

ORGANIZACE ZNALOSTÍ: KLÍČOVÁ TÉMATA

HELENA KUČEROVÁ

KAROLINUM



Organizace znalostí: klíčová témata

Helena Kučerová

Recenzovali: doc. Ing. Vilém Sklenák, CSc.
Mgr. Marta Špániová, Ph.D.

Vydala Univerzita Karlova, Nakladatelství Karolinum
Ovocný trh 560/5, 116 36 Praha 1
www.karolinum.cz
Redakce Františka Jirousová
Grafická úprava Jan Šerých
Sazba DTP Nakladatelství Karolinum

© Univerzita Karlova, 2017
© Helena Kučerová, 2017

Kniha je dílčím výstupem řešení projektu NAKI DF13P010VV013 Znalostní báze pro obor organizace informací a znalostí, realizovaného v letech 2013–2015 na ÚISK FF UK v Praze.

ISBN 978-80-246-3587-3
ISBN 978-80-246-3597-2 (pdf)



Univerzita Karlova
Nakladatelství Karolinum 2019

www.karolinum.cz
ebooks@karolinum.cz

OBSAH

Úvod	7
1 Organizace znalostí – staré téma v novém kontextu	15
1.1 Historičtí předchůdci oboru organizace znalostí	17
1.2 Problémy a paradigmaty organizace znalostí	21
1.2.1 Konstantní problémy organizace znalostí	23
1.2.2 Nová paradigmaty pro 21. století	26
1.3 Organizace znalostí jako obor	38
1.4 Pojmová základna a pojmová mapa organizace znalostí	46
1.5 Přístup k organizaci znalostí v tuzemském kontextu – poznámky (nejen) k české terminologii	51
1.6 Shrnutí	55
2 Systémový přístup k organizaci znalostí: cíle, prvky, procesy, kontext	57
2.1 Cíl organizace znalostí – proč se organizuje?	58
2.2 Kontext organizace znalostí – kdo, kdy, kde organizuje?	64
2.3 Prvky organizace znalostí – co se organizuje?	69
2.3.1 Jednotky organizace znalostí: dokument – informační zdroj – kolekce ..	70
2.3.2 Znalosti? Informace? Data?	77
2.3.3 Filozofický přístup k vymezení jednotky organizace znalostí	86
2.3.4 Vymezení jednotky organizace znalostí	92
2.4 Procesy organizace znalostí – co se děje při organizování?	94
2.4.1 Proces organizace znalostí	96
2.4.2 Proces přístupu ke znalostem	98
2.5 Shrnutí	103
3 Principy organizace znalostí	105
3.1 Principy intelektuální organizace znalostí	107
3.1.1 Ekvivalence a identifikace	108
3.1.2 Kategorizace a klasifikace	115
3.1.3 Analýza a její principy	120
3.1.4 Vícekriteriální analýza aneb fasety	128
3.1.5 Definování	132
3.1.6 Řazení	135
3.2 Technologické postupy v průběhu organizace znalostí	137
3.2.1 Obecné techniky organizace – umístění a označení	138
3.2.2 Metafory technik organizace – seznam, kartotéka, index	146
3.2.3 Techniky přístupu ke zdrojům – navigace, množinový přístup a přístup prostřednictvím indexu	148

3.3	Empirické principy a pravidla organizace znalostí	150
3.3.1	Pravidla slovníkového katalogu Charlese A. Cuttera a Mezinárodní principy katalogizace	151
3.3.2	Sayersovy kánony klasifikace	153
3.3.3	Principy klasifikace pro knihovny Henryho E. Blisse	154
3.3.4	Zákony, kánony, principy a postuláty Shiyali R. Ranganathana	155
3.3.5	Gruberova kritéria návrhu ontologií	159
3.4	Shrnutí	160
4	Vztahy v organizaci znalostí	163
4.1	Formální vlastnosti vztahů	164
4.2	Způsoby vyjádření a implementace vztahů	167
4.2.1	Možnosti vyjádření vztahů	168
4.2.2	Vztahy v datových strukturách	170
4.3	Vztahy v jazyce	175
4.3.1	Paradigmatické a syntagmatické vztahy	175
4.3.2	Vztahy významu pojmů a slov	178
4.4	Přístup informační vědy	184
4.5	Vztahy v systémech organizace znalostí	191
4.5.1	Vztahy v Mezinárodním desetinném třídění	191
4.5.2	Vztahy v tezaurech podle standardu ISO 25964	197
4.6	Shrnutí	198
5	Sémantická problematika organizace znalostí	201
5.1	Teoretická východiska sémantické problematiky organizace znalostí	202
5.2	Problematika označování obsahu informačních zdrojů	211
5.2.1	Pojem <i>aboutness</i> v logice	212
5.2.2	Pojem <i>aboutness</i> v informační vědě	216
5.2.3	Model <i>aboutness</i> založený na principu neomezené semiózy	223
5.3	Shrnutí	230
	Závěr	231
	Příloha 1: Notace UML používaná v diagramech	235
	Literatura	241
	Summary	247
	Věcný rejstřík	248
	Jmenný rejstřík	267

Úvod

Řád je duše věcí ... Rozvažujeme-li, co je to, co zachovává naše veškerenstvo i jednu každou věc o sobě v náležitém stavu, shledáme, že to není nic, naprosto nic jiného než řád. Řád je spořádání věcí přednějších i zadnějších, vyšších i nižších, větších i menších, podobných i nepodobných, podle místa, času, počtu, míry a váhy, jak každé patří a hodí se. Proto kdosi nazval řád duší věcí; pěkně a pravdivě. Neboť všechno, co je spořádáno, zachovává svůj stav neporušený potud, pokud zachovává řád; opustí-li řád, chřadne, kolísá, padá a boří se.¹

