

Univerzita Karlova v Praze

Filozofická fakulta

Ústav informačních studií a knihovnictví

Informační věda – Informační studia a knihovnictví

Mgr. Andrea Fojtů

**Strategie, návrh, řízení a administrace rozsáhlých
digitálních knihoven a archivů**

**Strategy, design, management and administration of large
Digital Libraries and Archives**

Disertační práce

Vedoucí práce – doc. PhDr. Richard Papík, Ph.D.

2014

Prohlášení

Prohlašuji, že jsem disertační práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze, 1.8.2014

.....
podpis

Poděkování

Tímto chci především poděkovat vedoucímu své disertační práce doc. PhDr. Richardu Papíkovi, Ph.D. za cenné připomínky při psaní této práce.

Dále mé poděkování patří i všem respondentkám a respondentům z Ústavu knihovnických a informačních studií na Univerzitě Karlově v Praze za jejich otevřenost a ochotu podělit o své názory a zkušenosti. Stejně tak velké díky patří zastupcům větších knihoven, kteří si našli čas ke sdílení pracovních postupů a přístupů ke strategiím, plánům, administraci a managementu svých digitálních knihoven, repozitářů a archivů.

Poděkování náleží i mé rodině, hlavně mamince Aleně a snoubenci Joãovi, bez jejichž podpory a trpělivosti by tato práce nevznikla.

Motto

„If we can put a man on the Moon and sequence the human genome, we should be able to devise something close to a universal digital public library.“

„Pokud dokážeme dostat člověka na Měsíc a rozluštit lidský genom, měli bychom být schopni navrhnout něco jako celosvětovou digitální veřejnou knihovnu.“

Peter Singer

PROHLÁŠENÍ	2
PODĚKOVÁNÍ	3
MOTTO	4
ANOTACE A KLÍČOVÁ SLOVA	8
ABSTRAKT (CZ)	9
ABSTRACT (EN)	10
PŘEDMLUVA	11
ÚVOD	13
1. PŘEDMĚT A CÍL PRÁCE	16
1.1 ZKUŠENOSTNÍ ZÁKLADNA PRO VYPRACOVÁNÍ DISERTAČNÍ PRÁCE – PŘEHLED	19
1.1.1 <i>Vzdělání a zaměstnání</i>	20
1.1.2 <i>Přehled odborných aktivit</i>	20
1.1.3 <i>Účast na projektech</i>	20
1.1.4 <i>Výuka</i>	21
1.1.5 <i>Účast na konferencích (výběr)</i>	21
1.1.6 <i>Publikační činnost (výběr)</i>	23
1.2 VYMEZENÍ ZÁKLADNÍCH POUŽÍVANÝCH TERMÍNŮ	26
2. DIGITÁLNÍ REPOZITÁŘE, KNIHOVNY A ARCHIVY Z POHLEDU INFORMAČNÍ ARCHITEKTURY	33
2.1 POŘÁDÁNÍ ZNALOSTÍ	36
2.1.1 <i>Implementace systémů pro řízení znalostí do digitálních repozitářů</i>	37
3. MODELY DIGITÁLNÍCH KNIHOVEN	39
3.1 KAHN-WILENSKÉHO ARCHITEKTURA	40
3.2 DÉLOS DIGITAL LIBRARY REFERENCE MODEL	44
3.2.1 <i>Třívrstvá konstrukce univerza digitální knihovny</i>	45
3.2.1.1 <i>Klíčové koncepty konstrukce</i>	46
3.2.2 <i>Aktéři digitální knihovny</i>	47
3.3 OPEN ARCHIVAL INFORMATION SYSTEM REFERENCE MODEL - OAIS	53
3.3.1 <i>OAIS prostředí</i>	55
3.3.2 <i>Informační model</i>	56
3.3.3 <i>Funkční model a entity</i>	61
3.3.3.1 <i>Příjem (Ingest)</i>	63
3.3.3.2 <i>Správa dat</i>	65
3.3.3.3 <i>Permanентní úložiště</i>	66
3.3.3.4 <i>Přístup</i>	68
3.3.3.5 <i>Administrace</i>	69
3.3.3.6 <i>Plánování dlouhodobé ochrany</i>	72
3.3.4 <i>Společné služby</i>	75

