

Digitálny informačný priestor

Milan Konvit

KMKD FHV Žilinská univerzita

Abstrakt

Kľúčovými fenoménmi dnešnej doby sú digitalizácia života a život v/na sieti. Tento príspevok sumarizuje základné vlastnosti digitálneho informačného priestoru z pohľadu jeho vhodnosti na sprístupňovanie digitálnych objektov kultúrneho dedičstva.

1. Koncept informačného priestoru

Pojem "informačný priestor" v sebe spája dva pojmy: "priestor" a "informácia". Aj keď ide o pojmy, ktoré dnes patria do tzv. bežnej slovnej zásoby, vysvetlenie ich obsahu je netriviálnou úlohou. Aj preto sú vo filozofickej literatúre kategórie "priestor" a „informácia“ jednými z najviac diskutovaných. Vezmime ako axiómu tvrdenie, že fyzikálny priestor je naplnený hmotnými objektmi a platia v ňom prírodné zákony. Takéto videnie priestoru sa opiera o Descartov koncept priestoru – priestor existuje len za podmienky, že je naplnený objektmi. Analogicky predpokladajme, že informačný priestor je, podobne ako fyzikálny priestor, zaplnený informačnými objektmi¹. Na rozdiel od reálnych objektov ale informačným objektom chýbajú Euklidovské rozmery. Preto ich musíme popisovať inými parametrami. Každý informačný objekt je nositeľom informácie², ktorá môže byť prezentovaná v rôznej podobe (audio, obraz, video, písané/tlačené symboly, atď.). Ide o priestor, ktorý, hoci je vybudovaný nad hmotným substrátom (nosičom informácie), sa stáva vnímateľným až aplikáciou definovaných procedúr na informačné objekty. Pritom informačné objekty môžu na seba naberať rôzne podoby: informačným objektom je dnes najčastejšie dokument v tej či onej forme, ale za určitých okolností sa takto môže javiť celá inštitúcia, po prípade jednotlivci. Podľa nášho názoru najvšeobecnejším konceptom informačného priestoru je Newbyho koncept:

„súbor objektov/konceptov a relácií medzi nimi udržiavaných (informačným) systémom“³

Tu sa uberáme skôr Leibnitzovským smerom – „priestor je súborom vzťahov (relácií) medzi objektmi“ aj keď pojmy vzdialenosť a smer musíme nanovo definovať.

Koncept digitálneho informačného priestoru

¹ Podľa TDKIV je informačný objekt „Informace nebo skupina informací tvořících jednotný celek bez ohledu na typ nebo formát.“ Informačný objekt plní v informačnom priestore rolu stimulu, ktorý má schopnosť produkovať informáciu pre používateľa – viď Gregory B. Newby Cognitive Space and Information Space <http://www.petascale.org/papers/cogstats-withimages.pdf> „we consider documents and other information objects to be stimuli with the ability to produce information (i.e., resolve ambiguity) in the user.“

² Pojem „informácia“ budeme, pre účely tejto kapitoly, chápať ako zmysel, vznikajúci interpretáciou abstraktných symbolov z dopredu dohodnutého súboru symbolov človekom.

³ Pôvodne Newby používal termín objekt aby ho neskôr nahradil pojmom koncept. NEWBY, G. B.: *The necessity for information space mapping for information retrieval on the semantic web. Information Research*, 7 (4).2002. Dostupné na <http://InformationR.net/ir/7-4/paper137.html>. [Citované 24.2.2012].

