2.1.3). Softwarové inženýrství má k dispozici propracované metodiky datové analýzy (např. ER modelování, diagramy tříd) a metodiky životního cyklu softwarového projektu, např. *RUP – Rational Unified Process*, *UP – Unified Process*, *Scrum*, *Crystal*.

Metodiky knihovní a informační vědy se zpravidla zaměřují na dosažení širokého konsenzu, jejich ambicí je tedy na rozdíl od softwarového inženýrství opakovatelná použitelnost (angl. *reusability*). K tomuto účelu mají často vypracovaný velmi důkladný a obsáhlý soubor pravidel, uplatňovaných například při tvorbě slovníku. Využitelné metodiky jsou jak kvantitativní (bibliometrie, citační analýza), tak kvalitativní, založené na analýze obsahu. Významným přínosem informační vědy jsou fasetová analyticko-syntetická metoda (viz kapitola 2.3), metoda doménové analýzy (o ní blíže v kapitole 2.2), v praxi ověřené metody automatické obsahové analýzy a indexace<sup>175</sup> a četné metodiky tvorby tezauru. V neposlední řadě přispěla knihovní věda k propracování principu označovaného jako tzv. *warrant* (viz kapitola 4.2), který by bylo možné interpretovat jako ‚ospravedlnění‘ přijatých ontologických závazků.

### 2.1.1 Principy analýzy a metodika *OntoClean*

Obecným rámcem všech typů metod ontologického inženýrství zmiňovaných v předchozí části je systémový přístup a jejich obecnou metodologií je systémová analýza. Význam systémového přístupu pro konceptualizaci byl již konstatován v kapitole 1.5.3, nyní se zaměříme na rozšíření pohledu na celou oblast ontologického inženýrství.

Význam systémového přístupu, jeho teoretické báze v podobě systémové vědy a aplikace v podobě systémové analýzy konstatují četní autoři. Univerzální charakter systémové analýzy umožňuje její aplikaci prakticky v jakémkoli oboru a je tak významnou podporou interdisciplinarity a transdisciplinarity. Systémová věda, jež má sama interdisciplinární charakter, tvoří společný sjednocující rámec, umožňující komunikaci mezi speciálními vědními disciplínami. Kenneth Boulding obrazně nazývá systémovou vědu kostrou vědy, jejímž cílem je „poskytnout rámec nebo strukturu systémů, na něž se zavěsí svalový a oběhový systém jednotlivých disciplín a konkrétních předmětů do řádného a koherentního souboru znalostí“. Zároveň upozorňuje na to, že systémová věda se může stát i ‚kostlivcem ve skříni‘ – „skříni v tomto případě je neochota vědy přiznat velmi nízkou úroveň svých úspěchů v systematizaci a její tendence zavírat dveře před problémy a předměty, které nezapadají snadno do jednoduchých mechanických schémat.“<sup>176</sup> Jaroslav Habr a Jaromír Vepřek rovněž konstatují, že „...roste vliv teorie obecných systémů na jiné systémové disciplíny, které tato teorie ovlivňuje jak terminologicky, tak metodologicky. Pro všechny systémové disciplíny mají značný význam práce vedené v teorii obecných systémů na přípravě formálních metajazyků vhodných pro popis pojmů a vztahů mezi nimi.“<sup>177</sup> Je zjevné, že pojmová báze informatických ontologií je odvozena z pojmového aparátu systémové vědy. Jak bylo uvedeno, systémová věda poskytuje i teoretický a metodologický základ pro konceptualizaci v rámci ontologického inženýrství. Význam systémové vědy pro obor organizace znalostí vyzdvihuje Douglas J. Foskett, který je přesvědčen, že spojením obsahové

---

<sup>175</sup> SCHWARZ, Josef. Současný stav a trendy automatické indexace dokumentů: přehledová studie. *Ikaros* [online]. 2003, 7(3) [cit. 2018-01-30]. ISSN 1212-5075. Dostupné z: <http://www.ikaros.cz/node/1300>.

<sup>176</sup> BOULDING, Kenneth E. General systems theory – the skeleton of science. In: *Management science*. April 1956, 2(3), s. 208.

<sup>177</sup> HABR, Jaroslav, VEPŘEK, Jaromír. *Systémová analýza a syntéza: zdokonalování a projektování systémů*. 2. přeprac. vyd. Praha: SNTL, 1986, s. 52.

analýzy a klasifikace se systémovou vědou lze pokrýt veškeré univerzum poznání.<sup>178</sup>

Proces analýzy, který je jádrem systémové analýzy, se na nejobecnější úrovni chápe jako „základní myšlenkový postup, který rozkládá vymezený celek na jeho prvky ... Klasicky se rozlišuje mezi analýzou ‚látkovou‘ (např. v chemii), analýzou ‚gramatickou‘ a analýzou ‚ideovou‘ či ‚mentální‘ (analýzou nějakého pojmu). V širším smyslu nazýváme analýzou každou metodu, jež se snaží předmět nebo událost vysvětlit z jeho složek.“<sup>179</sup> Způsoby správného analytického uvažování na vysoké úrovni abstrakce tradičně zkoumá logika. V průběhu používání analytické metody v nejrůznějších oblastech se vyprofilovala řada pravidel, jejichž dodržování zvyšuje pravděpodobnost kvality a využitelnosti výsledků analýzy. Jedná se v podstatě o požadavky na ideální třídy/kategorie/pojmy, jež by měly být jasně ohraničené, co do způsobu dělení úplné (vyčerpávající) a vzájemně se vylučující, co do míry souvislosti zahrnutých prvků soudržné, vytvořené podle jednotného kritéria členění, a homogenní, tj. se stejnou úrovní granularity a stejnou úrovní abstrakce. Při uplatňování principů správné analýzy v oboru ontologického inženýrství jde nejčastěji o jejich aplikaci v průběhu návrhu páteřní taxonomie, jejímž základem je konceptualizace (tj. tvorba ontologických tříd) a definování rododruhových vztahů, tj. generické hierarchie. Požadavky na ‚ideální‘ systém lze kromě výše uvedených požadavků odvozených z logiky ještě doplnit o následující pragmatičtější stanovené vlastnosti: interoperabilita, sémantická síla, jednoduchá struktura, flexibilita.

### Jasně ohraničení

Tento požadavek se týká určení hranic analyzovaného celku i jeho částí. Každou třídu je třeba jednoznačně definovat v prostoru (rozsahem) a v čase. Hledáme tedy odpovědi na otázky: Které prvky do třídy patří a které ne? V jakém období / při jaké události je prvek součástí třídy? Například pro třídu ‚Čtenář‘ mohou být k dispozici tato vymezení: Čtenář je ten, kdo umí číst. Čtenář je ten, kdo právě čte. Čtenář je ten, kdo je registrovaným uživatelem knihovny. Čtenář je ten, kdo předčítá text audioknihy. Je zjevné, že každé z vymezení ohraničí jinou množinu objektů. Tradičním nástrojem jednoznačného vymezení jsou definice, v ontologickém inženýrství plní tuto úlohu axiomy.

Při striktním vymezení členství ve třídě lze na správně definované třídy uplatnit požadavek splnění logické podmínky nutnosti a logické podmínky postačitelnosti. Podmínka nutnosti znamená: A platí jen tehdy, platí-li současně B. B je tedy nutnou podmínkou pro A. B implikuje A,  $B \Rightarrow A$ . Podmínka postačitelnosti znamená: B platí jen tehdy, platí-li současně A. Pokud jsou splněny obě podmínky současně, je vymezení třídy tzv. úplné. A je ekvivalentní B,  $A \Leftrightarrow B$ . Třídy splňující podmínky nutnosti a postačitelnosti jsou někdy označovány jako monotetické, na rozdíl od tříd polytetických, v nichž tyto striktní podmínky nejsou uplatněny.

### Úplnost

Části mají v úplnosti zachycovat celý rozsah analyzovaného celku. Úplné (angl. *complete, exhaustive*) členění zaručuje, že jsou vyjmenovány všechny prvky, které mohou příslušet do dané třídy. Jinými slovy, všechny instance či exemplifikace třídy musí být obsaženy v některé z jejích částí. V neúplném (angl. *incomplete*) členění může nastat i výskyt takového prvku, který není

---

<sup>178</sup> FOSKETT, Douglas J. Systems theory and its relevance to documentary classification. In: *International classification: a journal devoted to concept theory, organization of knowledge and data, and to systematic terminology*. 1980, 7(1), 2-5.

<sup>179</sup> DUROZOI, Gérard, ROUSSEL, André. *Filozofický slovník*. Přeložili Jan Binder aj. 1. vyd. Praha: EWA, 1994, s. 11. ISBN 80-85764-07-5.

předem definován.

Aplikace úplného a neúplného členění v ontologickém inženýrství vyplývá z přijetí logického předpokladu uzavřeného nebo otevřeného světa, již zmíněného v kapitole 1.5.5. Předpoklad uzavřeného světa v logice považuje vše, co není vyjmenováno, za nepravdivé. Předpoklad otevřeného světa připouští, že mohou existovat ještě jiné prvky třídy než ty, které jsou v ní uvedeny, a považuje je z hlediska pravdivosti za nerozhodnutelné. To má své důsledky ve způsobu odvozování (usuzování) v takto koncipovaném systému – v tzv. monotónním a nemonotónním usuzování. Zatímco monotónní usuzování vyžaduje při doplnění nových prvků zachovat beze změny stávající znalosti, nemonotónní usuzování připouští, že na základě nově přidávaných prvků mohou být dosavadní prvky revidovány jak z hlediska svého obsahu, tak co do své příslušnosti ke třídě.

Pragmatickým způsobem řešení problému neúplného členění v praxi je doplnění třídy ‚ostatní‘, ‚jiné‘ či ‚různé‘, určené k zatřídění všech nevyjmenovaných prvků.

### **Soudržnost**

Členění má být sémanticky homogenní a konzistentní. V každé třídě mají být jen prvky, jejichž vlastnosti spolu těsně souvisejí. Soudržnost (koheze, koherence, bezespornost) má panovat jak uvnitř částí, tak mezi nimi navzájem, tj. ve vztazích.

### **Jednotné kritérium členění**

Požadavku na jednotné kritérium členění je zpravidla obtížné vyhovět u komplexních objektů. Jeho splnění u složitého analyzovaného celku umožňuje technika fasetové analýzy (viz kapitola 2.3).

### **Vzájemně se vylučující části**

Fenomén vylučujícího (disjunktního, angl. *disjoint, mutually exclusive, non-overlapping*) a překrývajícího se členění má své teoretické základy ve formální logice, v booleovské algebře a v množinové teorii. Vychází z principů, na nichž jsou založeny logické spojky disjunkce (konkrétně tzv. vylučující disjunkce označovaná symbolem  $\vee$ , resp. vzájemně disjunktní množiny, tj. ty, jejichž průnik označovaný symbolem  $\cap$  představuje prázdná množina) a konjunkce (označovaná symbolem  $\wedge$ ). V disjunktním členění platí pravidlo ‚buď/nebo‘ (označované symbolem XOR) – prvek může být členem pouze jedné třídy.

V rámci požadavku na vylučující se části lze ještě rozlišovat mezi vylučováním obsahovým (intenzí) a rozsahovým (extenzí). V prvním případě požadujeme vylučující se definice obsahu tříd, vlastnosti jedné třídy nemá mít žádná jiná. Požadavek na nepřekrývající se rozsah spočívá v tom, že prvek, zařazený do jedné třídy, už nezařazujeme do žádné jiné – požadujeme vylučující se extenze tříd.

Překrývající se členění umožňuje, aby jeden prvek byl zařazen do více než jedné třídy. Platí zde pravidlo ‚a zároveň‘ (označované symbolem AND). Jinými slovy – jedna instance může být členem více tříd.

### **Stejná úroveň granularity**

V rámci tohoto pravidla je požadována homogenita rozsahu (extenze) tříd. Podle míry podobnosti vlastností či velikosti částí, nejčastěji však podle počtu částí, na něž je členěn celek, se hovoří o klasifikacích homogenních a heterogenních. Zatímco homogenní klasifikace je tvořena na každé úrovni a v každé větvi stejným počtem prvků, heterogenní klasifikace má

rozdílný počet prvků. Z hlediska potenciálních změn v budoucnu je homogenní klasifikace lépe předvídatelná a plánovatelná, heterogenní klasifikace ale zpravidla lépe odpovídá skutečné povaze analyzované reality.

### **Stejná úroveň abstrakce**

Podmínkou dodržení tohoto pravidla je homogenita obsahu (intenze) tříd na jedné úrovni hierarchie i v rámci různých úrovní členění. Obdobou tohoto pravidla je požadavek minimální syntaktické vzdálenosti mezi pojmy na stejné úrovni hierarchie, který stanoví, aby každý z těchto prvků byl definován pomocí stejných primitivů (konstrukčních prvků) a vzorů.

Již bylo uvedeno, že v ontologickém inženýrství se pravidla správné analýzy uplatňují především v procesu návrhu ontologií. Následující principy doplňují k těmto obecným pravidlům ještě charakteristiky umožňující efektivní používání a správu ontologie v praxi.

### **Interoperabilita**

Tento požadavek vychází z všeobecně přijímaného faktu, že nelze vytvořit univerzální ontologii, použitelnou pro všechny potenciální oblasti a případy užití, a že je tedy i do budoucna třeba počítat s koexistencí různých systémů. Interoperabilita je požadována jak na úrovni formátů (struktur, syntaxe), tak na úrovni sémantiky, a to v rámci různých oborů i v rámci různých prostředí. Požadavek je v souladu s kritériem minimální závislosti na znakovém systému, stanoveným Thomasem Gruberem. Komplikací na úrovni konceptualizace představuje, že vyjádření sémantiky založené na přirozeném jazyce je vždy ovlivněno specifiky daného jazykového, kulturního či geopolitického prostředí, v jehož kontextu se používá, a nelze u něj očekávat takovou míru sémantické unifikace jako kupříkladu u vysoce formalizovaných logických výroků.

### **Sémantická (vyjadřovací) síla**

Termín sémantická síla označuje vlastnosti znakové soustavy umožňující vyjádřit přesně a v úplnosti obsah entity. Zpravidla je doprovázena nástroji umožňujícími zachytit co nejširší a nejrozmanitější spektrum vztahů mezi jednotkami obsahu. Lze konstatovat, že čím vyšší sémantická síla, tím větší použitelnost daného systému lze očekávat.

### **Jednoduchá struktura**

Ontologie s jednoduchou strukturou představuje přínos jak pro své tvůrce, tak pro uživatele. Uživatelům umožní snadné používání a správcům snadnou správu a údržbu. Systém s jednoduchou strukturou tedy vykazuje vyšší míru opětovné použitelnosti. Tento požadavek je však v rozporu s předchozím, stejně legitimním požadavkem, aby ontologie svým slovníkem, strukturou i gramatikou co nejvíce modelovala pojmy a v konečném důsledku i jimi odráženou (složitou a proměnlivou) realitu. Platí tedy inverzní poměr použitelnost/znovupoužitelnost.<sup>180</sup>

### **Flexibilita**

Požadavek flexibility je diktován dynamikou univerza poznání, na jehož vývoj by měla být ontologie schopna reagovat. V systémech organizace znalostí je tato vlastnost někdy označována jako ‚pohostinnost‘ (angl. *hospitality*). Anthony C. Foskett ještě doplňuje rozdíl mezi tzv. otevřenými a uzavřenými systémy organizace znalostí spočívající v tom, zda umožňují provádět změny (například vkládání nových pojmů či změny jejich významu) jejich uživateli. Pokud je to

---

<sup>180</sup> Poznámka: Problém sladění požadavků na jednoduchost a na sémantickou sílu již byl diskutován v závěrech předchozí kapitoly (viz kapitola 1.6.1).



možné, označuje takový systém za otevřený a ten systém, jehož změny a aktualizace jsou vyhrazeny pouze jeho tvůrcům, označuje za uzavřený.<sup>181</sup>

Opět se jedná o ambivalentní požadavek – schopnost schématu pružně reagovat na změny znamená zpravidla, že se i sám systém změní, což ohrožuje konzistenci (např. stejný obsah je v různých časových etapách indexován různými metadaty).

### **Metodika *OntoClean***<sup>182</sup>

První návrh metodiky *OntoClean* byl publikován v roce 2000. Autoři Nicola Guarino a Christopher Welty konstatují, že metodika je založena na obecných ontologických pojmech převzatých z filozofie. Připomínají, že v praxi modelování světa pro účely návrhu informačního systému lze vidět paralely s filozofií, jež se pokouší popsat svět formalizovaným a logickým způsobem. Lze konstatovat, že principy, na nichž je metodika založena, jsou v souladu s výše uvedenými obecnými pravidly formálně správné analýzy, odvozenými z logiky.

Metodika je zaměřena na testování ontologie – kontroluje správnost a konzistenci navržené ontologie testováním vlastností jejích tříd. Vlastnosti jsou v rámci této metodiky chápány formalizovaně jako unární predikáty predikátové logiky prvního řádu. Autoři tedy pracují s vlastností jako s významem (intenzí) výroku o třídě typu ‚být čtenářem‘, ‚mít krevní skupinu‘. Podstata metodiky spočívá v určení čtyř metavlastností pro vlastnosti každé třídy: esencialita, rigidita, identita, jednota. Následně se tyto metavlastnosti aplikují na pravidla generické hierarchie. Přínosem metodiky je především to, že umožňuje na dostatečně formalizované úrovni zdůvodnit správnost zvolených řešení v konstrukci ontologie.

Esencialita (angl. *essence*) označuje takovou vlastnost, kterou entita musí mít. Přesněji řečeno, esenciální vlastnost musí mít každá instance třídy, pro niž je vlastnost definována (jež je jejím definičním oborem). Tato vlastnost je tedy v souladu s výše uvedenou logickou podmínkou nutnosti.

Rigidita je speciální typ esenciality. Vlastnost je rigidní, pokud je esenciální ve všech svých případech. Jinými slovy, tuto vlastnost musí mít každá instance po celou dobu své existence. Rigidita koresponduje s aristotelským pojmem podstatné vlastnosti (viz kapitola 1.5.4). Všechny třídy, jež mají rigidní vlastnosti, tvoří tzv. páteřní taxonomii ontologie.

Identita slouží k rozlišení instancí téže třídy. Každá entita v ontologii by měla mít určité kritérium identity a rigidní vlastnosti, které toto kritérium identity popisují. Tato vlastnost má stejnou funkci, jakou plní v prostředí relačních databází tzv. primární klíč.

Jednota určuje, v jakém vztahu jsou části určité entity k celku – slouží k rozlišení části instance od zbytku světa pomocí sjednocující relace. Třída má tuto vlastnost, pokud existují partitivní vztahy pouze mezi jejími instancemi a žádná instance není částí externího celku. Tato vlastnost je v souladu s doporučením, aby partitivní vztahy byly definovány jako monohierarchické (viz např. doporučení v normě ISO 25964-1, uvedené v kapitole 1.5.4).

Při uplatnění metodiky se postupuje následovně: V prvním kroku se každé vlastnosti přidělí metavlastnosti. Druhý krok se zaměřuje na ty vlastnosti, jež byly označeny jako rigidní, a na

---

<sup>181</sup> FOSKETT, Anthony Charles. *The subject approach to information*. 5th ed. London: Library Association Publishing, 1996, s. 111. ISBN 978-1-85604-048-8.

<sup>182</sup> GUARINO, Nicola, WELTY, Christopher A. An overview of *OntoClean*. In: STAAB, Steffen, STUDER, Rudi, ed. *Handbook on ontologies*. 2nd ed. Dordrecht: Springer, 2009, s. 201-220. [https://doi.org/10.1007/978-3-540-92673-3\\_9](https://doi.org/10.1007/978-3-540-92673-3_9). ISBN 978-3-540-70999-2 (Print). ISBN 978-3-540-92673-3 (Online).

tvorbu páteřní taxonomie. Ve třetím kroku se na páteřní taxonomii uplatní principy stanovené pro metavlastnosti: kontrola identity a kontrola jednoty. Čtvrtý krok posuzuje nerigidní vlastnosti. V pátém kroku se do upravené páteřní taxonomie doplní další třídy a vztahy.

### 2.1.2 Metodika UPON

Metodika UPON (*Unified Process for Ontology*) byla vyvinutá v rámci projektů *Interop NoE* (2003–2007) a *Athena IP*, jež byly součástí 6. rámcového programu EU.<sup>183</sup> Autory jsou Antonio De Nicola a Michele Missikoff z *Istituto di Analisi dei Sistemi ed Informatica, Consiglio Nazionale delle Ricerche* v Římě a Roberto Navigli (*Dipartimento di Informatica, Università di Roma „LaSapienza“*). Z hlediska využití softwaru je tato metodika převážně intelektuální, z hlediska postupu řešení využívá postup *middle-out*, životní cyklus je iterativní. Z hlediska účelu ji lze považovat za plánovací i konstrukční, co do formátu jde o technický standard a zdrojem je softwarové inženýrství. Je určena pro návrh rozsáhlých i menších doménových ontologií, určených pro oblast přesně vymezenou konkrétními případy užití. Při návrhu ontologie se počítá s dvěma klíčovými typy odborníků – znalostními (ontologickými) inženýry (*knowledge engineer – KE*) a doménovými specialisty (*domain expert – DE*), jimž jsou v procesu vývoje určeny aktivity, za něž zodpovídají. V každé fázi se propojuje postup *top-down* uplatněný při modelování ontologie s postupem *bottom-up* při tvorbě slovníku.

Metodika UPON nabízí podrobně definovaný postup pro celý proces tvorby ontologie a lze ji proto označit za ‚těžkou‘. V roce 2016 publikovali autoři odlehčenou verzi metodiky, určenou k agilnímu návrhu a rychlému prototypování.<sup>184</sup> Metodika UPON Lite nabízí postup tvorby doménové ontologie, použitelný pro doménové specialisty a koncové uživatele ontologie. Na rozdíl od metodiky UPON už neobsahuje aktivity, k nimž by bylo zapotřebí znalostního (ontologického) inženýra.

Zdrojem pro metodiku UPON je strukturovaná metodika *Unified Process (UP)*<sup>185</sup>, která je považována za standard ve vývoji rozsáhlých softwarových aplikací. Charakteristickými rysy metodiky UP jsou: orientace na uživatele (vývoj systému řízený případy užití), iterativní a inkrementální vývoj, prototypování, opětovné použití existujících návrhů, vzorů i komponent. Nabízí relativně komplikovaný životní cyklus, který ovšem právě díky své komplexnosti lépe odpovídá reálnému průběhu softwarových projektů, než je tomu například u jednoduchého vodopádového postupu. Strukturu softwarového projektu podle metodiky UP tvoří cykly (angl. *cycles*), členěné na čtyři fáze (angl. *phases*). Fáze se nazývají zahájení (angl. *inception*), rozpracování (angl. *elaboration*), konstrukce (angl. *construction*) a zavedení (angl. *transition*). Každá fáze se vnitřně člení na stejných pět pracovních postupů (angl. *workflow*): požadavky (angl. *requirements*), analýza (angl. *analysis*), návrh (angl. *design*), implementace (angl. *implementation*), testování (angl. *testing*). Jak cykly, tak jednotlivé fáze mohou být prováděny

---

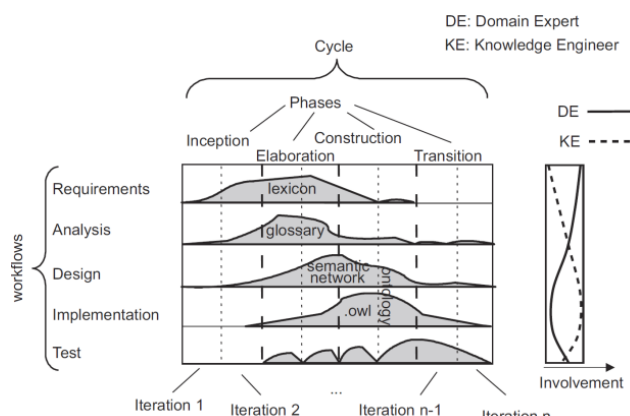
<sup>183</sup> DE NICOLA, Antonio, MISSIKOFF, Michele, NAVIGLI, Roberto. A proposal for a Unified Process for Ontology Building: UPON. In: Kim Viborg Andersen, John Debenham, Roland Wagner, ed. *Database and Expert Systems Applications: 16th International conference, DEXA 2005, Copenhagen, Denmark, August 22-26, 2005: Proceedings*. Berlin: Springer, 2005, s. 655-664. Lecture Notes in Computer Science, no. 3588. doi:10.1007/11546924\_64. ISSN 0302-9743.

<sup>184</sup> DE NICOLA, Antonio, MISSIKOFF, Michele. A lightweight methodology for rapid ontology engineering. In: *Communications of the ACM*. March 2016, 59(3), 79-86. doi:10.1145/2818359. ISSN 0001-0782 (Print). ISSN 1557-7317 (Online).

<sup>185</sup> JACOBSON, Ivar, BOOCH, Grady, RUMBAUGH, James. *The Unified Software Development Process*. Reading: Addison-Wesley Professional, 1999. 512 s. ISBN 978-0201571691.

opakovaně (iterativně). Iterace na úrovni celého cyklu vede k vývoji nové verze.

Metodika *UPON* přebírá beze změny strukturu procesu *UP* včetně názvů fází a pracovních postupů, které však naplňuje konkrétním obsahem přizpůsobeným pro účely návrhu ontologie (viz obrázek 6). Iterativní charakter metodiky se projevuje v postupném vytváření a doplňování výstupů jednotlivých pracovních postupů. Výstupem postupu Požadavky je slovník (angl. *lexicon*) ve formě seznamu termínů. Výstupem analýzy je výkladový slovník (angl. *glossary*), tj. termíny doplněné definicemi. Produktem návrhu je sémantická síť (angl. *semantic network*), v níž jsou do slovníku doplněny vztahy zahrnutých pojmů. Výsledkem implementace je doménová ontologie, jejíž testování je realizováno v následujícím pracovním postupu.



Obr. 6 Obecný rámec metodiky UPON<sup>186</sup>

Rámcový přehled a charakteristika pracovních postupů :

#### Požadavky

Cílem je specifikace sémantických potřeb uživatelů a uživatelského pohledu na znalosti, jež mají být zachyceny v ontologii. Vstupem jsou rozhovory s doménovými experty a dokumenty týkající se aplikační oblasti ontologie. Výstupem jsou kompetenční otázky, případy užití a aplikační slovník. Zodpovědnost za tento pracovní postup mají doménoví specialisté.

#### Analýza

Cílem je detailní rozpracování a strukturování požadavků. Vstupem je aplikační slovník a existující externí zdroje (dokumenty, standardy, ontologie). Výstupem jsou diagramy tříd, diagramy aktivit a referenční výkladový slovník. Zodpovědnost za tento pracovní postup mají jak doménoví specialisté, tak znalostní inženýři.

#### Návrh

Cílem je vypracování ontologické struktury pro položky referenčního výkladového slovníku. Vstupem je referenční výkladový slovník a diagramy tříd a aktivit. Výstupem je sémantická síť a její převod do formátu ontologie. Za tento pracovní postup odpovídají znalostní inženýři.

#### Implementace

Cílem je zakódování ontologie do rigorózního formálního jazyka. Výstupem je ontologie v jazyce *OWL*. Proces implementace zahrnuje výběr formálního jazyka a formalizaci ontologie. Je v zodpovědnosti znalostních inženýrů.

#### Testování

Cílem je kontrola splnění požadavků na ontologii a posouzení sémantické a pragmatické kvality (užitečnosti pro uživatele) ontologie. Za tento pracovní postup zodpovídají společně doménoví a znalostní inženýři.

<sup>186</sup> Obrázek převzat z: DE NICOLA, Antonio, MISSIKOFF, Michele, NAVIGLI, Roberto. A software engineering approach to ontology building. In: *Information systems*. April 2009, **34**(2), s. 260. doi:10.1016/j.is.2008.07.002. ISSN 0306-4379.

Detailní přehled činností v rámci pracovních postupů je shrnut v následující tabulce.

Pracovní postup	Aktivita	Činnosti v rámci aktivity
Požadavky	Určení zájmové domény a rozsahu ontologie	Vymezení té části reality, na niž se ontologie zaměří. Pokud je doména rozsáhlá, určí se subdomény. Nástrojem jsou pojmy a ontologické závazky.
	Definování účelu nebo motivačního scénáře s uživateli a jejich záměry	Vymezení uživatelů ontologie, jejich typologie a úloh, jež bude ontologie pro uživatele řešit.
	Vypracování kresleného scénáře	Načrtnutí pořadí aktivit, jež se odehrávají v rámci konkrétního scénáře.
	Vytvoření aplikačního slovníku	Shromáždění terminologie od doménových expertů a z dokumentů používaných v aplikační doméně. Možnost využití kreslených scénářů a nástrojů pro extrakci znalostí z textových dokumentů.
	Identifikace kompetenčních otázek	Formulace otázek na pojmové úrovni, jež má být ontologie schopna zodpovídat. Hlavními typy jsou otázky směřující k vyhledávání (informačních) zdrojů a otázky, jejichž cílem je dosažení sémantické dohody o významu zdrojů.
	Modelování případů užití	Určení funkcionalit, jež má ontologie poskytovat, aby umožnila zodpovězení kompetenčních otázek. Autoři doporučují použít diagramy případů užití v jazyce UML.
Analýza	Získání doménových zdrojů a tvorba doménového slovníku	Doménový slovník tvoří terminologie používaná v zájmové doméně, převážně získaná analýzou existujících externích dokumentových zdrojů, jako jsou zprávy, manuály, standardy, slovníky, tezaury a dostupné ontologie. Tvorba doménového i aplikačního slovníku může být podpořena automatizovanými nástroji pro <i>text mining</i> .
	Vytvoření referenčního slovníku	Sloučení vybraných termínů z aplikačního a doménového slovníku.
	Modelování aplikačních scénářů	Doplnění zpracovaných diagramů případů užití o diagramy aktivit a diagramy tříd UML.
	Vytvoření referenčního výkladového slovníku	Zpracování neformálních definic v přirozeném jazyce pro termíny zahrnuté v referenčním slovníku.
Návrh	Modelování pojmů	Pro kategorizaci pojmů autoři nabízejí vlastní metodiku OPAL <sup>187</sup> . V průběhu konceptualizace se pojmy rozčlení do kategorií (aktéři, objekty, procesy, zprávy a atributy).
	Modelování hierarchických a doménově specifických vztahů	Při definování hierarchických vztahů je doporučen postup <i>middle-out</i> od nejméně frekventovaných pojmů k obecným a k specifickým. Základní generická taxonomie se doplní o další typy vztahů – partitivní a asociativní. Výsledná sémantická síť je graficky znázorněna diagramem tříd UML.
Implementace	Výběr formálního jazyka	Zvažuje se vyjadřovací síla, výpočetní složitost usuzovací metody a akceptování jazyka v rámci dané komunity. Autoři doporučují jazyk OWL.
	Formalizace ontologie	Do formálního jazyka se převede množina definovaných pojmů, vztahy mezi pojmy a množina sémantických axiomů.
Testování	Kontrola konzistence	Využití automatických kontrol provedených specializovaným softwarem (angl. <i>reasoner</i> ), kontrolujícím správnost definovaných logických axiomů a pravidel (např. disjunktnost hierarchie).
	Verifikace pokrytí	Testuje se, zda je možné sémanticky anotovat zdroj spadající do aplikační domény s použitím pojmů obsažených v ontologii.
	Odpovědi na kompetenční otázky	Testuje se, zda je možné kompetenční otázky, zformulované v rámci pracovního postupu Požadavky, zodpovědět prostřednictvím obsahu ontologie.

Tab. 5 Činnosti v rámci pracovních postupů metodiky UPON

<sup>187</sup> D'ANTONIO, Fulvio, MISSIKOFF, Michele, TAGLINO, Francesco. Formalizing the OPAL eBusiness ontology design patterns with OWL. In: *Enterprise interoperability II: new challenges and approaches*. Proceedings of the Third International conference on interoperability for enterprise applications and software, I-ESA 2007. London: Springer, 2007, s. 345-356. doi:10.1007/978-1-84628-858-6\_38. ISBN 978-1-84628-858-6.

### 2.1.3 Návrhové vzory

Koncept návrhových vzorů pochází původně z oblasti architektury, kam ho uvedl Christopher Alexander (1936).<sup>188</sup> Svě pojetí architektury založil Alexander na přesvědčení, že svět se skládá z opakujících se instancí různých vzorů. Vzor je obecně chápán jako významná a opakující se konstanta mezi pozorovanými daty, objekty či procesy, jež je buď vytvořena uměle nebo se vyskytuje přirozeně. Je založen na induktivním principu zobecněné empirie. Při jeho tvorbě či rozpoznání se tedy uplatňuje princip *bottom-up*, a vzor se následně aplikuje směrem *top-down* na rozpoznáný analogický problém. Koncept vzorů se postupně rozšířil i do dalších disciplín. Kupříkladu matematika uplatňuje vysoce abstraktní a formalizované pojetí vzorů v teorii fraktálů nebo v aplikačních oblastech *data miningu* či zpracování přirozeného jazyka.

Ontologické inženýrství převzalo a aplikovalo již zavedený přístup k návrhovým vzorům z oblasti softwarového inženýrství. V analýze a návrhu softwarových systémů a aplikací představují návrhové vzory (angl. *design patterns*) zobecněná řešení často se vyskytujících a typických problémů. Pro účely opětovného použití jsou formalizovány, zdokumentovány a publikovány.<sup>189</sup> Často se chápou jako šablony, do nichž se v konkrétní implementaci vkládají data typická pro daný případ. Kupříkladu v kapitole 1.1 zmíněná specifikace platformy *MDA* ukládá v průběhu tvorby modelu nezávislého na platformě (*PIM*) provést výběr vzorů, které se pak automaticky transformují podle zvolené platformy do modelu *PSM* (model specifický pro určitou platformu).

Krzysztof Janowicz se spoluautory v úvodu k monografii *Ontologické inženýrství s ontologickými návrhovými vzory* konstatují, že v současné době stoupá zájem o tvorbu ontologických návrhových vzorů i o jejich opětovné používání. Zdůvodňují to aktuálním stavem vývoje sémantického webu. Trend propojených dat zpřístupnil na webu enormní množství dat, jež ovšem postrádají sémantiku. Doménová heterogenita těchto dat prakticky znemožňuje řešit jejich sémantiku pomocí některé ze základních ontologií a řešení je spatřováno v sémantické interoperabilitě doménových ontologií. A právě sémantická interoperabilita je podle autorů silnou stránkou návrhových vzorů.<sup>190</sup> Autoři jsou toho názoru, že většina ontologií „sdílí relativně malou množinu společných problémů týkajících se jejich modelování a publikování, jež lze řešit společnou strategií, využitím příkladů dobré praxe nebo kombinací existujících stovebních bloků“.

Popularitu a míru rozšíření ontologických návrhových vzorů dokládají i již zpracované typologie. Ve výše uvedené monografii jsou návrhové vzory členěny do dvou skupin. První skupinu tvoří kritické přehledy alternativních způsobů modelování (alternativní metody), do druhé skupiny spadají opětovně použitelné komponenty. Významným zdrojem návrhových vzorů pro ontologické inženýrství je portál *Ontology Design Patterns*<sup>191</sup>, provozovaný *Společností pro návrh ontologií a vzory (ODPA)*. Návrhové vzory uložené na tomto portále jsou rozčleněny do šesti základních kategorií, v některých případech ještě členěných na podkategorie: návrhové vzory

---

<sup>188</sup> ALEXANDER, Christopher, ISHIKAWA, Sara, SILVERSTEIN, Murray. *A pattern language: towns, buildings, construction*. New York: Oxford University Press, 1977. xliv, 1171 s. ISBN 0-19-501919-9.

<sup>189</sup> GAMMA, Erich, HELM, Richard, JOHNSON, Ralph, VLISSIDES, John. *Návrh programů pomocí vzorů: Stavební kameny objektově orientovaných programů*. Praha: Grada, 2003. 386 s. ISBN 80-247-0302-5.

<sup>190</sup> HITZLER, Pascal et al., ed. *Ontology engineering with ontology design patterns: foundations and applications*. Berlin: AKA, [2016], 2016, s. xii-xiii. Studies on the semantic web, 025. ISBN 978-3-89838-715-6.

<sup>191</sup> *Ontology design patterns.org (ODP)* [online]. Association for Ontology Design & Patterns, 2009- [cit. 2018-01-30]. Dostupné z: <http://ontologydesignpatterns.org/>.

strukturální, korespondenční (pro reengineering a slučování ontologií), obsahové, vzory pro automatické usuzování a odvozování, prezentační a lexikálně-syntaktické.

## 2.2 Doménová analýza

*Nejplodnějším horizontem informační vědy je studium znalostních domén jako myšlenkových nebo diskurzivních komunit, jež jsou součástí společenské dělby práce.*<sup>192</sup>  
Birger Hjørland, 2017

Ontologické inženýrství a organizace znalostí se sice uvažují jako samostatné oblasti, už na minimální úrovni zobecnění je však patrné, že mají významný průnik cílů, předmětu zájmu/zkoumání a samozřejmě i metod. Právě doménovou analýzu považují obě disciplíny za významnou součást svého metodologického aparátu.

Termín doména pochází z francouzštiny a původně byl spojen s fyzickou (zpravidla geografickou) oblastí. Označuje kupříkladu rozsah panství, ve virtuálním světě Internetu se v současné době používá pro označení rozsahu množiny internetových adres. V přeneseném slova smyslu se používá i pro označení oblastí vymezených intenzionálně, tj. vnitřními vlastnostmi, například pro oblasti činnosti nebo zkoumání.

V podobném smyslu jako ‚doména‘ se používá termín ‚univerzum diskurzu‘, zmíněný již v kapitole 1.2. Aplikací tohoto obecného pojmu jsou potom nejrůznější doménová univerza a komunity, například bibliografické univerzum, univerzum poznání, informační univerzum, diskurzivní komunita apod. V logice je s obdobným cílem používán koncept možných světů, zavedený již G. W. Leibnizem. Spočívá v začlenění uvažované části reality do konkrétního prostoru a času, v jehož rámci je analyzovaný výrok pravdivý.

Metoda doménové analýzy používá termín doména ve významu obor působnosti, odborná oblast nebo diskurzivní komunita. Za diskurzivní komunitu se obvykle považuje skupina lidí, spjatých společnými zájmy a s nimi spojenou komunikací. Specifickými diskurzivními komunitami v oblasti vědy jsou tzv. neviditelné univerzity<sup>193</sup>, v současné etapě síťové komunikace jsou často diskutovanými komunitami sociální sítě. I když se doménová analýza neztříká kvantitativních metod (viz např. níže uvedená Hjørlandova hlediska 4 a 5), je považována za metodu kvalitativního výzkumu.

Výsledkem doménové analýzy je model domény (doménový model). Tvorba takového modelu je kupříkladu zahrnuta mezi pravidla tvorby aplikačních profilů *Dublin Core*.<sup>194</sup> Za specifické doménové modely je možné považovat i doménové ontologie (viz kapitola 1.4). Je vhodné upozornit na to, že pojem doména ve smyslu, v němž je používán v metodě doménové analýzy, vyjadřuje nikoli oblast nebo předmět zájmu samotný, ale konstruovaný artefakt. Hanne Albrechtsenová zdůrazňuje, že domény „*nejsou nějaká území tam venku, čekající, až budou*

---

<sup>192</sup> HJØRLAND, Birger. Domain analysis [online]. Version 1.3. 2017-07-09 [cit. 2018-01-30]. In: *ISKO Encyclopedia of knowledge organization*. ISKO, 2017. Dostupné z: [http://www.isko.org/cyclo/domain\\_analysis](http://www.isko.org/cyclo/domain_analysis).

<sup>193</sup> JONÁK, Zdeněk. Neviditelná univerzita. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2018-01-30]. Dostupné z: [http://aleph.nkp.cz/F/?func=direct&doc\\_number=000003515&local\\_base=KTD](http://aleph.nkp.cz/F/?func=direct&doc_number=000003515&local_base=KTD).

<sup>194</sup> COYLE, Karen, BAKER, Thomas. *Guidelines for Dublin Core Application Profiles* [online]. Dublin (Ohio): DCMI, 2009-05-18 [cit. 2018-01-30]. Dostupné z: <http://dublincore.org/documents/profile-guidelines/>.

popsána a analyzována několika málo zasvěcenci. V podstatě je lze všechny vytvořit“<sup>195</sup>. To, že domény neexistují apriori, samozřejmě neznamená, že je nelze vymezit. Jako každý artefakt se i doména vyznačuje určitou stabilitou a je tedy použitelná k analýze a výzkumu.<sup>196</sup>

Uplatnění metody doménové analýzy není vázáno na konkrétní vědní disciplínu. Zcela v duchu svých principů je použitelná v různých doménách, na jejichž specifické rysy se adaptuje a nabývá tak v různých oblastech použití různých podob. Aplikačními oblastmi doménové analýzy relevantními pro ontologické inženýrství jsou především softwarové inženýrství a informační věda a její hraniční obory. Birger Hjørland a Hanne Albrechtsenová vyjmenovávají tyto významné oblasti aplikace doménové analýzy v hraničních oborech informační vědy: filozofie a teorie vědy, pedagogický výzkum, kognitivní psychologie (způsob, jak člověk přemýšlí, je ovlivněn tím, o čem přemýšlí), biologie (behaviorální ekologie hledání zdrojů), psycholingvistika a sociolingvistika (směry, které jsou v opozici vůči abstraktní textové analýze a formálnímu strukturalismu).<sup>197</sup>

### **Doménová analýza v softwarovém inženýrství**

Doménová analýza v softwarovém inženýrství se začala diskutovat v průběhu 80. let 20. století zejména v souvislosti s výzkumy metod a technik opětovné použitelnosti softwarových návrhů a komponent. Jako první zřejmě použil a definoval termín James M. Neighbors ve své disertaci věnované metodě návrhu opětovně použitelných softwarových komponent. Podle Neighborse spočívá doménová analýza v identifikaci objektů a operací v třídě podobných systémů v konkrétní problémové doméně. Tím doménovou analýzu přibližuje metodě objektově orientované analýzy a návrhu. Její specifika oproti systémové analýze vidí v tom, že se „*netýká konkrétních akcí v konkrétním systému. Zajímá se místo toho o to, jaké akce a objekty se vyskytují ve všech systémech v aplikační oblasti (problémové doméně)*“.<sup>198</sup>

K propracování metody doménové analýzy významně přispěl Ruben Prieto-Díaz, který se rovněž orientuje na opakovatelnou použitelnost softwaru. Termín doménová analýza použil pro označení způsobu, jímž dospěl k návrhu fasetové klasifikace softwaru. Z pohledu informační vědy je zajímavé, že inspirací a teoretickou základnou se mu stala teorie fasetové analýzy S. R. Ranganathana (viz kapitola 2.3.1). Prieto-Díaz definuje doménovou analýzu jako „*proces identifikace, záznamu a strukturování informací používaných při vývoji softwarových systémů s cílem učinit je opětovně použitelnými*“<sup>199</sup>. Chápe doménovou analýzu především jako proces kumulace zkušeností a poznatků, jež se po dosažení určitého množství syntetizují a abstrahují,

---

<sup>195</sup> ALBRECHTSEN, Hanne. This is not domain analysis. In: *Knowledge organization*. 2015, **42**(8), s. 561. ISSN 0943-7444.

<sup>196</sup> Poznámka: Dovedeno do důsledků je tedy sama doména modelem a doménový model je ,modelem modelu' neboli metamodelem.

<sup>197</sup> HJØRLAND, Birger, ALBRECHTSEN, Hanne. Toward a new horizon in information science: domain analysis. In: *Journal of the American Society for Information Science*. July 1995, **46**(6), 404-409. doi: 10.1002/(SICI)1097-4571(199507)46:6<400::AID-ASIS2>3.0.CO;2-Y. ISSN 0002-8231 (Print). ISSN 1097-4571 (Online).

<sup>198</sup> NEIGHBORS, James M. *Software construction using components* [online]. Ph.D. Dissertation, Technical Report UCI-ICS-TR-160. Irvine: Department of Information and Computer Science, University of California, Irvine, ©1980, 1981, 1996 [cit. 2018-01-30], s. 6. Dostupné z: <http://www.bayfronttechnologies.com/l02draco.htm#diss80>.

<sup>199</sup> PRIETO-DÍAZ, Rubén. Domain analysis: an introduction. In: *ACM SIGSOFT Software engineering notes (SEN)*. April 1990, **15**(2), s. 47. doi:10.1145/382296.382703. ISSN 0163-5948.

aby mohly být opětovně použitelné. Konstatuje, že doménová analýza může a má používat osvědčené nástroje jiných disciplín, například techniky získávání a reprezentace znalostí používané v umělé inteligenci a vyspělé techniky a ověřené metody systémové analýzy nebo technologie informačního managementu. Kromě toho Prieto-Díaz nabízí i vlastní procesní model doménové analýzy, založený na metodice tvorby fasetových klasifikací v oblasti knihovnictví. Jako zdroje doménových znalostí jmenuje technickou literaturu, existující implementace, uživatelské průzkumy a doporučení expertů. Pro vytvořené doménové modely doporučuje formu taxonomií, standardů, funkčních modelů a doménových jazyků. Zajímavým námětem je názor, že nelze čekat na přirozený vývoj domény, ale je nutné stimulovat proces jejího zrání systematickými aktivitami.

### **Doménová analýza v ontologickém inženýrství**

Rovněž ontologické inženýrství často používá doménovou analýzu, již většina metodik doporučuje provést v úvodní fázi návrhu ontologie. Kupříkladu metodika Natalye Noyové a Debory McGuinnessové zařazuje doménovou analýzu do prvního kroku, jenž zahrnuje určení domény a rozsahu ontologie. Autorky zdůrazňují pragmatický aspekt doménové analýzy, jejíž výsledek ovlivňuje především stanovený účel a uživatelé ontologie. „*Neexistuje žádný správný způsob modelování domény – vždy existují životaschopné alternativy. Nejlepší řešení téměř vždy závisí na aplikaci, kterou máte na mysli, a na rozšířeních, která očekáváte.*“<sup>200</sup>

### **Doménová analýza v informační vědě**

V novodobém oborovém knihovnictví byly aktivity doménové analýzy vždy implicitně obsaženy jako jeho zpravidla nedokumentované *know-how*. Joseph Tennis připomíná, že ranou variantou byl už princip *literary warrant*<sup>201</sup> (blíže o tomto principu v kapitole 4.2). V odborných a vědeckých knihovnách tradičně pracují na úsecích zpracování fondů a služeb oboroví knihovníci / doménoví specialisté. ‚Založení‘ doménové analýzy pro informační vědu v polovině 90. let 20. století je tedy spíše snahou o explicitní vyjádření a zformulování jejích principů. Obousměrnou linii vzájemné inspirace informační vědy a softwarového inženýrství při rozvoji doménové analýzy zosobňuje Hanne Albrechtsenová, vyučující na *Královské škole knihovnictví a informační vědy (RSLIS)* Univerzity v Kodani a v letech 1996–1998 předsedkyně *Mezinárodní společnosti pro organizaci znalostí ISKO*, která pracovala v letech 1987–1991 na projektech EU zaměřených na znovupoužitelnost softwaru (*ESPRIT – European strategic program on research in information technology*). V rámci projektu *Practitioner* se podílela na tvorbě tezauru pro obor softwarového inženýrství, při jehož návrhu byla použita metoda doménové analýzy. V roce 1995 spolu s Birgerem Hjørlandem publikovala článek *K novému horizontu informační vědy: doménová analýza*<sup>202</sup>, který je považován za počáteční podnět k explicitnímu vymezení doménové analýzy jako klíčové metody informační vědy.

---

<sup>200</sup> NOY, Natalya F., McGUINNESS, Deborah L. *Ontology development 101: a guide to creating your first ontology* [online]. March 2001 [cit. 2018-01-30]. 25 s. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880. Dostupné z: [http://protege.stanford.edu/publications/ontology\\_development/ontology101.html](http://protege.stanford.edu/publications/ontology_development/ontology101.html).

<sup>201</sup> TENNIS, Joseph T. What does a domain analysis look like in form, function, and genre? In: *Brazilian journal of information science*. Jan./Jun. 2012, 6(1), s. 3. ISSN 1981-1640.

<sup>202</sup> HJØRLAND, Birger, ALBRECHTSEN, Hanne. Toward a new horizon in information science: domain analysis. In: *Journal of the American Society for Information Science*. July 1995, 46(6), 400-425. doi: 10.1002/(SICI)1097-4571(199507)46:6<400::AID-ASI2>3.0.CO;2-Y. ISSN 0002-8231 (Print). ISSN 1097-4571 (Online).



V rámci informační vědy se doménová analýza obvykle chápe jako informační analýza domény, resp. zkoumání aspektů komunikace informací a znalostí v rámci domény. David Bawden a Lyn Robinsonová přímo používají termín informační doména, který definují jako „...soubor informačních systémů, zdrojů, služeb a procesů spojených se skupinou uživatelů se společnými zájmy a společným hlediskem, sdílející společnou terminologii.“<sup>203</sup>

Komplexní pohled doménové analýzy v informační vědě má souvislost s informační ekologií, jejímž zájmem je vztah člověka a informačního prostředí.<sup>204</sup> Zatímco cílem doménové analýzy je poznání informačního prostředí, cílem informační ekologie je nastavit interakci člověka s informačním prostředím (informační zdroje, technologie, aktéři – tvůrci, uživatelé, poskytovatelé).

Podle Birgera Hjørlanda je doménová analýza „teorie a přístup ke knihovní a informační vědě a organizaci znalostí... Doménová analýza se zabývá systémy a procesy organizace znalostí z kombinované sociologické a epistemologické perspektivy a zdůrazňuje význam oborových znalostí.“<sup>205</sup> Svým důrazem na sociální aspekty se doménová analýza liší od technicky (information retrieval, textová analýza) nebo behavioristicky zaměřených metod informační vědy (např. zkoumání informačního chování).

Hjørland definuje tři dimenze domény: ontologickou (tj. předmět zájmu v doméně), epistemologickou (tj. znalosti v doméně) a sociologickou (tj. lidé a skupiny v doméně). Joseph Tennis v roce 2003 přidal k Hjørlandovým dimenzím další dvě osy: oblasti modulace, umožňující pojmenování a extenzionální vymezení domény a stupně specializace, zaměřené na intenzionální definování.<sup>206</sup> V článku z roku 2012 Tennis doplňuje, že považuje za dva základní typy doménové analýzy analýzu popisnou a instrumentální. Popisná doménová analýza například ve formě bibliometrických map je určena pro základní výzkum. Instrumentální doménová analýza slouží k návrhu a vývoji systémů organizace znalostí.<sup>207</sup>

V roce 2002 odvodil Birger Hjørland z výše uvedených tří dimenzí (ontologie, epistemologie, sociologie) jedenáct způsobů a hledisek doménové analýzy pro informační vědu, jež zahrnují celý životní cyklus informace a znalosti. Lze tedy odvodit, že Hjørland chápe doménovou analýzu jako specifický typ vícekriteriální analýzy. Svůj návrh hledisek zformuloval následovně:

1. Průvodce zdroji a předmětové brány
2. Speciální klasifikace a tezaury
3. Indexace a vyhledávání informací v speciálních oborech
4. Empirické výzkumy uživatelů
5. Bibliometrické studie

---

<sup>203</sup> BAWDEN, David, ROBINSON, Lyn. *Úvod do informační vědy*. 1. vyd. Doubravník: Flow, 2017, s. 133. ISBN 978-80-88123-10-1.

<sup>204</sup> Viz např. STEINEROVÁ, Jela. Informačná ekológia – využívanie informácií srdcom. In: *ITlib: informačné technológie a knižnice* [online]. 2009, 9(2) [cit. 2018-01-30]. ISSN 1335-793X. Dostupné z: [http://itlib.cvtisr.sk/archiv/2009/2/informacna-ekologia-vyuzivanie-informacii-srdcom.html?page\\_id=1108](http://itlib.cvtisr.sk/archiv/2009/2/informacna-ekologia-vyuzivanie-informacii-srdcom.html?page_id=1108).

<sup>205</sup> HJØRLAND, Birger. Domain analysis [online]. Version 1.3. 2017-07-09 [cit. 2018-01-30]. In: *ISKO Encyclopedia of knowledge organization*. ISKO, 2017. Dostupné z: [http://www.isko.org/cyclo/domain\\_analysis](http://www.isko.org/cyclo/domain_analysis).

<sup>206</sup> TENNIS, Joseph T. Two axes of domains for domain analysis. In: *Knowledge organization*. 2003, 30(3-4), 191-195. ISSN 0943-7444.

<sup>207</sup> TENNIS, Joseph T. What does a domain analysis look like in form, function, and genre? In: *Brazilian journal of information science*. Jan./Jun. 2012, 6(1), 5-6. ISSN 1981-1640.

6. Historické studie
7. Výzkum dokumentů a žánrů
8. Epistemologické a kritické studie
9. Terminologické studie, analýzy speciálního jazyka a diskurzu
10. Struktury a instituce ve vědecké a profesní komunikaci.
11. Odborné poznání a reprezentace znalostí v počítačové vědě a v umělé inteligenci.

Tato hlediska Hjørland doporučil navzájem kombinovat. Současně je označil za specifické odborné kompetence informačního vědce a profesionála, jimiž se odlišuje od oborového specialisty, například od výzkumného pracovníka v daném oboru. Nejvýznamnějším hlediskem doménové analýzy podle Hjørlanda je epistemologický výzkum (tj. výzkum znalostí v doméně – ve výše uvedeném seznamu je zastoupen v bodě 8). Hjørland je přesvědčen, že toto hledisko musí být zastoupeno v každé doménové analýze, protože bez něj by se všechna ostatní hlediska stala povrchními.<sup>208</sup>

Již bylo uvedeno, že způsob uplatnění metody doménové analýzy v různých disciplínách se často liší. I v rámci informační vědy lze zaznamenat různě stanovené cíle analýzy, jež ovlivňují výběr technik a následně i podobu doménového modelu. Jako příklady mohou posloužit studie, jež zpracovali Anders Ørom, Lyn Robinsonová a Jiří Stodola.

Birger Hjørland<sup>209</sup> označuje za modelový příklad doménové analýzy studii organizace znalostí v oboru dějin umění, již zpracoval Anders Ørom.<sup>210</sup> Analýza je zaměřená na možnosti organizace znalostí v rámci této domény. Konstatuje existenci několika ‚tradičních‘ diskurzivních paradigmat (kulturní dějiny, ikonografie, stylové paradigma, materialismus) a změnu diskurzu v doméně v 70. letech 20. století (‚nové‘ paradigma). Posuzuje, zda významné systémy organizace znalostí (*Knihovnicko-bibliografické třídění, Deweyho desetinné třídění, Třídění Kongresové knihovny, Mezinárodní desetinné třídění*) jsou schopny reprezentovat dokumenty vytvořené jak v rámci ‚nového‘, tak ‚tradičního‘ diskurzu.

Lyn Robinsonová si ve své studii vytkla za cíl vytvořit pojmový model informační vědy.<sup>211</sup> Tradiční metodu výzkumu procesů komunikačního řetězce zkombinovala s nástroji doménové analýzy. Výsledkem je vysoce abstraktní model předmětu zkoumání informační vědy, tvořený třemi vzájemně kombinovatelnými úrovněmi: 1. komponenty komunikačního řetězce (tj. procesy tvorby, šíření, organizace, indexace, uložení a užití), 2. hlediska doménové analýzy (tj. výše uvedených 11 hledisek Birgera Hjørlanda) a 3. kontext (rozsah – jedinec, skupina, společnost a média, tj. typy informačních zdrojů). Ve smyslu tohoto modelu je pak informační věda definována jako „vědní obor, jehož předmětem zájmu jsou humánní zaznamenané informace se zaměřením na komponenty informačního řetězce, zkoumané pomocí doménové analýzy“.<sup>212</sup> Autorka předpokládá, že model bude možné rozšířit do formy fasetového klasifikačního

<sup>208</sup> HJØRLAND, Birger. Domain analysis in information science: eleven approaches – traditional as well as innovative. In: *Journal of documentation*. 2002, **58**(4), 422-462. doi:10.1108/00220410210431136. ISSN 0022-0418.

<sup>209</sup> HJØRLAND, Birger. Domain analysis [online]. Version 1.3. 2017-07-09 [cit. 2018-01-30]. In: *Encyclopedia of knowledge organization*. ISKO, 2017. Dostupné z: [http://www.isko.org/cyclo/domain\\_analysis](http://www.isko.org/cyclo/domain_analysis).

<sup>210</sup> ØROM, Anders. Knowledge organization in the domain of art studies – history, transition and conceptual changes. In: *Knowledge organization*. 2003, **30**(3/4), 128-143. ISSN 0943-7444.

<sup>211</sup> ROBINSON, Lyn. Information science: communication chain and domain analysis. In: *Journal of documentation*. 2009, **65**(4), 578-591. doi:10.1108/00220410910970267. ISSN 0022-0418.

<sup>212</sup> Tamtéž, s. 587.

schématu oblasti informační vědy.

V úvodu kapitoly bylo konstatováno, že téma doménové analýzy zatím není diskutováno v české odborné literatuře. Ojedinelou výjimkou je článek Jiřího Stodoly, zaměřený na dvě paradigmatu informační vědy, jež nazývá kognitivní a sociálně-doménové. Autor se snaží o integrovaný přístup – dle jeho názoru jsou tato paradigmatu „vzájemně slučitelná, když se z nich odstraní jejich extrémní pozice (*empirismus a racionalismus z kognitivismu, kolektivismus ze sociálně-doménového přístupu*) a ponechají se jejich přednosti (*individualismus u kognitivního a důraz na komunikaci u sociálně-doménového paradigmatu*).“<sup>213</sup> Možnosti integrace obou paradigmatu demonstruje Stodola na vlastním návrhu doménového modelu informační vědy, vycházejícím z modelu FRBR.

### **Doménová analýza v organizaci znalostí**

O tom, že je doménová analýza považována za významnou metodu organizace znalostí, svědčí to, že jejím otázkám byla věnovaná dvě zvláštní čísla časopisu *Knowledge organization* (2003/3-4, 2015/8). Důkazem zralosti metody je i vydaná příručka (učebnice) doménové analýzy<sup>214</sup>. Vspělost metody a její stěžejní význam pro informační vědu a organizaci znalostí potvrdila rovněž bibliometrická studie z roku 2015,<sup>215</sup> v níž Richard Smiraglia analyzoval 100 doménově analytických studií 80 autorů z 23 zemí, publikovaných v období 1995–2014. Smiraglia doménovou analýzou rozumí sadu technik pro identifikaci specifické poznatkové základny a její význam pro organizaci znalostí zdůvodňuje následovně:

*„Doménová analýza je jádrem organizace znalostí, neboť bez ní bychom neměli žádný ontologický materiál pro tvorbu našich systémů organizace znalostí (SOZ). Coby v podstatě empirická taktika zahrnující analýzu specifického prostředí pro zjišťování jeho jazyka, kultury a činností tak doménová analýza využívá všech epistemických přístupů ve své interpretaci a analýze doménových ontologií.“*<sup>216</sup>

Klíčovým procesem organizace znalostí je bezesporu obsahová analýza a indexace (viz kapitola 4.4). Metoda doménové analýzy použitá k analýze obsahu zdůrazňuje jeho sociální a kulturní kontext. Jens-Erik Mai to shrnuje slovy: „*Ústřední myšlenkou doménové analýzy je, že organizace a reprezentace informací má začínat analýzou kontextu, diskurzu a aktivit.*“<sup>217</sup>

---

<sup>213</sup> STODOLA, Jiří. Individua, sociální sítě a poznání: možnosti a limity kognitivního a sociálně-doménového paradigmatu v informační vědě a realistický model domény. In: *ProInflow: časopis pro informační vědy* [online]. 2012, 4(1) [cit. 2018-01-30], s. 56. ISSN 1804-2406. Dostupné z: <http://www.phil.muni.cz/journals/index.php/proinflow/article/view/808>.

<sup>214</sup> SMIRAGLIA, Richard P. *Domain analysis for knowledge organization: tools for ontology extraction*. 1st ed. Waltham: Elsevier Chandos Pub, 2015. 105 s. ISBN 978-0-08-100150-9.

<sup>215</sup> SMIRAGLIA, Richard P. Domain analysis of domain analysis for knowledge organization: observation on an emergent methodological cluster. In: *Knowledge organization*. 2015, 42(8), 602-611. ISSN 0943-7444.

<sup>216</sup> SMIRAGLIA, Richard P. Epistemology of domain analysis. In: SMIRAGLIA, Richard P., LEE, Hur-Li, ed. *Cultural frames of knowledge*. Würzburg: Ergon-Verlag, 2012, s. 111. ISBN 978-3-89913-918-1.

<sup>217</sup> MAI, Jens-Erik. Analysis in indexing: document and domain centered approaches. In: *Information processing and management*. May 2005, 41(3), s. 605. ISSN 0306-4573.

## 2.3 Fasetová analýza

*S enumerativními systémy se skutečně snadno pracuje, a pro statické univerzum znalostí jsou nejlepší volbou. Bohužel žádné univerzum znalostí, které by bylo statické, neexistuje.*<sup>218</sup>  
Mohinder Partap Satija a Daniel Martínez-Ávila, 2015

Metoda fasetové analýzy je příkladem metody ontologického inženýrství, jejíž zdroje lze spatřovat v knihovní a informační vědě. Podle Vandy Broughtonové mají informatické ontologie „konceptuálně velmi blízko k fasetové struktuře a ... fasetová analýza má hodně co nabídnout jako metoda pro tvorbu ontologií.“<sup>219</sup> Obdobně A. Neelameghan v komentáři k postulátům a normativním principům S. R. Ranganathana konstatuje podobnost databázového datového modelu a fasetové struktury s tím, že „je pozoruhodné, že existuje paralela ve vývoji modelů klasifikačních a indexačních jazyků (například hierarchických, fasetových, relačních, volně fasetových) s vývojem datových modelů (hierarchických, síťových a relačních)“.<sup>220</sup> Společné rysy spatřuje i v postupu návrhu<sup>221</sup>: jak návrh klasifikačních schémat, tak návrh databázových systémů lze členit do tří úrovní, jež Ranganathan pojmenoval jako úroveň idejí, úroveň slov a úroveň notace. Neelameghan uzavírá, že fasetová teorie má hodně společného s principy systémové analýzy a syntézy, na nichž je založen návrh informačních systémů a softwarových aplikací. Lze doplnit, že pojem faseta vykazuje příbuznost s pojmy ‚multisystém‘ a ‚multistruktura‘, jež vystihují situaci, kdy se na témže objektu identifikuje současně více systémů.<sup>222</sup> Tím souvisí rovněž s konceptem multiplikativních ontologií (viz kapitola 1.5.4), které na rozdíl od ontologií redukcionistických připouštějí v jednom časoprostoru existenci více entit. Podnětem k teoretickému rozvoji fasetové analýzy se stala nová fáze vývoje vědy ve 20. století, během níž byla dosavadní specializace vystřídaná inter- a transdisciplinaritou. Komplexnost zkoumaných problémů, při jejichž analýze už nepostačuje jedno hledisko speciální disciplíny, vyvolala nutnost vícekritériálního přístupu doprovázeného důrazem na zachycení vzájemných vztahů.

Klíčovým pojmem fasetové analýzy je faseta. Tento termín pochází z latinského *facies*, tj. tvář. V přeneseném slova smyslu jej používají klenotníci pro označení jedné z vybroušených plošek drahokamu. O další přenos významu slova faseta do oblasti práce s informacemi a znalostmi se zasloužil Shiyali Ramamrita Ranganathan, který jej použil pro označení jednoho z více možných aspektů zkoumaného obsahu dokumentu. Takovým aspektem může být kupříkladu téma, místo či doba děje nebo jazyková forma. V současných textech je termín faseta používán poměrně volně. Někdy označuje samotnou kategorii (často jednu ze základních kategorií v nějaké doméně, například ‚téma‘, ‚místo‘ – viz přehled základních kategorií v kapitole 1.5.3), někdy je používán pro označení jednoho z možných kritérií členění (například ‚podle tématu‘, ‚podle místa‘). Kromě

---

<sup>218</sup> SATIJA, Mohinder Partap, MARTÍNEZ-ÁVILA, Daniel. Features, functions and components of a library classification system in the LIS tradition for the e-environment. In: *Journal of information science theory and practice*. 2015, 3(4), 68. doi:<http://dx.doi.org/10.1633/JISTaP.2015.3.4.5>. ISSN 2287-9099 (Print). ISSN 2287-4577 (Online).

<sup>219</sup> BROUGHTON, Vanda. The need for a faceted classification as the basis of all methods of information retrieval. In: *Aslib proceedings*. 2006, 58(1/2), s. 66. doi:10.1108/00012530610648671. ISSN 0001-253X.

<sup>220</sup> NEELAMEGHAN, A., compiled. *S. R. Ranganathan's postulates and normative principles: applications in specialized databases design, indexing and retrieval*. Bangalore: Sarada Ranganathan Endowment for library science, 1997, s. 163.

<sup>221</sup> Tamtéž, s. 176-177.

<sup>222</sup> HABR, Jaroslav, VEPŘEK, Jaromír. *Systémová analýza a syntéza: zdokonalování a projektování systémů*. 2. přeprac. vyd. Praha: SNTL, 1986, s. 32-33.

systémů s předem stanovenými fasetami se lze setkat i s tzv. volně fasetovou klasifikací (angl. *freely faceted classification*) s ad hoc definovanými fasetami přizpůsobenými konkrétnímu organizovanému materiálu.<sup>223</sup>

Fasetová analýza vychází ze základního způsobu lidského myšlení, kterým je konceptualizace. Podle Vandy Broughtonové je založena na základních kategoriích, má tedy přímou souvislost s filozofickou problematikou řešící tento problém (viz kapitola 1.5.3). Jejím specifikem je, že je založena na jednom konkrétním typu ‚umělé‘ konceptualizace, jímž je kategorizace na základě společných vlastností, tj. uplatňuje objektivní, systémový přístup. Fasetová analýza tudíž vykazuje obdobné rysy jako systémová analýza. V obou případech se klade důraz na rozpoznání struktury, funkce a především vzájemných vztahů komponent. Fasetová analýza umožňuje provést vědecky správnou analýzu s jednotným kritériem členění a vzájemně se vylučujícími prvky a zároveň efektivní syntézu, tj. vícedimenzionální reprezentaci složených témat a komplexních objektů.

Již bylo uvedeno, že fasety se uplatňují i v návrhu informačních systémů a v softwarovém inženýrství. Ukazuje se, že vícerozměrný analyticko-syntetický fasetový přístup je v souladu s tím, jak se postupuje při analýze a návrhu složitých datových objektů za použití objektově orientovaného přístupu. Na rozdíl od většiny oblastí použití, kde je za fasetu považována kategorie, která obsahuje další kategorie (pojmy), se v softwarovém inženýrství obvykle uplatňuje jemnější granularita faset. Fasety reprezentují jednotlivé atributy entit/tříd. Kupříkladu fasetami třídy ‚Kniha‘ tak mohou být atributy ‚ISBN‘, ‚přírůstkové číslo‘ a ‚název‘, naplněné konkrétními hodnotami instancí třídy ‚Kniha‘.

