

# PROBLEMATIKA ELEKTRONICKÝCH PUBLIKACÍ

---

Adolf Knoll  
Národní knihovna ČR

## Elektronická publikace

Termín *elektronická publikace* obvykle označuje publikaci, k níž je umožňován přístup pomocí výpočetní techniky. Tento termín by však mohl být interpretován i jinak a případně zahrnout i analogové publikace, které lze zpřístupnit pouze pomocí elektronických přístrojů (například analogové video). Na druhé straně však je z oblasti chápán pod pojmem elektronické publikace velmi často vylučováno digitální audio CD. Někdy též bývá používán termín *elektronické zdroje*, který však tento problém neřeší, pouze snad reaguje na skutečnost, že může jít i o něco dalšího, jako jsou databáze, software nebo počítačové hry. Jsme toho názoru, že lepším termínem je *digitální publikace*, neboť spočívá zcela zřetelně na přesně definovaném způsobu kódování informací, nebo šířeji termín *digitální dokument*. V poslední době se také velmi často vyskytuje termín *digitální objekt* jako reakce na různorodost základních stavebních jednotek digitálního světa.

Dále je třeba si uvědomit, že digitální publikace je digitální dokument zveřejněný pomocí formálního publikačního kanálu. K našemu účelu je vhodné používat termín digitální publikace, resp. digitální dokument jako vnitřně uspořádaný soubor elementárních digitálních objektů. Elementárním digitálním objektem rozumějme pro naši potřebu jednotlivý jednoduchý datový (nebo metadatový) soubor jako základní jednotku výstavby digitálního dokumentu - v krajním případě se může digitální dokument rovnat právě jednomu souboru.

Pokud byl digitálním objektem textový soubor, nelišil se způsob chápání tohoto objektu příliš od způsobu chápání běžného klasického dokumentu s tím rozdílem, že byla poměrně brzy objevena základní a filozoficky nejdůležitější vlastnost takového objektu, již je variabilita jeho konkrétní jevové realizace ze stále totožného základního zdroje. V podstatě se tím digitální objekt neliší například od klasického objektu, neboť i on může procházet těmito proměnami (různě velké zvětšeniny ze stejného negativu, různě upravený identický text v různých vydáních atp.). Změnou je pouze rychlost a relativní snadnost těchto formálních změn.

Poměrně brzy se objevuje schopnost digitálního kódování obrazu a zvuku, posléze i pohyblivého obrazu. S tím vzniká problém, který si člověk dosud uvědomoval pouze nedostatečně, totiž možná odrzedenost těchto speciálnějších digitálních objektů od jejich dalšího kontextu, nejčastěji opět textových informací.

Například samotný soubor s digitální fotografií sice ukazuje něco, co do jisté míry chápeme, ale pokud zapomeneme kontext, nejsme z tohoto souboru již nadále schopni identifikovat místo, osoby nebo čas pořízení atp.

## Autonomie a závislost digitálního objektu

Skutečnost, že řada odborníků mluví o digitálních objektech, je signálem, že právě nezačleněnost těchto objektů do vyšších struktur, které o nich jsou schopny vypovídat, je velkým problémem.

To v praxi přináší dvě možnosti řešení a dva okruhy otázek:

1. Popisná data (metadata) existují a jsou uvedena mimo tyto objekty většinou v záznamech zpřístupňovaných relačními databázemi (ze záznamu se dostaneme k objektu, který sám o sobě není schopen vypovídat).
2. Popisná data jsou součástí komplexnějších dokumentů, které zahrnují tyto jednoduché objekty a vypovídají o nich jiným způsobem, než toho je schopen sám objekt způsobem jemu vlastním (čili ne žádným „metazpůsobem“: například některé obrazové formáty jako GIF umožňují zapsat i popisná data do hlavičky souboru).

Uvedená řešení spojuje otázka, jakým způsobem zabezpečit co možno nejotevřenější jednoznačně přiřazenou metainformaci o digitálním objektu.

První možnost má dva kritické momenty: zachování jednoznačné relace mezi metainformací a digitálním objektem (jsou uloženy odděleně) a dále zachování funkčnosti relační databáze.

Druhá možnost zvyšuje nezávislost digitálního objektu, tím však klade otázku otevřeného formátu onoho komplexnějšího dokumentu - nevylučuje však použití relační databáze pro zpřístupnění digitálních objektů.