Pôvodný koncept informačného priestoru je „humanocentrický“. Ideu informačného priestoru vzťahuje výlučne na človeka. Cesta k vzniku digitálneho informačného priestoru tak vedie od priestoru, udržiavaného len v hlavách ľudí a dopĺňaného/aktualizovaného rečou, spevom, rituálnymi tancami a pod., cez priestor, doplnený o externé informačné objekty (kresby, písmo, neskôr tlač) k priestorom, podporovaným IKT. IKT zmenili procesy prehľadávania priestoru, dlhodobého uchovania informácií, vyhľadávania a získavania informácií. Všeobecne sa v DIP nachádzajú dva typy informácie – pôvodná digitálna a digitalizovaná. Aktuálne trendy v digitálnom informačnom priestore je možno zosumarizovať nasledovne:

- Exponenciálny nárast počtu informačných objektov
- Rastúca prevaha dát nad informáciami. (potreba interpretačnej vrstvy IO)
- Rast zložitosti informácií (po stránke obsahu, formy, štruktúry, ...)
- Rast previazanosti informácií (prelinkovanie)
- Rast neusporiadanosti priestoru a z toho vyplývajúca potreba dolovania informácií z dát.
- Priemernému blogerovi verí viac ľudí ako priemernému žurnalistovi,
- O výbere informačného zdroja viac rozhoduje provokatívny nadpis, ako vierohodnosť zdroja,
- Filozofia a la Facebook: vaši priatelia vedia lepšie, čo je skutočne dôležité ako šéfredaktor uznávaného média.
- Súčasná idea reality pochádza z druhej, resp. až z n-tej ruky. Idea reality je dnes editovateľná, filtrovateľná a všeobecne manipulovateľná jednotlivcami, médiami i online komunitami.
- Občianske (amatérske) napĺňanie informačného priestoru má stúpajúcu tendenciu (chaty, videá, blogy, ...).
- Monopolizácia vlastníctva informačného priestoru - globálny online informačný priestor bude vlastniť minimálny počet hráčov typu Google, Facebook.
- Snaha o vytvorenie softwarového personálneho asistenta, vyhodnocujúceho vlastné aktivity jednotlivca a aktivity jeho priateľov (vaši priatelia práve čítajú/majú radi toto) s prihladením k afinite (hustote kontaktov), relatívnej váhe obsahu (korelácia aktivít jednotlivca a priateľa) a aktuálnosti aktivít (čím staršia aktivita, tým menšia váha).
- Jednotlivec do informačného priestoru čoraz viac vstupuje ako člen komunity/skupiny a podieľa sa na vkladaní a vytváraní informácií kolaboratívnym spôsobom.

2. Digitálny informačný objekt

Iné je skúmať informáciu ako koncept, a iné je skúmať informáciu v kontexte priestoru. V druhom prípade berie informácia na seba navonok podobu objektu. Jej vlastnosti/atribúty/charakteristiky potom popisujeme analogicky ako atribúty objektu.

Základným stavebným prvkom digitálneho informačného priestoru je digitálny objekt. „Digitálny informačný objekt je diskretnou zbierkou informácií a dát spoločne s metadátami a referenčnou informáciou potrebnou na reprezentáciu zbierky ako izolovanej konceptuálnej entity a vrátane unikátneho identifikátora.“⁴

Fyzikálne digitálny informačný objekt predstavuje zápis znakov na danom médiu.

⁴ http://itlaw.wikia.com/wiki/Digital_information_object

Najdôležitejšími atribútmi digitálneho objektu sú⁵:

- Editovateľnosť.
- Interaktívnosť (cez aktivovateľnosť v ňom zabudovaných funkcií).
- Prístupnosť (pre používateľa i iné informačné objekty).
- Modifikovateľnosť (pre používateľa i iné informačné objekty).
- Otvorenosť (v zmysle definície otvoreného systému).
- Reprogramovateľnosť.
- Distribuovanosť (viacnásobné kópie toho istého objektu/časti objektu na rôznych miestach).

3. Projektovanie digitálneho informačného priestoru

Pri koncipovaní nového digitálneho informačného priestoru treba odpovedať na nasledovné otázky:

- aký má byť účel priestoru?
- Aké typy informácie má obsahovať?
- Aký typ navigácie je pre daný priestor optimálny?
- Aký je optimálny počet interných/externých liniek?
- Ako čo najviac rešpektovať zásady informačnej architektúry?
- Ako redukovať heterogenitu z hľadiska obsahu, formy, podoby informácie, atď. ?