Kromě analýzy a návrhu se fasety uplatňují i při prezentaci informací v informačních systémech. Fasetová konstrukce uživatelského rozhraní softwarových aplikací umožňuje uživatelům přístup ke zdrojům a službám založený na prohlížení (fasetové listování nabídkou je typické např. pro e-shopy) a zvyšuje efektivnost navigace. Díky omezení prohledávaného souboru lze obvykle umožnit i listování ve výsledcích vyhledávání.

Princip faset patří rovněž ke stěžejním zásadám v informační architektuře, což je obor specializovaný na organizaci webového obsahu. Fasetové uspořádání se stává de facto normou v organizaci obsahu webových sídel a v katalozích elektronických obchodů.

Dalším příkladem aktuálního využití fasetového principu členění podle více hledisek jsou metody vícekritériálního rozhodování a hodnocení variant, doplněné formalizovaným matematickým aparátem, jež jsou uplatňované zejména v operačním výzkumu.

### **2.3.1 Fasetová analýza v informační vědě a v organizaci znalostí**

‚Předfasetová‘ vícekritériální analýza v organizaci znalostí používala jednu linii členění a různá kritéria umísťovala na různé hierarchické úrovně. Takový přístup kodifikuje například norma ISO 7154 pro pořádání bibliografických informací.<sup>224</sup> Pro kritérium členění norma používá označení ‚pořádací kritérium‘ (angl. *filing criterion*), označení ‚pořádací úroveň‘ (angl. *filing level*) se používá pro určení relativního postavení tzv. pořádací jednotky (angl. *filing unit*) v pořádací

---

<sup>223</sup> GNOLI, Claudio, HONG, Mei. Freely faceted classification for Web-based information retrieval. In: *New review of hypermedia and multimedia*. June 2006, **12**(1), 63-81. doi:10.1080/13614560600758944. ISSN 1361-4568 (Print), 1740-7842 (Online).

<sup>224</sup> ČSN ISO 7154 (01 0141). *Dokumentace – Zásady bibliografického pořádání*. Praha: Český normalizační institut, 2003. 15 s.

hierarchii. Norma stanoví pět hierarchických úrovní pořádaní bibliografických údajů, přičemž každá následující úroveň se uplatní tehdy, pokud se na předchozí úrovni vyskytnou stejné údaje (např. v případě stejného jména autora se záznamy dále mohou řadit podle názvu, v případě stejného názvu podle roku vydání atd.).

Za průkopníka fasetové analýzy v informační vědě je považován Shiyali Ramamrita Ranganathan. Je autorem *Dvojtečkového třídění*, univerzálního klasifikačního schématu založeného na fasetovém principu. Termínem faseta označuje Ranganathan množinu všech pojmů, jež se nacházejí v konkrétní kategorii klasifikačního schématu. Souběžně s prací na vývoji *Dvojtečkového třídění* zformuloval rozsáhlou a zevrubnou teorii klasifikace. Kompletní popis Ranganathanovy klasifikační teorie ve znění publikovaném v třetím vydání díla *Prolegomena to library classification* (Úvod do knihovnické klasifikace)<sup>225</sup>, tvoří jedenáct normativních principů, čtyřicet pět kánonů, třináct postulátů, dvacet jeden princip a devět prostředků (nástrojů). Principy, jež Ranganathan formuloval, členil do několika úrovní obecnosti, jež označil čísly. Na nejobecnější úrovni 1 jsou základní zákony procesu myšlení, úroveň 2 tvoří podstatné zákony knihovní vědy. Úroveň 3 a úroveň 4 se vztahují k návrhu klasifikačního schématu: úroveň 3 zahrnuje kányony klasifikace a úroveň 4 principy nápomocného pořadí v rámci pole. Úroveň 5 tvoří postuláty fasetové analýzy a principy fasetové sekvence, používané v průběhu klasifikační činnosti.<sup>226</sup> Podrobná, přesná a jednoznačná znění principů obsahové a fasetové analýzy formuloval Ranganathan s úmyslem, aby tyto principy byly použitelné i při strojovém zpracování.

Fasetový přístup k organizaci znalostí se mezi odborníky diskutoval v průběhu celého 20. století. Nejvýznamnějšími představiteli oboru, kteří se zasloužili o rozvoj teoretických základů fasetové analýzy, byli kromě Shiyali R. Ranganathana Henry E. Bliss a členové britské *Skupiny pro výzkum klasifikace (CRG)*<sup>227</sup>. V praxi se ale fasetový princip jevil dlouho jako implementačně problematický – pro lineární fyzické uspořádání např. knih na regále a pro vizualizaci takto koncipovaného systému organizace znalostí byl těžko použitelný. V současném virtuálním síťovém prostředí webu jsou však vytvořeny technologické podmínky pro využití jeho předností.

Využití faset v organizaci znalostí je skutečně všestranné. Fasetová analýza se používá při návrhu systémů organizace znalostí, zejména klasifikačních schémat, tezaurů a ontologií. Další oblastí použití je indexace dokumentů, kdy fasety umožňují uplatnit vícehlediskovou obsahovou analýzu. V této souvislosti se lze setkat jak s fasetami, jež umožňují zařadit indexovaný dokument pouze do jedné kategorie (např. žánr), tak s vícehodnotovými fasetami (např. téma, uživatelské určení). Stejně jako v softwarových aplikacích a v informačních systémech se fasety rovněž používají při vyhledávání a prohlížení, jakož i při prezentaci organizovaných zdrojů. Fasetová prezentace výsledků je typická pro současné online katalogy knihoven, jež kombinují vyhledávání a prohlížení: výsledky strukturovaného i plnotextového vyhledávání se v

---

<sup>225</sup> RANGANATHAN, Shiyali Ramamrita. *Prolegomena to library classification*. Assisted by M. A. GOPINATH. 3rd ed. London: Asia Publishing House, 1967. 640 s. Ranganathan series in library science, 20. ISBN 978-0-210-22731-2.

<sup>226</sup> Tamtéž, s. 113-114.

<sup>227</sup> Classification Research Group. The need for a faceted classification as the basis of all methods of information retrieval. In: CHAN, Lois Mai, RICHMOND, Phyllis A., SVENONIUS, Elaine, ed. *Theory of subject analysis: a sourcebook*. Littleton: Libraries Unlimited, 1985, s. 157-167. ISBN 0-87287-489-3. – Původně publikováno jako: UNESCO (IAC Doc. Ter. PAS) document 320/5515, 26 May 1955.

knihovních systémech typu *discovery* prezentují fasetovým rozhraním.

Efektivnost fasetové analýzy se projevuje v úsporném vyjádření složených pojmů a témat. Není nutné předem vyjmenovávat všechny varianty (např. dřevěné nebytové budovy, dřevěné bytové budovy, betonové bytové budovy ad.), ale vyjádří se podle potřeby kombinací obsahu jednotlivých faset (například ‚materiál‘, ‚účel‘, ‚objekt‘). Tato analyticko-syntetická metoda má společné rysy s některými sémantickými a lingvistickými metodami, zejména s teorií pojmové analýzy, sémantického rozkladu a sémantických primitivů (viz např. koncept elementárních sémantických jednotek Anny Wierzbické<sup>228</sup>). Ve všech případech je cílem rozložit složitý obsah na jednoduché, už dále nedělitelné jednotky významu. Vyjádření složeného tématu se pak děje jejich kombinací.

### 2.3.2 Metodika fasetové analýzy Williama Dentona a Louise Spiteriové

Ranganathanovy kánony, postuláty a principy mají bezesporu vysokou teoretickou úroveň, jejich komplikovanost ale brání bezprostřední aplikaci v praxi. Je v nich patrný vliv autorova kulturního prostředí, především východní filozofie a náboženství. Obtíže při studiu Ranganathanových textů navíc činí ne vždy jasná a konzistentní zdůvodnění principů, složitost formulací a originalita terminologie, která se ovšem právě kvůli tomu může stát překážkou odborné komunikace. Proto Louise F. Spiteriová v roce 1998 publikovala zjednodušenou metodiku fasetové analýzy určenou původně pro účely výuky studentů knihovní a informační vědy na *Dalhousieově univerzitě* v Halifaxu.<sup>229</sup> Metodika je založena na kompilaci vybraných kánonů, postulátů a principů S. R. Ranganathana a principů fasetové analýzy zpracovaných *Skupinou pro výzkum klasifikace (CRG)*. V roce 2003 dopracoval William Denton tuto metodiku do podoby vhodné pro tvorbu fasetové klasifikace určené pro publikování webového obsahu.<sup>230</sup>

Východiskem jsou čtyři fáze tvorby fasetové klasifikace, jež v roce 1960 definoval Brian Campbell Vickery<sup>231</sup>:

1. Roztřídění termínů v dané oblasti do homogenních, navzájem se vylučujících faset, z nichž každá je odvozena z nadřazeného univerza podle jednoho kritéria členění;
2. Určení pořadí, v němž budou fasety použity ke konstrukci složených hesel;
3. Vytvoření takové notace pro dané schéma, jež dovolí plně flexibilní kombinaci termínů a jež rozmístí předměty do preferovaného pořadí (angl. *filling order*);
4. Použití fasetového schématu takovým způsobem, že umožní jak specifické odkazy, tak požadovanou úroveň obecného průzkumu.

---

<sup>228</sup> WIERZBICKA, Anna. *Sémantika: elementární a univerzální sémantické jednotky*. Praha: Karolinum, 2014. 549 s. ISBN 978-80-246-2289-7.

<sup>229</sup> SPITERI, Louise F. A simplified model for facet analysis: Ranganathan 101. In: *Canadian journal of information and library science*. April-July 1998, 23(1-2), 1-30. ISSN 1195-096X (Print). ISSN 1920-7239 (Online).

<sup>230</sup> DENTON, William. *How to make faceted classification and put it on the web* [online]. [Personal web site]: November 2003 [cit. 2018-01-30]. Dostupné z: <http://www.miskatonic.org/library/facet-web-howto.html>.

<sup>231</sup> VICKERY, B. C. *Faceted classification: a guide to construction and use of special schemes*. London: Aslib, 1960, s. 12-13. Cit. dle Williama Dentona (viz předchozí citace).

V metodice navržené Williamem Dentonem jsou Vickeryho fáze doplněny o další tři fáze. Ústřední fáze – vytvoření a uspořádání faset, jsou založeny na zjednodušeném modelu fasetové analýzy, který vypracovala Louise Spiteriová. Ve vodopádovém životním cyklu se uplatňují četné prvky iterace, v jednotlivých fázích se střídavě užívají postupy *top-down* a *bottom-up*. Klíčové fáze a činnosti, jež probíhají v jejich rámci, jsou znázorněny v tabulce 6.

Fáze	Činnosti v rámci fáze
Shromáždění entit	Shromáždění reprezentativního vzorku entit pro danou doménu; u malých kolekcí je doporučeno zahrnout entity z celé domény.
Popis entit	Vytvoření popisu jednotlivých entit. Sémantický rozklad popisů – rozdělení vět na části, přeuspořádání slov, rozpoznání základních pojmů v jednotlivých částech vět a jejich izolování.
Vytvoření faset	Prozkoumání výsledných termínů a nalezení obecných kategorií na nejvyšší úrovni, vyskytujících se ve všech entitách. Omezení nalezených kategorií na množinu vzájemně se vylučujících a ve svém celku vyčerpávajících faset, použitelných pro všechny termíny vytvořené v předchozí fázi. V rámci výběru faset se uplatňují následující principy: <ul style="list-style-type: none"> <li>▪ diferenciace – použité kritérium členění má vést k jasnému rozlišení jednotlivých částí;</li> <li>▪ relevance – zvolené fasety mají odpovídat účelu, předmětu a rozsahu klasifikace;</li> <li>▪ průkaznost a permanence – fasety mají reprezentovat trvalé a měřitelné charakteristiky klasifikovaných objektů;</li> <li>▪ homogenita a disjunktnost – každá faseta smí reprezentovat pouze jednu charakteristiku členění;</li> <li>▪ základní kategorie mají být přizpůsobeny specifickým vlastnostem a kontextu klasifikovaných objektů.</li> </ul>
Uspořádání faset	Naplnění faset termíny vytvořenými ve fázi Popis entit, tak aby bylo možné zatřídit všechny shromážděné entity. Uspořádání termínů uvnitř fasety podle zvoleného principu (abecední, hierarchický, chronologický ad.). Volba preferovaných termínů, tvorba řízeného slovníku.
Určení pořadí faset	Rozhodnutí o standardním (defaultním) uvádění pořadí faset (např. na domovské stránce).
Klasifikace	Zatřídění celého obsahu domény. Analýza všech entit za použití faset a použití správných termínů pro popis každé entity.
Revize, testování, správa	Při jakýchkoli problémech zjištěných v etapě klasifikace následuje iterace – návrat k předchozí fázi, jež je relevantní k vyřešení problému. Následuje uživatelské testování a pravidelná správa klasifikace (změna terminologie, nové entity, změny v doméně).

Tab. 6 Činnosti v rámci klíčových fází metodiky Williama Dentona

Dentonova metodika se stala de facto standardem tvorby jednoduchých fasetových klasifikací v prostředí webu a dočkala se i aktualizace a přepracování do podoby ontologického návrhového vzoru.<sup>232</sup>

<sup>232</sup> RODRIGUEZ-CASTRO, Bene, GLASER, Hugh, CARR, Leslie. How to reuse faceted classification and put it on the semantic web. In: Peter F. PATEL-SCHNEIDER et al., ed. *The Semantic Web – ISWC 2010: 9th International Semantic Web Conference, ISWC 2010, Shanghai, China, November 7-10, 2010: Revised Selected Papers, Part I*. Berlin: Springer, © 2010, s. 663-678. Lecture Notes in Computer Science, no. 6496, ISSN 0302-9743. doi:10.1007/978-3-642-17746-0\_42. ISBN 978-3-642-17745-3 (Print). ISBN 978-3-642-17746-0 (Online).



### 2.3.3 Metoda 5W1H

Šestice kategorií označovaná jako 5W1H je příkladem ‚zlidovělé‘ fasetové kategorizace vycházející z aristotelského kategoriálního systému.<sup>233</sup> Zkratka je vytvořena z prvních písmen anglických tázacích zájmen *who* – kdo, *what* – co, *where* – kde, *when* – kdy, *why* – proč, *how* – jak. Stejně jako u dalších lidových výtvorů nelze jednoznačně určit autora a dobu vzniku. Metodu 5W1H lze považovat za ‚naivní‘ podobu fasetové analýzy. Její víceméně intuitivní přístup k celostnímu zkoumání však právě díky své názornosti a snadné aplikovatelnosti našel uplatnění v celé řadě odvětví.

Metoda je populární zejména v prostředí žurnalistiky a sociální komunikace, ale i v dalších oblastech společenského výzkumu. Příkladem z oblasti žurnalistiky je doporučené pevné pořadí ‚kdo – co – kde – kdy – proč – jak‘, stanovící optimální strukturu novinové zprávy. Metodu použili například Harold Dwight Laswell při definování principů politiky („*Who gets what, when, how*“ – Kdo získá co, kdy a jak) a sociální komunikace („*Who says what to whom in what channel with what effect*“ – Kdo říká co komu jak s jakým výsledkem) nebo Rick Szostak<sup>234</sup> při klasifikaci vědeckých teorií a metod. Obdobně Ladislav Tondl charakterizuje způsob chápání a interpretace smyslu sdělení v širším kontextu prostřednictvím otázek „kdo, kdy, proč, pro koho“ nebo „k čemu“<sup>235</sup>.

### 2.3.4 Teorie integrativních úrovní

Princip úrovní je známým a v četných variantách se vyskytujícím konceptem jak ve filozofii, tak v přírodních, společenských i technických vědách (viz např. vrstevnatá síťová architektura nebo vrstvy sémantického webu). Je založen na myšlence postupně vzrůstající složitosti. Objekty zájmu jsou řazeny lineárně tak, že objekt na vyšší úrovni zahrnuje všechny vlastnosti objektu na nižší úrovni a přidává ještě své specifické vlastnosti. Obdobně jako systémová analýza nabízí obecný rámec, použitelný napříč vědními disciplínami, a je tudíž vhodnou metodou pro interdisciplinární a transdisciplinární obory, jakož i pro mezioborový a mezikulturní výzkum.

James K. Feibleman ve své eseji z roku 1954 nazvané *Teorie integrativních úrovní*<sup>236</sup> zformuloval 12 zákonů integrativních úrovní:

1. Každá úroveň organizuje úroveň nebo úrovně pod ní a jednu vznikající kvalitu.
2. Složitost úrovní se směrem vzhůru zvyšuje.
3. V každé organizaci závisí vyšší úroveň na nižší úrovni.
4. V každé organizaci je nižší úroveň řízena úrovní vyšší.
5. Pro organizaci na kterékoliv úrovni je její mechanismus na úrovni níže a její účel na výše uvedené úrovni.

---

<sup>233</sup> Poznámka: Na tuto souvislost poukazuje například člen severoanglické *Skupiny pro výzkum klasifikace (Northern Classification Study Group)* R. Moss v článku: MOSS, R. Categories and relations: origins of two classification theories. In: *American documentation*. October 1964, **15**(4), 296-301.

<sup>234</sup> SZOSTAK, Rick. Classifying scholarly theories and methods. In: *Knowledge organization*. 2003, **30**(1), 20-35. ISSN 0943-7444.

<sup>235</sup> TONDL, Ladislav. *Půl století poté: pohledy na problémy sémantiky a sémiotiky v posledních desetiletích*. 1. vyd. Praha: Karolinum, 2006, s. 40. ISBN 80-246-1207-0 (Print). ISBN 978-80-246-2556-0 (pdf).

<sup>236</sup> FEIBLEMAN, James K. Theory of integrative levels. In: CHAN, Lois Mai, RICHMOND, Phyllis A., SVENONIUS, Elaine, ed. *Theory of subject analysis: a sourcebook*. Littleton: Libraries Unlimited, 1985, s. 136-143. ISBN 0-87287-489-3. – Původně publikováno v: *British journal for the philosophy of science*. 1954, 5, 55-66.

6. Porucha zavedená do organizace na jakékoli úrovni se projevuje na všech nižších úrovních, které pokrývá.
7. Čas potřebný pro změnu v organizaci se zkrátí, když postoupíme na vyšší úroveň.
8. Čím vyšší je úroveň, tím menší je počet instancí.
9. Není možné snížit vyšší úroveň na nižší.
10. Organizace na jakékoli úrovni je zkreslením níže uvedené úrovně.
11. Události na kterékoli úrovni ovlivňují organizace na jiných úrovních.
12. Cokoliv je ovlivněno jako organizace, má nějaký vliv jako organizace.

Z formulací zákonů integrativních úrovní je patrný těsný vztah této teorie k procesu organizace. Díky tomu teorie našla odezvu i v oboru organizace znalostí. V 60. letech 20. století se členové britské *Skupiny pro výzkum klasifikace* pokoušeli využít teorii integrativních úrovní k vytvoření univerzálního fasetového systému, který by umožnil propojit dílčí oborové fasetové klasifikace. Z filozofických principů teorie integrativních úrovní vychází fasetová klasifikace *ICC (Information coding classification)*<sup>237</sup> a z ní odvozená klasifikace *KO Literature* Ingetraut Dahlbergové (blíže viz kapitola 4.5.2) a *Integrative levels classification* Claudia Gnoliho<sup>238</sup>.

Zajímavým příkladem propojení metody *5W1H* s teorií integrativních úrovní je doménová ontologie podnikového systému Johna Zachmana, známá pod názvem *Zachmanův rámec (Zachman's framework)*. První návrh byl pod názvem *Architektura informačního systému (Information system architecture – ISA)* publikován v roce 1987<sup>239</sup>, aktuální podoba byla vytvořena v r. 1992 ve spolupráci s Johnem F. Sowou. Řádky dvojrozměrné tabulky představují integrativní úrovně a označují aktéry, kteří se postupně podílejí na tvorbě informačního systému: plánovač, majitel, návrhář, ‚stavitel‘, subkontraktor. Ve sloupcích jsou uvedeny entity *5W1H* charakterizující podnikový systém: ‚co‘ – data, ‚jak‘ – proces, ‚kde‘ – místo, ‚kdo‘ – role, ‚kdy‘ – časování, ‚proč‘ – motivace. Zachmanův rámec je propojen se standardem *OMG Business Motivation Model*.<sup>240</sup> Verze 3.0 z roku 2011 už má doplněn podnázev *Podniková ontologie (The enterprise ontology)*, jenž explicitně vyjadřuje jeho ontologický charakter.

---

<sup>237</sup> DAHLBERG, Ingetraut. The Information coding classification (ICC): a modern, theory-based fully-faceted, universal system of knowledge fields. In: *Axiomathes*. June 2008, **18**(2), 161-176. doi:10.1007/s10516-007-9026-8. ISSN 1122-1151 (Print). ISSN 1572-8390 (Online).

<sup>238</sup> GNOLI, Claudio. Categories and facets in integrative levels. In: *Axiomathes*. June 2008, **18**(2), 177-192. doi:10.1007/s10516-007-9022-z. ISSN 1122-1151.

<sup>239</sup> ZACHMAN, John A. A framework for information systems architecture. In: *IBM systems journal*. 1987, **26**(3), 276-292. doi:10.1147/sj.263.0276. – Další vývoj Zachmanova rámce je zdokumentován na: <https://www.zachman.com/ea-articles-reference/54-the-zachman-framework-evolution> [cit. 2018-01-30].

<sup>240</sup> Object Management Group. *Business motivation model (BMM)* [online]. Version 1.3. Needham (MA, USA): Object Management Group, May 2015 [cit. 2018-01-30]. 104 s. OMG Document Number: formal/2015-05-19. Dostupné z: <http://www.omg.org/spec/BMM/>.

## 2.4 Diskuse

### 2.4.1 Omezení systémového a analytického přístupu

Stejně jako v případě teoretických základů ontologického inženýrství, i v případě jeho metodik lze konstatovat, že obecným rámcem všech metod je systémový přístup a jejich obecnou metodologií je systémová analýza. Jejich přednosti byly konstatovány v kapitole 2.1.1 a potvrzují je četné úspěšné případy aplikace v průběhu analýzy a návrhu informatických ontologií. V této části jsou diskutována omezení systémového a analytického přístupu se zřetelem na jejich aplikaci v klíčovém procesu ontologického inženýrství, jímž je konceptualizace.

Jaroslav Habr a Jaromír Vepřek upozorňují na svého druhu paradox systémového přístupu. Snaha pohlížet na systém jako celek je současně překážkou jeho dekompozice a klasifikace.<sup>241</sup> Z pohledu v opačném směru lze limity metody systémové analýzy vyjádřit známým úslovím, že celek je víc, než součet jeho částí. Když Ludwig Wittgenstein uvažoval o možnostech definování pojmu ‚hra‘, dospěl k závěru, že je to pojem s nejasným ohraničením (mluví o „rozplývavých okrajích“), a klade si otázky:

*„Ale je rozplývavý pojem vůbec **pojmem**? – Je neostrá fotografie vůbec obrazem nějakého člověka? Ba může být neostrý obraz vždycky s výhodou nahrazen obrazem ostrým? Není často právě neostrý obraz tím, co potřebujeme?“<sup>242</sup>*

V jistém smyslu lze zaznamenat až určitou diskreditaci systémového přístupu v ‚tvrdé‘ podobě, v níž byl v polovině 20. století často naivně uplatňován na ‚měkký‘ materiál humánní, společenské komunikace. Je jistě žádoucí uplatnit v ontologickém inženýrství systémový přístup, je však zapotřebí respektovat výrazný lidský aspekt konceptualizace tak, jak to například nabízí metodologie měkkých systémů Petera Checklanda.<sup>243</sup> Požadavky na správnou analýzu, jejíž výsledky by byly počítačově zpracovatelné, vycházejí ze zásad formální logiky a jsou relevantní v rámci klasické, dle George Lakoffa „tradiční“ teorie konceptualizace (viz charakteristika systémového přístupu ke konceptualizaci v kapitole 1.5.3). Jejich dodržování přináší tu výhodu, že systémy konstruované podle těchto pravidel budou zpracovatelné počítačovými programy a do jisté míry jim budou i srozumitelné, tj. umožní ‚inteligentní‘ přístup a využití, například odvozování. Analyzovaná skutečnost se však často nedá začlenit do takového formálního rámce a vynucuje si pragmatická řešení. Četné příklady z různých aplikačních oblastí dokládají, že principy analýzy uvedené v kapitole 2.1.1 představují ideál, v reálném prostředí nedosažitelný. Některé požadavky dokonce vytvářejí navzájem se vylučující kombinace – například požadavek jasného ohraničení a úplnosti je v rozporu s požadavkem na flexibilitu nebo již několikrát zmíněný požadavek na sémantickou sílu koliduje s požadavkem na jednoduchost.

Navíc je potřeba si připomenout, že klasická formální logika, na níž jsou založeny obecně přijímané a aplikované principy analýzy, se zajímá o pravdivost, nikoli o obsah výroků. Její použitelnost pro obsahovou analýzu a konceptualizaci, jež je jádrem tvorby ontologií, je tedy

<sup>241</sup> HABR, Jaroslav, VEPŘEK, Jaromír. *Systémová analýza a syntéza: zdokonalování a projektování systémů*. 2. přeprac. vyd. Praha: SNTL, 1986, s. 15.

<sup>242</sup> WITTGENSTEIN, Ludwig. *Filosofická zkoumání*. Přel. J. Pechar. Praha: Filosofický ústav AV ČR, 1993, s. 48. ISBN 80-7007-040-4.

<sup>243</sup> CHECKLAND, Peter, POULTER, John. *Learning for action: a short definitive account of soft systems methodology and its use for practitioners, teachers and students*. Chichester: John Wiley & Sons, 2006. xxiii, 200 s. ISBN 0-470-02554-9.

problematická. Problémy se projevují zejména v obsahové analýze znalostí vyjádřených v přirozeném jazyce, v němž se vyskytují metafory, umělecké obrazy a další specifické jazykové jevy. Zákony intenzionální logiky, jež se zabývá sémantikou a mohla by tedy nabídnout potřebná řešení, jsou zatím ve fázi formování a diskusí.

K tomuto problému se přidává problematika složených pojmů a s tím spojených analyticko-syntetických operací. Shiyali R. Ranganathan popisuje svou inspiraci k vytvoření analyticko-syntetické fasetové klasifikace v dětské stavebnici *Meccano*, umožňující vytvořit z malého počtu základních stavebních prvků libovolně složité a rozmanité celky. Předmětem fasetové analýzy jsou ale pojmy vyjádřené jazykovými výrazy. Analogie mechanických komponent stavebnice a strukturních prvků jazyka je tudíž sporná. I když lze jazyk dekomponovat na jednotlivé strukturní části (slova, syntaktická pravidla apod.), je zřejmé, že výsledný význam jazykového výrazu konstruovaného z těchto částí ovlivňují četné další faktory. Lze tedy uzavřít, že systémová analýza je pouze jednou z metod konceptualizace, byť metodou vysoce efektivní a postavenou na exaktních základech.

#### 2.4.2 Vztah metodik ontologického inženýrství a metod organizace znalostí

Jedním z výsledků provedeného průzkumu metodik ontologického inženýrství byl poznatek, že nejnáročnější a současně v metodikách nejméně specifikovanou fází tvorby ontologie je konceptualizace. Vzhledem k již několikrát konstatované příbuznosti metod ontologického inženýrství a metod organizace znalostí proto není překvapivé, že někteří autoři nabízejí možnost využití nahromaděných zkušeností ze sféry paměťových institucí, tak jak je zobecnily a teoreticky zpracovaly obory informační věda a organizace znalostí. Příznačné v tomto smyslu je stanovisko Ingertraut Dahlbergové v úvodu ke knize *Organizace znalostí: vývoj, úlohy, využití, budoucnost*. Dahlbergová s trochou nadsázky konstatuje, že „...kolegům informatikům leckdy chybí znalost toho, co již přece během staletí bylo vyvinuto v duchovních vědách na poli klasifikace. Právě zjistili, že jejich data mají také obsah, a rádi používají klasifikační systémy a tezaury, vytvořené informačními vědci, chybí jim ovšem znalost jejich [teoretických] základů.“<sup>244</sup>

Průnik používaných metodik ontologického inženýrství a organizace znalostí se již v plné míře projevuje v metodách doménové analýzy a fasetové analýzy, jež jsou v praxi sdílené oběma disciplínami. V kapitole 2.3 již byly citovány názory Vandy Broughtonové a A. Neelameghana na možnosti použití metody fasetové analýzy v ontologickém inženýrství. Další možností je využití empirických principů organizace znalostí, shromážděných během tvorby knihovních klasifikačních schémat (podrobněji jsou představeny v monografii *Organizace znalostí: klíčová témata*<sup>245</sup>). Tyto principy zformulované významnými představiteli teorie a praxe organizace znalostí se obvykle vyvinuly z příkladů nejlepší praxe, jež se osvědčily v nějakém konkrétním prostředí, často v knihovnách. Jejich výstižnou charakteristiku podává Henry Evelyn Bliss:

„Principy lze definovat jako zobecnění nebo stálé vztahy či shody v rámci akcí, postupů a metod. Dáváme přednost pojmenování principy, protože nezahrnují závislosti nebo nezbytnost zákonů a

---

<sup>244</sup> DAHLBERG, Ingetraut. *Wissensororganisation: Entwicklung, Aufgabe, Anwendung, Zukunft*. Würzburg: Ergon Verlag, 2014, s. 9. Textbooks for knowledge organization, vol. 3. ISSN 0944-8152. ISBN 978-3-95650-065-7.

<sup>245</sup> Empirické principy a pravidla organizace znalostí. In: KUČEROVÁ, Helena. *Organizace znalostí: klíčová témata*. 1. vyd. Praha: Karolinum, 2017, kap. 3.3, s. 150-160. ISBN 978-80-246-3587-3 (brož). ISBN 978-80-246-3597-2 (pdf).

*kánonů; přitom jsou obecnější jak ve formulacích, tak v použitelnosti než pravidla nebo předpisy. Jsou odvozené z příslušných oblastí skutečnosti a zkušeností, spíše než by určovaly skutečnosti nebo jednání; jsou spíše praktické a použitelné než účinné; jsou popisným shrnutím ... spíše než usměrňující směrnici nebo nařizující formulací.*"<sup>246</sup>

Shiyali R. Ranganathan používá v obdobném smyslu termín postulát a doplňuje, že jde o výrok, o němž nelze prohlásit, zda je správný nebo chybný, může být pouze prohlášen za nápomocný nebo za neužitečný.<sup>247</sup> Mnohé principy a pravidla organizace znalostí vykazují díky své induktivní povaze podobnost s návrhovými vzory (viz kapitola 2.1.3), jež se dnes používají v ontologickém inženýrství.

Na možnosti synergie zkušeností a kompetencí knihovníků a informatiků jsme již upozornili v kapitole 1.1. V kapitole 2.2 bylo konstatováno, že organizace znalostí a ontologické inženýrství se sice uvažují jako samostatné oblasti, je však patrné, že mají významný počet společných cílů, shodují se i v předmětech svého zájmu a v principech a metodách jejich zkoumání. Lze předpokládat, že obousměrná integrace metodologií obou disciplín se může stát účinným stimulem jejich dalšího rozvoje.

---

<sup>246</sup> BLISS, Henry Evelyn. *The organization of knowledge in libraries and the subject-approach to books*. New York: H. W. Wilson, 1933, s. 21-22.

<sup>247</sup> RANGANATHAN, Shiyali Ramamrita. Facet analysis: fundamental categories. In: CHAN, Lois Mai, RICHMOND, Phyllis A., SVENONIUS, Elaine, ed. *Theory of subject analysis: a sourcebook*. Littleton: Libraries Unlimited, 1985, s. 88. ISBN 0-87287-489-3. – Původně publikováno v rámci: RANGANATHAN, Shiyali Ramamrita. *The elements of library classification*. 3rd ed. Bombay: Asia Publishing House, 1962. 168 s.

## Kapitola 3 První verze ontologie navržená pro znalostní bázi organizace znalostí

*Tvorba počítačových programů (programování) je pro kognitivní vědy příliš užitečná, než aby byla přenechána pouze umělé inteligenci. Existuje řada nesporných předností, jež programy přinášejí teoretikovi: úžasně koncentrují mysl; transformují mysticismus na zpracování informací tím, že nutí teoretika explicitně vyjádřit intuice a přeložit vágní terminologii do konkrétních návrhů; zajišťují zaručený test konzistence teorie a tím umožní bezpečně zkompletovat složité interaktivní komponenty [...] Pro kognitivního vědce by největší přednost programování neměla spočívat v samotném dokončeném programu či v tom, co program dělá, ale spíše v činnostech v průběhu jeho vývoje.<sup>248</sup>*

Philip N. Johnson-Laird, 1980

Předmětem této kapitoly je shrnutí výsledků výzkumu provedeného v letech 2013–2015 v rámci grantového projektu *Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) DF13P01OVV013 Znalostní báze pro obor organizace informací a znalostí*<sup>249</sup>. Výzkum byl zaměřen na tvorbu funkčního prototypu znalostní báze, což zahrnovalo i tvorbu první verze ontologie organizace znalostí. Výzkumu jsem se účastnila po celou dobu řešení jako členka řešitelského týmu.

Kapitola je členěna na pět částí. V první části jsou specifikovány cíle, harmonogram a řešitelé projektu. Druhá část představuje architekturu znalostní báze. Ve třetí části je popsána první verze ontologie organizace znalostí, tak jak je implementována v prototypu znalostní báze. Čtvrtá část popisuje technologická řešení a procesy implementace znalostní báze. V páté části jsou shrnuty výsledky projektu a jsou uvedeny dílčí závěry, které obsahují kritické zhodnocení první verze ontologie.

### 3.1 Cíle, harmonogram a řešitelé projektu

Podnětem k podání grantové přihlášky projektu znalostní báze organizace znalostí byly výsledky orientačního srovnávacího průzkumu stavu poznatkové základny oboru organizace znalostí v České republice a ve světě. Průzkum byl proveden v letech 2010–2012 a původně byl zamýšlen pro účely konkretizace tématu disertace, tj. ontologie organizace znalostí. Ukázal však, že kromě nekompatibilní odborné terminologie, k jejímuž řešení má přispět ontologie, chybí v tuzemském prostředí aktuální vědecké monografie a učebnice. Bylo proto rozhodnuto rozšířit rozsah výzkumu a kromě ontologie vytvořit i znalostní bázi organizace znalostí, jejíž obsah by bylo možné využít jako podkladový materiál pro publikační a vzdělávací aktivity.

Projekt měl stanoveny šest dílčích cílů:

Dílčí cíl 1 – Shromáždění a systemizace poznatků v oboru organizace znalostí s ohledem na lokální kulturní a jazyková specifika českého prostředí.

Dílčí cíl 2 – Návrh pojmového modelu znalostní báze ve formě ontologie.

---

<sup>248</sup> JOHNSON-LAIRD, Philip N. Mental models in cognitive science. In: *Cognitive science*. 1980, 4(1), 109-110. doi:10.1207/s15516709cog0401\_4. ISSN 0364-0213.

<sup>249</sup> Informace o projektu jsou dostupné v Centrální evidenci projektů (CEP) Informačního systému výzkumu, experimentálního vývoje a inovací z: <https://www.rvvi.cz/cep?s=jednoduche-vyhledavani&ss=detail&n=0&h=DF13P01OVV013> [cit. 2018-01-30].

Dílčí cíl 3 – Naplnění znalostní báze instancemi.

Dílčí cíl 4 – Zpřístupnění znalostní báze online ve formátu propojených otevřených dat.

Dílčí cíl 5 – Posouzení možnosti znovupoužití znalostní báze jako prototypu pro doménové znalostní báze v jiných oborech.

Dílčí cíl 6 – Aktualizace české odborné terminologie.

V návaznosti na určené dílčí cíle byly stanoveny následující předpokládané výsledky projektu:

1. Veřejně online dostupná znalostní báze pro obor organizace znalostí s využitím zejména v těchto oblastech:

- Podkladový materiál pro aktuální vědeckou monografii zachycující komplexně problematiku organizace znalostí.
- Obohacení české odborné terminologie z oboru organizace znalostí.
- Zdroje pro výuku a vzdělávání v oboru informační věda a knihovnictví. Využití znalostní báze pro studijní účely jak v kurzech v rámci vysokoškolských studijních programů, tak pro studium a další vzdělávání informačních profesionálů v České republice.
- Znovupoužití znalostní báze jako prototypu pro doménové znalostní báze v jiných oborech.

2. Průběžné informace o řešení projektu pro odbornou veřejnost. Články publikované v recenzovaných odborných časopisech a vystoupení na konferencích.

3. Vědecká studie shrnující výsledky provedeného výzkumu.

Teoretická východiska řešení projektu se kromě reprezentativní odborné literatury, zjištěné bibliografickým průzkumem, opírala o mezinárodní standardy ISO/TR 9007, ISO 25964, ISO 704, o pojmový model *FRBR* a o standardy sémantického webu (*RDF*, *RDFS*, *OWL*).

Na základě aktualizovaného průzkumu metodik ontologického inženýrství z roku 2011<sup>250</sup> byla pro návrh znalostní báze zvolena strukturovaná metodika *UPON*, jejíž detailní popis je podán v kapitole 2.1.1. Zpětně lze konstatovat, že ve fázi specifikace požadavků a ve fázi analýzy, zejména v průběhu konceptualizace, byla řešiteli implicitně uplatňována metoda doménové analýzy, její postup však nebyl dokumentován. Ve fázi návrhu ontologie byla při definování faset využita metodika Williama Dentona a Louise Spiteriové (viz kapitola 2.3.2) v kombinaci s metodou *5W1H*, charakterizovanou v kapitole 2.3.3. V klíčových etapách řešení projektu (akvizice znalostí, jejich analýza, zpracování a uložení do znalostní báze) byly tyto metodiky specifikovány a byly navrženy a uplatněny vlastní postupy, jež jsou popsány v části 3.4.1 této kapitoly.

Práce na projektu v letech 2013–2015 byly realizovány podle schváleného předběžného rámcového harmonogramu, jehož základem se staly fáze a pracovní postupy vycházející z metodiky *UPON*. Harmonogram je znázorněn níže v tabulce 7. V průběhu jednotlivých let řešení byl v návaznosti na dosavadní postup řešení upřesňován a konkretizován (viz skutečné termíny splnění činností, znázorněné v tabulce 7).

---

<sup>250</sup> KUČEROVÁ, Helena. Metodiky ontologického inženýrství. In: *Ikaros* [online]. 2011, 15(5) [cit. 2018-01-30]. URN-NBN:cz-ik6801. ISSN 1212-5075. Dostupné z: <http://www.ikaros.cz/node/6801>.

Převažující typ výzkumu			2013		2014		2015	
			1 – 6	7 – 12	1 – 6	7 – 12	1 – 12	
			Zahájení	Rozpracování	Konstrukce	Zavedení		
základní (ZV)		rozhovory s experty		x				
			rešerše zdrojů	x	✓			
aplikovaný (AV)	9.1	Specifikace požadavků	kompetenční otázky		x			
			případy užití		✓			
			seznam termínů		x			
			seznam zdrojů pro excerpci	x	✓			
			excerpce zdrojů		x	✓		
			diagram tříd a diagram aktivit	1. verze	x		✓	
	9.2	Analýza	výkladový slovník	1. verze		x	✓	
				2. verze			x	
			sémantická síť	1. verze			x	
				2. verze			✓	
	9.3	Návrh	sémantická síť	1. verze			x	
				2. verze			✓	
vývoj (VV)	9.4	Implementace ontologie	1. verze			x		
			2. verze			✓		
			3. verze			x		
	9.5	Testování				x	x	
	9.6	Naplnění znalostní báze instancemi				✓	✓	
9.7	Návrh uživatelského rozhraní a zpřístupnění báze					x	✓	
						✓	✓	
základní (ZV)	9.8	Vědecká studie	1. verze			x	✓	
			finální verze				x	✓
								✓

x	Plánované činnosti podle schváleného rámcového harmonogramu
✓	Skutečný termín splnění činnosti

Tab. 7 Harmonogram řešení projektu NAKI a přehled jeho plnění

### Řešitelé a jejich podíl na projektu

Řešitelka PhDr. Eva Bratková, Ph.D. zodpovídala za administrativu projektu. Garantovala řešení a zpracování modulu bibliografických entit znalostní báze. Soustředila se na problematiku systémů organizace znalostí z hlediska jejich typologie a struktury popisu.

Spoluřešitelka PhDr. Helena Kučerová zodpovídala za návrh architektury a datového modelu znalostní báze a za návrh ontologie organizace znalostí. Vedle řešitelský tým spolupracovníků, kteří se v jednotlivých fázích projektu podíleli na dílčích aktivitách a řešení. Jádrem týmu tvořili PhDr. Marcela Buřilová, PhDr. Linda Jansová, Ph.D., Ing. David Klimánek, Ph.D. a Mgr. Josef Schwarz.



Marcela Buřilová se podílela na zpracování výroků pro znalostní bázi. Provedla klasifikaci bibliografických entit podle klasifikačního schématu *KO Literature*. Podílela se na shromáždění a definování termínů pro výkladový slovník a na stanovení jejich vzájemných vztahů.

Josef Schwarz provedl testování správnosti a úplnosti výkladového slovníku. Navrhl strukturu faset a typologii predikátů pro ontologii organizace znalostí.

David Klimánek navrhl a implementoval databázové řešení editačního modulu znalostní báze včetně uživatelského rozhraní. Provedl experimentální ověření možností automatické extrakce termínů z textu a automatické klasifikace a klastrování, možností automatizovaných metod sémantické analýzy shromážděných popisných metadat a vybraných plných textů odborných publikací a jejich rejstříků.

Linda Jansová navrhla a realizovala zpřístupnění prototypu znalostní báze ve formátu propojených otevřených dat v systému *Drupal*. Navrhla schéma a algoritmus pro konstrukci *URI* a trojic *RDF*, navrhla aplikaci pro transformaci dat, navrhla a implementovala způsob zpřístupnění dat.

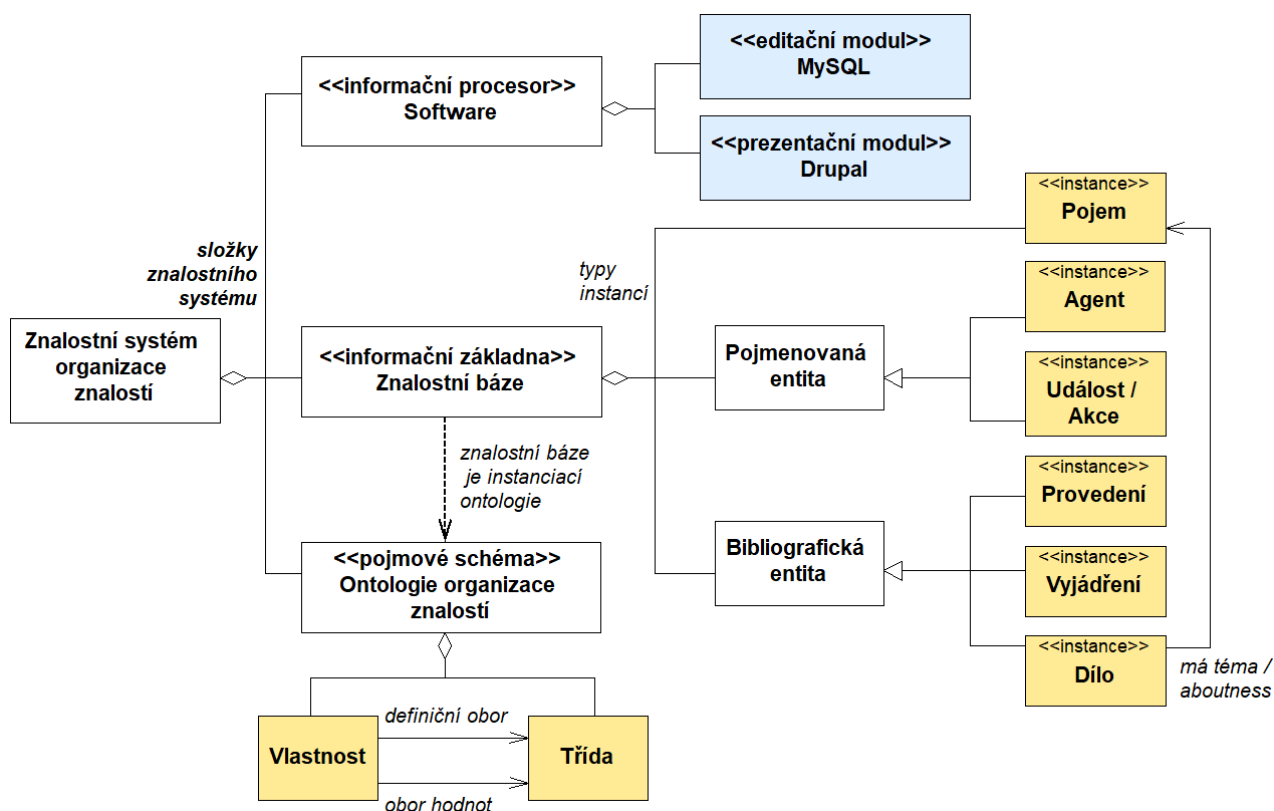
Nejdůležitější milníky a výstupy projektu spolu s údaji o odpovědných řešitelích jsou shrnuty v tabulce 8.

2013	Rozhovory s doménovými experty.	Helena Kučerová
	Bibliografická rešerše.	Eva Bratková
	Formulace kompetenčních otázek a případů užití.	Helena Kučerová
	Seznam termínů.	Marcela Buřilová
	Excerpce zdrojů.	Helena Kučerová
	1. verze pojmového modelu znalostní báze (diagram tříd, diagram aktivit).	Helena Kučerová
	Experimentální ověření možností automatické extrakce termínů z textu a automatické klasifikace a klastrování.	David Klimánek
	Návrh editačního modulu znalostní báze.	
2014	1. a 2. verze výkladového slovníku.	Marcela Buřilová Helena Kučerová
	Workshop k terminologickým otázkám oboru organizace znalostí.	Eva Bratková Helena Kučerová
	1. a 2. verze sémantické sítě.	Helena Kučerová
	1. verze ontologie.	Helena Kučerová Josef Schwarz
	Testování ontologie.	Josef Schwarz
	Plnění znalostní báze instancemi – 1. fáze.	Eva Bratková Marcela Buřilová Helena Kučerová
2015	Návrh uživatelského rozhraní a zpřístupnění báze.	Linda Jansová
	Plnění znalostní báze instancemi – dokončení.	Eva Bratková Marcela Buřilová Helena Kučerová
	Zpracování rukopisu vědecké studie.	Helena Kučerová

Tab. 8 Milníky a výstupy projektu *NAKI* a zodpovědní řešitelé

## 3.2 Architektura znalostní báze organizace znalostí

Ve znalostní bázi organizace znalostí je uplatněn předpoklad otevřeného světa (viz kapitola 1.5.5). Architektura znalostní báze je proto modulární, takže ji bude možné v budoucnu rozšiřovat a eventuálně měnit v reakci na nové požadavky na její využití. Východiskem konstrukce architektury se stala obecná koncepce informačního systému podle normy ISO/TR 9007, jež je popsána v kapitole 1.2. Na obrázku 7 jsou ve formátu diagramu tříd v jazyce UML znázorněny klíčové moduly znalostní báze včetně jejího začlenění do kontextu znalostního systému.



Obr. 7 Architektura znalostní báze organizace znalostí, založená na principech standardu ISO/TR 9007

Znalostní systém organizace znalostí tvoří datová část a procesní část (tj. software, který je v názvosloví ISO/TR 9007 označen jako informační procesor). Procesní část tvoří dva moduly – editační modul v relačním databázovém systému *MySQL* a prezentační modul v systému správy obsahu *Drupal* (viz kapitola 3.4.3). Datová část je v souladu s normou ISO/TR 9007 členěna na pojmové schéma a informační základnu.

Modul pojmového schématu tvoří doménová ontologie organizace znalostí, jejíž detailní popis je podán v kapitole 3.3.

Modul informační základny naplňuje znalostní báze organizace znalostí. Jejím obsahem jsou instance ve formě výroků, jež tvoří jednotky znalostí nebo popisná metadata. Množina instancí ve znalostní bázi se člení na tři navzájem propojené moduly: modul pojmů, modul pojmenovaných entit a modul bibliografických entit. Rozsah jednotlivých modulů zachycuje tabulka 9.

<b>Pojmy</b>		<b>443</b>			
<b>Pojmenované entity</b>	záznamy osob a korporací	<b>952</b>			
	záznamy akcí	<b>84</b>			
<b>Bibliografické entity</b>		<b>1188</b>	<b>Díla</b>	<b>Vyjádření</b>	<b>Provedení</b>
			986	1077	1188

Tab. 9 Rozsah znalostní báze

V modulu Pojem jsou obsažené ‚čisté znalosti‘, instance mají formát termínů a výroků vztahujících se k jejich významu.

Instance v modulu pojmenovaných entit mají formát popisných údajů o objektech reálného světa. Pojmenované entity jsou ve stávající verzi znalostní báze představovány dvěma moduly: Agent a Událost/Akce. Modul Agent zahrnuje osoby a instituce, jež mohou být asociovány s libovolnou entitou vztahem vytvoření, produkce, distribuce apod. Příkladem instance v modulu Agent je ‚Melvil Dewey‘ (osoba) nebo ‚Classification Research Group‘ (výzkumná skupina). Modul Událost/Akce zachycuje údaje o konání konferencí, projektů, o výzkumných aktivitách, publikování významných děl. Příkladem instance je ‚International Society for Knowledge Organization. Conference (13 : 2014 : Krakow, Polsko)‘. V další etapě budování báze se počítá s doplněním modulů pro nástroje, pomůcky a metodiky organizace znalostí (software, jazyky, formáty, pravidla, technologie), aktivity a procesy (tj. pro jednotky procedurálních znalostí, potřebných pro zdárný průběh aktivit a procesů realizovaných během organizace znalostí).

V modulu bibliografických entit jsou instance, jež mají formát popisných údajů (deskriptivních metadat) o dokumentech. Dokumentové zdroje, jejichž metadata jsou významnou součástí znalostní báze organizace znalostí, jsou v souladu s referenčním pojmovým modelem *FRBR* (resp. *IFLA LRM*) analyzovány a popisovány ze dvou pohledů – abstraktního a konkrétního (fyzického). Abstraktní úroveň reprezentují entity *FRBR* Dílo a Vyjádření. Dílo zachycuje ‚čistý‘ obsah, tj. sémantiku znalostí obsažených ve zdroji, a Vyjádření reprezentuje ‚čistou‘ formu, tj. obecný formát, jehož prostřednictvím jsou znalosti komunikovány. Entita Provedení popisuje konkrétní fyzické vlastnosti zdroje. Trojice entit Dílo, Vyjádření a Provedení představuje vzájemně propojený a logicky strukturovaný celek údajů o popisovaných zdrojích. Osamostatnění informací o díle a o vyjádření umožňuje jednak explicitně vyjádřit obsah (předmět) popisovaných dokumentů a začlenit jej do znalostní báze, jednak opakovaně použít tyto informace pro více zdrojů. Entity jsou vzájemně spojeny asociacemi vyjadřujícími skutečnost, že dílo je realizováno pomocí vyjádření a vyjádření je ztělesněno v provedení. Tomu, že jedno dílo může mít více vyjádření a jedno vyjádření více provedení, odpovídají i kvantitativní údaje v tabulce 9.

Ukázku hierarchie entit podle modelu *FRBR* a strukturovaných metadat pro bibliografické entity ve znalostní bázi obsahuje obrázek 8 na následující straně, na němž je znázorněna vizualizace údajů v prostředí *Drupal*.

## Dílo Dewey Decimal Classification (1876)

- Vyjádření Dewey Decimal Classification | Zkrácené vydání | Anglicky (1895)
  - Provedení Abridged Dewey decimal classification... (2012)
  - Provedení WebDewey... (2002)
  - Provedení Abridged Dewey Decimal Classification... (2004)
  - Provedení Abridged Dewey Decimal Classification... (1997)
  - Provedení Abridged Dewey Decimal Classification... (1990)
  - Provedení Abridged Dewey Decimal Classification... (1979)
  - Provedení Abridged Dewey Decimal Classification... (1971)
  - Provedení Dewey Decimal Classification... (1953)
  - Provedení Abridged Decimal Classification... (1945)
  - Provedení Abridged Decimal Classification... (1936)
  - Provedení Abridged decimal classification... (1929)
- Provedení Abri
- Provedení Abri
- Provedení Abri
- Vyjádření Dewey
  - Provedení Clas
- Vyjádření Dewey
  - Provedení Dew
  - Provedení Web
- Vyjádření Dewey
  - Provedení Dew

## Znalostní báze pro obor organizace informací a znalostí

Můj účet Odhl

Domů

Provedení Dewey Decimal Classification / Linked Data (2002)

Zobrazit

Upravit

**Hlavní název:** Dewey Decimal Classification / Linked Data

**Rok vydání:** 2002

**Bibliografický záznam (podle ČSN ISO 690):**

DEWEY, Melvil. Dewey Decimal Classification / Linked Data [online]. Dublin (Ohio): OCLC Online Computer Library Center, 2002- , [cit. 2014-03-17]. Zahrnuje zatím vybrané třídničky z 23. vydání. Dostupné z: <http://dewey.info/>.

**Poznámka:**

Tento záznam je záznamem volně dostupné aplikace OCLC, ve které systém experimentálně nabízí (v rámci prototypu databáze, <http://dewey.info/>) zpřístupňování záznamů třídniček Deweyho desetinného třídění v několika jazykových variantách a v různých formátech: HTML, zejména ale ve struktuře RDF/XML, Turtle, JSON aj. v rámci propojených otevřených dat (Linked Open data, LOD). Například: <http://dewey.info/class/6/2009-03/about.de.html> a <http://dewey.info/class/6/2009-03/about.de.rdf>. Vyhledávat lze záznamy vybraných třídniček 23. vydání DDT. Hierarchické procházení je zatím možné jenom do 3. úrovně. Každý záznam třídničky má vlastní URI. Kódování je v souladu se specifikací SKOS.

**Vydavatel:** OCLC

**Dostupnost:**

Dostupný volně v licenci CC BY-NC-ND z: <http://dewey.info/>.

**Vztahuje se k vyjádření:**

[Vyjádření Dewey Decimal Classification | Propojená data | Anglicky \(2002\)](#)

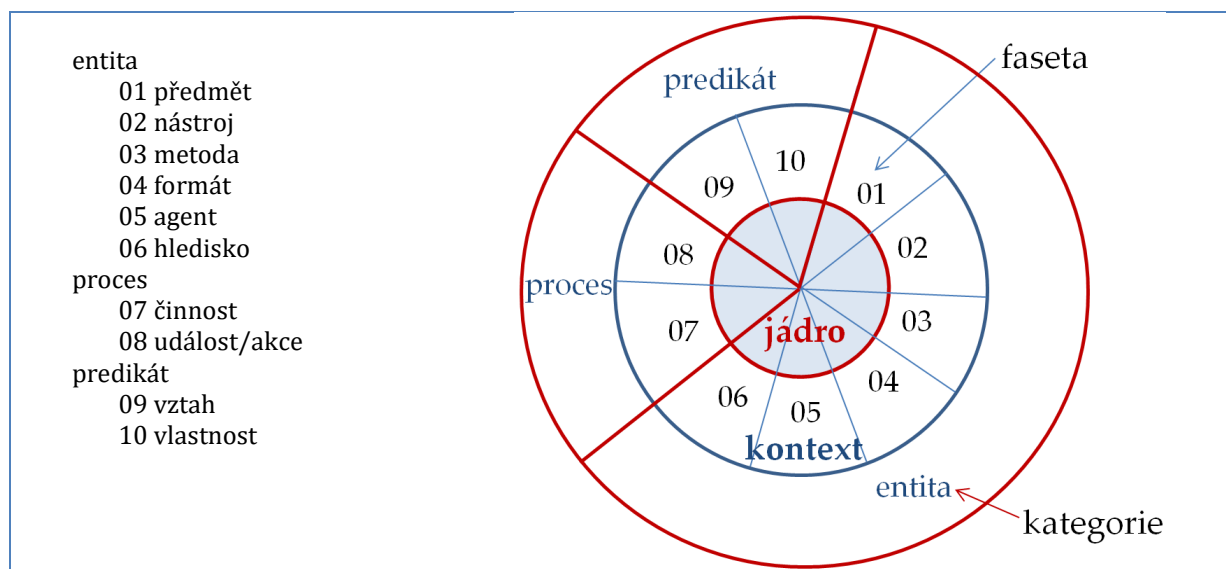
Obr. 8 Instance v modulu bibliografických entit

### 3.3 Ontologie pro znalostní bázi organizace znalostí

Tvorbu ontologie uložil řešitelům dílčí cíl 2 projektu, který předpokládá její využití ve funkci pojmového modelu znalostní báze. Cílem tedy bylo vymezit třídy objektů (entit), o nichž jsou ve znalostní bázi organizace znalostí zaznamenávány informace/znalosti, a specifikovat vlastnosti a vzájemné vztahy těchto tříd formou predikátů. Na obrázku 7 v kapitole 3.2 jsou tyto základní komponenty vyznačeny v rámci architektury znalostní báze jako modul tříd a modul vlastností. Třídy a predikáty jsou označeny dvěma typy označení. Označení vhodné pro strojové použití v sémantickém webu má formu *URI*, pro označení pojmu v mezilidské komunikaci bylo použito slovní pojmenování – termín (např. pojem označený termínem ‚automatická indexace‘ má *URI* <http://ko.cuni.cz/taxonomy/term/370>). Třídy mají svou intenzi popsanou textovou charakteristikou a definicí.

Ontologie tvoří pojmový základ pro všechny ostatní komponenty znalostní báze. Veškeré prvky znalostní báze lze považovat za instance některé z tříd či vlastností (predikátů) obsažených

v ontologii organizace znalostí. V případě popisných údajů o objektech reálného světa se jedná o standardní ontologický vztah třída–instance (tj. vztah instanciacie/ilustrace, popsany v kapitole 1.5.4). V případě formátu metadat o dokumentech je navíc uplatněn vztah třída–obsah / téma, známý z knihovnických systémů organizace znalostí jako indexace (viz kapitola 1.5.3 – konkrétně jde o asociativní vztah mezi instancemi z modulů ‚Dílo‘ a ‚Pojem‘). Základní obsahová struktura ontologie je znázorněna na obrázku 9.



Obr. 9 Páteřní taxonomie (pojmová struktura) ontologie

## Modul Třída

Struktura modulu tříd je tvořena tříúrovňovou hierarchií, kterou tvoří kategorie, fasety a pojmy/třídy.

Na nejvyšší úrovni jsou definovány tři základní kategorie: entita, proces, predikát. Tyto kategorie se odlišují mírou své existenční samostatnosti. Jediná entita může existovat samostatně. Proces potřebuje ke své existenci entitu (něco lze dělat pouze s něčím). Může existovat pouze ve spojení s nějakou entitou, ať už v pasivním nebo aktivním způsobu: například katalogizace (proces) dokumentů (entita) nebo automatická indexace, tj. indexace (proces) pomocí počítače (entita). Součástí definice pojmů z kategorie proces je tedy povinně vždy určení pojmů z kategorie entita, spojených s procesem. Predikát opět nemůže existovat samostatně a je vždy vztažen k nějaké entitě nebo procesu, případně k jinému predikátu.

Kategorie se člení na fasety. Při jejich definování byl použita metoda *5W1H*. Kategorie entita se člení na šest faset: fasety předmět, nástroj, metoda, formát, agent, hledisko. Kategorie proces obsahuje fasety činnost a událost/akce. Kategorie predikát obsahuje fasety vztah a vlastnost.

Faseta předmět vyjadřuje, ‚co‘ a ‚proč‘ se organizuje. Třídy v této fasetě zahrnují pojmy reprezentující předmět organizace (tj. zaznamenané znalosti – zdroje, bibliografické entity) včetně jeho prvků (např. lingvistické entity) a produktů (organizované zdroje, metadata). Významnými instancemi pojmů z této fasety jsou metadata z modulů Dílo, Vyjádření a Provedení v modulu Bibliografické entity znalostní báze.

Faseta nástroj se věnuje tomu, ‚čím‘ se organizuje. Třídy této fasety zachycují nástroje organizace (věci) včetně systémů organizace znalostí, jazyků (úroveň slov), softwaru, technologií.

Faseta metoda („jak“ se organizuje) obsahuje třídy popisující metody, pravidla a principy organizace (procesní pohled).

Faseta formát zahrnuje struktury, formáty a modely organizace. Obsahuje třídy vyjadřující strukturní a konceptuální prvky organizovaných objektů a produktů organizace na úrovni pojmů (např. „entita“, „třída“, „kategorie“).

Faseta agent („kdo“ / „co“ organizuje) se zaměřuje na subjekty organizace. Instance pojmů z této fasety zahrnují pojmenované entity z modulu Agent znalostní báze.

Faseta hledisko (hledisko / přístup) slouží ke specifikaci rámce, v němž daný pojem (resp. jeho význam) platí. Používá se i pro vymezení oborů a domén systémů organizace znalostí.

Faseta činnost („co“ se dělá v průběhu organizace znalostí) zahrnuje procesy a činnosti.

Faseta událost/akce („co“, „kde“ a „kdy“ se děje v průběhu organizace znalostí) se zaměřuje na události spojené s historií a rozvojem oboru, např. konference, výzkumné projekty a další.

Faseta vztah zahrnuje pojmy vyjadřující binární relace/predikáty. Vztah může existovat pouze mezi entitami, procesy nebo vlastnostmi.

Faseta vlastnost zahrnuje pojmy vyjadřující unární relace/predikáty. Vlastnost je vždy vztažena k nějaké entitě, procesu nebo vztahu.

Fasety obsahují třídy, jejichž aktuální množina je k dispozici na <http://ko.cuni.cz/seznam-pojmu>. Jádrem struktury každé fasety je generická hierarchie tříd/pojmů. Pojmy jsou definovány prostřednictvím nejbližšího nadřazeného pojmu a druhových specifik (vlastností a vztahů). Pojmy na nejvyšší úrovni jsou definovány axiomatically. Kromě své příslušnosti do faset jsou pojmy navíc rozděleny na dvě skupiny – jádro a kontext oboru organizace znalostí. Do jádra jsou zařazeny pojmy, jež jsou zkoumány výhradně v rámci oboru organizace znalostí, kontext tvoří pojmy, jež jsou předmětem zájmu i jiných vědních disciplín (např. lingvistika, informační věda). Výjimku tvoří nejobecnější pojmy, které jsou filozofickými kategoriemi, například „data (údaj)“, „informace“, „znalost“, „vztah“, „význam“ – ty jsou zařazeny do jádra, protože tvoří jakýsi svorník pro specializovanou pojmovou základnu oboru, bez nějž by jednotlivé pojmy nedávaly smysl. Přehled kvantitativního zastoupení tříd v jednotlivých fasetách podává tabulka 10.

	Počet tříd	z toho jádro
faseta předmět	50	41
faseta nástroj	75	66
faseta metoda	24	13
faseta formát	69	11
faseta agent	15	9
faseta hledisko	62	8
<b>celkem kategorie entita</b>	<b>295</b>	<b>148</b>
faseta činnost	84	58
faseta událost / akce	5	5
<b>celkem kategorie proces</b>	<b>89</b>	<b>63</b>
faseta vztah	22	7
faseta vlastnost	37	8
<b>celkem kategorie predikát</b>	<b>59</b>	<b>15</b>
<b>celkem všechny kategorie</b>	<b>443</b>	<b>226</b>

Tab. 10 Kvantitativní přehled pojmové struktury ontologie

## Modul Vlastnost

Modul vlastností obsahuje 87 predikátů. Základní strukturu tohoto modulu tvoří níže znázorněná několikaúrovňová hierarchie. Na nejvyšší úrovni se predikáty člení na skupinu paradigmatických a syntagmatických vztahů. Paradigmatické vztahy se dále člení na vztahy asociativní, ekvivalenční a hierarchické. Asociace jsou vnitřně členěny na symetrické a asymetrické. Významnou úlohu mají vztahy bibliografických entit, tj. realizace a ztělesnění. V závorkách jsou kurzívou uvedeny role v daném vztahu – pokud je vztah symetrický, je uvedena jen jedna role, u vztahů asymetrických jsou uvedeny dvě role.

### paradigmatický vztah

#### **asociace** (*má asociovanou entitu*)

asymetrická asociace (*má asymetrickou asociaci s*)

popis (*popisuje, má popis*)

předmět (*má předmět, je předmětem*)

vztahy bibliografických entit

realizace (*dílo–vyjádření: má realizaci, je realizací*)

ztělesnění (*vyjádření–provedení: má ztělesnění, je ztělesněním*)

symetrická asociace (*má symetrickou asociaci s*)

#### **ekvivalence** (*má ekvivalentní entitu*)

mapování (*je mapováno s*)

#### **hierarchie** (*má nadřazenou entitu, má podřazenou entitu*)

generický vztah (*je druhem, má druh*)

instanční vztah (*má instanci, je instancí*)

partitivní vztah (*je částí, má část*)

### syntagmatický vztah

hledisko/doména (*má hledisko/doménu, je hlediskem/doménou pro*)

role agenta (*má vztah k entitě v roli, má vztah k agentovi v roli*)

## 3.4 Implementace znalostní báze

Znalostní báze organizace znalostí je vytvořena jako ‚tradiční‘ expertní systém, jehož obsah je založený na manuální akvizici znalostí a na intelektuálním posuzování jejich relevance a hodnoty. V průběhu prvních dvou let řešení proběhlo experimentální ověření možnosti automatické extrakce termínů z textu a automatické klasifikace a klastrování. Byly ověřovány i možnosti automatizovaných metod sémantické analýzy shromážděných popisných metadat a vybraných plných textů odborných publikací a jejich rejstříků. Po vyhodnocení výsledků pilotního průzkumu řešitelský tým rozhodl dát přednost intelektuálním metodám akvizice i analýzy znalostí. Rovněž pro popis jednotek znalostí a výběr sémanticky relevantních interních i externích propojení obsahu byla zvolena metoda intelektuálního zpracování.

V kapitole 1.2 byly charakterizovány tři typy znalostních systémů: ‚tradiční‘ systémy umělé inteligence, systémy správy znalostí a sémantický web. Znalostní báze organizace znalostí je koncipována v souladu se stanovenými cíli projektu (viz dílčí cíl 4) pro využití v prostředí sémantického webu. Proto je ontologie znalostní báze založena na klíčové technologii sémantického webu, jíž je jazyk pro reprezentaci znalostí *RDF*. V souladu se standardem *RDF* jsou instance ontologických tříd používány jako subjekty a objekty výroků a ve funkci predikátů se používají instance vlastností.

