

Data, informace, znalosti

Ondřej Šlapák

Katedra informačních technologií

Vysoká škola ekonomická

slapak@vse.cz

Aquasoft, spol. s r.o.

o.slapak@aquasoft.cz

***Abstrakt:** Tento článek vysvětluje pojmy data, informace a znalosti. Vychází z dosavadních poznatků, které dává do souvislostí a rozšiřuje je. Problematikou se zabývá i pragmaticky, z ekonomického hlediska.*

Klíčová slova: data, informace, znalosti, peníze, rozhodování, informační systém

1. Úvod

Informační společnost, o které se v současnosti tolik hovoří a říká se, že v ní žijeme, můžeme charakterizovat stručně i tak, že **postupy dosahování zisku jsou v ní založeny na intenzivním a inteligentním používání informací**, což myslím vyjadřuje podstatu ekonomického nazírání na informační společnost, které osobně řadím tzv. „až na první místo“.

Hovoříme-li o používání informací, měli bychom přesně vědět, co tento pojem znamená. Další text článku již bude laděn spíše filosoficky a odchýlí se tak od pragmatičnosti mé osobní „ekonomické definice“ informační společnosti. Definice pojmů jako jsou data, informace a znalosti se dnes může zdát jako obehnaná písnička. Dovolte mi přesto přidat další sloku. Tento článek přispívá k vysvětlení těchto pojmů, ve kterých se stále ještě chybí. Věřím ale, že článek bude přínosem i těm, kteří v těchto pojmech mají jasno.

2. Informace

Claude Shannon v matematické teorii komunikace vymezil informaci jako statistickou pravděpodobnost určitého signálu či znaku, který je na vstupu určitého systému. Čím je menší pravděpodobnost daného znaku, tím větší má takzvanou informační hodnotu (pokud má nějaký znak pravděpodobnost velkou, je spíše očekáván a jeho výskyt příjemce tolik nepřekvapí). Tím, že systém signál zpracoval, dostal se na nižší úroveň nejistoty (entropie), tedy do stavu s vyšší mírou uspořádanosti.

Na informaci lze nahlížet též ze statického hlediska. Sémiotika – vědní obor zabývající se studiem znakových systémů – informaci vymezuje volně řečeno jako působek, vjem, který splňuje tři požadavky. Prvním je syntaktická relevance. Subjekt, který vjem přijímá, musí být schopen vjem

vůbec detekovat a rozumět mu. Druhým požadavkem je sémantická relevance. Subjekt musí vědět, co vjem znamená, co vypovídá o něm a jeho okolí. Třetím požadavkem je pak pragmatická relevance. Působek musí mít pro přijímající subjekt nějaký význam.

Shannonova definice informace vyžaduje z pohledu ekonomického využití informace rozšíření. Dříve než ji modifikuji podle požadavků sémiotiky, je nutné ještě připomenout další přístup k pojetí informace¹: Subjekty jsou vystaveny působení signálů. Některé zachytí a porozumí jim. To je pro subjekt to, co nazýváme data. Data může subjekt ignorovat či je uložit pro pozdější zpracování, transformovat do jiné podoby atd. Jsou vyjádřena fyzickým nosičem, ať už jde o inkoust a papír, elektrické signály či elektromagnetické záření. Data jako taková mají vypovídací schopnost. Podle sémiotiky jsou na úrovni sémantické relevance. [VODL97-01] data řadí pouze na syntaktickou úroveň a na úroveň sémantickou umísťuje zprávy jako obsah přiřazený datům. S tímto rozdělením souhlasím, nicméně pro další výklad není příliš podstatné. V ekonomické praxi je datům (jako pojmu) běžně přisuzován význam zpráv. V dalším výkladu proto zůstanu u tohoto širšího významu². Jestliže subjekt data momentálně používá k rozhodování, stávají se pro něj informací (pragmatická relevance). Subjekt datům přiřazuje význam, smysl. Subjekty za tímto účelem data shromažďují. Proto je někdy datům přiřazován nejen význam zpráv, ale také informace. Můžeme říci, že data jsou potenciálními informacemi. Tím se ocitám v podobné situaci jako s daty a zprávami. Pojem informace se používá jak pro skutečné informace, tedy data v běžícím procesu rozhodování, tak pro data, která představují pouze potenciální informaci.³