Jan Amos Komenský, *Velká didaktika*, 1627–1638

Organizace znalostí je zapotřebí a proniká všude. Existuje mnoho oblastí, úkolů a funkcí, kde může organizace znalostí změnit svět k lepšímu tím, že podpoří každodenní činnosti (léčba pacientů, poskytnutí půjčky), učení a porozumění, tvorbu významu a smyslu, rozhodování, odvozování a objevování nových znalostí lidmi i počítačovými programy.²

Dagobert Soergel, 2015

-
- 1 Cit. dle vydání z r. 1958: Jan Amos Komenský. *Vybrané spisy Jana Amose Komenského. Svazek 1, Slovo úvodní, Navržení krátké o obnovení škol v Království českém, Velká didaktika, Informatorium školy mateřské, Brána jazyků otevřená*. Jan Patočka, red. 1. vyd. Praha: Státní pedagogické nakladatelství, 1958, s. 112.
 - 2 Dagobert Soergel. Unleashing the power of data through organization: structure and connections for meaning, learning and discovery. *Knowledge organization*. 2015, 42 (6), s. 402.

*Proč zanikla Římská říše?
Co znamená zkratka IFLA?
Co znamená slovo pozitivismus?
Ve kterém českém slově se po „r“ nepíše tvrdé „y“?
Jak je venku? Bude zítra pršet?
Je vesmír konečný?
Jak vyřešit dopravní zácpy ve velkoměstech?
Jaký je dnešní kurs eura vůči americkému dolaru?
Kde se lze dočíst něco o historii ostrova Kréta?
Jaké je chemické složení léku Paralen?
Ve kterém dramatu W. Shakespeara se praví, že „celý svět je scéna a muži, ženy, všichni jsou jen herci“?
Jak vypadají nejnovější snímky pořízené Hubbleovým dalekohledem?
Jakou melodii začíná Dvořákova Novosvětská symfonie?
Jak komentovaly německé deníky průběh parlamentních voleb v Rusku?
Jak zastavit krvácení z tepny?
Kolik obyvatel má San Francisco?
O čem je film Tisícročná včela?
Kdo je výkonným ředitelem společnosti Microsoft?
Jak nejlépe pečovat o pelargonie pěstované v truhlíku na okně?
Která firma nabízí nejlevnější inkoustové tiskárny v Čechách?*

Lidská touha po poznání nemá hranic a nesčetné jsou i druhy toho, co si člověk přeje poznat. Už zběžný pohled na výše uvedené otázky připomene existenci různých typů údajů, informací, poznatků, know-how, znalostí či vědomostí, v nichž by se mohly nacházet odpovědi. Paralely mezi hodnocením průběhu voleb a mezi chemickým složením léku se skutečně nalézají jen těžko. Některé dotazy lze zodpovědět stručnými a jednoznačnými faktografickými údaji („ano/ne“, „Mezinárodní federace knihovnických sdružení a institucí“, „cca 800 000“). Jiné, například dotaz na zastavení krvácení z tepny, budou vyžadovat obsírnější popis či návod. V dalších případech narazíme na to, že existují různé možnosti řešení (kupříkladu u dopravní zácpy) nebo různá pojetí či přístupy k chápání (například názory různých autorů na konečnost vesmíru či na příčiny pádu Římské říše, odlišné interpretace děje filmu). Jindy dokonce odpověď přinese nepředpokládané rozšiřující či vysvětlující poznatky (ti, kdo si lámou hlavu vzpomínáním na české slovo s měkkým „i“ po r, budou možná překvapeni odpovědí, že mezi slova odpovídající zadání patří i „mrkev“). Pro některé typy znalostí se ani nedostává slov – snímky z Hubbleova dalekohledu jsou vyjádřeny obrazem, melodie Novosvětské zvukem. Žádanými typy poznatků jsou i takové, jež místo odpovědí ukazují jen cestu k nim v podobě odkazu na zdroj, kde se potřebná znalost nachází. Kromě typických dotazů typu literatura o historii Kréty to může být výhodné i pro znalosti závislé na čase (měnové kurzy, počasí, ceny tiskáren, jméno výkonného ředitele firmy). Někdy ani nepožadujeme explicitní odpověď, protože ji ještě nikdo nevytvořil, ale jde

nám o fakta, texty a údaje, s jejichž využitím budeme schopni sami odpověď zformulovat (například subjektivně vyhodnotit aktuální údaje o oblačnosti, síle větru, teplotě a vlhkosti vzduchu a prohlásit, že venku je hezky), případně originálním způsobem objevit nebo vynalézt nový poznatek.

Různé typy otázek jsou diktovány různými typy lidských kognitivních problémů, nazvaných Nicholasem Belkinem „anomální stav vědění“.³ Existují dva typy anomálního stavu vědění: nedostatek informací neboli informační potřeba a naopak nadbytek informací, tzv. informační zahlcení či přesycenost informacemi. Obě tyto anomálie narušují pravidlo efektivního života s přiměřenými informacemi, zformulované již Norbertem Wienerem, a obě se řeší stejnými nástroji: výběrem toho, co odpovídá zjištěným či stanoveným informačním potřebám a odfiltrováním toho, co jim neodpovídá. Co se týče informačních potřeb, lze u nich pozorovat obdobnou diverzitu jako v případě znalostí: sahají od požadavku opětovného nalezení již známého přes výzkum, vyhledávání neznámého až k objevování nového. K uspokojení některých informačních potřeb bude zapotřebí provést úplný hloubkový průzkum, někdo může naopak vyžadovat pouze nejnovější přírůstky poznatků z oblasti, kterou již zevrubně ovládá.

