

Tomáš Samek

Všechna tvrzení humanitních věd se nakonec opravdu týkají skutečnosti "vnitřní zkušenosti": oblasti bytí, na niž se nevztahuje kategorie "vysvětlení", nýbrž kategorie "porozumění". [2, s.15]

Tezaurus v lingvistice a informační vědě

Chceme-li zkoumat filozofické zakotvení problematiky tezaurů, je nejprve třeba ujasnit základní pojmy "tezaurus" a "selekční jazyky", jak je chápe současná informační věda jakožto mezioborová disciplína studující informační komunikaci ve společnosti. Informačně-vědné pojetí je totiž odlišné od poněkud známějšího pojetí jazykovědného: termín "tezaurus" vymezují obě disciplíny odlišně, a tato polysémie namnoze vede k pojmoslovným konfuzím a interdisciplinárním nedorozuměním.

Jazykověda chápe tezaurus ve dvojitým základním smyslu:

1. jako úplný souhrn slovní zásoby národního jazyka jisté (minulé nebo současné) epochy. V české lexikografii nejvíce proslul pokus o tezaurus uskutečněný J. A. Komenským;
2. jako ideografický slovník Rogetova typu, tedy slovník sémaziologický, v němž se na rozdíl od běžných slovníků onomaziologických nepostupuje při výkladu od výrazu k jeho významům, nýbrž naopak od významu k nejrozličnějším možným výrazům, které ho mohou vyjadřovat, a to v daném kontextu nebo oboru úvahy. Toto pojetí extenzionálně přibližuje význam termínu informační vědě.

V ní se totiž tezaurus definuje jako řízený a měnitelný slovník výrazů selekčního jazyka sloužící k popisu, ukládání a vyhledávání dokumentů (nebo jejich částí) v informačních systémech. Tato definice obsahuje několik dalších odborných výrazů, které je třeba vysvětlit, aby byla nejen přesná, nýbrž i jasná. Vždyť co se míní "selekčním jazykem" a co jsou ony "informační systémy"? Ačkoli v knihovnické a informační vědě dodnes nepanuje mezi jednotlivými autory terminologická jednota v odpovědích na tyto zdánlivě jednoduché otázky, přece však jsou tyto odpovědi, ač různé, v jakési základní typologii známy. Nepovažuj i tedy za nutné rekapitulovat různá terminologická řešení dosud předložená nebo navrhovaná odbornou knihovnicko-informační obcí. Spíš bych se rád pokusil tyto tradiční odpovědi nazřít z poněkud širší a, jak doufám, i hlubší perspektivy, než je ta, kterou běžně nabízí scientistně redukcionistické paradigma charakteristické pro hlavní proud soudobé informační vědy.

Zkusme se podívat na ony elementární pojmy způsobem, který do centra pozornosti opět vrací toho, jenž se v

přístupu informační vědy, často orientované na efektivitu informačních strategií, poněkud vytrácí: totiž člověka a jeho jedinečnost. Takové překračující hledání nás zřejmě zavede na půdu, na níž již nevystačíme s habituálním pojmoslovím speciálních věd - na nejistou, riskantní půdu filozofického tázání. Snad jedině filozofie má odvalu fundamentálním způsobem problematizovat jistoty, které ostatní vědy zpochybňovat nejen nechťejí, nýbrž ve své dobrovolně si uložené metodické kázni často ani nemohou.

Informační systém

"Informačním systémem" se zpravidla chápe velmi obecně jako jakékoli objektivně existující a relativně ohraničitelné místo (množina prvků), jehož základní funkcí je optimalizovat společenskou informační komunikaci; informační systém je takovým místem transformace informací, které zvyšuje stupeň uspořádanosti a přehlednosti v dané škále společenských informačních toků. Jestliže soubor informačních toků vstupujících do informačního systému vykazuje jistý stupeň entropie, je souborná entropie informačních toků, které ze systému vystupují, podstatně snížena právě působením transformačních procesů uvnitř informačního systému. Navzdory obecnosti takto chápaného informačního systému se tímto pojmem v praxi nejčastěji rozumí knihovnicko-informační instituce, tedy knihovna, databanka, případně subsystémy těchto systémů (např. jednotlivá databáze).