Báze je zpřístupněna na webu ve formě propojených otevřených dat, která splňují všechny požadavky pětihvězdičkové struktury otevřených dat<sup>251</sup> v sémantickém webu:

<sup>251</sup> BERNERS-LEE, Tim. *Linked data – design issues* [online]. 2006-07-27, last change 2009-06-18 [cit. 2018-01-30]. Dostupné z: <http://www.w3.org/DesignIssues/LinkedData.html>. – *Stupně otevřenosti*

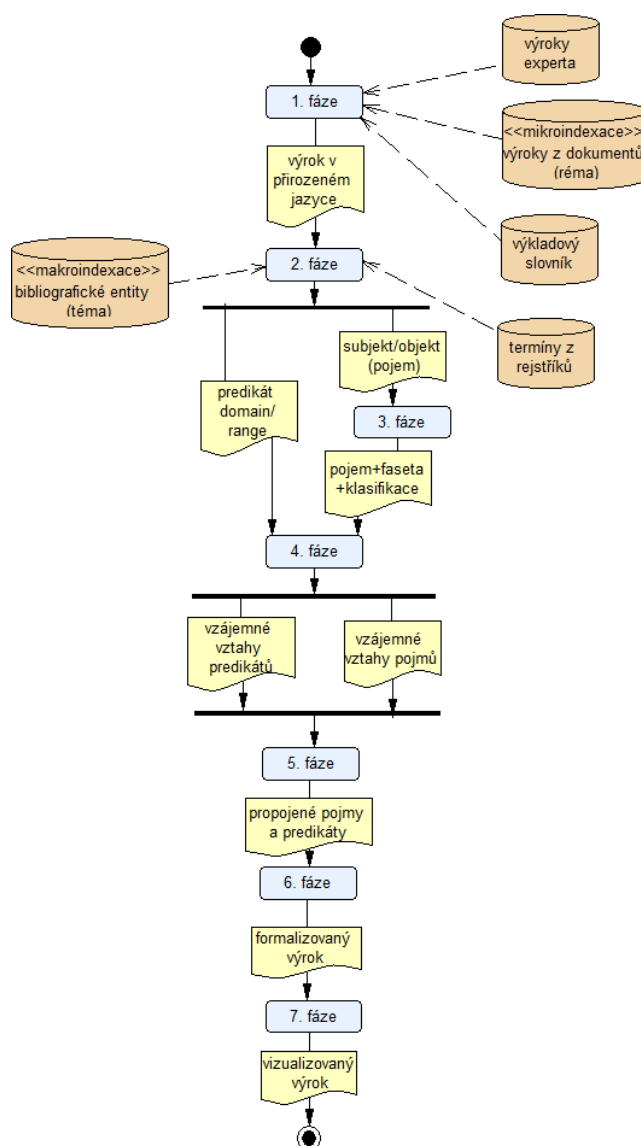


- data jsou dostupná na webu pod otevřenou licencí *Creative Commons BY-NC-ND*
- data jsou poskytována ve strojově čitelném strukturovaném formátu
- formát dat je otevřený (*RDF*)
- jako identifikátory objektů jsou použity *URI*
- data jsou pomocí odkazů propojena na jiná související data

Díky propojení obsahu a struktury znalostní báze s externími zdroji mají uživatelé zajištěn přístup i ke znalostem, jejichž zdroje se nacházejí vně znalostní báze (podrobně jsou jednotlivé typy propojení popsány v kapitole 3.4.2).

### 3.4.1 Pracovní postup pro naplnění znalostní báze instancemi

Pracovní postup (*workflow*) byl vytvořen přizpůsobením metodiky *UPON* pro specifika znalostní báze organizace znalostí, vyplývající ze stanovených cílů a specifikace požadavků. Jednotlivé fáze a jejich posloupnost znázorňuje obrázek 10.



Obr. 10 Pracovní postup pro naplnění znalostní báze instancemi

*otevřených dat* [online]. Praha: Ministerstvo vnitra ČR, 2015 [cit. 2018-01-30]. Dostupné z: <http://www.mvcr.cz/clanek/stupne-otevrenosti-otevrenych-dat.aspx>.



### 1. fáze – akvizice znalostí

Pro každý typ znalostních jednotek byla použita specifická metodika jejich akvizice a (intelektuálního) zpracování. Dokumentové zdroje byly shromážděny standardními rešeršními postupy a byly detailně zpracovány s využitím bibliografické heuristiky. Znalosti ve formě výroků byly manuálně excerpovány z vybraných odborných textů a ze záznamů rozhovorů s doménovými experty.

Proces: Vyjádření znalostí prostřednictvím výroků v přirozeném jazyce.

Výstup: Výrok v přirozeném jazyce.

Zdroje:

- výroky zpracované doménovým expertem
- výroky odvozené z výsledků obsahové analýzy vybraných odborných dokumentů (uplatnění principu *literary warrant*)
- definice termínů ve výkladovém slovníku (definice z *TDKIV*, vlastní formulace)

Kroky:

- Vytvoření výroku. Vyjádření znalostí větami přirozeného jazyka (např. *„Funkcí systému organizace znalostí je podpora procesů organizace znalostí a přístupu ke znalostem“*).
- Určení typu výroku: deklarativní – procedurální. Další fáze se týkají deklarativních výroků.

### 2. fáze – analýza znalostí

Proces: Nalezení a vyjádření subjektů, objektů a predikátů.

Výstup: Standardizované vyjádření pojmů a predikátů.

Zdroje:

- výroky v přirozeném jazyce
- termíny z rejstříků odborných publikací

Kroky:

- Nalezení subjektů a objektů (např. *„organizace znalostí“*).
- Doplnění seznamu subjektů a objektů o termíny z rejstříků odborných publikací.
- Standardizace vyjádření subjektů a objektů (1. pád sg, přirozený slovosled u víceslovných výrazů).
- Výběr preferovaných termínů, jež budou označovat pojmy.
- Nalezení predikátů.
- Standardizace vyjádření predikátů (určení definičního oboru a oboru hodnot).

### 3. fáze – indexace

Proces: Kategorizace a zatřídění pojmů a predikátů do faset a do tříd *Klasifikačního systému literatury oboru organizace znalostí* (klasifikace *KO Literature*). Rozdělení pojmů na 2 skupiny: 1. 'jádro' oboru organizace znalostí, 2. kontext oboru organizace znalostí.

Kroky:

- Definování faset (využití metodiky Denton – Spiteri).
- Kategorizace pojmů (tj. preferovaných termínů) do faset.
- Kategorizace pojmů do skupin – jádro a kontext.
- Klasifikace pojmů (tj. preferovaných termínů) do klasifikace *KO Literature*.
- Kategorizace predikátů.

Výstup:

Definované vztahy pojem–faseta a pojem–klasifikace.

Vytvoření dvou skupin pojmů – jádro a kontext oboru organizace znalostí.

Definovaný vztah predikát–kategorie.

Začlenění subjektů a objektů do ontologických tříd a jejich indexace podle klasifikačního systému *KO Literature*.

### 4. fáze – specifikace vztahů

Proces: Určení vzájemných vztahů pojmů a vzájemných vztahů predikátů: hierarchie, asociace.

Kroky:

- Definování generických hierarchických vztahů mezi pojmy z jádra.
- Definování asociativního vztahu k pojmu z jádra pro každý pojem z kontextu.
- Definování ostatních potřebných vztahů mezi pojmy z jádra a z kontextu.
- Definování vzájemných vztahů predikátů.
- Zápis vztahů do databáze.
- Automatické vytvoření odvozených vztahů.

Výstup:

Definované vzájemné vztahy pojmů a predikátů.

Propojení subjektů a objektů pomocí formalizovaných predikátů

## 5. fáze – propojení

Proces: Propojení pojmů a predikátů (podrobněji viz kapitola 3.4.2).

Vnitřní propojení:

- Pojem–Pojem, Pojem–Dílo, Pojem–Agent ad.

Vnější propojení:

- Propojení na externí pojmy a zdroje z relevantních jmenných prostorů
- Propojení na externí predikáty z relevantních jmenných prostorů

Výstup:

Začlenění pojmů a predikátů do linků *RDF* s informací o jejich vnitřním a vnějším propojení.

Propojení subjektů a objektů na externí zdroje z relevantních jmenných prostorů

## 6. fáze – formalizace

Proces: Tvorba formalizovaných výroků

Kroky:

Nahrazení provizorních definic (ty jsou přesunuty do poznámky) formalizovanými výroky ve formátu *RDF*, sestavenými z definovaných pojmů (ve funkci subjektů/objektů) a predikátů.

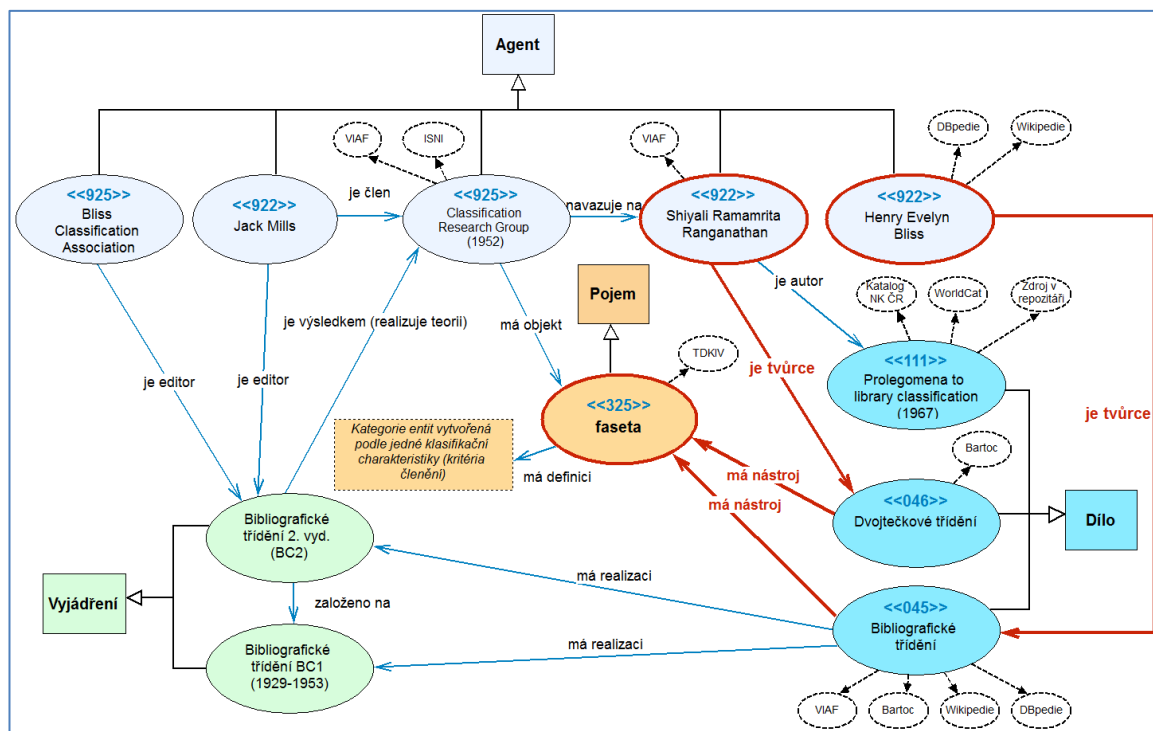
Výstup:

Formalizované výroky.

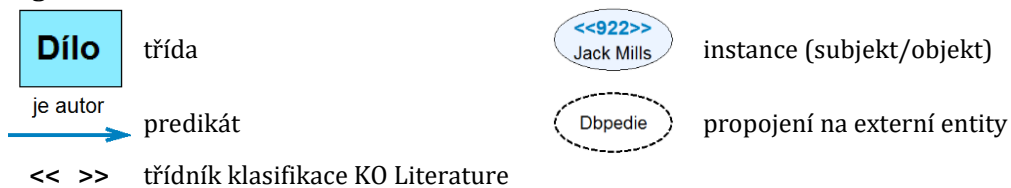
Tyto výroky jsou připraveny pro vložení do znalostní báze a po vygenerování identifikátorů *URI* jsou transformovány na trojice *RDF*. Výroky na sebe vzájemně navazují a vytvářejí tak *RDF* graf. Ilustrační ukázka části grafu je na obrázku 11. Tato ukázka zachycuje ve formalizované podobě obsah následujících výroků:

*Shiyali Ramamrita Ranganathan je autorem díla Prolegomena to library classification (1967) a je tvůrcem Dvojtečkové klasifikace, založené na principu faset. Na jeho myšlenky navazuje CRG (Classification Research Group), která byla ustavena v roce 1952 a věnovala se zkoumání fasetové analýzy. Henry Evelyn Bliss (1870–1955) je tvůrcem Bibliografického třídění. Člen CRG Jack Mills a Bliss Classification Association (BCA) jsou editoři 2. revidovaného vydání Blissova Bibliografického třídění (BC2), jež je založeno na principu faset a stalo se praktickou realizací (je výsledkem) teorií CRG. Faseta je kategorie entit vytvořená podle jedné klasifikační charakteristiky (principium divisionis). Obsah fasety je specifická a samostatná množina pojmů, patřících k jedné základní kategorii, jež umožňuje organizovat obsah podle kritéria členění daného těmito základními kategoriemi.*

Kromě výše uvedených výroků jsou na obrázku 11 znázorněna i externí propojení, jež budou podrobně popsána v kapitole 3.4.2. Červeně zvýrazněné subjekty, objekty a predikáty naznačují možnosti inference v grafu. Konkrétně znázorňují odvození asociace mezi dvěma tvůrci, jejichž díla jsou založena na stejném principu fasetové analýzy.



**Legenda:**



Obr. 11 Část grafu znázorňujícího formalizované výroky ve znalostní bázi

**7. fáze - vizualizace**

Proces: Vizualizace znalostí

Výstup: Uživatelsky přizpůsobené zobrazení zaznamenaných znalostí.

Ukázka aktuálního formátu uživatelsky přístupné vizualizace výroků v prostředí *Drupal* je na obrázku 12.

organizace znalostí (proces)
<b>Anglický ekvivalent:</b> knowledge organization
<b>Definice:</b> Jakákoli záměrná činnost, spočívající v zavádění struktury do existujících zaznamenaných znalostí s cílem usnadnit k nim přístup a umožnit jejich využití.
<b>Autor definice:</b> Kučerová, Helena
<b>Faseta:</b> faseta 7
<b>Typ termínu:</b> preferovaný
<b>Typ pojmu:</b> jádro
<b>Vztah nepreferovaný termín:</b> pořádání informací (proces)
<b>Vztah nadřazenost:</b> organizace (proces)
<b>Vztah podřazenost:</b> přístup správa (proces) uložení vyjádření (proces) zaznamenání / záznam (proces) zpracování (v užším slova smyslu)
<b>Termín v České terminologické databázi knihovnictví a informační vědy (TDKIV):</b> <a href="http://aleph.nkp.cz/F/?func=direct&amp;doc_number=000001604&amp;local_base=KTD">http://aleph.nkp.cz/F/?func=direct&amp;doc_number=000001604&amp;local_base=KTD</a>
<b>Termín v souboru národních autorit:</b> <a href="http://aleph.nkp.cz/F/?func=direct&amp;doc_number=000805421&amp;local_base=AUT">http://aleph.nkp.cz/F/?func=direct&amp;doc_number=000805421&amp;local_base=AUT</a>
<b>Znak z třídění Knowledge Organization Literature:</b> 113
<b>Vztah hledisko/doména:</b> informační věda
<b>Má cíl:</b> přístup použitelnost
<b>Má nástroj:</b> systém organizace znalostí
<b>Má výsledek:</b> struktura metadata
<b>Má objekt:</b> explicitní znalost
<b>Je objektem:</b> organizace znalostí (obor)

Obr. 12 Ukázka vizualizace formalizovaných výroků v prostředí *Drupal*

### 3.4.2 Propojení / mapování / linkování ve znalostní bázi

Pojmy, bibliografické entity i pojmenované entity ve znalostní bázi jsou propojeny jak navzájem, tak s dalšími entitami ve znalostní bázi i s externími entitami dostupnými na webu. Explicitní propojení představuje nejen významnou přidanou hodnotu k shromážděným znalostním jednotkám, ale lze jej v mnoha případech využít jako infrastrukturu pro odvozování nových poznatků.

Propojení uplatněná ve znalostní bázi se dělí do dvou základních skupin:

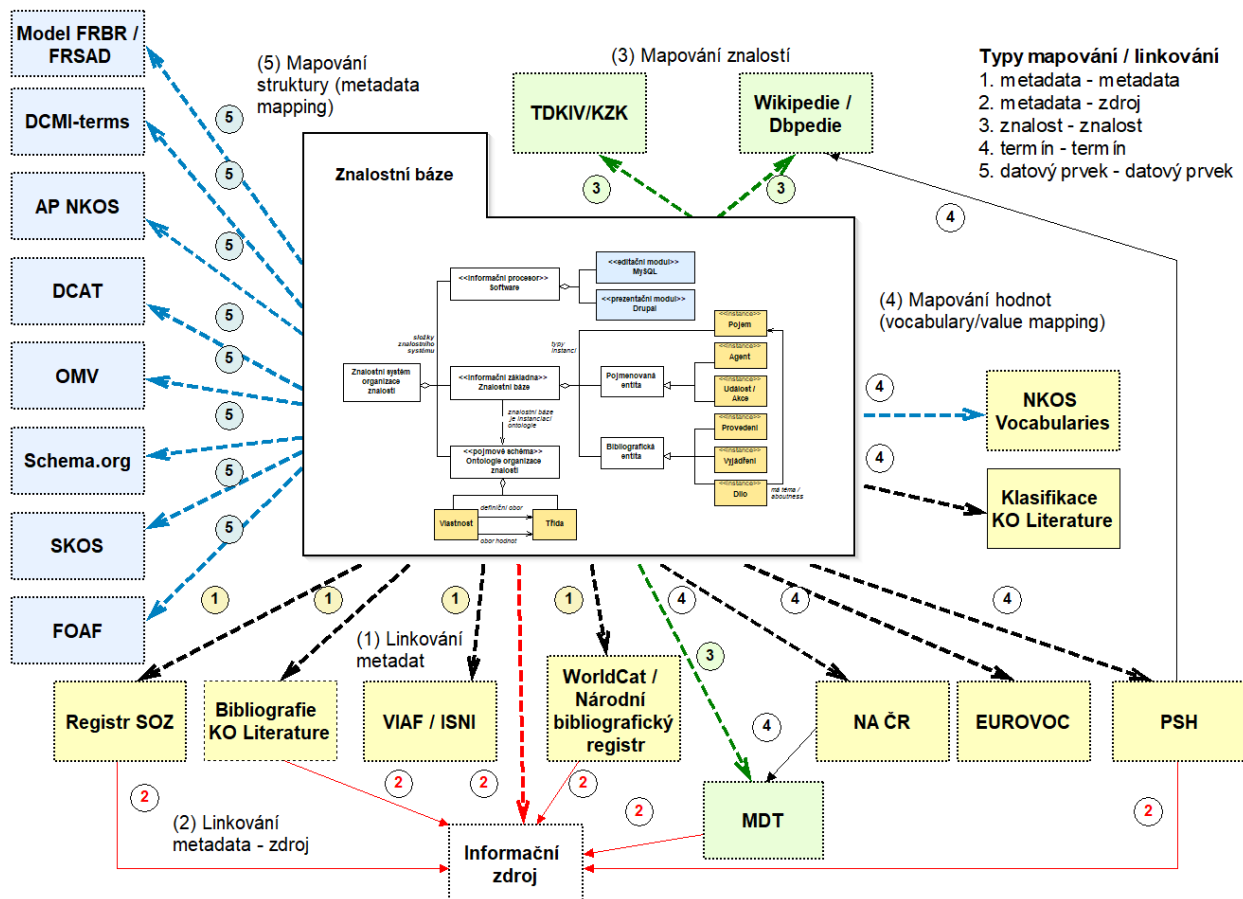
1. Interní propojení obsahu jednotlivých modulů znalostní báze (specifikace obsahu modulů je uvedena v kapitole 3.2).
2. Propojení obsahu modulů znalostní báze s externími entitami.

Nejvýznamnějšími typy interního propojení jsou propojení bibliografických entit podle modelu *FRBR* a propojení ontologických tříd s instancemi ve znalostní bázi.

Interní propojení podle modelu *FRBR* umožňuje vyjádření vzájemných vztahů mezi entitami 1. skupiny (dílo – vyjádření – provedení) prostřednictvím predikátů ‚realizace‘ a ‚ztělesnění‘ a vztahů mezi entitami 1. a 2. skupiny (entity se vztahem odpovědnosti k entitám 1. skupiny) prostřednictvím predikátu ‚role agenta‘. K dispozici jsou rovněž predikáty pro vzájemné propojení entit typu dílo–dílo (např. informace o tom, že dánská klasifikace *Decimalklassedeling* je založena na *DDT*, je instanciována pomocí predikátu ‚odvození‘), vyjádření–vyjádření (např. predikát ‚překlad‘ umožňuje vyjádřit, že úplné vydání *DDT* v němčině je překladem anglického úplného vydání) a provedení–provedení (např. o článku Melvila Deweyho *Decimal Classification beginnings*, publikovaném v roce 1920 v časopise *Library journal*, je pomocí predikátu ‚má ekvivalentní entitu‘ sděleno, že je obsahově ekvivalentní se stejnojmenným článkem, přetištěným v roce 1990 v témže časopise). Kromě agentů (entity 2. skupiny *FRBR*) mohou být bibliografické entity propojeny i s dalšími pojmenovanými entitami (např. článek Cyrila Cleverdona *The Cranfield tests on index language devices* je spojen predikátem ‚má asociovanou entitu‘ s akcí ‚Cranfieldské experimenty‘).

Propojení ontologických tříd s instancemi ve znalostní bázi je uplatněno u instancí modulů pojmenovaných entit a bibliografických entit, jež jsou s ontologickými třídami propojeny prostřednictvím termínů v modulu Pojem. Všechny osoby, instituce, události, akce a díla, o nichž jsou ve znalostní bázi uložena popisná metadata, byly obsahově analyzovány a výsledkem této analýzy je propojení popisných metadat s třídami v ontologii organizace znalostí. Používá se predikát ‚předmět‘.

Druhá skupina propojení, jež spojují instance ve znalostní bázi s externími zdroji, doplnila do báze další typ znalostních jednotek. Propojení byla vytvářena manuálně na základě intelektuálního posouzení hodnoty a věcné relevance každého zdroje. Na obrázku 13 je znázorněna znalostní báze organizace znalostí ve formě komponenty balíček v jazyce *UML*. Prostřednictvím přerušovaných šipek vedoucích ve směru od znalostní báze k externím entitám (třídy v *UML*) je vyznačeno pět typů propojení obsahu znalostní báze s externími zdroji/daty: 1. mapování metadat (interní metadata – externí metadata), 2. linkování metadat a zdrojů (interní metadata – externí zdroj), 3. mapování znalostí (interní znalost – externí znalost), 4. mapování termínů (interní termín – externí termín) a 5. mapování struktury (interní datový prvek – externí datový prvek).



Obr. 13 Typy externího mapování a propojování obsahu znalostní báze

## 1. Mapování metadat

Tato skupina mapování propojuje deskriptivní metadata z modulů Bibliografické entity a Pojmenované entity ve znalostní bázi se sémanticky ekvivalentními metadaty v externích zdrojích, jež aktuálně zahrnují tyto zdroje: souborný katalog *WorldCat*, národní bibliografické registry, registry systémů organizace znalostí, databáze autorit *VIAF* a *ISNI*. K propojení se používá predikát ‚má ekvivalentní entitu‘ (<http://ko.cuni.cz/taxonomy/term/5596>).

Specifickým typem propojení je propojení bibliografických metadat s třídíky specializovaného klasifikačního systému literatury oboru organizace znalostí (klasifikace *KO Literature*) prostřednictvím predikátu ‚má předmět‘ (<http://ko.cuni.cz/taxonomy/term/19>). Bibliografické entity jsou tedy kromě propojení s ontologickými třídami ještě indexovány podle této klasifikace. K duálnímu indexování bylo přikročeno proto, že zvolené nástroje se vzájemně doplňují. Pojmy v ontologii umožňují zachytit detailně réma (to, co je ve zdroji nové a specifické) a klasifikace umožňuje zařadit téma zdroje do širších skupin (kategorií).<sup>252</sup>

Příklady: Výroky vyjadřující propojení deskriptivních metadat o *Deweyho desetinném třídění (DDT)* ve znalostní bázi se záznamy *DDT* v registrech *Bartoc* (1), *TaxoBank* (2) a *VEST* (3) a se záznamem 9. vydání *DDT* v katalogu *WorldCat* (4) a indexace *DDT* pomocí klasifikace *KO*

<sup>252</sup> Východiskem pro toto pojetí se stal přístup Johna W. Hutchinse: HUTCHINS, W. John. The concept of 'aboutness' in subject indexing. In: *Aslib proceedings*. January 1978, 30(5), 172-181. doi:10.1108/eb050629. ISSN 0001-253X.

## Literature (5).

- (1) <http://ko.cuni.cz/node/255> <<http://ko.cuni.cz/taxonomy/term/5596>> <http://bartoc.org/en/node/241>
- (2) <http://ko.cuni.cz/node/255> <<http://ko.cuni.cz/taxonomy/term/5596>> <http://www.taxobank.org/content/dewey-decimal-classification>
- (3) <http://ko.cuni.cz/node/255> <<http://ko.cuni.cz/taxonomy/term/5596>> <http://aims.fao.org/vest-registry/vocabularies/dewey-decimal-classification-0>
- (4) <http://ko.cuni.cz/node/255> <<http://ko.cuni.cz/taxonomy/term/5596>> <http://www.worldcat.org/oclc/39096695>
- (5) <http://ko.cuni.cz/node/255> <<http://ko.cuni.cz/taxonomy/term/19>> <http://ko.cuni.cz/taxonomy/term/636>

## 2. Linkování metadat a zdrojů

Tento typ mapování zajišťuje přímé a nepřímé propojení deskriptivních metadat z modulu bibliografických entit ve znalostní bázi se zdroji (týká se především volně dostupných online zdrojů). Ve výrocích se používá predikát ‚popisuje‘ (<http://ko.cuni.cz/taxonomy/term/20>).

Příklady: Výroky v *RDF* vyjadřující údaje o dostupnosti digitalizované verze 1. vydání *DDT* (1), odkaz na přístup ke klasifikaci *Decimalklassedeling* ve formátu propojených otevřených dat (2).

- (1) <http://ko.cuni.cz/node/2262> <<http://ko.cuni.cz/taxonomy/term/20>> <http://archive.org/details/classificationan00dewerich>
- (2) <http://ko.cuni.cz/node/3151> <<http://ko.cuni.cz/taxonomy/term/20>> <http://opensource.dbc.dk/linked-data/dk5-linked-data>

## 3. Mapování znalostí (intenze pojmů)

Tato skupina mapování se věnuje mapování ontologických pojmů na externí zdroje. Zahrnuje propojení definic ontologických tříd a predikátů ve znalostní bázi na následující sémanticky ekvivalentní zdroje obsahující alternativní výklady a definice: terminologická báze *TDKIV*, databáze knihovnických zkratk *KZK*, články v české *Wikipedii*, heslo v české *DBpedii*. Používají se predikáty ‚má ekvivalentní entitu‘ (<http://ko.cuni.cz/taxonomy/term/5596>) a ‚je předmětem‘ (<http://ko.cuni.cz/taxonomy/term/5531>).

Příklady: Výroky v *RDF* vyjadřující propojení pojmu ‚umělá inteligence‘ s článkem ve *Wikipedii* (1), s heslem v *DBpedii* (2) a se záznamem v *TDKIV* (3).

- (1) <http://ko.cuni.cz/taxonomy/term/344> <<http://ko.cuni.cz/taxonomy/term/5531>> [https://cs.wikipedia.org/wiki/Um%C4%9Bl%C3%A1\\_inteligence](https://cs.wikipedia.org/wiki/Um%C4%9Bl%C3%A1_inteligence)
- (2) <http://ko.cuni.cz/taxonomy/term/344> <<http://ko.cuni.cz/taxonomy/term/5531>> [http://cs.dbpedia.org/page/Um%C4%9Bl%C3%A1\\_inteligence](http://cs.dbpedia.org/page/Um%C4%9Bl%C3%A1_inteligence)
- (3) <http://ko.cuni.cz/taxonomy/term/344> <<http://ko.cuni.cz/taxonomy/term/5596>> [http://aleph.nkp.cz/F/?func=direct&doc\\_number=000000137&local\\_base=KTD](http://aleph.nkp.cz/F/?func=direct&doc_number=000000137&local_base=KTD)

## 4. Mapování termínů

Mapování v této skupině rovněž propojuje ontologii na externí zdroje, tentokrát ovšem na úrovni termínů. Cílem je propojení termínů v modulu *Pojmy* s termíny označujícími významově ekvivalentní pojmy v následujících systémech organizace znalostí: národní autority ČR, *Polytematický strukturovaný heslář (PSH)*, tezaurus *EUROVOC*, klasifikace *KO Literature*, *NKOS Vocabularies*. Používá se predikát ‚má ekvivalentní entitu‘ (<http://ko.cuni.cz/taxonomy/term/5596>).

Příklady: Výroky v *RDF* vyjadřující propojení pojmu ‚umělá inteligence‘ s heslem v *PSH* (1), s termínem v tezauru *EUROVOC* (2), s heslem v souboru národních autorit (3) a s třídítkem

klasifikace *KO Literature* (4).

(1) <http://ko.cuni.cz/taxonomy/term/344> <<http://ko.cuni.cz/taxonomy/term/5596>>  
<https://psh.techlib.cz/skos/PSH12516>

(2) <http://ko.cuni.cz/taxonomy/term/344> <<http://ko.cuni.cz/taxonomy/term/5596>>  
<http://eurovoc.europa.eu/3030>

(3) <http://ko.cuni.cz/taxonomy/term/344> <<http://ko.cuni.cz/taxonomy/term/5596>>  
[http://aleph.nkp.cz/F/?func=direct&doc\\_number=000116536&local\\_base=Tem](http://aleph.nkp.cz/F/?func=direct&doc_number=000116536&local_base=Tem)

(4) <http://ko.cuni.cz/taxonomy/term/344> <<http://ko.cuni.cz/taxonomy/term/5596>>  
<http://www.isko.org/kolit.php?cl=157>

## 5. Mapování struktury

Předmětem mapování v této skupině jsou datové prvky konceptuálního modelu znalostní báze. Ty jsou mapovány na datové prvky následujících relevantních schémat: *FRBR*, *FRSAD*, *DCMI-terms*, *AP NKOS*, *DCAT*, *OMV*, *OWL*, *RDFS*, *Schema.org*, *FOAF*, *SKOS*. Používá se predikát ,je mapováno s' (<http://ko.cuni.cz/taxonomy/term/5597>).

Příklady: Výroky v *RDF* vyjadřující mapování predikátu ,má podřazenou entitu' s predikátem *skos:broader* (1) a mapování predikátu ,má předmět' s predikátem *dc:subject* (2).

(1) <http://ko.cuni.cz/taxonomy/term/41> <<http://ko.cuni.cz/taxonomy/term/5597>>  
<http://www.w3.org/2004/02/skos/core#broader>

(2) <http://ko.cuni.cz/taxonomy/term/19> <<http://ko.cuni.cz/taxonomy/term/5597>>  
<http://purl.org/dc/elements/1.1/subject>

### 3.4.3 Technologické řešení implementace znalostní báze

Pro všechny tři typy znalostních jednotek/instancí v modulu pojmů, v modulu pojmenovaných entit i v modulu bibliografických entit je použita jednotná technologie zpracování a publikování. Pro zpracování (zápis a editaci) je používána relační databázová technologie, konkrétně je použita databáze *MySQL* (verze 5.5.44-*MariaDB*) s webovým rozhraním tvořeným skripty v jazyce *PHP*. Toto řešení, jehož konceptuální model vychází z architektury znalostní báze, popsané v kapitole 3.2, poskytuje dostatečně robustní a zároveň flexibilní základnu pro uložení dat. Zajišťuje kontrolu validity dat a jejich vzájemných vazeb s minimalizovanou redundancí (propojení využívá referenční integritu s dvojicemi primární klíč – cizí klíč). Řešení je plně funkční a již od roku 2014 slouží k záznamu a k uchování údajů ve znalostní bázi.

Pro pohodlné vkládání a editaci dat, jež jsou v databázi rozložena do více tabulek, je vytvořeno uživatelské rozhraní, umožňující vyhledávání, třídění a filtrování údajů. Formuláře a reporty (sestavy) jsou vytvořené v jazyce *PHP* a po připojení k příslušnému *SQL* serveru jsou použitelné v jakémkoli webovém prohlížeči. Obsahují editační pole, rozbalovací seznamy povolených hodnot a pole pro zadávání dotazů. Umožňují zápis, kontroly správnosti, úplnosti a konzistence editovaných údajů. Slouží k zápisu pojmů a predikátů, k editaci výroků o bibliografických entitách a k tvorbě a editaci číselníků hodnot. Reporty propojují instance z různých modulů znalostní báze a instance s třídami v ontologii.<sup>253</sup>

Pro publikování a zpřístupnění ve formátu propojených otevřených dat je vytvořen funkční

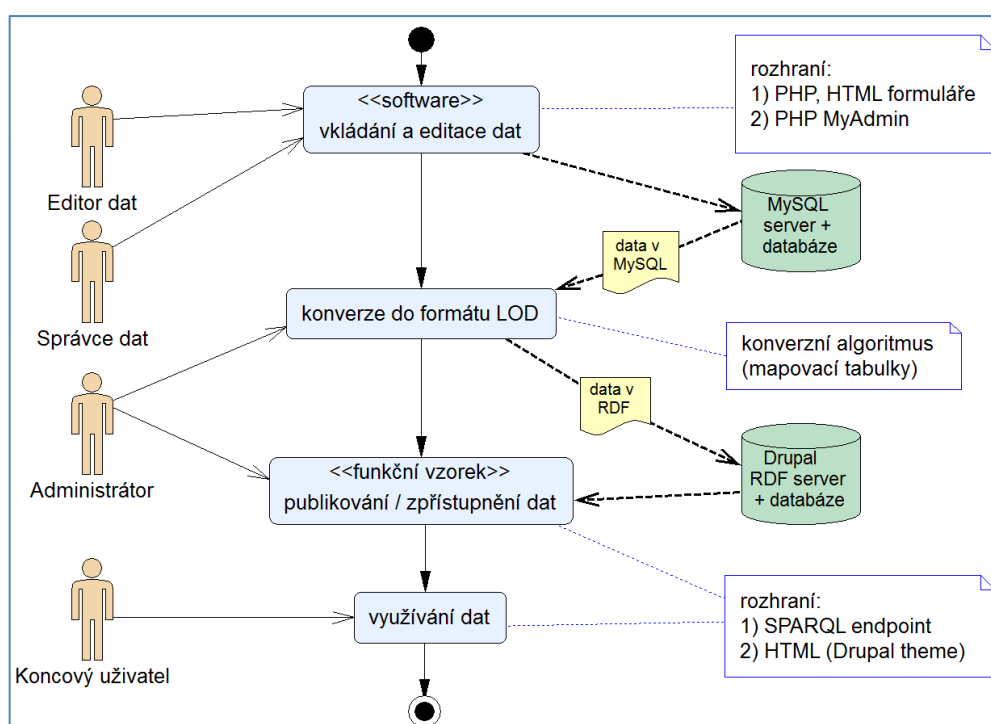
---

<sup>253</sup> KUČEROVÁ, Helena, KLIMÁNEK, David. *Uživatelské rozhraní pro zápis a editaci znalostních jednotek do znalostní báze pro obor organizace informací a znalostí* [software]. Praha: Univerzita Karlova v Praze, 2015-. Připraveno v systému *MySQL*, verze 5.5.8. Dostupné na základě přihlašovacích údajů z: <http://edit.ko.cuni.cz/>.



prototyp implementovaný v systému pro správu obsahu *Drupal*.<sup>254</sup> Při konverzi do formátu propojených otevřených dat (tj. do *RDF*) bylo třeba vytvořit *URI* pro každý objekt znalostní báze. Tato úloha byla realizována automaticky nástroji systému *Drupal*. Dále bylo zapotřebí ‚instanciovat‘ (explicitně definovat) vztahy, jež jsou implicitně vyjádřeny v relační databázi *MySQL* prostřednictvím referenční integrity (např. vztahy: Dílo <je realizováno pomocí> Vyjádření, Dílo <má jako předmět (je o)> Pojem). Další úlohou bylo vygenerování *RDF* linků pro mapování vztahů i datových prvků do relevantních externích jmenných prostorů.

Prototyp znalostní báze je hostován na webu *Univerzity Karlovy* v Praze. Technologické řešení báze je založeno na volně dostupných open source softwarových komponentách. Pro zpřístupnění znalostní báze byl vytvořen virtuální server s doménou *ko.cuni.cz* (*CentOS 7*, *Apache 2.4.6*, *PHP 5.4.16*, *MariaDB 5.5.44*, *Drupal 7.39*, přístup přes *SSH*). Prezentační modul je k dispozici na adrese <http://ko.cuni.cz/>. Jednotlivé komponenty a jejich návaznost na klíčové procesy tvorby a zpřístupnění znalostní báze jsou schematicky znázorněny na obrázku 14.



Obr. 14 Pracovní postup tvorby a zpřístupnění znalostní báze

Prototyp obsahuje všechny klíčové funkce stanovené ve specifikaci požadavků na znalostní bázi. Kromě vyhledávání a prohlížení obsahu umožňuje i editaci jednotek znalostní báze (funkce pro administrátory). Dostupnost dat pro stroje je zajištěna pomocí koncového bodu *SPARQL* (<http://ko.cuni.cz/sparql>). Dostupnost dat pro lidské uživatele zajišťuje standardní *HTML* rozhraní systému pro správu obsahu *Drupal*, v němž je znalostní báze uložena (<http://ko.cuni.cz/>).

<sup>254</sup> KALOU, Aikaterini K., KOUTSOMITROPOULOS, Dimitrios A., SOLOMOU, Georgia D. CMSs, linked data and semantics: a linked data mashup over Drupal for personalized search. In: Emmanouel Garoufallo, Jane Greenberg, ed. *Metadata and semantics research: 7th Research Conference, MTSR 2013, Thessaloniki, Greece, November 19-22, 2013. Proceedings*. Springer, 2013, s. 48–59. Communications in Computer and Information Science (CCIS), v. 390. doi:10.1007/978-3-319-03437-9\_6. ISSN 1865-0929. ISBN 978-3-319-03436-2 (Print). ISBN 978-3-319-03437-9 (Online).

## 3.5 Shrnutí výsledků projektu

### 3.5.1 Výsledky projektu

Projekt DF13P01OVV013 *Znalostní báze pro obor organizace informací a znalostí* byl úspěšně ukončen 31. 12. 2015 s hodnocením U – Uspěl podle zadání (s publikovanými či patentovanými výsledky atd.).

Byl vytvořen plně funkční prototyp znalostní báze naplněný daty, která poskytla zdrojový materiál pro zpracování původní české monografie<sup>255</sup> a pro aktualizaci české terminologie. Shromážděné znalosti představují reprezentativní a relativně úplnou množinu poznatků oboru, v jejímž rámci jsou akcentována lokální kulturní a jazyková specifika. Znalostní báze umožňuje ukládání, prohlížení a vyhledávání zaznamenaných doménových znalostí. Instance ve formě deskriptivních metadat popisují systémy organizace znalostí, osoby, instituce, události, aktivity, nástroje a další relevantní zdroje. Znalostní báze je průběžně aktualizována. Aktuálně je naplněna deklarativními znalostmi, vzhledem k jejímu modulárnímu charakteru bude ovšem v budoucnu možné připojit i modul pro znalosti procedurální. Architektura znalostní báze a prakticky otestované pracovní postupy pro naplnění znalostní báze instancemi jsou připraveny k opětovnému použití při návrhu doménových znalostníchází v jiných oborech.

O postupu prací při návrhu a implementaci znalostní báze byla průběžně informována odborná veřejnost na národní<sup>256</sup> i mezinárodní úrovni.<sup>257</sup> Vytvořená znalostní báze je veřejně zpřístupněna na webu ve formátu propojených otevřených dat, která splňují všechny požadavky pětihvězdičkové struktury otevřených dat v sémantickém webu. Díky propojení obsahu a struktury znalostní báze s externími zdroji mají uživatelé zajištěn přístup i ke znalostem, jejichž zdroje se nacházejí vně znalostní báze.

Kromě svého hlavního účelu jako zdroje pro zpracování původní české monografie slouží znalostní báze i jako východisko pro aktualizaci české odborné terminologie. Již v prvním roce řešení projektu dospěl řešitelský tým k zřetelnému vymezení obsahu i rozsahu pojmu organizace znalostí a souvisejících pojmů. Návrhy na aktualizaci české terminologie oboru organizace znalostí byly publikovány v odborné studii<sup>258</sup>, byly předloženy na redakční radě

---

<sup>255</sup> KUČEROVÁ, Helena. *Organizace znalostí: klíčová témata*. 1. vyd. Praha: Karolinum, 2017. 269 s. ISBN 978-80-246-3587-3 (brož). ISBN 978-80-246-3597-2 (pdf).

<sup>256</sup> KUČEROVÁ, Helena, BRATKOVÁ, Eva. Znalostní báze pro obor organizace znalostí. In: Vojtěch SVÁTEK, Ondřej ZAMAZAL, eds. *Znalosti 2014: Exhibice, Edukace a nacházení Expertů: 13. ročník konference, Jasná pod Chopkom, Nízke Tatry, 26. - 27. září 2014: Sborník příspěvků = Exhibition, Education and Expert finding: 13th Annual Conference, Jasná pod Chopkom, Nízke Tatry, Slovakia, 26th - 27th September 2014: Proceedings* [online]. 1. vyd. Praha: Oeconomica, 2014, s. 70-73. ISBN 978-80-245-2054-4. Dostupné též z: <http://hdl.handle.net/10760/24440> [cit. 2018-01-30].

<sup>257</sup> KUČEROVÁ, Helena, BRATKOVÁ, Eva. Knowledge base for subject area of knowledge organization: Issues, challenges, and preliminary outcomes of a research project [poster]. 2014. In: *Knowledge organization in the 21st century: Between historical patterns and future prospects: 13th International Conference (ISKO 2014)*, Jagiellonian University, Institute of Information and Library Science, Kraków (Poland), May 19th - 22nd, 2014. (Unpublished). Dostupné z: <http://hdl.handle.net/10760/23349> [cit. 2018-01-30].

<sup>258</sup> KUČEROVÁ, Helena. České termíny pro věcné pořádání informací po 40 letech: příspěvek k terminologické diskusi. In: *ProInflow: časopis pro informační vědy* [online]. 2013, 5(Speciál), [cit. 2018-01-30], s. 1-19. ISSN 1804-2406. Dostupné z: <http://www.phil.muni.cz/journals/index.php/proinflow/article/view/749/>.

*Terminologické databáze knihovnictví a informační vědy (TDKIV)* a byly projednány se zúčastněnými odborníky na workshopu, uspořádaném v roce 2014<sup>259</sup>.

Další oblastí využití znalostní báze je výuka a vzdělávání v oboru informační věda a knihovnictví. Například ve studijním roce 2016/2017 byl na *Ústavu informačních studií a knihovnictví FF UK* realizován seminář *Znalostní báze – návrh, implementace, správa* (kód předmětu AISV141).

### 3.5.2 Kritická analýza první verze ontologie

Po ukončení projektu byla první verze ontologie podrobena kritické analýze. Cílem bylo posoudit, do jaké míry je ontologie použitelná z hlediska stanoveného cíle disertace, tj. specifikace pojmové báze domény organizace znalostí. Bylo konstatováno, že při návrhu první verze ontologie bylo nutné respektovat skutečnost, že program *NAKI* má charakter aplikovaného výzkumu, který předpokládá konkrétní v praxi použitelné výstupy. V harmonogramu řešení proto nebylo možné vyčlenit dostatek času pro teoretická zkoumání. Výzkumné aktivity se v souladu se zadáním soustředily především na naplnění znalostní báze a při návrhu ontologie byla přijímána pragmatická řešení směřující k rychlému vytvoření prototypu.

Omezení první verze ontologie je dáno jejím úzce specializovaným zaměřením na funkci pojmového modelu znalostní báze organizace znalostí. Ze čtyř obecných oblastí použití ontologií, charakterizovaných v kapitole 1.3, se řešitelský tým soustředil na druhou oblast, tj. na využití pro opakované použití a organizaci znalostí. Toto omezení bylo v souladu se stanovenými cíli projektu znalostní báze a umožnilo v plánovaném čase úspěšně dokončit projekt. Na druhou stranu vedlo ke zúžení použitelnosti ontologie. První verze ontologie je využitelná jako pojmový model znalostní báze organizace znalostí, nikoli už ovšem jako pojmový model domény organizace znalostí.

Kvantitativní rozbor rozdělení tříd v ontologii na jádro a širší kontext (viz tabulka 10 v kapitole 3.3) poukázal na to, že přes zamýšlený doménový (tj. oborově specializovaný) charakter ontologie bylo nutné zahrnout do ní velké množství obecných pojmů – třídy označené jako kontext tvoří cca 50 % z celkového rozsahu. Tyto obecné pojmy, například ‚abstrakce‘, ‚analýza‘, ‚informační zdroj‘, ‚jazyk‘, ‚kategorie‘, ‚množina‘, ‚pojmem‘, ‚vztah‘ ad. byly řešiteli vnímány jako významné pro doménu organizaci znalostí, bylo je ovšem velmi obtížné definovat. Dále bylo konstatováno, že kategorizace pojmů v ontologii, popsána v kapitole 3.3, byla provedena víceméně intuitivním způsobem, bez významnější opory v některém teoretickém směru konceptualizace (viz kapitola 1.5.3). A konečně řešení vícestupňové hierarchie v páteřní taxonomii prostřednictvím dvou typů tříd – tzv. kategorií a tzv. faset, se v praxi neukázalo jako užitečné, resp. nebyly zaznamenány jeho přínosy.

Tyto zjištěné problémy bylo rozhodnuto v druhé verzi ontologie řešit prostřednictvím napojení na některou ze základních ontologií, jež řešení uvedených problémů nabízejí na potřebné teoretické úrovni ve formě jednoznačně zformulovaných axiomů.

---

<sup>259</sup> Informace o workshopu jsou dostupné v Centrální evidenci projektů (CEP) Informačního systému výzkumu, experimentálního vývoje a inovací z: <https://www.rvvi.cz/riv?s=jednoduche-vyhledavani&ss=detail&n=0&h=RIV%2F00216208%3A11210%2F14%3A10314683> [cit. 2018-01-30].

## Kapitola 4 Doménová analýza organizace znalostí

*Epistemologickým imperativem pro organizaci znalostí je přijetí doménově-analytického paradigmatu.*<sup>260</sup>

Richard Smiraglia, 2012

*Pátrání po jistotě, po bezpečné základně znalostí, musí být opuštěno.*<sup>261</sup>

Karl Raimund Popper, 1979

V této kapitole je dokumentován postup a jsou předloženy výsledky doménové analýzy organizace znalostí. Metoda doménové analýzy byla v kapitole 2.2 charakterizována jako metoda kvalitativního výzkumu, umožňující vytvořit komplexní model stanovené oblasti zájmu, použitelný pro poznávací nebo pragmaticky zaměřený účel. Důvody, které vedly k volbě doménové analýzy jako klíčové metody návrhu druhé verze ontologie organizace znalostí, jsou následující: 1) Komplexní problémově orientovaný pohled doménové analýzy odpovídá současnému trendu vědeckého výzkumu, kdy se přesouvá zájem od specializovaných vědních disciplín k multi-, inter- a transdisciplinarity. 2) Celostní pohled a dialektický charakter doménové analýzy je rovněž v souladu se systémovým přístupem. 3) Díky svému důrazu na sociální a kulturní kontext je doménová analýza považována za vhodnou metodu jak pro ontologické inženýrství, tak pro informační vědu a organizaci znalostí. 4) Výsledkem doménové analýzy je model domény, jehož konceptualizací a formalizací lze zkonstruovat doménovou ontologii. V kapitole 2.2 bylo konstatováno, že doménová ontologie je specifickým typem doménového modelu. Doménová analýza je tedy logickým předstupněm návrhu doménové ontologie.

Cílem provedené doménové analýzy je vytvoření pojmového modelu domény organizace znalostí. Předmětem zájmu je proces organizace znalostí a jeho realizační a kognitivní kontext, jehož součástí je i ustavující se obor organizace znalostí. Provedenou doménovou analýzu lze označit za instrumentální – slouží k tvorbě ontologie. Má umožňovat popis domény organizace znalostí, *as is*, nikoli formovat ji stylem *to be*, její charakter je tedy deskriptivní, nikoli preskriptivní.

Pravidla metody doménové analýzy jsou stanovena velmi volně – používá se jak pro návrh a tvorbu artefaktů (například tezaurů nebo ontologií), tak pro hodnocení, interpretaci a výzkum fenoménů či objektů již existujících. Námí použitý postup při doménové analýze kombinuje systémový přístup a vybraná hlediska doménové analýzy zformulovaná Birgerem Hjørlandem<sup>262</sup> (viz kapitola 2.2). Materiál pro analýzu poskytl rozsáhlý literární průzkum. Poznatky získané studiem a kritickou analýzou literárních zdrojů byly doplněny empirickými poznatky, zejména zkušenostmi z návrhu a implementace znalostní báze a první verze ontologie. Další zkušenosti byly načerpány při výuce oborově relevantních předmětů *Organizace informací a znalostí*, *Věcné pořádání informací a znalostí*, *Pořádání a vyhledávání informací*, *Organizace znalostí* a *Informační architektura* na *Vyšší odborné škole informačních služeb*, na *Ústavu informačních studií a knihovnictví FF UK* a na *Katedře informačních a znalostních technologií FIS VŠE*. Lze konstatovat,

<sup>260</sup> SMIRAGLIA, Richard P. Epistemology of domain analysis. In: SMIRAGLIA, Richard P., LEE, Hur-Li, ed. *Cultural frames of knowledge*. Würzburg: Ergon-Verlag, 2012, s. 120-121. ISBN 978-3-89913-918-1.

<sup>261</sup> POPPER, Karl Raimund. *Objective knowledge: an evolutionary approach*. Revised ed. Oxford: Oxford University Press, 1979, s. 37. ISBN 0-19-875024-2.

<sup>262</sup> HJØRLAND, Birger. Domain analysis in information science: eleven approaches – traditional as well as innovative. In: *Journal of documentation*. 2002, **58**(4), 422-462. doi:10.1108/002204102110431136. ISSN 0022-0418.

že aktivity realizované v závěrečné fázi doménové analýzy měly charakter teoretického výzkumu.

Pro literární průzkum byl využit obsah modulu bibliografických entit ve znalostní bázi organizace znalostí. V rámci řešení projektu *NAKI* provedli jeho řešitelé obsahovou analýzu zahrnutých zdrojů, jejíž výsledky jsou zachyceny v popisných metadatech (podrobněji o projektu viz kapitola 3, záznamy zdrojů jsou k dispozici na <http://ko.cuni.cz/node/3251>). Modul bibliografických entit zahrnuje cca 1 000 zdrojů v retrospektivě cca 30 let od počátku 80. let do současnosti, ze staršího období jsou zařazeny nejvýznamnější historické tituly. Jazykový profil tvoří z 67 % angličtina, 14 % čeština, 3 % francouzština, 3 % němčina, 2 % slovenština, 1 % ruština, 1 % španělština, 9 % ostatní jazyky. Snahou bylo zachytit v úplnosti zdroje v češtině, ostatní jazyky jsou pokryty výběrově. Podrobná analýza dostupnosti zdrojů, jejichž metadata jsou uložena ve znalostní bázi, byla publikována v samostatné studii Evy Bratkové *Otevřený přístup k vědecké literatuře: případová studie pro obor organizace informací a znalostí*<sup>263</sup>, která podává rozbor stavu k datu 1. 9. 2013 zaměřený na způsob přístupu.

Tato kapitola je členěna na sedm částí. V části 4.1 je v souladu s principy systémové analýzy specifikován cíl organizace znalostí. Část 4.2 je shrnutím hlavních témat, paradigmat a principů organizace znalostí. Svým zaměřením na analýzu znalostí v doméně odpovídá hledisku doménové analýzy, jež Birger Hjørland nazývá ‚epistemologické a kritické studie‘ a jež považuje za nejvýznamnější součást doménové analýzy. Tomuto významu odpovídá i poměrně značný rozsah dané části. Část 4.3 obsahuje pracovní vymezení jednotky organizace znalostí, část 4.4 je věnována procesu organizace znalostí a jeho klíčovému subprocesům. V části 4.5 jsou popsány nejvýznamnější komponenty realizačního kontextu procesu organizace znalostí: část 4.5.1 se zaměřuje na aktéry, čímž zhruba odpovídá Hjørlandovu hledisku ‚struktury a instituce ve vědecké a profesní komunikaci‘, část 4.5.2 analyzuje zdroje v doméně organizace znalostí a odpovídá tak hledisku ‚výzkum dokumentů a žánrů‘. Samostatná část 4.5.3 je věnována nejvýznamnějšímu doménovému zdroji, jímž je systém organizace znalostí. V šesté části je vymezen kognitivní kontext procesu organizace znalostí (část 4.6.1) a problematika jeho terminologie (část 4.6.2), v níž je aplikováno Hjørlandovo hledisko ‚terminologické studie, analýzy speciálního jazyka a diskurzu‘. Závěrečná část shrnuje výsledky doménové analýzy formou postupně formalizovaného modelu: v části 4.7.1 je textový model domény, část 4.7.2 obsahuje jeho semiformalizovanou verzi ve formě výroků a fasetové klasifikace klíčových pojmů. V části 4.7.3 jsou vztahy klíčových pojmů vizualizovány formou pojmové mapy.

Výsledky doménové analýzy organizace znalostí byly již v roce 2017 publikovány v monografii *Organizace znalostí: klíčová témata*.<sup>264</sup> Konkrétně obsah části 4.1 byl publikován v kapitole 2.1 *Cíl organizace znalostí – proč se organizuje?* (s. 58-64). Obsah části 4.2 je zahrnut v kapitole 1.2 *Problémy a paradigmat organizace znalostí* (s. 21-38) a v kapitole 3 *Principy organizace znalostí* (s. 105-161). Text a obrázek v části 4.3 byl publikován v kapitole 2.3.4 *Vymezení jednotky organizace znalostí* (s. 92-93), text a obrázek v části 4.4 byl publikován v kapitole 2.4 *Procesy organizace znalostí – co se děje při organizování?* (s. 94-103). Text a obrázek v části 4.5.1 lze najít

---

<sup>263</sup> BRATKOVÁ, Eva. *Otevřený přístup k vědecké literatuře: případová studie pro obor Organizace informací a znalostí*. In: *ProInflow: časopis pro informační vědy* [online]. 2013 [cit. 2018-01-30], 5(Speciál), s. 40-62. ISSN 1804-2406. Dostupné z: <http://www.phil.muni.cz/journals/index.php/proinflow/article/view/751>.

<sup>264</sup> KUČEROVÁ, Helena. *Organizace znalostí: klíčová témata*. 1. vyd. Praha: Karolinum, 2017. 269 s. ISBN 978-80-246-3587-3 (brož.). ISBN 978-80-246-3597-2 (pdf).

v kapitole 2.2 *Kontext organizace znalostí – kdo, kdy, kde organizuje?* (s. 64-69). Obsah části 4.6 byl publikován v kapitole 1.3 *Organizace znalostí jako obor* (s. 38-45) a v kapitole 1.4 *Pojmová základna a pojmová mapa organizace znalostí* (s. 46-50). Pro pojmovou mapu v části 4.7.3 bylo zvoleno stejné grafické řešení jako pro pojmovou mapu organizace znalostí na obrázku 1 na s. 47 monografie.

#### 4.1 Cíl organizace znalostí

Jak bude ukázáno v kapitole 4.4, patří organizace znalostí mezi „nepovinné“ procesy, které je možné z životního cyklu znalosti vypustit, aniž by se znemožnila komunikace. Organizace znalostí je jedním typem zpracovatelských procesů a jako každý takový proces vyžaduje pro svou realizaci určité zdroje – materiální, personální, finanční, časové. Je na místě ptát se, jaké přínosy (výstupy) má tento proces přinášet, abychom mohli mluvit o efektivním vynaložení často velmi nákladných zdrojů.

Účel organizace obecně je spatřován v tom, že usnadňuje přístup k organizovaným věcem. Cíl organizace znalostí, tak jak ho chápe většina autorů, specifikuje obecný účel organizace následovně: umožnit 1. přístup ke znalostem a 2. využitelnost znalostí. V eseji *Two kinds of power* (Dva druhy moci) používá Patrick Wilson pro tyto dva účely termíny, jež označují procesy umožňující daných cílů dosáhnout: popisování (angl. *describing*) a využívání (angl. *exploiting*).<sup>265</sup> Wilson považuje z hlediska rozvoje poznání samozřejmě za významnější druhý cíl, tj. využívání existujících znalostí a odvozování znalostí nových. Tato aktivita je ovšem podmíněna zajištěním přístupu. Z tohoto pohledu je tedy možné přístup ke znalostem považovat za primární cíl organizace znalostí.

Využitelnost znalosti lze spatřovat ve třech oblastech: poznání, řešení problémů a tvorba nových znalostí. Využití znalostí k poznání spočívá v jejich osvojení (internalizaci) subjektem a v jejich zařazení do kontextu znalostí, jimiž subjekt disponuje. Využití znalostí k řešení problémů se týká problémů praktických i teoretických a zahrnuje i využití znalostí k rozhodování. Využití existujících znalostí pro tvorbu nových znalostí a inovace (tj. odvození nových znalostí z toho, co bylo zaznamenáno) poukazuje na kreativní charakter znalosti – znalost není pouhou kopií vnímaného, ale je aktivně vytvářena subjektem.

Za přístupné a využitelné jsou považovány ty znalosti, jež vykazují alespoň jednu z následujících vlastností: nalezitelnost/vyhledatelnost, srozumitelnost a komunikovatelnost.

Pořadí, v němž jsou vlastnosti organizovaných znalostí uvedeny, není náhodné. Nalezitelnost je nejdůležitější: abychom mohli posoudit srozumitelnost a komunikovatelnost znalosti, musíme ji nejdříve mít k dispozici – tj. musíme být schopni ji rozpoznat a odlišit ji od ostatních, které nás obklopují, eventuálně ji vyhledat z úložišť, do nichž máme přístup. Nalezitelnost je vlastností umožňující jak efektivní zpětné vyhledávání znalostí uložených v nějaké kolekci (například v knihovně), tak efektivní prezentaci znalostí v nějakém kontextu (například na výstavě, v novinách, v dokumentu).

Srozumitelnost je sémantickou podmínkou pro přístupnost a využitelnost znalosti.

Srozumitelnost znalosti je možné podpořit vysvětlením jejího významu, například označením, definováním nebo rozšiřujícím výkladem. Další možností je uspořádání (například knih v knihovně, jmen zaměstnanců v seznamu, slov ve větě). Významným faktorem podporujícím

---

<sup>265</sup> WILSON, Patrick. *Two kinds of power: an essay on bibliographical control*. Berkeley: University of California Press, 1978, © 1968. 155 s. California library reprint series. ISBN 978-0-520-03515-7.

srozumitelnost znalosti je začlenění do kontextu, čímž se explicitně vyjádří její souvislost s jinými jevy.

Komunikovatelnost je technologickou podmínkou pro přístupnost a využitelnost znalosti. Zahrnuje umožnění komunikace organizovaných znalostí v nejširším slova smyslu, tj. přenos v prostoru, uchování v čase i sdílení znalostí, případně spolupráci (společnou činnost) umožněnou sdílenými znalostmi. Podmínkou komunikovatelnosti je, aby znalost byla k dispozici v nějaké fyzické formě, ať už analogové nebo digitální.

## 4.2 Hlavní témata organizace znalostí

Za reprezentativní souhrn témat organizace znalostí lze považovat encyklopedii organizace znalostí *ISKO (ISKO Encyclopedia of Knowledge Organization)*<sup>266</sup>, jejíž vydávání bylo zahájeno na podzim 2016. Záměrem autorů je pokrýt pojmy, disciplíny, organizace, systémy, standardy a teorie v rámci domény organizace znalostí. Encyklopedie je hostována na webovém sídle společnosti *ISKO* (viz kapitola 4.5.1) a je online dostupná v režimu otevřeného přístupu. Jejím producentem je vědecký poradní sbor *ISKO*, šéfredaktorem je Birger Hjørland a editorem je Claudio Gnoli. Články v encyklopedii jsou redigovány, počítá se s jejich aktualizací a verzováním. První verze každého článku je vždy publikována i v nově zřízené rubrice přehledů pojmů z organizace znalostí (*Reviews of concepts in knowledge organization*) časopisu *Knowledge organization*. K 11. 1. 2018 jsou zpracovány celkem 22 články, pokrývající mj. i dvě témata, jež Birger Hjørland<sup>267</sup> označuje za hlavní témata organizace znalostí – organizaci znalostí a systém organizace znalostí.

Pro vymezení tematické struktury domény organizace znalostí je zajímavé předběžné stanovení kategorií článků, s jejichž postupným naplněním editoři počítají:

Všeobecné a historické problémy organizace znalostí

Obor a příbuzné disciplíny

Hnutí a organizace

Biografické články (pouze nežijící osoby)

Klíčové pojmy organizace znalostí

Teoretické pojmy

Specifické typy dokumentů, žánrů a médií

Systémy organizace znalostí (SOZ)

Všeobecné problémy SOZ

Druhy SOZ

Specifické typy SOZ (všeobecné, univerzální)

Specifické typy SOZ (doménově specifické)

SOZ ve specifických doménách

Standardy a formáty pro reprezentaci dat

Procesy organizace znalostí

Metody, přístupy, filozofie

Organizace znalostí v různých kontextech a aplikacích

Výzkum, který provedli Kristie Saumureová a Ali Shiri kvalitativní obsahovou analýzou vybraného vzorku literatury evidované v databázi *LISTA* v letech 1966–2006, vymezuje obor prostřednictvím rozpoznávaných trendů v průběhu dvou období, nazvaných autory ‚před webem‘,

<sup>266</sup> *ISKO Encyclopedia of knowledge organization (IEKO)* [online]. HJØRLAND, Birger, GNOLI, Claudio, ed. ISKO Scientific Advisory Council, 2016- [cit. 2018-01-30]. Dostupné z: <http://www.isko.org/cyclo/>.

<sup>267</sup> HJØRLAND, Birger. Knowledge organization (KO). In: *Knowledge organization*. 2016, **43**(6), s. 475. ISSN 0943-7444.

tj. do roku 1993, a ‚po webu‘. V období 1966–1992 byla zaznamenána tato témata: indexace a tvorba abstraktů, tezaury, organizace znalostí s pomocí počítačů, katalogizace a klasifikace, kognitivní modely, vzdělávání. V letech 1993–2006 přibýly následující tematické okruhy: organizace podnikových informací, knihovny jako organizátoři webu, interoperabilita, klasifikace informací na webu, digitální uchovávání a digitální knihovny, metadata.<sup>268</sup>

Obdobné výsledky přinesl výzkum Fidelie Ibekwe-SanJuan a Erica SanJuan, realizovaný metodou automatické textové analýzy za použití text miningového nástroje *TermWatch*.<sup>269</sup> Automatický výběr témat a jejich shlukování bylo provedeno na vzorku 931 názvů a abstraktů článků, rozdělených na dvě dekády: 1988–1997 a 1998–2008. Společná témata obou dekád byla zastoupena pojmy znalost, informace, informační věda, organizace znalostí, tezaurus a klasifikace. Specifická témata první dekády představovaly bibliografické databáze, dokumentace a indexace. V druhé dekádě je vystřídal tématy metadat a webu.

V kapitole 2 jsme upozornili na to, že ontologické inženýrství a organizace znalostí sdílí analogické metody, konkrétně to bylo konstatováno o metodách doménové a fasetové analýzy. Toto tvrzení je možné rozšířit i na řešené teoretické problémy. Klíčová témata organizace znalostí na nejobecnější teoretické úrovni jsou analogická s těmi, jež byla popsána v kapitole 1.5 jako teoretická a filozofická témata informatických ontologií: podstata znalosti a informace, (ko)existence různých kategoriálních/klasifikačních systémů, různé přístupy k řešení/chápaní reprezentace a konceptualizace (včetně problému *aboutness*), určení typů, vlastností a vztahů organizovaných entit.

V souladu se základním východiskem systémové analýzy, kterým je analýza problému<sup>270</sup>, se při vlastní analýze témat v doméně rovněž zaměříme na klíčové problémy organizace znalostí a témata odvodíme z řešených problémů. Teoretické a praktické problémy a z nich vyplývající klíčová témata organizace znalostí rozčleníme do čtyř skupin, z nichž první dvě tvoří konstantní problémy a principy jejich řešení a druhou dvojicí jsou dobově podmíněná paradigmaty, konkrétně nově se formující paradigmaty pro počátek 21. století, a vznikající modely jejich řešení.

První skupina představuje trvalé problémy, jež bylo třeba řešit v každé historické etapě a budou zřejmě relevantní i v budoucnu: obtížné určování efektivnosti, nejednotnost a subjektivnost v chápaní obsahu/potřeby znalosti, proměnlivost obsahu/potřeby znalosti v čase, závislost na (přirozeném) jazyce a na společenském a kulturním kontextu, pragmatický aspekt organizace versus teoretické principy. Tato témata je možné považovat za konstantní charakteristiky domény.

Ve druhé skupině jsou uvedeny principy a techniky, které jsou trvalou metodologickou

---

<sup>268</sup> SAUMURE, Kristie, SHIRI, Ali. Knowledge organization trends in information and library studies: a preliminary comparison of the pre- and post-web eras. In: *Journal of information science*. 2008, **34**(5), 651-666. doi:10.1177/0165551507084300. ISSN 0165-5515 (Print). ISSN 1741-6485 (Online).

<sup>269</sup> IBEKWE-SANJUAN, Fidelia, SANJUAN, Eric. Knowledge organization research in the last two decades: 1988-2008. In: *Paradigms and conceptual systems in knowledge organization: Proceedings of the Eleventh International ISKO Conference, 23-26 February 2010, Rome, Italy*. Organized by the Italian chapter of ISKO and the Faculty of Philosophy, Sapienza University, Rome. Ed. by Claudio GNOLI and Fulvio MAZZOCCHI. Würzburg: Ergon Verlag, © 2010, s. 115-121. Advances in knowledge organization, vol. 12, ISSN 0938-5495. ISBN 978-3-89913-746-0.

<sup>270</sup> HABR, Jaroslav, VEPŘEK, Jaromír. *Systémová analýza a syntéza: zdokonalování a projektování systémů*. 2. přeprac. vyd. Praha: SNTL, 1986, s. 86.



základnou organizace znalostí: principy ekvivalence a identifikace, kategorizace a klasifikace, definování a řazení, techniky umístění a označení.

Paradigmatem nazývá Thomas S. Kuhn vymezení a způsob řešení dobově podmíněných problémů: „*Za paradigma považuji obecně uznávané vědecké výsledky, které v dané chvíli představují pro společenství odborníků model problémů a model jejich řešení.*“<sup>271</sup> Podle Kuhna tedy existují dva typy paradigmat, jež lze vysledovat i v oblasti organizace znalostí: 1) představa o tom, jak věci fungují (model problémů – gnozeologické paradigma), a 2) metody a techniky, jak věci dělat (model řešení problémů – inženýrské paradigma). Obdobný přístup uplatňuje Jiří Cejpek, který hovoří o dvou paradigmatech informační vědy: model problémů označuje jako „*paradigma existenciální, ontogenetické či holistické (celostní)*“ a model řešení problémů jako „*instrumentálně užitelné paradigma*“.<sup>272</sup> Paradigmata modelující problémy jsou obvykle formulována v teoretických textech významných odborníků a paradigmata řešení problémů lze často najít v normativních dokumentech, tj. ve standardech a pravidlech. Pro oblast organizace znalostí představují v současnosti taková paradigmata řešení problémů katalogizační principy a pravidla, standardy pro obsahovou analýzu a terminologickou práci a standardy pro vývoj a správu tezaurů (viz kapitola 4.5.2).