Jednou z mála nezpochybnitelných jistot je, že k řešení obou typů anomálního stavu vědění pomáhá organizace. Je prokázáno, že přístup k jakýmkoli znalostem v organizovaných zdrojích je efektivnější než ve zdrojích neorganizovaných. Přes širokou variabilitu typů znalostí a informačních potřeb platí jedno obecné reciproční pravidlo: čím více úsilí věnujeme organizaci znalostí, tím méně ho budeme muset vynaložit při jejich vyhledávání a využívání - a naopak.

Škála aktérů, kteří se účastní procesu přístupu ke znalostem, ať už jako tvůrci, zpracovatelé, správci nebo uživatelé, je stejně pestrá jako znalosti samy. Kromě toho, že nový poznatek získá člověk sám přímo, ať už z empirických údajů, nebo „z vlastní hlavy“, jsou k dispozici ještě početné nepřímé možnosti získání znalosti od nějakého zprostředkovatele. Tím může být profesionál, tj. některý ze zástupců znalostních profesí (učitel, právník, inženýr, finanční poradce, lékař...), ale i laik, jehož „neodbornost“ je mnohdy vyvážena vysokou důvěryhodností (od souseda s nádhernými pelargoniiemi na okně si rádi necháme poradit, jak o ně pečovat). Významné institucionální aktéry představují paměťové a fondové instituce (knihovny, archivy, muzea, galerie), podniky a organizace, sféra školství a výzkumu a média. Instituce mají často velmi propracovanou dělbu práce mezi profesemi specializovanými na některou z aktivit podporujících přístup ke znalostem. Kromě lidí a institucí vstupují do hry i informační a komunikační technologie. Zajímavým a stále častějším fenoménem získávání znalostí je interakce člověka a počítače, v níž hrají počítače, resp. softwarové aplikace čím dál aktivnější roli.

3 Nicholas J. Belkin, Robert Norman Oddy, Henry Mason Brooks. ASK for information retrieval: Part I. Background and theory. *Journal of documentation*. 1982, 38 (2), s. 61-71.

O tom, že odpovědi na otázky z úvodu existují, není pochyb. Ale kde je najít? Opět se nabízejí rozmanité typy zdrojů. Možná odpověď známe sami, protože ji dokážeme zjistit (například počasí venku), či ji máme uloženou ve své vnitřní paměti (třeba si vzpomeneme, že výrok o světě jako divadle je ze hry *Jak se vám líbí*). Nebo znalost získáme přímou komunikací s jinými osobami (konzultace s lékařem, rozhovor s dopravním dispečerem, poslech výkladu učitele). Další možností je vyhledávání ve vnější paměti, kdy znalost najdeme zaznamenanou v nějakém externím zdroji, například v učebnici. Na takový způsob přístupu, jehož podmínkou jsou organizované zaznamenané znalosti, se zaměříme v této knize.

Za informační zdroje, jak jsou zaznamenané znalosti nazývány, už zdaleka nejsou považovány jen sbírky knihoven a archivů. Patří k nim i znalostní báze, bibliografické databáze, autoritní báze, tezaury a další pomůcky k vyhledávání, e-knihy, elektronické časopisy, emailové archivy, dokumenty vygenerované pomocí kancelářských balíků (korespondence, zprávy, výkazy, tabulky), záznamy v transakčních databázích a datových skladech podnikových informačních systémů či studijních informačních systémů ve školách a četné další zdroje. Co do typů mohou zahrnovat jak text, tak i obrázky, audio, video, numerická data, primární zdroje i metadata v různých formátech. Mohou obsahovat informace „zdarma“ (volně přístupné, veřejné) i informace za poplatek (licencované zdroje). Zdroje mohou být interní (podnikový intranet nebo institucionální repozitář) i externí (např. online jízdní řády), místní i vzdálené (např. v cloudu, dostupné prostřednictvím webových služeb). Mohou to být soubory uložené na síťových serverech, na lokálních discích nebo na přenosných paměťových médiích, ale i vysílané či „tekoucí“ v prouděch dat, zdroje volně dostupné na internetu neboli povrchový web i tzv. neviditelný web profesionálních a institucionálních informačních zdrojů a databází. Potenciálně relevantní zdroje mohou být vysoce formalizované, ale i zcela neformální, vytvářené a využívané bez předem daných pravidel.

A právě toto ohromné množství typů zdrojů informací, doprovázené dramaticky se zvyšujícím počtem dostupných „jednotek znalostí“ v každém typu, se paradoxně stává první překážkou na cestě k jejich využití. Například na dotaz „pozitivismus“ vyhledal 6. 11. 2016 Google 27 000 dokumentů, po zadání termínu v angličtině, tj. „positivism“, ohlásil 2 910 000 dokumentů. Studie *How much information* (Kolik je informací), publikovaná v roce 2003 Kalifornskou univerzitou v Berkeley,⁴ přináší ohromující údaje o množství a dynamice růstu dat zaznamenaných na papíře, filmu a na magnetických a optických médiích. V roce 1999 byly podle této studie vyprodukovány 2 exabyty (1 exabyte je 10^{18} bytů) nových informací, v roce 2002 už to bylo 5 exabytů, přičemž 92 % bylo uloženo na magnetickém médiu. Vzrůstá

4 Peter Lyman, Hal R. Varian. *How much information? 2003* [online]. Berkeley: School of Information Management and Systems, 2003 [cit. 7. 11. 2016]. Dostupné z: <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>.

i počet organizovaných zdrojů. Zatímco první seznam veřejně dostupných databází *Computer-readable databases* (CRD), který sestavila v roce 1976 Martha E. Williamsová, obsahoval 301 položek, jeho aktuální verze ve formě 38. vydání seznamu *Gale directory of databases* z roku 2015 už zahrnuje přes 14 000 databází od více než 3 000 producentů a poskytovatelů. Je zřejmé, že orientace v takové rozmanitosti zdrojů, doprovázené jejich obrovským rozsahem, vyžaduje nějakou organizaci.