Míru nárůstu negentropie (tj. míru snížení entropie) informačních toků po průchodu systémem je tedy možno chápat jako kritérium efektivity daného informačního systému: oč uspořádanější (a tudíž pro uživatele snáze dostupné) budou informace obsažené v dokumentech, které knihovna zpracovává, o to účinnější je působení této knihovny při optimalizaci komunikace společenského poznání v daném oboru. Tento kyberneticko-technický pohled na věc je svůdný, avšak přesnost takových konstatování je jen zdánlivá a při bližším pohledu vyžaduje jistou filozofickou korekci. Zejména proto, že ve všech těchto procesech se přímo či nepřímo angažuje člověk; tyto ryze synchronní funkční modely však neberou v potaz lidskou subjektivitu ani historický čas, v němž se odehrávají. Než se dostaneme k těmto bodům, bude třeba se zmínit o jednom z míst, jímž historický a kulturní čas do informačních systémů vniká. Takovým místem jsou veškerá jazyková rozhraní informačního systému. Pokusme se nyní věnovat jednomu z nich - selekčnímu jazyku.

Selekční jazyk

"Selekční jazyk" je možno definovat jako každý ustálený znakový systém, jenž slouží k formální nebo obsahové identifikaci dokumentů {nebo informací v nich obsažených} a kterého se užívá pro potřeby zpracování, ukládání a vyhledávání informací.

Snažíme-li se zachytit, popsat obsah určitého informačního dokumentu (tj. hledáme-li odpověď na otázku

“0 čem pojednává daná kniha, článek apod.?”), provádíme jeho obsahovou analýzu. Její výsledek je třeba nějak znakově fixovat, nejlépe v takové podobě, která pokud možno co nejméně podléhá subjektivitě toho kterého analytika.¹⁾ Právě jako praktický nástroj této objektivující fixace si archiváři a knihovníci v průběhu dějin vytvořili nejrůznější popisné soustavy čili selekční jazyky, např. Mezinárodní desetinné třídění, hesláře předmětových katalogů a ovšem také tezaury slov (a slovních spojení) preferovaných při popisování obsahu dokumentů a jejich následném vyhledávání.²⁾

Třebaže některé ze selekčních jazyků jsou již poměrně starobylé, nacházejí - někdy v poněkud modifikované podobě - podstatné uplatnění i v době supermoderních počítačů a nejvyspělejších informačních a komunikačních technologií. Ani ty totiž nemohou opominout či dokonce zcela popřít základní fakt, že ne-li všechno, tedy významná část našeho poznání je vždy nějak vázána na jazyk. Poznání musí tedy být koneckonců převoditelné do znaků přirozeného jazyka, aby bylo sdělitelné v sociální komunikaci.

Lze si ovšem představit, že jistá část porozumění probíhá na jakési podprahové, subverbální, intuitivní úrovni a je jen velmi obtížně (pokud vůbec) postižitelná slovy a tím méně chladně neosobním pojmovým diskurzem speciálních věd. Rovněž informační věda provádí rozsáhlou redukci: nejen že se při obsahové analýze dokumentu/informace programově nezajímá o tuto vrstvu porozumění, jakoby zasutou pod nánosem slovních formulací, nýbrž dokonce i na samotné slovní rovině metodicky odhlíží téměř od všeho, co z ní není uchopitelné aparátlem pojmového myšlení. To, co nelze abstrahovat v pojem, se při obsahové analýze, jak ji předepisuje tradiční knihovnická metodika, z dokumentu vytrácí: do notace selekčního jazyka se - až na výjimky³⁾ převádí jen nociónální rovina obsahu dokumentu/informace. Posléze je tato rovina opět zredukována, a to přetlumočením do příslušného *selekčního* jazyka, kteréžto přetlumočení, jak sám název jazyka napovídá, je nutně (protože metodicky) *selektivní*, a tudíž fragmentární.

Chápeme-li *smysl* textu hermeneuticky, ne tedy jako pouhou formálně-sémantickou *intenzi*, nýbrž jako *Ty* textu, s nímž se může bytostně setkat mé interpretující *Já*, pak shledáváme, že *smysl* se přinejmenším trojnásobně odcizil sám sobě. Překládat dokument do selekčního jazyka znamená totiž postupně:

1. vybrat několik nejhrubších sémantických zlomků, v něž byla roztržena původní významová celistvost dokumentu,
2. abstrahovat je ve víceméně izolované paradigmatické pojmy, a ty pak
3. vtěsnat do cizorodých pojmových schémat, která, jak potvrdí každý indexátor, nám mnohdy vnucují pojmovou hierarchii neslučitelnou s pojmovou strukturou původního dokumentu.