V tomto odstavci se vracím k předeslané modifikaci Shannonovy definice informace. Podle mého názoru **je informace působek na subjekt (systém) splňující všechny požadavky sémiotiky, který mění pravděpodobnosti jednotlivých možných cest, kterými se subjekt může vydat (stavů, do kterých se systém může dostat).**

3. Informace a snížení nejistoty

Z pohledu výše uvedené modifikace je nutné odpovědět na otázku, zda informace vždy snižuje nejistotu. Ve svém velmi zajímavém výkladu pojetí a významu informace v [MARW95-01] autor uvádí, že například informování občanů o tom, že se blíží hurikán, rozhodně nepřispívá ke snížení jejich nejistoty. Myslím, že jde o nesprávné pojmenování jevů. Jak jsem zmínil, datům je přiřazován význam zpráv i informací. Fakt, že se blíží hurikán, není informací, ale zprávou. Jestliže však se některý subjekt rozhodoval, zda se v dané lokalitě zdrží ještě několik hodin, daná zpráva je pro něj skutečně informací, neboť značně snížila pravděpodobnost toho, že subjekt zůstane. Pro lidi, kteří v lokalitě bydlí, tato zpráva znamená naopak zvýšení nejistoty. Tato zpráva přesunula subjekty do zcela jiného stavu. Generuje nutnost rozhodování.

Jestliže se na celou situaci podíváme v širším pojetí, dojdeme k jinému závěru. Lidé, kteří v lokalitě žijí, již vědí, že občas nějaký hurikán přijde. V období mezi hurikány se ovšem chovají, jako by hurikány nikdy nebyly. V mezích časového úseku mezi hurikány zpráva o blížícím se živlu skutečně zvyšuje nejistotu. Objektivně je ale tento stav jeden z možných. Zpráva o hurikánu snižuje pravděpodobnost setrvání ve stavu charakteristickém pro období mezi hurikány a zvyšuje pravděpodobnost přechodu ke stavu evakuace. Zvyšuje v lidech přesvědčení (jistotu), že je nutné přijmout příslušná opatření. Toto přesvědčení je v lidech v normálním stavu bez hurikánu velmi potlačeno.

¹ Spíše jde o doplnění sémiotiky.

² Pokud bude potřeba, výslovně rozdíl označím.

³ v dalších kapitolách budu používat pojem informace právě i ve významu potenciální informace. I zde platí, že bude-li to nutné, výslovně upozorním na skutečný význam pojmů.

Autor měl zřejmě na mysli poněkud odlišnou interpretaci. Každý subjekt má nějaký myšlenkový obraz světa. V rámci tohoto myšlenkového obrazu přechází z jednoho stavu ke druhému, a to na základě informací. Možné stavy i použité informace jsou víceméně očekávané a do jisté míry známé či předvídatelné. Situaci lze ilustrovat hledáním dat o nějakém produktu na Internetu z důvodu být informován v procesu rozhodování při jeho koupi. Prohledávající subjekt dosud ví, že je možné rozhodovat se podle určitých kritérií. Webové stránky mu podávají informaci, že produkt s takovou a takovou jakostí (podle známých kritérií) je dostupný v určitém obchodě. Pak ovšem subjekt klikne na odkaz, který zobrazí stránku popisující produkt pomocí dalších kritérií, které dosud neznal. Subjekt je s rozhodováním opět na začátku, v nejistotě. Tuto problematiku lze zobecnit na celý myšlenkový obraz, matematicky pak znázornit vektorovým prostorem. Myšlenkový obraz jednotlivce je jen podmnožina universa, tedy všeho, co jest. Universum je matematicky n -rozměrný vektorový prostor, přičemž n neznáme a může být „rovno“ nekonečno). Jestliže některá zpráva způsobí průlom hranice myšlenkového obrazu subjektu, dojde k jeho rozšíření, avšak jednotlivé nové možné stavy subjektu nejsou vůbec známy (natož jejich pravděpodobnosti). V pojetí vektorového prostoru je to jako rozšíření z dvourozměrného do třírozměrného prostoru, či lépe z čtyřrozměrného do pětirozměrného, což z našeho časoprostorového chápání výstižněji charakterizuje vznik nejistoty. **Z pohledu universa je každá příchozí zpráva informací a snižuje tak nejistotu subjektu. Z pohledu subjektu příchozí konkrétní zpráva může nejistotu zvýšit. Další zprávy pak tuto zvýšenou nejistotu snižují. Subjekt se tak dostává objektivně stále k vyšší (širší) úrovni jistoty, neboť se snižuje pravděpodobnost (je-li n menší než nekonečno), že přijde zpráva, která opět nejistotu zvýší.**

