

Elektronické publikování pro e-learning

Petr Sojka

Masarykova univerzita v Brně, Fakulta informatiky
sojka@fi.muni.cz

Abstrakt Nejčastějším způsobem komunikace při e-výuce a aktivitách typu blended learning zůstává stále *dokument* či *e-dokument*, ať již to je skriptum, prezentace, webová stránka nebo e-mail. Přesto kvalita dnešního elektronického publikování zaostává za kvalitou, kterou u klasických publikací dosahovali mistři černého řemesla před sto lety.

V článku analyzujeme možné příčiny tohoto stavu, i nové dosud málo využívané možnosti dnešního elektronického publikování, a ty demonstrujeme na přípravě a zpracování tohoto sborníku.

1 Co je šémem elektronického publikování pro e-learning?

Jedna z verzí příběhu o Golemovi praví, že rabi Yehudah Löw ben Bezalel přivedl k životu prach země tak, že na pergamen napsal magická slova a tento dokument (šém) vložil do úst z hlíny uhněteného Golema. Pergamen dnes již není hlavním médiem pro přenos (magických, oživujících) slov. Přestože klasická kniha zůstane i nadále ověřenou cestou podávání informací a znalostí a jejich šíření ve výuce, nikdo nepochybuje o tom, že paralelně bude nabývat důležitosti elektronické publikování, kdy dokumenty a knihy (e-book) budou šířeny elektronickou cestou a budou navrženy tak, aby i jejich četba mohla probíhat na digitálních zařízeních.

Současné snahy jsou napnuty ke standardizaci elektronicky přenášených výukových objektů (shareable content objects, reusable content objects), mezi nimiž kromě zvukových a video-komponent hrají dokumenty stále významnou roli. V tomto případě jde vlastně o *e-dokumenty*, tedy dokumenty přenášené elektronickou cestou, a často konzumované na digitálních monitorech, personálních digitálních asistentech či mobilech třetí generace.

Dokument je dán svým *obsahem* a *formou*. Obsah dokumentu je samozřejmě primární – samotný obsah však nestačí. Přístup “Content is king” z konce minulého století vystřídalou vystřízlivění na začátku jedenadvacátého století [Skyme, 2002]. Pouze obsah nestačí. Je třeba jít kus cesty dále vstříc autorovi: důležitá je nejen forma, design přizpůsobený konkrétním čtecím zařízením, ale také možnostem navigace v e-dokumentu, vyhledávání či časování přístupu k informacím. A to je pouze vrchol ledovce rozsáhlé problematiky přípravy e-dokumentů. Pověstný šém ale musíme hledat na této cestě ke čtenáři k jeho potřebám, zvykům, a možnostem čtecího zařízení.

2 Začínáme na zelené louce?

Historie publikování se měří na tisíce let. Za zlomový okamžik se považuje polovina minulého tisíciletí, kdy došlo k zásadní změně možností šíření dokumentů. Vynález

knihtisku umožnil mnohem snazší a levnější šíření dokumentů. Podobně významná revoluce je tu dnes. Internet je chápán jako nové přenosové médium, které stejně nebo ještě výrazněji snižuje bariéru šíření tištěných dokumentů danou fyzikálními vlastnosti dosud užívaného přenosového materiálu, tedy papíru.

Teorie automatizace zpracování dokumentů sahá do šedesátých let minulého století, kdy Charles Goldfarb, Ed Mosher a Ray Lorie navrhli první strukturní *značkovací jazyk*, GML (Generalized Markup Language) [Goldfarb, 2003]. Nutnost standardizace dokumentového metajazyka a nástrojů s ním pracujících vedla v roce 1974 k návrhu SGML, Standard Generalized Markup Language, normy ISO 8879:1986. Metajazyk, tedy formalismus pro definici formálních jazyků dokumentů SGML se používal zejména tam, kde šlo o zpracování obrovských objemů textových dat – patentové databáze, databáze právnických předpisů, dokumentace velkých projektů typu návrh letadel či atomových elektráren. V těchto aplikacích plně vynikla hlavní myšlenka značkovacích jazyků, a to *oddělení formy od obsahu* a z toho vyplývajícího usnadnění editace textu, jeho ukládání, indexování a vyhledávání a násobné publikování v různých podobách. Dokumenty uložené v tomto *otevřeném* formátu se dají editovat libovolným editorem, a velcí nakladatelé si SGML zvolili pro dlouhodobé ukládání textů již před třemi dekádami. Jak prozívá tato volba byla se ukazuje dnes, kdy pouhou definicí formy odpovídající logickým entitám gramatiky dokumentu mohou publikovat ze stejných dat dokumenty přizpůsobené dnešním zobrazovacím zařízením.