Nutnou podmienkou na to, aby sa digitálny informačný priestor stal pre konzumenta prínosom je možnosť vyhľadávať, prehliadať/prezerať dokumenty v ňom umiestnené. Vzhľadom na rozmery aktuálneho informačného priestoru, ktorého väčšiu časť dnes tvorí web, treba nájsť nový špecifický spôsob jeho prehliadania. Do úvahy prichádzajú dve principiálne odlišné metódy – náhodný výber dokumentov z informačného priestoru (vzorkovanie) alebo povrchné prezeranie teoreticky všetkých dokumentov, alebo aspoň tej časti, ktorú vo vymedzenom čase obsiahneme. Obe naposledy uvedené metódy spadajú pod anglický termín „browsing“, pôvodne používaný pri popise spôsobu získavania/výberu potravy živočíchmi⁶. Termín „povrchný“ tu nemožno stotožňovať s nedbanlivosťou, ale naopak, s efektívnosťou a rýchlosťou. Ide o to, aby sme v krátkom čase vyhodnotili relevanciu maximálneho množstva dokumentov vo vzťahu k danej téme/problému. Pripomeňme si, že normalizovaná stránka formátu A4 obsahuje 30 riadkov textu, pričom v každom riadku je 60 znakov. Na jednej A4 strane je tak, pri 8bitovom kódovaní znakov, $1800 \times 8 = 14\,400$ bitov informácie. Oproti tomu, pri rozlíšení 1280x900 pixelov a čierno-bielom zobrazení bez odtieňov obsahuje jedna obrazovka 1,08Mbit informácie. Pokiaľ ide o množstvo informácie je rozdiel medzi jednou stránkou textu a jedným čiernobielym obrazom cca dva rády, t.j. obraz ponúka n.100 bitov informácie viac ako text a pri farebnom obraze je tento rozdiel ešte väčší. Aby takéto množstvo informácie mozog dokázal spracovať, potrebuje čas a koncentráciu. Pri „surfovaní“ na webe sa mu nedostáva ani jedného, ani druhého. Ak teda obraz na monitore len preletíme bez toho, aby sme sa dlhšie zdržali skúmaním istého detailu, resp. bez vracania sa k tomu, čo sme už videli, akoby sme ho

⁵ Jannis Kallinikos, Aleksii Aaltonen, and Attila Marton. A theory of digital objects. *First Monday*, Volume 15, Number 6 - 7 June 2010. <http://firstmonday.org/ojs/index.php/fm/article/view/3033/2564>

Manovich, L. 2001. *The language of new media*. Cambridge, Mass.: MIT Press.

⁶ najbližší slovenský výraz na tento druh činnosti je „paberkovanie“.

čítali technikou rychločítania. No i napriek tomu dochádza ku kognitívnemu preťaženiu⁷. Kognitívne zaťaženie konzumenta možno, bez zmenšenia rozmeru informačného priestoru, znížiť zvýšením transparentnosti informačného priestoru. Transparentnosť informačného priestoru zvyšuje:

- Jasné vymedzenie hraníc priestoru.
- Jasné vymedzenie obsahu priestoru.
- Rozčlenenie priestoru na subpriestory podľa predmetu.
- Jasné pravidlá preskúmania priestoru.
- Platnosť zákona aditivnosti v informačnom priestore.
- Možnosť objektívne posúdiť kvalitu, vierohodnosť a aktuálnosť informácií.
- Nezávislosť od technológií.
- Vizualizácia – analýza obsahu dokumentu pomocou softwarových nástrojov.

V ostatnom čase sa DIP čoraz viac optimalizuje pre „tuláka“⁸. Tulák po informačnom priestore sa podobá na Adamsovo sprievodcu po Galaxii⁹. Tulák je vnímavý na krásu, poeticky naladený, zároveň je kritický a kreatívny. Aby sme mu vyhovelí, musí byť informačný priestor prezentovaný cez vhodné rozhranie, ukazujúce ho v graduálnom kontinuu od detailu až do abstrakcie. Detail je reprezentovaný náhľadom dokumentu (text, klip, zvuková ukážka). Abstrakcia stavia na vizualizácii – objekty sú situované na časovej osi, ponúkajú sa mapy priestoru, grafy v rôznych podobách a pod.