Z hlediska hermeneutického rozumění (které pragmatická informační věda nepovažuje za relevantní, ale o něž

nám zde jde právě proto, abychom ukázali, oč se tím

ochuzuje) je setkání s textem bytostným setkáním ducha se sebou samým [srv. 2, s. 23]. Je zřejmé, že setkání s textem, k jakému dochází mezi informačním profesionálem a dokumentem v průběhu rutinního indexování, má radikálně jinou povahu a je do jisté míry popřením oné bytostnosti, o níž hovoří hermeneutika.

Odhlédneme nyní od utilitárního hlediska každodenního knihovnického provozu, pro který je právě takové popření bytostného setkání jakousi zcela nezbytnou daní za produktivitu práce, bez níž by každá knihovnicko-informační instituce brzy zkrachovala; ptejme se raději, proč je požadavek pracovní výkonnosti (alespoň částečně) v tak radikálním rozporu s hermeneutickými principy. Hned při zběžném pohledu nalézáme dvě základní překážky, jež brání hlubšímu porozumění textu:

1. knihovník nemá *čas* důkladně prostudovat celý dokument. Na seznámení se s jeho obsahem má v nejlepším případě jen několik málo minut, i nemůže se než soustředit na nejhrubší obsahové rysy, které při setkání vnímajícího vědomí s textem vystupují na povrch jako první a nejsnáze dostupné znaky: obsah s názvy kapitol, úvod, resumé, informace na záložce obálky apod. Cílem jeho práce není hluboký vhled do daného textu ani jakékoli intimnější porozumění, nýbrž toliko rychlé, letmé a nutně povrchní sevření dokumentu do několika abstraktních a extenzionálně často příliš širokých pojmů, které zhruba charakterizují jeho téma a které jsou - a tak se dostáváme ke druhé překážce
2. *přeloženy* do daného selekčního jazyka. Je to specifický překlad: jestliže při translaci z jednoho přirozeného jazyka do jiného dochází ke ztrátě oněch kompenzičních prvků originálního textu, které se z nejrůznějších příčin jeví jako nepřeložitelné, pak při překladu z přirozeného jazyka do jazyka selekčního je tato ztráta daleko podstatnější, mj. i proto, že je do jisté míry metodicky požadovaná.

Informační věda ovšem přesně rozlišuje, co je *žádoucí informační redukce* obsahu dokumentu v jeho selekčním obrazu (např. eliminace tzv. nerelevantních obsahových prvků, marginálních témat apod.) a co je naopak *redukce nežádoucí* ~např. je-li ústřední téma dokumentu netransponovatelné do sémanticko-syntaktických struktur toho kterého selekčního jazyka, takže selekční obraz dokumentu je zkreslený a tudíž zavádějící při následném vyhledávání informací).

O redukci prvního typu se hovořívá jako o *informační komprimaci*, zatímco u druhé se mluví o *informačních šumech a bariérách*. Praxe však ukazuje, že mezi oběma krajními póly existuje celá škála přechodů a hranice mezi nimi mohou být nejen spojité, ale přímo “fuzzy”: závisejí totiž na hledisku, z něhož se provádí následná selekce informací, neboli na aktuálních informačních potřebách, jež nejsou neměnné a vyvíjejí se v čase. Co bylo včera ve věcném popisu dokumentů považováno za žádoucí komprimaci, může být již dnes informačním šumem. Tato ambivalence se podle mého soudu v informační vědě dosud málo zdůrazňuje, a to možná právě proto, že nám

často chybí celostní pohled, který ani při studiu informačních procesů neztrácí ze zřetele člověka a jeho jedinečnost. Právě filozofické přístupy zde mohou nabídnout plodnou korekci jednostranností úzce chápané informační vědy.

Vidíme, že co je žádoucí z hlediska fungování knihovny a tzv. optimalizace informačních toků, nemusí být nutně žádoucí z hlediska člověka v jeho existenciální hloubce a integritě.