Z této analýzy je zřejmé, že bezpečnější je druhé řešení, zatímco první řešení lze doporučit spíše jako nadstavbu pro lepší zpřístupňování komplexních datových objektů.

## Data a metadata

Existují různá pojetí především termínu *metadata*. Vždy jde zejména o účel tohoto pojetí ve vztahu k obecnému tvrzení, že metadata jsou veškerá data o datech, přičemž je tento vztah relativní a vždy závisí na postoji a interpretaci tvůrce digitálního dokumentu. Dle účelu lze vymezit v současné době tři základní oblasti, kde se užití termínu metadata jeví jako účelné:

1. identifikace digitálního dokumentu - tímto lze rozumět buď jen jakýsi identifikátor (unique identifier, handler) typu DOI (Digital Object Identifier) nebo URN (Uniform Resource Name), tj. jednoznačnou a případně nezávisle ověřitelnou sekvenci nikdy se neměnicích znaků, nebo různé stupně popisu dokumentu jako celku, například popis typu katalogizačního (bibliografického) záznamu at' má charakter MARC, charakter různě kódovaného Dublin Core (podrobněji v následujícím příspěvku E. Bratkové) aj.
2. popis různých objektů v dokumentu a kontrola jeho integrity (důležitá zejména v případě velké komplexnosti digitálního dokumentu) - tento přístup obecně zahrnuje i přístup z bodu 1, jenž je jeho podmnožinou
3. popis vlastností především externího prostředí, v němž lze dokument zpřístupnit a zajistit jeho původní chování (metadata pro budoucí emulaci softwarových a hardwarových platform).

Každá z těchto oblastí se musí chovat standardizovaně, tzn. současně je třeba zajistit:

1. stejný formální zápis metadat
2. jednoznačné vymezení popisovaných objektů.

Dále platí, že systém zápisu metadat a systémy vymezení popisovaných objektů mají být na sobě nezávislé, přičemž první systém by měl být značně obecný, ale zároveň i velmi rigorózní při formalizaci vymezených objektů. Popisované objekty při splnění výše uvedených podmínek jsou pak zásadně vymežovány externími standardy, definicemi nebo praktikami.

### **Objektově orientovaný přístup**

Jak již je zřetelné z výše uvedeného, lze hovořit o dokumentu jako celku nebo o jakýchkoli jeho částech jako o objektech. Hned na začátku je třeba říci, že objekty jsou vymezeny pomocí externích standardů, definic či úzu tvůrcem dokumentu. Objekty, z nichž se dokument skládá, mohou mít různý stupeň sémantické zřetelnosti. Některé jsou obsahově velmi slabé a jsou vymezeny spíše formálním, a tím i obsahově nepřesným, způsobem. Například odstavec v textu obsahově snad pouze znamená další množinu myšlenek, ale dále jej nelze jinak identifikovat než jako zřetelně formálně oddělenou část textu. Naproti tomu vymezení určité části informací v dokumentu jako jméno jeho fyzického původce zahrnuje v sobě již značnou sémantickou informaci, neboť z hlediska obsahu daný objekt zařazuje do určité kategorie. Dále platí, že též objekt může být zařazen jak do kategorie obsahové, tak i současně do určité kategorie formální.

Jestliže v dokumentu přiřadíme sekvenci znaků *Cervantes* příznak určité obsahové kategorie (autor) a současně i příznak určité formální kategorie (například odstavec), pak tím říkáme, že tato sekvence znaků je autor dokumentu a že má být zobrazen v samostatném odstavci (na samostatném řádku).

V praxi je výhodné přiřazovat vždy celé obsahově bohaté objektové kategorii onu formální kategorii jako celku, přičemž je žádoucí, aby byly obsahově bohaté kategorie na formálních kategoriích zcela nezávislé. To pak umožňuje různé formátování množin objektů na výstupu na základě různých přiřazení. Ve zjednodušené formě to známe například z formátování bibliografického záznamu do různých výstupů, přičemž tento princip platí obecně i pro nakládání s primárními dokumenty. Můžeme také pro větší přehlednost využít názvů objektů jako návěstí formálních výstupů.

Každá kategorie objektů je definována určitou svou specifikací (má svou podstatu). Touto specifikací může být určitý normativní předpis, který sám o sobě může být založen na různých obsahově orientovaných standardech. Například ve své specifikaci mohou mít, že objekt *autor* je sám o sobě definován vnějším standardem typu určitých katalogizačních pravidel nebo že objekt *iluminace* je definován územ obecně přijímaným při popisu grafických vlastností rukopisů.