Specialisty v elektronickém publikování nepřekvapil výběr metajazyka SGML Tim Berners-Lee pro formální popis prvního jazyka internetu – HTML, a následný boom technologie značkovacích jazyků (WML, DocBook, TEI, XHTML, MathML) a standardů a technologií od nich se odvíjejících (CSS, RDF, Topic Maps). “The web is broken and I broke it.” přiznává guru webového designu David Siegel, který stavěl na specifických nekompatibilních rozšířeních internetových prohlížečů a ty využíval k sice graficky efektním, ale důsledně odlišením obsahu a formy neumožňujícím, prezentacím. Kéž by podobně prožili uživatelé textových procesorů typu Microsoft Word, kterým dělá problém v textu vyznačit jednoznačným a konzistentním způsobem například kapitolu do knihy nebo článek do sborníku pro jednotné přesazení specialistou.

Dnes složité SGML nahradilo jednodušší XML. Indexování slovních tvarů, bez lematizace a sémantického indexování má své limity. Nastala vize *sémantického webu*, kdy indexovanými objekty dokumentů již nebudou jednotlivé slovní tvary, ale významy. Jak pěkná představa, ptát se na významy, jenže jak? Jak přinutit autory, aby své dokumenty značkovali sémanticky a vytvářeli potřebná metadata, když neumí ani říct co je nadpis příslušné úrovně? Slibný je tedy spíše postup sémantická metadata e-dokumentů *generovat* z plného textu s využitím všech informací v textu dostupných, včetně použití technik zpracování přirozeného jazyka [Gregar a Pitner, 2005]. Podobně to dělá Google, který přihlíží k tomu, zda je e-dokument formálně správný, *validní*, a z jeho struktury odhaduje význam dokumentu z významů slov v titulku, nadpisech, začátku dokumentu apod. Dále metodou vycházející ze standardní citační analýzy počítá kredit e-dokumentu tzv. *PageRank*, který slouží k určení relevance e-dokumentu k dotazu.

3 Pár příkladů, námětů a zkušeností

Myšlenky důsledného oddělení obsahu od formy, kvalitního logického značkování a aplikace metod zpracování přirozeného jazyka byly použity v řadě velkých publikačních projektů [Všech pět pohromadě, 1999, DVD 10@FI, 2004, Sojka, 1998, Sojka, 2004]. Ohlasy studentů na sjednocení studijních materiálů stovky předmětů do přenositelného formátu [DVD 10@FI, 2004] jsou veskrze pozitivní. Po opakované používání studijních materiálů a generování různých variant e-dokumentů pro různá výstupní zařízení [Došlá et al., 1999, Došlá et al., 2002] je nezbytné užití dobře značkování dokumentu v DocBooku [Bařínka a Vojtěch, 2005] nebo \LaTeX u. Využití XML jako rozhraní pro výměnu výukových komponent je jistě perspektivní, byť zatím není v dohledu shoda na nějaké všeobecně uznávané ontologii pro práci s významem textů, byť vyjadřovací možnosti jazyků a technologií XML pro zpracování textů tu jsou [Pavlovič et al., 2005].

U e-dokumentů, zejména výukových, je často žádaným prvkem interaktivita a animace. To nás vedlo k vývoji portabilních řešení interaktivních animací přímo ve formátu PDF [Holeček a Sojka, 2004]. Jejich užití v [Došlá et al., 2002] ukazuje na možnosti skriptování ve formátu PDF při využití volně šířeného Acrobat Readeru, a setkal se taktéž pozitivním přijetím u studentů. Umožňuje uvést do praxe heslo Bena Shneidermana “Collect relate create donate” [Shneiderman, 2000].

Zanedbávanou částí přípravy výukových e-dokumentů je forma. S přechodem zpracování textů od specialistů na stůl autora se zvýšily jednak možnosti přímého ovlivnění formy autorem, ale také požadavky na znalosti autora v oblastech digitální typografie a zpracování textů. Ty jsou stále velmi mizivé, jak ukazovaly i práce na tomto sborníku. Jak byl sborník zpracováván je popsáno v kolofónu na straně 243 této knihy. Pokud jej držíte v rukou, podařilo se ujednotit formu všech 42 příspěvků a tak zase udělat jeden z krůčků k čtenáři; příložený CD-ROM zase vychází vstříc těm, kteří chtějí učební text rychle prohledávat či číst na obrazovce. Další zanedbávanou oblastí pozornosti při přípravě rozsáhlých sad e-dokumentů je navigace čtenáře při jejich procházení [Neveřilová a Sojka, 2005].