Pestrosti typů zdrojů odpovídají i rozmanité možnosti přístupu k nim. Opět můžeme konstatovat, že při získávání znalostí lze uplatnit všechny typy komunikace – přímou (intrapersonální, interpersonální, skupinovou, masovou) i nepřímou, zprostředkovanou nějakým médiem. Tradičním typem přístupu v knihovnických a bibliografických systémech je přístup ke znalostem prostřednictvím prohledávání předem organizovaných, vysoce formálně strukturovaných a sémanticky anotovaných zdrojů, vyžadující aktivní roli zpracovatele (poskytovatele, autora, organizátora) i uživatele. Typickým příkladem takových zdrojů je bibliografická databáze MEDLINE nebo znalostní báze DBpedia. V zásadě opačným typem je přímý přístup ke znalostem prostřednictvím plnotextového prohledávání neorganizovaných zdrojů počítačovou technologií. Odborníci označují tento způsob vyhledávání anglickým termínem *information retrieval*, běžný uživatel internetu by pro něj dnes nejspíš použil sloveso „vygooglit“ či „vygooglovat“, podle anglického vzoru *google*⁵, metonymicky využívajícího názvu největší vyhledávací služby internetu. Další typy přístupu doplňují to, co plnotextové vyhledávání nedokáže: listování zdroji a personalizaci jejich aktivní nabídky. Listování či prohlížení již existujících struktur umožňuje uživateli „vynechat“ formulaci dotazu a zaměřit se jen na posouzení relevance prohlížených zdrojů. Další možností je přístup prostřednictvím „nabídky“ citační nebo sociální sítě (odkaz, citace, doporučení poskytnuté a vypracované lidmi, s nimiž jsme v sociálním kontaktu). Do této skupiny patří jak sdílení zdrojů na sociálních sítích, tak tradiční seznamy literatury uváděné v závěru odborných textů. Do hry vstupuje i přístup obrazně nazývaný dobývání znalostí z dat, který je plně v režii počítačových technologií.

Současný terén organizace znalostí je tedy nesmírně členitý a brání se sevření do jednotně vymezených hranic. V souvislosti s tím se diskutuje řada otázek: Je organizace znalostí vůbec možná? Neboli – lze uspořádat něco, co vzniká zcela spontánně? Nelze-li ovlivnit proces vytvoření znalosti, jaká je šance, že se podaří zorganizovat jeho produkty? A pokud je organizace znalostí možná, je i nutná? Tedy – je vůbec v éře Googlu třeba organizovat informace (resp. jejich zdroje)? Nestačí mít k dispozici výkonné vyhledávače? A je-li

5 Sloveso *google* je ve slovníku Merriam-Webster definováno následovně: „Používat (coby osoba) vyhledávač Google k získání informací na síti world wide web (to use the Google search engine to obtain information about /as a person/ on the World Wide Web)“. Zdroj: Google. In: *Merriam-Webster Online Dictionary* [online]. Merriam-Webster, 2011 [cit. 7. 11. 2016]. Dostupné z: <http://www.merriam-webster.com/dictionary/google>.

organizace znalostí nutná, jsou pro její realizaci zapotřebí specialisté – odborníci? Není lepší (a lacinější) spolehnout se na počítačové technologie anebo na „moudrost davu“, podporovanou v prostředí Webu 2.0 a sociálních sítí informačními technologiemi? A i kdybychom si na všechny předchozí otázky odpověděli kladně a shodli se na tom, že organizace znalostí je nejen možná, ale i nutná a že je k její realizaci zapotřebí školených odborníků, zbývá ještě z hlediska zaměření této knihy klíčová otázka: Je možné vytvořit pro tuto specializovanou činnost obecnou teorii? Je zřejmé, že pro různé typy informačních potřeb, pro různé zdroje a pro různé typy přístupů k nim je zapotřebí používat i různé metody a techniky organizace znalostí. Je proto také potřeba mít různé teorie, například teorii organizace znalostí pro knihovny, specializovanou teorii pro archivy, pro podniky? A můžeme pokračovat ještě dále – potřebujeme opravdu teorii? Není tou nejschůdnější cestou pro organizaci znalostí právě eklektičnost a ad hoc kombinování vhodných nástrojů pro ten který účel? Tváří v tvář existenci tolika rozdílných motivací pro to, aby lidé (a nejen oni) organizovali znalosti, enormnímu množství a především diverzitě potřeb a zdrojů znalostí se skutečně nabízí pokusení vzdát se snahy vytvořit nějakou obecnou teorii organizace znalostí. K hnacím motorům lidských intelektuálních aktivit však vždy patřilo usilovat o nedosažitelné a hledat cesty k potlačení entropie uspořádáním neustále se měnící materie poznatků o světě. Motto v úvodu svědčí o tom, že tento názor zjevně zastává i jeden ze zakladatelů oboru organizace znalostí Dagobert Soergel, jenž si přes hranice věků podává ruku s Janem Amosem Komenským. Svým skromným dílem se chce na této snaze podílet i tato kniha.