Jako taková musí mít tato kategorie objektů také své naprosto jednoznačné specifické označení, ať je jím neza-

měnitelné označení pole a podpole v MARC nebo jiný způsob, jako je třeba *markup* v SGML (podrobněji dále v textu) a odvozených aplikacích. Dále by měla mít taková kategorie jméno, které se může lišit pro jednotlivé jazykové oblasti. K tomu pro nás může být důležité specifikovat nějaké další vlastnosti (atributy) označeného objektu v rámci dané kategorie. Můžeme například chtít označit, že daná hodnota objektu (tzn. jeho konkrétní realizace v dokumentu) se liší od případných jiných hodnot tím, že je vyjádřena číslem, a ne textem (to může být dobře zhodnoceno při zpracovávání dat).

### **Uspořádání objektů**

Jestliže tvrdíme, že digitální dokument se skládá z objektů, pak také tvrdíme, že těmito objekty je onen dokument vytvářen, to znamená, že mezi těmito objekty existují určité vztahy: jedny mohou být nadřazeny druhým, některé mohou být zahrnuty (*nested*) dovnitř definovaných objektů, některé mohou ukazovat na další externí objekty ať datového nebo metadatového charakteru, a ty samotné mohou též být složitě strukturovány.

Toto znamená, že potřebujeme nástroj, který by tyto vztahy řídil a zajišťoval tak integritu celého digitálního dokumentu. Výkonná součást tohoto nástroje by měla být součástí digitálního dokumentu, v současné době již není udržitelné jiné pojetí. Rovněž je zcela zřetelné doporučovaná platforma SGML.

Pomocí nástrojů SGML lze vyjádřit jakýkoli objektově orientovaný přístup. Vzhledem k tomu, že je SGML obecné, neboť říká, jak označovat objekty, vyjadřovat jejich vlastnosti a vztahy mezi nimi, je vždy nezbytná jeho konkretizace. Tato konkretizace je vyjádřena vlastními nástroji SGML pomocí tzv. DTD (definice typu dokumentu), která stanoví základní stavební jednotky toho kterého typu dokumentu dle našeho zadání.

Většinou se užívá vícestupňového definování příslušného SGML prostředí. Mezi konkrétně zakódovaným dokumentem a SGML bývá často určitá přesně definovaná konkretizace SGML pro dané účely. Z tohoto pohledu se pracuje s konkretizacemi, jako jsou HTML, TEI (Text Encoding Initiative) nebo naše DOBM.

Některé tyto konkretizace jsou již využívány jako takové (například HTML), jiné mají v sobě další stupně jemnějších konkretizací pro různé účely. Vždy však používají své vlastní nástroje. Jestliže například HTML neumožňuje *markup* (značení) objektů z hlediska jejich obsahu, TEI i DOBM to umožňují, přičemž DOBM je značně obecnější.

V praxi to znamená, že tzv. objektová mapa dokumentu je pro konkrétní dokument v TEI externí ve vztahu k tomuto dokumentu (je jí určité DTD), zatímco v DOBM je součástí tohoto konkrétního dokumentu (je jí zvláštní SGML soubor, který též stanovuje, který soubor je třeba pokládat za kořenový - root file).

Tyto vlastnosti aplikací SGML jsou důležité zejména pro popis, výstavbu a kontrolu integrity tzv. složených (compound) dokumentů, jež jsou zároveň i značně složitě.

Tyto dokumenty vznikají například při digitalizaci klasických dokumentů, kdy jsou generovány na základě jed-

noho klasického dokumentu často stovky i tisíce souborů, které se skládají každý uvnitř z různých objektů a které spolu musí vytvářet jeden hierarchicky strukturovaný (ale zároveň též rozložitelný do lineární sekvence) celek. Navíc je charakter souborů uvnitř složeného dokumentu obecně různorodý, neboť nejenže v něm lze najít množiny metadatových a datových souborů, ale též různé kvalitové hladiny datových souborů zobrazujících tentýž objekt. Dále mohou být datové soubory různého charakteru, a to i v jednom dokumentu: textové, obrazové, audio, video. A samozřejmě v různých datových formátech pro různý uživatelský účel.

Vezmeme-li toto vše na vědomí, nelze se bez SGML zodpovědně obejít.