4 Závěrem

I při pokroku dnešních informačních technologií nelze spoléhat na jejich samospasitelnost, ani dosud není všeobecně akceptován žádný formát pro rozhraní mezi výukovými sdílitelnými objekty. Pro přípravu kvalitních e-výukových objektů je třeba přizvat specialisty na design pro dnešní výstupní zařízení, ale také zvyšovat kulturnost i vzdělanost pedagogů zapojených do e-learningu.

Doufejme, že kvalitně zpracované e-dokumenty pro e-výuku budou mít takovou životadárnou sílu jako Golemův šém a že se nám příprava e-dokumentů pro e-learning nevyvíjí z rukou tak, jak se to stalo v příběhu o Golemovi.

Poděkování

Tato esej vznikla v rámci řešení projektu „E-learning v kontextu sémantického webu“ národního programu Informační společnost, grant číslo 1ET208050401.

Reference

- Bařinka a Vojtěch, 2005. Bařinka, L. a Vojtěch, L. (2005). DocBook jako cesta k e-learningu. V [Sojka a Pitner, 2005], strany 33–38.
- Došlá et al., 1999. Došlá, Z., Plch, R., a Sojka, P. (1999). Matematická analýza s programem Maple: 1. Diferenciální počet funkcí více proměnných. CD-ROM, <http://www.math.muni.cz/~plch/mapm/>.
- Došlá et al., 2002. Došlá, Z., Plch, R., a Sojka, P. (2002). Matematická analýza s programem Maple: 2. Nekonečné řady. CD-ROM, <http://www.math.muni.cz/~plch/nkpm/>.
- DVD 10@FI, 2004. DVD 10@FI (2004). DVD 10@FI. DVD-ROM, <http://10.fi.muni.cz/>.
- Goldfarb, 2003. Goldfarb, C. F. (2003). The SGML History Niche. <http://www.sgmlsource.com/history/index.htm>.
- Gregar a Pitner, 2005. Gregar, T. a Pitner, T. (2005). Využití nástrojů pro zpracování přirozeného jazyka v e-learningu. V [Sojka a Pitner, 2005], strany 45–50.
- Holeček a Sojka, 2004. Holeček, J. and Sojka, P. (2004). Animations in a pdf \TeX -generated PDF. *TUGboat*, 25:35–41.
- Pavlovič et al., 2005. Pavlovič, J., Pitner, T., Kubásek, M., and Svoboda, L. (2005). Searching and E-learning System above Digital Sources. In [Sojka a Pitner, 2005], pages 39–44.
- Neveřilová a Sojka, 2005. Neveřilová, Z. a Sojka, P. (2005). XML-Based Flexible Visualisation of Networks: Visual Browser 5 pages, submitted for publication.
- Shneiderman, 2000. Shneiderman, B. (2000). Creating creativity: User interfaces for supporting innovation. *ACM Transactions on Computer-Human Interaction*, 7(1):114–138.
- Skyme, 2002. Skyme, D. J. (2002). Is Content King? http://www.skyrme.com/updates/u59_f2.htm.
- Sojka, 1998. Sojka, P. (1998). Publishing Encyclopaedia with Acrobat using \TeX . In *Towards the Information-Rich Society. Proceedings of the ICC/IFIP conference Electronic publishing '98*, pages 217–222, Budapest, Hungary. ICC Press.
- Sojka, 2004. Sojka, P. (2004). Technologie multimediálního publikování na DVD: zkušenosti z projektu DVD 10@fi. V Sedláček, J., editor, *Sborník příspěvků ze semináře a soutěže eLearning 2004*, strany 198–206, Hradec Králové.
- Sojka a Pitner, 2005. Sojka, P., Pitner T., editoři (2005). *SCO 2005 – Sharable Content Objects*, Brno, Czech Republic. Masarykova univerzita v Brně, ČR.
- Všech pět pohromadě, 1999. Všech pět pohromadě (1999). Všech pět pohromadě. CD-ROM, <http://nlp.fi.muni.cz/projekty/vsech5/>.