4. Vzdialenosť v DIP

Najvhodnejšou mierou vzdialenosti v DIP je podobnosť. Tá ale môže mať celý rad rozmerov:

- Podobnosť metadátová,
- Podobnosť obsahová (sémantická),
- Podobnosť popisu (hlavne u netextovej informácie),
- Podobnosť abstraktu,
- Podobnosť formy vyjadrenia informácie,
- Časová

5. Personalizácia DIP

- Vlastným výberom, kedy si používateľ DIP podľa svojej potreby/vkusu vymedzuje svoj DIP a to buď fyzicky, ukladaním kópií vybraných informačných objektov napr. na vlastný disk, alebo symbolicky, ukladaním adres objektov
- Vnútená, kedy vyhľadávací stroj, „v dobrej viere“ a na základe analýzy vyhľadávacieho profilu¹⁰ používateľa, personalizuje výsledky vyhľadávania.

⁷ Kognitívne preťaženie je definované ako „úsilie a koncentrácia vynaložené navyše, aby sme dokázali vykonávať/sledovať viacero úloh/ciest súčasne.“ Viac pozri Conklin, J. (1987). Hypertext-an introduction and survey. IEEE Computer, Sept. 20, 17-41.

⁸ Idea tuláka nie je nová. Vid' napr. knihu J. Londona Tulák po hviezdach. Ak by J. London tvoril dnes, pravdepodobne by napísal „Tulák po webe“

⁹ Adams, D. The hitchhiker's guide to the Galaxy, Ballantine books 1995

¹⁰ Vyhľadávacím profilom používateľa rozumieme profil, vytvorený analýzou histórie jeho požiadaviek na vyhľadávanie za určité obdobie, napr. jeden mesiac.

6. Charakteristiky DIP

DIP je, i napriek jeho exponenciálnemu rastu, stále iba podpriestorom globálneho IP a objektívne tomu tak bude dovtedy, kým budú existovať nedigitálne súčasti globálneho informačného priestoru. Jeho výhodou i prekliatím zároveň je, že vznikol v Petriho miske informačno-komunikačných technológií a bez nich prestáva existovať. IKT dávajú digitálnemu informačnému priestoru nové vlastnosti v oblasti prenosu, uchovávaní, odovzdávania, prezentácie, sprístupňovania informácie, nový obsah získava aj koncept interaktivity a koncept siete. Asi najcharakteristickejším znakom DIP je heterogenita – heterogenita z pohľadu kvality, aktuálnosti, vierohodnosti informačných objektov, spôsobu spracovania, heterogenita tematická, štýlová, heterogenita formy, atď.

Charakteristiky DIP môžeme zosumarizovať nasledovne:

- Materializovateľný len cez jeho reprezentáciu (najčastejšie na obrazovke).
- Obsahuje informácie v podobe postupností bitov, navonok prístupných vo forme informačných objektov.
- Človek má tendenciu vnímať ho na jednej strane ako miesto, na druhej strane ako neohraničený priestor.
- Pohyb v DIP vyžaduje existenciu navigácie.
- Navigácia vyžaduje existenciu organizácie priestoru (z pohľadu IS aj z pohľadu človeka.)
- K organizovanému priestoru sa pristupuje cez rozhranie – rozhodujúca je preto prezentácia priestoru na rozhraní.
- Cez rozhranie je prezentovaný, ako špička ľadovca, vždy len výsek DIP.
- Realizáciou konceptu hypertextu¹¹ v DIP sa vytvorila technika prístupu k teoreticky nekonečnému množstvu informácií rozpriestranenému do teoreticky nekonečného počtu smerov/rozmerov.
- Nelineárnosť usporiadania informácie vnesená do DIP hypertextom nás núti opustiť geometrickú predstavu vnímania priestoru cez perspektívu a vytvoriť novú, založenú na postupnom (a nutne neukončenom) vnímaní DIP cez pozorovanie jeho výsekov, pričom výber výsekov je viac – menej náhodný.