V souladu s tímto členěním paradigmat obsahuje třetí skupina model problémů, které aktuálně představuje změna společenského kontextu a komunikačních technologií využívaných při organizaci znalostí. Současnou vývojovou etapu organizace znalostí lze v souladu s teorií Thomase S. Kuhna označit za krizi paradigmatu: paradigmata oboru, zformovaná v 19.–20. století, se vlivem změněného společenského kontextu a nových komunikačních technologií již jeví jako nedostatečná a obor se formuje na nových základech. Faktory určující podobu nového paradigmatu budou v praktické oblasti představovat znalostní ekonomika, globalizace a multikulturalismus, síťový model komunikace, elektronické dokumenty a propojená data.

Ve čtvrté skupině jsou charakterizovány změny modelů řešení těchto problémů: změny v aktivitách, změny aktérů a změny metodologických přístupů. V aktivitách se klade důraz na dostupnost zdrojů, aktivní návrh nových systémů, řešení problémů informačního zahlcení a důvěryhodnosti zdrojů. V teoretické rovině půjde o přechod od pragmatického řešení každodenních problémů k vědeckým metodám zkoumání problematiky, doprovázený rozšířením zájmu na všechny typy institucí i jednotlivců a snahou o integraci a interoperabilitu. Specifickým problémem k řešení je transdisciplinarita, jež se týká ve stejné míře jak organizace znalostí jako oboru, tak současného způsobu zkoumání a poznávání světa obecně.

Přehled jednotlivých problémů a jejich typů zachycuje tabulka 11 na následující straně. V následujícím přehledu jsou jednotlivá témata komentována podrobněji.

---

<sup>271</sup> KUHN, Thomas Samuel. *Struktura vědeckých revolucí*. Dotisk 1. vyd. Praha: OIKOYMENEH, 2008. Předmluva, s. 10. ISBN 80-86005-54-2.

<sup>272</sup> CEJPEK, Jiří. Knihovnictví v širších souvislostech. In: *Národní knihovna*. 1999, **10**(6), s. 276. ISSN 0862-7487.

a) Konstantní problémy organizace znalostí	1. Obtížné určování efektivity	
	2. Nejednotnost v chápání obsahu/potřeby znalosti	
	3. Proměnlivost obsahu/potřeby znalosti v čase	
	4. Lingvistické a sémantické problémy	
	5. Pragmatický aspekt organizace versus teoretické principy	
b) Řešení konstantních problémů organizace znalostí	6. Princip ekvivalence a identifikace	
	7. Princip kategorizace a klasifikace	
	8. Princip definování	
	9. Princip řazení	
	10. Technika umístění	
	11. Technika označení	
c) Aktuální problémy organizace znalostí	Společenský kontext	12. Od industriální ekonomiky ke znalostní ekonomice, od informačních zdrojů ke zdrojům ekonomickým
	Komunikační technologie	13. Od idejí národního státu ke globalizaci a multikulturalismu
		14. Od lineárních komunikačních a datových modelů k modelům síťovým
		15. Od analogových dokumentů k digitálním
		16. Od monolitických statických dokumentů k propojeným datům, od makroorganizace k mikroorganizaci, od dokumentů k informacím
		17. Od koncentrace na uchování a ochranu dokumentů k zajištění dostupnosti dokumentů/znalostí
d) Řešení aktuálních problémů organizace znalostí	Převažující aktivity	18. Od pasivního popisu k aktivnímu návrhu nových systémů
		19. Od nedostatku informací k řešení problému informačního zahlcení a důvěryhodnosti
		20. Od technicistního a kulturně historizujícího přístupu k vědeckému přístupu, od praxe k teorii
		21. Rozšíření okruhu osob i institucí, jež se věnují organizaci znalostí
Aktéři	Metodika	22. Od izolovaných specifických aktivit k integrovanému přístupu
		23. Od hledání jednoho ‚nejlepšího‘ systému k propojení systémů stávajících
		24. Od specializovaných vědních disciplín k transdisciplinaritě

Tab. 11 Problémy organizace znalostí

## a) Konstantní problémy organizace znalostí

### 1. Obtížné určování efektivity

Organizace znalostí se vždy realizuje v konkrétním ekonomickém prostředí a v jeho rámci je zvykem při určování efektivity poměřovat náklady versus přínosy. Zatímco ovšem proces organizace spotřebovává snadno rozpoznatelné a kvantifikovatelné hmotné zdroje, jeho přínosy jsou nehmotné a obtížně definovatelné. Navíc je potřeba je hledat nikoli na výstupu z procesu organizace, ale minimálně ‚o proces dále‘, ve výsledcích procesu vyhledávání nebo využívání znalostí, eventuálně dokonce až v procesu vytvoření nové znalosti. Další problém představuje stanovení ukazatelů efektivity. Ty mohou být kvantitativní nebo kvalitativní, přičemž i v případě kvalitativních ukazatelů je žádoucí jejich vyjádření pomocí kvantitativních údajů.

Obecnými fyzikálními veličinami umožňujícími kvantitativně vyjádřit efektivity organizace jsou prostor a čas. Prostor umožňuje lokalizovat a kvantifikovat, kde (jak daleko) je to, co hledáme, čas určuje, za jak dlouho se lze k hledanému dostat. Tyto ukazatele jsou ovšem použitelné pouze pro materiální entity.

Nejčastěji používaným kvalitativním způsobem určení efektivity organizačního systému jsou ukazatele, poprvé použité v průběhu Cranfieldských experimentů<sup>273</sup> – míra přesnosti a míra

<sup>273</sup> CLEVERDON, Cyril W., MILLS, Jack. The testing of index language devices. In: *Aslib proceedings*. April 1963, 15(4), 106-130. doi:10.1108/eb049925. ISSN 0001-253X.

úplnosti dosahované při vyhledávání. Oba ukazatele jsou založené na posuzování relevance získaných informací. Přesnost (angl. *precision*, někdy zjednodušeně označovaná jako relevance) označuje podíl relevantních jednotek na vyhledané množině. Úplnost (angl. *recall*) se snaží o vyjádření podílu vyhledaných relevantních jednotek z celkového počtu relevantních jednotek přítomných v prohledávaném souboru. Přirozené přání dosáhnout co nejvyšší přesnosti a současně co nejvyšší úplnosti naráží na prokázaný inverzní poměr těchto dvou veličin – čím vyšší je přesnost, tím nižší úplnost a naopak. Obdobně dopadne snaha o dosažení maximální efektivnosti zpracování: požadavek ‚vyčerpávající‘ indexace (zachycení všech dílčích témat zdroje, angl. *exhaustivity*) vede k vysoké úplnosti a nízké přesnosti při vyhledávání; požadavek na vysokou specifičnost indexace (zachycení ‚nejuzšího‘ pojmu vystihujícího obsah zdroje, angl. *specificity*, v češtině se někdy označuje jako hloubka indexace) zvýší přesnost a sníží úplnost.<sup>274</sup>

## 2. Nejednotnost v chápání obsahu/potřeby znalosti

Organizace znalostí usnadňuje přístup ke znalostem tehdy, když podporuje nalezení shody mezi obsahem znalosti a její potřebou. Avšak vzhledem k tomu, že chápání významu je ovlivněno i subjektivními a kontextovými faktory, je nalezení takové shody obtížné. Kupříkladu v Laswellově lineárním modelu sociální komunikace jsou zahrnuty tři různé ‚obsahy‘ sdělení – smysl sdělení pro komunikátora, smysl sdělení pro komunikanta a ‚objektivní‘ komuniké. I výše uvedené způsoby měření relevance jsou přes svou zdánlivou exaktnost zatížené subjektivností v klíčové fázi určování relevance. Navíc konzistence je obtížné dosáhnout nejen mezi vyjádřeným obsahem a potřebou znalosti, ale i mezi přístupy různých tvůrců, zpracovatelů i uživatelů.

Chápání významu ovlivňuje i společenské a kulturní prostředí. Souvisejícím problémem je tzv. zaujetí (angl. *bias*). Mnozí autoři tvrdí, že se mu nelze vyhnout a že při organizaci znalostí je vždy nutné zaujmout nějaký postoj, resp. pohled z určitého hlediska.

## 3. Proměnlivost obsahu/potřeby znalosti v čase

Změny informačních potřeb a změny chápání obsahu informačních zdrojů v čase jsou objektivním faktem. Otázkou je, zda na tyto změny reagovat reorganizací zdrojů (spojenou s dalšími náklady), či zda ponechat uspořádání v historické podobě už neodpovídající současným potřebám. Ani kompromisní řešení v podobě zachování historických principů pro stávající kolekci a nové organizace aktuálních přírůstků není z hlediska uživatele ideální, protože komplikuje přístup ke zdroji nutností používat různé přístupové techniky pro jeho dílčí části.

Navíc – na rozdíl od klasifikace věd, která modeluje stav poznání aktuální v době svého vytvoření, by měl systém organizace znalostí být připraven i na pořádání znalostí, jež v době jeho vzniku ještě neexistují. A. C. Foskett upozorňuje, že „v každé předmětné oblasti bude nějaký akceptovaný korpus znalostí, ale každý dokument, který indexujeme, ho může změnit“.<sup>275</sup>

## 4. Lingvistické a sémantické problémy

Tato sada problémů úzce souvisí se sémiotickou problematikou vztahu věcí, pojmů a jejich označení, tj. s problémy reprezentace a konceptualizace, diskutovanými v kapitolách 1.5.2 a 1.5.3. Klíčovou technikou organizace zaznamenaných znalostí je provedení obsahové analýzy a vyjádření jejích výsledků v nějakém jazyce. Použití přirozeného jazyka k těmto účelům je na jedné straně výhodné pro uživatele, který tento jazyk zná, na druhé straně je komplikováno

<sup>274</sup> BROUGHTON, Vanda. *Essential classification*. London: Facet, 2004, s. 70-76. ISBN 978-1-85604-514-8.

<sup>275</sup> FOSKETT, Anthony Charles. *The subject approach to information*. 5th ed. London: Library Association Publishing, 1996, s. 29. ISBN 978-1-85604-048-8.

nejednoznačností (synonymie, polysémie a homonymie) a vágností významu. Další sémantický problém představují složená témata, vyjádřená složenými pojmy. K tomu se přidává problém mnohोजazyčnosti, existence velkého množství přirozených jazyků. Tvorba a používání umělého (formalizovaného) jazyka pak opět vyžaduje dodatečné investice jak ze strany jeho tvůrců, tak ze strany uživatelů.

## 5. Pragmatický aspekt organizace versus teoretické principy

Informační zdroje odrážejí realitu, jež se ve své rozmanitosti mnohdy vymyká z formálních pravidel, na nichž jsou založeny organizační principy. Tento fenomén zahrnutí teoreticky ‚nesprávných‘ jevů do praxe organizace znalostí je v anglické terminologii označován slovem *warrant* (zdůvodnění, ospravedlnění, oprávnění). Tedy z reality jsou přijímány jen ty jevy a pravidla, jež potřebujeme pro organizaci znalostí a potřebujeme-li pro organizaci znalostí jevy nebo pravidla, jež neexistují v realitě nebo nejsou považovány za pravdivé (například geocentrismus v současné astronomii), doplníme je. Ve filozofické rovině rozpracoval pojem *warrant* Alvin Plantinga<sup>276</sup>, do praxe organizace znalostí jej uvedl v roce 1911 Edward Wyndham Hulme<sup>277</sup>. Tento princip koresponduje s konceptem deskriptivních ontologií (viz kapitola 1.5.4).

Nejznámějším typem je tzv. *literary warrant*, zahrnující do systémů organizace znalostí nikoli jevy reality, ale témata, o nichž se píše v literatuře. Představuje vlivnou koncepci oboru, doplňující princip základu organizace znalostí ve vědní klasifikaci. Příkladem společenského zdůvodnění je Cutterův princip „uživatelé obecně přijímaného“ označení předmětu<sup>278</sup> nebo požadavek Henryho E. Blisse, aby knihovní klasifikace byly vytvářeny na základě „vědeckého a vzdělávacího konsenzu“, tj. aby byly v souladu s obecně přijímanou klasifikací vědění. Míra odlišnosti systému organizace znalostí od reality je samozřejmě předmětem diskuse. Praxe informačního profesionála spočívá v hledání míry mezi striktním dodržováním formálně logických pravidel a ohledem na skutečnou povahu organizovaných entit.

### b) Řešení konstantních problémů organizace znalostí – principy a techniky organizace znalostí

Klíčovou metodou organizace je analýza, jejíž efektivnost podporuje dodržování pravidel odvozených z logiky: jasné ohraničení analyzovaného celku i částí, úplnost členění, soudržnost, jednotné kritérium členění, vzájemně se vylučující části, stejná úroveň granularit i abstrakce.

## 6. Princip ekvivalence a identifikace

V rámci obecného principu jakékoli organizace, včetně organizace znalostí, lze rozpoznat dva vzájemně se doplňující procesy: seskupování na základě ekvivalence a odlišení (identifikace). I když mají odlišný účel, jejich realizace spočívá na stejných principech. Vzájemná souvislost ekvivalence a identifikace je v souladu se systémovým přístupem a je vyjádřena i v definici množiny, již zakladatel množinové teorie Georg Cantor formuloval následovně: „*Množinou rozumíme sloučení v jedno nějakých objektů, které jsou dobře rozlišitelné naší intuicí nebo naší*

---

<sup>276</sup> PLANTINGA, Alvin. *Warrant: the current debate*. New York: Oxford University Press, 1993. ix, 228 s. doi:10.1093/0195078624.001.0001. ISBN 978-0-19-507861-9 (váz.). ISBN 978-0-19-507862-6 (brož.).

<sup>277</sup> HULME, Edward Wyndham. *Principles of book classification*. Reprinted from the Library Association Record ... London: Association of Assistant Librarians, 1950. 25 s.

<sup>278</sup> CUTTER, Charles Ammi. Subjects. In: *Rules for a dictionary catalog*. 4th ed., rewritten. Washington: Government Printing Office, 1904, s. 12, 70.

myslí.<sup>279</sup> Ekvivalence slouží k zařazení do skupiny, odlišnost slouží k identifikaci a k uspořádání (třídění) v rámci dané skupiny. Stejně jako v ontologickém inženýrství, je i v organizaci znalostí nejvýznamnějším typem ekvivalence založená na stejných vlastnostech (viz kapitola 1.5.3).

Identifikování objektu spočívá v odlišení od všech ostatních objektů a dále v potvrzení, že jde skutečně o požadovaný objekt. Obojí se děje s využitím specifických atributů, obvykle nazývaných identifikátory. Identifikátor je atribut, který svou existencí odlišuje členy různých tříd ekvivalentních objektů a svou hodnotou odlišuje jednotlivé členy třídy mezi sebou. Funkce identifikace v informačních systémech mohou být ještě širší: kromě rozlišení či rozpoznání entity může identifikátor zajišťovat přístup k dalším informacím o entitě (,rozklíčování', angl. *resolution, resolving, dereference*), řízení přístupu entity k systému prostřednictvím autorizace a autentizace a tzv. vysledovatelnost (angl. *traceability*), typicky identifikaci produktu v logistickém řetězci.

Omezení identifikátorů spočívá v tom, že svou funkci plní pouze v těch oblastech použití, pro něž byly určeny. Pro tyto oblasti se používá název jmenný prostor (angl. *namespace*), což je soubor jedinečných pojmenování, který se používá pro identifikaci objektů v rámci vymezené domény.

## 7. Princip kategorizace a klasifikace

Skupinu objektů, považovaných z určitého pohledu nebo pro daný účel za ekvivalentní, nazýváme kategorie či třída. Procesy kategorizace a klasifikace v organizaci znalostí jsou založeny na stejných principech jako proces konceptualizace v ontologickém inženýrství, který je popsán v kapitole 1.5.3. Vícekriteriální kategorizaci a klasifikaci umožňuje fasetová analýza, která je rovněž vhodnou metodou pro analyticko-syntetickou organizaci složených pojmů a témat (viz kapitola 2.3).

## 8. Princip definování

Podmínkou efektivního organizování je poznání, tj. vymezení (definování) organizovaných entit. Význam organizovaných entit, tak jak byly pochopeny lidskou myslí, reprezentují pojmy. Za definici se tradičně považuje relativně úplné, závazné a postačující vystižení obsahu pojmu (lat. *definiendum*) jinými pojmy (lat. *definiens*).<sup>280</sup>

K dispozici jsou dva typy metod umožňující definování: extenzionální metoda založená na empirii a intenzionální metoda, jejímž základem je abstrakce. Extenzionálního vymezení obsahu pojmu se dosáhne výčtem (seznamem, vyjmenováním, enumerací) všech zahrnutých prvků, případně vymezením (ohraňčením) rozsahu. Výsledkem intenzionálního definování založeného na určení obsahových charakteristik jsou deskripce (popis) a ,klasická' intenzionální definice.

Deskripce je možné chápat jako pojmenování popisem. Na rozdíl od definice nemá deskripce charakter výroku, tj. nemá formu oznamovací věty.<sup>281</sup> Definice v užším slova smyslu (tj. intenzionální definice) je výrok, který popisuje vlastnosti definovaných entit (a jejich případné hodnoty) a určuje jejich kontext (vztah k ostatním prvkům). Klasický návod pro správně

---

<sup>279</sup> CANTOR, Georg. *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*. Berlin: Julius Springer, 1932, s. 282. – Překlad z němčiny cit. dle KOŘÍNEK, Vladimír. *Teorie množin, její vznik a vývoj*. In: *Pokroky matematiky, fyziky a astronomie*. 1965, 10(3), s. 132.

<sup>280</sup> NOVÁK, Lukáš, DVOŘÁK, Petr. *Úvod do logiky aristotelské tradice*. 1. vyd. České Budějovice: Teologická fakulta Jihočeské univerzity v Českých Budějovicích, 2007, s. 94. Opuscula; 5. ISBN 978-80-7040-959-6.

<sup>281</sup> Podrobněji viz CMOREJ, Pavel. *Úvod do logické syntaxe a sémantiky*. Praha: Triton, 2002. Kap. 2.3, Individuové deskripce a popisy, s. 42-44. ISBN 80-7254-294-X.

zkonstruovanou intenzionální definici zformuloval už Aristoteles: „*Definitio per genus proximum et differentia specifica*“ – definice prostřednictvím zařazení prvku do nejbližší nadřazené kategorie a určením specifických vlastností, jež definovaný prvek odlišují od dalších prvků zařazených do téže kategorie.

## 9. Princip řazení

Jakmile analýza dospěje ke kategorii obsahující více než jeden prvek, je zapotřebí vyřešit způsob uspořádání obsažených prvků, obvykle pro účely jejich uložení a prezentace. Řazení je uspořádání organizovaných entit podle zvoleného kritéria. Jeden soubor (množinu) entit lze uspořádat různými způsoby: s uplatněním různé míry granularity (podrobnosti členění) nebo podle různých kritérií členění.

Pravidla pro řazení bibliografických entit, například bibliografických záznamů v katalogích nebo v bibliografických soupisech, shrnuje mezinárodní norma ISO 7154.<sup>282</sup> Kromě tradičního abecedního nebo numerického řazení je uživatelsky nejžádanější a zároveň nejobtížněji realizovatelný způsob řazení podle relevance. Jedním z příkladů je vzorec *APUPA* Shiyali R. Ranganathana, který vychází z empirie a stanoví schéma řazení zdrojů v rámci stejného předmětu podle relevance. Zdroje jsou v rámci předmětu zájmu rozděleny podle toho, jak těsně jejich obsah s daným předmětem koresponduje. Nerelevantnější zdroje jsou označeny jako U, okrajově související jako P a písmeno A označuje zdroje, jež s předmětem zájmu čtenáře nesouvisí, resp. nejsou předmětem jeho požadavku.<sup>283</sup>

Obdobným typem členění podle relevance, který je založen na statistickém vyhodnocení počtu relevantních článků v odborných časopisech, je tzv. Bradfordovo rozdělení. Mezi počtem časopisů a mezi počtem v nich publikovaných relevantních článků je inverzní vztah – maximální počet relevantních článků je soustředěn v minimálním počtu časopisů tvořících tzv. jádro oboru.<sup>284</sup> Dalším typem řazení podle relevance je technika řazení výsledků vyhledávání zdrojů Internetu pomocí vyhledávače *Google* založená na algoritmu *PageRank*. Zatímco k Bradfordovu rozdělení vede cesta přes intelektuální určování relevance, řazení podle tohoto algoritmu vychází z výsledků statistického vyhodnocení odkazů směřujících na danou webovou stránku.

Specifický problém představuje tzv. citační pořadí při lineární prezentaci metadat, jež jsou výsledkem procesu organizace znalostí. Mezi klasické pomůcky patří formule *PMEST* Shiyali R. Ranganathana, řadící základní kategorie jeho *Dvojtečkového třídění* (viz kapitola 1.5.3) v pořadí od konkrétního k abstraktnímu, tj. ‚osobnost‘, ‚materiál‘, ‚energie‘, ‚prostor‘, ‚čas‘.

## 10. Technika umístění

Umístění dává odpověď na otázku, kde je to, co potřebujeme. Technika umístění může být tedy uplatněna pouze na materiální objekty. Vychází z fyzické podstaty organizovaných objektů, jejichž vnitřní vlastností je na rozdíl od abstraktního označení vždy jednoznačné určení konkrétním prostorem. Tato technika má teoretické základy ve fyzice. Souvisí s fyzikálními

---

<sup>282</sup> ČSN ISO 7154. *Dokumentace – Zásady bibliografického pořádkání*. Praha: Český normalizační institut, 2003. 15 s. Třídící znak 01 0141.

<sup>283</sup> RANGANATHAN, Shiyali Ramamrita. *Classification and communication*. 1st ed. Delhi: University of Delhi, 1951, s. 227-229. Dostupné též z: <http://hdl.handle.net/10150/105279> [cit. 2018-01-30]

<sup>284</sup> Poznámka: Konkrétně se postupuje tak, že se seřadí tituly časopisů sestupně podle počtu relevantních článků. Seznam se rozdělí na tři skupiny, v nichž je vždy stejný počet článků. Hodnota čísla ‚n‘ se získá vydělením počtu titulů v 1. zóně počtem titulů obsažených v jádru.

veličinami prostoru a času, které nejen podmiňují tento způsob organizace, ale určují i jeho efektivnost. Platí, že prostorové parametry korelují s časovými, tj. čím blíže je požadovaný objekt, tím kratšího času bude zapotřebí k jeho dosažení.

Umístění má tu výhodu, že pro umístěný objekt není nutné ‚vymýšlet‘ jméno. Pro tento způsob organizace nejsou zapotřebí jiná metadata než popis lokace/adresa objektu. Z toho vyplývá, že takové uspořádání není optimální pro movité objekty, protože při přemístění objektu se logicky změní i jeho adresa. Pro to, aby technika umístění fungovala, je zapotřebí, aby lokace objektu byla stabilní v relevantním časovém úseku.

Problémem umístění je fyzické omezení místem (prostorem). Z něj vyplývá, že na jednom místě může být v jednom časovém okamžiku vždy jen jedna věc a stejné věci („jedna věc“) mohou být jen na jednom místě. Zatímco tyto fyzikální limity jsou užitečné pro jednoznačnou identifikaci, mohou se stát překážkou rychlého přístupu k organizovaným objektům. Již bylo uvedeno, že umístění (poloha) ovlivňuje čas potřebný k přístupu, fyzikální omezení ovšem neumožňují umístit všechny relevantní objekty do bezprostřední blízkosti. To znamená, že k různým objektům povede v závislosti na jejich umístění různě dlouhá cesta, což pochopitelně ovlivní dobu potřebnou pro přístup k nim.

Techniky umístění lze členit podle toho, jaký proces podporují, na techniku lokace (adresování) a kolokace. Lokace je technika umožňující proces přístupu, proces odkazu na konkrétní objekt a proces odlišení, tj. identifikaci umístěním. Kolokace podporuje proces seskupování, tj. vyjádření ekvivalence a kategorizaci umístěním. Spočívá v tom, že ekvivalentní objekty jsou seskupeny na jednom místě.

## 11. Technika označení

Účelem označení je pomoci při zjištění, co je či co znamená to, co je označováno. Znalost označení umožňuje referenci neboli odkaz na označovaný objekt, tj. smysluplnou komunikaci. Funkci označení může plnit text, ale i zvuk, obraz, gesto či jiný symbol, případně jejich kombinace. Označení je abstraktní (virtuální, logické), označit lze cokoli, materiální i abstraktní objekty. Pro pojmenování procesu označení obsahu organizovaných zdrojů se v prostředí paměťových institucí tradičně používá termín indexace. Technika označení umožňuje organizaci prostřednictvím metadat. Pořádku se tedy dosahuje přidáním dalších údajů k organizovaným objektům, jež samozřejmě po svém zaznamenání zabírají v organizovaném systému další místo.

Princip označení spočívá v reprezentaci organizovaných objektů, má tedy teoretické základy v sémiotice. Výhodou je, že označený objekt může měnit své umístění, což je vhodné pro mobilní objekty. Navíc jedno označení může být současně použito pro více objektů bez ohledu na jejich umístění. Tím je umožněno i virtuální ‚seskupování‘: objekty, které mají stejné označení, patří do stejné skupiny. Na rozdíl od umístění umožňuje označení organizovat i fyzicky nedostupné věci, například již nežijící osoby či abstraktní pojmy.

Označení je vždy vnější, dodatečně připojený umělý atribut objektu, a vazba objekt – označení se tudíž může přerušit, resp. není trvalá. Označení tedy není stabilní, může se měnit v čase.

Označení je spojeno s jazykovými a sémantickými problémy. Vztah objekt – označení je ze své podstaty nejednoznačný, tj. pro jeden objekt může být více označení a zároveň jedno označení může být použito pro více různých věcí. Nevýhodou označení ve funkci identifikátoru je obtížné hledání jedinečných jmen, zejména pokud je soubor označovaných objektů rozsáhlý.

Techniky označení lze členit na dva typy – označení objektu/pojmu a zjednoznačnění. Již bylo

uvedeno, že označení objektu/pojmu je ze své podstaty nejednoznačné. To činí potíže jak při kategorizaci, tak při identifikaci, kdy se vyžaduje jednoznačnost. Technika zjednoznačnění se snaží vyřešit tento problém a stanovit taková označení, jež v daném kontextu umožní jednoznačnou referenci na označený objekt a tím ho odlišit od ostatních. Obvyklým nástrojem explicitního zjednoznačnění je formalizovaný či přímo umělý jazyk. V praxi organizace znalostí jsou nejčastěji používanými typy formalizovaného jazyka řízený slovník, notace a identifikátory. Další možností, jež ovšem už vyžaduje určitý podíl inteligence, je odvození jednoznačného významu z kolokace, tj. z kontextu daného objektu.

### **c) Dobově podmíněná paradigmatata organizace znalostí**

#### **12. Od industriální ekonomiky ke znalostní ekonomice, od informačních zdrojů ke zdrojům ekonomickým**

Agrární ekonomika využívala pro práci se znalostmi bibliografii, v éře industriální ekonomiky byl ustaven obor dokumentace. Nástup post-industriální, znalostní ekonomiky je spojen s obory informačního a znalostního managementu. Na rozdíl od industriální ekonomiky jsou klíčovými zdroji znalostní ekonomiky znalosti. S tím souvisí zvyšování podílu znalostních pracovníků v organizacích. Počet profesionálů, kteří jsou ve své každodenní činnosti závislí na přísunu informací, dramaticky stoupl. Nutnost organizovat znalosti tak proniká ze sféry vědy a výzkumu a z tradičních paměťových a fondových institucí do podnikové a obchodní sféry. To přináší jednak rozšíření záběru, jednak zcela nová témata a problémy k řešení: převod implicitních znalostí na explicitní, reprezentace znalostí (tj. převod znalostí na informace), sdílení a transfer znalostí, integrace znalostí do pracovních procesů.

#### **13. Od idejí národního státu ke globalizaci a multikulturalismu**

Pro praxi organizace znalostí představuje tento fenomén obtížně řešitelný problém. Pokud platí, že kategorizaci a klasifikaci významně ovlivňuje i specifický kulturní kontext tvůrce a uživatele kategorizace, je nasnadě, že v globálním multikulturním prostředí je třeba hledat ‚jednotu v roznořnosti‘. Obhájcem tohoto přístupu je David Weinberger, který v knize *Everything is miscellaneous* (Vřechno patří do různého)<sup>285</sup> dokládá, jak digitalizace umožňuje uskutečnit myřlenku různých pohledů a rozmanitosti pořádaní.

S ústupem idejí národního státu souvisí i změny v oblasti institucionální/státní podpory ‚národním‘ institucím a projektům. Nyní se pro takové projekty hledají nové způsoby zdůvodnění existence a zdroje nákladů na jejich fungování. Symbolicky tento posun od národních entit ke globalizaci vyjadřuje název programu *UNESCO Paměť světa*, vyhlášeného v roce 1992, či název projektu digitální kolekce *Europeana*, spolufinancovaného Evropskou unií.

#### **14. Od lineárních komunikačních a datových modelů k modelům síťovým**

Příkladem tradičního modelu vědecké a technické komunikace je komunikační model *UNISIST*<sup>286</sup>, vytvořený v 70. letech 20. století. Jeho lineární charakter, založený na Shannon-

---

<sup>285</sup> WEINBERGER, David. *Everything is miscellaneous: the power of the new digital disorder*. 1st ed. New York: Times Books, 2007. 277 s. ISBN 978-0-8050-8043-8.

<sup>286</sup> *UNISIST: Study report on the feasibility of a World Science Information System*. Paris: Unesco, 1971. 161 s. Dostupné z: <http://unesdoc.unesco.org/images/0006/000648/064862eo.pdf>. – FJØRDBACK SØNDERGAARD, T., ANDERSEN, J., HJØRLAND, B. Documents and the communication of scientific and scholarly information. Revising and updating the UNISIST model. In: *Journal of documentation*. 2003, 59(3), 278-320.



Weaverově transmisivním modelu (zdroj – kodér – sdělení/kanál – dekodér – příjemce)<sup>287</sup>, však byl kritizován už v době svého vzniku. Hlavní nedostatky lineárního modelu spočívají v tom, že vyhovuje spíše pojetí komunikace jako přenosu než jako spojení (tj. komunikace ve smyslu sdílení zdrojů, komunikace jako společná činnost), navíc vyhovuje spíše pro znázornění přenosu v prostoru než v čase, nebere v úvahu kontext a prakticky neumožňuje znázornit interakci. Komunikace současnosti je podle Dana Brickleyho oproti tomu založena na třech typech informačních sítí: hypertextový graf s propojenými dokumenty, sociální graf propojující lidi a faktový (angl. *factual*) graf s propojenými daty.<sup>288</sup> K této trojici se přidává rychle se rozrůstající síť propojených komunikujících věcí (tzv. *Internet věcí*). Infrastrukturu pro všechny typy sítí tvoří technologie *Internetu*. Z toho důvodu bude zapotřebí konstrukce nového, pravděpodobně síťového modelu komunikace, který bude lépe odpovídat skutečné povaze sociální komunikace a bude i v souladu s povahou dnešního komunikačního prostředí.

Specifika síťové komunikace tvoří především její dynamika – jeden zdroj může vytvořit více sdělení, jedno sdělení může být vyprodukováno více zdroji, jedno sdělení může být určeno více příjemcům, jeden příjemce může obdržet více sdělení. Síť ve větší míře než jiné typy komunikace podporují interakci – příjemce se snadno může změnit ve zdroj a naopak. K tomu se přidává dynamika v čase: v síti mohou přibývat nové uzly/vazby a existující uzly/vazby mohou mizet. Ty, které zůstávají, mohou měnit svůj obsah. Z toho vyplývá, že síťový model komunikace je obtížněji organizovatelný než model lineární. Případné snahy o jeho řízení a správu budou vyžadovat zcela nové přístupy.

Změny, které přinesl nástup elektronické komunikace, se týkají i samotných komunikovaných zdrojů. Zatímco tradiční analogové dokumenty prakticky nebylo možné organizovat jinak než lineárně, v sekvenční posloupnosti, pro digitální dokumenty a záznamy je možné zvolit i komplikovanější způsob organizace, který lépe odpovídá realitě, například hierarchické nebo síťové struktury. Tato možnost se týká jak organizace dokumentů v rámci nějakého souboru, tak vnitřní organizace obsahových částí dokumentů (příkladem jsou hypertextové či relační databázové struktury).

## 15. Od analogových dokumentů k digitálním

V současné době se živě diskutuje o problematice elektronických dokumentů, u nichž jsou pozorovány některé specifické vlastnosti, odlišující je od ‚tradičních‘ dokumentů. Jiří Cejpek dokonce tvrdí, že „*digitální či digitalizovaný dokument přestává být dokumentem s vlastnostmi a významem, který mu na konci minulého století přisoudili belgičtí právníci P. Otlet a H. Lafontaine*“.<sup>289</sup> Specifika digitálních informačních zdrojů byla popsána v kapitole 1.5.1, v této části doplníme ty vlastnosti, jež mají vliv na proces organizace znalostí.

Bity tvoří jednotnou platformu pro všechny formáty dokumentů – text, audio, grafiku, video i softwarové aplikace, a sjednocují tak postupy tvorby, zpracování i komunikace složených

---

<sup>287</sup> SHANNON, Claude E. A mathematical theory of communication. In: *The Bell System technical journal*. 1948, **27**(3, July), 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x; **27**(4, October), 623–656. doi:10.1002/j.1538-7305.1948.tb00917.x. ISSN 0005-8580.

<sup>288</sup> BRICKLEY, Dan. *Classification, collaboration and the Web of data*. Přednáška na semináři Classification and ontology: formal approaches and access to knowledge: proceedings of the International UDC Seminar, 19-20 September 2011, The Hague, The Netherlands. Prezentace a audiozáznam přednášky dostupné z: <http://seminar.udcc.org/2011/programme.htm> [cit. 2018-01-30].

<sup>289</sup> CEJPEK, Jiří. Knihovnictví v širších souvislostech. In: *Národní knihovna*. 1999, **10**(6), 274. ISSN 0862-7487 (Print). ISSN 1214-0678 (Online).

dokumentů a heterogenních kolekcí (příkladem je digitální knihovna jako komplex digitálních objektů, databází s uloženými metadaty, řízených slovníků a vyhledávacího softwaru). Zatímco ještě v druhé polovině 20. století platilo, že v počítači (tj. v digitálním formátu) jsou zpravidla pouze metadata o primárních zdrojích, které jsou k dispozici v analogovém formátu a tudíž mimo počítač, modelem pro 21. století jsou sekundární i primární informace v jednotné digitální formě. Digitalizují se nejen zdroje, ale i procesy. K e-byznysu a e-learningu přibyl fenomén e-science – proces tvorby nového vědeckého poznání v počítači. Jsou k dispozici softwarové nástroje usnadňující tvorbu, editaci a publikování digitálních dokumentů, jež přinesly i řešení některých problémů organizace. Dokumenty sestavené z bitů jsou přímo zpracovatelné počítačem. To ovšem neznamená, že jsou počítači srozumitelné a mohou být předmětem automatické obsahové analýzy. Zpravidla se jedná ‚pouze‘ o významné zrychlení některých mechanických operací, typicky řazení a reprodukce.

Navzdory technologickému potenciálu digitálních dokumentů, jenž umožňuje jejich prakticky neomezenou trvanlivost, je prokázáno, že v praxi je tzv. ‚poločas rozpadu‘ zejména dokumentů online dostupných na webu mnohem kratší než u analogových dokumentů.<sup>290</sup> Zřetel na dostupnost se tak doplňuje o nutnost zajistit dlouhodobou archivaci digitálních dokumentů a dat. Ke krátké životnosti se navíc přidává vysoká frekvence změn obsahu, vyvolaná zřejmě snadnou editovatelností elektronických dokumentů. To vyvolává potřebu správy verzí. Mnohem rychleji se také mění formáty zdrojů a parametry zařízení, jež jsou schopná je zpracovávat. Všechny tyto faktory jsou důvodem toho, že současný informační prostor je mnohem dynamičtější, a tím i obtížněji organizovatelný, než byla dosavadní ‚analogová univerza‘ zaznamenaného poznání v kterékoli z vývojových etap minulosti.

## **16. Od monolitických statických dokumentů k propojeným datům, od makro-organizace k mikro-organizaci, od dokumentů k informacím**

V případě elektronických dokumentů se uvolňují nejen původní pevné vazby obsahu tradičních dokumentů s fyzickým nosičem, ale stále častěji i s aplikačním softwarem a uživatelským rozhraním. Princip sítí propojených dat (angl. *linked data*) spočívá v tom, že do repozitářů se už neukládají potenciálně relevantní statické dokumenty a záznamy jako celek, ale před svým uložením jsou dekomponovány na jednotlivé obsahové části. Uživatel má pak možnost nechat si na základě svého individuálního požadavku vygenerovat ad hoc dynamický dokument ‚poskládaný‘ z jednotlivých věcně relevantních distribuovaných faktů, představujících obsah. Náznornou představu o současném rozsahu a obsahu sítí propojených otevřených dat podává *Linked open data cloud diagram*, který zpracovali Richard Cyganiak a Anja Jentzschová (dostupný z <http://lod-cloud.net/>).

Tyto změny informačního prostředí vedou k tomu, že dosavadní dokumentově orientované postupy organizace znalostí jsou nahrazovány přístupy zaměřenými na jemnější granularitu zpracovávaného obsahu, a to až do úrovně jednotlivých faktů. K jemnější granularitě směřovala už technologie plnotextového vyhledávání, umožňující přístup až do úrovně jednotlivého slova v textu. Zatímco ovšem plnotextové vyhledávání dosáhlo potřebné jemnosti granularity pouze na syntaktické úrovni, databázové technologie a zejména technologie sémantického webu se snaží o totéž na úrovni sémantické. Současně vzniká požadavek na řešení problémů vyplývajících z distribuovanosti obsahu, především ztráty kontextu a (centrální) kontroly nad kvalitou

---

<sup>290</sup> Viz např. longitudinální výzkum Wallace Koehlera: OGUZ, Fatih, KOEHLER, Wallace. URL decay at year 20: a research note. In: *Journal of the American Society for Information Science and Technology*. 2015, 66(2), 477-479. doi:10.1002/asi.23561. ISSN 1532-2882 (Print). ISSN 1532-2890 (Online).

obsahu. Thomas Baker navrhuje, aby se ‚rozbočovači‘ (převodními kolečky) v takové síti staly pojmy specifikované v systémech organizace znalostí, identifikované prostřednictvím *URI*.<sup>291</sup>

#### **d) Řešení aktuálních problémů organizace znalostí**

##### **17. Od koncentrace na uchovávání a ochranu dokumentů k zajištění dostupnosti dokumentů/znalostí**

Jestliže je obecným cílem organizace znalostí přístup ke znalosti/informaci, musíme jako jednu z paradigmatických změn vnímat otevřený přístup (angl. *open access*), software s otevřeným kódem (*open source*), otevřená data, veřejné licence (*creative commons*) a další novinky v principech přístupu ke znalostem. Příkladem aktivit v tomto směru jsou iniciativy nadace *Open Knowledge Foundation* (<https://okfn.org/>), založené v roce 2004. Podle její definice „*znalosti jsou otevřené, pokud k nim může kdokoliv volně přistupovat, používat je, upravovat a sdílet je – za podmíněk, které nejvíce zabezpečí zachování původu, tedy vlastnictví, a otevřenosti.*“<sup>292</sup> Procesy zajištění dostupnosti a zajištění ochrany se navzájem nevyklučují. Archivace a ochrana je ale v novém paradigmatu chápána jako podmínka dostupnosti, nikoli jako cíl organizace sám o sobě.

##### **18. Od pasivního popisu k aktivnímu návrhu nových systémů**

Tento trend zaznamenává např. Ryan Shaw, který upozorňuje na změnu funkčnosti tradičních nástrojů organizace znalostí v novém kontextu:

*„Nástroje jako tezaury a klasifikační schémata se používaly především pro účely indexování a vyhledávání. Ale v éře ohromných digitálních archivů a výkonných vyhledávacích algoritmů je klíčovým problémem organizace informací konstrukce systémů, jež podporují porozumění, kontextualizaci a sebeorientaci v mase zdrojů. Budování takových systémů se méně podobá organizaci informací v tradičním chápání, a více vytváření stručných shrnutí nebo produkci výstav.“*<sup>293</sup>

Patrick Lambe vyzývá k aktivnější roli samotných informačních profesionálů a spojuje tuto otázku s etikou organizace znalostí. Současné prostředí organizace znalostí určují podle Lambeho dvě dimenze s krajními póly: pořádek – nepořádek a vlastnictví (*sequestration*) – sdílení (*commons*). Místo informačního profesionála vidí nikoli na některém z krajních bodů, ale v ohnisku jejich vzájemného působení. Jeho úlohou pak není zachování nebo popis rozpoznaného statu quo, ale návrh nových řešení.<sup>294</sup>

##### **19. Od nedostatku informací k řešení problému informačního zahlcení a důvěryhodnosti**

Kvantitativní nárůst potenciálně dostupných informací je tak veliký, že tato kvantita přechází v novou kvalitu: informační zahlcení je už vnímáno jako větší problém než nedostatek informací.

---

<sup>291</sup> BAKER, Thomas. The concepts of knowledge organization systems as hubs in the web of data. In: *Classification and ontology: formal approaches and access to knowledge: proceedings of the International UDC Seminar*, 19-20 September 2011, The Hague, The Netherlands. Prezentace a audiozáznam přednášky dostupné z: <http://seminar.udcc.org/2011/programme.htm> [cit. 2018-01-30]

<sup>292</sup> *Open definition* [online]. Open Knowledge International. Dostupné z: <http://opendefinition.org/od/2.0/cz/> [cit. 2016-11-07].

<sup>293</sup> SHAW, Ryan. Information organization and the philosophy of history. In: *Journal of the American Society for Information Science and Technology*. June 2013, **64**(6), s. 1100. doi:10.1002/asi.22843. ISSN 1532-2882 (Print). ISSN 1532-2890 (Online).

<sup>294</sup> LAMBE, Patrick. From cataloguers to designers: Paul Otlet, social impact and a more proactive role for knowledge organization professionals. In: *Knowledge organization*. 2015, **42**(6), 445-455. ISSN 0943-7444.

Kvanta údajů generovaná denně počítačovými aplikacemi už nesou přímo název ‚velká data‘. Příčinou této exploze v dostupnosti informací se staly počítačové aplikace pro zpracování a digitalizaci dokumentů a počítačové sítě, které revolučním způsobem usnadnily vytváření a publikování obsahu. Zatímco však informační a komunikační technologie účinně zasáhly do úvodní publikační fáze komunikačního procesu, dosud nejsou k dispozici obdobné nástroje, jež by stejně výrazně usnadnily zpracování a organizaci znalostí. Jsou zapotřebí jednak nástroje pro zredukování (filtrování) dostupných informací jen na ty relevantní, jednak (na první pohled možná poněkud paradoxně) nástroje pro tvorbu přidáných informací<sup>295</sup> v podobě metadat. Organizace tohoto typu se stále dělá ‚postaru‘, ‚manuálně‘, což v tomto případě znamená za využití lidského intelektu, a to je drahá a vzácná komodita. Očekávaným řešením problému by se měly stát softwarové aplikace, které budou buď podporovat a v některých případech i nahrazovat intelektuální aktivity při organizaci znalostí, nebo budou schopné prohledávat neorganizovaný obsah. V obou případech bude zapotřebí podpořit sémantické aktivity softwaru ontologiemi a systémy organizace znalostí.

Problém důvěryhodnosti je vyvolán rostoucím počtem dostupných zdrojů, u nichž nejsou k dispozici tradiční indikátory důvěryhodnosti, například renomovaný vydavatel či časopisecký titul. Dnešní uživatel tak navíc k problémům získání informace a rozhodnutí o její relevanci stále častěji řeší problém, zda může důvěřovat jejímu obsahu. Názorným příkladem, o němž se aktuálně diskutuje ve sféře vědeckého publikování, jsou tzv. predátorské zdroje, jež zneužívají principy otevřeného přístupu. K tomu se přidává relativní snadnost elektronické komunikace pod falešnou identitou. Sféra zájmu organizace znalostí se tak rozšiřuje: k původnímu schématu ‚popis + obsahová analýza + kontext‘ přibyla další položka – ‚posouzení důvěryhodnosti‘.<sup>296</sup> Jean L. Graefová tuto záležitost vztahuje i na systémy organizace znalostí: „*Kromě plnění základních tří funkcí [tj. identifikace, objevování, získání] by taxonomie také měla podněcovat důvěru. Uživatel by měl cítit jistotu, že mu taxonomie pomůže najít informaci, již hledá – pokud existuje.*“<sup>297</sup>

## **20. Od technicistního a kulturně historizujícího přístupu k vědeckému přístupu, od praxe k teorii**

Pronikání počítačových technologií do praxe organizace znalostí v 2. polovině 20. století celkem pochopitelně vyvolalo soustředění aktivit odborníků na praktické zvládnutí jejich implementace do stávajících procesů. V některých případech dokonce možnosti technologie vyvolaly představy, že technika sama vyřeší problémy organizace. Nehledě na to, že takové názory měly své oponenty už v době svého vzniku (příkladem je polemika Birgera Hjørlanda s Karen Spärck Jonesovou o možnostech plnotextového vyhledávání a automatické indexace<sup>298</sup>), v současné

---

<sup>295</sup> „...řešením nadbytku informací je více informací“. WEINBERGER, David. *Everything is miscellaneous: the power of the new digital disorder*. New York: Times Books, 2007, s. 13.

<sup>296</sup> Viz například navrhované postupy a metody k odhalení predátorských zdrojů v: PLCH, Lukáš, KRATOCHVÍL, Jiří. *Predátorské časopisy* [online]. 1. vyd. Brno: Knihovna univerzitního kampusu Masarykovy univerzity, 2016, aktualizace 2017-02-21 [cit. 2018-01-30]. Dostupné z: <https://kuk.muni.cz/vyuka/materialy/predatori/>.

<sup>297</sup> GRAEF, Jean L. Managing taxonomies strategically [online]. In: *Montague Institute Review*. March 2001, s. 2. [cit. 2018-01-30]. ISSN 1554-303X. Dostupné z: <http://www.montague.com/review/articles/taxonomy3.pdf>.

<sup>298</sup> HJØRLAND, Birger, PEDERSEN, Karsten Nissen. A substantive theory of classification for information retrieval. In: *Journal of documentation*. September 2005, 61(5), 582-597. doi:10.1108/00220410510625804. ISSN 0022-0418. – SPÄRCK JONES, Karen. Revisiting classification for

době se opět pozornost vědců přesouvá na intelektuální aktivity organizace znalostí.

Armando Silva a Fernanda Ribeirová ve své analýze paradigmat informační vědy, již lze přenést i na oblast organizace znalostí, dále konstatují, že stávající metody informační vědy se často omezují na historický přístup, kdy porozumění přítomnosti je odvozováno z pochopení proměn kulturních vzorců minulosti. Ten je podle nich v současnosti zapotřebí nahradit komplexním vědeckým přístupem.<sup>299</sup> Rovněž Birger Hjørland vyjadřuje přesvědčení, že obor organizace znalostí je zapotřebí opřít o pevný vědecký základ, a navrhuje, aby se takovým základem stala teorie poznání. Vědecký přístup ovšem bezpodmínečně vyžaduje, aby byl jasně vymezen předmět zkoumání a aby byly k dispozici teoretické principy a metody zkoumání, specifické pro daný obor. Richard P. Smiraglia se domnívá, že takovéto pevné základy zatím obor organizace znalostí nemá, nicméně konstatuje, že v tomto směru již bylo dosaženo velkého pokroku.<sup>300</sup>

## 21. Rozšíření okruhu osob i institucí, jež se věnují organizaci znalostí

Aktuální spektrum institucionálních aktérů organizace znalostí a jejich měnící se vztahy popisuje kapitola 4.5.1. Proměna paradigmatu tvorby znalostí (poznání) způsobila i změnu paradigmatu v zodpovědnosti za jejich organizaci. Na pořadu dne jsou změny spočívající v převzetí některých aktivit samotnými autory (tak jako již autoři začali přebírat publikační aktivity) nebo uživateli (princip *Webu 2.0* – uživatel je v roli tvůrce a organizátora obsahu, objevují se nové fenomény tagování, folksonomie, „naivní“ klasifikace). Další významnou skupinou, která se začíná integrovat s dosud oddělenými komunitami autorů, institucionálních producentů a zpracovatelů informací, se tak stávají uživatelé a čtenáři.

Laičtí uživatelé začali do oblasti donedávna vyhrazené profesionálům pronikat s nárůstem počtu zdrojů informací dostupných online. Model druhé poloviny 20. století, kdy pro ně vyhledávání v odborných elektronických zdrojích zajišťovali profesionální rešeršéři, nahradil všudypřítomný *Internet*, v jehož prostředí si každý může hledat informace sám. Vpád laiků do oblasti donedávna vyhrazené specialistům způsobil dramatickou změnu požadavků na způsob vyhledávání v informačních zdrojích – uživatelé vyžadují snadný přístup a simultánní prohledávání všech typů zdrojů z jednoho místa, tak jak si na něj navykli při vyhledávání v *Internetu* prostřednictvím *Google*. Softwarové aplikace a počítačové sítě jim již usnadnily publikování dokumentů, nyní od nich totéž očekávají v přístupu k informacím. Náléhavě se projevuje potřeba řešit problém tradičních systémů organizace znalostí (klasifikací a tezurů), jež byly původně určeny úzké skupině specializovaných profesionálů a jejichž použití vyžadovalo intenzivní školení a praxi. Laikové evidentně dávají přednost rychlé odezvě a intuitivnímu ovládnutí systému před kvalitou podmíněnou tréninkem. Uživatelská přívětivost se stává faktorem, který v praxi dostává přednost před „objektivními“ kvalitami systému. V oblasti webového designu tento imperativ vystihl Steve Krug v titulu své knihy: „*Nenuťte uživatele přemýšlet*“.<sup>301</sup>

---

retrieval. In: *Journal of documentation*. September 2005, **61**(5), 598-601.  
doi:10.1108/00220410510625813. ISSN 0022-0418.

<sup>299</sup> SILVA, Armando Malheiro da, RIBEIRO, Fernanda. Documentation / information and their paradigms: characterization and importance in research, education, and professional practice. In: *Knowledge organization*. 2012, **39**(2), 111-124. ISSN 0943-7444.

<sup>300</sup> SMIRAGLIA, Richard P. *The elements of knowledge organization*. Springer, 2014, s. 2. ISBN 978-3-319-09356-7. ISBN 978-3-319-09357-4 (eBook). doi:10.1007/3-319-09357-4.

<sup>301</sup> KRUG, Steve. *Web design – nenuťte uživatele přemýšlet! 2.*, aktualiz. vyd. Brno: Computer Press, 2006. 167 s. ISBN 80-251-1291-8.

## 22. Od izolovaných specifických aktivit organizace znalostí k integrovanému přístupu

Dosavadní polarita ‚identifikační / věcné pořádkání‘ a ‚předmětové / systematické selekční jazyky‘, ustavená např. v českém prostředí Blahoslavem Kovářem (podrobněji viz část 4.6.2), se v kontextu měnícího se společenského i technologického paradigmatu jeví jako stále méně vyhovující. Jsou patrné snahy uplatnit na tyto fenomény integrovaný přístup. Ukázkou pokusu o překonání dosavadní duality ‚předmětové / systematické selekční jazyky‘ je představa plynulé škály přechodu mezi jednotlivými systémy organizace znalostí Gail Hodgeové<sup>302</sup> nebo Heather Heddenové<sup>303</sup> (blíže viz kapitola 4.5.3).

## 23. Od hledání jednoho ‚nejlepšího‘ systému k propojení systémů stávajících

Richard P. Smiraglia konstatuje, že nastal obrat od marného hledání univerzální klasifikace k vzájemnému porozumění a interoperabilitě.<sup>304</sup> Tomuto požadavku odpovídá současný trend propojování dat a interoperability heterogenních systémů. Požadované sjednocení způsobu přístupu k heterogenním zdrojům má zároveň zachovat jejich samostatnost a umožnit jim další nezávislý vývoj podporující jejich specifika. Řečeno slovy Susan Feldmanové, současné vyhledávací technologie by měly „indexovat všechno a přitom to ponechat na svém místě, uspokojit potřebu získat vše z jednoho místa, a přitom umožnit expertům, aby pokračovali v používání svých specializovaných nástrojů“.<sup>305</sup>

## 24. Od specializovaných vědních disciplín k transdisciplinaritě

V literatuře se o tomto jevu píše jako o multidisciplinarity, interdisciplinarity nebo transdisciplinarity. Typickým znakem je vždy předmět zkoumání přesahující hranice kterékoli vědní disciplíny a používání teoretických principů a metod z více oborů. Multidisciplinarita zpravidla označuje používání prostředků více vědních disciplín, jež si zachovávají svou samostatnost. Interdisciplinarita se navíc vyznačuje integrací použitých nástrojů, přičemž vzniká nová kvalita. Termín transdisciplinarita se používá k označení situace, kdy se propojují nástroje a principy nejen vědních disciplín, ale i všech relevantních profesionálních i neprofesionálních aktivit.<sup>306</sup> Pro organizaci znalostí jsou významné dva aspekty transdisciplinarity: za prvé je to transdisciplinární povaha samotného oboru, a za druhé se očekává, že procesy a systémy organizace znalostí budou transdisciplinarity ve všech oblastech poznání podporovat.

Rizikem transdisciplinarity je potenciální eklekticismus a pojmová nekonzistence v takové doméně. Transdisciplinarita, již je třeba považovat nikoli za znak přechodného počátečního stadia vývoje, ale za klíčový atribut organizace znalostí, komplikuje extenzionální vymezení

---

<sup>302</sup> HODGE, Gail. *Systems of knowledge organization for digital libraries: beyond traditional authority files*. Washington: The Digital Library Federation, Council on Library and Information Resources, 2000. vi, 37 s. ISBN 978-1-933645-06-3. Dostupné též z: <http://www.clir.org/pubs/reports/pub91/pub91.pdf> [cit. 2018-01-30].

<sup>303</sup> „...*The range of knowledge organization systems should be thought of as a spectrum*“. HEDDEN, Heather. *The accidental taxonomist*. Medford: Information Today, 2010, s. 1.

<sup>304</sup> SMIRAGLIA, Richard P. *The elements of knowledge organization*. Springer, 2014, s. 30. ISBN 978-3-319-09356-7. ISBN 978-3-319-09357-4 (eBook). doi:10.1007/3-319-09357-4.

<sup>305</sup> FELDMAN, Susan. The answer machine: are we there yet? In: *Searcher*. January/February 2011, **19**(1), 18-27. ISSN 1070-4795.

<sup>306</sup> SZOSTAK, Rick, GNOLI, Claudio, LÓPEZ-HUERTAS, María. *Interdisciplinary knowledge organization*. Springer, 2016, s. 7-8. ISBN 978-3-319-30147-1. ISBN 978-3-319-30148-8 (eBook). doi:10.1007/978-3-319-30148-8.

hranic domény. Navíc kupříkladu Susan Herbstová v článku věnovaném výzkumu sociální komunikace dokonce mluví nikoli o transdisciplinaritě, ale o postdisciplinaritě: tedy o situaci, kdy nemá smysl určovat mezioborové vztahy, protože nelze jasně vymezit hranice jednotlivých disciplín.<sup>307</sup> Pokud by měla tato hypotéza o mizení hranic oborů platit, znamenalo by to konec klasifikace věd. Dlouho zastávané pravidlo, že knihovnický a bibliografický systém organizace znalostí by měl být založen na klasifikaci vědních disciplín, by tak přestalo být aplikovatelné. Přes tyto problémy je zřejmé, že interdisciplinarita a transdisciplinarita jsou typickými projevy současného vědeckého výzkumu a organizace znalostí je musí být schopna podporovat.

### 4.3 Jednotka organizace znalostí

Pro úspěšnou realizaci procesu organizace je nezbytné mít přesně definovanou základní jednotku organizace. V případě organizace znalostí se lze setkat s různými označeními takové jednotky. June Abbasová mluví o znalostních objektech (angl. *knowledge object*)<sup>308</sup>, Ingetraut Dahlbergová rází pojmenování jednotka znalosti (něm. *Wissenseinheit*, angl. *knowledge unit*).<sup>309</sup> Helena Brožová a Milan Houška používají rovněž termíny znalostní jednotka/jednotka znalosti a kromě toho i termín znalostní kontejner, což je podle nich „*jakýkoli hmotný nebo nehmotný subjekt včetně lidského vědomí, který obsahuje znalost*“.<sup>310</sup>

Ze slovní podoby termínu organizace znalostí se zdá být na první pohled zřejmé, že materiálem, který se organizuje, jsou znalosti. Ve skutečnosti je termín znalost používán spíše jako metafora, navíc mnohoznačná tím, že může zastupovat různá pojetí a přístupy (viz poznámky k podstatě znalosti a informace v kapitole 1.5.1). Při vymezování základní jednotky organizace tak máme na výběr z velmi širokého spektra sahajícího od ‚čisté znalosti‘ až po fyzické předměty (např. knihy, CD disky). Robert J. Glushko nabízí rozšířený pohled na organizované zdroje prostřednictvím pojmů: fyzická věc – informace o fyzické věci – digitální věc – informace o digitální věci.<sup>311</sup> V rámci bibliografického univerza různorodost jednotek organizace reflektuje model *IFLA LRM* prostřednictvím čtyřúrovňové hierarchie entit Dílo – Vyjádření – Provedení – Jednotka, doplněné o entity Agent, Místo, Časové rozpětí a Jméno. Birger Hjørland se pokusil o svým způsobem všezahrnující vymezení jednotky organizace a tvrdí, že „*organizace znalostí je o pojmech a jejich sémantických vztazích (a současně i o reálném světě)*“.<sup>312</sup> Hjørland tedy konstatuje klíčový význam pojmů v organizaci znalostí, poukazuje ovšem i na jejich sepětí s reálnými objekty.

Vlastní analýzu směřující k určení jednotky organizace založíme na analýze univerza poznání. Na obrázku 15 na následující straně jsou schematicky znázorněny možnosti členění univerza

<sup>307</sup> HERBST, Susan. Disciplines, intersections, and the future of communication research. In: *Journal of communication*. December 2008, **58**(4), 603-614. doi:10.1111/j.1460-2466.2008.00402.x. ISSN 0021-9916.

<sup>308</sup> ABBAS, June. *Structures for organizing knowledge: exploring taxonomies, ontologies, and other schemas*. New York: Neal-Schuman Publishers, © 2010, s. 206. ISBN 978-1-55570-699-9.

<sup>309</sup> DAHLBERG, Ingetraut. How to improve ISKO's standing: ten desiderata for knowledge organization. In: *Knowledge organization*. 2011, **38**(1), s. 69. ISSN 0943-7444.

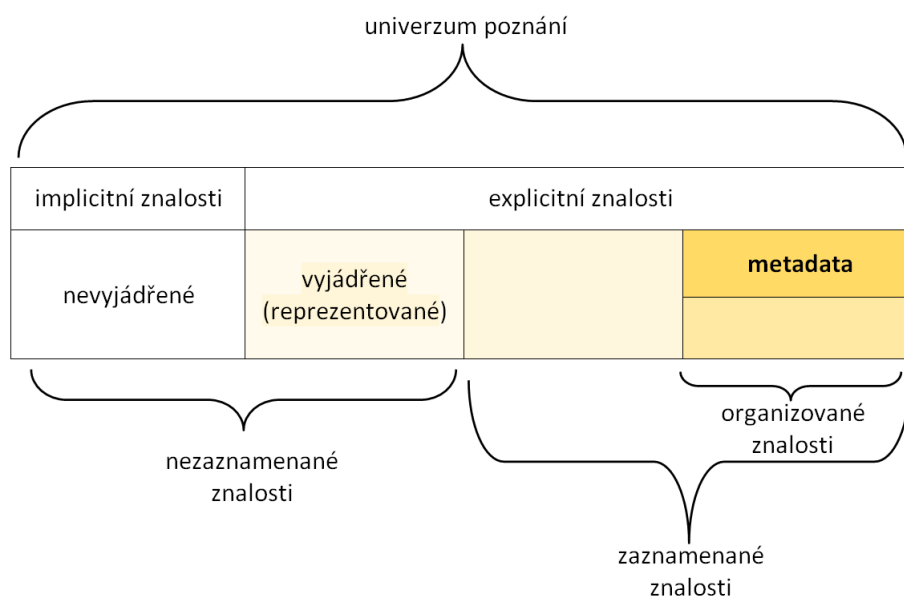
<sup>310</sup> BROŽOVÁ, Helena, HOUŠKA, Milan. *Modelování znalostí*. 1. vyd. Praha: Professional Publishing, 2011, s. 69. ISBN 978-80-7431-069-0.

<sup>311</sup> GLUSHKO, Robert J., ed. *The discipline of organizing*. Cambridge(MA): MIT Press, 2013, s. 2-3. ISBN 978-0-262-51850-5.

<sup>312</sup> HJØRLAND, Birger. Theories of knowledge organization – theories of knowledge. In: *Knowledge organization*. 2013, **40**(3), s. 169. ISSN 0943-7444.

lidského poznání jednak na znalosti implicitní a explicitní a jednak na znalosti zaznamenané a nezaznamenané. Z obrázku je patrné, že množiny zaznamenaných a explicitních znalostí, jakož i množiny znalostí nezaznamenaných a implicitních, se výrazně překrývají, nejsou však zcela totožné. Rozdíl mezi nimi působí množina reprezentovaných a nezaznamenaných znalostí.

Panuje obecná shoda v tom, že předmětem teorie i praxe organizace znalostí nejsou subjektivní implicitní znalosti, které jsou předmětem kognitivních procesů, ale objektivní explicitní znalosti. Pro účely naší analýzy tedy vyloučíme z předmětu zájmu organizace znalostí organizaci implicitních znalostí, čímž vyloučíme i určitou skupinu jednotek organizace, specifickou pro organizaci znalostí v lidské mysli. Tím zúžíme pojetí základní jednotky organizace znalostí na artefakty (lidské výtvořiny), vytvořené za účelem reprezentace znalostí a ukládané v externích pamětech. I v tomto případě se ovšem mezi odborníky vede diskuse o tom, co za takové artefakty považovat. Ingetraut Dahlbergová se přiklání k tomu, že artefaktem reprezentujícím znalost jsou už pojmy, reprezentující znalosti na abstraktní úrovni.<sup>313</sup> Další autoři se orientují na znaková vyjádření pojmů například v podobě textů, nebo až na fyzické objekty vzniklé ‚zhmotněním‘ znalostí zakódovaných do nějakého jazyka a zaznamenaných na hmotném nosiči, tj. dokumenty.



Obr. 15 Typologie znalostí z hlediska jejich komunikace a organizace

Je možné konstatovat, že zaznamenané znalosti jsou podmnožinou explicitních znalostí. V prostředí knihoven se tato podmnožina označuje termínem bibliografické univerzum. Podmnožinou zaznamenaných znalostí jsou znalosti organizované, tj. ty, jež jsou záměrně strukturované. Produktem organizace a zároveň specifickou podmnožinou organizovaných znalostí jsou metadata.

Již několikrát zmíněný termín metadata navzdory svému znění, jež navozuje výklad ve smyslu ‚data o datech‘, případně ‚informace o informacích‘, lze považovat za označení specifického typu informačního zdroje. Termín metadata se začal používat s nástupem digitálních zdrojů. V éře

<sup>313</sup> DAHLBERG, Ingetraut. *Wissensorganisation: Entwicklung, Aufgabe, Anwendung, Zukunft*. Würzburg: Ergon Verlag, 2014, s. 37. Textbooks for knowledge organization, vol. 3. ISSN 0944-8152. ISBN 978-3-95650-065-7.



analogových zdrojů se pro takový typ dokumentů používalo označení sekundární dokument. Tedy i samotná metadata, generovaná v procesu organizace a představující jeden z výstupů tohoto procesu, jsou zdroje, které je rovněž třeba organizovat na mikro- i makroúrovni. Specifikem metadat není jejich forma, ale obsah, který se vždy vztahuje k nějakému informačnímu zdroji. Uvažujeme-li dvě základní techniky organizace, popsané v kapitole 4.2 jako umístění a označení, lze metadata považovat za specifický případ označení (přesně řečeno označení informačního zdroje). Metadata mohou být fyzickou součástí zdroje (např. tiráž, ex libris, metadata v *HTML* dokumentu), nebo na něj odkazovat (např. bibliografická citace, katalogizační záznam, deskriptory, třídník *MDT*). Typickým účelem metadat je náhrada (surogát) zdroje. Často slouží jako odkaz na zdroj, tj. jako přístupový bod, mohou ovšem obsahovat i informace, doplňující obsah zdroje. Podle toho, jakého atributu nebo jaké funkce zdroje se metadata týkají, se rozlišují metadata obsahová, identifikační, administrativní a další. Nejvýznamnější z hlediska organizace znalostí jsou metadata umožňující přístup ke zdrojům. V souladu s tím, jaký princip organizace zdrojů byl zvolen, obsahují přístupová metadata buď adresu (lokaci) nebo označení zdroje. Stejně jako ve světě fyzických objektů je tedy nástrojem fyzického přístupu k organizovaným zdrojům mapa jejich lokací v podobě indexu, katalogu či rejstříku. Analogickou ‚mapou‘ označení, resp. mapou sítě jejich vzájemných vztahů, jsou systémy organizace znalostí (viz kapitola 4.5.3).

Výsledkem naší analýzy je následující pracovní vymezení jednotky organizace znalostí: Jednotkou organizace znalostí jsou zaznamenané znalosti. Při určování atributů pojmu zaznamenané znalosti dojdeme k charakteristikám velmi podobným těm, jimiž lze charakterizovat dokument. I zaznamenaná znalost je hmotně fixovaný zdroj informací, tvořící samostatnou jednotku z hlediska obsahu. Podmínkou zaznamenaní znalosti je určitě existence nějaké jednotky obsahu, již lze zaznamenat, a dále existence hmotného nosiče, na/do nějž je záznam proveden.

Tento způsob vymezení jednotky organizace, založený na filozofickém základě, umožňuje vyhnout se extenzionálnímu vymezování výčtem a jeho intenzionální povaha poskytuje platformu, jež je dostatečně široká pro zachycení všech současných forem zaznamenaných znalostí. Pojem zaznamenaná znalost může podle našeho názoru sloužit jako zastřešující pojem pro všechny tradičně uvažované a praxí prověřené entity: pro informační zdroje, dokumenty i tzv. bibliografické entity ve smyslu první skupiny entit modelu *IFLA LRM* – díla, vyjádření, provedení, jednotky. Je použitelný jak pro organizaci znalostí na makroúrovni, tak pro organizaci na mikroúrovni. Zahrnuje jak analogové, tak digitální objekty. Domníváme se, že je i dostatečně flexibilní pro začlenění nových formátů zaznamenaných znalostí, jež přinese technologický vývoj v budoucnu. Je samozřejmě zapotřebí počítat s tím, že vymezení jednotky organizace na filozofické úrovni neumožní přímou aplikaci, ale bude vyžadovat tvůrčí přístup a interpretaci.