Tato publikace je jedním z výsledků výzkumu realizovaného v rámci řešení projektu NAKI DFPO1OVV013 „Znalostní báze pro obor organizace informací a znalostí“ v letech 2013–2015. Cílem projektu bylo navrhnout znalostní bázi pro nově se ustavující disciplínu organizace znalostí a naplnit ji daty, jež mimo jiné poskytla zdrojový materiál pro zpracování této monografie a pro návrhy na aktualizaci české terminologie. V knize je podán přehled významných aktuálních problémů oboru, přičemž přednostní pozornost je věnována „obsahové“ organizaci, založené na sémantice organizovaných jednotek. Problematika „formální“, „deskriptivní“ (v českém jazykovém prostředí často nazývané „identifikační“) organizace je záměrně sledována spíše pasivně s vědomím toho, že se jí v současnosti věnuje velké množství jiných výzkumných a publikačních projektů, především bibliografické modely FRBR (*Funkční požadavky na bibliografické záznamy*) a BIBFRAME a pravidla popisu RDA (*Resource description and access*). Sémantickou organizaci považujeme jednak za významnější, jednak za podstatně méně prozkoumanou ve srovnání s organizací formální. Řečeno slovy Karen Coyleové:

Spolu s novými modely pro bibliografický popis nebyly navrženy žádné nové věcné (obsahové) metody ... Může to být historická náhoda nebo to může souviset s proveditelností úloh ... Ať už z jakéhokoli důvodu, nacházíme naši profesi uprostřed aktivní

diskuse o popisné bibliografii, s velmi malou pozorností zaměřenou na úlohu usnadnění tematického (obsahového) přístupu.⁶

Text knihy je členěn do pěti kapitol. První kapitola vymezuje obor organizace znalostí a jeho pojmovou a terminologickou základnu. Je nastíněna historická dimenze organizace znalostí a jsou uvedeny obory, jež tuto úlohu řešily a zkoumaly už před ustavením oboru organizace znalostí. Po shrnutí konstantních problémů organizace znalostí je charakterizována změna paradigmat, vyvolaná změněným společenským a technologickým kontextem. Popis ustavení oboru organizace znalostí je doplněn přehledem přístupů jeho významných představitelů k vymezení předmětu zkoumání a hranic oboru. Prostřednictvím pojmové mapy je definováno pojetí oboru uplatněné v textu této knihy. Závěr kapitoly komentuje problematiku terminologického aparátu organizace znalostí především v českém jazykovém prostředí.

Ve druhé kapitole je na problematiku organizace znalostí uplatněn systémový přístup – je definován její cíl, prvky, procesy a je vymezen její kontext. Pozornost je věnována především vymezení jednotky organizace, v jehož rámci je aplikována teorie poznání Karla R. Poppera. Následující kapitoly jsou věnovány třem významným tématům organizace znalostí: principům organizace, vztahům a sémantice v organizaci znalostí.

Teoretické a technologické principy organizace znalostí shrnuje třetí kapitola. V části věnované principům intelektuálních procesů organizace znalostí je popsána problematika ekvivalence, identifikace, kategorizace, fasetové analýzy, definování a řazení. V části věnované technologickým principům je vysvětlena úloha indexového souboru v zajištění přístupu k analogovým i digitálním zdrojům a jsou popsány klíčové techniky organizace a přístupu ke zdrojům – umístění, označení, navigace a množinový přístup. Následuje přehled empirických principů a pravidel organizace znalostí, formulovaných významnými představiteli oboru.

Čtvrtá kapitola se věnuje problematice vztahů v organizaci znalostí. Na teoretické úrovni jsou popsány formální vlastnosti vztahů a možnosti jejich vyjádření, paradigmatické a syntagmatické vztahy v jazyce a další lexikálně sémantické vztahy. Sledovanými aplikačními oblastmi jsou především terminologické systémy a systémy organizace znalostí, datové struktury a datové modely.

V páté kapitole je prezentována sémantická problematika organizace znalostí. Po přehledu teoretických východisek je pozornost věnována problematice označování obsahu informačních zdrojů (*aboutness*) a v závěru kapitoly je předložen vlastní návrh modelu pojmu *aboutness*, založený na principu neoomezené semiózy Charlese S. Peirce.

6 Karen Coyle. *FRBR before and after: a look at our bibliographic models*. Chicago: ALA Editions, an imprint of the American Library Association, 2016, s. 6.

/1/

Organizace znalostí – staré téma v novém kontextu

Organizace znalostí (OZ) ... zahrnuje aktivity jako popis dokumentů, indexace a klasifikace prováděné v knihovnách, databázích, archivech atd. Tyto aktivity vykonávají knihovníci, archiváři, oboroví specialisté a rovněž počítačové programy. OZ jako vědní obor se zabývá povahou a kvalitou procesů organizace znalostí (např. tvorbou taxonomií a ontologií), právě tak jako systémy organizace znalostí (SOZ), používanými k organizování dokumentů, reprezentací dokumentů a pojmů.⁷

Wikipedie, 2016

Motto pro tuto kapitolu bylo zvoleno z více důvodů. Jednak výstižně poukazuje na dvě stěžejní oblasti organizace znalostí – samotný proces organizace zaznamenaných znalostí a jeho nástroj, systém organizace znalostí. Druhým důvodem je snaha předložit názorný příklad tvrzení z názvu kapitoly, že organizace znalostí je „starým tématem v novém kontextu“. Text propojuje informaci o tradici organizace znalostí (knihovny, archivy, popis, indexace, klasifikace dokumentů) se současným kontextem, jež představují

⁷ Knowledge organization [online]. In: *Wikipedia: the free encyclopedia*, 2007- [cit. 7. 11. 2016]. Dostupné z: http://en.wikipedia.org/wiki/Knowledge_organization.