3.4	OD REFERENČNÍCH, KONCEPTUÁLNÍCH A PROCESNÍCH MODELŮ K DŮVĚRYHODNÝM DIGITÁLNÍM KNIHOVNÁM, REPOZITÁŘŮM A ARCHIVŮM.....	79
3.4.1	<i>Atributy důvěryhodných repozitářů</i>	80
3.4.2	<i>Faktory zodpovědných důvěryhodných repozitářů</i>	82
4.	DLOUHODOBÁ OCHRANA DIGITÁLNÍCH OBJEKTŮ.....	85
4.1	ŽIVOTNÍ CYKLUS DIGITÁLNÍCH DOKUMENTŮ	87
4.1.1	<i>DCC model životního cyklu dlouhodobé ochrany</i>	87
4.2	PLATTER – PLÁN DŮVĚRYHODNÉHO DIGITÁLNÍHO REPOZITÁŘE.....	89
4.3	PLATO PRESERVATION PLANNING TOOL	91
4.3.1	<i>Signifikantní vlastnosti</i>	92
4.3.2	<i>Testování pomocí nástroje PLATO</i>	97
4.3.2.1	Hlavní implementované prvky nástroje PLATO ve SW verzi 4.2	97
4.3.2.2	Práce s nástrojem Plato	99
4.4	KOMERČNÍ SYSTÉMY PRO DLOUHODOBOU OCHRANU	106
4.4.1	<i>Rosetta</i>	106
4.4.2	<i>SDB - Safety Deposit Box</i>	108
4.4.2.1	Preservica	109
4.5	OPEN SOURCE SYSTÉMY A NÁSTROJE PRO DLOUHODOBOU OCHRANU	109
4.5.1	<i>Archivematica</i>	110
4.5.2	<i>RODA & CRIB</i>	110
4.5.3	<i>AIDA</i>	112
5.	AKTIVITY NA ORGANIZAČNÍ ÚROVNI	114
5.1	METRIKY, CERTIFIKAČNÍ A AUDITNÍ NÁSTROJE.....	114
5.1.1	<i>Základní certifikace - Data Seal of Approval</i>	117
5.1.2	<i>Rozšířená a formální certifikace</i>	120
5.1.2.1	NESTOR - Network of Expertise in Long-Term Storage and Long-Term availability of Digital Resources in Germany a DIN 31644 Criteria for trustworthy digital archives	120
5.1.2.2	TRAC - Trustworthy Repositories Audit & Certification: Criteria and Checklist, ISO 16363	121
5.1.3	<i>Další auditní a certifikační nástroje</i>	123
5.1.3.1	DRAMBORA - Digital Repository Audit Method Based on Risk Assessment	123
5.2	EKONOMICKÉ MODELY	126
5.2.1	<i>LIFE (Life Cycle Information for E-Literature)</i>	130
5.2.1.1	Metodologie LIFE	132
5.2.2	<i>4C - Collaboration to Clarify the Costs of Curation</i>	135
6.	POŽADAVKY UŽIVATELŮ NA DIGITÁLNÍ KNIHOVNY, REPOZITÁŘE A ARCHIVY	139
6.1	ZPĚTNÁ VAZBA – DOTAZNÍKOVÉ ŠETŘENÍ	141
7.	SOUČASNÝ STAV PROBLEMATIKY V ČR	143
8.	BEST PRACTICES	149

8.1	SOUČASNÝ STAV, DEFINICE REPOZITÁŘE.....	150
8.2	HARDWAROVÉ A SOFTWAREVÉ ZABEZPEČENÍ, DATOVÝ PLÁN	151
8.2.1	<i>Datový plán</i>	152
8.3	AKVIZIČNÍ PLÁN REPOZITÁŘE	153
8.4	PRÁCE S OBJEKTY, DLOUHODOBÁ OCHRANA.....	156
8.5	FINANČNÍ ZABEZPEČENÍ REPOZITÁŘE	157
8.6	PERSONÁLNÍ ZAJIŠTĚNÍ REPOZITÁŘE	158
8.7	PLÁN ZPŘÍSTUPŇOVÁNÍ.....	158
8.8	KRIZOVÝ PLÁN A MANAGEMENT RIZIK	159
8.9	KONTINUITA REPOZITÁŘE	161
8.10	PROJEKTOVÉ ŘÍZENÍ	162
8.11	AUDIT, CERTIFIKACE A POLITIKY REPOZITÁŘE.....	163
	ZÁVĚR.....	165
	SLOVNÍK ZKRATEK A ACRONYMŮ	169
	SEZNAM OBRÁZKŮ	173
	SEZNAM TABULEK	175
	SEZNAM CITOVANÉ LITERATURY	176
	SEZNAM STUDOVANÉ LITERATURY	183
	PŘÍLOHA Č. 1 - PŘÍKLAD MYŠLENKOVÉ MAPY (FORMÁT JP2)	189
	PŘÍLOHA Č. 2 - PŘEHLED STRATEGICKÝCH PLÁNŮ, ODPOVĚDNOSTÍ A CÍLŮ PLATTER.....	190
	PŘÍLOHA Č. 3 - PŘEHLEDNÝ POPIS FUNKČNÍCH ENTIT A ROLÍ OAIS MODELU	195
	PŘÍLOHA Č. 4 - PŘÍKLAD JEDNODUCHÉ POLITIKY A PŘÍSLUŠNÝCH POSTUPŮ PODLE RLG REPORT	198
	PŘÍLOHA Č. 5 - PŘÍKLAD KRITÉRIA Č. 14 V DATA SEAL OF APPROVAL.....	199
	PŘÍLOHA Č. 6 - SEZNAM KRITÉRIÍ KATALOGU NESTOR.....	200
	PŘÍLOHA Č. 7 - SEZNAM NEJDŮLEŽITĚJŠÍCH STANDARDŮ, METODIK A DOKUMENTŮ PRO EFEKTIVNÍ SPRÁVU REPOZITÁŘŮ, KNIHOVEN A ARCHIVŮ	202
	PŘÍLOHA Č. 8 - DOTAZNÍK PRO INSTITUCE (VZOR)	205
	PŘÍLOHA Č. 9 - TABULKA SOUHRNNÝCH ÚDAJŮ DOTAZNÍKU PRO INSTITUCE.....	207
	PŘÍLOHA Č. 10 - DOTAZNÍK PRO STUDENTY (VZOR)	214
	PŘÍLOHA Č. 11 - TABULKA SOUHRNNÝCH ÚDAJŮ DOTAZNÍKU PRO STUDENTY	217