#### **4.4 Proces organizace znalostí**

Proces organizace znalostí může probíhat i v mysli (přesně řečeno v mozku) – tato problematika je předmětem zájmu psychologie, neurologie a kognitivní vědy. Předmětem zájmu oboru organizace znalostí je ovšem výhradně organizace znalostí v externích pamětech (dokumentech, úložištích, kolekcích). Na obrázku 16 je znázorněn s použitím Erikssonovy-Penkerovy notace pojmový model procesu organizace znalostí a přístupu ke znalostem. V modelu jsou zachyceny kromě posloupnosti jednotlivých procesů i jejich cíle a artefakty představující jejich vstupy a výstupy. Do modelu jsou zahrnuti i aktéři, jejichž úloha je neméně významná jako specifika

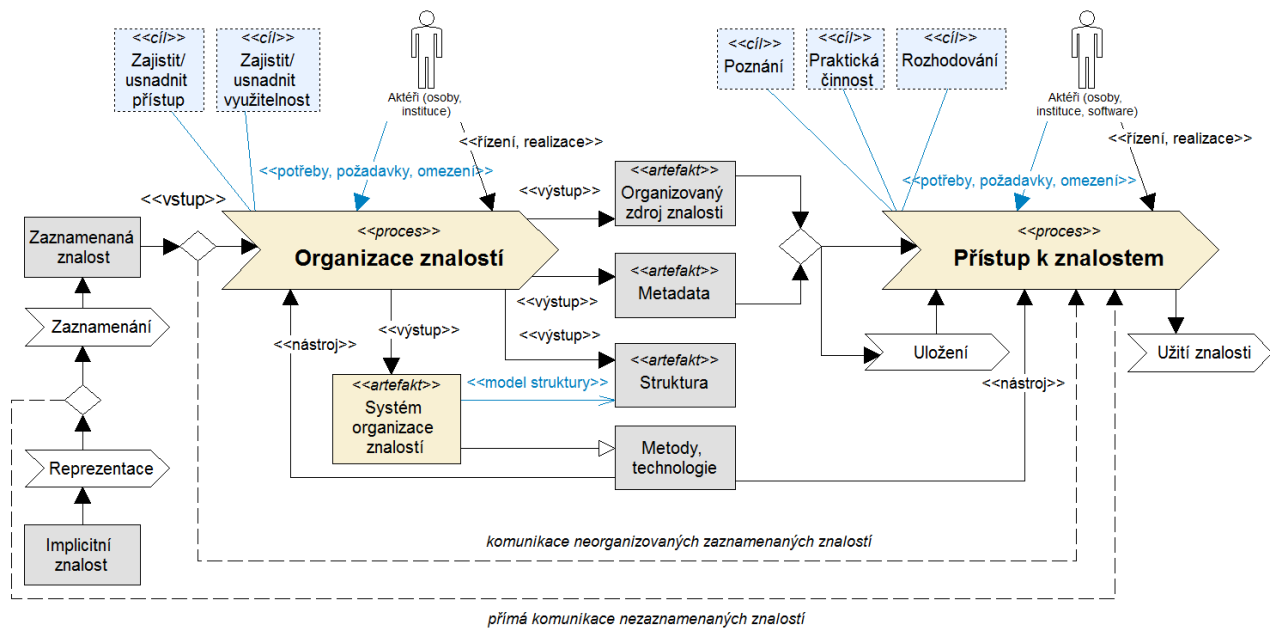
organizovaných zdrojů. Oni (resp. jejich potřeby a požadavky) určují, která z možných variant uspořádání se bude v daném případě realizovat.

Organizace znalostí je typem zpracovatelského procesu, jenž transformuje vstup v podobě zaznamenaných znalostí na výstup, kterým je organizovaný zdroj znalostí. Proces organizace se projevuje snahou vytvořit pořádek (organizaci, uspořádání, strukturu) místo dosavadního chaosu, přičemž se uplatňuje princip ekvivalence. Cílem transformace zaznamenaných znalostí je zajistit, aby vykazovaly vlastnosti vedoucí k dosažení cíle organizace, tj. umožnění přístupu a využití (viz kapitola 4.1).

Pohled na organizaci znalostí z časového hlediska umožňuje určit ‚směry‘ na časové ose vzniku organizovaného zdroje znalostí: Z hlediska času je možné organizaci v externích pamětech vnímat ve směru časové osy vzniku organizovaného zdroje znalostí: ‚dopředná‘ organizace probíhá už při vzniku zdroje, před jeho publikováním nebo zařazením do kolekce či do jiného kontextu; ‚zpětný‘ směr označuje (re)organizaci již existujících zdrojů a jejich kolekcí. Cílem těchto aktivit může být pořádek na místo původního chaosu, nebo reorganizace – jiné uspořádání již uspořádané množiny informací.

Je zřejmé, že má-li proces organizace znalostí být efektivní, musí být prováděn uvědoměle, tj. s vědomím toho, pro jaké procesy budou jeho výstupy představovat vstupy. Proto při jeho analýze věnujeme pozornost i procesům přístupu ke znalostem a jejich využívání. Typologie přístupu ke znalostem ovlivňuje způsob, jak budou znalosti organizovány, a má rovněž vliv na volbu výstupů z procesu organizace. V současné době jsou nejdůležitějšími typy přístupu strukturovaný přístup s metadaty, plnotextové vyhledávání, listování, přístup prostřednictvím sociálních sítí a přístup zajišťovaný softwarovými algoritmy (např. *data mining*). Každý z těchto přístupů má své přednosti a nedostatky, volba optimálního typu závisí na konkrétním prostředí, typu organizovaných zdrojů, předpokládaných uživatelích a účelu, kterému mají organizované znalosti sloužit.

Kromě komunikace organizovaných znalostí jsou v modelu ještě vyznačeny alternativy přímé komunikace, které rovněž umožňují přístup ke znalostem a jejich využití. To lze uskutečnit dvěma způsoby. Jednak formou přímé komunikace reprezentovaných, ale nezaznamenaných explicitních znalostí (ústní podání, vysílání), jednak přímou komunikací zaznamenaných explicitních znalostí bez jejich organizování, například sdílením v rámci sociálních sítí. Tyto varianty upozorňují na to, že proces organizace znalostí není povinnou fází komunikačního cyklu znalosti.



Obr. 16 Pojmový model procesu organizace znalostí a procesu přístupu ke znalostem

Pro úplnost je na obrázku 16 uveden i proces reprezentace znalostí, tj. konverze implicitních (tacitních) znalostí na znalosti explicitní neboli komunikovatelné, i když jeho zkoumání se obor organizace znalostí nevěnuje. Tento proces zkoumají kognitivní vědy a umělá inteligence (viz kapitola 1.1).

Po procesu reprezentace následuje buď již zmíněný alternativní proces přímého přístupu k reprezentovaným znalostem, nebo proces zaznamenání znalosti. V průběhu procesu zaznamenání probíhá tzv. mikroorganizace, tj. tvorba struktury zaznamenané znalosti uspořádáním jejích obsahových i formálních elementů. Především u elektronických zdrojů může být navíc už ve fázi zaznamenání explicitně prostřednictvím metadat vyjadřováno to, co u ‚tradičních‘ dokumentů bývá výsledkem informační analýzy v rámci procesu organizace, tj. současně s vytvořením dokumentu proběhne jeho indexace (např. přidělením autorských tagů) jako součást makroorganizace.

Východiskem pro uvažování o procesu organizace znalostí se pro nás stane pracovní vymezení jednotky organizace jako zaznamenané znalosti, jež představuje vstup procesu organizace. Proces organizace znalostí se realizuje s využitím nástrojů v podobě metod a technologií. Abstraktním vyjádřením výstupu procesu organizace znalostí je pojem struktura – tj. soubor vzájemných vztahů organizovaných prvků.

Ke konkrétnímu produktu organizace znalostí v podobě organizovaného zdroje znalostí jsou doplněny ještě další nepovinné, ale významné výstupy – metadata, tj. informace umožňující uživateli organizované zdroje vyhledávat a využívat, a systém organizace znalostí, využívaný jako nástroj organizace i jako nástroj přístupu ke znalostem. Na obrázku je vyznačena i funkce systému organizace znalostí jako modelu struktury organizovaného celku (blíže viz kapitola 4.5.3). Výstupy procesu organizace znalostí mohou být buď uloženy, nebo se stát bezprostředním vstupem procesu přístupu k znalostem.

I samotný proces organizace je samozřejmě možné dále dekomponovat, přičemž ani jeden z podprocesů není povinný. Za nejvýznamnější jsou označovány procesy identifikace (viz kapitola 4.2), obsahové analýzy a indexace (charakteristika těchto procesů byla podána v rámci kapitoly

1.5.3 věnované konceptualizaci) a začlenění do kontextu. Začlenění do kontextu se děje s cílem vyjádřit vztah organizované jednotky k jiným entitám v témže nebo v jiném informačním zdroji: zařazením do nějaké skupiny (kategorizace, klasifikace) nebo začleněním do sémantické nebo citační sítě (odkazy, citační vazby).

K dalším typickým podprocesům organizace znalostí patří:

- Fyzické uspořádání, tj. umístění, seskupení/kolokace a řazení (viz kapitola 4.2).
- Popis (formální, povrchová analýza) – určení a označení jednotek formy či formálních vlastností zdroje. Popis typicky spočívá v rozpoznání vlastností popisovaného objektu a v určení jejich hodnoty.
- Analýza struktury – určení prvků obsahu a formy a jejich vztahů.
- Explicitní vyjádření výsledků analýzy: jazykovými výrazy, schématem struktury, odkazy/směrováním – mezi zdroji, příp. mezi zdroji a metadaty.
- Hodnocení (kvality, relevance) zdroje.
- Kategorizace – logické seskupení souvisejících jednotek do tříd ekvivalence.
- Klasifikace – kromě logického uspořádání souvisejících jednotek do tříd ještě určuje vzájemný vztah tříd.
- Reorganizace – opakované provedení organizace.

Mnohé podprocesy spolu vzájemně souvisejí, navazují na sebe v čase (například procesu vyjádření výsledků analýzy předchází proces samotné analýzy, účelem popisu je identifikace entity), nebo jsou seskupitelné do větších celků (například klasifikace je součástí procesu začlenění do kontextu). Avšak přestože se o to někteří autoři<sup>314</sup> pokoušeli, v jednom případě dokonce na úrovni mezinárodní normy ISO 5963, je třeba konstatovat, že plná algoritimizace většiny procesů organizace znalostí je krajně obtížná, ne-li nemožná. V plné míře to platí pro klíčový proces indexace obsahu, tj. „*pracovní postup popisování nebo identifikace dokumentu ve vztahu na jeho věcný obsah*“<sup>315</sup>. Norma ISO 5963 v části 4.3 specifikuje tři fáze procesu indexace: fáze analýzy dokumentu a určení jeho věcného obsahu, fáze identifikace pojmů obsažených v určeném předmětu, fáze jazykového vyjádření pojmů. Nehledě na to, že i v samotné normě se připouští, že se tyto procesy mohou v praxi překrývat, je linearita představeného modelu v rozporu se současným chápáním komunikace jako síťového procesu. Navíc empirické výzkumy<sup>316</sup> ukázaly, že v praxi nejsou tyto procesy vnímány jako oddělené a neprobíhají ve stanovené sekvenci.

---

<sup>314</sup> WELLISCH, Hans H. A flow chart for indexing with a thesaurus. In: *Journal of the American Society for Information Science*. May/June 1972, 23(3), 185-194. ISSN 0002-8231.

<sup>315</sup> ČSN ISO 5963 (01 0174). *Dokumentace. Metody analýzy dokumentů, určování jejich obsahu a výběru lexikálních jednotek selekčního jazyka*. Praha: Český normalizační institut, 1995, s. 5.

<sup>316</sup> RONDEAU, Sophie. The life and times of aboutness: a review of the library and information science literature. In: *Evidence based library and information practice*. 2014, 9(1), 14-35. ISSN 1715-720X.

## 4.5 Realizační kontext procesu organizace znalostí

### 4.5.1 Aktéři – osobnosti, institucionální a komunikační základna

Na nejobecnější úrovni lze za klíčové aktéry organizace znalostí označit tvůrce obsahu, jeho zprostředkovatele a koncové uživatele, přičemž v každé z rolí může vystupovat jak člověk, tak počítačový program. Rozšíření okruhu aktérů organizace znalostí bylo již v kapitole 4.2 uvedeno jako jedno z paradigmat řešení aktuálních problémů organizace znalostí. Birger Hjørland zmiňuje tyto oblasti organizace znalostí, jež rozlišuje podle technologické platformy: analogové prostředí (fyzických) knihoven, archivů a muzeí a digitální klasické bibliografické databáze a *Internet*.<sup>317</sup> Minimálně od doby, kdy proběhl přechod od industriální ke znalostní ekonomice, je zřejmé, že tradiční paměťové instituce už zdaleka nejsou jedinými typy institucí, jež na profesionální bázi organizují znalosti. *Pracovní skupina Kongresové knihovny pro budoucnost bibliografické kontroly*<sup>318</sup> ve studii *On the record* („O záznamu“, ale i „K zveřejnění“) z roku 2008 formuluje jako dva ze tří základních principů nutnost redefinovat bibliografickou kontrolu a redefinovat bibliografické univerzum (třetím principem bylo redefinování role *Kongresové knihovny*). V obou případech navrhuje rozšířené vnímání těchto pojmů, které byly v minulosti často chápány jako synonyma pro katalogizaci v knihovnách. Konstatuje, že oddělování komunit tvůrců, vydavatelů, prodejců, distributorů a producentů databází, knihoven, archivů a muzeí, už není udržitelné ani funkční, a vyzývá k chápání bibliografického univerza v tomto rozšířeném pojetí.<sup>319</sup> Účastníci panelové diskuse uspořádané v roce 2014 k 25. výročí společnosti *ISKO*<sup>320</sup> rovněž konstatují, že je nezbytné, aby obor organizace znalostí zahrnul všechny oblasti lidských i strojových aktivit, organizující zaznamenané znalosti.

#### Institucionální aktéři

Obrázek 17 ve zjednodušené formě dokládá šíři spektra institucionálních aktérů a zároveň poukazuje na interdisciplinární a transdisciplinární charakter domény. Zahrnuje okruh informací v institucích a organizacích: podnikovou informatiku, informatiku veřejné správy a vědeckou informatiku. Jsou uvedeni i aktéři vytvářející a spravující veřejné informační kolekce a publikační a distribuční sektor. Donedávna platilo, že producenti, vydavatelé, distributoři, agregátoři a zpracovatelé představují zřetelně oddělené oblasti, propojené pevným řetězcem sekvenčně následujících aktivit. Tyto vztahy se však v současnosti mění v důsledku paradigmatických změn, charakterizovaných v kapitole 4.2. Významnou změnu přinesly vyhledávače internetového obsahu, umožňující vyhledat jakýkoli online zpřístupněný zdroj bez ohledu na jeho producenta. I když neprodukují žádné vlastní zdroje, poskytují vyhledávače jednotné vyhledávací rozhraní a pro značnou část koncových uživatelů jsou primárním přístupovým bodem k online zdrojům. Integrovaní procesy zasahují i další aktivity, zejména publikování a zpracování. Vycházejí z toho, že současné technologie poskytují jednotnou bázi jak

---

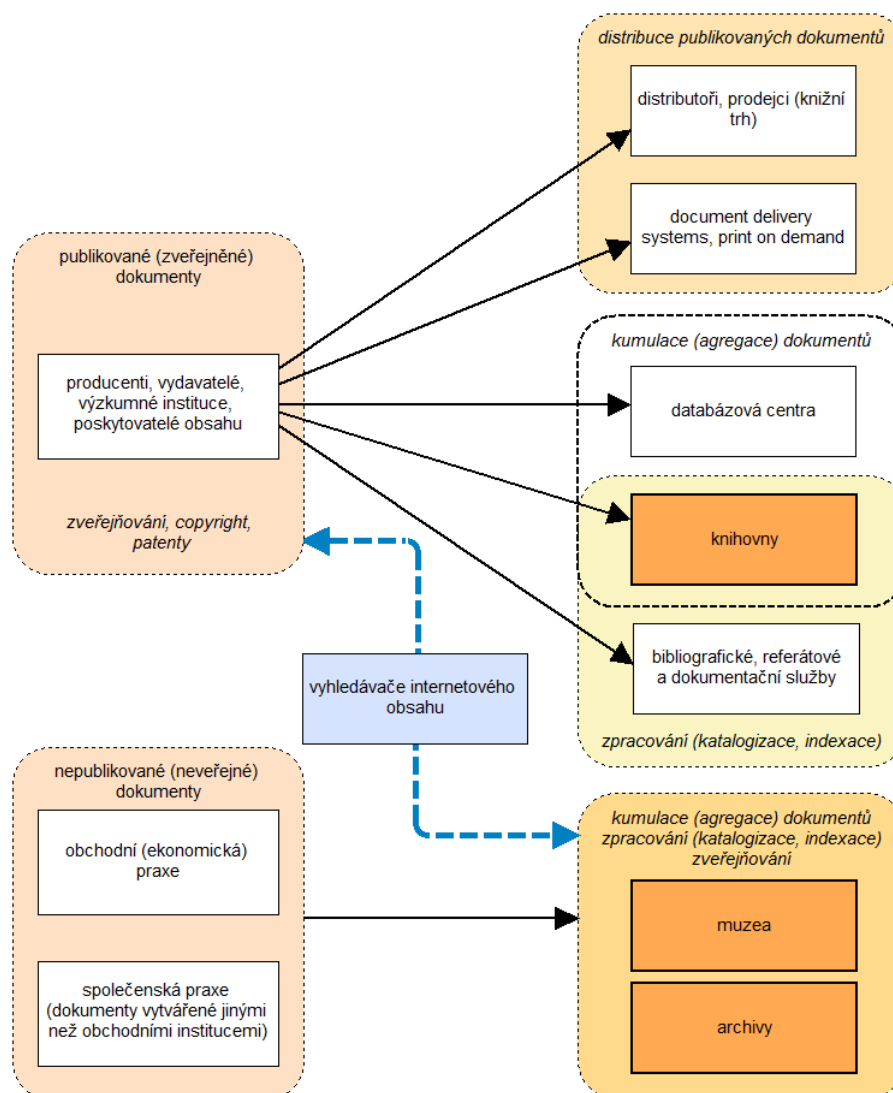
<sup>317</sup> HJØRLAND, Birger. Knowledge organization (KO). In: *Knowledge organization*. 2016, **43**(6), 478-482. ISSN 0943-7444.

<sup>318</sup> Domovská stránka: <http://www.loc.gov/bibliographic-future/> [cit. 2018-01-30].

<sup>319</sup> Library of Congress. *On the record: report of The Library of Congress Working Group on the Future of Bibliographic Control* [online]. Washington: Library of Congress, 2008 [cit. 2016-11-07]. 44 s. Dostupné z: <https://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>.

<sup>320</sup> GREEN, Rebecca. ISKO and Knowledge organization's 25<sup>th</sup> anniversary: the future of Knowledge organization and ISKO panel discussion. In: *Knowledge organization*. 2014, **41**(4), 327-331. ISSN 0943-7444.

pro odborné (vědecké) informace, tak pro informace ze sféry byznysu, státní správy a samosprávy, a samozřejmě i pro osobní informační systémy odborníků i laiků.



Obr. 17 Aktéři a kontext organizace znalostí

## Organizace a odborné komunity

Nejvýznamnějším sdružením věnujícím se v období po druhé světové válce organizaci znalostí se stala *Skupina pro výzkum klasifikace – Classification Research Group (CRG)*. Představovala volné seskupení především britských klasifikačních teoretiků a praktiků, jež bylo ustaveno v roce 1952. Skupina se zaměřovala především na fasetové doménově specifické systémy organizace znalostí. Členové pokračovali v rozvíjení teorie fasetové analýzy S. R. Ranganathana a podíleli se na 2. vydání Blissova *Bibliografického třídění*, přepracovaném na plně fasetovém principu.

Na mezinárodní úrovni je institucionální základnou oboru mezinárodní společnost pro organizaci znalostí – *International Society for Knowledge Organization (ISKO)*, která již od svého založení v roce 1989 usiluje o to, stát se platformou pro odborníky z oboru organizace znalostí. Vydává stěžejní oborový časopis *Knowledge organization* (byl založen v roce 1974, do roku 1993 vycházel pod názvem *International classification*). Zpracovává průběžnou oborovou bibliografii

*Knowledge organization literature* (Literatura z organizace znalostí, dostupné z <http://www.isko.org/lit.html>), pro jejíž účely byla vytvořena stejnojmenná klasifikace *KO Literature*. *ISKO* je vydavatelem ediční řady *Advances in knowledge organization* (Pokroky v organizaci znalostí) a v dvouletých intervalech pořádá mezinárodní konference.<sup>321</sup>

Nejvýznamnější výzkumné i aplikační aktivity v oblasti systémů organizace znalostí v současné době vyvíjí komunita *NKOS* (*Networked Knowledge Organization Systems/Services/Structures – Systémy/služby/struktury organizace znalostí v síťovém prostředí*)<sup>322</sup>, jež je volným mezinárodním sdružením odborníků. Pořádá pravidelné workshopy a její pracovní skupina je tvůrcem návrhu aplikačního profilu *Dublin Core* pro systémy organizace znalostí<sup>323</sup>.

V oblasti organizace znalostí vyvíjí aktivity rovněž *Mezinárodní federace knihovnických sdružení a institucí IFLA*, založená v roce 1929. V současné době se problematice věnují především *Sekce IFLA pro katalogizaci*<sup>324</sup> a *Sekce IFLA pro obsahovou analýzu a přístup* (do roku 2016 *Sekce pro klasifikaci a indexaci*).<sup>325</sup> Sekce pro obsahovou analýzu a přístup se zaměřuje na metody a zajištění věcného přístupu v katalozích, bibliografiích a rejstřících k dokumentům všeho druhu, včetně elektronických dokumentů. Slouží jako fórum pro knihovny a informační služby všech typů, jakož i pro výrobce a uživatele nástrojů klasifikace a indexace. Prosazuje standardizaci a jednotné uplatňování nástrojů klasifikace a indexace institucemi, které vytvářejí nebo využívají bibliografické záznamy. Podporuje výzkum a jeho výsledky šíří prostřednictvím otevřených setkání a publikací. Členové sekce se podíleli na tvorbě referenčního pojmového modelu věcných autorit *FRSAD* a vícejazyčného terminologického slovníku katalogizace *Multilingual dictionary of cataloguing terms and concepts (MulDiCat)*.

*Konsorciium pro Mezinárodní desetinné třídění (UDC Consortium, UDCC)* je nezisková organizace se sídlem v Haagu, Nizozemsko. Jeho účelem je správa a distribuce *Mezinárodního desetinného třídění* a podpora jeho používání a vývoje. V dvouletých intervalech pořádá odborné konference zaměřené na problematiku klasifikace a systémů organizace znalostí. Konference má formát semináře (přednášky následované odbornou diskusí) s názvem *UDC Seminar*.

V USA vyvíjí aktivity v oblasti organizace znalostí *Zájmová skupina pro výzkum klasifikace (SIG/CR, Special Interest Group for Classification Research)*, ustavená při *Sdružení pro informační vědu a technologie (Association for Information Science and Technology, ASIST)*. Skupina pořádá každoročně v rámci výroční konference *ASIST* workshop věnovaný problematice organizace znalostí; příspěvky jsou publikovány v sborníku *Advances in classification research* (Pokroky ve výzkumu klasifikace).

---

<sup>321</sup> Přehled struktury a regionálních aktivit *ISKO* viz NOWAKOVÁ, Jana. *Aktivity, úkoly a cíle International Society for Knowledge Organization (ISKO)*. Praha, 2013. 57, xvii s. Bakalářská práce (Bc.). Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví. Vedoucí bakalářské práce Richard Papík.

<sup>322</sup> Dokumentace činnosti je dostupná z: <http://nkos.slis.kent.edu/> [cit. 2018-01-30].

<sup>323</sup> Dokumentace činnosti je dostupná z: <http://dublincore.org/groups/nkos/> [cit. 2018-01-30].

<sup>324</sup> Dokumentace činnosti sekce je dostupná z: <https://www.ifla.org/cataloguing> [cit. 2018-01-30].

<sup>325</sup> Dokumentace činnosti sekce je dostupná z: <https://www.ifla.org/subject-analysis-and-access> [cit. 2018-01-30].

## Osobnosti

Výběr nejvýznamnějších osobností organizace znalostí jsme se snažili objektivizovat přihlédnutím k počtu děl, jejichž metadata jsou uvedena ve znalostní bázi organizace znalostí (viz <http://ko.cuni.cz/node/3282>) a k počtu odkazů na jednotlivé autory ve jmenném rejstříku monografie *Organizace znalostí: klíčová témata*<sup>326</sup>. Ve velmi hrubém historickém přehledu je možné zaznamenat tyto tři skupiny:

1. ‚Zakladatelská‘ generace oboru z přelomu 19. a 20. století, jejímiž příslušníky byli zakladatelé moderního knihovnictví a informační vědy. Jejými představiteli jsou Melvil Dewey (1851–1931), Charles Ammi Cutter (1837–1903), Henry Evelyn Bliss (1870–1955), William Charles Berwick Sayers (1881–1960), zakladatel dokumentalistiky Paul Otlet (1868–1944) a Shiyali Ramamrita Ranganathan (1892–1972).
2. Generace, jež svůj přínos formulovala v souvislosti s nástupem počítačů v 50.–80. letech 20. století, v době předcházející rozšíření *Internetu*. Jejými představiteli byli především členové *Skupiny pro výzkum klasifikace*: Derek Austin (1920–2001), Eric Coates, Jason Farradane (1906–1989), Robert Arthur Fairthorne (1904–2000), Douglas John Foskett (1918–2004), Barbara Kyleová, Derek Langridge, Jack Mills (1918–2010), Bernard I. Palmer, Brian C. Vickery (1918–2009), A. Jack Wells. Zásadní úlohu při formování oboru sehrála Ingetraut Dahlbergová (1927–2017), zakladatelka časopisu *International classification* a spoluzakladatelka a první předsedkyně *ISKO* (v letech 1989–1996). V kontextu tehdejšího Československa formoval obor na moderních základech Blahoslav Kovář (1926–2002).
3. K významným představitelům oboru v současnosti, kteří jsou aktivní v oblasti výzkumné, pedagogické a publikační, patří Clare Beghtolová (*Fakulta informací Torontské univerzity, Kanada*), Vanda Broughtonová (*University College London, Velká Británie*), Robert J. Glushko (*Informační škola Kalifornské univerzity v Berkeley, USA*), Claudio Gnoli (*Univerzita v Pavii, Itálie*), editorka *DDT* Rebecca Greenová (*OCLC, USA*), Birger Hjørland (*Královská škola knihovny a informační vědy, Dánsko*), současný šéfredaktor časopisu *Knowledge organization* Richard P. Smiraglia (*USA*), autorka významné publikace *The intellectual foundation of information organization* (Intelektuální základy organizace informací)<sup>327</sup> Elaine Svenoniová (*USA*), Marcia Lei Zengová (*Kentská státní univerzita, USA*).

### 4.5.2 Typologie zdrojů v doméně organizace znalostí

Představu o struktuře zdrojů v doméně umožnila konkretizovat množina bibliografických entit ve znalostní bázi organizace znalostí, při jejíž tvorbě se řešitelé projektu *NAKI* snažili o zachycení reprezentativního vzorku celosvětové produkce (viz přehled struktury modulu bibliografických entit v úvodu kapitoly 4).

Bibliografické entity zahrnuté ve znalostní bázi mají následující obsahovou strukturu: zhruba dvě třetiny zdrojů (58 %) pokrývají oblast organizace znalostí, pětina (18 %) zachycuje nezbytný teoretický a filozofický kontext (lingvistika, sémantika, sémiotika, logika, kognitivní

---

<sup>326</sup> KUČEROVÁ, Helena. *Organizace znalostí: klíčová témata*. 1. vyd. Praha: Karolinum, 2017, s. 267–269. ISBN 978-80-246-3587-3 (brož). ISBN 978-80-246-3597-2 (pdf).

<sup>327</sup> SVENONIUS, Elaine. *The intellectual foundation of information organization*. Cambridge (Mass): MIT Press, 2000. xiv, 255 s. Digital Libraries and Electronic Publishing Series. ISBN 0-262-19433-3. ISBN 978-0-262-19433-4.



vědy), cca 14 % zdrojů je z oblasti počítačových disciplín a umělé inteligence, zbylých 10 % tvoří zdroje z oborů informační věda, znalostní management a podniková informatika.

Co do zastoupených typů dokumentů představují téměř polovinu (46 %) články z časopisů, pětinu (21 %) monografie, desetinu (11 %) systémy organizace znalostí, zbývající pětina zdrojů připadá na ostatní typy (8 % příspěvky ve sborníku, 4 % výzkumné a technické zprávy, 2 % kapitoly z knih, 2 % encyklopedická hesla, 2 % vysokoškolské kvalifikační práce, 2 % sborníky, 1 % standardy a 1 % učební materiály).

Na základě vyhodnocení tohoto reprezentativního vzorku lze konstatovat, že v doméně organizace znalostí se používají obdobné typy zdrojů jako v analogických doménách zaměřených na komunikaci informací ve společnosti. Za doménově specifické zdroje lze prohlásit systémy organizace znalostí, kterým bude věnována následující kapitola. Poměrně vysoký podíl monografií lze interpretovat jako indikátor zralosti disciplíny. Zároveň lze zastoupení četných příruček a učebnic v rámci této skupiny chápat jako potvrzení inženýrského charakteru oboru (viz kapitola 4.6.1).

Výsledky teoretického výzkumu se komunikují prostřednictvím odborných periodik, monografií a sborníků z konferencí. Ty nejvýznamnější byly již uvedeny v předchozí kapitole: Stěžejními oborovými zdroji jsou časopis *Knowledge organization* (ISSN 0943-7444) a sborníky z konferencí, jejichž pořadatelé jsou společnosti *ISKO*, *Konsorcium pro MDT* a *ASIST* (viz kapitola 4.5.1).

V kapitole 1.2 byla charakterizována znalostní báze jako klíčová komponenta znalostního systému, již tvoří množina výroků, reprezentujících tvrzení o univerzu diskurzu. Znalostní bázi lze tedy rovněž považovat za specifický typ zdroje, charakteristický pro doménu organizace znalostí. Představuje vysoce formalizovaný produkt procesu organizace znalostí, jehož jednotkou jsou výroky vyjádřené v jazyce pro reprezentaci znalostí.

### **Oborové standardy a pravidla**

Proces organizace znalostí je regulován standardy a pravidly. Stěžejní význam pro tvorbu popisných metadat mají *Mezinárodní principy katalogizace (International cataloguing principles, ICP)*<sup>328</sup>, jejichž aktuální verze z roku 2009 je založena na principech formulovaných Charlesem Ammi Cutterem již v roce 1876<sup>329</sup> a je kompatibilní s modelem bibliografického univerza *IFLA LRM*. Na obecných principech katalogizace jsou založena pravidla; v paměťových institucích jsou v současné době nejpoužívanějšími katalogizační pravidla pro popis a zpřístupňování zdrojů *RDA (Resource description and access)*<sup>330</sup>. Pro metodiku procesu indexace byla v roce 1985 přijata mezinárodní norma ISO 5963 *Metody analýzy dokumentů, určování jejich obsahu a výběru*

---

<sup>328</sup> TILLET, Barbara B., CRISTÁN, Ana Lupe, ed. *IFLA cataloguing principles: the statement of international cataloguing principles (ICP) and its glossary: in 20 languages*. München: Saur, 2009. 304 s. IFLA series on bibliographic control; 37. ISBN 978-3-598-24285-4.

<sup>329</sup> CUTTER, Charles Ammi. *Rules for a printed dictionary catalogue*. Washington: Government Printing Office, 1876. 89 s.

<sup>330</sup> Viz např. LICHTENBERGOVÁ, Edita. *Katalogizace podle RDA ve formátu MARC 21 – tištěné a elektronické monografie – katalogizace na úrovni minimálního/doporučeného záznamu* [online]. Praha: Národní knihovna ČR, 2016 [cit. 2018-01-30]. Dostupné z: <http://www.nkp.cz/o-knihovne/odborne-cinnosti/zpracovani-fondu/katalogizacni-politika/katalogizace-podle-rda-ve-formatu-marc-21-tistene-a-elektronicke-monografie-katalogizace-na-urovni-minimalniho-doporuceneho-zaznamu>.

*lexikálních jednotek selekčního jazyka* <sup>331</sup>.

Klíčovým standardem pro vývoj, používání a správu nástrojů pro proces organizace znalostí je mezinárodní norma věnovaná tezaurům a dalším řízeným slovníkům ISO 25964<sup>332</sup>, jejíž historie sahá do počátku 70. let 20. století. Aktuální vydání z let 2011 a 2013 připravila subkomise technické komise ISO/TC 46 *Informace a dokumentace*. Norma reflektuje změny, k nimž došlo zavedením informačních technologií do této oblasti: softwarové aplikace pro tvorbu a využívání tezaurů, technologie plnotextového vyhledávání ad. Od původního striktního zaměření na návrh tezaurů se záběr rozšířil směrem k obecnějšímu pojetí, použitelnému pro širší spektrum typů řízených slovníků a jiných systémů organizace znalostí. V druhé části normy, věnované interoperabilitě, jsou samostatné kapitoly věnovány klasifikačním schémátům, taxonomiím, předmětovým heslářům, ontologiím, terminologiím, seznamům jmenných autorit a seznamům synonym.

Vzhledem k těsné souvislosti organizace znalostí s jazykem hrají významnou roli standardy pro terminologickou práci, především ISO 1087-1 a ISO 704. Terminologická norma ISO 1087-1 obsahuje slovník s výkladem pojmů z oblasti terminologie. Byla vydána v roce 2000 a český překlad spolu s národními poznámkami vyšel v roce 2002<sup>333</sup>. Její textová část je doprovázena v informativní příloze pojmovými diagramy, jež grafickou formou znázorňují významové vztahy zahrnutých termínů.

Norma ISO 704 v aktuálním třetím revidovaném vydání z roku 2009<sup>334</sup> se věnuje principům a metodám terminologické práce. V normě jsou explicitně jmenovány tyto aktivity, jež zhruba odpovídají klíčovému subprocessu organizace znalostí, jímž je indexace: identifikace pojmů a pojmových vztahů, analýza a návrh pojmových systémů (tj. množin pojmů strukturovaných na základě jejich vzájemných vztahů), reprezentace pojmových systémů pojmovými diagramy, definování pojmů, označování pojmů termíny. Uvedené metody práce při tvorbě a správě terminologických systémů vycházejí z obecně teoretických lingvistických a sémantických principů a jsou dle tvůrců normy aplikovatelné i v dalších oblastech.

### **Oborové systémy organizace znalostí**

Stejně jako v jiných oborech je i pojmová základna organizace znalostí zachycena v systémech organizace znalostí, především v terminologických slovnících. Protože obor organizace znalostí má významný průnik s informační vědou (viz kapitola 4.6.1), jsou pro tuto doménu relevantní i zdroje, původně koncipované pro doménu informační vědy, resp. jejich části. Terminologickou základnu informační vědy a oblasti informačních technologií zachycují mezinárodní normy ISO

---

<sup>331</sup> ČSN ISO 5963 (01 0174). *Dokumentace. Metody analýzy dokumentů, určování jejich obsahu a výběru lexikálních jednotek selekčního jazyka*. Praha: Český normalizační institut, 1995. 10 s.

<sup>332</sup> ISO 25964-1:2011. *Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval*. 1st ed. Geneva: International Organization for Standardization, 2011-08-08. 152 s. – ISO 25964-2:2013. *Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies*. 1st ed. Geneva: International Organization for Standardization, 2013-03-04. 99 s.

<sup>333</sup> ČSN ISO 1087-1 (01 0501). *Terminologická práce – Slovník – Část 1: Teorie a aplikace*. Praha: Český normalizační institut, 2002. 38 s.

<sup>334</sup> ISO 704:2009. *Terminology work – Principles and methods*. 3rd ed. Geneva: International Organization for Standardization, 2009. 65 s.

5127 *Information and documentation – Foundation and vocabulary*<sup>335</sup> a ISO/IEC 2382 *Information technology – Vocabulary*<sup>336</sup>. Českou terminologii v rozsahu přes 3 000 termínů obsahuje *Česká terminologická databáze knihovnictví a informační vědy (TDKIV)*.<sup>337</sup>

Na pokrytí domény informační vědy v úplnosti aspiruje rovněž fasetový *Tezaurus ASIST pro informační vědu, technologii a knihovnictví (ASIS&T thesaurus of information science, technology, and librarianship)*.<sup>338</sup> Primárním účelem tezauru je indexace a vyhledávání zdrojů v oborových databázích a digitálních knihovnách a proto se soustředí na tematické pojmy a předměty (tj. na takové termíny, jež mohou sloužit k označení obsahu indexovaných zdrojů). Jeho třetí vydání z roku 2005 obsahuje 1 970 termínů, členěných do 18 faset s 10 subfasetami. Pro doménu organizace znalostí jsou významné především termíny z faset ‚aktivity a operace‘ (*activities and operations*), ‚atributy informací a dat‘ (*attributes of information and data*) a ‚znalosti a informace‘ (*knowledge and information*).

Jediným systémem organizace znalostí specializovaným na doménu organizace znalostí je klasifikace *KO Literature*. Sestavila ji v roce 1974 Ingetraut Dahlbergová pro účely indexace oborových zdrojů v průběžné oborové bibliografii (bibliografie původně vycházela jako příloha časopisu *International Classification/Knowledge organization*, od roku 2009 je volně dostupná na webu *ISKO*), v roce 1993 ji mírně upravili Gerhard Riesthuis a Ia C. McIlwainová. Od roku 2012 je klasifikace k dispozici online na webu *ISKO*.<sup>339</sup> Klasifikace koncepčně navazuje na univerzální klasifikační systém *Information Coding Classification (ICC)*<sup>340</sup>, jehož autorkou je rovněž Ingetraut Dahlbergová. *ICC* je fasetovým systémem a vychází z teorie integrativních úrovní Jamese K. Feiblemana, představené v kapitole 2.3.4.

Klasifikace *KO Literature* obsahuje 660 tříd členěných do tří hierarchických úrovní. Nejvyšší úroveň tvoří 10 tříd označených arabskými číslicemi 0-9.<sup>341</sup>

0 Formální oddíly/členění

1 Teoretické základy a obecné problémy

2 Klasifikační systémy a tezaury, struktura a konstrukce

3 Metodika klasifikace a indexace

4 Univerzální klasifikační systémy a tezaury – o nich

---

<sup>335</sup> ISO 5127:2017 *Information and documentation – Foundation and vocabulary*. 2nd ed. Geneva: International Organization for Standardization, 2017-05. 353 s. Části normy jsou dostupné z: <https://www.iso.org/obp/ui/#iso:std:iso:5127:ed-2:v1:en> [cit. 2018-01-30].

<sup>336</sup> ISO/IEC 2382:2015. *Information technology – Vocabulary* [online]. 1st ed. Geneva: International Organization for Standardization, 2015-05 [cit. 2018-01-30]. Dostupné z: <https://www.iso.org/obp/ui/#iso:std:iso-iec:2382:ed-1:v1:en>.

<sup>337</sup> *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online databáze]. Praha: Národní knihovna České republiky, 2003- [cit. 2018-01-30]. Dostupné z: <http://aleph.nkp.cz/cze/ktd>.

<sup>338</sup> REDMOND-NEAL, Alice, HLAVA, Marjorie M. K., ed. *ASIS&T thesaurus of information science, technology, and librarianship*. 3rd ed. Medford: Information Today on behalf of the American Society for Information Science and Technology, 2005. xiii, 255 s. ASIST monograph series. ISBN 978-1-57387-243-0. ISBN 1-57387-244-X (s CD-ROM).

<sup>339</sup> *Classification System for Knowledge Organization Literature* [online]. ©ISKO 2011-2016, last update 2016-08-03 [cit. 2018-01-30]. Dostupné z: <http://www.isko.org/scheme.php>.

<sup>340</sup> DAHLBERG, Ingetraut. ICC – Information Coding Classification: principles, structure and application possibilities. In: *International classification*. 1982, 9(2), 87-93. ISSN 0340-0050.

<sup>341</sup> Poznámka: Názvy tříd jsou uvedeny ve znění pracovního překladu klasifikace, který byl zpracován v rámci řešení projektu *NAKI* (dostupné z: <http://ko.cuni.cz/node/3286>).

- 5 Klasifikace (systémy organizace znalostí) speciálních objektů – o nich
- 6 Specializované (oborové) klasifikace a tezaury – o nich
- 7 Reprezentace znalostí prostřednictvím jazyka a terminologie
- 8 Aplikované třídění a indexace
- 9 Prostředí organizace znalostí

Třída 0 je určena pro vyjádření formy indexovaných zdrojů, zbývajících devět ‚obsahových‘ tříd je členěno do tří faset. První faseta (třídy 1–3) zahrnuje teorii (pořádání, klasifikační teorie, teorie pojmů, matematika, systémová teorie, psychologie), základní pojmy a metody (obsahová analýza, techniky, metody a procesy klasifikace a indexace). Druhá faseta (třídy 4–6) zachycuje aplikaci pojmů z první fasety na specifika oboru, tj. na systémy organizace znalostí (struktura, konstrukce, typologie). Třetí faseta (třídy 7–9) pokrývá kontext oboru, především jazykové (lingvistická a terminologická problematika reprezentace znalostí), sociální, právní a ekonomické aspekty.

### 4.5.3 Systém organizace znalostí

V úvodu ke kapitole 4.2 bylo zmíněno, že mnozí autoři považují systém organizace znalostí (zkratka SOZ) spolu s procesem organizace znalostí za stěžejní předmět domény. Je jediným typem zdroje, o němž lze prohlásit, že je specifický pro doménu organizace znalostí. Z toho důvodu se mu v této kapitole budeme věnovat podrobněji.<sup>342</sup>

Navzdory abstraktní povaze všech tří slov, z nichž je vytvořen termín ‚systém organizace znalostí‘, neoznačuje tento termín žádnou abstrakci, ale konkrétní artefakt, tj. záměrně za určitým účelem vytvořenou věc, přesně řečeno pomůcku pro realizaci procesů organizace znalostí a přístupu ke znalostem. Na obrázku 16 v kapitole 4.4 je znázorněn dvojitý vztah systému organizace znalostí k procesu organizace znalostí: představuje jeho výstup a zároveň jeden z jeho nástrojů, tj. vstup. Toto těsné sepětí procesu a nástroje organizace znalostí se projevuje i v tom, že na SOZ lze vztáhnout i všechny konstantní problémy organizace znalostí představené v kapitole 4.2, tj. obtížné určování efektivnosti, sémantickou nejednotnost v prostoru i v čase (subjektivita, kulturní rozdíly, zaujetí, vývoj poznání), lingvistické a sémantické problémy, a samozřejmě i dilema teoretických a pragmatických řešení. V plné míře se SOZ týká i konstatovaná krize paradigmatu.

Fulvio Mazzocchi analogicky k rozlišování mezi širokým a úzkým pojetím organizace znalostí u Birgera Hjørlanda<sup>343</sup> uvažuje široké a úzké pojetí SOZ.<sup>344</sup> Konstatuje, že v nejširším slova smyslu by bylo možné za systém organizace znalostí považovat jakoukoli množinu uspořádaných znalostí (jako příklad uvádí encyklopedie, knihovny, bibliografické databáze, pojmové modely, teorie, vědní obory). Takové pojetí zhruba odpovídá termínu *organizing system* (organizační systém), který Robert J. Glushko definoval jako jakoukoli záměrně uspořádanou kolekci

---

<sup>342</sup> Poznámka: Text této části je zkrácenou a aktualizovanou verzí článku BRATKOVÁ, Eva, KUČEROVÁ, Helena. Systémy organizace znalostí a jejich typologie. In: *Knihovna: knihovnická revue*. 2014, **25**(2), 5-29. ISSN 1801-3252 (Print). ISSN 1802-8772 (Online). Typologie SOZ představená v této kapitole byla oproti publikované verzi mírně upravena.

<sup>343</sup> HJØRLAND, Birger. Knowledge organization (KO). In: *Knowledge organization*. 2016, **43**(6), s. 482. ISSN 0943-7444.

<sup>344</sup> MAZZOCCHI, Fulvio. Knowledge organization system (KOS) [online]. Version 1.1. 2017-07-13 [cit. 2018-01-30]. In: *ISKO Encyclopedia of knowledge organization*. ISKO, 2017. Dostupné z: <http://www.isko.org/cyclo/kos>.

zdrojů<sup>345</sup>. V kontextu oboru organizace znalostí je však obvyklé používat termín SOZ v užším smyslu k označení nástroje organizace ve formě organizačního schématu, definujícího strukturu organizované množiny zaznamenaných znalostí.

### Typologie systémů organizace znalostí

Pro vymezení pojmu systém organizace znalostí má typologie velký význam. Jak ukázal literární průzkum, všeobecně akceptovaná intenzionální definice SOZ zatím chybí a mnozí autoři ji nahrazují extenzionální charakteristikou spočívající ve výčtu typů. Zevrubný aktuální přehled stávajících typologií SOZ podává Fulvio Mazzocchi v článku *Knowledge organization system (KOS)* v encyklopedii organizace znalostí *ISKO*.<sup>346</sup> Často citovaný seznam zpracovaný Gail Hodgeovou je koncipován obdobně jako definice ontologie zpracovaná konsorciem *OMG*, uvedená v úvodu kapitoly 1.1. Zahrnuje „klasifikační schémata, jež organizují materiál na všeobecné úrovni (jako například knihy na regále), předmětová hesla, jež umožňují detailnější přístup, soubory autorit, jež slouží k řízení variantních verzí klíčových informací (jako například geografických jmen nebo osobních jmen) [...] a rovněž méně tradiční schémata jako sémantické sítě a ontologie“<sup>347</sup>. Obdobně v datovém slovníku *SKOS* se používá pro vymezení SOZ následující výčet: „tezaury, klasifikační schémata, systémy předmětových hesel a taxonomie“.<sup>348</sup>

Následující vlastní typologie SOZ znázorněná v tabulce na obrázku 18 je kompilací existujících typologií, přičemž jako hlavní zdroj byly využity typologie Marcii Lei Zengové<sup>349</sup> a Gail Hodgeové<sup>350</sup>. Výsledkem je škála, jež zahrnuje 16 typů SOZ. Základním kritériem členění je jejich sémantická síla, jež byla v kapitole 2.1.1 definována jako schopnost znaku vyjádřit přesně a v úplnosti obsah entity. Obrazné pojmenování sémantická síla označuje komplex struktury a funkce SOZ, toto kritérium je tedy založeno na systémovém přístupu. Při řazení jednotlivých typů bylo postupováno v souladu s teorií integrativních úrovní (viz kapitola 2.3.4), tj. podle jejich „složitosti“. Dilema sladění požadavků na sémantickou sílu a zároveň jednoduchost byla již diskutována v kapitole 1.6.1 v souvislosti s reprezentací a v kapitole 2.1.1 v souvislosti s požadavky na inforatické ontologie. Obdobně v kontextu SOZ se předpokládá, že spolu s tím, jak stoupá komplexnost SOZ, stoupá i jeho vyjadřovací potenciál a tím i míra přesnosti a úplnosti vyhledávání realizovaného s jeho pomocí. Zároveň však vzrůstá i obtížnost a pracnost konstrukce takového systému a do jisté míry i obtížnost jeho používání.

Přehled začíná nejjednodušším typem – seznamy slov, a uzavírají ho terminologické ontologie – nástroje umožňující reprezentovat nejbohatší strukturu obsahů a vztahů mezi nimi. S určitým

---

<sup>345</sup> GLUSHKO, Robert J., ed. *The discipline of organizing*. Cambridge (MA): MIT Press, 2013, s. 1-2. ISBN 978-0-262-51850-5 (brož.).

<sup>346</sup> MAZZOCCHI, Fulvio. *Knowledge organization system (KOS)* [online]. Version 1.1. 2017-07-13 [cit. 2018-01-30]. In: *ISKO Encyclopedia of knowledge organization*. ISKO, 2017. Dostupné z: <http://www.isko.org/cyclo/kos>.

<sup>347</sup> HODGE, Gail. *Systems of knowledge organization for digital libraries: beyond traditional authority files*. Washington: The Digital Library Federation, Council on Library and Information Resources, 2000, s. 3. ISBN 978-1-933645-06-3. Dostupné též z: <http://www.clir.org/pubs/reports/pub91/> [cit. 2018-01-30].

<sup>348</sup> *SKOS: Simple knowledge organization for the web* [online]. Cambridge (MA): World-Wide Web Consortium, 2012 [cit. 2018-01-30]. Dostupné z: <https://www.w3.org/2004/02/skos/intro>.

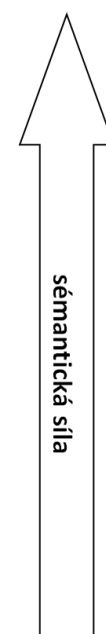
<sup>349</sup> ZENG, Marcia Lei. *Knowledge organization systems (KOS)*. In: *Knowledge organization*. 2008, 35(2-3), s. 160-182. ISSN 0943-7444.

<sup>350</sup> HODGE, Gail. *Systems of knowledge organization for digital libraries: beyond traditional authority files*. Washington: The Digital Library Federation, Council on Library and Information Resources, 2000, s. 4-7.

zjednodušením by se daly vztahy postupného narůstání komplexnosti mezi skupinami typů označenými v tabulce písmeny vyjádřit následovně: skupina A = seznam slov; skupina B = A + definování významu slov; skupina C = B + seskupení významově příbuzných slov (vztahy ekvivalence); skupina D = C + určení preferovaných termínů; skupina E = D + určení hierarchických vztahů; skupina F = E + určení asociativních vztahů; skupina G = F + možnost odvozování a usuzování.

Na obrázku 18 je dále vyznačeno, že typy 1–9 se zaměřují na organizaci slov (termínů), zatímco typy 10–16 jsou nástrojem organizace na pojmové úrovni. S použitím mírně upravené trojice kategorií Gail Hodgeové<sup>351</sup> je aplikováno ještě kritérium členění založené na typologii struktur a vyjadřovaných vztahů, jehož výsledkem je trojice kategorií: slovníky s jednoduchou lineární strukturou, klasifikace se stromovou strukturou a pojmové sítě.

kategorie		typ	anglicky		
pojmové sítě síťová struktura	organizace pojmnů	16 (terminologická) ontologie	<i>(terminological) ontology</i>	G	F + odvozování
		15 sémantická síť	<i>semantic network</i>	F	E + asociace
		14 tezaurus	<i>thesaurus</i>		
klasifikace stromová struktura	organizace pojmnů	13 klasifikační schéma / systém	<i>classification scheme</i>	E	D + hierarchie
		12 taxonomie (biologie)	<i>taxonomy</i>		
		11 kategorizační schéma / systém	<i>categorization scheme</i>		
		10 předmětový heslář	<i>subject heading scheme</i>		
slovníky lineární struktura	organizace termínů	9 seznam jmenných autorit	<i>name authority list/file</i>	D	C + preferované termíny
		8 slovník zeměpisných názvů	<i>gazetteer</i>		
		7 seznam synonym	<i>synonym ring, synset</i>	C	B + ekvivalence
		6 terminologický slovník	<i>terminology</i>	B	A + definice
		5 (výkladový) slovník, glosář	<i>glossary</i>		
		4 číselník (nomenklatura, seznam, kódovník)	<i>list, code list</i>		
		3 (řízený) slovník	<i>(controlled) vocabulary, dictionary</i>	A	seznam slov
		2 negativní slovník (stop slova, zakázaná slova)	<i>stopwords, exclusion list</i>		
		1 nabídkový seznam (menu)	<i>pick list</i>		



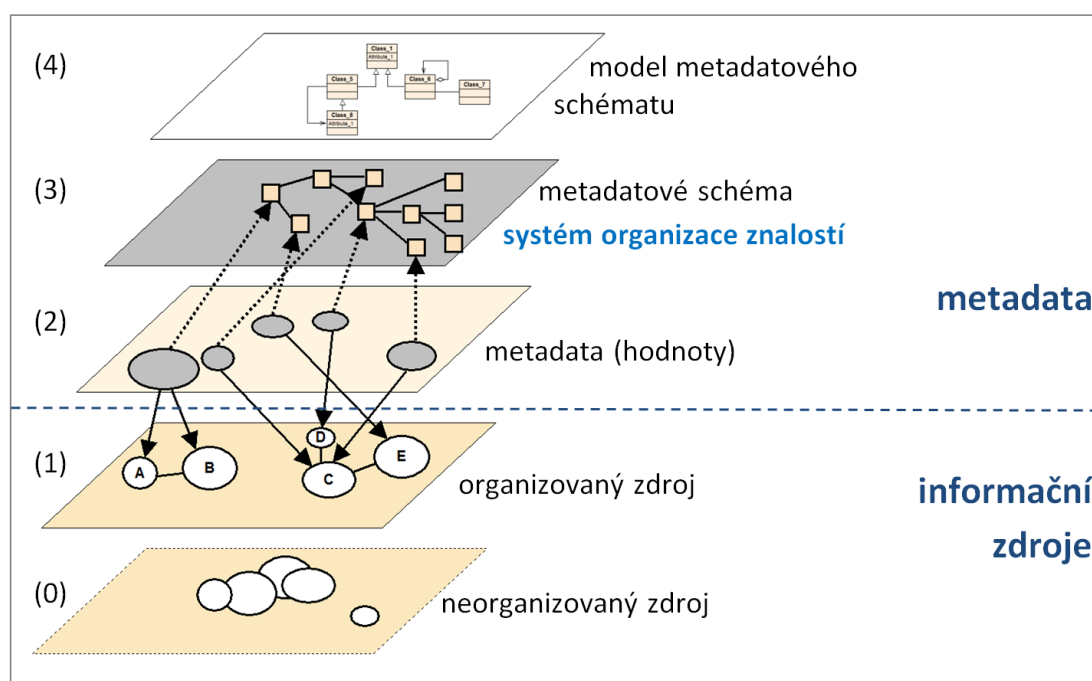
Obr. 18 Typologie systémů organizace znalostí

Již bylo konstatováno, že pojem SOZ zatím postrádá obecně akceptovanou intenzionální definici. Z toho důvodu provedeme vlastní analýzu a s využitím jejích výsledků zformulujeme pracovní definici SOZ. V naší analýze uplatníme dva komplementární přístupy k systému organizace znalostí, založené na systémovém přístupu: 1) SOZ jako model metadat a 2) SOZ jako řízený slovník.

<sup>351</sup> Tamtéž, s. 5.

## System organizace znalostí jako model metadat

Na obrázku 19<sup>352</sup> je schematicky znázorněn systém organizace znalostí ve funkci modelu metadat. I v tomto případě je uplatněn systémový přístup, přičemž pozornost je soustředěna na charakteristiku funkce SOZ, tj. na jeho úlohu ve funkci nástroje procesu organizace znalostí. Tento pohled zároveň odkazuje na metaforu systému, jež je použita v označení tohoto nástroje. Jednotlivé vrstvy na obrázku 19 oddělují významné subsystemy organizovaného zdroje znalostí a jeho kontextu.



Obr. 19 Systém organizace znalostí jako model metadat

V nejnižší vrstvě (0) se nachází neuspořádaná množina zdrojů, představující neorganizované zaznamenané znalosti. I přístup k neorganizovaným zdrojům je technicky uskutečnitelný. Pokud by například byly k dispozici ve formě digitalizovaného textu, třeba po zpracování technologií *OCR*, bylo by možné provést tzv. plnotextové vyhledávání. Takovéto vyhledávání, jakkoli rychlé, může ale zajistit přístup jen na základě porovnání slov (termínů) v dotazu a ve zdroji (angl. *term matching*).

Ve vrstvě (1) je znázorněna uspořádaná množina zdrojů (typicky zachycených v dokumentech či v záznamech). Základem tohoto typu uspořádání je struktura, která vznikla uplatněním techniky umístění – lokací (identifikace umístěním) či kolokací (kategorizace umístěním). Přístup k relevantním objektům je možný navigací v organizovaném prostoru.

Nad vrstvami (0) a (1), znázorňujícími neorganizované a organizované informační zdroje, jsou metadatové vrstvy. První z nich je vrstva (2), která představuje neuspořádanou množinu metadat. Metadata, jež jsou typickým produktem procesu organizace znalostí (např. indexace), jsou výsledkem uplatnění techniky označení. Vrstvy (1) a (2) je možné považovat za SOZ v širším

<sup>352</sup> Grafické řešení obrázku je inspirováno grafikou použitou v článkách Steve Peppera o mapách námětů (např. PEPPER, Steve. *The TAO of topic maps: finding the way in the age of infoglut* [online]. Oslo: Ontopia, April 2000 [cit. 2018-01-30]. Dostupné z: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.32.5473>).

pojetí, tak jak byl charakterizován v úvodu této kapitoly.

Vrstva (3) představuje SOZ v užším pojetí, uplatňovaném v rámci domény organizace znalostí. Znázorňuje uspořádanou množinu metadat, jež je oproti neustále se dynamicky proměňující množině metadat ve vrstvě (2) relativně trvalá. Její struktura vznikla uplatněním techniky zjednoznačnění a začlenění do kontextu. Jednoznačného významu metadat je dosaženo jejich definicemi. Začlenění do kontextu se děje vyjádřením vztahů mezi obsahy metadat. Výsledná struktura je využitelná jako referenční množina, díky níž je možné reprezentovat nejen jednotlivé komponenty organizovaných znalostí, ale i vztahy mezi nimi, a to jak v rámci jednoho zdroje, tak i mezi nimi navzájem. Na tuto referenční množinu se pak odkazují informační zdroje prostřednictvím metadat, vytvořených při analýze jejich obsahu. Díky její relativně stabilní sémantické struktuře je možné využít pro přístup k organizovaným zdrojům nejen hodnoty metadat, ale i vztahy mezi metadatovými prvky a jejich případné vlastnosti, definované v tzv. metadatovém schématu<sup>353</sup>. Až na této úrovni je možné uplatnit přístup ke zdrojům na základě porovnávání pojmů (angl. *concept matching*). SOZ je tedy nástrojem sémantického přístupu ke znalostem obsaženým ve zdrojích.

Vrstva (4) je schematickým znázorněním struktury a funkce SOZ (např. klasifikace, tezauru, seznamu autorit). Jak vrstvu (3), tak vrstvu (4) lze označit jako pojmový model či pojmové schéma. Vrstva (4) se odlišuje tím, že představuje metamodel – ‚model modelu‘, ‚strukturu struktury‘ metadat. Pojmové modely systémů organizace znalostí (například níže uvedený datový model *SKOS* nebo datový model tezauru v normě ISO 25964<sup>354</sup>) zobecňují jejich strukturu a umožňují tak teoretický výzkum a zejména implementaci systémů organizace znalostí do současné informační infrastruktury síťového prostředí.

### **System organizace znalostí jako řízený slovník**

V této části se soustředíme na charakteristiku struktury SOZ, již tvoří pojmy, jejich označení a vzájemné vztahy. Tento pohled se zaměřuje na ty rysy SOZ, jež jsou obvykle vyjadřovány metaforou jazyka (například v anglickém termínu *indexing language* nebo v českém termínu selekční jazyk, případně v označení řízený nebo strukturovaný slovník).

K popisu klíčových prvků a vlastností SOZ využijeme vybrané prvky datového slovníku *SKOS* (*Simple knowledge organization system*). *SKOS* je určený pro takový popis systémů organizace znalostí, který umožňuje jejich sdílení a propojování na webu. Jeho ambicí je definovat minimální množinu prvků, společných všem typům SOZ v celém rozsahu jejich škály, jež byla charakterizována výše. V roce 2009 byl *SKOS* přijat jako standard konsorcia *W3C*<sup>355</sup> a v současné době představuje nejvýznamnější a v praxi nejpoužívanější obecný metamodel systému organizace znalostí. Má charakter ontologie, formalizované v jazyce *OWL*, a jako takový je tvořen třídami a jejich vlastnostmi (predikáty).

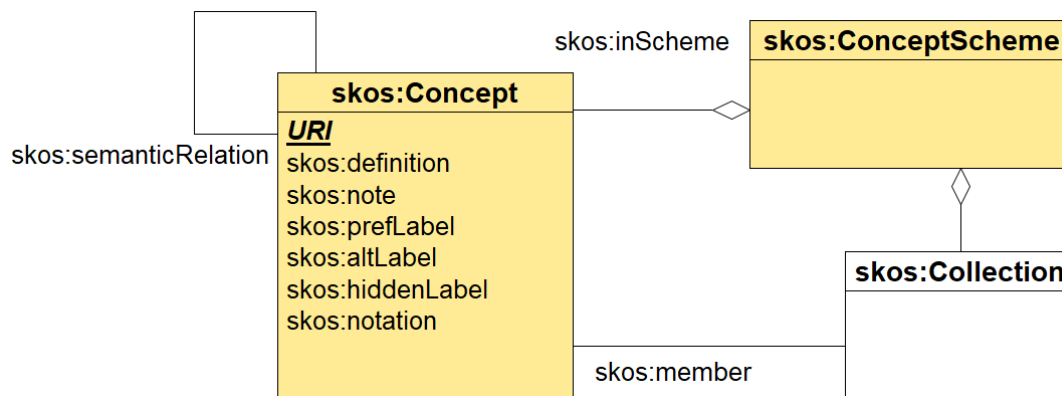
---

<sup>353</sup> Poznámka: Používané pojmenování ‚metadatové schéma‘ navozuje představu, že se bude jednat o schéma nějaké složité struktury. Ve skutečnosti může jít i o zcela jednoduchou lineární strukturu abecedně uspořádaných termínů z nějakého řízeného slovníku.

<sup>354</sup> Dostupné z [http://www.niso.org/schemas/iso25964/Model\\_2011-06-02.jpg](http://www.niso.org/schemas/iso25964/Model_2011-06-02.jpg) [cit. 2018-01-30].

<sup>355</sup> MILES, Alistair, BECHHOFFER, Sean, ed. *SKOS Simple Knowledge Organization System Reference* [online]. W3C Recommendation 18 August 2009 [version]. Cambridge (MA): World-Wide Web Consortium, © 2009 [cit. 2018-01-30]. Dostupné z: <http://www.w3.org/TR/2009/REC-skos-reference-20090818/>.





Obr. 20 Struktura systému organizace znalostí podle SKOS

Na obrázku 20 jsou schematicky znázorněny vybrané prvky SKOS a jejich vzájemné vztahy ve formě diagramu tříd UML. Základní třídou SKOS je pojem (skos:Concept), který je identifikován svým URI. Pojmy lze agregovat do pojmových schémat, tj. do systémů organizace znalostí (třída skos:ConceptScheme, používá se predikát skos:inScheme), například do tezaurů nebo klasifikačních schémat. Další možností je seskupování pojmů v rámci daného SOZ do kolekcí (třída skos:Collection, používá se predikát skos:member), například do faset nebo mikrotezaurů.

Vlastnosti (predikáty) SKOS umožňují formálně vyjádřit pojmy a jejich vzájemné vztahy. Sémantiku pojmů pro lidské uživatele pomáhají upřesnit a definovat různé typy poznámek (skos:definition, skos:note ad.).

Skupina vlastností pro označení pojmů umožňuje každý pojem označit jazykovými výrazy (skos:prefLabel – preferované označení, skos:altLabel – alternativní označení, např. synonyma, skos:hiddenLabel – nezobrazované označení) nebo notací (skos:notation).

Nejdůležitějšími vlastnostmi jsou ty, jež umožňují vyjádřit sémantické vztahy mezi pojmy (skos:semanticRelation). Člení se na dvě vzájemně propojené skupiny: vztahy pojmů v rámci jednoho SOZ a vztahy pojmů z různých SOZ, nazývané mapování (skos:mappingRelation).

Vztahy pojmů v kontextu jednoho SOZ jsou asociace (skos:related), obecná hierarchie (skos:broader – širší pojem, skos:narrower – užší pojem) a hierarchie, umožňující tranzitivitu (skos:broaderTransitive, skos:narrowerTransitive).

Při mapování pojmů z různých SOZ je možné vyjádřit různé stupně ekvivalence: skos:exactMatch, skos:closeMatch. Navíc lze mezi pojmy z různých SOZ definovat asociativní (skos:relatedMatch) a hierarchické vztahy (skos:broadMatch, skos:narrowMatch), jež jsou významově ekvivalentní analogickým vztahům v rámci jednoho SOZ. Jediným vztahem mapování, který umožňuje tranzitivitu, je skos:exactMatch.

## Pracovní definice systému organizace znalostí

S využitím závěrů provedené analýzy byla zformulována následující pracovní definice<sup>356</sup>:

Systém organizace znalostí je schéma, modelující strukturu (tj. prvky a vzájemné vztahy) organizované množiny zaznamenaných znalostí. Funkcí systému organizace znalostí je podpora procesů organizace znalostí a přístupu ke znalostem. Základním strukturálním prvkem systému organizace znalostí je pojem. Jádrem fyzické reprezentace každého systému organizace znalostí je slovník, tj. formální vyjádření pojmů. Ten je používán jak pro vyjádření sémantiky, tak syntaxe organizovaného celku, případně i pravidel, určujících používání struktury.

## Vztah SOZ a informatických ontologií

Porovnáme-li pojmový obsah pracovní definice SOZ a pracovní definice informatické ontologie, zformulované v kapitole 1.6.3, jsou patrné významné analogie a v mnoha případech i shoda pojmů, na nichž jsou tyto definice založeny. V obou případech se jedná o specifický typ pojmového modelu, základním strukturálním prvkem SOZ i informatických ontologií je pojem. Lze tedy konstatovat, že oba pojmy mají společný nejbližší nadřazený rod (*genus proximum*), jímž je pojmový model. Shodu lze najít i v rozlišujících druhových rozdílech (*differentia specifica*), které oba tyto specifické typy odlišují od jiných typů pojmových modelů. Jak SOZ, tak informatické ontologie jsou inženýrské artefakty, jejichž funkcí je podpora procesů organizace znalostí, oba jsou opakovaně použitelné a sdílené rozsáhlejším okruhem uživatelů.

Vzájemný vztah systémů organizace znalostí a ontologií je vnímán různými způsoby.<sup>357</sup> Výše uvedené společné rysy SOZ a informatických ontologií vedou některé autory zejména z oblasti informatiky k tomu, že používají označení ‚ontologie‘ souhrnně jak pro informatické ontologie, tak pro jednotlivé typy SOZ. Takové pojetí tedy chápe SOZ jako podmnožinu informatických ontologií. V naší analýze se pokusíme předložit alternativní pohled, v jehož rámci nahlédneme SOZ a informatické ontologie jako dvě sice rozdílné, ale částečně se překrývající množiny.