databáze a počítačové programy. Lokalizace tohoto textu na Wikipedii navíc poukazuje na aktuální konvergenci klasických forem vědeckého zkoumání a laických aktivit, realizovaných v obou případech s podporou stejných technologií a se stejným cílem – tvorba nových poznatků, jejich správa a zpřístupnění. Heslo *Knowledge organization* (Organizace znalostí) bylo na anglické Wikipedii vytvořeno 15. 2. 2007 a k datu 6. 11. 2016 ho editovalo 38 lidských autorů (zásahy robotů nejsou započítávány). Při porovnání textu ve Wikipedii a článku Birgera Hjørlanda *What is knowledge organization* (Co je organizace znalostí) z roku 2008⁸, uveřejněného ve speciálním čísle časopisu *Knowledge organization*, věnovaném problematice vymezení organizace znalostí, lze konstatovat významovou i formulační shodu zde uvedených definic⁹ (uživatel Wikipedie BirgerH, který v letech 2008, 2011 a 2012 heslo editoval, je evidentně totožný s Birgerem Hjørlandem¹⁰). Další konvergenci lze sledovat na poli „tradičních“ formátů a elektronického publikování. Zatímco v roce 2008 Hjørland publikoval nejprve svůj text v klasickém tištěném časopise a poté v něm obsažené vymezení organizace znalostí umístil na Wikipedii, v roce 2016 se pořadí obrací. Hjørlandův článek *Knowledge organization (KO)*¹¹, který je aktualizovanou a rozšířenou verzí textu z roku 2008, byl nejprve publikován jako součást online encyklopedie organizace znalostí (*ISKO Encyclopedia of knowledge organization*)¹² a až poté byl publikován v tištěném médiu. Tento konkrétní příklad naznačuje rostoucí význam a zároveň i stoupající důvěryhodnost nových médií a technologií v organizaci znalostí.

Text této kapitoly je členěn na pět částí. V části 1.1 je nastíněna historická dimenze organizace znalostí a jsou uvedeny obory, jež tuto úlohu řešily a zkoumaly už před ustavením oboru organizace znalostí. Část 1.2 shrnuje problémy organizace znalostí a charakterizuje změnu paradigmat vyvolanou změnám společenským a technologickým kontextem. Část 1.3 popisuje vznik oboru organizace znalostí a podává přehled přístupů jeho významných představitelů k vymezení předmětu zkoumání a hranic oboru. V části 1.4 je prostřednictvím pojmové mapy definováno pojetí oboru uplatněné v textu této knihy. Část 1.5 komentuje problematiku terminologického aparátu organizace znalostí a jeho specifika v českém jazykovém prostředí.

8 Birger Hjørland. What is knowledge organization (KO)? *Knowledge organization*. 2008, 35 (2-3), s. 86-101.

9 Je určitou ironií, že v tištěném článku uvádí Hjørland ve výčtu aktérů organizace znalostí i laiky, ti však už ve Wikipedii uvedeni nejsou.

10 Viz <https://en.wikipedia.org/wiki/User:BirgerH> [cit. 7. 11. 2016].

11 Birger Hjørland. Knowledge organization (KO). *Knowledge organization*. 2016, 43 (6), s. 475-484. – Na odvozenost tištěné verze z verze elektronické výslovně upozorňuje poznámka v úvodu článku.

12 ISKO. *ISKO Encyclopedia of Knowledge Organization* [online]. Birger Hjørland, ed. ISKO Scientific Advisory Council, 2016– [cit. 7. 11. 2016]. Dostupné z: <http://www.isko.org/cyclo/>.

1.1 HISTORIČTÍ PŘEDCHŮDCI OBORU ORGANIZACE ZNALOSTÍ

*Věda jsou organizované znalosti. Moudrost je organizovaný život.*¹³

Immanuel Kant (1724–1804)

Jak bude uvedeno v kapitole 1.3, obor organizace znalostí byl sice „oficiálně ustaven“ až na přelomu 20. a 21. století, to ovšem neplatí o procesech, jež jsou předmětem jeho zájmu a jež se realizovaly v rámci různých nazývaných aktivit prakticky po celou dobu lidské historie. Nejvýznamnějšími oblastmi byla tvorba encyklopedií, klasifikace věd, techniky duševní práce, knihovnické a bibliografické pořádací systémy. Historicky nejmladším oborem je technologie počítačového zpracování dat.

Encyklopedistika

K historicky nejstarším pokusům o uspořádání veškerého lidského vědění patří encyklopedie. Slovo pocházející z řečtiny a původně označující úplné vzdělání se od 16. století začalo používat pro označení zdrojů, jež toto úplné vzdělání zahrnují. Encyklopedie má stejnou ambici jako knihovny a další paměťové a fondové instituce (muzea, archivy, galerie), tj. soustředit na jednom místě původně rozptýlené zdroje. Místo samotných informačních zdrojů ovšem encyklopedie na jednom místě shromažďuje znalosti. S paměťovými institucemi má encyklopedie společný i aspekt shromažďování všeho, co bylo zaznamenáno, nikoli jen vědecky potvrzených faktů (existují encyklopedie mystiky nebo encyklopedie ras a bytostí z filmu *Star Wars*). Výstižnou charakteristikou encyklopedie jsou závazná pravidla Wikipedie: nezaujatý úhel pohledu, ověřitelnost a především „žádný vlastní výzkum, za který jsou ... považovány všechny dosud nepublikované teorie, data, sdělení, pojmy, argumenty nebo myšlenky a všechny nepublikované analýzy nebo syntézy již publikovaných dat, sdělení, teorií, argumentů nebo myšlenek“.¹⁴ Encyklopedie je tedy uspořádaná množina informací o zaznamenaných poznacích; není to systém organizace znalostí v tom smyslu, jak je definován v kapitole 1.4, nicméně nějaký systém organizace potřebuje, ať už ve formě hesláře, nebo kategoriálního či klasifikačního schématu.¹⁵ Jednotkou organizace je reprezentace

13 Poznámka: Tyto výroky, tradičně připisované Immanuelu Kantovi, nebyly v citovaném znění v žádném jeho díle publikovány. Jedná se o parafrázi, kterou charakterizoval Kantův světonázor Will Durant ve svém přehledu dějin filozofie: „Vjemy jsou neorganizovaným stimulem, vnímání je organizovaným vjemem, pojem je organizované vnímání, věda je organizované vědění, moudrost je organizovaný život.“ Citováno dle Will Durant. *Příběh filozofie: životy a myšlenky největších filozofů*. Z angl. orig. *Story of philosophy* přel. Pavel Kaas. Praha: Pragma, [2003], s. 245.