Anotace a klíčová slova

Anotace: Cílem disertační práce je sestavit obecně použitelnou strategii pro návrh, administraci a řízení rozsáhlých digitálních knihoven, repozitářů a archivů.

Anotation: The aim of the presented dissertation thesis lies in creating a general applicable strategy for design, administration and management large digital libraries, repositories and archives.

Klíčová slova: digitální dokument, digitální objekt, životní cyklus digitálních dokumentů, digitální knihovna, rozsáhlá digitální knihovna, digitální repozitář, digitální archiv, Referenční model OAIS, DELOS Referenční model digitálních knihoven, dlouhodobá ochrana, certifikace, audit, Best Practices

Keyterms: digital document, digital object, digital document life-cycle, digital library, Large Digital Library, VLDL, digital repository, digital archive, Reference model OAIS, The DELOS Digital Library Reference Model, digital preservation, long-term preservation, certification, audit, Best Practices

Abstrakt (CZ)

Předložená disertační práce „Strategie, návrh, řízení a administrace rozsáhlých digitálních repozitářů (knihoven) a archivů“ se, jak vypovídá samotný název, soustřeďuje na přehled standardních a nejčastějších postupů či strategií při tvorbě a řízení digitálních knihoven, repozitářů a archivů nikoli z technického, nýbrž z knihovnicko-informačního hlediska.

Teoretická část práce nabízí základní orientaci v komplexní problematice. Tu ulehčují nejrůznější typy modelů - od referenčních, přes objektové a procesní modely, až po ty funkční. Zachycují budoucí stav, ke kterému chceme dospět a představují abstraktní model reálného systému. Proto práce podrobně rozebírá dva nejdůležitější modely: ISO standard 14721:2003 Otevřený archivní informační systém a de facto standard DELOS Referenční model digitálních knihoven, jež zastupují část návrhů a současně i možné strategické přístupy u digitálních knihoven, repozitářů a archivů.

Ve fázi návrhů je nutné myslet i na datový model, datový plán a respektovat pravidla informační architektury, jenž se dotýkají i systémů pro organizaci znalostí (KOS) a do značné míry ovlivňují výslednou podobu uživatelského rozhraní, které má být dostatečně přívětivé, přehledné a mělo by se ním snadno pracovat.

To vše musí směřovat k důvěryhodným repozitářům a archivům, přičemž důležité je porozumět životnímu cyklu digitálních dat s využitím dostupných „open source“ či komerčních systémů pro řízení a administraci digitálních knihoven, repozitářů a archivů. Jen důvěryhodné systémy dokáží zabezpečit dlouhodobý přístup a ochranu svých digitálních objektů. Průkaznost důvěryhodnosti je postavena na auditu a následné certifikaci, čímž završují (nikoli ale nutně ukončují) práce na návrzích a administraci digitálních knihoven, repozitářů a archivů.

Praktická část disertační práce má podobu „Best Practices“, které popisují 10 základních okruhů pro rychlé zorientování se v dané oblasti. Je pro ně charakteristická maximální obsahová výtěžnost a co nejstručnější forma se zaměřením na praxi. Vychází z průzkumů a osobních rozhovorů se zástupci několika větších českých repozitářů, archivů a knihoven. Jejím cílem je poskytnout zejména managementu a netechnickým (i knihovnickým a informačním) pracovníkům digitálních repozitářů a archivů instrukce pro jejich budoucí či současný informační systém.