Můžeme rekapitulovat: Při obsahové analýze a indexování dokumentů znaky selekčního jazyka vzniká selekční obraz dokumentu, který se liší od hlubšího setkání s textem v hermeneutickém smyslu, a to v důsledku celé soustavy bariér redukujících původního ducha textu. Jde o bariéry:

1. časovou: nedostatek času k interpretaci;
2. nociónální: eliminace všeho nepojmového v textu;
3. metodické redukce: požadavek extrakce toliko hlavních ("relevantních") informací a jejich komprimace;
4. překladovou: podstatná výrazová a významová inkongruence mezi přirozeným a selekčním jazykem (běžná mezi přirozenými jazyky) je násobena selektivitou selekčního jazyka.

Ani tento výčet není úplný a nezahrnuje zdaleka všechny problematizující faktory, jež se uplatňují při proceduře přeměny obsahu dokumentu v jeho selekční obraz za pomoci tezauru. Všimněme si např. celé palety problémů spjatých s bariérou, již jsem nazval nociónální: co vše ztrácíme převodem obsahu textu do pojmových hierarchií vůbec a hierarchií tezauru zvlášť?

Pojmy versus čas čili antropologická chudoba pojmů

Ve svém rozboru Bergsona píše M. Petříček o kardinálním rozporu, do něhož se pojmové myšlení dostává v okamžiku, kdy se chce zmocnit času: "Jakmile vědecké myšlení naráží na kontinuum, tj. na cosi, co plyne anebo trvá, pak takové jevy buď vůbec eliminuje, anebo je převádí na rozčleněné soubory věcí. To proto, protože pouze pokud je skutečnost takto segmentována, pouze tehdy může být uchopena pojmy. (. . .) Věda není s to uchopit nepřetržitý pohyb a tím méně čas" [6, s. 40-41]. Platí-li toto tvrzení o vědě a vědeckých textech, platí ještě koncentrovaněji o tezaurech, jež navzdory všem moderním pokusům o jejich sémantickou dynamizaci jsou v okamžiku obsahové analýzy dokumentu víceméně statickými pojmovými soustavami.⁴⁾

Jestliže nám ovšem uniká čas, dostáváme se rovněž do problému, jak zachytit *dějinnost*. Třebaže právě dějinné vědomí hraje klíčovou úlohu v lidském poznávání a sebepoznání, synchronní perspektiva tvůrců tezaurů se s ním dostává do rozporu: pojmové hierarchie reflektované tezaurem jsou vždy historicky relativní, a tedy drive či později se zákonitě musejí dostat do konfliktu s postupujícím stavem poznání, jež je zachyceno v dokumentech popisovaných tím kterým tezaurem. Od určitého okamžiku tezaurus zastarává a stává se brzdou právě oné efek-

tivní komunikace společenského poznání, kterou měl původně urychlovat a již měl napomáhat. V tom podle mého názoru spočívá obrovská paradoxie věčných selekčních jazyků: že ač je jejich základní funkcí optimalizovat společenskou komunikaci, tím, že nové poznání opakovaně pořádají do pojmových schémat historicky staršího poznání, mohou v některých případech neotřelým pojmovým vzorcům bránit, aby se nově prosadily a nahradily tak zastarávající pojmová schémata.

Tím, že odrážejí dějinné vědomí, jsou tezaury stejně jako ono samo historicky relativní. Tezaury jsou ovšem determinovány nejen dobou, nýbrž také místem svého vzniku, tedy společností a její kulturou. Od dějinného uvědomění vede tedy přímá cesta k uvědomění sociálnímu a kulturnímu, a s dějinnou relativitou souvisí kulturní relativismus. A právě kulturní relativita je prostor, kde do teorie selekčních jazyků vstupuje kulturní antropologie. Nicméně tím se již dostáváme k další samostatné kapitole, k níž tyto poznámky měly být pouhým úvodem: kapitole o kulturně antropologických aspektech indexování.