IKT otvorili v DIP Pandorinu skrinku datovizácie priestoru, robiacu DIP čoraz viac prostredím strojov ako prostredím pre človeka. Zjednodušenie procesu vytvárania nových digitálnych objektov má tiež dve tváre: tá priaznivá ukazuje nárast dostupných informácií, tá odvrátená predstavuje ekologickú záťaž: DIP je znečistený chybnými informáciami a duplicitami viac, ako ktorýkoľvek iný informačný priestor, prejavuje sa efekt nadbytku informácií (aj v dôsledku masovej digitalizácie a redundancie informačných objektov).

Z hľadiska času sleduje DIP zákon pôsobenia entropie t.j. smeruje k stále väčšej neusporiadanosti. Môžeme mu priradiť niekoľko ne“ prívlastkov:

¹¹ Koncept ako prvý použil T. Nelson v roku 1963

- netransparentný,
- neštruktúrovaný,
- neohraničený,
- nelineárny,
- nehmotný.

Akoby DIP trpel detskými chorobami. Ich najtypickejšími syndrómami sú:

- mcDonaldizácia priestoru: stále viac je DIP zapĺňaný „sériovými“ informačnými objektmi obsahujúcimi povrchné, predžuvané informácie, ktoré musia byť navyše „cool“.
- syndrom poštového koňa: kedysi poštový kôň automaticky zastavil pred každou poštou a okolo ostatných stavaní prešiel bez povšimnutia. V DIP sa vytvárajú „pošty“, pri ktorých sa väčšina používateľov automaticky zastaví. Tým sa iné objekty dostávajú do virtuálneho tieňa „pošty“.
- Minimálneho minima: v snahe preskúmať čo najviac zaujímavých miest v DIP sa pri riešení pracovného problému používateľ snaží nájsť čo najrýchlejšie minimálne minimum informácií, postačujúce na jeho vyriešenie.
- Maximálneho maxima: pri oddychu/zábave sa používateľ snaží z DIP vyťažiť maximálne maximum pôžitku a to často bez ohľadu na čas.
- Nasledovateľa: nasleduj proroka
- Amerického turistu: v minulom storočí známy americký film „Ak je štvrtok, musíme byť v Paríži“ parodoval americký spôsob poznávania krajín podľa dopredu nastavenej šablóny. Podobný prístup sa uplatňuje v DIP.
- Slepého kurčaťa: známa ľudová múdrosť optimisticky tvrdí, že aj slepé kurča si nakoniec nájde svoje zrnko. Aplikované na DIP: aj informačne negramotný môže nakoniec nájsť čo potrebuje.

Záver

Vďaka digitalizácii prebieha agresívna remediácia orálneho a písomného informačného priestoru do DIP doteraz nevídanou rýchlosťou a v doteraz nevídaných objemoch. Rovnako nevídanou rýchlosťou a v doteraz nevídaných objemoch vzniká i nový, pôvodný digitálny obsah. Každú minútu sa¹²:

- objaví 48 hodín nových príspevkov na YouTube
- pošle 285 000 000 mailov
- zaznamená 2 000 000 dopytov na vyhľadávanie dokumentu
- vygeneruje 35 000 „lajkov“
- Stiahne 47 000 aplikácií
- Objaví 570 nových www sídel, 3 100 nových fotografií, cca 28 000 príspevkov do blogov
- ...

¹²Podľa štatistík dostupných online na adrese <http://www.visualnews.com/2012/06/19/how-much-data-created-every-minute/?view=infographic>

To sú čísla, ktoré dokumentujú nielen rýchlosť rozpínania DIP ale aj potrebu aplikácie princípov architektúry naň.

Tento článok vznikol s podporou projektu „Pamäť Slovenska – Národné centrum excelentnosti výskumu, ochrany a sprístupňovania kultúrneho a vedeckého dedičstva“ (ITMS:26220120061) v rámci OP Výskum a vývoj spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.