### **Typy digitálních dokumentů**

#### *Podle charakteru datových souborů*

Podle charakteru datových souborů lze rozdělit digitální dokumenty na textové, obrazové, zvukové, video a smíšené. Za zvláštní skupinu lze považovat software. Toto rozdělení vychází z charakteru základní stavební jednotky digitálního dokumentu, ne z charakteru zpřístupňování těchto jednotek. Nevylučuje to zpřístupňování množin těchto dokumentů prostřednictvím databázových nástrojů. Například i bibliografická databáze je z tohoto pohledu zpřístupňovací mechanismus vnitřně strukturovaných textových dokumentů. Samozřejmě může být i ona sama na své úrovni považována za velmi komplexní dokument nebo zdroj.

Výše uvedené skupiny dokumentů se v praxi vyznačují užitím různých struktur a datových formátů. Problémem je vždy změna datového prostředí a s tím často přicházející změna filozofie výstavby dokumentů.

Většina z nás si pamatuje složitý přechod a konverze z formátu T602, který se vyznačoval absencí objektově orientované výstavby (zde absence tzv. stylů), do objektově orientovaných formátů typu MS WORD. A to přitom šlo převážně jen o ryze formální objekty, nikoli popis a vnitřní značení orientované obsahově - to již zvládá jen SGML.

V prostředí sdělování textových informací jednoznačně vede HTML, MS WORD, případně RTF a PDF (tam, kde chci předat publikaci, aby byla věrně vytištěna). Někteří dobří vydavatelé také používají čisté SGML na základě svých vlastních DTD.

V případě obrazových formátů je patrný vliv Internetu především rozšířením formátů GIF a JPEG a užíváním (i když zatím ne častým) velmi dobrého a doporučeného PNG. Obecně se velmi používají různé modifikace TIFF, ať nekomprimované nebo komprimované různými technologiemi. Další formáty se používají ve speciálnějších případech nebo jako vnitřní formáty aplikací (například BMP, WMF nebo PSP). Dochází k vývoji formátů nových zejména za účelem dosažení větší komprese novými algoritmy při zachování kvality zobrazení, například DJVU.

V oblasti audia kraluje CD audio formát (CDA) a WAVE, přičemž vliv Internetu je cítit opět ve snaze dosáhnout podstatné komprese dat při zachování kvalitního poslechu (například MP3 nebo Real Audio). Obdobně ten-

dence jsou patrné i v přenosu digitálních videodat. Smíšené dokumenty jsou pak obvykle kombinací nezřídka všech výše uvedených datových formátů.

Tento vývoj bude pokračovat a není jiná možnost, než jej monitorovat a převádět data ze stárnoucích formátů do formátů nových, přičemž může v některých případech docházet ke ztrátám informací, resp. kvality.

Mnohem složitější je však situace ve strukturování dokumentů sestávajících z více než jednoho datového souboru.

#### *Podle provenience*

Typy digitálních dokumentů lze též rozdělit podle zdroje, z něhož do digitální knihovny přicházejí. Rozlišují se tyto tři typy:

1. dokumenty vyrobené digitalizací předloh uvnitř instituce nebo velmi obecně vlastní digitální dokumenty,
2. dokumenty publikované na fyzických nosičích a získané digitální knihovnou,
3. dokumenty, které jsou k dispozici ve vnějších sítích a mohou být knihovnou zpřístupněny.

Každý z těchto typů klade různé nároky na zodpovědnost instituce za jejich především dlouhodobé zpřístupnění v případě, že jde o dokumenty domácí provenience. U národních knihoven pak jde o zvážení vlastních sil a možností pro jejich dlouhodobé uchování, přičemž se obecně soudí, že konzervace těchto dokumentů je nemožná, zatímco ochrana jejich obsahu v podstatě ano, pokud se vyřeší řada problémů.

#### *Ochrana a dlouhodobé zpřístupnění*

Pro ochranu informací obsažených v digitálních dokumentech jsou možné tři přístupy, případně jejich kombinace:

1. skladování dokumentů na originálních nosičích spolu s budováním funkčního softwarového a hardwarového muzea; v tomto případě jediným možným opatřením je organizované monitorování kvality nosičů;
2. skladování dokumentů po jejich převodu do vhodného archivního formátu, který se knihovna zavázala realizovat spolu s dílčími převody souvisejících datových souborů - toto je vzhledem k charakteru produkce, vyznačující se celosvětově absencí jednotného přístupu a neaplikací otevřených standardů, finančně i pracovníčně fakticky nemožné;
3. skladování dokumentů v originálních formátech a strukturách spolu s dalšími informacemi o softwarovém a hardwarovém prostředí, v němž lze dokument zpřístupnit se všemi jeho funkcemi. Toto řešení předpokládá v budoucnu vývoj emulátorů příslušných typových prostředí, které umožní dokument zpřístupnit. Zde je nezbytná standardizace zápisu příslušných informací - v současné době jsou náznaky řešení - ale otevřená zůstává vůle budoucích generací se tímto problémem zabývat (a také jej finančně podporovat). Rovněž je nejasný fakt, zda pro budoucnost budou námi představitelné informace pro emulaci postačovat (přece jen všichni vycházíme ze současného chápání věci).

## Směrování k digitální knihovně

Otázka způsobu budování digitální knihovny a priorit daných jednotlivým zdrojům vychází z možností jednotlivých institucí. Lze celkem zřetelně pozorovat, že instituce, které nemají nebo mají jen nevelkou produkci vlastních digitálních dat zejména v oblasti převodu složitých originálních dokumentů, se věnují většinou úvahám o zpřístupnění, případně uchování, digitálních dokumentů, které jsou produkovány ve vnějším prostředí, ať je jím Internet nebo nakladatelství vydávající digitální publikace na fyzických nosičích.

Při zajišťování dlouhodobé dostupnosti digitálních dokumentů jsou pozorovatelné tyto tendence:

- snaha o řešení zápisu dat pro emulaci (NEDLIB - Networked European Deposit Library, některé aktivity v USA) - zde je nebezpečí faktického odložení různých dílčích, ale pozitivních řešení ve jménu globálního řešení, nicméně tento výzkum je nezbytný;
- snaha sbírat vše nebo místy selektivně s tím, že je aplikováno řešení založené na monitoringu fyzických nosičů (což je velmi neperspektivní, ale ještě na hranici finanční únosnosti, například Bibliothèque de France a některé další národní knihovny);
- snaha omezit se na vybrané produkty, resp. datové nebo strukturální formáty (Norsko - PDF, Nizozemí - produkce Elsevier a Kluwer) - k tomuto řešení se přiklání stále více institucí;
- snaha o lepší zvladatelnost a dostupnost internetových zdrojů - poměrně rozsáhlé zastoupena v knihovnických (většinou pouze knihovnických) kruzích - vyznačuje se snahou vnutit Internetu knihovnické nástroje a myšlení; v praxi jde:
  - o katalogizaci vybraných digitálních dokumentů, což lze vítat, pokud tyto dokumenty mají velkou stabilitu a filozofii analogickou klasickému publikování;
  - o snahu přimět autory internetových dokumentů ukládat do jejich hlavičky jakousi katalogizaci v publikaci ve formě tzv. Dublin Core (problémem však je nejen to, že tzv. search engines to neberou v potaz, ale i fakt, že toto doporučení není respektováno a že vzhledem ke značnému knihovnickému zájmu se struktura Dublin Core a požadavky na něj stávají stále složitějšími) - toto řešení není teoreticky špatné, ale nezohledňuje filozofii bouřlivého rozvoje technologií, které si říkají o novou kvalitu myšlení a přístupu k tomuto problému;
- snaha o archivaci WWW - je třeba ji považovat za extrémní (nebo spíše extremistickou) a neproduktivní (některé aktivity v USA, Švédsko), vychází z nepochopení charakteru digitálních publikací, resp. objektů;
- nově se objevující tzv. hlasy rozumu neboli „informed skepticism“ - přicházejí z míst s velkou zkušeností s digitálními dokumenty (USA, Kanada – počítáme se k nim i my) a postupně jim začíná být nasloucháno i jinde. Tento přístup je zřetelný od roku 1998. Informed skepticism neztrácuje digitální technologie, jen

priznává obklopujícímu nás světu jeho různorodost včetně dílčích nedokonalostí, a tím i předjímá velkou diversifikaci možných řešení od nejjednodušších (vytištění digitálního dokumentu na papír tam, kde je to vhodné) po ta nejsložitější (migrace dat a struktur, emulace platform – zase jen tam, kde je to vhodné). Informed skepticism též rehabilituje mikrofilm jako médium pro ochranu řady klasických dokumentů - nicméně podporuje digitální přístup k datům včetně těch na mikromédiích.