Při porovnání vyjdeme z typologie informatických ontologií, představené v kapitole 1.4, a z typologie SOZ zformulované v této kapitole. Konkrétně využijeme členění ontologií na lehké (terminologické) a těžké (znalostní) a členění SOZ na slovníky, klasifikace a pojmové sítě. V obou případech lze konstatovat, že ve skutečnosti se nejedná o zřetelně ohraničené a oddělené skupiny, ale spíše o orientační body v rámci plynulého spektra, v němž jsou jednotlivé typy řazeny podle své sémantické síly.

S využitím této typologie lze uvažovat tři stupně sémantické síly, k nimž lze jednotlivé typy přiřadit:

V první skupině jsou SOZ s nejnižší sémantickou silou, v nichž převažuje důraz na funkce slovníku či terminologie. Tato skupina zahrnuje SOZ ze skupiny označené jako slovníky (tj. typy 1–9 v tabulce 18).

---

<sup>356</sup> Srovnání s dalšími definicemi lze nalézt v: MAZZOCCHI, Fulvio. Knowledge organization system (KOS) [online]. Version 1.1. 2017-07-13 [cit. 2018-01-30]. In: *ISKO Encyclopedia of knowledge organization*. ISKO, 2017. Dostupné z: <http://www.isko.org/cyclo/kos>.

<sup>357</sup> Viz např. úvodník k monotematickému číslu časopisu *Applied ontology*, věnovanému vztahu ontologií a terminologických systémů: GRABAR, Natalia, HAMON, Thierry, BODENREIDER, Olivier. Ontologies and terminologies: continuum or dichotomy? In: *Applied ontology*. 2012, 7(4), 375-386. doi:10.3233/AO-2012-0119. ISSN 1570-5838 (Print). ISSN 1875-8533 (Online).

Druhá skupina se střední sémantickou silou zahrnuje SOZ orientované na pojmy – klasifikace a pojmové sítě (tj. typy 10–16 v tabulce 18). Do této skupiny řadíme rovněž lehké (terminologické) ontologie.

Třetí skupina zahrnuje těžké (znalostní) ontologie s nejvyšší sémantickou silou. Jak bylo uvedeno v kapitole 1.4, tento typ ontologií disponuje kromě uvedených společných rysů se SOZ ještě specifickými vlastnostmi. Ontologické pojmy (třídy) mají explicitně definované vlastnosti a axiomaticky definovanou sémantiku. Funkce těžkých ontologií jsou oproti lehkým ontologiím rozšířeny o komunikaci, opakované využití znalostí, popis skutečnosti pro počítačové zpracování a především automatické odvozování.

Členění na tři výše uvedené skupiny demonstruje naše chápání vztahu SOZ a informatických ontologií jako průniku dvou množin: první skupina zahrnuje výhradně SOZ, druhá skupina představuje průnik množin SOZ a informatických ontologií a ve třetí skupině jsou opět výhradně informatické ontologie.

## 4.6 Kognitivní kontext procesu organizace znalostí

### 4.6.1 Vědní disciplíny a aplikační oblasti

Organizace znalostí je proces, jehož historie je stejně stará jako historie lidského poznávání světa. Nejvýznamnějšími oblastmi, jež se organizaci znalostí věnovaly v minulosti a jež dodnes významně ovlivňují její metody a techniky, jsou encyklopedistika, klasifikace věd a vědecké klasifikace, techniky duševní práce, knihovnické a bibliografické pořádací systémy, nejnověji též technologie zpracování a komunikace dat v počítačích a počítačových sítích.

V tabulce 12 jsou disciplíny tvořící kognitivní kontext organizace znalostí seskupeny do několika okruhů. V rámci každého okruhu jsou názvy jednotlivých disciplín řazeny abecedně.

filozofie	speciální vědy		
	teoretické		aplikované / inženýrské obory
	jednooborové	interdisciplinární	
filozofie vědy	lingvistika	informační věda	informační architektura
klasifikace věd	logika	kognitivní věda	informační/znalostní inženýrství
teorie poznání	psychologie	počítačová věda	informační/znalostní management
věda o vědě	sémiotika	systemová věda	metody a techniky duševní práce
	sémantika	umělá inteligence	ontologické inženýrství
			organizace znalostí
			organizační inženýrství
			podniková informatika
			softwarové inženýrství
			systemové inženýrství

Tab. 12 Kognitivní kontext procesu organizace znalostí

Nejobecnější okruh představují filozofické disciplíny. Jejich předností je maximální úroveň zobecnění prezentovaných poznatků, jež ovšem právě z tohoto důvodu nebývají bezprostředně aplikovatelné v praxi. Richard P. Smiraglia nicméně konstatuje, že vědci v oboru organizace znalostí převzali četné filozofické přístupy, jež použili zajímavým a často inovativním

způsobem.<sup>358</sup> Pro organizaci znalostí jsou relevantní filozofie vědy, věda o vědě a klasifikace věd, resp. mapování vědy. Nejvýznamnější filozofickou disciplínou je bezesporu teorie poznání (gnozeologie, epistemologie, noetika), již mnozí autoři považují za teoretickou bázi organizace znalostí.

Okruh speciálních věd je možné rozčlenit na teoretické a aplikované vědy. Historicky starší jednooborové teoretické vědy, významné pro organizaci znalostí, jsou především obory, jež se nějakým způsobem dotýkají obsahu organizovaných jednotek a možností jeho vyjádření: sémiotika, sémantika, lingvistika, logika, psychologie.

Do skupiny interdisciplinárních teoretických věd, jež se začaly formovat v 20. století, řadíme informační vědu, která se zaměřuje na zkoumání komunikace informací ve společnosti a jejíž aplikační doménou jsou paměťové a fondové instituce, zejména knihovny. Typickou interdisciplinární vědou je kognitivní věda, která zkoumá procesy a struktury organizace znalostí v mysli, přesně řečeno v mozku. Ivan M. Havel ji charakterizuje jako obor na průniku psychologie, neurovědy, biologie, kybernetiky, informatiky, lingvistiky a filozofie.<sup>359</sup> Základem je výzkum lidské psychiky, zejména pak krátkodobé a dlouhodobé paměti, aplikační oblasti jsou kupříkladu metodiky a techniky učení a duševní práce. Dalšími významnými interdisciplinárními obory, jejichž teoretické principy využívá organizace znalostí, jsou systémová věda a umělá inteligence. Technologický základ organizace znalostí současnosti je založen na principech a teoriích počítačové vědy (informatiky).

Nejvýznamnější součástí kognitivního kontextu procesu organizace znalostí je v současné době samozřejmě obor organizace znalostí. Ve smyslu vědní disciplíny organizaci znalostí jako první anticipovali editoři časopisu *International classification* v redakčním úvodníku prvního čísla v roce 1974. Jako první ze tří bodů programu zakládaného časopisu je uveden návrh na mezinárodní úrovni „*upozornit na existenci nového, autonomního vědního oboru*“. Obor, zatím bez názvu, je charakterizován prostřednictvím svého účelu, předmětu a metod: „*jeho účelem je zavést pořádek, jeho předmětem jsou pojmy a pojmové systémy a za jeho metody lze považovat tvorbu pojmových systémů a přiřazování pojmů z takových systémů prvkům reality a naopak*.“<sup>360</sup>

Vymezení vztahu organizace znalostí k ostatním vědním disciplínám je předmětem diskusí. Dle názoru Richarda P. Smiraglii obor zatím postrádá svoji specifickou teorii<sup>361</sup>, je tedy problematické jednoznačně stanovit jeho vztahy k již etablovaným, ‚zralým‘ disciplínám. Nabízejí se možnosti považovat organizaci znalostí za subdisciplínu informační vědy, případně chápat ji jako samostatný obor, který má s informační vědou významný průnik co do předmětu a metod zkoumání. Ingetraut Dahlbergová ovšem zastává alternativní názor a považuje organizaci znalostí za součást vědy o vědě.<sup>362</sup> Hranice organizace znalostí vymezuje v souladu s obsahovým

---

<sup>358</sup> SMIRAGLIA, Richard P. *The elements of knowledge organization*. Springer, 2014, s. 30. ISBN 978-3-319-09356-7. ISBN 978-3-319-09357-4 (eBook). doi:10.1007/3-319-09357-4.

<sup>359</sup> HAVEL, Ivan M. Věda o duši. In: *Vesmír*. 2000, 79(7), 363-364. ISSN 0042-4544.

<sup>360</sup> Editorial. Why this journal? In: *International classification: Journal on theory and practice of universal and special classification systems and thesauri*. 1974, 1(1), 1-2.

<sup>361</sup> SMIRAGLIA, Richard P. *The elements of knowledge organization*. Springer, 2014, s. 2. ISBN 978-3-319-09356-7. ISBN 978-3-319-09357-4 (eBook). doi:10.1007/3-319-09357-4.

<sup>362</sup> V klasifikaci *Information coding classification* (ICC), jejíž autorkou je I. Dahlbergová, je třída organizace znalostí (814 *Wissensorganisation/knowledge organization*) součástí třídy věda o vědě (81 *Wissenschaftswissenschaft/science of science*). Viz DAHLBERG, Ingetraut. *Wissensorganisation: Entwicklung*,

zaměřením časopisu *Knowledge organization*: teorie pojmů, systematická terminologie, organizace znalostí.

Přístup Ingetraut Dahlbergové k organizaci znalostí jako k teoretické vědní disciplíně, jakkoli důkladně věcně argumentovaný, je spíše menšinovým názorem. Většina autorů je toho mínění, že organizace znalostí patří mezi inženýrské disciplíny a řídí se pravidly takových oborů. Za inženýrství (angl. *engineering*) je obecně považována systematická aplikace vědeckých znalostí při návrhu a tvorbě efektivních řešení praktických problémů. Uplatňuje se při tvorbě rozsáhlých, složitých a technologicky náročných artefaktů, tj. objektů vytvořených člověkem. Typická teoretická základna inženýrské disciplíny je převážně eklektická, převzatá z jiných oborů. Jejím cílem není další rozvoj příslušných teoretických oborů, ale využití jejich poznatků při řešení reálných problémů. V inženýrství se často uplatňují metody projektového managementu: vychází se z předem určeného cíle, existujících omezení a z požadované spolehlivosti. Inženýrské řešení respektuje existující standardy a je dokumentováno. Na rozdíl od ‚čisté vědy‘ má výrazný etický rozměr.

Aktuálním příspěvkem v tomto směru je publikace *The discipline of organizing* (Obor či disciplína organizování, organizační věda), zpracovaná kolektivem autorů pod vedením Roberta J. Glushka.<sup>363</sup> Je příkladem extrémně širokého pojetí organizace znalostí, vedoucího dokonce až k postulování nového oboru ‚organizace všeho‘. Glushko zavádí pojem organizační systém (angl. *organizing system*), který definuje jako jakoukoli záměrně uspořádanou kolekci zdrojů<sup>364</sup>, čímž rozšiřuje předmět zájmu oboru na organizaci jakýchkoli zdrojů použitelných k nějakému účelu, tj. nejen znalostí. Poukazuje na multioborový charakter disciplíny, syntetizující poznatky z knihovni a informační vědy, informačního managementu, podnikové informatiky, kognitivní vědy, systémové analýzy, softwarového a znalostního inženýrství a informatiky. Zdůrazňuje záměrnost a cílevědomost organizačních aktivit (vylučuje tedy samoorganizační systémy) a podtrhuje význam interakce aktérů s organizačním systémem.

Pojetí organizace znalostí jako inženýrské disciplíny je zjevně blízké i Patricku Lambeovi, který vyjadřuje explicitně názor, že v průběhu organizace znalostí mají být vytvářeny nejen deskriptivní modely popisující to, co je, ale především modely ‚to-be‘, plánující a přinášející inovace.<sup>365</sup>

Zařazením do skupiny inženýrských věd se organizace znalostí ocitá v sousedství četných disciplín, jež mají často obdobný předmět zájmu a řeší ho nástroji přizpůsobenými pro specifický okruh problémů či pro danou uživatelskou komunitu. Předmětem zájmu informačního a znalostního managementu je organizace informací, resp. explicitně vyjádřených znalostí, s aplikací ve sféře institucionální a podnikové ekonomiky. Širší kontext zahrnuje organizační inženýrství a podniková informatika. Softwarové a systémové inženýrství, informační a znalostní inženýrství a nejnověji ontologické inženýrství řeší organizaci dat

---

*Aufgabe, Anwendung, Zukunft*. Würzburg: Ergon Verlag, 2014, s. 96. Textbooks for knowledge organization, vol. 3. ISSN 0944-8152. ISBN 978-3-95650-065-7.

<sup>363</sup> GLUSHKO, Robert J., ed. *The discipline of organizing*. Cambridge(MA): MIT Press, 2013. 475 s. ISBN 978-0-262-51850-5 (brož.).

<sup>364</sup> Tamtéž, s. 1-2.

<sup>365</sup> LAMBE, Patrick. From cataloguers to designers: Paul Otlet, social impact and a more proactive role for knowledge organization professionals. In: *Knowledge organization*. 2015, 42(6), 445-455. ISSN 0943-7444.

v počítačových systémech. Na specifické problémy organizace webového obsahu se soustředí informační architektura. Aplikační oblast těchto ‚počítačových‘ disciplín není definována institucionálně, ale prostřednictvím produktů: databázové a znalostní systémy, navrhované s využitím principů těchto oborů, jsou používány ve všech oblastech společenské praxe.

Formující se teoretickou platformu organizace znalostí tvoří kromě převzatých teoretických principů výše uvedených vědních disciplín empirické principy a zobecněné příklady nejlepší praxe z oblastí paměťových a fondových institucí, podnikové informatiky, informatiky veřejné správy a vědecké informatiky. Významnou roli hrají poznatky shromážděné v prostředí knihoven a zformulované v 19. a 20. století. K nejvýznamnějším patří *Pravidla slovníkového katalogu* Charlese Ammi Cuttera<sup>366</sup>, *Kánony klasifikace* W. C. B. Sayerse<sup>367</sup>, principy klasifikace pro knihovny Henry E. Blisse<sup>368</sup> a zákony, kánony, principy a postuláty Shiyali R. Ranganathana<sup>369</sup>. Příkladem principů formulovaných v oblasti umělé inteligence jsou kritéria návrhu ontologií Thomase Grubera (viz kapitola 2.1). Další inspiraci pro rozvoj oboru je možné najít v pracích autorů, kteří od 50. let 20. století formovali nové paradigma práce s informacemi s podporou počítačů (viz osobnosti uvedené v kapitole 4.5.1), a jejichž teoretické texty se v období překotné implementace digitálních a síťových technologií po nástupu *Internetu* očitly na periferii zájmu ve prospěch pragmatických řešení. V současné době lze konstatovat, že poté, co byly vyřešeny technické a syntaktické problémy digitální komunikace, je příležitost současně s postupem o úroveň výše k řešení sémantických problémů vrátit se znovu k těmto pozapomenutým teoretickým úvahám z období před *Internetem*. K významným autorům takových úvah patří zejména Jesse Hauk Shera (1903–1982), Bertram Claude Brookes (1910–1991), Patrick G. Wilson (1927–2003), Robert Fugmann (1927), Dagobert Soergel (1940) a členové britské *Skupiny pro výzkum klasifikace*.

Konstatování z kapitoly 4.2, že organizace znalostí se v současnosti nachází v etapě změny paradigmatu, vede k úvahám, jak tuto skutečnost interpretovat, zda na ni pohlížet jako na výhodu či nevýhodu. Vyjdeme-li z tezí Thomase S. Kuhna, mohla by to být přednost, protože každá změna paradigmatu poskytuje příležitost k inovacím. Navíc následně s ustalováním (dle Kuhna normalizací) zkoumané reality lze očekávat, že se ustálí i teoretická základna oboru. V případě oboru organizace znalostí, který zatím teorii srovnatelnou s historicky staršími obory nemá, se samozřejmě v průběhu změny paradigmatu neočekávají revoluční změny stávajících teorií, protože ty nejsou k dispozici, ale vytvoření nové plnohodnotné teorie a její normalizace.

---

<sup>366</sup> CUTTER, Charles Ammi. *Rules for a dictionary catalog*. 4th ed., rewritten. Washington: Government Printing Office, 1904. 173 s.

<sup>367</sup> SAYERS, W. C. Berwick. *Canons of classification applied to „the subject“, „the expansive“, „the decimal“, and „the Library of Congress“ classifications: a study in bibliographic classification method*. London: Grafton, 1915. 173 s.

<sup>368</sup> BLISS, Henry Evelyn. *A bibliographic classification extended by systematic auxiliary schedules for composite specification and notation*. New York: H. W. Wilson, 1940-1953. 4 sv.

<sup>369</sup> RANGANATHAN, Shiyali Ramamrita. *Prolegomena to library classification*. Assisted by M. A. GOPINATH. 3rd ed. London: Asia Publishing House, 1967. 640 s. Ranganathan series in library science, 20. ISBN 978-0-210-22731-2.

## 4.6.2 Terminologie organizace znalostí

Rozsáhlý realizační kontext (viz kapitola 4.5) je příčinou terminologické nejednotnosti v doméně organizace znalostí, protože s procesem organizace znalostí se lze setkat všude, kde se pracuje s informacemi a znalostmi. V současné době je nicméně alespoň v anglickém jazykovém prostředí možné pozorovat šířící se konsenzus v označování dvou klíčových pojmů – organizace znalostí a systém organizace znalostí. Další problém byl konstatován už během úvodního průzkumu předcházejícího řešení návrhu ontologie: nekompatibilita české a světové oborové terminologie, která brání mezinárodní komunikaci.

### Termín *knowledge organization*

Jazyková podoba termínu ‚organizace znalostí‘ napovídá, že se jedná o označení činnosti nebo procesu. Toto označení ovšem také slouží jako název vědní disciplíny (viz kapitola 4.6.1).<sup>370</sup> Anglický termín *knowledge organization* (organizace znalostí, angl. zkratka *KO*) v dnešním slova smyslu se podle Birgera Hjørlanda objevoval už počátkem 20. století v pracích zakladatelů moderního pojetí organizace znalostí, používali ho Charles A. Cutter, E. C. Richardson, W. C. B. Sayers.<sup>371</sup> V prvních dvou desetiletích se však jednalo spíše o ojedinělé případy použití, např. Sayers ve své práci *Kánony klasifikace* dává přednost termínům *knowledge classification / classification of knowledge* (klasifikace znalostí), případně *knowledge order / order of knowledge* (pořádání či soustava znalostí). Systematické používání sousloví *organization of knowledge* lze zaznamenat až od konce 20. let 20. století u Henryho E. Blisse<sup>372</sup>, v 50. letech 20. století ho začal používat jeden ze zakladatelů informační vědy Jesse Hauk Shera<sup>373</sup>. K rozšíření termínu *knowledge organization* mezi širokou odbornou veřejností došlo poměrně nedávno. Bezprostředním podnětem se stalo založení *Mezinárodní společnosti pro organizaci znalostí ISKO* (viz kapitola 4.5.1) v roce 1989. Podle Ingetraut Dahlbergové v té době v německé *Společnosti pro klasifikaci (Gesellschaft für Klassifikation)*, založené v roce 1977, začalo převažovat pojetí klasifikace ve striktně matematickém a statistickém pojetí. ‚Pojmově‘ či na ‚obsah‘ orientovaní členové se rozhodli založit novou společnost, která by se i názvem odlišila od matematicky orientované *Společnosti pro klasifikaci*. Zvolili název, který odkazoval právě na termíny *organization of knowledge* v názvech děl Henryho E. Blisse. Místo sousloví *organization of knowledge* však použili kratší variantu v podobě doslovného překladu německého *Wissensorganisation*, tj. *knowledge organization* (mezi devíti zakladateli *ISKO* bylo sedm německých odborníků).<sup>374</sup> V roce 1993 pak následovalo přejmenování stěžejního oborového časopisu *International classification* na *Knowledge organization*. I v relativně konzervativním prostředí vysokoškolských studijních oborů dochází k postupnému nahrazování dosavadních názvů obsahově relevantních kurzů termínem organizace znalostí (viz např. kurz *Knowledge*

---

<sup>370</sup> Anglický termín *knowledge organization* může ještě navíc označovat tzv. znalostní organizaci, tuto denotaci ovšem český překlad postrádá.

<sup>371</sup> HJØRLAND, Birger. What is knowledge organization (KO)? In: *Knowledge organization*. 2008, **35**(2-3), s. 97. ISSN 0943-7444.

<sup>372</sup> BLISS, Henry Evelyn. *The organization of knowledge and the system of the sciences*. With an introduction by John Dewey. New York: Henry Holt, [© 1929]. xx, 433 s. – BLISS, Henry Evelyn. *The organization of knowledge in libraries and the subject-approach to books*. New York: H. W. Wilson, 1933. xvi, 335 s.

<sup>373</sup> SHERA, Jesse Hauk. *Documentation and the organization of knowledge*. 1st ed. London, 1966. 185 s.

<sup>374</sup> DAHLBERG, Ingetraut. International Society for Knowledge Organization (ISKO). In: Marcia J. BATES, Mary Niles MAACK, ed. *Encyclopedia of library and information sciences*. 3rd ed. Boca Raton (Florida): CRC Press, © 2010, s. 2941.

organization na Torontské univerzitě<sup>375</sup>).

### **Termín *knowledge organization system***

Pro systémy organizace znalostí, charakterizované v kapitole 4.5.3 jako nástroje procesu organizace znalostí ve formě metadatového schématu, lze v literatuře najít nejen různé přístupy k vymezení, ale i k pojmenování. Renato Souza et al. pro ně používají obecné označení „artefakty reprezentující znalosti“ (angl. *knowledge representation artifacts*)<sup>376</sup>, Vanda Broughtonová et al. je nazývají „sémantické nástroje“ (angl. *semantic tools*)<sup>377</sup>. Časté je nahrazování zobecňujícího pojmenování metonymickým využitím názvu některého z konkrétních typů SOZ – například slovníky, tezaury, klasifikační schémata, taxonomie a dokonce i ontologie. Termín *knowledge organization system* (systém organizace znalostí, angl. zkratka KOS) se začal široce používat až v 90. letech 20. století. Podle Gail Hodgeové byl vytvořen na ustavujícím setkání *Pracovní skupiny pro síťově propojené systémy organizace znalostí NKOS (Networked Knowledge Organization Systems Working Group)*, pořádaném 27. 6. 1998 v rámci konference *ACM Digital Libraries '98* v Pittsburghu<sup>378</sup>. Termín využili a v komunitě sémantického webu zpopularizovali tvůrci datového slovníku *SKOS (Simple knowledge organization system, doslova ‚jednoduchý systém organizace znalostí‘)*.

### **Česká terminologie**

Reprezentací současného tuzemského pojetí terminologie organizace znalostí jsou termíny a jejich výklady v *České terminologické databázi knihovní a informační vědy (TDKIV)*. Jejich vymezení je založeno na pojetí Blahoslava Kováře, zformovaném v 70. letech 20. století. Pro klíčové pojmy v doméně – (proces) organizace znalostí a systém organizace znalostí se v českém prostředí používají tradičně termíny pořádání informací a selekční jazyk. Jak procesy, tak jazyky se člení na identifikační a věcné, jež se ještě dále člení na systematické a předmětové.<sup>379</sup>

Podrobně je problematika české terminologie analyzována v samostatné studii, v níž byl předložen návrh na nahrazení stávajících českých termínů ‚pořádání informací‘ a ‚selekční jazyk‘ doslovnými překlady anglických termínů *knowledge organization* a *knowledge organization system*, tedy termíny ‚organizace znalostí‘ a ‚systém organizace znalostí‘.<sup>380</sup> Návrhy na aktualizaci

---

<sup>375</sup> *Knowledge organization*. Toronto: Faculty of Information (iSchool) University of Toronto, 2018 [cit. 30-01-2018]. Dostupné z: <https://ischool.utoronto.ca/course/knowledge-organization/>.

<sup>376</sup> SOUZA, Renato Rocha, TUDHOPE, Douglas, ALMEIDA, Maurício Barcellos. Towards a taxonomy of KOS: dimensions for classifying knowledge organization systems. In: *Knowledge organization*. 2012, **39**(3), s. 180. ISSN 0943-7444.

<sup>377</sup> BROUGHTON, Vanda, HANSSON, Joacim, HJØRLAND, Birger, LÓPEZ-HUERTAS PÉREZ, María José. Knowledge organization. In: *European curriculum reflections on library and information science education*. Leif KAJBERG, Leif LØRRING, ed. Copenhagen: Royal School of Library and Information Science, 2005, [chap.] 7, s. 147. ISBN 87-7415-292-0.

<sup>378</sup> HODGE, Gail. *Systems of knowledge organization for digital libraries: beyond traditional authority files*. Washington: The Digital Library Federation, Council on Library and Information Resources, © 2000, s. 3. ISBN 978-1-933645-06-3.

<sup>379</sup> KOVÁŘ, Blahoslav. České termíny pro pojmenování procesu věcného pořádání a jeho druhů. In: *Knihovnictví a bibliografie*. 1973, **2**(2), 37-44; **2**(3), 55-59. Příl. čas. Čtenář, 1973, **24**(2); **24**(5).

<sup>380</sup> KUČEROVÁ, Helena. České termíny pro věcné pořádání informací po 40 letech: příspěvek k terminologické diskusi. In: *ProInflow: časopis pro informační vědy* [online]. 2013, **5**(Speciál), [cit. 2018-01-30], s. 1-19. ISSN 1804-2406. Dostupné z: <http://www.phil.muni.cz/journals/index.php/proinflow/article/view/749/>.



české terminologie byly předloženy i na redakční radě *TDKIV* a byly projednány se zúčastněnými odborníky na workshopu, uspořádaném v roce 2014<sup>381</sup>. Předmětem diskuse na workshopu byly termíny, uvedené v tabulce 13, a typologie systémů organizace znalostí (viz kapitola 4.5.3).

anglicky	česky		bez českého / anglického ekvivalentu
	ekvivalence významu	částečné překrývání významu	
<b>knowledge organization</b>		pořádání informací	
indexing	indexace / indexování		
subject indexing	věcné pořádání		
			systematické pořádání
			předmětové pořádání
descriptive indexing (cataloging)	identifikační pořádání		
indexing language		selektivní jazyk	
			<b>knowledge organization system</b>
subject language	věcný selektivní jazyk		
descriptive language	identifikační selektivní jazyk		
			systematický selektivní jazyk
			předmětový selektivní jazyk

Tab. 13 Srovnání české a anglické terminologie organizace znalostí

Diskuse účastníků workshopu se zaměřila na hledání odpovědi na otázku: ‚S kterými českými termíny propojit anglické termíny *knowledge organization* a *knowledge organization system* při jejich převodu do českého jazykového prostředí?‘ V průběhu diskuse byly vyjádřeny dva navzájem neslučitelné názory na postup při převádění anglické terminologie do češtiny: snaha o co nejpřesnější, ‚doslovný‘ překlad, co nejbližší původnímu jazyku, a snaha najít pro nové termíny co neadekvátnější český ekvivalent, který bude v souladu se specifiky českého jazyka a jeho dosavadní terminologie. Pro termín *knowledge organization* dala většina účastníků workshopu přednost českému ekvivalentu ‚pořádání informací‘. Analogicky k této preferenci navrhla většina přítomných pro termín *knowledge organization system* český ekvivalent ‚systém pořádání informací‘, případně ‚pořádací systém‘.

Na závěr lze konstatovat, že diskuse k české terminologii k datu psaní této práce není uzavřena. V souladu se zásadami terminologické práce je možné předpokládat, že k postupnému řešení kompatibility české a světové terminologie organizace znalostí v budoucnu přispěje větší množství publikovaných textů v češtině, tematicky zpracovávajících aktuální problematiku domény.

#### 4.7 Model domény organizace znalostí

Výsledkem doménové analýzy je model domény organizace znalostí, který se stal zdrojem pro návrh druhé verze ontologie. Model byl vytvořen v několika stupních formalizace: textový model domény (viz kapitola 4.7.1), semiformalizovaný model ve formě výroků a fasetové klasifikace

<sup>381</sup> Informace o workshopu jsou dostupné v Centrální evidenci projektů (CEP) Informačního systému výzkumu, experimentálního vývoje a inovací z: <https://www.rvvi.cz/riv?s=jednoduche-vyhledavani&ss=detail&n=0&h=RIV%2F00216208%3A11210%2F14%3A10314683> [cit. 2018-01-30].

klíčových pojmů (viz kapitola 4.7.2), vizualizace pojmů a jejich vztahů v pojmové mapě (viz kapitola 4.7.3).

### 4.7.1 Textový model domény organizace znalostí

#### Cíl organizace znalostí

Obecným cílem organizace je nastolit pořádek místo chaosu (snížit entropii) prostřednictvím zavedení struktury do organizovaného zdroje. Organizace znalostí specifikuje tento obecný účel do dvou cílů: přístup ke znalostem a využití znalostí. Přístup je primárním cílem z hlediska technologie, podmiňuje využití, což je významnější cíl z hlediska kognitivního. Vlastnosti organizovaných znalostí jsou nalezitelnost/vyhledatelnost, srozumitelnost, umožnění komunikace. Organizované znalosti lze využít k poznání, k řešení problémů a rozhodování a k tvorbě nových znalostí.

#### Hlavní témata organizace znalostí

##### a) Konstantní problémy organizace znalostí

Konstantní problémy řešené v rámci procesu organizace znalostí se člení na lingvistické, sémantické a pragmatické. Lingvistické problémy jsou synonymie, homonymie a slovní spojení. Nástrojem řešení lingvistických problémů jsou řízený slovník a kontext. Sémantické problémy jsou nejednotnost v chápání obsahu/potřeby znalosti, *aboutness*/obsah a složená témata. Pragmatické problémy jsou obtížné určování efektivnosti, proměnlivost obsahu/potřeby znalosti v čase, zkreslení / zaujetí (*bias*) a *warrant*.

##### b) Řešení konstantních problémů organizace znalostí – principy a techniky

Principem pořádku je ekvivalence a identifikace, jež jsou realizovány těmito procesy: seskupení podmínečně ekvivalentních objektů a odlišení objektů navzájem. K dispozici jsou dvě techniky umožňující seskupení a odlišení: umístění a označení. Jejich charakteristiky jsou symetricky protikladné a vzájemně se doplňují.

Technika umístění má dva typy – lokaci a kolokaci. Lokace umožňuje procesy přístupu, odkazu a odlišení, tj. identifikaci umístěním. Kolokace podporuje proces seskupování, tj. ekvivalenci umístěním (kategorizaci).

Technika označení má dva typy – označení objektu/pojmu, které je nejednoznačné a zjednoznačnění, jež umožňuje v rámci daného kontextu jednoznačnost, tj. identifikaci označením. Definice je výrok, který umožňuje identifikaci tím, že popisuje vlastnosti definovaných entit a určuje jejich kontext, tj. vztah k ostatním entitám.

Výsledek seskupení objektů, jež jsou považovány za ekvivalentní, se nazývá kategorie. Nejčastěji používaným kritériem ekvivalence při organizaci znalostí je kategorizace založená na společných vlastnostech. Řazení je uspořádání organizovaných objektů podle zvojeného kritéria. Vícekriteriální kategorizaci a klasifikaci umožňuje fasetová analýza.

##### c) Aktuální problémy organizace znalostí

Aktuální problémy organizace znalostí představují změny ve společenském kontextu (přechod od industriální ekonomiky ke znalostní ekonomice, od informačních zdrojů ke zdrojům ekonomickým, od idejí národního státu ke globalizaci a multikulturalismu) a změna komunikačních technologií (přechod od lineárních komunikačních a datových modelů k modelům síťovým, od analogových dokumentů k digitálním, od monolitických statických

dokumentů k propojeným datům, od makroorganizace k mikroorganizaci, od dokumentů k informacím).

#### d) Řešení aktuálních problémů organizace znalostí

Způsoby řešení aktuálních problémů přinášejí změny v samotném procesu organizace znalostí (od koncentrace na uchování a ochranu dokumentů k zajištění dostupnosti dokumentů/znalostí, od pasivního popisu k aktivnímu návrhu nových systémů, od nedostatku informací k řešení problému informačního zahlcení a důvěryhodnosti, od technicistního a kulturně historizujícího přístupu k vědeckému přístupu, od praxe k teorii), rozšíření okruhu aktérů, tj. osob i institucí, jež se věnují organizaci znalostí, a změny v metodické oblasti (od izolovaných specifických aktivit k integrovanému přístupu, od hledání jednoho ‚nejlepšího‘ systému k propojení systémů stávajících, od specializovaných vědních disciplín k transdisciplinaritě).

#### **Jednotka organizace znalostí**

Jednotkou organizace znalostí je zaznamenaná znalost. Význam pojmu zaznamenaná znalost je významovou podmnožinou pojmu explicitní znalost. Lze konstatovat, že takto definovaná jednotka organizace znalostí se významně obsahově protíná s pojmem dokument.

#### **Proces organizace znalostí**

Proces organizace znalostí je jakákoli záměrná činnost, spočívající v zavádění struktury do existujících zaznamenaných znalostí. Vstupem jsou zaznamenané znalosti, výstupem (tj. produktem organizace) jsou organizované zdroje, struktura a metadata. Procesu organizace znalostí předcházejí procesy reprezentace a zaznamenání znalosti, následuje případné uložení, přístup ke znalostem a užití znalosti. Typologie přístupu ke znalostem ovlivňuje způsob, jak budou znalosti organizovány, a má rovněž vliv na volbu výstupů z procesu organizace. Významnými podprocesy organizace znalostí jsou identifikace a popis, obsahová analýza a začlenění do kontextu.

Technologické postupy v průběhu organizace znalostí mají za úkol podporovat průběh procesů přístupu, odkazování, seskupování a odlišení.

Technologické nástroje procesu organizace znalostí v současnosti představují technologie *Internetu* a počítačové technologie. Technologie *Internetu* se aktuálně soustředí na problémy sémantického webu, počítačové technologie aktuálně představují databázové a znalostní systémy.

#### **Kontext procesu organizace znalostí**

Realizační kontext procesu organizace znalostí představují informační zdroje, používané metody a nástroje a aktéři.

Zdroje používané v doméně organizace znalostí jsou obdobného typu jako zdroje používané v analogických doménách zaměřených na komunikaci informací ve společnosti. Jediným doménově specifickým typem zdroje je systém organizace znalostí. Ten plní kromě funkce doménového zdroje současně i funkci nástroje procesu organizace znalostí.

Typologie systémů organizace znalostí zahrnuje tyto typy: nabídkový seznam, negativní slovník, (řízený) slovník, číselník, výkladový slovník, terminologický slovník, seznam synonym, slovník

zeměpisných názvů, seznam jmenných autorit, předmětový heslář, kategorizační schéma, taxonomie, klasifikační schéma, tezaurus, sémantická síť, (terminologická) ontologie.

Klíčovou metodou organizace je analýza, jejíž efektivnost podporuje dodržování pravidel odvozených z logiky: jasné ohraničení analyzovaného celku i částí, úplnost členění, soudržnost, jednotné kritérium členění, vzájemně se vylučující části, stejná úroveň granularity i abstrakce.

Aktéry organizace znalostí jsou tvůrci obsahu, jeho zprostředkovatelé a koncoví uživatelé, přičemž v každé z rolí může vystupovat jak člověk, tak počítačový program.

Proces organizace znalostí má rozsáhlý kognitivní kontext, který tvoří filozofické disciplíny i speciální teoretické a aplikované obory. Obor organizace znalostí byl ustaven v 90. letech 20. století. Předmětem zkoumání oboru organizace znalostí je proces organizace znalostí a jeho realizační kontext, tj. zdroje, jež jsou v procesu organizace transformovány (jednotky organizace), používané metody a nástroje a produkty, jež jsou tímto procesem vytvářeny, včetně zúčastněných aktérů – osob, institucí, technologií.

#### 4.7.2 Semiformalizovaný model domény organizace znalostí (výroky)

Výsledkem doménové analýzy problematiky organizace znalostí je následující soubor semiformalizovaných výroků, jež byly odvozeny z textového modelu:

- Cílem organizace znalostí je pořádek (snížení entropie).
- Pořádek umožňuje přístup ke znalostem a využívání znalostí.
- Znalosti je možné využít k poznání, řešení problémů a tvorbě nových znalostí.
- Jednotkou organizace znalostí je zaznamenaná znalost.
- Vlastností organizovaných entit je naležitelnost/vyhledatelnost, srozumitelnost a komunikovatelnost.
- Obor organizace znalostí zkoumá proces organizace znalostí a jeho realizační kontext.
- Realizační kontext procesu organizace znalostí představují zdroje, metody a nástroje, produkty a aktéři.
- Aktéři procesu organizace znalostí jsou tvůrci obsahu, zpracovatelé a koncoví uživatelé.
- Proces organizace znalostí má rozsáhlý kognitivní kontext, který tvoří filozofické disciplíny i speciální teoretické a aplikované obory.
- Klíčové podprocesy organizace znalostí jsou identifikace, obsahová analýza, začlenění do kontextu, kategorizace/klasifikace a řazení.
- Produktem (výstupem) procesu organizace znalostí je struktura.
- Produktem (výstupem) procesu organizace znalostí jsou metadata.
- Problémy řešené v rámci procesu organizace znalostí jsou lingvistické, sémantické a pragmatické.
- Lingvistické problémy jsou synonymie, homonymie a slovní spojení.
- Sémantické problémy jsou *aboutness*/obsah a složená témata.
- Pragmatické problémy jsou zkreslení / zaujetí (*bias*) a *warrant*.
- Znalostní báze je jedním z typů produktů organizace znalostí.
- Jednotkou organizace ve znalostní bázi je výrok jako specifický typ zaznamenané znalosti.
- Nástroje organizace znalostí jsou založeny na principu ekvivalence.
- Ekvivalenci je možné vyjádřit umístěním nebo označením.
- Nástrojem řešení lingvistických problémů organizace znalostí je řízený slovník a kontext.
- Systém organizace znalostí je produktem (výstupem) i nástrojem procesu organizace znalostí.
- Jednotkou organizace v systému organizace znalostí jsou pojmy.
- Formou vyjádření pojmů v systému organizace znalostí je slovník.
- Technologie používané v procesu organizace znalostí jsou databázové a znalostní systémy a technologie sémantického webu.

Z výše uvedených semiformalizovaných výroků byly odvozeny následující pojmy (jsou uvedeny v abecedním pořadí):

<i>aboutness</i>	organizace znalostí (obor)	synonymie
databázový/znalostní systém	organizace znalostí (proces)	systém organizace znalostí
ekvivalence	označení	tvůrce
homonymie	pojem	umístění
identifikace	pořádek	výrok
indexace (zdrojů)	přístup (ke znalostem)	využívání (znalostí)
kategorizace / klasifikace	řazení	<i>warrant</i>
komunikovatelnost	řízený slovník	začlenění do kontextu
koncový uživatel	slovní spojení	zkreslení / zaujetí ( <i>bias</i> )
kontext (kognitivní / realizační)	slovník	zaznamenaná znalost
metadata	složené téma	znalostní báze
nalezitelnost / vyhledatelnost	srozumitelnost	zprostředkovatel obsahu
obsahová analýza	struktura	

Úplnost pojmového pokrytí domény byla testována fasetovým způsobem pomocí metody *5W1H*. Zvolené fasety koncepčně vycházejí ze základních kategorií této metody. Faseta ‚proč‘ se zaměřuje na cíl/účel organizace znalostí. Ve fasetě ‚co‘ jsou jednotky a výstupy organizace znalostí. Faseta ‚jak‘ zahrnuje procesy, nástroje a metody a faseta ‚kdo‘ aktéry organizace znalostí. Fasety ‚kde‘ a ‚kdy‘ charakterizují místo a čas organizace znalostí a jejich použití vedlo k doplnění čtyř dalších pojmů: organizace v externí paměti, organizace v mysli, dopředná organizace, zpětná organizace. V tabulce 14 jsou jednotlivé termíny ve fasetách řazeny formálně podle abecedy.

organizace znalostí					
proč	co	jak	kdo	kde	kdy
komunikovatelnost	kategorizace/klasifikace	databázový / znalostní systém	koncový uživatel tvůrce obsahu	organizace v externí paměti	dopředná organizace
lingvistické problémy	kontext	ekvivalence	zprostředkovatel obsahu	organizace v mysli	zpětná organizace
▪ synonymie	metadata	identifikace			
▪ homonymie	pojem	indexace			
▪ slovní spojení	slovník	obsahová analýza			
nalezitelnost/vyhledatelnost	struktura	označení			
pořádek	výrok	řazení			
pragmatické problémy	zaznamenaná znalost	řízený slovník			
▪ zkreslení /zaujetí		sémantický web			
▪ <i>warrant</i>		systém organizace znalostí			
přístup		umístění			
sémantické problémy		znalostní báze			
▪ <i>aboutness</i>					
▪ složené téma					
srozumitelnost					
využívání					

Tab. 14 Fasetová klasifikace klíčových pojmů organizace znalostí



## Kapitola 5 Návrh druhé verze ontologie organizace znalostí

*Práce s digitálními objekty a systémy v žádném případě nezmenšila význam terminologie a terminologické přesnosti pro předmětnou oblast. Naopak, tato důležitost se zvýrazňuje, jelikož pojmy a termíny reprezentující objekty musí být ještě přesnější v případě, kdy se provádí rozsáhlé automatické zpracování dat.<sup>382</sup>*  
ISO 5127:2017

Po vyhodnocení první verze ontologie organizace znalostí, navržené a implementované v rámci řešení projektu *NAKI Znalostní báze pro obor organizace informací a znalostí* (viz závěry kapitoly 3), bylo v roce 2016 rozhodnuto provést její reengineering. Pro reengineering byla zvolena metoda doménové analýzy a opětovného použití základní ontologie. Výsledkem je návrh druhé verze ontologie, který je dokumentován v této kapitole. Kapitola je členěna na pět částí. V části 5.1 je stanovena specifikace požadavků na druhou verzi ontologie. Část 5.2 je věnována popisu architektury ontologie. V části 5.3 je popsán pojmový model ontologie. Její součástí je popis pojmové struktury základní ontologie *DOLCE*, jež byla využita při návrhu druhé verze ontologie organizace znalostí. Část 5.4 popisuje postup návrhu a implementace ontologie. V části 5.5 je zhodnoceno splnění stanovených cílů prostřednictvím specifikace uskutečnění stanovených případů užití ontologie.

### 5.1 Specifikace požadavků

Formulování specifikace požadavků předcházela fáze obecných úvah o typu druhé verze ontologie, jež stanovila nejobecnější úroveň ontologických závazků. Cílem bylo zachovat charakter doménové ontologie i stávající oblast užití, již pokrývá první verze, tj. opakované použití a organizaci doménových znalostí v prostředí znalostní báze organizace znalostí. Druhá verze má rozšířit možnosti použití i na další tři oblasti užití inforatických ontologií, charakterizované v kapitole 1.3: komunikace doménových znalostí, popis domény umožňující zpracování a podporu počítačovými systémy, automatické odvozování.

Typ znalostí reprezentovaných v druhé verzi ontologie bylo rozhodnuto ponechat beze změny. Stejně jako první verze se zaměřuje na reprezentaci deklarativních znalostí z domény organizace znalostí, reprezentace procedurálních znalostí není požadována. Ontologické třídy mají zahrnovat jak 3D (třírozměrné), tak 4D (čtyřrozměrné) entity, tj. objekty i události. Protože doména organizace znalostí zahrnuje velké množství abstraktních entit bez prostorových a časových vlastností, bude ontologie obsahovat i posibilistické třídy (pro entity, jejichž existence není dokazatelná).

Nejvýznamnějším rozhodnutím bylo určení typu konceptualizace z hlediska účelu, tj. rozhodnutí, zda ontologie bude deskriptivní nebo revizionistická. I s vědomím toho, že v praxi je často obtížné určit hranici mezi gnozeologickou deskripcí a ontologickou restrikcí, jsme se při určení charakteru ontologie organizace znalostí přiklonili k deskripci. Druhá verze ontologie má tedy zahrnovat deskriptivní entity, jež vycházejí z běžného chápání skutečnosti za použití přirozeného jazyka.

V souladu s těmito ontologickými závazky byla zformulována specifikace požadavků a z nich odvozených případů užití. Každému případu užití byl přiřazen aktér, který reprezentuje klíčovou uživatelskou roli, tj. uživatele dané funkce.

---

<sup>382</sup> ISO 5127:2017. *Information and documentation – Foundation and vocabulary*. 2nd ed. Geneva: International Organization for Standardization, 2017-05.

### **Požadavek 1 – Tvorba výroků umožňujících odvozování**

Komentář: Ontologie bude využitelná pro tvorbu výroků pro znalostní bázi organizace znalostí. Tento požadavek se zaměřuje na zajištění kontinuity a na zachování funkce, kterou ontologie plnila už ve své první verzi. Využití výroků konkretizují dva případy užití. Ty jsou s ohledem na nutnost údržby a dalšího vývoje znalostní báze organizace znalostí zaměřeny na tvorbu a editaci výroků pro znalostní bázi a na možnost automatického odvozování. Ontologie má umožnit tvorbu instancí ve formě výroků pro všechny tři moduly znalostní báze, tj. pro modul pojmů, modul bibliografických entit i pro modul pojmenovaných entit. Výroky budou tvořeny podle obecného schématu daného jazykem *RDF*: [subjekt] <predikát> [objekt]. Tato formalizace umožní splnit požadavek na zajištění funkce popisu domény pro její zpracování a podporu počítačovými programy.

#### **Případ užití 1.1**

Ontologie bude nástrojem opětovného použití zaznamenaných doménových znalostí.

Aktér: Uživatel znalostní báze.

#### **Případ užití 1.2**

Ontologie bude nástrojem odvozování, objevování a tvorby nových znalostí v rámci oborového výzkumu.

Aktéři: Odvozovací software (zpracovatel), uživatel znalostní báze.

### **Požadavek 2 – Indexace doménových zdrojů umožňující přístup**

Komentář: Ontologie bude využitelná pro indexaci zdrojů v doméně organizace znalostí. Tento požadavek vychází z potřeby zajistit funkci opětovného použití doménových znalostí. Využití výroků konkretizují dva případy užití, odvozené z deskriptivního charakteru ontologie – metadatový popis zdrojů a vyhledávání zdrojů založené na metadatach. Výroky budou tvořeny podle schématu: [zdroj/dotaz] <má předmět> [ontologická třída]. Ontologie umožní uživatelům upřesnit nebo rozšířit formulaci dotazu podle nabídky: zahrnutí nadřazených / podřazených termínů v souladu s taxonomickou strukturou ontologie, zahrnutí ekvivalentních a asociovaných termínů v souladu s definovanými restrikcemi / axiomy. Odpovědi na dotazy budou buď odkazy na zdroje, nebo výroky představující „odpověď“ znalostního systému.

#### **Případ užití 2.1**

Vyjádření obsahu doménových zdrojů.

Aktér: Editor znalostní báze (zpracovatel), uživatel znalostní báze.

#### **Případ užití 2.2**

Vyjádření obsahu uživatelských dotazů (tj. formulace informační, resp. znalostní potřeby) v jazyce *SPARQL*.

Aktéři: Uživatel znalostní báze, vyhledávací software (zpracovatel).

### **Požadavek 3 – Sémantické propojení s dalšími zdroji umožňující komunikaci**

Komentář: Ontologie bude propojena s věcně relevantními ontologiemi a s dalšími externími zdroji. Tento požadavek vychází z potřeby zajistit komunikační funkci ontologie jak pro lidské uživatele, tak pro stroje v rámci sémantického webu.

#### **Případ užití 3.1**

Sémantické mapování ontologických tříd a predikátů na věcně relevantní ontologie.

Aktér: Odvozovací software



### Případ užití 3.2

Formalizace ontologie, umožňující začlenění do cloudu propojených dat.

Aktér: Vyhledávací software.

#### Specifikace požadavků na základní ontologii

Vzhledem k tomu, že bylo rozhodnuto využít při návrhu druhé verze ontologie metodu opětovného použití základní ontologie, stala se součástí specifikace požadavků i sada kritérií pro výběr základní ontologie. Při výběru základní ontologie byla uplatněna tato rámcová kritéria: 1) ontologie má svým obsahem a ontologickými závazky korespondovat se specifiky domény organizace znalostí a 2) ontologie má být k dispozici v režimu otevřeného přístupu ve formátu propojených otevřených dat. Tato kritéria byla aplikována jednak během vlastního orientačního seznámení s pojmovým obsahem vrcholových tříd vybraných základních ontologií *BFO*, *DOLCE*, *KR Ontology*, *OpenCYC*, *SUMO*, jednak za použití nástroje *ONSET*<sup>383</sup>, který umožňuje výběr z ontologií *BFO*, *DOLCE*, *GFO* a *SUMO*.

V charakteristice nástroje *ONSET* v kapitole 2.1 bylo konstatováno, že nabízí podporu pro vícekritériální rozhodování, přičemž uživatel má možnost zvolit pouze pro něj relevantní kritéria. Kritéria jsou členěna do pěti kategorií a uživatel může přiřadit těmto kategoriím vlastní váhy (0–5). V našem výběru jsme použili tato kritéria s následujícími váhami (v závorce jsou uvedeny požadované hodnoty):

#### 1. kategorie ontologických závazků – váha 3

- deskriptivní nebo realistická ontologie (deskriptivní)
- multiplikativní nebo redukcionistická ontologie (multiplikativní)
- aktualistická nebo posibilistická ontologie (posibilistická)
- konkrétní nebo abstraktní entity (obojí)

#### 2. kategorie jazyka pro reprezentaci znalostí – váha 5

- jazyk pro reprezentaci znalostí (OWL 2 DL)

#### 3. kategorie vlastností odvozených ze softwarového inženýrství – váha 5

- modularita (ano)
- dostupnost ‚lehkých‘ a ‚těžkých‘ verzí (ano)
- oddělení endurantů a perdurantů (ano)
- moduly pro funkce a role (ano)

#### 4. kategorie pokrytí věcných domén – váha 0 (kritéria z této kategorie nebyla použita)

#### 5. kategorie aplikačních oblastí – váha 3

- informační systémy řízené ontologiemi (ano)
- sémantický web (ano)
- vědecký výzkum (ano)

Na základě požadovaných hodnot zvolených kritérií nástroj *ONSET* doporučil základní ontologii *DOLCE*. Výsledky jsou dokumentovány v příloze 2. Doporučení získané vícekritériální analýzou pomocí nástroje *ONSET* bylo podpořeno výsledky vlastního šetření. Ty potvrdily, že *DOLCE* svým obsahem a ontologickými závazky skutečně nejlépe odpovídá specifikům domény organizace znalostí.

---

<sup>383</sup> Nástroj a dokumentace jsou dostupné z: <http://www.meteck.org/files/onset/index.html> [cit. 2018-01-30].

Ontologie se zaměřuje na problematiku lingvistiky, sémiotiky a komunikace, jež tvoří teoretickou bázi oboru organizace znalostí. Obsahuje mj. i návrhové vzory pro řešení problému identity webových zdrojů a pro vyjádření sémantiky. Má propracovaný a na filozofických základech vytvořený systém abstraktních tříd (tj. tříd, jež nezaujímají místo v prostoru a čase, nemají hmotu ani energii). Tím umožňuje propojení s pojmovou základnou organizace znalostí, již tvoří v převažující většině abstraktní pojmy.

Ontologie má všeobecný a deskriptivní charakter, což je v souladu se záměrem vytvořit doménovou ontologii jako deskriptivní. Autoři explicitně stanovili, že jejich cílem je vytvořit základní ontologii „zachycující ontologické kategorie, jež jsou základem přirozeného jazyka a běžného chápání skutečnosti“.<sup>384</sup> Uvádějí, že tyto kategorie „jsou myšleny jako kognitivní artefakty, jež v konečném důsledku závisejí na lidském vnímání, kulturních vzorcích a společenských konvencích (jakási ‚kognitivní‘ metafyzika)“.<sup>385</sup> Dalším rysem DOLCE je multiplikativnost, spočívající v kolokalizaci více entit v jednom časoprostoru.

Požadavek na dostupnost v režimu otevřeného přístupu ve formátu propojených otevřených dat je rovněž splněn. Ontologie je volně dostupná ve formátu *RDF/OWL*. Ontologie je relativně dobře známá v odborné komunitě a má již četné aplikace (*DOLCE* je mimo jiné mapována do terminologické ontologie *WordNet*). Modulární charakter ontologie umožňuje rozšiřování a další úpravy. Je perspektiva jejího dalšího vývoje; autoři předpokládají mapování *DOLCE* do dalších základních ontologií, což dále rozšíří její aplikační potenciál v rámci sémantického webu.

## 5.2 Architektura ontologie organizace znalostí

Při návrhu architektury byly uplatněny obecně přijímané principy ontologického inženýrství, jež byly představeny v kapitole 2.1, především:

- Maximalizace opětovného použití již existujících ontologií
- Modulární struktura
- Flexibilita, připravenost na budoucí změny.

Ontologie organizace znalostí je integrovaným celkem, který čerpá ze tří zdrojů: základní ontologie *DOLCE*, výsledky doménové analýzy, první verze ontologie organizace znalostí.

Architekturu tvoří dva navzájem propojené moduly, rozšířené o opětovně použité pojmy z externí základní ontologie: jádrová (*core*) ontologie a terminologická ontologie.

Jádrová a terminologická ontologie se liší v míře formalizace. Jádrová ontologie je semiformální, terminologická ontologie je vyjádřena přirozeným jazykem (k typologii ontologií viz kapitola 1.4) a má strukturu tezauru dle ISO 25964<sup>386</sup>. Jádrová a terminologická ontologie mají rozdílné případy užití. Výroky vytvořené za použití jádrové ontologie umožňují automatické odvozování,

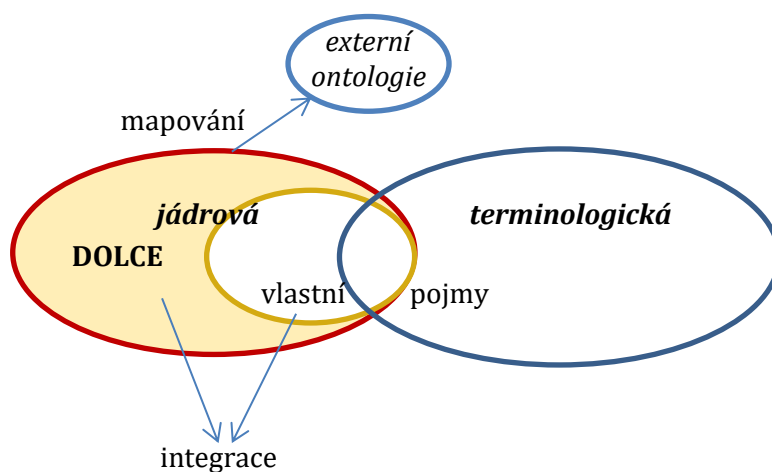
---

<sup>384</sup> MASOLO, Claudio, BORGIO, Stefano, GANGEMI, Aldo, GUARINO, Nicola, OLTRAMARI, Alessandro. *WonderWeb deliverable D18: Ontology library (final)* [online]. Version 1.0. ISTC-CNR, 2003-12-31 [cit. 2018-01-30], s. 13. Dostupné z: <http://wonderweb.man.ac.uk/deliverables/D18.shtml> a z <http://www.loa.istc.cnr.it/old/Papers/D18.pdf>.

<sup>385</sup> Tamtéž, s. 8.

<sup>386</sup> ISO 25964-1:2011. *Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval*. 1st ed. Geneva: International Organization for Standardization, 2011-08-08. 152 s.

výroky používající terminologickou ontologii mohou dosáhnout významové obsažnosti a přesnosti.



Obr. 22 Schéma propojení modulů ontologie

### Základní ontologie

Oproti způsobu řešení uplatněnému v první verzi ontologie, do níž byly zahrnuty pouze vlastní pojmy získané excerpcí odborné literatury a od doménových expertů, byla při návrhu druhé verze ontologie zvolena varianta založení doménové ontologie na základní ontologii. Přínosem opětovného použití pojmů ze základní ontologie je inference (zdědění) bohatých axiomů definovaných v této ontologii a jejich aplikace na pojmy v ontologii organizace znalostí. Opětovné použití pojmů umožnilo převzít již zformulované axiomy a definice významu těchto pojmů a rovněž ‚zdědit‘ jejich predikáty (tj. atributy a vztahy).

Základní ontologie ve verzi *DOLCE+DnS Ultralite* (viz kapitola 5.3.1) se svými rozšiřujícími moduly *LMM* a *IOL* je importována do jádrové ontologie organizace znalostí a její třídy, predikáty a axiomy tvoří vrchní nejobecnější úroveň ontologie organizace znalostí.

### Jádrová ontologie

Jádrová ontologie organizace znalostí obsahuje nejobecnější třídy a predikáty, doplněné importovanými pojmy z ontologie *DOLCE*. Třídy i predikáty jsou mapovány do externích informačních a ‚lehkých‘ terminologických ontologií a do relevantních systémů organizace znalostí. Jádrová ontologie je formalizovaná v jazyce *OWL* a *RDF* a umožňuje tudíž odvozování. Lze ji považovat za těžkou ontologii.

### Terminologická ontologie

Terminologická ontologie organizace znalostí obsahuje detailní terminologii oboru. Každá nejvyšší třída (*top term*) terminologické ontologie je podtřídou některé ze tříd jádrové ontologie, což zajišťuje jejich propojení. Predikáty a axiomy jsou převzaty (děděny) z jádrové ontologie. Tím je dáno, že oba moduly se částečně překrývají.

### Mapování na externí ontologie

Dalším krokem směrem k interoperabilitě ontologie organizace znalostí je mapování zahrnutých ‚originálních‘ pojmů z jádra do relevantních ontologií a dalších zdrojů s vyjádřením stupně ekvivalence. Mapování umožňuje interoperabilitu v rámci sémantického webu. K mapování byly

využity predikáty z datového slovníku *SKOS* (viz kapitola 4.5.3). Při volbě zdrojů pro mapování byla zvažována dvě kritéria: obsahová relevance a dostupnost ve formátu propojených otevřených dat.

## 5.3 Pojmový model ontologie

### 5.3.1 Základní ontologie *DOLCE*

*DOLCE: a Descriptive Ontology for Linguistic and Cognitive Engineering* (Deskriptivní ontologie pro lingvistiku a znalostní inženýrství) je základní ontologie, vyvinutá v rámci výzkumného projektu *WonderWeb*, realizovaného na půdě italského institutu *ISTC-CNR* (Řím) v letech 2001–2004. Na návrhu *DOLCE* se podíleli Claudio Masolo, Stefano Borgo, Aldo Gangemi, Nicola Guarino a Alessandro Oltramari. Projekt byl součástí 5. rámcového programu Evropské komise *IST (Information Society Technologies, IST-2001-33052)*.<sup>387</sup> Jeho cílem bylo navrhnout knihovnu základních ontologií (*WonderWeb Foundational Ontologies Library – WFOL*). Ontologie *DOLCE* se stala prvním modulem této knihovny. Původně byla zamýšlena jako nástroj porovnání existujících ontologií a lingvistických zdrojů (slovníků) a objasnění jejich ontologických závazků. V současné době patří mezi nejpoužívanější základní ontologie při tvorbě doménových ontologií. Obsahuje cca 100 pojmů a cca 100 axiomů.

Ontologie *DOLCE* je podle svých autorů minimálně ‚zatížena‘ ontologickými závazky, jež by ji spojovaly s některým filozofickým směrem. Základními filozofickými koncepty jsou orientace na jednotliviny (instance) a členění zahrnutých objektů podle vztahu k času na tzv. *perduranty* a *enduranty* (viz kapitola 1.5.4). *DOLCE* má polyhierarchickou strukturu. Jedna třída tedy může být podřazena více jiným třídám (příkladem jsou třídy Informační objekt, Popis a Situace, znázorněné na obrázku 23). Tato skutečnost na jedné straně komplikuje ‚přehlednost‘ struktury ontologie, na druhé straně polyhierarchie lépe odpovídá reprezentaci často komplikované a nejednoznačné skutečnosti.

*DOLCE* lze označit za ‚těžkou‘ ontologii, protože vykazuje vysoký stupeň formalizace a umožňuje tak automatické odvozování a usuzování. Axiomy jsou konstruovány pomocí výroků modální logiky. Při kontrole logické správnosti *DOLCE* byla použita metodika *OntoClean* (viz kapitola 2.1.2).

#### Verze a moduly

Ontologie má modulární charakter, což se projevuje ve dvou směrech: 1. *DOLCE* je součástí (tj. jedním z modulů) knihovny základních ontologií. 2. Samotná ontologie je k dispozici v několika variantách a samostatných modulech: *DOLCE Full*, *DOLCE Lite-Plus (DLP)*, *Description and Situation framework / Descriptions and situations (D&S)* a *DOLCE+DnS Ultralite (DUL)* s několika dalšími zásuvnými moduly.

Úplná verze ontologie *DOLCE Full* je k dispozici ve formě textového souboru v rámci dokumentace projektu *WonderWeb*<sup>388</sup> a její koncepce je popsána v samostatné kapitole příručky

<sup>387</sup> Informace o projektu jsou dostupné z: [http://cordis.europa.eu/project/rcn/60325\\_en.html](http://cordis.europa.eu/project/rcn/60325_en.html) [cit. 2018-01-30].

<sup>388</sup> MASOLO, Claudio, BORGO, Stefano, GANGEMI, Aldo, GUARINO, Nicola, OLTRAMARI, Alessandro. *WonderWeb deliverable D18: Ontology library (final)* [online]. Version 1.0. ISTC-CNR, 2003-12-31 [cit. 2018-01-30], s. 129-164. Dostupné z: <http://wonderweb.man.ac.uk/deliverables/D18.shtml> a z <http://www.loa.istc.cnr.it/old/Papers/D18.pdf>.

*Handbook of ontologies*<sup>389</sup>. Je formalizována v ontologickém jazyku *KIF (Knowledge Interchange Format)*, není však k dispozici ve formátu propojených dat.

*DOLCE Lite-Plus (DLP)* je zjednodušením úplné verze a jejím překladem do několika ontologických jazyků: *KIF3.0 (PL)*, *DAML+OIL* a *OWL-DL*.

*Description and Situation framework / Descriptions and situations (D&S)* je ontologie, která funguje jako zásuvný rozšiřující modul pro *DOLCE*. Její jádro tvoří třídy *Situace (Situation)* a *Popis (Description)*, jež slouží k popisu kontextů (referenčních rámců) a situací. Ontologie je založena na teorii ontologického kontextu, která umožňuje modelovat nefyzické objekty (společenské instituce, plány, organizace, normy, role, parametry), jejichž význam nevyplývá z jejich vlastností, ale z výroků, tj. z kombinace jiných entit. Autoři Aldo Gangemi a Peter Mika se při tvorbě této ontologie snažili o integraci lingvistické a sémiotické teorie.<sup>390</sup>

*DOLCE+DnS Ultralite* vytvořil Aldo Gangemi jako zjednodušení a současně rozšíření určitých částí ontologie *DOLCE Lite-Plus* a ontologie *Descriptions and situations*. Tato verze *DOLCE* používá intuitivnější jména tříd a vztahů, zjednodušené časové a prostorové vztahy a jednodušší axiomatické konstrukce. Z hlediska zaměření ontologie organizace znalostí je významné rozšíření o třídu Informační entita (*InformationEntity*) s podtřídami Informační objekt (*InformationObject*) a Informační realizace (*InformationRealization*). Aktuální verze 3.31 obsahuje 75 tříd a 112 predikátů.

Pro použití v ontologii organizace znalostí jsou významné dva zásuvné moduly, jež byly vytvořeny jako rozšíření ontologie *DOLCE+DnS Ultralite: IOL – Information Objects Ontology* (Ontologie informačních objektů) a *LMM – Linguistic Meta-Model* (Lingvistický metamodel).

*IOL – Information Objects Ontology* je zásuvný modul umožňující modelovat informační a lingvistické objekty a zdroje. Obsahuje 46 tříd a 31 predikátů. Zavádí významnou třídu Lingvistický objekt (*LinguisticObject*), jež je typem informačního objektu a zahrnuje jazykové nástroje komunikace (slovo, text, věta ad.).

*LMM – Linguistic Meta-Model* má formu návrhového vzoru, který umožňuje konstruovat pojmy související se sémiotickými a lingvistickými úlohami a se zpracováním přirozeného jazyka.<sup>391</sup> Člení se na dva dílčí moduly *L1* a *L2*. Modul *L1 (Lexical MetaModel Level 1)* obsahuje 8 tříd a 9 predikátů. Modul *L2 (Lexical MetaModel Level 2)* obsahuje 21 tříd a 7 predikátů.

## **Základní třídy**

Pojmový základ ontologie *DOLCE* tvoří čtyři základní třídy (autoři je nazývají kategorie): *Endurant (Endurant)*, *Perdurant (Perdurant/Occurence)*, *Kvalita (Quality)* a *Abstraktní entita (Abstract)*. Jak je v základních ontologiích obvyklé, ani v ontologii *DOLCE* nelze význam tříd

---

<sup>389</sup> BORGIO, Stefano, MASOLO, Claudio. Foundational choices in DOLCE. In: STAAB, Steffen, STUDER, Rudi, ed. *Handbook on ontologies*. 2nd ed. Dordrecht: Springer, 2009, s. 361-381. doi:10.1007/978-3-540-92673-3\_16. ISBN 978-3-540-70999-2 (Print). ISBN 978-3-540-92673-3 (Online).

<sup>390</sup> GANGEMI, Aldo, MIKA, Peter. Understanding the Semantic web through Descriptions and Situations. In: R. Meersman, Z. Tari, D. C. Schmidt, ed. *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE 2003*. Berlin, Heidelberg: Springer-Verlag, 2003, s. 689-706. doi: [https://doi.org/10.1007/978-3-540-39964-3\\_44](https://doi.org/10.1007/978-3-540-39964-3_44). Lecture notes in computer science, no. 2888, ISSN 0302-9743 (Print). ISSN 1611-3349 (Online). ISBN 978-3-540-20498-5 (Print). ISBN 978-3-540-39964-3 (Online).

<sup>391</sup> PICCA, Davide, GLIOZZO, Alfio Massimiliano, GANGEMI, Aldo. LMM: an OWL metamodel to represent heterogeneous lexical knowledge. In: *Proceedings of the International Conference on Language Resources and Evaluation (LREC)*. 26 May-1 June 2008, Marrakech, Morocco, s. 2413-2419.

odvodit z rozboru významu slov, jež byla použita pro jejich označení, a je nutné se seznámit s jejich definicí a kontextem. Názvy jsou často voleny tak, aby nekolidovaly s již zavedenými termíny v jiných ontologiích.

Třídy Endurant a Perdurant zhruba odpovídají koncepci 3D a 4D tříd, jak byly představeny v kapitole 1.5.4. Základním vztahem dvojice tříd Endurant a Perdurant v *DOLCE* je účast: endurant existuje (a může se měnit) v čase tím, že se účastní nějakého perdurantu (tj. procesu). V *DOLCE* se používá jako synonymum pro označení perdurantu termín *occurence* (výskyt).