A tak otevřel-li tuto esej citát filozofa, mohl by ji ukončit citát kulturního antropologa: ne však jako uzavření daného problému, nýbrž spíše jako otevření jeho další kapitoly, v jejímž záhlaví by, jak se domnívám, mohl dobře být a která by se zabývala dalšími nenomoteckými aspekty paradigmatu informační vědy: "*Člověk je bytostí, která je zachycena do sítě významů, kterou si sama vytvořila. Tuto síť můžeme považovat za kulturu. Její zkoumání proto není experimentální vědou, která hledá zákony, nýbrž vědou interpretující, která hledá významy.*" [3]

Je pozoruhodné, že tento výrok, ačkoliv otevírá již další kapitolu, se vrací k témuž východisku jako úvodní citát Gadamerův.

Poznámky:

- 1) Takové překonání subjektivity bývá však úspěšné pouze částečně a v omezené míře, a to z principiálních epistemologických důvodů: domnívám se, že z žádné interpretace nikdy zcela nevymažeme osobnost interpreta. Kde je interpretace, je i interpret. M. Petříček v této souvislosti poznamenává: "Náš svět je vždy již nejrůznějším způsobem *interpretovaný*, jinak řečeno: náš přístup ke světu je ve většině případů vždy již zprostředkovaný." [6, s. 22]
- 2) Obsahová analýza dokumentu je ve skutečnosti složitý poznávací proces, zde jen velmi hrubě a zjednodušeně načrtnutý: odehrává se v interakci mezi obsahem dokumentu/informace, konkrétním selekčním jazykem a analytikem (indexátorem), v jehož vědomí, vždy již podmíněném přirozeným jazykem, se postupně vytváří tzv. mentální selekční obraz dokumentu, který se snaží v redukované podobě postihnout jeho podstatné obsahové charakteristiky a posléze je převést do pokud možno významově ekvivalentních výrazů daného selekčního jazyka.
- 3) V případě tezaurů mohou být takovou výjimkou tzv. identifikátory, jež mohou postihnout jedinečného člo-

věka - ovšem jen za předpokladu, že je sám tématem daného dokumentu. Jestliže však identifikace osoby je jednoznačná (zpravidla ji reprezentuje jméno a data narození, případně úmrtí, někdy ještě stručná zmínka o profesionální afilaci), neznamená ani zdaleka respekt k jakýmkoli osobním zvláštnostem člověka. Ty jsou metodicky eliminovány jakožto *informačně irrelevantní*. I jedinečnost je tedy odosobňována, abstrahována v jakýsi univerzální vzorec, v němž z původního člověka zbývá jen značka, jakési onomatologické ukazovátka, jehož účel je výlučně přidělit jednoznačné jméno objektu, který je sice jedinečný, ale jehož jedinečnost nás vlastně nezajímá. Jde zkrátka o to přidělit jedinečnosti jakési minimální jedinečné jméno, minimální v tom smyslu, že nám nejde o jedinečnost samu, nýbrž o její co nejkratší pojmenování, které již je jednoznačné: jednoznačnost jména je důležitější než jedinečnost člověka jím označeného. Vztah mezi člověkem a jeho identifikátorem se tak podobá vztahu mezi spisovou značkou a úředním spisem.

- 4) Do procesu věcného popisu (indexování) pomocí tezauru vnáší jistou flexibilitu, a tedy i sémantickou dynamičnost *sui generis*, nepochybně sám fakt postkoordinovanosti selekčního jazyka deskriptorového typu. To je však jev poněkud jiného řádu a nesouvisí primárně se směřováním našeho argumentu o relativní rigiditě tezaurových pojmových rámců.

Literatura:

1. Čermák, F.: Jazyk a jazykověda. Praha, Pražská imaginace 1994. 251 s.
2. Gadamer, H.-G.: Problém dějinného vědomí. Praha, Filosofie 1994. 53 s.
3. Geertz, C.: The Interpretation of Cultures. New York, Basic Books 1973. 470 s.
4. Hendl, J.: Hermeneutika - metodologické poznámky. In: Hogenová, A. (ed.): Hermeneutika sportu. Praha, Univerzita Karlova 1998, s. 33-40.
5. Murphy, R. F.: Úvod do kulturní a sociální antropologie. Praha, Sociologické nakladatelství 1998. 267 s.
6. Petříček, M.: Úvod do (současné) filosofie. Praha, Herrmann & synové 1997. 178 s.
7. Rowley, J.: The controlled versus natural indexing languages debate revisited : a perspective on information retrieval practice and research. Journal of Information Science, 20 (2), 1994, s. 108-119.