Obecně nelze většinu přístupů pouze zatracovat nebo pouze vyzdvihovat. Je však třeba si v nich najít rozumné metody a důsledně je dodržovat, byť třeba jen na části celé národní produkce digitálních dokumentů.

V této oblasti víc než v jiné platí obecné zásady, které platí i ve světě klasických papírových publikací, kde si je však neuvědomujeme vzhledem k dlouhé životnosti těchto publikací ať v oblasti nosiče nebo způsobu kódování informací (ten je ještě trvanlivější).

Hlavní zásada filozofického charakteru je, že **lidstvo potřebuje především kontinuitu idejí**. Proto jsou informace sdělovány, šířeny, přijímány a vyhodnocovány. Děje se tak v různých dobách různými technologiemi.

Dávné národy memorovaly významné texty a středověcí písaři přenášeli hlavní ideje v rukopisech. V obou případech nešlo o doslovné předání textu, ale o předání idejí, proto se různé verze folklórní tvorby od sebe liší, proto je tentýž původní text v různých rukopisech ztvárněn různým způsobem a představuje různé verze.

Knihtisk přinesl unifikaci a produkci identických kopií jedné předlohy. Tím se rozšířila cirkulace identických textů a došlo svým způsobem k duplicitnímu šíření idejí (identický text v různých vydáních). Rukopis i tištěná kniha obecně představují hodnotu jako artefakt. Sbíráni artefaktů přineslo vážné problémy sběratelům, kteří se stali zodpovědnými za jejich dlouhodobé uchování a dlouhodobou dostupnost informací v nich obsažených. Došlo k významnému posunu od základního zabezpečení kontinuity idejí k vytváření rozsáhlých kulturních muzeí. Avšak digitální technologie dohnaly lidstvo k dalšímu myšlenkovému přelomu, který vychází z neschopnosti zvládnout obrovské množství proměnlivých informací pomocí filozofie kulturních muzeí.

Znovu se objevuje otázka základního smyslu celého tohoto počínání a odpověď je celkem nasnadě: je třeba oddělit vytváření kulturních muzeí od základního zabezpečení kontinuity idejí. Ukazuje se, že přístup středověkého písaře může být nosný i dnes. Tak jako on, i my máme dnes (relativně) velmi chatrné a omezené možnosti. Měli bychom tedy (tak jako on) zvolit rozumný a pragmatický přístup, který zabezpečí základní kontinuitu naší přidané hodnoty ke světové kultuře v širokém slova smyslu.

### Možná opatření v Národní knihovně ČR

Výše uvedené úvahy jsou jen krátkou charakteristikou myšlenkových proudů v dané oblasti, nicméně postačují k formulování několika možných opatření v Národní

knihovně ČR (dále NK). NK musí zaujmout stanovisko ke každému zdroji digitálních publikací, resp. dokumentů. Specifikem NK je však ten fakt, že je velkým producentem digitálních dokumentů hlavně prostřednictvím digitalizace a v blízké budoucnosti bude ještě větším. Její zkušenosti tedy nezahrnují jen oblast užití, ale i oblast výroby a uložení v mnoha dílčích otázkách. Z tohoto pohledu je NK v dané oblasti mnohem pokročilejší než většina jiných knihoven ve světě.

### ***Vlastní digitalizace***

Nastoupená cesta je perspektivní, NK zvládla, resp. zvládne celý proces dlouhodobého zpřístupnění digitálních dokumentů. Opírá se o progresivní řešení na bázi SGML.

Je třeba dále prohlubovat nejen vlastní produkci, ale též organizaci a technické zabezpečení digitálního archivu, online zpřístupňování dokumentů včetně monitorování jejich integrity, přepisu a převodu. Jako pozitivní je třeba též vyzdvihnout tu skutečnost, že je tato oblast podložena výzkumným záměrem.

Vzhledem k roli, jakou NK v této oblasti hraje na mezinárodní scéně, jí vzniká povinnost udržovat a rozvíjet standard DOBM a jeho konkrétní aplikace, případně zajistit konverzi tohoto formátu do jiného.

Kromě toho je třeba sledovat vývoj datových formátů zejména v oblasti obrazu a zvuku s vědomím možné využitelnosti pro potřeby NK (včetně případných užití formátů digitálního videa, pokud tato povinnost vyplyne z našich potřeb nebo úkolů).