4. Znalosti

S informacemi neodmyslitelně souvisejí znalosti, jak ostatně vyplývá i z předchozího odstavce. Abychom se problematikou informace mohli vůbec zabývat, data musí být svázána s nějakým systémem, na který působí, což lze vyvodit ze sémiotiky i Shannonovy definice informace. V [VODL97-01] autoři zprostředkovávají schéma od imunologa J. E. Salka. Schéma zobrazuje tři kvality systémů: fyzikální, biosféra a metabiosféra. Kvality jsou vymezeny emergentními vlastnostmi a vznikají v průběhu evoluce.

Tyto emergentní vlastnosti jsou přímo spojeny s pojmy data, informace a znalosti. Fyzikální sféra je charakteristická tím, že působky deterministicky mění stav systémů (pokud se nepohybujeme na úrovni, kde převládají kvantové jevy). Stejně postupy vedou ke stejným výsledkům. Systémy biosféry se navíc dokáží adaptovat. Jestliže se systém dostane do téměř stejné situace, ve které se již nacházel, ke změně svého stavu již nepotřebuje další informaci zvenčí. Systémy z metabiosféry nepotřebují dodatečnou informaci ani tehdy, když se nacházejí v situaci, kterou dosud neprošly⁴, neboť jsou schopny na základě svých znalostí odvodit si další postup. Otázkou zůstává, do jaké míry jsou deterministické i systémy biosféry a metabiosféry. Zatímco u systémů biosféry bych se přikláněl k tomu, že deterministické do jisté míry jsou, u systémů metabiosferických o tom přesvědčen nejsem, což neznamená, že bych to odmítal.

Co to znalosti jsou? Podle [VODL97-01] jsou znalosti výsledkem aktivního učení se. V této literatuře i v [MARW95-01] nalezneme i názor Petera F. Druckera, podle kterého jsou znalosti potřebné pro konverzi dat v informace. [MARW95-01] uvádí také definici Roberta M. Hayese. Ten říká, že znalosti jsou výsledkem porozumění informací, která byla právě sdělena, a její integrace s dřívějšími informacemi. Můj názor je ten, že stanovení, co jsou znalosti, je nerozlučně spjata s popisem jejich používání, které je dále uchopitelné jen přes studium vyšší mozkové činnosti – viz např. [PSTK98-01] či [NOVM93-01] – a pravděpodobně i se studiem emocí, neboť myšlení je jimi ovlivňováno. Protože tato tematika není předmětem dalšího postupu v tomto článku, pro přehled zcela postačuje schéma

⁴ Nejedná se o každou situaci, ale o situaci, která je (v pojetí vektorového prostoru uvedeném výše) prvkem stejného vektorového prostoru jako situace, kterými již systém prošel.

uvedené v [VODL97-01], v němž to, co člověk zná o sobě a svém okolí, používá pro mentální proces zpracování vstupu ve výstup. Na základě působení tohoto výstupu na okolí i na něj samotného upravuje své dosavadní poznatky, které rozšiřuje i cíleným učením se. Svými slovy bych řekl, že znalosti subjektu jsou obrazem světa, ve kterém subjekt žije. Obraz světa je reprezentován systémem dat jako obrazem jednotlivých percepčí. Vztahy mezi daty jsou obrazem skutečných vztahů mezi zobrazenými percepčemi. Tento systém je neustále ovlivňován novými percepčemi, k jejichž zpracování je využit. Tento systém sám není schopen fungování, je jen nástrojem řídicí jednotky, která ho může cíleně rozšiřovat. **Znalosti lze charakterizovat také jako informace o tom, jak využít jiné informace a data (a to i ve vzájemných kombinacích) v různých životních situacích.**