14 Dostupné z: https://cs.wikipedia.org/wiki/Kategorie:Wikipedie:Závazná_pravidla [cit. 7. 11. 2016].

15 Viz například d'Alembertova a Diderotova klasifikace oborů, zveřejněná v 1. díle francouzské Encyklopedie (https://upload.wikimedia.org/wikipedia/commons/5/58/ENC_SYSTEME_FIGURE.jpeg [cit. 7. 11. 2016]) nebo kategoriální systém Wikipedie na: https://cs.wikipedia.org/wiki/Kategorie:Hlavní_kategorie [cit. 7. 11. 2016].

zaznamenané znalosti ve formě hesla či článku, jež obsahuje definici a zpravidla i další charakteristiky zahrnutých pojmů.

Encyklopedické zdroje byly pod nejrůznějším označením vytvářeny už ve starověkém Egyptě. Jejich historie sahá od takzvané Amenemopovy encyklopedie nazvané *Učebnice, která činí chytrým a poučuje nevědomé, jak poznat vše, co existuje* z konce druhého tisíciletí př. n. l.¹⁶ přes antiku (např. *Naturalis historiae libri XXXVII* Gaia Plinia Secunda staršího z roku 77) až ke středověku. Pokračuje pansofickými snahami baroka, jejichž významným představitelem byl Jan Amos Komenský, přes osvícenské ideály francouzských encyklopedistů k tištěným projektům dalších velkých národních encyklopedií, jejichž symbolem se v českém prostředí staly Ottovy slovníky naučné¹⁷. Encyklopedické aktivity se nepřestávají vyvíjet. Encyklopedický charakter měly utopické vize Paula Otleta (částečně realizované v projektu Mundanea) i eseje H. G. Wellse na téma „mozek světa“ (*World Brain, The Idea of Permanent World Encyclopaedia*). K vizím patří i memex Vannevara Bushe, jehož prorocká slova z roku 1945 o tom, že se objeví „zcela nové formy encyklopedií s již předpřipravenou sítí asociačních řetězců, jež bude možné vložit do memexu a v něm rozvíjet“¹⁸, v současnosti našla svoji aktuální podobu ve Wikipedii.

Klasifikace věd a vědecké klasifikace

Jednotkou organizace vědecké klasifikace jsou předměty zkoumané příslušnou vědní disciplínou. Prominentním příkladem vědecké klasifikace je periodická tabulka prvků, díky jejíž teoretické koncepci bylo možné i předvídat v době jejího vzniku dosud neobjevené chemické prvky. Specifickou oblastí s vlastní bohatou tradicí je přírodovědná taxonomie, která se soustředí na vědecké určování příbuznosti organismů založené na jejich vlastnostech. Klasická taxonomie organismů Carla Linného je vytvořena na základě jejich fyzických a morfologických charakteristik, každý organismus má své rodové a druhové jméno. Fylogenetické klasifikace třídí organismy podle jejich vývojové příbuznosti, tzv. molekulární taxonomie využívá pro systematiku organismů analýzu DNA.

Jednotkou organizace v klasifikaci věd je vědní obor, disciplína. Nad utříděním vědních oborů uvažovali mnozí filozofové, kteří prostřednictvím klasifikace věd obvykle vyjadřovali svůj názor na způsob uspořádání světa, jež dané vědy zkoumají. K nejvýznamnějším systémům, jež ovlivnily i pořádací systémy knihoven, patří Porfyriův strom vědění – dichotomický systém zpracovaný jako komentář k Aristotelovu dílu Kategorie, dále středověký systém

16 Jevgenij Ivanovič Šamurin. *Geschichte der bibliothekarisch-bibliographischen Klassifikation* I. Leipzig: Bibliographisches Institut, 1964, s. 11-12.

17 Přehled vývoje české encyklopedistiky viz: Dagmar Hartmanová. *Historie Československé encyklopedistiky do roku 1945. Národní knihovna*. 2000, 11 (1), s. 15-21. – Dagmar Hartmanová. *Historie Československé encyklopedistiky 1945-1992. Národní knihovna*. 2000, 11 (2-3), s. 80-88.

18 Vannevar Bush. *As we may think. Atlantic monthly*. 1945, 176 (7), s. 108.

sedmi svobodných umění a především systém Francise Bacona, který se stal ideovou základnou Deweyho desetinného třídění a byl převzat i pro Mezinárodní desetinné třídění. Teze o bezprostřední souvislosti systémů organizace znalostí v knihovnách a klasifikace věd, formulovaná např. Henry E. Blissem a u nás Blahoslavem Kovářem, patřila k vlivným teoretickým principům organizace znalostí v 19. i ve 20. století a má své zastánce i dnes.