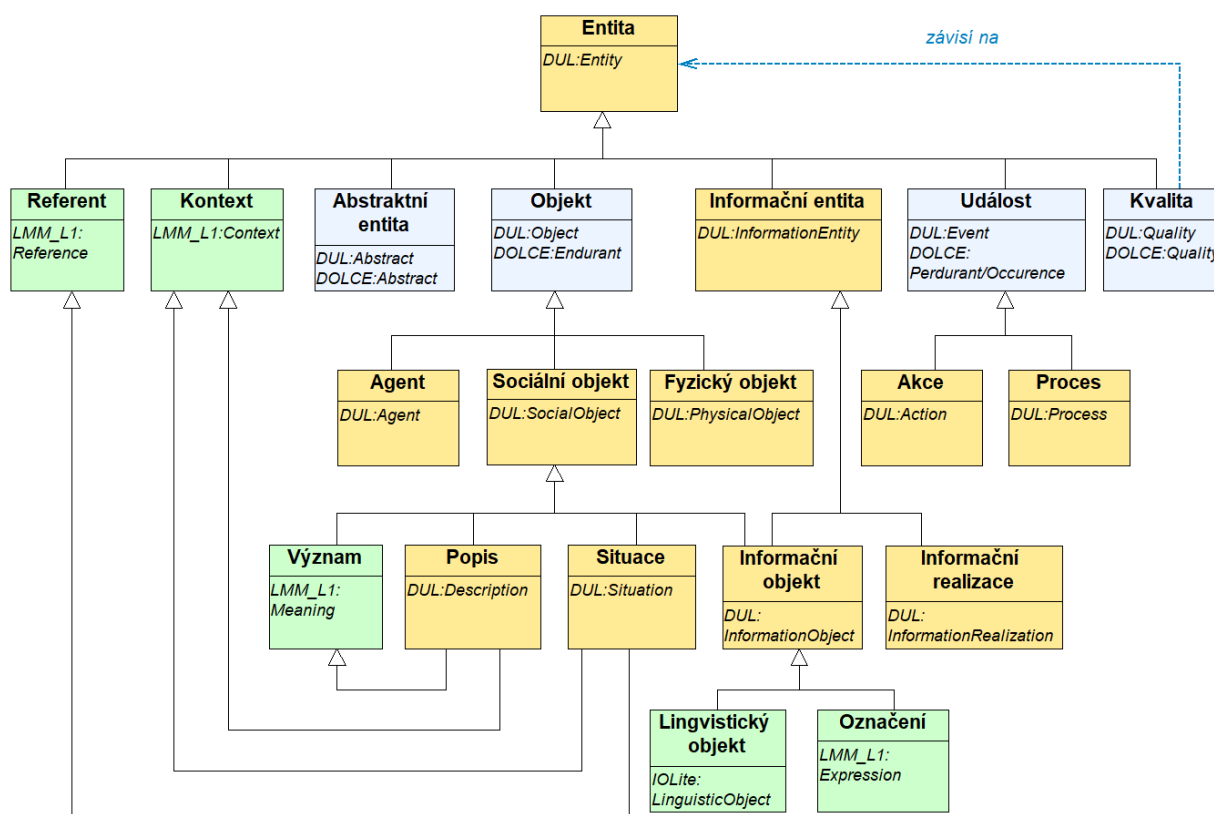
Třídu Kvalita lze zhruba chápat jako vlastnost, v *DOLCE* má ovšem specifický význam – stejně jako všechny ostatní třídy je chápána jako jednotlivina. Je to entita, již lze vnímat nebo měřit (barva, hmotnost, včetně vlastností časových a prostorových), a vždy je inherentní (závislá) vzhledem k nějaké entitě.

Abstraktní entity se vyznačují tím, že nemají prostorové ani časové vlastnosti. Jejich existence v čase a prostoru je možná jen tehdy, jsou-li fyzicky zakódovány nebo ztělesněny.

Na obrázku 23 je znázorněna struktura vrcholových tříd *DOLCE*, konkrétně její verze *DOLCE+DnS Ultralite*, doplněné zásuvnými moduly *IOL* a *LMM*, jež byly implementovány v druhé verzi ontologie organizace znalostí. Třídy jsou opatřeny názvy v češtině (v ontologii je označení implementováno prostřednictvím predikátu `rdfs:label xml:lang="cs"`) a jejich původní označení v angličtině jsou doplněna prefixy, jež označují příslušnou verzi *DOLCE*, z níž byly převzaty.

DOLCE: *DOLCE Full*  
 DUL: *DOLCE+DnS Ultralite*  
 IOLite: *IOL – Information Objects Ontology*  
 LMM\_L1: *LMM –Lexical MetaModel Level 1*

Detailní údaje o používaných jmenných prostorech jsou v příloze 1.



Obr. 23 Třídy nejvyšší úrovně převzaté z ontologie *DOLCE*

Na nejvyšší úrovni hierarchie je zastřešující třída Entita (*DUL:Entity*), která zahrnuje cokoli skutečného, možného nebo imaginárního, o čem by tvůrce modelu chtěl za nějakým účelem uvažovat.

Ke čtyřem základním třídám Objekt, Událost, Kvalita a Abstraktní entita, převzatým z *DOLCE Full*, je v *DOLCE+DnS Ultralite* doplněna třída Informační entita (*DUL:InformationEntity*) a její podtřídy Informační objekt (*DUL:InformationObject*) a Informační realizace (*DUL:InformationRealization*). Tyto tři „informační“ třídy hrají významnou roli při vyjádření abstraktních i konkrétních entit z domény organizace znalostí. Informační entita je definována jako „určité množství informace, ať už je nebo není konkrétně realizováno“. Informační objekt je popsán jako „určité množství informace, jako je hudební kompozice, text, slovo, obraz, nezávisle na tom, jak je konkrétně realizován“. Toto vymezení zhruba odpovídá entitě Vyjádření z referenčního modelu *FRBR*. Informační realizace je specifikována jako „konkrétní realizace informačního objektu, např. písemný dokument obsahující text zákona“, což koresponduje s chápáním obsahu entity Provedení z modelu *FRBR*.

Třída Endurant z verze *DOLCE Full* je ve verzi *DOLCE+DnS Ultralite* přejmenována na Objekt (*DUL:Object*) a třída Perdurant na Událost (*DUL:Event*) a jejich rozsah je ve srovnání s koncepty endurantu a perdurantu rozšířen (tj. zahrnují více typů entit). Jako pomůcka pro rozlišení objektu a události je stanoveno pravidlo, že objekt existuje primárně v prostoru a událost existuje primárně v čase.

Třída Událost má dvě podtřídy: Akce (*DUL:Action*) a Proces (*DUL:Process*).

Objekty jsou členěny na tři navzájem propojené skupiny – fyzické objekty, sociální objekty a agenty. Pro tyto účely slouží tři podtřídy: Agent (*DUL:Agent*), Sociální objekt (*DUL:SocialObject*) a Fyzický objekt (*DUL:PhysicalObject*).

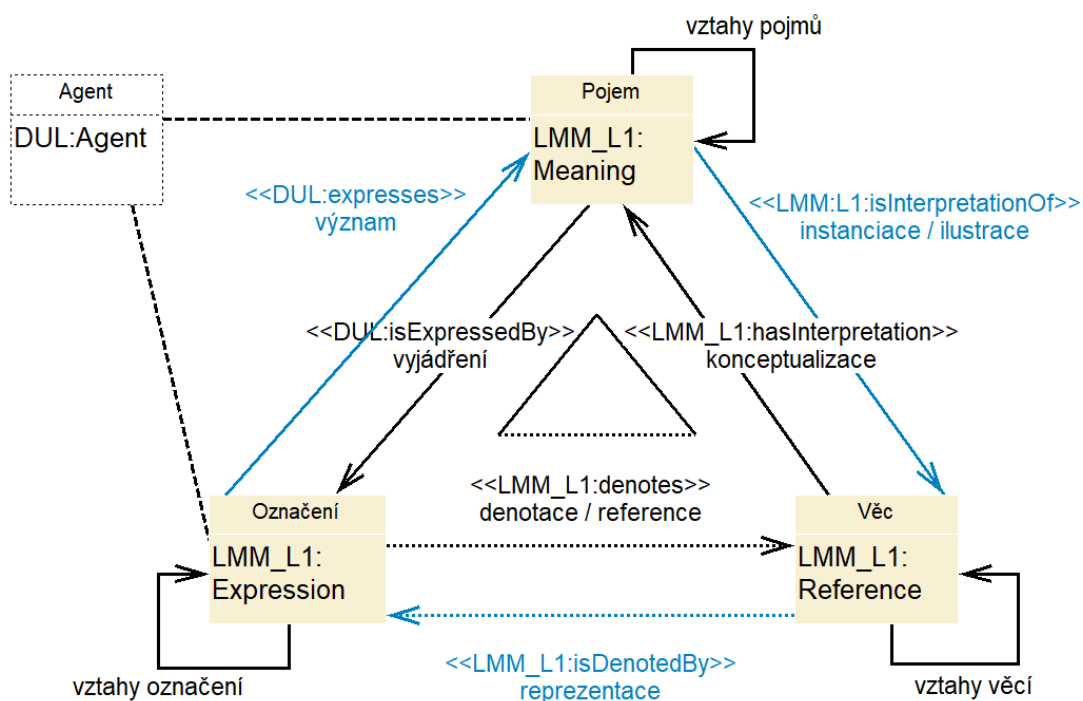
Další důležité třídy Popis (*DUL:Description*) a Situace (*DUL:Situation*) jsou převzaty z modulu *D&S* a jsou do ontologie *DOLCE+DnS Ultralite* integrovány jako podtřídy třídy Sociální objekt. Třída Situace zachycuje podrobný kontext, v němž se vyskytují entity a události: předběžné podmínky a důsledky, aktivity, aktéry, objekty apod. Je reifikací výroků konstruovaných s použitím prvků základní ontologie *DOLCE*. Třída Popis popisuje situaci a její intenze je konstruována z popisných entit, jež tvoří opět prvky základní ontologie. Tyto skutečnosti se odrážejí v polyhierarchických vztazích obou tříd.

Další vrcholovou třídou, doplněnou z modulu *LMM*, je Kontext (*LMM\_L1:Context*). Tato třída je zobecněním všech typů kontextů, jak materiálních (například místo jako prostorový kontext entity), tak nemateriálních (například kulturní, vztahový, sociální, informační kontext).

Klíčový význam v modulu *LMM* mají třídy, jež reprezentují pojmy sémiotického trojúhelníku. Tyto pojmy byly na obecné úrovni diskutovány v kapitole 1.5.2. Níže uvedený obrázek 24 je replikou obrázku 5 z této kapitoly a pro možnost srovnání použité terminologie obsahuje i termíny použité v této kapitole: věc, označení a pojem. Ontologie *DOLCE* používá pro označení klíčových pojmů sémiotického trojúhelníku tyto termíny: Referent (*LMM\_L1:Reference*, věc), Označení (*LMM\_L1:Expression*, označení), Význam (*LMM\_L1:Meaning*, pojem).

Třída Referent je podtřídou obecné třídy Entita, což umožňuje, aby se předmětem označení či vyjádření stal jakýkoli fyzický či abstraktní objekt, včetně znaku či pojmu samotného. Třída Význam je zařazena do kategorie Sociální objekt a třída Označení je podtřídou třídy Informační objekt. Vztahy mezi těmito třídami jsou definovány standardním způsobem ve formě predikátů. Vztahy mezi třídami Význam a Označení vyjadřuje dvojice predikátů vyjadřuje (*expresses*,

význam) / je vyjádřen (*isExpressedBy*, vyjádření). Třídy Označení a Referent spojují predikáty označuje (*denotes*, denotace/reference) / má označení (*isDenotedBy*, reprezentace). Vztah tříd Význam a Referent je podle autorů vztahem interpretace a je realizován predikáty je interpretací (*isInterpretationOf*, instance/ilustrace) / má interpretaci (*hasInterpretation*, konceptualizace).



Obr. 24 Interpretace sémiotického trojúhelníku v modulu LMM

### 5.3.2 Jádrová ontologie organizace znalostí

Integrací vybraných tříd základní ontologie *DOLCE* a tříd odvozených z vlastního modelu domény organizace znalostí (viz závěry kapitoly 4) byla vytvořena páteřní taxonomie jádrové ontologie. Celkem bylo do jádrové ontologie převzato 83 tříd a 137 predikátů včetně s nimi souvisejících axiomů z ontologie *DOLCE*. Ty jsou integrovány s vlastními třídami a predikáty (celkem 71 tříd a 64 predikátů). Úplný přehled tříd a predikátů je uveden v přílohách 3 a 4. Strukturu tříd a predikátů z hlediska jejich původu zachycuje tabulka 15.

prefix	počet tříd	počet predikátů
ko:	71	64
DUL:	44	91
IOLite:	28	30
LMM_L1:	5	9
LMM_L2:	6	7
<b>celkem</b>	<b>154</b>	<b>201</b>

Tab. 15 Kvantitativní přehled tříd a predikátů v jádrové ontologii

Východiskem pro začlenění vlastních tříd do struktury jádrové ontologie se staly jednak vztahy rozpoznané při tvorbě doménového modelu, jednak kategorizace pojmů při jejich fasetové



analýze (viz kapitoly 4.7.2 a 4.7.3). V páteřní taxonomii jsou zachovány vztahy polyhierarchie, jež jsou uplatňovány v základní ontologii *DOLCE*.

### Vlastní třídy v jádrové ontologii

Třídy v níže uvedeném seznamu jsou seskupeny podle toho, do které z vrcholových tříd základní ontologie *DOLCE* byly začleněny. Třídy založené na pojmech z pojmové mapy (viz kapitola 4.7.3) jsou zvýrazněny tučným písmem.

#### Abstraktní entita (DUL:Abstract)

jednotka organizace	ko:Jednotka_organizace
kategorie	ko:Kategorie
<b>pořádek</b>	<b>ko:Pořádek</b>
<b>struktura</b>	<b>ko:Struktura</b>

#### Informační entita (DUL:InformationEntity)

<b>(databázový) / znalostní systém</b>	<b>ko:Znalostní_systém</b>
bibliografický zdroj (FRBR)	ko:Bibliografická_entita_FRBR
dílo (FRBR)	ko:Dílo_FRBR
dokument	ko:Dokument
folksonomie	ko:Folksonomie
indexový soubor / index	ko:Indexový_soubor
informace	ko:Informace
informatická ontologie	ko:Informatická_ontologie
jednotka (FRBR)	ko:Jednotka_FRBR
klasifikační schéma	ko:Klasifikační_schéma
provedení (FRBR)	ko:Provedení_FRBR
<b>řízený slovník</b>	<b>ko:Řízený_slovník</b>
<b>slovník</b>	<b>ko:Slovník</b>
<b>systém organizace znalostí</b>	<b>ko: Systém_organizace_znalostí</b>
tezaurus	ko:Tezaurus
vyjádření (FRBR)	ko:Vyjádření_FRBR
<b>výrok</b>	<b>ko:Výrok</b>
<b>zaznamenaná znalost</b>	<b>ko:Zaznamenaná_znalost</b>
znalost	ko:Znalost
<b>znalostní báze</b>	<b>ko:Znalostní_báze</b>

#### Objekt (DUL:Object)

<b>koncový uživatel</b>	<b>ko:Koncový_uživatel</b>
<b>tvůrce</b>	<b>ko:Tvůrce</b>
<b>zprostředkovatel obsahu</b>	<b>ko:Zprostředkovatel_obsahu</b>

#### Událost (DUL:Event)

analýza	ko:Analýza
doménová analýza	ko:Doménová_analýza
fasetová analýza	ko:Fasetová_analýza
<b>identifikace</b>	<b>ko:Identifikace</b>
<b>indexace (zdrojů)</b>	<b>ko:Indexace</b>
<b>kategorizace / klasifikace</b>	<b>ko:Kategorizace_klasifikace</b>
konceptualizace	ko:Konceptualizace
<b>obsahová analýza</b>	<b>ko:Obsahová_analýza</b>
odkaz / reference	ko:Odkaz_reference
odlišení (zdrojů)	ko:Odlišení
<b>organizace znalostí (proces)</b>	<b>ko:Organizace_znalostí_proces</b>

popis (proces)	ko:Popis_proces
<b>přístup (ke znalostem)</b>	<b>ko:Přístup_ke_znalostem</b>
reprezentace (znalostí)	ko:Reprezentace_znalostí
<b>řazení</b>	<b>ko:Řazení</b>
seskupování (zdrojů)	ko:Seskupování
syntéza	ko:Syntéza
<b>umístění</b>	<b>ko:Umístění</b>
využívání (znalostí)	ko:Využívání_znalostí
<b>začlenění do kontextu</b>	<b>ko:Začlenění_do_kontextu</b>

### Kvalita (DUL:Quality)

<b>aboutness</b>	<b>ko:Aboutness</b>
asociace	ko:Asociace
<b>ekvivalence</b>	<b>ko:Ekvivalence</b>
generická hierarchie	ko:Generická_hierarchie
hierarchie	ko:Hierarchie
<b>homonymie</b>	<b>ko:Homonymie</b>
instanční hierarchie	ko:Instanční_hierarchie
<b>komunikovatelnost</b>	<b>ko:Komunikovatelnost</b>
<b>nalezitelnost / vyhledatelnost</b>	<b>ko:Nalezitelnost</b>
paradigmatický vztah	ko:Paradigmatický_vztah
partitivní hierarchie	ko:Partitivní_hierarchie
relevance	ko:Relevance
<b>srozumitelnost</b>	<b>ko:Srozumitelnost</b>
<b>synonymie</b>	<b>ko:Synonymie</b>
syntagmatický vztah	ko:Syntagmatický_vztah
vlastnost	ko:Vlastnost
vztah	ko:Vztah
<b>warrant</b>	<b>ko:Warrant</b>
<b>zkreslení / zaujetí (bias)</b>	<b>ko:Zkreslení_zaujetí</b>

### Kontext (LMM\_L1:Context)

<b>kognitivní kontext</b>	<b>ko:Kognitivní_kontext</b>
<b>metadata</b>	<b>ko:Metadata</b>
organizace znalostí (obor)	ko:Organizace_znalostí_obor
popis (zdroj)	ko:Popis_zdroj
<b>realizační kontext</b>	<b>ko:Realizační_kontext</b>

### Logická hierarchie predikátů

Hierarchie predikátů sleduje hierarchii tříd v jádrové ontologii, jež představují jejich definiční obor. Predikáty byly dále pro účely prezentace logicky strukturovány do tří skupin:

1. Sémiotické vztahy – vztahy v sémiotickém trojúhelníku, tj. vztahy reprezentace, konceptualizace a vyjádření.

<kódování>  
 <interpretace>  
 <konceptualizace>  
 <vyjádření>  
 <předmět>  
 <realizace>  
 <typ>  
 <významová souvislost>

2. Kontextové vztahy – v této skupině jsou sloučeny vlastní syntagmatické vztahy a kontextové vztahy z *DOLCE*.

- <členství>
- <hledisko\_doména>
- <místo>
- <následnost>
- <nastavení\_uspořádání>
- <účast>
- <uvedení>
- <verze>

3. Paradigmatické vztahy – skupina zahrnuje relativně trvalé, ‚věcné‘, ‚nesémiotické‘ vztahy.

- <asociace>
  - <asymetrická\_asociace>
    - <adaptace>
    - <cíl>
    - <funkce>
    - <jazyk>
    - <kvalita>
    - <metoda>
    - <nástroj>
    - <objekt>
    - <odvození>
    - <opětovné\_použití>
    - <popis>
    - <poznámka>
    - <princip>
    - <problém>
      - <lingvistický\_problém>
      - <pragmatický\_problém>
      - <sémantický\_problém>
      - <technologický\_problém>
    - <produkt\_artefakt>
    - <sumarizace>
    - <vztahy\_bibliografických\_entit>
      - <ilustrace>
      - <realizace>
      - <ztělesnění>
    - <časový\_údaj>
    - <autorství>
    - <schéma>
  - <symetrická\_asociace>
- <ekvivalence>
  - <preferance\_označení>
  - <překlad>
- <hierarchie>
  - <generický\_vztah>
  - <instanční\_vztah>
  - <partitivní\_vztah>

## 5.4 Postup návrhu a implementace

V průběhu návrhu a implementace ontologie byl aplikován iterativní postup, během něžž byl kombinován postup ‚zdola nahoru‘ (*bottom-up*) s postupem ‚shora dolů‘ (*top-down*).

Úvodní analytická fáze zahrnuje provedení doménové analýzy, která je dokumentována v kapitole 4. Výsledný model domény byl použit jako zdroj pro vytvoření seznamu vlastních pojmů pro jádrovou ontologii.

K implementaci ontologie byl použit editor *Protégé* v desktopové verzi 5.2.0, jenž umožňuje formalizaci v jazyce *OWL*. Po vložení vlastních pojmů do *Protégé* byl proveden import základní ontologie *DOLCE*, konkrétně její verze *Lexical MetaModel Level 2*, jež zahrnuje i ontologie *DOLCE+DnS Ultralite*, *Information Objects Ontology* a *Lexical MetaModel Level 1*. Následně byl proveden výběr tříd *DOLCE* relevantních pro doménu organizace znalostí a vlastní pojmy byly integrovány s pojmy základní ontologie. Byly posouzeny sémantické vztahy pojmů slučovaných ontologií a v případě významové ekvivalence byla zpravidla zrušena vlastní třída a byla ponechána třída ze základní ontologie. Současně s tím bylo provedeno mapování tříd a predikátů na externí ontologie a další zdroje pomocí predikátů datového slovníku *SKOS*.

Po dokončení strukturování jádrové ontologie byl doplněn modul terminologické ontologie. Klíčovým zdrojem pojmů pro terminologickou ontologii byla první verze ontologie organizace znalostí. Terminologická ontologie je rovněž implementována v editoru *Protégé* a její relativní samostatnost je vyjádřena i vlastním jmenným prostorem. Každá třída nejvyšší úrovně v terminologické ontologii je definována jako podtřída jedné ze tříd jádrové ontologie.

K prezentaci ontologie je použita online aplikace *WebProtégé* na <https://webprotege.stanford.edu/>. Ukázky implementace ontologie v obou verzích *Protégé* jsou v příloze 5.

## 5.5 Dílčí závěry

### 5.5.1 Realizace případu užití 1.1 – opětovné použití zaznamenaných doménových znalostí

Pro účely testování tohoto případu užití byla zpracována verifikační případová studie, jejímž cílem bylo ověřit možnost použití ontologie k tvorbě výroků ve znalostní bázi organizace znalostí. Ověřována byla úplnost (rozsah) ontologických tříd a predikátů a jejich vyjadřovací (sémantická) síla.

Pro verifikační studii byly zvoleny stejné výroky, jež byly v kapitole 3.4.1 použity pro demonstraci pracovního postupu při naplnění znalostní báze instancemi. Cílem bylo umožnit srovnání první a druhé verze ontologie. Výroky jsou voleny tak, aby obsahovaly instance ze všech tří modulů znalostní báze, tj. výroky týkající se bibliografických entit, pojmenovaných entit i abstraktní znalosti zahrnuté v modulu pojmů.

Neformálně vyjádřené výroky:

*Shiyali Ramamrita Ranganathan je autorem díla Prolegomena to library classification (1967) a je tvůrcem Dvojtečkové klasifikace, založené na principu faset. Na jeho myšlenky navazuje CRG (Classification Research Group), která byla ustavena v roce 1952 a věnovala se zkoumání fasetové analýzy. Henry Evelyn Bliss (1870–1955) je tvůrcem Bibliografického třídění. Člen CRG Jack Mills a Bliss Classification Association (BCA) jsou editoři 2. revidovaného vydání Blissova Bibliografického*

třídění (BC2), jež je založeno na principu faset a stalo se praktickou realizací (je výsledkem) teorií CRG. Faseta je kategorie entit vytvořená podle jedné klasifikační charakteristiky (principium divisionis). Obsah fasety je specifická a samostatná množina pojmů, patřících k jedné základní kategorii, jež umožňuje organizovat obsah podle kritéria členění daného těmito základními kategoriemi.

Poznámka: Výše uvedené výroky jsou ukázkou principu AAA (*anyone can say anything about anything* – viz kapitola 1.5.5) – o stejném tématu jsou zpravidla k dispozici různá tvrzení, jež nemusí být navzájem konzistentní, ale nemusí se ani vzájemně vylučovat.

Výroky v přirozeném jazyce byly následně analyzovány, což zahrnovalo dekompozici na subjekty, predikáty a objekty, sjednocení terminologie, přiřazení subjektů a objektů k ontologickým třídám, přiřazení predikátů k ontologickým predikátům.

Subjekty, objekty a predikáty odvozené z výše uvedených výroků:

Bibliografické třídění	je instancí třídy
Bibliografické třídění BC1	má označení
Bibliografické třídění BC2	má ekvivalentní entitu
Bliss Classification Association (BCA)	je založeno na
Classification Research Group (CRG)	má realizaci
dílo	je autorem
Dvojtečkové třídění (CC)	je editorem
faseta	má předmět
fasetová analýza	má nástroj
Henry Evelyn Bliss	je členem
Jack Mills	má definici
klasifikační schéma	má poznámku
Prolegomena to library classification	mapování
Shiyali Ramamrita Ranganathan	

Následně byla vytvořena množina výroků ve formátu trojic *RDF*, v nichž byly použity třídy a predikáty z druhé verze ontologie organizace znalostí. Kompletní přehled formalizovaných výroků je v příloze 6.

Příklad formalizovaných výroků pro modul bibliografických entit:

Bibliografické třídění <je instancí třídy> Dílo  
 Bibliografické třídění <má realizaci> Bibliografické třídění BC2  
 Bibliografické třídění <má předmět> <http://www.isko.org/kolit.php?cl=045> (*URI klasifikačního znaku třídění KO Literature*)  
 Bibliografické třídění <je předmětem>  
[http://cs.dbpedia.org/page/Blissovo\\_t%C5%99%C3%ADd%C4%9Bn%C3%AD](http://cs.dbpedia.org/page/Blissovo_t%C5%99%C3%ADd%C4%9Bn%C3%AD) (*URI článku v DBpedii*)  
 Bibliografické třídění <má ekvivalentní entitu> <http://viaf.org/viaf/183040827> (*identifikátor díla Bibliografické třídění v bázi VIAF*)

Příklad formalizovaných výroků pro modul pojmenovaných entit:

Jack Mills <je členem> Classification Research Group  
 Jack Mills <je editorem> Bibliografické třídění BC2

Příklad formalizovaných výroků pro modul pojmů:

Bibliografické třídění <má (používá) nástroj> faseta

faseta <má definici> Faseta je kategorie entit vytvořená podle jedné klasifikační charakteristiky. klasifikační schéma <je mapováno> <https://psh.techlib.cz/skos/PSH6484> (URI pojmu v Polytematickém strukturovaném hesláři)

Výsledkem verifikační případové studie je potvrzení, že ontologie disponuje relevantními třídami i predikáty pro tvorbu všech tří typů výroků ve znalostní bázi.

### 5.5.2 Realizace případu užití 1.2 – odvozování a objevování nových znalostí

Možnosti a typy inference v druhé verzi ontologie jsou vymezeny stupněm její formalizace a typem reprezentovaných znalostí, kterým jsou deklarativní znalosti. Odvozování je tedy založeno na strukturních vztazích v modelu znalostní báze a na vlastnostech definovaných pro jednotlivé predikáty. V jádrové ontologii, jež je založena na formalizované a axiomatizované základní ontologii *DOLCE*, se uplatňuje jak tranzitivita vlastností, tak tranzitivita vztahů.

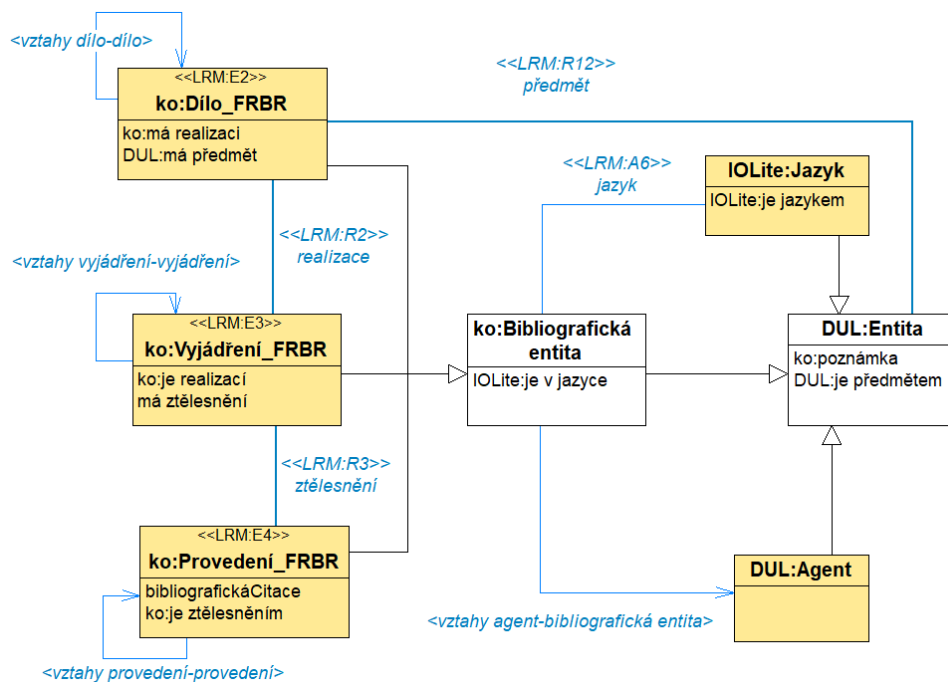
Tranzitivita vlastností spočívá v tom, že entity, které jsou podřazené v generickém vztahu, dědí vlastnosti entit nadřazených (např. vztah ‚je typ role‘ má kromě svých specifických vlastností i vlastnosti nadřazeného vztahu ‚je typ‘, třída ‚Dílo‘ má všechny vlastnosti třídy ‚Bibliografické entity‘).

Vztah, který je tranzitivní, platí i pro entity v řetězci následujícím za tímto vztahem (například provedení ‚ztělesňuje‘ nejen vyjádření, ale i dílo, jež je vyjádřením realizováno). Vztahy, které jsou tranzitivní, mají tuto vlastnost explicitně uvedenou ve své specifikaci. Na výročích z verifikační případové studie popsané v předchozí části lze dokumentovat, jak tranzitivita vztahů umožňuje odvozování. Díky tomu, že je tranzitivně vyjádřen vztah Dvojtečkového třídění a Bibliografického třídění, jež obě používají fasety, lze odvodit, že panuje vztah asociace i mezi tvůrci těchto klasifikací H. E. Blissem a S. R. Ranganathanem, i když znalostní báze neobsahuje explicitní informaci o tomto vztahu ve formě přímé asociace obou tvůrců.

Každý predikát má určen svůj definiční obor a obor hodnot, pro něž se rovněž uplatňuje tranzitivita. Použití daného predikátu ve výroku (tj. v trojici *RDF*) umožňuje odvodit, že definiční obor a obor hodnot stanovený pro daný predikát se vztahuje i na subjekty a objekty v této trojici. Příklad z verifikační studie: Protože definičním oborem predikátu ‚IOLite:je\_autorem‘ je třída *DUL:Agent* a oborem hodnot je třída *DUL:Informační entita*, lze z výroku ‚Shiyali Ramamrita Ranganathan <je autor> Dvojtečkového třídění‘ odvodit, že Shiyali Ramamrita Ranganathan <je instancí třídy> *Agent* a Dvojtečkové třídění <je instancí třídy> *Informační entita*.

### 5.5.3 Realizace případu užití 2.1 – vyjádření obsahu (indexace) zdrojů

Pro zajištění realizace tohoto případu užití byly vytvořeny dva návrhové vzory. V obou případech se řeší vyjádření obsahu, tj. sémantika. Vzory pro vyjádření obsahu popisovaných zdrojů a pro označení, kategorizaci a vyjádření časových údajů ve výročích pro znalostní bázi modelují tyto situace za použití tříd a predikátů ontologie *DOLCE* a současně jsou v souladu s koncepcí entit, atributů a vztahů v referenčním pojmovém modelu *IFLA LRM*. V diagramech na obrázcích 25 a 26 jsou třídy a predikáty uvedeny jednotně s českými názvy a jsou rozlišeny svými prefixy (např. *DUL:* nebo *ko:*).



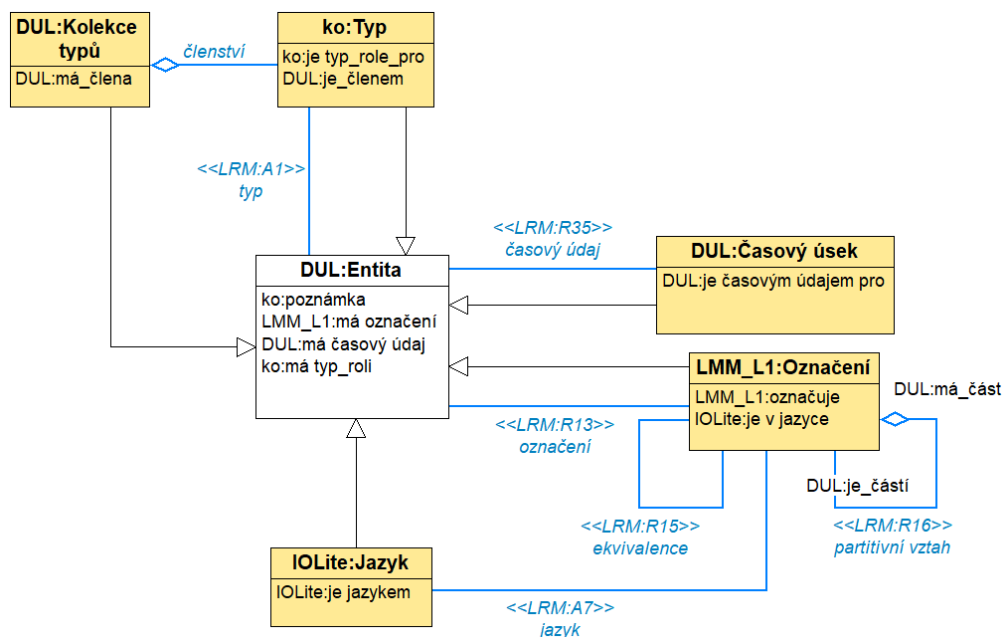
Obr. 25 Návrhový vzor pro popis zdrojů

Klíčovými třídami v návrhovém vzoru pro popis zdrojů jsou třídy založené na modelu *IFLA LRM*: dílo, vyjádření a provedení. Jak bylo uvedeno v kapitole 2.1, v praxi ontologického inženýrství jsou obvyklou součástí případů užití kompetenční otázky, jež formulují řešené problémy a dotazy, na něž má systém poskytovat odpověď. Pro ověření použitelnosti návrhového vzoru pro popis zdrojů byly použity tyto kompetenční otázky:

- Jaký je předmět díla X?
- Která díla zahrnují předmět Y?
- V jakém jazyce je bibliografická entita Z?
- Které bibliografické entity jsou k dispozici v jazyce W?
- S kterými bibliografickými entitami souvisí agent B?
- Kterí agenti souvisejí s bibliografickou entitou C?
- S kterými dalšími díly souvisí dílo X?
- S kterými dalšími vyjádřeními souvisí vyjádření V?
- S kterými dalšími provedeními souvisí provedení P?
- Která vyjádření realizují dílo X?
- Která provedení ztělesňují vyjádření V?
- Které vyjádření ztělesňuje provedení P?
- Které dílo realizuje vyjádření V?
- Jaké další popisné údaje jsou k dispozici ke zdroji s označením / identifikátorem?

Návrhový vzor pro označení, kategorizaci (typování) a časové údaje byl vytvořen pro umožnění zodpovězení následujících kompetenčních otázek:

- Jaké je označení entity E?
- Která entita je genericky nadřazená entitě E?
- Z jakých částí se skládá označení O?
- Do kterého označení patří část C?
- V jakém jazyce je označení O?
- Do které kolekce typů patří typ T?
- Jaké typy obsahuje kolekce typů T?



Obr. 26 Obecný návrhový vzor pro označení, typování a časové údaje

Funkčnost návrhového vzoru pro kategorizaci (typování) je založena na třídě Kolekce typů (DUL:TypeCollection), jež je v ontologii *DOLCE* definována jako kolekce, jejíž členové jsou maximální množinou individuí, jež sdílí stejný (pojmenovaný) typ, např. ‚editoři‘, ‚vydavatelé‘. Tato třída je velmi užitečná pro použití v návrhových vzorech typu ‚ClassesAsValues‘ (viz kapitola 1.5.3), kdy se instance třídy používají jako hodnoty predikátů.

Pro obecný způsob označení jakékoli entity byla použita třída Označení (LMM\_L1:Expression), která je součástí návrhového vzoru *LMM*, který modeluje klíčové komponenty sémiotického trojúhelníku (viz kapitola 5.3.1).

#### 5.5.4 Realizace případu užití 3.1 – sémantické mapování ontologických tříd a predikátů

Ontologie organizace znalostí je mapována do 16 externích ontologií/zdrojů, jejichž jmenné prostory jsou uvedeny v příloze 1.

Přehled mapovaných externích zdrojů je rozdělen do tří skupin: informační ontologie (tj. datové modely a metadatová schémata), ‚lehké‘ terminologické ontologie a další relevantní zdroje, které v současné době nemají charakter ontologie.

##### 1. informační ontologie

- Referenční pojmový model CIDOC-CRM
- Termíny Dublin core (DCMI metadata terms)
- Referenční pojmový model FRBR (FRBRer model)
- Datový model tezauru (ISO 25964 SKOS extension)
- Aplikační profil NKOS AP
- Metadatové schéma Schema.org
- Datový slovník Simple knowledge organization system (SKOS)



## **2. terminologické ontologie**

GoodRelations

MARC code list for relators

Multilingual dictionary of cataloguing terms and concepts (MulDiCat)

Polytematický strukturovaný heslář (PSH)

WikiData

## **3. ostatní (jiné zdroje než ontologie)**

Národní věcné autority

ISO 5127-2017 *Information and documentation – Foundation and vocabulary*

Česká terminologická databáze knihovnictví a informační vědy (TDKIV)

Klasifikace KO Literature

Celkem bylo provedeno 354 mapování za použití predikátů datového slovníku SKOS, z toho je 80 tranzitivních mapování pomocí predikátu skos:exactMatch. Další mapování použitelná v dotazech jsou provedena pomocí predikátů skos:mappingRelation (9), skos:broadMatch (22), skos:closeMatch (153), skos:narrowMatch (65) a skos:relatedMatch (16).

## Závěr

*Všechny modely jsou špatné; některé modely jsou užitečné.*

George Edward Pelham Box (1919–2013)

Cílem disertační práce bylo specifikovat pojmovou bázi domény organizace znalostí prostřednictvím návrhu doménové ontologie. Dílčí cíle práce byly stanoveny formou tří požadavků specifikujících oblasti užití ontologie organizace znalostí: tvorba výroků pro znalostní bázi organizace znalostí, indexace zdrojů v doméně a propojení s věcně relevantními ontologiemi.

Základní metodou použitou v provedeném teoretickém výzkumu byla kvalitativní metoda doménové analýzy. S využitím této metody byl vytvořen model domény organizace znalostí, který se stal obsahovým východiskem pro návrh doménové ontologie.

Aplikovaný výzkum směřující k návrhu ontologie probíhal ve dvou etapách. První etapa proběhla v letech 2013–2015 v rámci řešení grantového projektu *NAKI Znalostní báze pro obor organizace informací a znalostí* a jejím výsledkem byla první verze ontologie. Při konstrukci první verze ontologie byl uplatněn postup ‚od nuly‘ směrem *bottom-up*, tj. od shromáždění termínů přes konceptualizaci k strukturování a fasetizaci. Návrh byl realizován řešitelským týmem grantového projektu *NAKI* za použití strukturované metodiky *UPON*. V roce 2014 byla první verze ontologie implementována do funkčního prototypu znalostní báze organizace znalostí, který je online dostupný z <http://ko.cuni.cz>.

Druhou etapu v letech 2016–2017 jsem realizovala individuálně a v jejím průběhu byla navržena druhá verze ontologie s uplatněním metody reengineeringu první verze a opětovného použití základní ontologie *DOLCE*. Druhá verze ontologie byla vytvořena metodou *top-down*: konceptualizací výroků získaných doménovou a fasetovou analýzou organizace znalostí byly získány ‚nejvyšší‘ pojmy, které byly integrovány s pojmy základní ontologie *DOLCE*. Konkrétně byly implementovány její vybrané třídy nejvyšší úrovně a predikáty včetně s nimi spojených axiomů. Výsledkem jsou dva propojené moduly – axiomatizovaná jádrová ontologie a terminologická ontologie.

Přínosy práce lze rozdělit do tří okruhů. První okruh představuje vytvoření použitelného artefaktu – ontologie organizace znalostí. Druhým okruhem jsou přínosy v oblasti metodologie a třetí okruh tvoří teoretické přínosy.

### **Ontologie organizace znalostí (artefakt)**

Obsah navržené ontologie organizace znalostí představuje explicitní specifikaci pojmového modelu domény organizace znalostí. Všechny třídy mají označení v češtině, což je možné považovat za krok směrem k řešení kompatibility české a světové oborové terminologie. Ontologie je použitelná ve všech základních oblastech užití informatických ontologií: komunikace znalostí, opakované použití a organizace znalostí, popis skutečnosti pro její zpracování/podporu počítačovými systémy, automatické odvozování znalostí. Splnění požadavků na druhou verzi ontologie bylo testováno prostřednictvím realizace jednotlivých případů užití. Pro realizaci případu užití 2.1 byly vytvořeny dva návrhové vzory: návrhový vzor pro popis zdrojů a návrhový vzor pro označení, kategorizaci (typování) a časové údaje. Pro ověření realizace případu užití č. 1.1 byla zpracována verifikační případová studie. Ověřována byla úplnost (rozsah) ontologických tříd a predikátů a jejich vyjadřovací (sémantická) síla. Lze konstatovat, že ve všech případech testování byla potvrzena použitelnost ontologie pro

specifikovaný účel.

Ontologie má modulární strukturu, je formalizována v jazyce *OWL* a je tak připravena pro použití v síti propojených otevřených dat na sémantickém webu. Řešení architektury ontologie je opakovaně použitelné pro další doménové ontologie. Pro implementaci ontologie byl použit editor *Protégé*. Ontologie je online dostupná na <http://webprotege.stanford.edu> (username: O0Z, password: demo).

Architekturu ontologie tvoří dva vzájemně propojené moduly:

1. Jádrová ontologie s nejobecnějšími třídami a predikáty, jež jsou integrovány s importovanými třídami a predikáty základní ontologie *DOLCE* včetně s nimi souvisejících axiomů.
2. Terminologická ontologie obsahující detailní terminologii domény.

Jádrová ontologie umožňuje díky axiomům importovaným z *DOLCE* netriviální odvozování a lze ji tedy považovat za těžkou ontologii. Má rozsah 154 tříd (z toho je 83 tříd opětovně použitých ze základní ontologie *DOLCE* a 71 vlastních tříd) a 201 predikátů (z toho je 137 predikátů opětovně použitých ze základní ontologie *DOLCE* a 64 vlastních predikátů). Jádrová ontologie navíc díky mapování na externí ontologie umožňuje interoperabilitu v rámci sémantického webu. Je provedeno 345 mapování na externí entity z 16 jmenných prostorů.

Terminologická ontologie se soustředí na extenzivní zachycení pojmové báze oboru a spadá do kategorie lehkých ontologií. Má rozsah 300 tříd a predikáty sdílí s jádrovou ontologií. Vrcholové třídy terminologické ontologie jsou propojeny s třídami jádrové ontologie.

### **Metodologický přínos**

Zdokumentované uplatnění metody doménové analýzy je prvním případem uplatnění této metody na doménu organizace znalostí jak v českém prostředí, tak v mezinárodním měřítku. Použitý postup kombinuje systémový přístup a vybraná hlediska doménové analýzy zformulovaná Birgerem Hjørlandem. Výsledkem doménové analýzy je model domény organizace znalostí, který byl vytvořen v několika stupních formalizace: textový model domény, semiformalizovaný model ve formě výroků a fasetové klasifikace klíčových pojmů, vizualizace pojmů a jejich vztahů v pojmové mapě. Použitelnost metody doménové analýzy byla potvrzena začleněním jejího výsledku, tj. vytvořeného modelu, do modulu jádrové ontologie organizace znalostí.

V rámci řešení grantového projektu *NAKI*, zaměřeného na tvorbu znalostní báze organizace znalostí, byl implementován modul bibliografických entit znalostní báze o rozsahu 1 188 instancí. Tento modul představuje dílčí příspěvek k praktickému ověření aplikace referenčního pojmového modelu *FRBR*, resp. *IFLA LRM* na tvorbu popisných metadat. Významné zkušenosti přineslo zejména řešení metadatového popisu systémů organizace znalostí podle modelu *FRBR*, jehož výsledky byly publikovány v odborné studii.<sup>392</sup>

### **Teoretický přínos**

Za teoretický přínos práce, který je příspěvkem k rozvoji teorie oboru organizace znalostí a informační vědy, považujeme zpracovaný model domény organizace znalostí a dílčí teoretické poznatky, jež přinesl teoretický výzkum realizovaný v průběhu řešení tématu práce.

---

<sup>392</sup> BRATKOVÁ, Eva, KUČEROVÁ, Helena. K otázkám metadatového popisu systémů organizace znalostí. In: *Knihovna*. 2015, **26**(1), 5-36. ISSN 1801-3252 (Print). ISSN 1802-8772 (Online).

Hlavním přínosem práce v teoretické rovině je model domény organizace znalostí, který uplatňuje rozšířené pojetí domény přesahující hranice paměťových a fondových institucí. Model specifikuje cíl organizace znalostí jako přístup ke znalostem a jejich využití. Charakterizuje hlavní témata domény ve formě problémů a principů jejich řešení. Témata organizace znalostí jsou rozčleněna do čtyř skupin: a) konstantní problémy organizace znalostí a b) jejich řešení, c) aktuální problémy organizace znalostí a d) jejich řešení. Při stanovení těchto skupin byl převzat koncept paradigmat ve vědě z teorie Thomase S. Kuhna.<sup>393</sup> Je vymezena jednotka organizace znalostí jako zaznamenaná znalost. Model obsahuje výsledky analýzy procesu organizace znalostí a jeho realizačního a kognitivního kontextu.

Je možné konstatovat, že díky provedenému teoretickému i aplikovanému výzkumu byl v českém prostředí učiněn významný krok směrem k vyrovnání poznatkových a publikačních deficitů zjištěných úvodním průzkumem. Byly zpracovány návrhy na aktualizaci české terminologie organizace znalostí, jež byly publikovány v odborné studii<sup>394</sup>, byly předloženy na redakční radě *Terminologické databáze knihovnictví a informační vědy (TDKIV)* a byly projednány se zúčastněnými odborníky na workshopu, uspořádaném v roce 2014. Kromě již zmíněné studie věnované metadatovému popisu systémů organizace znalostí byly publikovány tři odborné studie ke klíčovým problémům organizace znalostí, jež zahrnují problematiku obsahové analýzy a *aboutness*<sup>395</sup>, vymezení vlastností a typologii systémů organizace znalostí<sup>396</sup> a komplexní pohled na vztahy v organizaci znalostí<sup>397</sup>. V roce 2017 byla v nakladatelství Karolinum vydána vědecká monografie *Organizace znalostí: klíčová témata*.

Teoretický výzkum realizovaný v průběhu řešení tématu práce přinesl dílčí nové poznatky, reagující na změnu paradigmatu v organizaci znalostí:

- 1) Pracovní definice procesu organizace znalostí, jež ho vymezuje jako záměrnou činnost, spočívající v zavádění struktury do existujících zaznamenaných znalostí.
- 2) Pracovní definice systému organizace znalostí vycházející ze dvou komplementárních pohledů založených na systémovém přístupu: 1) systém organizace znalostí jako model metadat a 2) systém organizace znalostí jako řízený slovník.
- 3) Vymezení vztahu systémů organizace znalostí a inforatických ontologií, jejichž extenze považujeme za dvě rozdílné, ale částečně se překrývající množiny. Jejich průnik s relativně ekvivalentními vlastnostmi představují lehké (terminologické) ontologie a systémy organizace znalostí orientované na pojmy (klasifikace a pojmové sítě).

---

<sup>393</sup> KUHN, Thomas Samuel. *Struktura vědeckých revolucí*. Dotisk 1. vyd. Praha: OIKOYMENEH, 2008. Předmluva, s. 10. ISBN 80-86005-54-2.

<sup>394</sup> KUČEROVÁ, Helena. České termíny pro věcné pořádání informací po 40 letech: příspěvek k terminologické diskusi. In: *ProInflow: časopis pro informační vědy* [online]. 2013, 5(Speciál), [cit. 2018-01-30], s. 1-19. ISSN 1804-2406. Dostupné z: <http://www.phil.muni.cz/journals/index.php/proinflow/article/view/749/>.

<sup>395</sup> KUČEROVÁ, Helena. Co analyzujeme při obsahové analýze dokumentů? K pojmu *aboutness* v organizaci znalostí. In: *Knihovna: knihovnická revue*. 2014, 25(1), 36-54. ISSN 1801-3252 (Print). ISSN 1802-8772 (Online).

<sup>396</sup> BRATKOVÁ, Eva, KUČEROVÁ, Helena. Systémy organizace znalostí a jejich typologie. In: *Knihovna: knihovnická revue*. 2014, 25(2), 5-29. ISSN 1801-3252 (Print). ISSN 1802-8772 (Online).

<sup>397</sup> KUČEROVÁ, Helena. Vztahy informačních zdrojů: pokus o mezioborovou syntézu. In: *Knihovna: knihovnická revue*. 2015, 26(2), 5-24. ISSN 1801-3252 (Print). ISSN 1802-8772 (Online).

4) Vymezení vztahu oboru organizace znalostí a informační vědy. Obor organizace znalostí chápeme jako samostatnou inženýrskou disciplínu, jejíž vztah k informační vědě je dán shodou v předmětu zkoumání a rozdíly v metodologii – informační věda zkoumá proces organizace znalostí metodami teoretického výzkumu a organizace znalostí se zabývá způsoby efektivního řešení problémů s tímto procesem souvisejících.

5) Vymezení vztahu oboru organizace znalostí a ontologického inženýrství. Zatímco ontologické inženýrství se zabývá způsoby efektivního návrhu a správy informatických ontologií, předmětem zájmu oboru organizace znalostí je proces organizace znalostí a jeho realizační kontext, tj. zdroje, jež jsou v procesu organizace transformovány (jednotky organizace), používané metody a nástroje a produkty, jež jsou tímto procesem vytvářeny, včetně zúčastněných aktérů – osob, institucí, technologií.

Specifikace ‚styčných ploch‘ organizace znalostí, informační vědy a ontologického inženýrství ukázala v mnoha případech na shodu v řešených teoretických problémech, metodách i produktech (informatické ontologie, systémy organizace znalostí). Organizace znalostí, informační věda a ontologické inženýrství se sice uvažují jako samostatné oblasti, je však patrné, že mají významný počet společných cílů, shodují se i v předmětech svého zájmu a v principech a metodách jejich zkoumání. V souladu s panujícím trendem transdisciplinárního výzkumu předpokládáme, že všestranná integrace teorií a metodologií těchto disciplín se může stát účinným stimulem jejich dalšího rozvoje.

## Příloha 1 Přehled používaných jmenných prostorů

### Jmenné prostory a prefixy ontologie organizace znalostí

**ko:** Jádrová ontologie organizace znalostí  
jmenný prostor: <http://ko.cuni.cz/ko.owl>

**to:** Terminologická ontologie organizace znalostí  
jmenný prostor: <http://ko.cuni.cz/to.owl>

### Jmenné prostory a prefixy integrovaných modulů základní ontologie *DOLCE*

**DUL:** DOLCE+DnS Ultralite  
jmenný prostor: <http://www.ontologydesignpatterns.org/ont/dul/DUL.owl>

**IOLite:** IOL – Information Objects Ontology  
jmenný prostor: <http://www.ontologydesignpatterns.org/ont/dul/IOLite.owl>  
Poznámka: Tato verze zahrnuje i *DOLCE+DnS Ultralite*.

**LMM\_L1:** Lexical MetaModel Level 1  
jmenný prostor: [http://www.ontologydesignpatterns.org/ont/lmm/LMM\\_L1.owl](http://www.ontologydesignpatterns.org/ont/lmm/LMM_L1.owl)

**LMM\_L2:** Lexical MetaModel Level 2  
jmenný prostor: [http://www.ontologydesignpatterns.org/ont/lmm/LMM\\_L2.owl](http://www.ontologydesignpatterns.org/ont/lmm/LMM_L2.owl)  
Poznámka: Tato verze zahrnuje i *DOLCE+DnS Ultralite*, *Information Objects Ontology* a *Lexical MetaModel Level 1*.

### Jmenné prostory a používané prefixy mapovaných ontologií a dalších zdrojů

#### 1. Informační ontologie

**cidoc-crm:** CIDOC-CRM  
jmenný prostor: <http://www.cidoc-crm.org/>

**dcterms:** DCMI metadata terms  
jmenný prostor: <http://purl.org/dc/terms/>

**frbrer:** FRBRer model  
jmenný prostor: <http://iflastandards.info/ns/fr/frbr/frbrer/>

**iso-thes:** ISO 25964 SKOS extension  
jmenný prostor: <http://purl.org/iso25964/skos-thes>

**nkos:** NKOS AP  
jmenný prostor: <http://w3id.org/nkos/nkostype#>

**schema:** Schema.org  
jmenný prostor: <http://schema.org/>

**skos:** Simple knowledge organization system (SKOS)  
jmenný prostor: <http://www.w3.org/2004/02/skos/core#>

## 2. Terminologické ontologie

**gr:** GoodRelations

jmenný prostor: <http://purl.org/goodrelations/v1#>

**mrel:** MARC code list for relators

jmenný prostor: <http://id.loc.gov/vocabulary/relators.html>

**mul:** Multilingual dictionary of cataloguing terms and concepts (MulDiCat)

jmenný prostor: <http://iflstandards.info/ns/muldicat>

**psh:** Polytematický strukturovaný heslář (PSH)

jmenný prostor: <https://psh.techlib.cz/skos/>

**wi:** WikiData

jmenný prostor: <https://www.wikidata.org/>

## 3. Ostatní (jiné zdroje než ontologie)

**aut:** Národní věcné autority

dostupné z: <http://authority.nkp.cz/vecne-autority/soubor-tematickych-autorit>

**iso-5127:** ISO 5127-2017

dostupné z: <https://www.iso.org/obp/ui/#iso:std:iso:5127:ed-2:v1:en:term:>

**tkd:** Česká terminologická databáze knihovnictví a informační vědy (TKDIV)

dostupné z: <http://tdkiv.nkp.cz/>

**kol:** KO Literature

dostupné z: <http://www.isko.org/scheme.php>

## Příloha 2 Dokumentace výběru základní ontologie pomocí nástroje *ONSET*

Ontological Commitments | Representation Language | Software Engineering Properties | Applications | Submit

Useful Tip:  
Make use of the 'Explain' button found throughout ONSET to learn more about what may apply to your ontology.

Back to Start Menu

Exit

### Ontological Commitments

You may skip unnecessary questions

Ontology of Universals/ Classes/Concepts or Particulars/Individuals? Explain

- Universals/ Classes/Concepts
- Particulars/ Individuals
- Both

Descriptive or Realist ontology? Explain

- Descriptive
- Realist (Prescriptive)
- Both

Multiplicative or Reductionist approach? Explain

- Multiplicative
- Reductionist

Actualism or Possibilism? Explain

- Actualism
- Possibilism

Concrete or Abstract Entities? Explain

- Concrete entities
- Abstract entities
- Both



Would you like to include the usage of properties and their values in your ontology?

Explain

Yes

No

Do you need to include a particular Mereology theory such as General Extensional Mereology (GEM) ?

Explain

Yes

No

Do you require a model for space and time?

Explain

Yes

No

Are temporal aspects required in your domain ontology?

Explain

Yes

No

Will your domain ontology address many different levels of granularity?

Explain

Yes

No

Do you require a foundational ontology with a three-layered architecture consisting of an abstract top level , abstract core level and basic level or a one-layered architecture consisting of a basic level only?

Explain

One-layered architecture

Three-layered architecture

Would you like to include the usage of situations and situoids in your ontology?

Explain

Yes

No

Submit

Ontological Commitments Representation Language Software Engineering Properties Subject Domain Applications Submit

Useful Tip:  
Make use of the 'Explain' button found throughout ONSET to learn more about what may apply to your ontology.

Back to Start Menu

Exit

## Representation Language

You may skip unnecessary questions

Which language will you use to represent your ontology?

- First order logic ( computer processable)
- KIF
- SUO-KIF
- OBO
- OWL DL
- OWL 2 DL
- OWL 2 EL
- OWL 2 QL
- OWL 2 RL

Submit

Ontological Commitments Representation Language **Software Engineering Properties** Subject Domain Applications Submit

Useful Tip:  
Make use of the 'Explain' button found throughout ONSET to learn more about what may apply to your ontology.

Back to Start Menu

Exit

## Software Engineering Properties

You may skip unnecessary questions

Do you require modularity?

Yes  
 No

What type of modularity would you prefer?

Availability of lighter/more expressive versions of a foundational ontology  
 Endurants and perdurants are separate  
 Modules for functions and roles  
 Built-in domain-specific ontologies

Do you intend to register your domain ontology with the OBO Foundry?

Yes  
 No

Explain

Explain

Submit

Ontological Commitments Representation Language Software Engineering Properties Applications Submit

Useful Tip:  
Make use of the 'Explain' button found throughout ONSET to learn more about what may apply to your ontology.

Back to Start Menu

Exit

## Applications

You may skip unnecessary questions

Will you use your domain ontology for any of the following application scenarios?

<input checked="" type="checkbox"/> Ontology driven information systems	Explain
<input type="checkbox"/> Database integration	Explain
<input checked="" type="checkbox"/> The semantic web	Explain
<input type="checkbox"/> Information Retrieval	Explain
<input checked="" type="checkbox"/> For scientific research	Explain
<input type="checkbox"/> To formally represent scientific theory	
<input type="checkbox"/> Ontologies for natural language processing	Explain
<input type="checkbox"/> To model methodologies and languages to be used in software applications making them more explicit	
<input type="checkbox"/> Ontological foundation of conceptual modelling	Explain
<input type="checkbox"/> Domain specific semantic wikis	
<input type="checkbox"/> Search	Explain

Submit

Ontological Commitments Representation Language Software Engineering Properties Applications **Submit**

### Submit All Answers

Calculate result

Clear result

### View Results

Based on your responses, the selected foundational ontology for you is DOLCE

Reasons why DOLCE is the selected ontology:

1. DOLCE is Descriptive in nature.
2. DOLCE takes on a Multiplicative approach - different objects may be co-localised in the same space-time.
3. DOLCE is based on Possibilism- objects are allowed independently of their actual existence.
4. DOLCE considers Concrete and abstract entities- objects that exist in time and space and objects that are outside time and space.
5. DOLCE recognises quality and quale as a separate hierarchy to represent properties and their values.
6. DOLCE is sensitive to different levels of granularity.
7. DOLCE may be represented in OWL 2 DL.
8. DOLCE offers modularity- the availability of lighter/expressive versions of DOLCE.
9. DOLCE offers modularity- endurants and perdurants are separate.
10. DOLCE offers modularity- the availability of lighter/expressive versions of DOLCE.
11. DOLCE offers modularity- endurants and perdurants are separate.
12. DOLCE has been applied to Ontology driven information systems.
13. DOLCE has been applied to The Semantic Web.
14. DOLCE has been applied to Scientific research.

### Conflicting Answers

A single foundational ontology doesn't cover all your requirements.

Features that are met by another foundational ontology

1. GFO offers modularity- modules for functions and roles.
2. GFO offers modularity- modules for functions and roles.

Features that are problematic for the selected foundational ontology

1. DOLCE doesn't offer modularity in terms of modules for functions and roles.
2. DOLCE doesn't offer modularity in terms of modules for functions and roles.

## Příloha 3 Seznam tříd jádrové ontologie organizace znalostí (páteřní taxonomie)

Poznámka: Dvojitě šipky (symbol  $\Rightarrow$ ) vyznačují vztahy polyhierarchie v páteřní taxonomii.

### DUL:Entita (*Entity*)

#### DUL:Abstraktní\_entita (*Abstract*)

ko:Jednotka\_organizace

ko:Pořádek

ko:Struktura

DUL:Formální\_entita (*FormalEntity*)

ko:Kategorie

DUL:Množina (*Set*)

DUL:Oblast (*Region*)

DUL:Kvantita (*Amount*)

DUL:Fyzická\_charakteristika (*PhysicalAttribute*)

DUL:Sociální\_charakteristika (*SocialObjectAttribute*)

DUL:Prostorová\_oblast (*SpaceRegion*)

DUL:Časoprostorová\_oblast (*SpatioTemporalRegion*)

DUL:Časový\_úsek (*TimeInterval*)

#### DUL:Informační\_entita (*InformationEntity*)

ko:Bibliografická\_entita\_FRBR

ko:Dílo\_FRBR

ko:Jednotka\_FRBR

ko:Provedení\_FRBR

ko:Vyjádření\_FRBR

ko:Informace

ko: Systém\_organizace\_znalostí

IO Lite:KOS  $\Rightarrow$  IO Lite:Datová\_struktura (*DataStructure*)

IO Lite:Lexicon  $\Rightarrow$  ko:Slovník

IO Lite:Thesaurus  $\Rightarrow$  ko:Tezaurus

ko:Folksonomie

ko:Klasifikační\_schéma

ko:Slovník

IO Lite:Lexicon  $\Rightarrow$  IO Lite:KOS

ko:Řízený\_slovník

ko:Tezaurus

IO Lite:Thesaurus  $\Rightarrow$  IO Lite:KOS

ko:Znalost

ko:Zaznamenaná\_znalost  $\Rightarrow$  DUL:Informační\_realizace (*InformationRealization*)

DUL:Informační\_objekt (*InformationObject*)  $\Rightarrow$  DUL:Sociální\_objekt (*SocialObject*)

IO Lite:Kód (*Code*)

IO Lite:Datová\_struktura (*DataStructure*)

ko:Indexový\_soubor

IO Lite:Databázové\_schéma (*DBSchema*)

IO Lite:KOS  $\Rightarrow$  ko: Systém\_organizace\_znalostí

IO Lite:Data (*Datum*)

IO Lite:Lingvistický\_objekt (*LinguisticObject*)

IO Lite:Složené\_téma (*Multiword*)

IO Lite:Slovní\_spojení (*Phrase*)

IO Lite:Slovo (*Word*)

IO Lite:Termín (*Term*)

IO Lite:Lexém (*Lexeme*)

LMM\_L2:Označení\_pojmu (*ConceptExpression*)  $\Rightarrow$

LMM\_L1:Označení (*Expression*)

LMM\_L2:Kontextové\_vyjádření (*ContextualExpression*)  $\Rightarrow$

LMM\_L1:Označení (*Expression*)

LMM\_L2:Jméno (*Name*)  $\Rightarrow$  LMM\_L1:Označení (*Expression*)

IO Lite:Text  
IO Lite:Věta (*Sentence*)  
ko:Výrok

ko:Typ

**LMM\_L1:Označení (*Expression*)**

LMM\_L2:Označení\_pojmu (*ConceptExpression*)  $\Rightarrow$  IO Lite:Termín (*Term*)

LMM\_L2:Vyjádření\_složeného\_pojmu  
(*PolyrheticConceptExpression*)

LMM\_L2:Vyjádření\_jednoduchého\_pojmu  
(*SimpleConceptExpression*)

LMM\_L2:Kontextové\_vyjádření (*ContextualExpression*)  $\Rightarrow$  IO Lite:Termín  
(*Term*)

LMM\_L2:Vyjádření\_enumerací (*EnumeratedExpression*)

LMM\_L2:Jméno (*Name*)  $\Rightarrow$  IO Lite:Termín (*Term*)

**DUL:Informační realizace (*InformationRealization*)**

ko:Dokument

IO Lite:Digitální zdroj (*DigitalResource*)

IO Lite:Digitální fotografie (*DigitalPhoto*)

IO Lite:Webová stránka (*WebPage*)

IO Lite:Multimediální objekt (*MultimediaObject*)

IO Lite:Webová stránka (*WebPage*)

ko:Informatická ontologie

IO Lite:Tělesný pohyb (*BodilyMotion*)

IO Lite:Gesto (*Gesture*)

IO Lite:Mluva (*Speech*)

ko:Zaznamenaná znalost  $\Rightarrow$  ko:Znalost

ko:Znalostní báze

ko:Znalostní systém

**LMM\_L1:Kontext (*Context*)**

ko:Kognitivní kontext

ko:Realizační kontext

DUL:Popis (*Description*)  $\Rightarrow$  DUL:Sociální objekt (*SocialObject*)

DUL:Místo (*Place*)  $\Rightarrow$  DUL:Sociální objekt (*SocialObject*)

DUL:Situace (*Situation*)  $\Rightarrow$  DUL:Sociální objekt (*SocialObject*)

LMM\_L1:Diskurzivní komunita (*KnowledgeCommunity*)  $\Rightarrow$  DUL:Komunita (*Community*)

**DUL:Kvalita (*Quality*)**

**ko:Vlastnost**

ko>Aboutness

ko:Komunikovatelnost

ko:Nalezitelnost

ko:Relevance

ko:Srozumitelnost

ko:Warrant

ko:Zkreslení zaujetí

**ko:Vztah**

ko:Paradigmatický vztah

ko:Asociace

ko:Ekvivalence

ko:Homonymie

ko:Synonymie

ko:Hierarchie

ko:Generická hierarchie

ko:Instanční hierarchie

ko:Partitivní hierarchie

ko:Syntagmatický vztah

**DUL:Objekt (*Object*)**

**DUL:Agent**

ko:Koncový uživatel

ko:Tvůrce  
 ko:Zprostředkovatel\_obsahu  
 DUL:Osoba (*Person*)  
 DUL:Fyzický\_agent (*PhysicalAgent*)  
 DUL:Sociální\_agent (*SocialAgent*) ⇒ DUL:Sociální\_objekt (*SocialObject*)  
     DUL:Kolektivní\_agent (*CollectiveAgent*)  
         DUL:Komunita (*Community*)  
             LMM\_L1:Diskurzivní\_komunita (*KnowledgeCommunity*)  
                 ⇒ LMM\_L1:Kontext (*Context*)  
         DUL:Skupina (*Group*)  
         DUL:Organizace (*Organization*)  
         DUL:Personifikace (*Personification*)

**DUL:Fyzický\_objekt (*PhysicalObject*)**  
 DUL:Fyzický\_agent (*PhysicalAgent*)

**DUL:Sociální\_objekt (*SocialObject*)**  
 DUL:Kolekce (*Collection*)  
     DUL:Kolektiv (*Collective*)  
     DUL:Konfigurace (*Configuration*)  
     DUL:Kolekce\_typů (*TypeCollection*)

**DUL:Pojem (*Concept*)** ⇒ LMM\_L1:Význam (*Meaning*)  
**DUL:Popis (*Description*)** ⇒ LMM\_L1:Kontext (*Context*), ⇒ LMM\_L1:Význam (*Meaning*)

ko:Popis\_zdroj  
 DUL:Cíl (*Goal*)  
 ko:Metadata  
 DUL:Metoda (*Method*)  
 DUL:Sociální\_norma (*Norm*)  
 DUL:Plán (*Plan*)  
     DUL:Workflow  
     IO Lite:Komunikativní\_funkce (*CommunicativeFunction*)  
 DUL:Teorie (*Theory*)  
     IO Lite:Lingvistická\_teorie (*LinguisticTheory*)  
     ko:Organizace\_znalostí\_obor  
 IO Lite:jazyk (*Language*)  
     IO Lite:Formální\_jazyk (*FormalLanguage*)  
     IO Lite:Ikonicný\_jazyk (*IconicLanguage*)  
     IO Lite:Přirozený\_jazyk (*NaturalLanguage*)

DUL:Informační\_objekt (*InformationObject*) ⇒ DUL:Informační\_entita  
 DUL:Místo (*Place*) ⇒ LMM\_L1:Kontext (*Context*)  
 DUL:Situace (*Situation*) ⇒ LMM\_L1:Kontext (*Context*), ⇒ LMM\_L1:Referent (*Reference*)  
 DUL:Sociální\_agent (*SocialAgent*) ⇒ DUL:Agent  
 LMM\_L1:Význam (*Meaning*)  
 DUL:Pojem (*Concept*) ⇒ DUL:Sociální\_objekt (*SocialObject*)  
 DUL:Popis (*Description*) ⇒ DUL:Sociální\_objekt

**LMM\_L1:Referent (*Reference*)**  
 DUL:Situace (*Situation*) ⇒ LMM\_L1:Kontext (*Context*), ⇒ DUL:Sociální\_objekt (*SocialObject*)

**DUL:Událost (*Event*)**  
**DUL:Akce (*Action*)**  
     IO Lite:Tělesný\_pohyb (*BodilyMotion*)  
     IO Lite:Gesto (*Gesture*)

**DUL:Proces (*Process*)**  
 ko:Organizace\_znalostí\_proces  
     ko:Analýza  
         ko:Doménová\_analýza  
         ko:Fasetová\_analýza  
         ko:Obsahová\_analýza  
 ko:Identifikace



ko:Indexace  
ko:Kategorizace\_klasifikace  
ko:Konceptualizace  
ko:Odkaz\_reference  
ko:Odlišení  
ko:Popis\_proces  
ko:Přístup\_ke\_znalostem  
ko:Reprezentace\_znalostí  
ko:Řazení  
ko:Syntéza  
    ko:Seskupování  
    ko:Začlenění\_do\_kontextu  
ko:Umístění  
ko:Využívání\_znalostí

## Příloha 4 Seznam predikátů ontologie organizace znalostí

V zájmu snazší čitelnosti seznamu predikátů do něj byla doplněna návěští (nadpisové položky) označující skupiny obsahově souvisejících predikátů. Návěští jsou uzavřena v ostrých závorkách a slouží pouze pro vizualizaci obsahové souvislosti predikátů, nemají *URI* ani jmenný prostor.