5. Emotivní podtext informace

Vedle požadavku pragmatické relevance je možné u informace sledovat i to, čemu říkám emotivní podtext. Vezměme si tři verze prohlášení politika:

- „Nezaměstnaných příliš nepřibývá.“
- „Nezaměstnaných stále neubývá.“
- „Nezaměstnaných je téměř stále stejně.“

První věta má nabudit optimismus. Druhá je stížností, ale v podstatě říká to samé. Třetí věta má potvrdit současný stav. **Interpretace zprávy, využití informací a znalostí a jejich tvorba jsou přímo ovlivněné emocemi na obou stranách komunikace. Subjekt, který sdělení předává, může navíc vedle přirozených emocí do zprávy vložit právě emotivní podtext s tím, aby ovlivnil posluchačovo vnímání sdělení.**

6. Informace, znalosti, peníze, řízení informací

V této pasáži se opět přiblížím k ekonomickému pohledu na celou problematiku. Setkal jsem se s názorem, že v informační společnosti jsou informace přinejmenším stejně důležité, jako peníze, možná důležitější. Nechci působit jako odpůrce moderního myšlení, ale řadit informace před peníze se mi zdá nadnesené. Jak uvádím v jiném svém článku ([ŠLAO03-01]), peníze jsou odrazem toho, co člověk přinesl společnosti. Jsou tedy odrazem minulosti a jsou jistotou (abstrahujeme nyní od krádeží, nedostatečných úrocích na termínovaných vkladech apod.). Přímo představují potenciální spotřebu jednotlivce. Zisk (a tedy spotřeba) je hnací silou veškerého ekonomického úsilí. Peníze člověk prostě má (pokud je má). Informaci může člověk za peníze získat. Proč? Protože od toho očekává více peněz, než kolik zaplatil. Výnos z použití informace je však nejistý. Informace je v tomto pohledu předobrazem budoucnosti. Samozřejmě záleží na konkrétní informaci. Podle toho by se toto výnosové aktivum zařadilo do investičního portfolia.

Informace, znalosti a peníze spolu úzce souvisí. V informační společnosti dochází k vyšší tvorbě zisku při využití stejných tradičních výrobních zdrojů. Katalyzátorem jsou právě informace a znalosti. Informace a znalosti můžeme za peníze koupit, pokud nemáme vlastní. Aby přinesly dodatečný zisk, je nutné jejich inteligentní využití spolu s novým stylem řízení. Managementem v informační společnosti se zabývá např. [VODL99-01], [TRUJ97-01], [TRUJ99-01].

Kromě nových manažerských přístupů, organizací podniku a způsobu práce s informacemi bych chtěl vyzvednout nutnost práce s informacemi a znalostmi nejen v primárním významu (jejich využití k tvorbě hodnoty), ale také ve významu jejich správy, řízení. Informace a znalosti jsou zdrojem. Jestliže podnik vede agendu majetku, proč ne agendu informací a znalostí? Je proto nutné vést **informační systém (IS) o informacích**. Takový IS jednak obsahuje primární informační systém poskytující obsah informací a znalostí, jednak přehled, která informace či spíše zpráva odkud pochází,

jak byla nákladná, kdy byla zaevidována, kdy byla k čemu použita, v jakých souvislostech, jaký byl výsledek jejího použití (znalostní i finanční), její relevance předmětová, časová i teritoriální atd. Aby bylo možné řízení informací a znalostí, informační systém musí umožnit sledovat vývoj přidané hodnoty každé evidované zprávy a znalosti.⁵