V současné době se snahy o vědní systematiku obou typů opírají o matematické metody. tzv. numerická klasifikace umožňuje s využitím matematických metod vytvářet automatizovaně klasifikace předmětů zkoumaných vědními disciplínami. Modelování vztahů vědních disciplín umožňují kvantitativní analýzy výstupů vědecké práce, typicky vědeckých publikací. Spíše než o klasifikaci věd se v tomto případě mluví o mapování vědy a poznání (angl. *knowledge mapping*). Zajímavou ukázkou je například vizualizace kategoriálního systému Wikipedie a její porovnání s rozložením tříd v Mezinárodním desetinném třídění na serveru Places & Spaces, který se zaměřuje na zveřejňování výsledků mapování vědy.¹⁹

Techniky duševní práce

Jak bude podrobně zdůvodněno v kapitole 2, organizace znalostí v lidské mysli není předmětem oboru organizace znalostí. Ten se soustředí na znalosti zaznamenané a uložené v externích pamětech. Je nicméně pravdou, že mnohé z technik duševní práce, například slovní asociační metody, metody vedoucí k zapamatování a rozšíření rozsahu paměti (tzv. paměťové paláce či metody loci – organizace znalostí jejich umísťováním do pomyslných budov či bytů²⁰, mnemotechnika ad.) nebo metody rychlého čtení či tzv. informační hygieny mohou poskytnout inspiraci jak pro procesy organizování v externích úložištích, tak pro procesy přístupu ke zdrojům v nich uložených. Není bez zajímavosti, že významný český autor publikací v této oblasti Jiří Toman²¹ (1913–1988) byl zároveň jedním ze zakladatelů československé soustavy VTEI, působil v knihovně Akademie věd, podílel se na tvorbě univerzálního systému organizace znalostí BSO (*Broad system of ordering*) a byl rovněž autorem publikací věnovaných organizaci informací²², mezi jinými též první české příručky pro tvorbu tezauru.²³

Knihovnické a bibliografické pořádací systémy

Knihovnictví je významnou oblastí, v níž se za tisíciletý vývoj nakumulovalo velké množství praktických zkušeností i teoretických poznatků o efektivní organizaci velkých kolekcí dokumentů. Dokument ve svých historicky se vyvíjejících

19 Dostupné z: http://scimaps.org/mapdetail/design_vs_emergence__127 [cit. 7. 11. 2016].

20 Viz např. David Weinberger. *Everything is miscellaneous: the power of the new digital disorder*. 1. vyd. New York: Times Books, 2007, s. 52.

21 Viz např. Jiří Toman. *Organizace a technika duševní práce*. 1. vyd. Praha: Svoboda, 1970.

22 Viz např. Jiří Toman. *Systémy pořádání informací: úvod do heslování a třídění*. Praha: Státní technická knihovna, 1964.

23 Jiří Toman. *Postup při sestavování tezauru*. Praha: Ústředí vědeckých, technických a ekonomických informací, Státní technická knihovna, 1968.

formátech zůstává i dnes základní jednotkou organizace v knihovnách. Už v Aššurbanipalově knihovně v Ninive ze 7. století př. n. l., jež byla objevena při vykopávkách v letech 1849–1854, předpokládal asyrolog Joachim Ménant systematické členění shromážděných hliněných destiček na obory jazykověda, historie, právo, přírodověda (členěná na zvířata, rostliny, minerály), zeměpis, matematika a astronomie, magie, náboženství, legendy.²⁴ *Bibliotheca universalis* Conrada Gesnera představuje první pokus o sestavení univerzální světové bibliografie, zahrnující na 12 000 děl (zaznamenání údajů o všech jejich vydáních bylo už sto let po vynálezu knihtisku považováno za neproveditelné). Soupis vydaný v roce 1545 byl řazen abecedně podle jmen autorů a v roce 1548 byl doplněn o systematické uspořádání (*Pandectarum sive partitionum universalium*).

Z dnešního pohledu bylo zlomovým obdobím 19. století, kdy se knihovny staly na základě širokého společenského konsenzu úložišti a správci zaznamenaného poznání, jehož dostupnost zajišťovaly tehdejší lístkové katalogy. Nejvýznamnějším teoretickým přínosem knihovnictví 19. století bylo vymezení cílů katalogu v *Pravidlech slovníkového katalogu* Charlese A. Cuttera²⁵, zaměřené na umožnění vyhledávání a přístupu uživatele k organizovaným zdrojům (blíže viz kapitola 3.3.1). Projekt celosvětové bibliografie veškeré literární produkce Paula Otleta a Henri La Fontaina z roku 1895 sice zůstal nenaplněn, jeho 16 milionů bibliografických záznamů však poskytlo impuls pro vytvoření Mezinárodního desetinného třídění a materiál pro vypracování principů dokumentační práce. Obojí je aktuální dodnes. Na podobu současného stavu organizace znalostí měly vliv rovněž významné univerzální knihovnické pořádací systémy 19. a počátku 20. století: Deweyho desetinné třídění, Mezinárodní desetinné třídění, Blissovo Bibliografické třídění, Dvojtečkové třídění S. R. Ranganathana, Knihovnicko-bibliografické třídění a *Broad system of ordering*. Knihovnické a bibliografické systémy se staly ve druhé polovině 20. století oblastmi testování a praktické aplikace nových informačních technologií. V teoretické rovině to byla bezesporu teorie fasetové analýzy, jež ve 20. století obohatila obor o vědecké principy organizace založené na systematickém vícekritériálním přístupu.

Technologie zpracování a komunikace dat v počítačích a v počítačových sítích

Dostupná technologie vždy představovala limity a zároveň nabízela podněty pro další způsoby organizace zaznamenaného poznání. Stačí jen připomenout, jak rozměr hliněných tabulek, papyrových svitků, rozsah kodexů, gramofonových desek nebo katalogizačních lístků ovlivňoval dobové vymezení jednotky organizace. V současné době jsou obdobně určující velikosti displeje osobních počítačů či mobilních zařízení, svou roli hrají i limity rozsahu textových zpráv v mobilních aplikacích nebo v sociálních sítích. V polovině 20. století

24 Joachim M. Ménant. *La bibliothèque du palais de Ninive*. Paris: E. Leroux, 1880.

25 Charles Ammi Cutter. *Rules for a printed dictionary catalogue*. Washington: Government Printing Office, 1876.