### owl:topObjectProperty

#### DUL:má\_asociovanou\_entitu (*associatedWith*)

##### <semiotický\_vztah>

##### <interpretace>

- LMM\_L1:má\_interpretaci (*hasInterpretation*)
- DUL:je\_klasifikováno (*isClassifiedBy*)
  - DUL:vykonává\_úlohu (*executesTask*)
  - DUL:má\_omezení (*hasConstraint*)
  - DUL:má\_rolí (*hasRole*)
    - ko:má\_cíl
    - ko:má\_funkci
      - LMM\_L2:má\_syntaktickou\_funkci (*hasSyntacticFunction*)
    - ko:má\_nástroj
    - DUL:je\_parametrizován (*isParametrizedBy*)
    - LMM\_L2:je\_instancí (*isInstanceOf*)
  - DUL:je\_pokryto (*isCoveredBy*)
  - DUL:má\_popis (*isDescribedBy*)
  - DUL:je\_členem (*isMemberOf*)
  - DUL:je\_sjednocován (*isUnifiedBy*)
  - DUL:splňuje (*satisfies*)
- LMM\_L1:je\_interpretací (*isInterpretationOf*)
- DUL:klasifikuje (*classifies*)
  - DUL:je\_omezením\_pro (*isConstraintFor*)
  - DUL:je\_vykonáván (*isExecutedIn*)
  - DUL:je\_rolí\_pro (*isRoleOf*)
    - ko:je\_cílem
    - ko:je\_funkcí
      - LMM\_L2:je\_syntaktickou\_funkcí (*isSyntacticFunctionOf*)
    - ko:je\_nástrojem\_pro
  - DUL:parametrizuje (*parametrizes*)
  - LMM\_L2:má\_instanci (*hasInstance*)
- DUL:pokrývá (*covers*)
- DUL:popisuje (*describes*)
- DUL:má\_člena (*hasMember*)
- DUL:je\_plněn (*isSatisfiedBy*)
- DUL:sjednocuje (*unifies*)

##### <kódování>

- IOLite:kóduje (*encodes*)
  - DUL:realizuje (*realizes*)
    - ko:je\_realizací
  - IOLite:je\_formálním\_termínem\_pro (*isFormalTermFor*)
  - IOLite:lexikalizuje (*lexicalizes*)
    - LMM\_L2:lexikalizuje\_ontologický\_prvek (*lexicalizesOE*)
  - IOLite:reprodukuje (*reproduces*)
    - IOLite:digitálně\_reprodukuje (*digitallyReproduces*)
- IOLite:je\_kódován (*isEncodedBy*)
  - DUL:realizuje\_se\_prostřednictvím (*isRealizedBy*)
    - ko:má\_realizaci
  - IOLite:má\_formální\_termín (*hasFormalTerm*)
  - IOLite:je\_lexikalizován (*isLexicalizedBy*)

LMM\_L2:ontologický\_prvek\_je\_lexikalizován (*OEisLexicalizedBy*)  
 IOLite:je\_reprodukován (*isReproducedBy*)  
 IOLite:je\_digitálně\_reprodukován (*isDigitallyReproducedBy*)

**<konceptualizace>**  
 DUL:konceptualizuje (*conceptualizes*)  
 DUL:je\_konceptualizován (*isConceptualizedBy*)

**<vyjádření>**  
 LMM\_L1:má\_interpretans (*hasInterpretant*)  
 DUL:konkrétně\_vyjadřuje (*concretelyExpresses*)  
 DUL:vyjadřuje (*expresses*)  
 IOLite:je\_formálním\_termínem\_pro (*isFormalTermFor*)  
 IOLite:lexikalizuje (*lexicalizes*)  
 LMM\_L2:lexikalizuje\_ontologický\_prvek (*lexicalizesOE*)  
 DUL:vyjadřuje\_pojem (*expressesConcept*)  
 LMM\_L1:je\_interpretans\_pro (*isInterpretantFor*)  
 DUL:je\_konkrétně\_vyjádřen (*isConcretelyExpressedBy*)  
 DUL:je\_vyjádřen (*isExpressedBy*)  
 IOLite:má\_formální\_termín (*hasFormalTerm*)  
 IOLite:je\_lexikalizován (*isLexicalizedBy*)  
 LMM\_L2:ontologický\_prvek\_je\_lexikalizován (*OEisLexicalizedBy*)  
 DUL:je\_vyjádřen\_pojmem (*isConceptExpressedBy*)

**<předmět>**  
 DUL:má\_předmět\_je\_o (*isAbout*)  
 IOLite:formálně\_reprezentuje (*formallyRepresents*)  
 IOLite:je\_formálně\_označen\_jako (*isAssignmentOf*)  
 LMM\_L1:označuje (*denotes*)  
 DUL:je\_předmětem (*isReferenceOf*)  
 IOLite:je\_formálně\_reprezentován (*isFormallyRepresentedIn*)  
 IOLite:je\_formálním\_označením (*isAssignedTo*)  
 LMM\_L1:má\_označení (*isDenotedBy*)

**<realizace>**  
 DUL:realizuje\_se\_prostřednictvím (*isRealizedBy*)  
 ko:má\_realizaci  
 DUL:je\_realizací\_informace\_o  
 DUL:realizuje (*realizes*)  
 ko:je\_realizací  
 DUL:realizuje\_informaci\_o

**<typ>**  
 ko:je\_typ\_role\_pro  
 ko:má\_typ\_rolí  
 LMM\_L1:související\_význam (*relatedMeaning*)

**<charakteristika>**  
 DUL:charakterizuje (*characterizes*)  
 DUL:je\_charakterizován (*isCharacterizedBy*)

**<kulturní\_základ>**  
 IOLite:má\_kulturní\_základ (*hasCulturalGrounding*)  
 IOLite:je\_kulturním\_základem\_pro (*isCulturalGroundingFor*)

**<pokrytí\_významu>**  
 DUL:pokrývá (*covers*)  
 DUL:je\_pokryto (*isCoveredBy*)

**<použití\_pojmu>**  
 DUL:je\_pojmem\_použitým\_v (*isConceptUsedIn*)  
 DUL:je\_definován\_v (*isDefinedIn*)  
 DUL:používá\_pojem (*usesConcept*)  
 DUL:definuje (*defines*)

**<specializace\_významu>**  
 DUL:je\_specializován (*isSpecializedBy*)  
 DUL:specializuje (*specializes*)

### <sjednocení\_významu>

DUL:je\_sjednocován (*isUnifiedBy*)  
DUL:sjednocuje (*unifies*)  
DUL:má\_vztah\_k\_pojmu (*isRelatedToConcept*)  
DUL:má\_parametr (*hasParameter*)  
DUL:má\_úlohu (*hasTask*)  
DUL:je\_parametrem\_pro (*isParameterFor*)  
DUL:je\_podřízen (*isSubordinatedTo*)  
DUL:je\_nadřízen (*isSuperordinatedTo*)  
DUL:je\_úlohou (*isTaskOf*)  
DUL:má\_vztah\_k\_popisu (*isRelatedToDescription*)  
DUL:rozšiřuje (*expands*)  
DUL:je\_rozšířen (*isExpandedIn*)  
IOLite:metaforicky\_spojeno\_s (*metaphoricallyBlendsWith*)  
LMM\_L2:má\_další\_interpretaci (*hasCoInterpreter*)

### <paradigmatický\_vztah>

#### <asociace>

#### <asymetrická\_asociace>

##### <adaptace>

ko:je\_adaptací  
ko:má\_adaptaci

##### <cíl>

ko:je\_cílem  
ko:má\_cíl

##### <funkce>

ko:je\_funkcí  
LMM\_L2:je\_syntaktickou\_funkcí (*isSyntacticFunctionOf*)  
ko:má\_funkci  
LMM\_L2:má\_syntaktickou\_funkci (*hasSyntacticFunction*)

##### <jazyk>

IOLite:je\_v\_jazyce (*hasRepresentationLanguage*)  
IOLite: je\_jazykem (*isRepresentationLanguageOf*)

##### <kvalita>

DUL:má\_kvalitu (*hasQuality*)  
DUL:je\_kvalitou (*isQualityOf*)

##### <metoda>

ko:je\_metodou\_pro  
ko:má\_metodu

##### <nástroj>

ko:je\_nástrojem\_pro  
ko:má\_nástroj

##### <objekt>

ko:je\_objektem  
ko:má\_objekt

##### <odvození>

ko:je\_založeno\_na  
IOLite:má\_základ (*hasGrounding*)  
ko:je\_základem\_pro  
IOLite:je\_základ\_pro (*isGroundingFor*)

##### <opětovné\_použití>

IOLite:je\_opětovně\_používán (*isReusedBy*)  
IOLite:opětovně\_používá (*reuses*)

##### <popis>

DUL:popisuje (*describes*)  
DUL:má\_popis (*isDescribedBy*)

##### <poznámka>

ko:je\_poznámkou\_k  
ko:má\_poznámku

##### <princip>

ko:je\_principem  
ko:má\_princip

**<problém>**  
ko:je\_problémem  
ko:je\_lingvistickým\_problémem  
ko:je\_pragmatickým\_problémem  
ko:je\_sémantickým\_problémem  
ko:je\_tecnologickým\_problémem  
ko:má\_problém  
ko:má\_lingvistický\_problém  
ko:má\_pragmatický\_problém  
ko:má\_sémantický\_problém  
ko:má\_tecnologický\_problém

**<lingvistický\_problém>**  
ko:je\_lingvistickým\_problémem  
ko:má\_lingvistický\_problém

**<pragmatický\_problém>**  
ko:je\_pragmatickým\_problémem  
ko:má\_pragmatický\_problém

**<sémantický\_problém>**  
ko:je\_sémantickým\_problémem  
ko:má\_sémantický\_problém

**<technologický\_problém>**  
ko:je\_tecnologickým\_problémem  
ko:má\_tecnologický\_problém

**<produkt\_artefakt>**  
ko:je\_produktem  
ko:má\_produkty

**<sumarizace>**  
ko:je\_sumarizací  
ko:má\_sumarizaci

**<vztahy\_bibliografických\_entit>**

**<ilustrace>**  
ko:je\_ilustrací  
ko:má\_ilustraci

**<realizace>**  
ko:je\_realizací  
ko:má\_realizaci

**<ztělesnění>**  
ko:je\_ztělesněním  
ko:má\_ztělesnění

**<časový\_údaj>**  
DUL:má\_časový\_údaj (*hasTimeInterval*)  
DUL:je\_časovým\_údajem\_pro (*isTimeIntervalOf*)

**<autorství>**  
IOLite:má\_autora (*hasAuthor*)  
IOLite:je\_autorem (*isAuthorOf*)

**<schéma>**  
IOLite:má\_schéma (*hasSchema*)  
IOLite:je\_schématem\_pro (*isSchemaOf*)

**<symetrická\_asociace>**  
DUL:překrývá\_se\_s (*overlaps*)  
ko:sdílí\_vlastnost

**<ekvivalence>**  
ko:má\_ekvivalentní\_entitu  
ko:je\_mapováno\_s  
ko:má\_antonymum  
IOLite:je\_kopii (*isCopyOf*)  
IOLite:je\_originálem (*isOriginalOf*)

<preference\_označení>

ko:má\_nepreferovaný\_termín

ko:má\_preferovaný\_termín

<překlad>

ko:je\_překladem

ko:má\_překlad

<hierarchie>

<generický\_vztah>

ko:je\_druhem

ko:má\_druh

<instanční\_vztah>

<element\_hodnota>

ko:je\_hodnotou

ko:má\_hodnotu

<ilustrace>

ko:je\_ilustrací

ko:má\_ilustraci

LMM\_L2:má\_instanci (*hasInstance*)

LMM\_L2:je\_instancí (*isInstanceOf*)

ko:má\_nadřazenou\_entitu

ko:je\_doplňkem

ko:je\_druhem

ko:je\_hodnotou

ko:je\_ukázkou

DUL:je\_částí (*isPartOf*)

DUL:je\_komponentou (*isComponentOf*)

DUL:je\_složkou (*isConstituentOf*)

DUL:je\_členem (*isMemberOf*)

DUL:speciáluje (*specializes*)

LMM\_L2:je\_instancí (*isInstanceOf*)

ko:má\_podřazenou\_entitu

ko:má\_doplňk

ko:má\_druh

ko:má\_hodnotu

ko:má\_ukázkou

DUL:má\_část (*hasPart*)

DUL:má\_komponentu (*hasComponent*)

DUL:má\_složku (*hasConstituent*)

DUL:má\_člena (*hasMember*)

DUL:je\_speciáluován (*isSpecializedBy*)

LMM\_L2:má\_instanci (*hasInstance*)

<partitivní\_vztah>

<doplňk>

ko:je\_doplňkem

ko:má\_doplňk

<ukázka>

ko:je\_ukázkou

ko:má\_ukázkou

DUL:má\_část (*hasPart*)

DUL:má\_komponentu (*hasComponent*)

DUL:má\_složku (*hasConstituent*)

DUL:má\_člena (*hasMember*)

DUL:je\_částí (*isPartOf*)

DUL:je\_komponentou (*isComponentOf*)

DUL:je\_složkou (*isConstituentOf*)

DUL:je\_členem (*isMemberOf*)

<kontextový\_vztah>

<členství>

DUL:je\_členem (*isMemberOf*)

DUL:má\_člena (*hasMember*)

<hledisko\_doména>  
ko:je\_hlediskem\_doménou\_pro  
ko:má\_hledisko\_doménu

<místo>  
DUL:je\_místem\_pro (*isLocationOf*)  
DUL:má\_umístění (*hasLocation*)  
DUL:je\_daleko\_od (*farFrom*)  
DUL:je\_blízko (*nearTo*)

<následnost>  
<pokračování>  
ko:je\_pokračováním  
ko:má\_pokračování  
DUL:následuje (*follows*)  
DUL:přímo\_následuje (*directlyFollows*)  
DUL:má\_předběžnou\_podmínku (*hasPrecondition*)  
DUL:je\_následnou\_podmínkou (*isPostconditionOf*)  
DUL:je\_podřízen (*isSubordinatedTo*)  
DUL:je\_předchůdcem (*precedes*)  
DUL:přímo\_předchází (*directlyPrecedes*)  
DUL:má\_následnou\_podmínku (*hasPostcondition*)  
DUL:je\_předběžnou\_podmínkou\_pro (*isPreconditionOf*)  
DUL:je\_nadřízen (*isSuperordinatedTo*)

<nastavení\_uspořádání>  
DUL:je\_nastavení\_uspořádání\_pro (*isSettingFor*)  
DUL:má\_nastavení\_uspořádání (*hasSetting*)  
DUL:je\_ve\_stejném\_nastavení\_uspořádání (*sameSettingAs*)

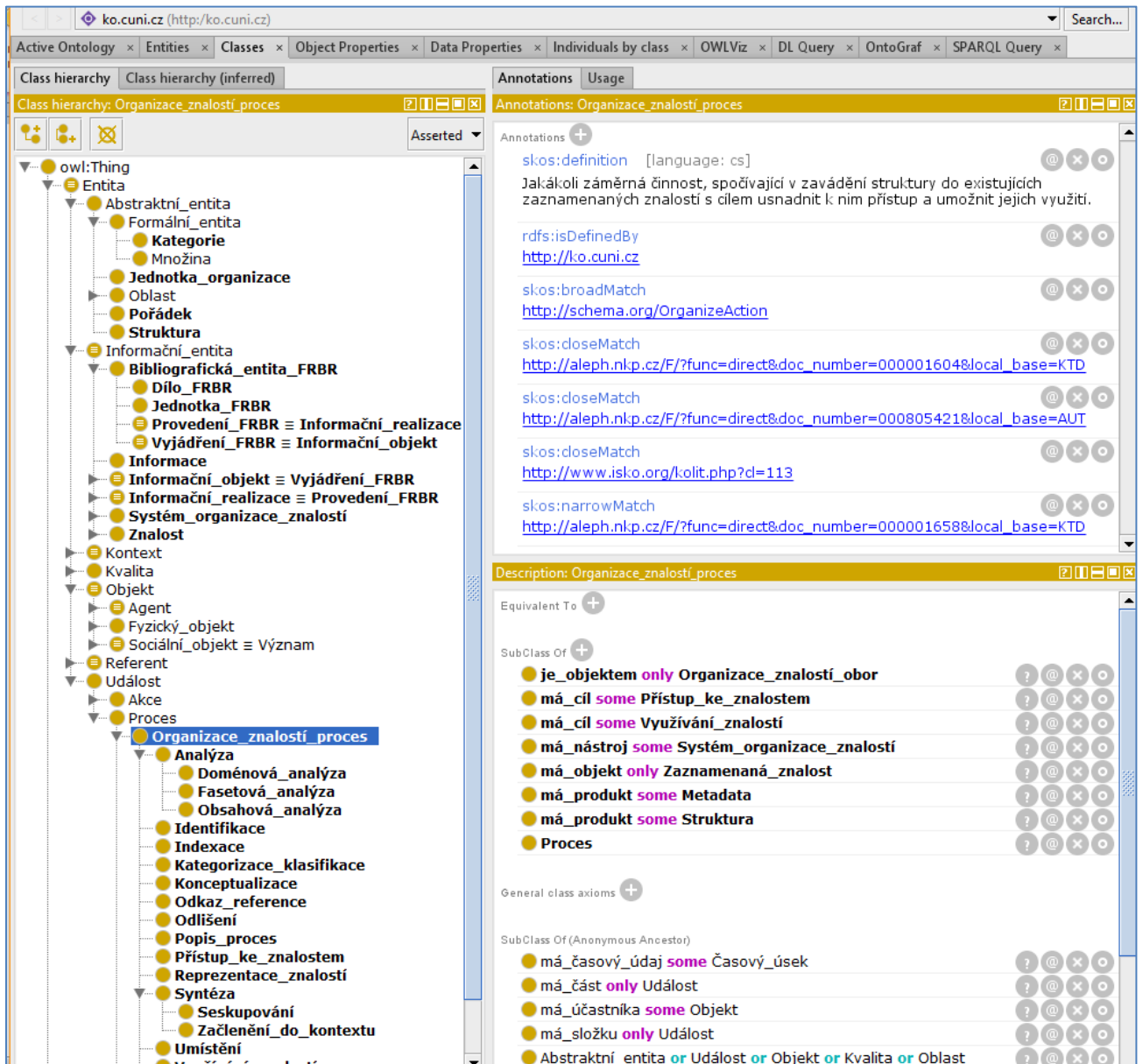
<účast>  
DUL:má\_účastníka (*hasParticipant*)  
DUL:zahrnuje\_agenta (*involvesAgent*)  
DUL:účastní\_se\_v (*isParticipantIn*)  
DUL:je\_agentem\_účastným\_v (*isAgentInvolvedIn*)  
DUL:účastní\_se\_s (*coparticipatesWith*)

<uvedení>  
DUL:je\_uveden (*isIntroducedBy*)  
DUL:uvádí (*introduces*)

<verze>  
ko:je\_verzí  
ko:má\_verzi  
LMM\_L1:je\_kontextem\_pro (*isContextOf*)  
DUL:má\_část (*hasPart*)  
DUL:má\_komponentu (*hasComponent*)  
DUL:má\_složku (*hasConstituent*)  
DUL:má\_člena (*hasMember*)  
DUL:používá\_pojem (*usesConcept*)  
DUL:definuje (*defines*)  
LMM\_L1:má\_kontext (*hasContext*)  
DUL:je\_částí (*isPartOf*)  
DUL:je\_komponentou (*isComponentOf*)  
DUL:je\_složkou (*isConstituentOf*)  
DUL:je\_členem (*isMemberOf*)  
DUL:je\_pojmem\_použitým\_v (*isConceptUsedIn*)  
DUL:je\_definován\_v (*isDefinedIn*)  
DUL:má\_společné\_hranice (*hasCommonBoundary*)

# Příloha 5 Ukázka implementace ontologie v editoru Protégé

## Implementace v Protégé desktop v. 5.2.0





## Zápis v jazyce OWL

```
<!-- http://ko.cuni.cz/Organizace_znalostí_proces -->
<owl:Class rdf:about="http://ko.cuni.cz/Organizace_znalostí_proces">
  <rdfs:subClassOf
rdf:resource="http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#Process"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="http://ko.cuni.cz/má_cíl"/>
      <owl:someValuesFrom rdf:resource="http://ko.cuni.cz/Přístup_ke_znalostem"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="http://ko.cuni.cz/má_cíl"/>
      <owl:someValuesFrom rdf:resource="http://ko.cuni.cz/Využívání_znalostí"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="http://ko.cuni.cz/má_nástroj"/>
      <owl:someValuesFrom
rdf:resource="http://ko.cuni.cz/Systém_organizace_znalostí"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="http://ko.cuni.cz/má_produkť"/>
      <owl:someValuesFrom rdf:resource="http://ko.cuni.cz/Metadata"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="http://ko.cuni.cz/má_produkť"/>
      <owl:someValuesFrom rdf:resource="http://ko.cuni.cz/Struktura"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="http://ko.cuni.cz/je_objektem"/>
      <owl:allValuesFrom
rdf:resource="http://ko.cuni.cz/Organizace_znalostí_obor"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="http://ko.cuni.cz/má_objekt"/>
      <owl:allValuesFrom rdf:resource="http://ko.cuni.cz/Zaznamenaná_znalost"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:isDefinedBy rdf:resource="http://ko.cuni.cz"/>
  <skos:broadMatch rdf:resource="http://schema.org/OrganizeAction"/>
  <skos:closeMatch rdf:resource=
"http://aleph.nkp.cz/F/?func=direct&doc_number=000001604&local_base=KTD"/>
  <skos:closeMatch
rdf:resource="http://aleph.nkp.cz/F/?func=direct&doc_number=000805421&local_base=AU
T"/>
  <skos:closeMatch rdf:resource="http://www.isko.org/kolit.php?cl=113"/>
  <skos:definition xml:lang="cs">Jakákoli záměrná činnost, spočívající v zavádění
struktury do existujících zaznamenaných znalostí s cílem usnadnit k nim přístup a umožnit
jejich využití.</skos:definition>
  <skos:narrowMatch
rdf:resource="http://aleph.nkp.cz/F/?func=direct&doc_number=000001658&local_base=KT
D"/>
</owl:Class>
```

## Implementace ve WebProtégé

Implementace je volně dostupná na <http://webprotege.stanford.edu> (username: O0Z, password: demo).

The screenshot shows the WebProtégé interface. On the left, the 'Class Hierarchy' pane displays a tree structure starting from 'owl:Thing', with 'Systém\_organizace\_znalostí' selected. On the right, the 'Class: Systém\_organizace\_znalostí' details pane shows the IRI, annotations (including skos:closeMatch and skos:definition), and classes (Informační\_entita).

## Zápis v jazyce OWL

```
<!-- http://ko.cuni.cz/Systém_organizace_znalostí -->
  <owl:Class rdf:about="http://ko.cuni.cz/Systém_organizace_znalostí">
    <rdfs:subClassOf
rdf:resource="http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#InformationEntity"/>
    <skos:closeMatch rdf:resource="http://www.w3.org/2004/02/skos/core#ConceptScheme"/>
    <skos:closeMatch rdf:resource="https://www.wikidata.org/wiki/Q6423319"/>
    <skos:definition xml:lang="cs">Schéma, modelující strukturu (tj. prvky a vzájemné
vztahy) organizované množiny znalostí. Funkcí systému organizace znalostí je podpora
procesů organizace znalostí a přístupu k znalostem. Základním strukturálním prvkem systému
organizace znalostí je pojem. Jádrem fyzické reprezentace každého systému organizace
znalostí je slovník, tj. formální vyjádření pojmů. Ten je používán jak pro vyjádření
sémantiky, tak syntaxe organizovaného celku, případně i pravidel, určujících používání
struktury.</skos:definition>
    <skos:relatedMatch
rdf:resource="http://aleph.nkp.cz/F/?func=direct&doc_number=000001625&local_base=KT
D"/>
  </owl:Class>
```

## Příloha 6 Ukázka formalizovaných výroků pro testování případu užití 1.1

Výroky byly pro snazší čitelnost zjednodušeny a obsahují pouze nejnужnější komponenty trojic *RDF*. Názvy tříd a predikátů jsou opět pro snazší čitelnost uváděny v české variantě.

### Bibliografické třídění (owl:NamedIndividual)

```
<!-- je instancí třídy Dílo -->
<rdf:type>ko:Dílo_FRBR
<!-- je instancí třídy Klasifikační schéma -->
<rdf:type>ko:Klasifikační_schéma
<IOLite:má_autora>Henry Evelyn Bliss
<!-- článek v DBpedii -->
<DUL:je_předmětem>http://cs.dbpedia.org/page/Blissovo_t%C5%99%C3%ADd%C4%9Bn%C3%AD
<!-- článek ve Wikipedii -->
<DUL:je_předmětem>https://cs.wikipedia.org/wiki/Blissovo_t%C5%99%C3%ADd%C4%9Bn%C3%AD
<!-- identifikátor v bázi VIAF -->
<ko:má_ekvivalentní_entitu>http://viaf.org/viaf/183040827
<!-- metadata v registru Bartoc -->
<ko:má_ekvivalentní_entitu>http://bartoc.org/en/node/462
<ko:má_nástroj>faseta
<!-- klasifikační znak třídění KO Literature -->
<DUL:má_předmět_je_o>http://www.isko.org/kolit.php?cl=045
<!-- dílo má realizace -->
<ko:má_realizaci>Bibliografické třídění BC1
<ko:má_realizaci>Bibliografické třídění BC2
<!-- má alternativní názvy -->
<skos:altLabel xml:lang="en">Bibliographic classification
<skos:altLabel xml:lang="en">BC
```

### Bibliografické třídění BC1 (owl:NamedIndividual)

```
<!-- je instancí třídy Vyjádření -->
<rdf:type>ko:Vyjádření_FRBR
<!-- vyjádření je realizací -->
<ko:je_realizací>Bibliografické třídění
<!-- má alternativní název -->
<skos:altLabel xml:lang="en">BC1
```

### Bibliografické třídění BC2 (owl:NamedIndividual)

```
<!-- je instancí třídy Vyjádření -->
<rdf:type>ko:Vyjádření_FRBR
<!-- vyjádření je realizací -->
<ko:je_realizací>Bibliografické třídění
<!-- má alternativní název -->
<skos:altLabel xml:lang="en">BC2
<ko:je_produktem>Classification Research Group
<ko:je_založeno_na>Bibliografické třídění BC1
```

### Bliss Classification Association (owl:NamedIndividual)

```
<!-- je instancí třídy Organizace -->
<rdf:type>DUL:Organizace
<je_editor>Bibliografické třídění BC2
<!-- klasifikační znak třídění KO Literature -->
<DUL:má_předmět_je_o>http://www.isko.org/kolit.php?cl=925
<!-- má alternativní název -->
<skos:altLabel xml:lang="en">BCA
```

### Classification Research Group (owl:NamedIndividual)

```
<!-- je instancí třídy Skupina -->
<rdf:type>DUL:Skupina
<DUL:má_časový_údj>1952
<ko:má_objekt>fasetová analýza
```

```

<ko:je_pokračováním>Shiyali Ramamrita Ranganathan
<!-- identifikátor v bázi ISNI -->
<ko:má_ekvivalentní_entitu>http://isni.org/isni/0000000110110690
<!-- identifikátor v bázi VIAF -->
<ko:má_ekvivalentní_entitu>http://viaf.org/viaf/133663593
<!-- klasifikační znak třídění KO Literature -->
<DUL:má_předmět_je_o>http://www.isko.org/kolit.php?cl=925
<!-- má alternativní název -->
<skos:altLabel xml:lang="en">CRG

```

#### Dvojtečkové třídění (owl:NamedIndividual)

```

<!-- je instancí třídy Dílo -->
<rdf:type>ko:Dílo_FRBR
<!-- je instancí třídy Klasifikační schéma -->
<rdf:type>ko:Klasifikační_schéma
<IOLite:má_autora>Shiyali Ramamrita Ranganathan
<ko:má_nástroj>faseta
<!-- klasifikační znak třídění KO Literature -->
<DUL:má_předmět_je_o>http://www.isko.org/kolit.php?cl=046
<!-- metadata v registru Bartoc -->
<ko:má_ekvivalentní_entitu>http://bartoc.org/en/node/862
<!-- má alternativní názvy -->
<skos:altLabel xml:lang="en">Colon Classification
<skos:altLabel xml:lang="en">CC

```

#### Faseta (owl:Class)

```

<!-- je podtřídou třídy Kategorie -->
<rdfs:subClassOf>ko:Kategorie
<!-- má definice -->
<skos:definition>Kategorie entit vytvořená podle jedné klasifikační charakteristiky
(kritéria členění).
<skos:definition>Obsahem fasety je specifická a samostatná množina pojmů, patřících k
jedné základní kategorii, jež umožňuje organizovat obsah podle kritéria členění
daného těmito základními kategoriemi.
<!-- heslo v TDKIV -->
<ko:má_ekvivalentní_entitu>
http://aleph.nkp.cz/F/?func=direct&doc_number=000001528&local_base=KTD
<!-- klasifikační znak třídění KO Literature -->
<DUL:má_předmět_je_o>http://www.isko.org/kolit.php?cl=325

```

#### fasetová analýza (owl:Class)

```

<!-- je podtřídou třídy Analýza -->
<rdfs:subClassOf>Analýza
<!-- má poznámku -->
<rdfs:comment>Dle ISO 25964-1: analýza věcných oblastí na podstatné pojmy seskupené
do faset a rozčlenění pojmů na užší pojmy podle specifikované charakteristiky
členění.
<!-- mapování -->
<skos:broadMatch>http://www.isko.org/kolit.php?cl=325
<skos:exactMatch>https://www.iso.org/obp/ui/#iso:std:iso:5127:ed-
2:v1:en:term:3.8.3.26

```

#### Henry Evelyn Bliss (owl:NamedIndividual)

```

<!-- je instancí třídy Osoba -->
<rdf:type>DUL:Osoba
<DUL:má_časový_údj>1870-1955
<!-- článek v DBpedii -->
<ko:je_předmětem>http://cs.dbpedia.org/page/Henry_Evelyn_Bliss
<!-- článek ve Wikipedii -->
<ko:je_předmětem>https://cs.wikipedia.org/wiki/Henry_Evelyn_Bliss
<IOLite:je_autorem>Bibliografické třídění
<!-- klasifikační znak třídění KO Literature -->
<DUL:má_předmět_je_o>http://www.isko.org/kolit.php?cl=922

```

#### Jack Mills (owl:NamedIndividual)

```

<!-- je instancí třídy Osoba -->

```

```
<rdf:type>DUL:Osoba
<DUL:je_členem>Classification Research Group
<je_editor>Bibliografické třídění BC2
<!-- klasifikační znak třídění KO Literature -->
<DUL:má_předmět_je_o> http://www.isko.org/kolit.php?cl=922
```

#### **Klasifikační schéma** (owl:Class)

```
<!-- je podtřídou třídy Systém organizace znalostí -->
<rdfs:subClassOf>ko:Systém_organizace_znalostí
<!-- má definice -->
<skos:definition>Nejčastěji se jedná o hierarchickou strukturu s menším počtem
vrcholových (kořenových) tříd a s větším počtem hierarchických úrovní. Základní
strukturní jednotka: vztah mezi pojmy.
<!-- má poznámku -->
<rdfs:comment>ISO 25964-1: Seznam pojmů a prekoordinovaných kombinací pojmů,
uspořádaný do tříd.
<!-- mapování -->
<skos:closeMatch>http://psh.ntkcz.cz/skos/PSH6484
<skos:closeMatch>
http://aleph.nkp.cz/F/?func=direct&doc_number=000001559&local_base=KTD
<skos:exactMatch>http://w3id.org/nkos/nkostype#classification_schema
<skos:exactMatch>https://www.iso.org/obp/ui/#iso:std:iso:5127:ed-
2:v1:en:term:3.8.5.02
```

#### **Prolegomena to library classification** (owl:NamedIndividual)

```
<!-- je instancí třídy Dílo -->
<rdf:type>ko:Dílo_FRBR
<IOLite:má_autora>Shiyali Ramamrita Ranganathan
<DUL:má_časový_údaj>1967
<!-- klasifikační znak třídění KO Literature -->
<DUL:má_předmět_je_o>http://www.isko.org/kolit.php?cl=111
<!-- metadata v katalogu WorldCat -->
<ko:má_ekvivalentní_entitu>http://www.worldcat.org/oclc/60497
<!-- metadata v katalogu NK ČR -->
<ko:má_ekvivalentní_entitu>
http://aleph.nkp.cz/F/?func=direct&doc_number=001036313&local_base=NKC
<!-- zdroj v repozitáři -->
<DUL:popisuje>http://hdl.handle.net/10150/106370
```

#### **Shiyali Ramamrita Ranganathan** (owl:NamedIndividual)

```
<!-- je instancí třídy Osoba -->
<rdf:type>DUL:Osoba
<IOLite:je_autorem>Prolegomena to library classification
<IOLite:je_autorem>Dvojtečkové třídění
<!-- klasifikační znak třídění KO Literature -->
<DUL:má_předmět_je_o>http://www.isko.org/kolit.php?cl=922
<!-- identifikátor v bázi VIAF -->
<ko:má_ekvivalentní_entitu>http://viaf.org/viaf/49268668
```

## Literatura

Následující soupis je výběrem nejvýznamnějších zdrojů, relevantních tematické disertace. Zdroje, jejichž texty či myšlenky byly použity v textu disertace, jsou uvedeny v poznámkovém aparátu. Odkazy na webové zdroje byly aktuální k 30. 1. 2018.

ABBAS, June. *Structures for organizing knowledge: exploring taxonomies, ontologies, and other schemas*. New York: Neal-Schuman Publishers, © 2010. 249 s. ISBN 978-1-55570-699-9.

AITCHISON, Jean, GILCHRIST, Alan, BAWDEN, David. *Thesaurus construction and use: a practical manual*. 4th ed. Chicago: Fitzroy Dearborn Publishers, © 2000. 218 s. ISBN 978-1-57958-273-9.

ALLEMANG, Dean, HENDLER, James. *Semantic web for the working ontologist: effective modeling in RDFS and OWL*. 2nd ed. Waltham (MA) Amsterdam: Morgan Kaufmann/Elsevier, © 2011. xiii, 354 s. ISBN 978-0-12-385966-2 (Online). ISBN 978-0-12-385965-5 (Print).

ANDERSON, James D. Organization of knowledge. In: John FEATHER, Paul STURGES, ed. *International encyclopedia of information and library science*. 2nd ed. London: Routledge, 2003, s. 471-490. ISBN 978-0-415-25901-9.

BAWDEN, David, ROBINSON, Lyn. *Úvod do informační vědy*. 1. vyd. Doubravník: Flow, 2017. 451 s. ISBN 978-80-88123-10-1.

BEAN, Carol A., GREEN, Rebecca, ed. *Relationships in the organization of knowledge*. Dordrecht: Kluwer Academic Publishers, 2011 (dotisk vyd. z r. 2001). 248 s. Information science and knowledge management, vol. 2. doi:10.1007/978-94-015-9696-1. ISBN 978-90-481-5652-8 (brož.). ISBN 978-94-015-9696-1 (Online).

BLISS, Henry Evelyn. *The organization of knowledge and the system of the sciences*. With an introduction by John Dewey. New York: Henry Holt, [© 1929]. xx, 433 s.

BLISS, Henry Evelyn. *The organization of knowledge in libraries and the subject-approach to books*. New York: H. W. Wilson, 1933. xvi, 335 s.

BORGIO, Stefano, MASOLO, Claudio. Foundational choices in DOLCE. In: STAAB, Steffen, STUDER, Rudi, ed. *Handbook on ontologies*. 2nd ed. Dordrecht: Springer, 2009, s. 361-381. doi:10.1007/978-3-540-92673-3\_16. ISBN 978-3-540-70999-2 (Print). ISBN 978-3-540-92673-3 (Online).

BROUGHTON, Vanda. *Essential classification*. London: Facet, 2004. x, 324 s. ISBN 978-1-85604-514-8.

BROUGHTON, Vanda. *Essential thesaurus construction*. London: Facet, 2006. v, 296 s. ISBN 978-1-85604-565-0.

Classification Research Group. The need for a faceted classification as the basis of all methods of information retrieval. In: CHAN, Lois Mai, RICHMOND, Phyllis A., SVENONIUS, Elaine, ed. *Theory of subject analysis: a sourcebook*. Littleton: Libraries Unlimited, 1985, s. 157-167. ISBN 0-87287-489-3. – Původně publikováno jako: UNESCO (IAC Doc. Ter. PAS) document 320/5515, 26 May 1955.

COYLE, Karen. *FRBR before and after: a look at our bibliographic models*. Chicago: ALA Editions, 2016. xv, 179 s. ISBN 978-0-8389-1345-1 (Print). ISBN 978-0-8389-1364-2 (PDF). ISBN 978-0-8389-1365-9 (ePub). ISBN 978-0-8389-1366-6 (Kindle).

CUTTER, Charles Ammi. *Rules for a dictionary catalog*. 4th ed., rewritten. Washington: Government Printing Office, 1904. 173 s.

ČSN ISO 1087-1 (01 0501). *Terminologická práce – Slovník – Část 1: Teorie a aplikace*. Praha: Český normalizační institut, 2002. 38 s.

ČSN ISO 5963 (01 0174). *Dokumentace. Metody analýzy dokumentů, určování jejich obsahu a výběru lexikálních jednotek selekčního jazyka*. Praha: Český normalizační institut, 1995. 10 s.

ČSN ISO 7154 (01 0141). *Dokumentace – Zásady bibliografického pořádku*. Praha: Český normalizační institut, 2003. 15 s.

ČSN ISO/TR 9007 (97 9702). *Systémy zpracování informací: Pojmy a terminologie pro pojmové schéma a informační základnu*. Praha: Český normalizační institut, 1996. 131 s.

DAHLBERG, Ingetraut. *Grundlagen universaler Wissensordnung: Probleme und Möglichkeiten eines universalen Klassifikationssystems des Wissens*. Pullach bei München: Verlag Dokumentation, 1974. xviii, 366 s. DGD-Schriftenreihe, Bd. 3. ISBN 978-3-7940-3623-3.

DAHLBERG, Ingetraut. International Society for Knowledge Organization (ISKO). In: Marcia J. BATES, Mary Niles MAACK, ed. *Encyclopedia of library and information sciences*. 3rd ed. Boca Raton (Florida): CRC Press, © 2010, s. 2941-2949. ISBN 0-8493-9711-1 (online). doi:10.1081/E-ELIS3-120044729.

DAHLBERG, Ingetraut. *Wissensorganisation: Entwicklung, Aufgabe, Anwendung, Zukunft*. Würzburg: Ergon Verlag, 2014. 175 s. Textbooks for knowledge organization, vol. 3. ISSN 0944-8152. ISBN 978-3-95650-065-7.

DE NICOLA, Antonio, MISSIKOFF, Michele, NAVIGLI, Roberto. A proposal for a Unified Process for Ontology Building: UPON. In: Kim Viborg Andersen, John Debenham, Roland Wagner, ed. *Database and Expert Systems Applications: 16th International conference, DEXA 2005, Copenhagen, Denmark, August 22-26, 2005: Proceedings*. Berlin: Springer, 2005, s. 655-664. Lecture Notes in Computer Science, no. 3588. doi:10.1007/11546924\_64. ISSN 0302-9743. ISBN 978-3-540-28566-3 (Print). ISBN 978-3-540-31729-6 (Online).

DENTON, William. *How to make faceted classification and put it on the web* [online]. [Personal web site]: November 2003 [cit. 2018-01-30]. Dostupné z: <http://www.miskatonic.org/library/facet-web-howto.html>.

FAIRTHORNE, Robert A. Content analysis, specification, and control. In: *Annual review of information science and technology*. Vol. 4, 1969. Carlos A. Cuadra, ed. Chicago: Encyclopaedia Britannica, 1969, s. 73-109.

FEIBLEMAN, James K. Theory of integrative levels. In: CHAN, Lois Mai, RICHMOND, Phyllis A., SVENONIUS, Elaine, ed. *Theory of subject analysis: a sourcebook*. Littleton: Libraries Unlimited, 1985, s. 136-143. ISBN 0-87287-489-3. – Původně publikováno v: *British journal for the philosophy of science*. 1954, 5, 55-66.

FOSKETT, Anthony Charles. *The subject approach to information*. 5th ed. London: Library Association Publishing, 1996 xv, 456 s. ISBN 978-1-85604-048-8.

FRICKÉ, Martin. *Logic and the organization of information*. New York: Springer, © 2012. xv, 312 s. ISBN 978-1-4614-3087-2 (váz.). ISBN 978-1-4614-3088-9 (Online).

FUGMANN, Robert. *Subject analysis and indexing: theoretical foundation and practical advice*. 1st ed. Frankfurt am Main: Indeks Verlag, 1993. xvi, 250 s. Textbooks for knowledge organization, vol. 1. ISBN 978-3-88672-500-7.

FURNER, Jonathan. FRASAD and the ontology of subjects of works. In: *Cataloging & classification quarterly*. 2012, 50(5-7), 494-516. ISSN 0163-9374. doi:10.1080/01639374.2012.681252.

GANGEMI, Aldo, PRESUTTI, Valentina. Ontology design patterns. In: STAAB, Steffen, STUDER, Rudi, ed. *Handbook on ontologies*. 2nd ed. Dordrecht: Springer, 2009, s. 221-243. [https://doi.org/10.1007/978-3-540-92673-3\\_10](https://doi.org/10.1007/978-3-540-92673-3_10). ISBN 978-3-540-70999-2 (Print). ISBN 978-3-540-92673-3 (Online).

GLUSHKO, Robert J., ed. *The discipline of organizing*. Cambridge(MA): MIT Press, 2013. 475 s. ISBN 978-0-262-51850-5 (brož.).

GOLUB, Koraljka. *Subject access to information: an interdisciplinary approach*. Santa Barbara: Libraries Unlimited, an imprint of ABC-CLIO, LLC, 2015. 165 s. ISBN 978-1-61069-577-0.

GÓMEZ-PÉREZ, Asunción, FERNÁNDEZ-LÓPEZ, Mariano, CORCHO, Oscar. *Ontological engineering: with examples from the areas of knowledge management, e-commerce and the Semantic Web*. 1st ed. London: Springer, 2004. xii, 403 s. Advanced information and knowledge processing, ISSN 1610-3947. doi:10.1007/b97353. ISBN 978-1-85233-551-9 (Print). ISBN 978-1-85233-840-4 (Online).

- GÓMEZ-PÉREZ, Asunción, ROYAS-AMAYA, Dolores. Ontological reengineering for reuse. In: Dieter Fensel, Rudi Studer, ed. *Knowledge Acquisition, Modeling and Management: 11th European Workshop, EKAW'99 Dagstuhl Castle, Germany, May 26–29, 1999 Proceedings*. Berlin, Heidelberg: Springer-Verlag, 1999, s. 139-156. doi: [https://doi.org/10.1007/3-540-48775-1\\_9](https://doi.org/10.1007/3-540-48775-1_9). Lecture notes in computer science, no. 1621, ISSN 0302-9743 (Print). ISSN 1611-3349 (Online). ISBN 978-3-540-66044-6 (Print). ISBN 978-3-540-48775-3 (Online).
- GRUBER, Thomas R. A translation approach to portable ontology specification. In: *Knowledge acquisition*. June 1993, 5(2), 199-220. doi:10.1006/knac.1993.1008. ISSN 1042-8143.
- GRUBER, Thomas R. Ontology. In: Ling LIU, M. Tamer ÖZSU, ed. *Encyclopedia of database systems*. London: Springer, 2009, s. 1963-1965. doi:10.1007/978-0-387-39940-9\_1318. ISBN 978-0-387-35544-3 (Print). ISBN 978-0-387-39940-9 (Online).
- GRUBER, Thomas R. Toward principles for the design of ontologies used for knowledge sharing. In: *International journal of human-computer studies*. November/December 1995, 43(5-6), 907-928. doi:10.1006/ijhc.1995.1081. ISSN 1071-5819.
- GUARINO, Nicola, WELTY, Christopher A. An overview of OntoClean. In: STAAB, Steffen, STUDER, Rudi, ed. *Handbook on ontologies*. 2nd ed. Dordrecht: Springer, 2009, s. 201-220. [https://doi.org/10.1007/978-3-540-92673-3\\_9](https://doi.org/10.1007/978-3-540-92673-3_9). ISBN 978-3-540-70999-2 (Print). ISBN 978-3-540-92673-3 (Online).
- GUIZZARDI, Giancarlo. *Ontological foundations for structural conceptual models*. Enschede: CTIT, Telematica Instituut, 2005. 416 s. Telematica Instituut Fundamental Research Series, no. 015 (TI/FRS/015). CTIT PhD Thesis Series, No. 05-74, ISSN 1381-3617. ISBN 90-75176-81-3. Dostupné z: <https://research.utwente.nl/en/publications/ontological-foundations-for-structural-conceptual-models> [cit. 2018-01-30].
- HABR, Jaroslav, VEPŘEK, Jaromír. *Systémová analýza a syntéza: zdokonalování a projektování systémů*. 2. přeprac. vyd. Praha: SNTL, 1986. 316 s.
- HITZLER, Pascal et al., ed. *Ontology engineering with ontology design patterns: foundations and applications*. Berlin: AKA, [2016], ©2016. xvi, 371 s. Studies on the semantic web, 025. ISBN 978-3-89838-715-6.
- HJØRLAND, Birger. Domain analysis [online]. Version 1.3. 2017-07-09 [cit. 2018-01-30]. In: *ISKO Encyclopedia of knowledge organization*. ISKO, 2017. Dostupné z: [http://www.isko.org/cyclo/domain\\_analysis](http://www.isko.org/cyclo/domain_analysis).
- HJØRLAND, Birger. Domain analysis in information science: eleven approaches – traditional as well as innovative. In: *Journal of documentation*. 2002, 58(4), 422-462. doi:10.1108/00220410210431136. ISSN 0022-0418.
- HJØRLAND, Birger. Knowledge organization (KO). In: *Knowledge organization*. 2016, 43(6), 475-484. ISSN 0943-7444. – Text odvozený z online článku HJØRLAND, Birger. Knowledge organization (KO) [online]. Version 1.0. 2016-08-01 [cit. 2018-01-30]. In: *ISKO Encyclopedia of knowledge organization*. ISKO, 2016. Dostupné z: [http://www.isko.org/cyclo/knowledge\\_organization](http://www.isko.org/cyclo/knowledge_organization).
- HJØRLAND, Birger, ALBRECHTSEN, Hanne. Toward a new horizon in information science: domain analysis. In: *Journal of the American Society for Information Science*. July 1995, 46(6), 400-425. doi:10.1002/(SICI)1097-4571(199507)46:6<400::AID-ASI2>3.0.CO;2-Y. ISSN 0002-8231 (Print). ISSN 1097-4571 (Online).
- HODGE, Gail. *Systems of knowledge organization for digital libraries: beyond traditional authority files*. Washington: The Digital Library Federation, Council on Library and Information Resources, 2000. vi, 37 s. ISBN 978-1-933645-06-3. Dostupné též z: <http://www.clir.org/pubs/reports/pub91/> [cit. 2018-01-30].
- CHAN, Lois Mai, RICHMOND, Phyllis A., SVENONIUS, Elaine, ed. *Theory of subject analysis: a sourcebook*. Littleton: Libraries Unlimited, 1985. 415 s. ISBN 0-87287-489-3.
- ICOM/CIDOC Special Interest Group. *Definition of the CIDOC Conceptual Reference Model* [online]. Produced by the ICOM/CIDOC Documentation Standards Group, continued by the CIDOC CRM Special Interest Group. Christian Emil ORE, Martin DOERR, Patrick Le BOEUF, Stephen STEAD, ed. Version 6.2. © 2003, version from May 2015 [cit. 2018-01-30]. 206 s. Dostupné z: <http://www.cidoc-crm.org/Version/version-6.2>.
- IFLA. *IFLA library reference model: a conceptual model for bibliographic information* [online]. Pat Riva,



Patrick LeBoeuf, Maja Žumer, ed. Hague: International Federation of Library Associations and Institutions, rev. August 2017 as amended and corrected through December 2017 [cit. 2018-01-30]. 101 s. Dostupné z: <https://www.ifla.org/publications/node/11412>.

IFLA Working Group on the Functional Requirements for Subject Authority Records. *Functional requirements for subject authority data (FRSAD): a conceptual model*. Marcia Lei ZENG, Maja ŽUMER, Athena SALABA, ed. Berlin: De Gruyter Saur, © 2011. vi, 74 s. IFLA series on bibliographic control, vol. 43, ISSN 1868-8438.

ISBN 978-3-11-025323-8 (Print). ISBN 978-3-11-026378-7 (Online).

*ISKO Encyclopedia of knowledge organization (IEKO)* [online]. HJØRLAND, Birger, GNOLI, Claudio, ed. ISKO Scientific Advisory Council, 2016- [cit. 2018-01-30]. Dostupné z: <http://www.isko.org/cyclo/>.

ISO 25964-1:2011. *Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval*. 1st ed. Geneva: International Organization for Standardization, 2011-08-08. 152 s.

ISO 25964-2:2013. *Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies*. 1st ed. Geneva: International Organization for Standardization, 2013-03-04. 99 s.

ISO 5127:2017. *Information and documentation – Foundation and vocabulary*. 2nd ed. Geneva: International Organization for Standardization, 2017-05. 353 s. Části normy jsou dostupné z: <https://www.iso.org/obp/ui/#iso:std:iso:5127:ed-2:v1:en> [cit. 2018-01-30].

ISO 704:2009. *Terminology work – Principles and methods*. 3rd ed. Geneva: International Organization for Standardization, 2009. 65 s.

ISO/IEC 2382:2015. *Information technology – Vocabulary* [online]. 1st ed. Geneva: International Organization for Standardization, 2015-05 [cit. 2018-01-30]. Dostupné z: <https://www.iso.org/obp/ui/#iso:std:iso-iec:2382:ed-1:v1:en>.

JOUDREY, Daniel N., TAYLOR, Arlene G., MILLER, David P. *Introduction to cataloging and classification*. 11th ed. Santa Barbara: Libraries Unlimited, 2015. 1048 s. ISBN 978-1-59884-857-1 (váz.). ISBN 978-1-59884-856-4 (brož.).

JOUDREY, Daniel N., TAYLOR, Arlene G., WISSER, Katherine M. *The organization of information*. 4th ed. Santa Barbara: Libraries Unlimited, 2018. 722 s. ISBN 978-1-59884-859-5 (váz.). ISBN 978-1-59884-858-8 (brož.).

JURČACKOVÁ, Zora. *Organizácia poznania v dokumentovej komunikácii*. Bratislava: Centrum VTI SR, 1996. 98 s. ISBN 80-85165-85-6.

KOVÁŘ, Blahoslav. České termíny pro pojmenování procesu věcného pořádkání a jeho druhů. In: *Knihovnictví a bibliografie*. 1973, **2**(2), 37-44; **2**(3), 55-59. Příl. čas. Čtenář, 1973, **24**(2); **24**(5).

KOVÁŘ, Blahoslav. *Problémy teorie procesu věcného pořádkání informací a selekčních jazyků*. Praha: Univerzita Karlova, 1976. 165 [1] s. Acta Universitatis Carolinae. Philosophica et historica. Monographia, sv. 65/1976. ISSN 0567-8307.

KOVÁŘ, Blahoslav. *Věcné pořádkání informací a selekční jazyky*. 2. vyd. Praha: ÚVTEI, Institut pro mimoškolní vzdělávání, 1984. 251 s. Učební texty ÚVTEI, sv. 22.

KUČEROVÁ, Helena. *Organizace znalostí: klíčová témata*. 1. vyd. Praha: Karolinum, 2017. 269 s. ISBN 978-80-246-3587-3 (brož.). ISBN 978-80-246-3597-2 (pdf).

LACASTA, Javier, NOGUERAS-ISO, Javier, ZARAZAGA-SORIA, Francisco Javier. *Terminological ontologies: design, management and practical applications*. New York: Springer, 2010. xvii, 197 s. Semantic Web and Beyond. Computing for Human Experience, vol. 9, ISSN 1559-7474. doi:10.1007/978-1-4419-6981-1. ISBN 978-1-4419-6980-4 (Print). ISBN 978-1-4419-6981-1 (Online).

LAMBE, Patrick. From cataloguers to designers: Paul Otlet, social impact and a more proactive role for knowledge organization professionals. In: *Knowledge organization*. 2015, **42**(6), 445-455. ISSN 0943-7444.

MAI, Jens-Erik. Analysis in indexing: document and domain centered approaches. In: *Information processing and management*. May 2005, **41**(3), 599-611. ISSN 0306-4573.

- MASOLO, Claudio, BORGIO, Stefano, GANGEMI, Aldo, GUARINO, Nicola, OLTRAMARI, Alessandro. *WonderWeb deliverable D18: Ontology library (final)* [online]. Version 1.0. ISTC-CNR, 2003-12-31 [cit. 2018-01-30]. 343 s. Dostupné z: <http://wonderweb.man.ac.uk/deliverables/D18.shtml> a z <http://www.loa.istc.cnr.it/old/Papers/D18.pdf>.
- MAZZOCCHI, Fulvio. Knowledge organization system (KOS) [online]. Version 1.1. 2017-07-13 [cit. 2018-01-30]. In: *ISKO Encyclopedia of knowledge organization*. ISKO, 2017. Dostupné z: <http://www.isko.org/cyclo/kos>.
- McGUINNESS, Deborah L. Ontologies come of age. In: FENSEL, Dieter, HENDLER, James A., LIEBERMAN, Henry, WAHLSTER, Wolfgang, ed. *Spinning the semantic Web: bringing the World Wide Web to its full potential*. 1st ed. Cambridge (Mass.): MIT Press, 2003, s. 171-196. ISBN 978-0-26256212-6 (brož.). ISBN 978-0-262-06232-9 (váz.).
- MEALY, George H. Another look at data. In: *Proceedings of the Fall Joint Computer Conference (AFIPS Fall '67), November 14-16, 1967*. New York: ACM, 1967, s. 525-534. doi:10.1145/1465611.1465682.
- MILES, Alistair, BECHHOFFER, Sean, ed. *SKOS Simple Knowledge Organization System Reference* [online]. W3C Recommendation 18 August 2009 [version]. Cambridge (MA): World-Wide Web Consortium, © 2009 [cit. 2018-01-30]. Dostupné z: <http://www.w3.org/TR/2009/REC-skos-reference-20090818/>.
- NEWELL, Allen. The knowledge level. In: *Artificial intelligence*. January 1982, **18**(1), 87-127. doi:10.1016/0004-3702(82)90012-1. ISSN 0004-3702.
- NOY, Natalya F., McGUINNESS, Deborah L. *Ontology development 101: a guide to creating your first ontology* [online]. March 2001 [cit. 2018-01-30]. 25 s. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880. Dostupné z: [http://protege.stanford.edu/publications/ontology\\_development/ontology101.html](http://protege.stanford.edu/publications/ontology_development/ontology101.html).
- OBITKO, Marek, ZAMAZAL, Ondřej, SVÁTEK, Vojtěch. Ontologie a sémantický web. In: Vladimír MAŘÍK, Olga ŠTĚPÁNKOVÁ, Jiří LAŽANSKÝ a kol. *Umělá inteligence*. 6. 1. vyd. Praha: Academia, 2013, kap. 2, s. 85-125. ISBN 978-80-200-2276-9.
- Object Management Group. *Ontology definition metamodel (ODM)* [online]. Version 1.1. Needham (MA, USA): Object Management Group, September 2014 [cit. 2018-01-30]. 348 s. OMG Document Number: formal/2014-09-02. Dostupné z: <http://www.omg.org/spec/ODM/>.
- PALEK, Bohumil, ed. *Sémiotika: Ch. S. Peirce, C. K. Ogden and I. A. Richards, Ch. W. Morris, H. B. Curry*. 2. přeprac. vyd. Praha: Karolinum, 1997. 335 s. ISBN 978-80-7184-356-6.
- PICCA, Davide, GLIOZZO, Alfio Massimiliano, GANGEMI, Aldo. LMM: an OWL metamodel to represent heterogeneous lexical knowledge. In: *Proceedings of the International Conference on Language Resources and Evaluation (LREC)*. 26 May-1 June 2008, Marrakech, Morocco, s. 2413-2419.
- PRIETO-DÍAZ, Rubén. Domain analysis: an introduction. In: *ACM SIGSOFT Software engineering notes (SEN)*. April 1990, **15**(2), 47-54. doi:10.1145/382296.382703. ISSN 0163-5948.
- RANGANATHAN, Shiyali Ramamrita. Facet analysis: fundamental categories. In: CHAN, Lois Mai, RICHMOND, Phyllis A., SVENONIUS, Elaine, ed. *Theory of subject analysis: a sourcebook*. Littleton: Libraries Unlimited, 1985, s. 88-93. ISBN 0-87287-489-3. – Původně publikováno v rámci: RANGANATHAN, Shiyali Ramamrita. *The elements of library classification*. 3rd ed. Bombay: Asia Publishing House, 1962. 168 s.
- RANGANATHAN, Shiyali Ramamrita. *Prolegomena to library classification*. Assisted by Malur Aji GOPINATH. 3rd ed. London: Asia Publishing House, 1967. 640 s. Ranganathan series in library science, 20. ISBN 978-0-210-22731-2.
- ROWLEY, Jennifer E., HARTLEY, Richard J. *Organizing knowledge: an introduction to managing access to information*. 4th ed. Aldershot (GB): Ashgate, © 2008. xxiv, 367 s. ISBN 978-0-7546-4431-6 (váz.). ISBN 0-7546-4431-6 (váz.).
- SAYERS, W. C. Berwick. *Canons of classification applied to „the subject“, „the expansive“, „the decimal“, and „the Library of Congress“ classifications: a study in bibliographic classification method*. London: Grafton, 1915. 173 s.
- SECHSER, Otto, MOJŽIŠEK, Josef, KÖNIGOVÁ, Marie. *Selektivní jazyk a jeho popis*. 1. vyd. Praha: ÚVTEI, 1968. 219 s.

- SHERA, Jesse Hauk. *Documentation and the organization of knowledge*. 1st ed. London, 1966. 185 s.
- SILVA, Armando Malheiro da, RIBEIRO, Fernanda. Documentation / information and their paradigms: characterization and importance in research, education, and professional practice. In: *Knowledge organization*. 2012, **39**(2), 111-124. ISSN 0943-7444.
- SMIRAGLIA, Richard P. *Domain analysis for knowledge organization: tools for ontology extraction*. 1st ed. Waltham: Elsevier Chandos Pub, 2015. 105 s. ISBN 978-0-08-100150-9.
- SMIRAGLIA, Richard P. *The elements of knowledge organization*. Springer, 2014. 101 s. doi:10.1007/3-319-09357-4. ISBN 978-3-319-09356-7. ISBN 978-3-319-09357-4 (eBook).
- SMITH, Barry. Beyond concepts: ontology as reality representation. In: Achille C. VARZI, Laure VIEU, ed. *Formal ontology in information systems: proceedings of the Third International Conference (FOIS 2004)*: [International Conference on Formal Ontology and Information Systems, Turin, November 4-6, 2004]. Amsterdam: IOS Press, © 2004, s. 73-84. *Frontiers in artificial intelligence and applications*, vol. 114. ISSN 0922-6389. ISBN 978-158603-468-9.
- SOERGEL, Dagobert. *Indexing languages and thesauri: construction and maintenance*. Los Angeles: Melville Publishing, 1974. xliii, 632 s. Information Science Series. ISBN 0-471-81047-9. ISBN 978-0-471-81047-6.
- SOERGEL, Dagobert. *Organizing information: principles of data base and retrieval systems*. Orlando (FA, USA): Academic Press, 1985. xiv, 450 s. Library and Information Science series. ISBN 978-0-12-654260-8 (váz.). ISBN 978-0-12-654261-5 (brož.).
- SOUZA, Renato Rocha, TUDHOPE, Douglas, ALMEIDA, Maurício Barcellos. Towards a taxonomy of KOS: dimensions for classifying knowledge organization systems. In: *Knowledge organization*. 2012, **39**(3), 179-192. ISSN 0943-7444.
- SOWA, John F. *Knowledge representation: logical, philosophical, and computational foundations*. Pacific Grove: Brooks/Cole, © 2000. xiv, 594 s. Computer Science Series. ISBN 978-0-534-94965-5.
- SPITERI, Louise F. A simplified model for facet analysis: Ranganathan 101. In: *Canadian journal of information and library science*. April-July 1998, **23**(1-2), 1-30. ISSN 1195-096X (Print). ISSN 1920-7239 (Online).
- STAAB, Steffen, STUDER, Rudi, ed. *Handbook on ontologies*. 2nd ed. Dordrecht: Springer, 2009. xix, 811 s. International Handbooks on Information Systems. doi:10.1007/978-3-540-92673-3. ISBN 978-3-540-70999-2 (Print). ISBN 978-3-540-92673-3 (Online).
- SVÁTEK, Vojtěch, VACURA, Miroslav. Ontologické inženýrství na sémantickém webu. In: Vladimír MAŘÍK, Olga ŠTĚPÁNKOVÁ, Jiří LAŽANSKÝ a kol. *Umělá inteligence*. 6. 1. vyd. Praha: Academia, 2013, kap. 3, s. 149-168. ISBN 978-80-200-2276-9.
- SVÁTEK, Vojtěch. Aktuální problémy a perspektivy sémantického webu. In: *Systémová integrace*. 2014, **21**(1-2), 7-29. ISSN 1210-9479 (Print). ISSN 1804-2716 (Online).
- SVÁTEK, Vojtěch. Ontologie a WWW. In: Dušan CHLAPEK, ed. *DATAKON 2002: sborník databázové konference: Brno, Česká republika, 19.-22. října 2002*. Brno: Masarykova univerzita, 2002, s. 27-55. ISBN 978-80-210-2958-3.
- SVENONIUS, Elaine. *The intellectual foundation of information organization*. Cambridge (Mass): MIT Press, 2000. xiv, 255 s. Digital Libraries and Electronic Publishing Series. ISBN 978-0-262-19433-4.
- SZOSTAK, Rick, GNOLI, Claudio, LÓPEZ-HUERTAS, María. *Interdisciplinary knowledge organization*. Springer, 2016. 227 s. ISBN 978-3-319-30147-1. ISBN 978-3-319-30148-8 (eBook). doi:10.1007/978-3-319-30148-8.
- ŠAMURIN, Jevgenij Ivanovič. *Očerki po istorii bibliotečno-bibliografičeskoj klassifikacii*. Tom 1. Moskva: Izdatel'stvo Vsesojuznoj knižnoj palaty, 1955. 393 s.
- ŠAMURIN, Jevgenij Ivanovič. *Očerki po istorii bibliotečno-bibliografičeskoj klassifikacii*. Tom 2. Moskva: Izdatel'stvo Vsesojuznoj knižnoj palaty, 1959. 562 s.
- TONDL, Ladislav. *Znalost a její lidské, společenské a epistemické dimenze*. Praha: Filosofia, 2002. 173 s. ISBN 978-80-7007-167-0 (brož.).

USCHOLD, Michael, GRÜNINGER, Michael. *Ontologies: principles, methods and applications* [online]. Edinburgh: February 1996 [cit. 2018-01-30]. 63 s. Technical Report, AIAI-TR-191. University of Edinburgh, School of Informatics, Artificial Intelligence Applications Institute. Dostupné z: <http://www.aiai.ed.ac.uk/publications/documents/1996/96-ker-intro-ontologies.pdf>.

WEINBERGER, David. *Everything is miscellaneous: the power of the new digital disorder*. 1st ed. New York: Times Books, 2007. 277 s. ISBN 978-0-8050-8043-8.

WILSON, Patrick. *Two kinds of power: an essay on bibliographical control*. Berkeley: University of California Press, 1978, © 1968. 155 s. California library reprint series. ISBN 978-0-520-03515-7.

ZDRÁHAL, Zdeněk. Ontologie: od filosofie k umělé inteligenci. In: Vladimír MAŘÍK, Olga ŠTĚPÁNKOVÁ, Jiří LAŽANSKÝ a kol. *Umělá inteligence*. 6. 1. vyd. Praha: Academia, 2013, kap. 1, s. 21-84. ISBN 978-80-200-2276-9.

ZDRÁHAL, Zdeněk. Reprezentace znalostí. In: Vladimír MAŘÍK, Jiří LAŽANSKÝ, Olga ŠTĚPÁNKOVÁ a kol. *Umělá inteligence*. 1. 1. vyd. Praha: Academia, 1993, kap. 4, s. 99-122. ISBN 978-80-200-0496-3.

ZENG, Marcia Lei, QIN, Jian. *Metadata*. 2nd edition. Chicago: Neal-Schuman, an imprint of the American Library Association, 2016. xxvii, 555 s. ISBN 978-1-55570-965-5.