Spolu s růstem intenzity využívání informací a znalostí pro tvorbu hodnot sledujeme také tlak na rychlost rozhodování. Praxe i literatura ([DYSE01-01], [GATB99-01]) říká, že důležitější je rozhodnutí spíše rychlé, než kvalitní. Osobně rychlost před kvalitou nepreferuji, nicméně je jasné, že doba pro rozhodnutí nesmí překročit čas, který je vymezen existencí problému. Pak může automaticky nastat taková varianta, která by byla horší, než ta, kterou bychom vybrali i bez hlubšího zvažování. Srovnání, který přístup je lepší, je možné jen přes změření jejich přínosů. Prakticky je to neproveditelné, neboť rozhodneme-li se pro rychlé řešení, nemůžeme současně provést řešení téhož problému kvalitně (pokud neprovádíme kvalitní řešení rychle, což by bylo ideální). Musíme jen doufat, že rychlá rozhodnutí přinesou v sumě vyšší přidanou hodnotu, než by mělo jedno kvalitní. Kvalitu a rychlost rozhodování pozitivně ovlivňují simulační systémy, s jejichž pomocí můžeme přece jen porovnání do jisté míry provést. V dnešním turbulentním prostředí však výsledek simulace nemůžeme brát se zárukou.

7. Shrnutí a závěr

Tento článek je jedním z těch, které se snaží vysvětlit základní pojmy data, informace a znalosti. Činí tak s odvoláním na informační společnost, které jsme snad součástí. Vysvětlování pojmů nezačíná postupně od dat, ale spíše ve vzájemných vztazích souvislostech dat, informací a znalostí. Využívá Shannonovy definice informace, poznatků sémiotiky i dalších zdrojů a pokouší se je dát do souvislostí. Článek se dále zabývá otázkou, zda informace (jak říká Shannonova definice) vždy přispívá ke snížení nejistoty. Nabízí vyvrácení i potvrzení tohoto tvrzení, a to z různých pohledů. Navazuje pak tématem znalostí a tématem emotivního podtextu informace. Poslední pasáž se na informace a jejich využívání dívá z ekonomického hlediska. Dává vysvětlení, že informace nelze řadit před peníze, neboť informace jsou jen zdrojem tvorby peněz. S tímto souvisí i krátké pojednání o vztahu informací a jejich zpracováním a využíváním v dnešním rozhodování.

Data, informace a znalosti zřejmě zůstanou i nadále živým tématem. Ať už je jejich definice jakákoliv, je jasné, že bez nich to půjde čím dál hůře. S nimi to může jít lépe, budeme-li schopni s nimi dobře pracovat.

Literatura

- [DYSE01-01] Dyson, E.: „Release 2.1 – vize života v digitálním věku“, Management Press, Praha, 2001, ISBN 80-7261-030-9
- [GATB99-01] Gates, B.: „Byznys rychlostí myšlenky, jak upět v digitálním věku“, Management Press, Ringier ČR, Praha, 1999, ISBN 80-85943-97-2
- [MARW95-01] Martin, W.J.: „The Global Information Society“, Aslib Gower, England, 1995, ISBN 0-566-0775-9
- [NOVM93-01] Novák, M., Faber, J., Kufudaki, O.: „Neuronové sítě a informační systémy živých organismů“, Grada, a.s., Praha, 1993, ISBN 80-85424-95-9

⁵ S tím může být v protikladu nemožnost vést informace a znalosti jako aktiva v účetnictví, jak uvádí [MARW95-01]. Legislativa by v tomto ohledu neměla ztrácet krok.

- [PSTK98-01] Pstružina, K.: „Svět poznávání“, Nakladatelství Olomouc, Olomouc, 1998, ISBN 80-7182-074-1
- [ŠLAO03-01] Šlapák, O.: „Totální elektronizace peněz“, časopis Bankovníctví, číslo 6-7 a 8 / 2003, Economia, a.s., 2003, ISSN 1212-4273
- [TRUJ97-01] Truneček, J.: „Management v informační společnosti“, VŠE, Praha, 1997, ISBN 80-7079-201-9
- [TRUJ99-01] Truneček, J.: „Systémy podnikového řízení ve společnosti znalostí“, VŠE, Praha, 1999, ISBN 80-7079-083-0
- [VODL97-01] Vodáček, L., Rosický, A.: „Informační management, pojetí, poslání a aplikace“, Management Press, Praha, 1997, ISBN 80-85943-35-2
- [VODL99-01] Vodáček, L., Vodáčková, O.: „Management, teorie a praxe v informační společnosti“, management Press, Praha, 1999, ISBN 80-85943-94-8