### ***Digitální dokumenty publikované mimo NK na fyzických nosičích***

Tato oblast je - nejen v ČR - značně nesourodá a vykazuje velkou absenci řešení na bázi otevřených standardů. Z toho důvodu je naprosto nereálné budovat archivní sbírku těchto digitálních dokumentů (výjimku tvoří formátově homogenní produkce zvukových nahrávek - NK má právo povinného výtisku a vzhledem ke standardizaci této oblasti zahrnuje tyto publikace do digitálního archivu). Nicméně je třeba se snažit celou oblast ovlivnit:

- je třeba vyhodnotit urgence povinných výtisků těchto publikací;
- je třeba se pokusit navázat spolupráci se seriózními vydavateli skutečně hodnotných publikací a snažit se o jejich získání pro oblast dlouhodobého zpřístupnění digitálních publikací - se všemi důsledky pro obě strany. Většina nyní vydávaných publikací nepřináší nové ideje, proto není třeba usilovat o jejich trvalé uchování. Nelze totiž převádět do archivního metaformátu každou z nich. Na takové konverze nemá nikdo na světě - natož NK - finanční prostředky;
- je třeba sledovat dění kolem emulačního řešení (NEDLIB, atp.);
- je třeba tyto dokumenty používat důsledně pro služby NK;

- je třeba se ve výzkumném záměru NK věnovat novým druhům médií a jejich vlastnostem - po CD v nejbližší době DVD.

### ***Zpřístupňování, resp. archivace dokumentů ze sítí***

Jako rozumné se jeví vypracovat, resp. zpřesnit kritéria selekce dokumentů ze sítí z hlediska jejich nesporné hodnoty. Pro takto vybrané dokumenty se doporučuje:

- začít uskutečňovat jejich bibliografickou kontrolu;
- zahájit dialog s jejich tvůrci o jejich dlouhodobém uchování (archivaci).

Pro ostatní dokumenty (nejen národní) je třeba neustále optimalizovat WWW služby NK jako vhodnou „gateway“ k těmto zdrojům. Dále je třeba sledovat současné trendy v této oblasti, tzn. že je třeba být ve styku s relevantními programy a institucemi.

Obecně je třeba zvyšovat vědomí producentů a uživatelů digitálních dokumentů o celé problematice, aby mohlo být časem dosaženo standardnějšího chování, a tím se zvýšila i manipulovatelnost těchto dokumentů včetně šance na jejich dlouhodobější uchování. Je třeba si uvědomit, že tato oblast je velmi mladá a je plná různorodých a často i protichůdných názorů. Ty je třeba průběžně vyhodnocovat, případně aplikovat nosná řešení, jichž je v praxi velký nedostatek anebo mají jen velmi dílčí charakter. NK je objektivně jednou z čelných institucí na digitální scéně a její názor má velkou váhu zejména proto, že je současně i velkým producentem komplexních digitálních dokumentů. A jedním z mála, kteří užívají otevřené standardy a současně zohledňují i povinnosti vyplývající z archivace dat.

Svět digitální komunikace je jiný než svět klasických dokumentů, nicméně není zásadně nový, je jen rychlejší a dynamičtější, a proto též těžce zvladatelný. V některých svých aspektech je i přirozenější. Chceme-li v něm žít, budeme se muset naučit rozhodovat o tom, co má a nemá hodnotu, budeme muset umět odhadnout, kde vedou hranice mezi formálním a neformálním, soukromým a veřejným. Budeme se muset naučit rozpoznávat podstatu věcí od jejich momentální interpretace, ať je jí interpretace určená pro sdělení nebo interpretace vzniklá na základě sdělení. Celkem jistě pak zpětně zjistíme, že i nedigitální svět se chová podle týchž zákonů - pouze se tak děje daleko pomaleji, v cyklech překračujících život jedné generace. Odtud pramení i řada našich postojů.

*Pozn.:*

*Zájemci o hlubší pohled mají k dispozici nejnovější přehlednou studii o této problematice, napsanou jako referát pro ABDOS '99 (Knoll, Adolf. *Electronic Publications: libraries and access mediation problems*. 16 s.) - studie je v RTF formátu (čitelná MS WORD) - viz URL [http://digit.nkp.cz/EnglishArticles/Read\\_Papers/Electronic\\_Publications\\_ABDOS99.rtf](http://digit.nkp.cz/EnglishArticles/Read_Papers/Electronic_Publications_ABDOS99.rtf) (obsahuje mj. další bibliografické odkazy na nejnovější informační prameny).*