Abstract (EN)

The presented dissertation thesis “Strategy, design, management and administration of large Digital Repositories (Libraries) and Archives”, as the very name of it reveals, provides an overview of the standard and most common techniques, strategies in the design and management of digital libraries, repositories or archives from a librarianship’s and information science’s point of view.

The theoretical part presents with basic knowledge of this complex issue and various types of models, from the reference, object and process models to the functional ones. They represent the future and wanted state and stand for an abstract model of a real system. Therefore, this thesis deals with the two most important models, the ISO standard 14721:2003 Open Archival Information System and de facto standard DELOS Reference Model of Digital Libraries, which represent part of the design processes and strategic approaches to digital libraries, repositories and archives.

In the design phase, it is necessary to come up with a good data model, data plan and respect the rules of information architecture, including knowledge management with Knowledge Organization Systems (KOS) and to a large extent influence the final form of the user interface, which should be sufficiently friendly, clear and it should be easy to work with.

This should lead to trusted repositories and archives as well as being able to understand the digital data life cycle, while using available open source or commercial systems for the management and administration of digital libraries, repositories and archives. Only the trusted systems can secure ongoing access and long-term preservation of their digital assets. The sufficient evidence of trust can be taken from audit processes and subsequent certification, which complete (but not necessarily finish) works on the design and administration of digital libraries, repositories and archives.

The practical part of the thesis takes the form of Best Practices and describes 10 basic points for a quick reference in the field. It is typical for practices to be to the point and closely focused on real issues. Therefore they are based on surveys and interviews with representatives of several large Czech repositories. Its aim is to provide a non-technical management in particular employees and library and information professionals of mainly digital repositories and archives with key instructions for their future or current information system.

Předmluva

Před koncem střední školy nastalo klasické, v mém případě i zdlouhavé rozhodování o budoucí škole, ideálně i povolání. Psal se rok 1995 a na pulty knihkupectví dorazila jedinečná a na svou dobu průlomová kniha s názvem „Sprievodca po Internete alebo Internet od A do Z“ od prof. Soni Makulové. Na obalu knihy se psalo, že autorka působí na Univerzitě Komenského v Bratislavě, na katedře Knižničná a informačná veda. A i když jsem si to tenkrát ještě neuvědomovala, o mé budoucnosti v knihovnictví a informační vědě bylo tímto rozhodnuto. Přes počáteční „menší boje“ na zkouškách z katalogizace a historie knihoven jsem se od třetího ročníku mohla plně zaměřit na informační vědu a svět internetu. Semestry utíkaly jako voda a byl čas na diplomovou práci. Mé tápání ve výběru tématu opět ukončila prof. Makulová, která mi navrhla prozkoumat doposud málo známou oblast – přístupnost webů pro znevýhodněné uživatele. Ve své kvalifikační práci „Analýza súčasného stavu a perspektívy riešenia v oblasti iniciatívy prístupnosti webových sídiel“ jsem obhajovala myšlenku Tima Berners-Leeho, zakladatele webu, že „Síla webu spočívá v jeho univerzalitě. Přístup k němu pro každého, nezávisle na schopnostech, je jeho základním aspektem.“ Práce byla rozdělena do deseti kapitol, v úvodních částech definovala nový pojem "přístupnost" a zaměřila se různé mýty a chybná přesvědčení v jeho pojmání. Poukazovala jsem na špatnou situaci na Slovensku a odkazovala na nejznámější světové metodiky přístupnosti webů (WCAG 1.0, WCAG 2.0, Section 508 a české Best Practice). Představila jsem také způsoby jakými znevýhodnění uživatelé (nevidomí, fyzicky handicapovaní apod.) přistupují na web. V závěru jsem na základě vlastní metodiky vyhodnotila nejznámější webová sídla knihoven na Slovensku a navrhla konkrétní řešení ke zlepšení. Práce měla velký ohlas a představila slovenskému knihovnímu světu důležitou oblast k řešení.

Už v posledním ročníku studia jsem začala pracovat ve firmě s mezinárodní působností - IBM International Ltd., kde jsem uplatnila své jazykové schopnosti v německém a anglickém jazyce. A protože jsem se díky své prezentaci na 33. Medzinárodnom informatickom sympóziu INFOS (příspěvek s názvem „Prístupnosť webu alebo prístup k informáciám pre každého“) seznámila s kolegy z pražského Ústavu výpočetní techniky na UK, věděla jsem, že hledají někoho na zprovoznění a následnou správu jejich institucionálního repozitáře. Výběrové řízení proběhlo pro mne úspěšně a já tak měla možnost rozšířit si své znalosti v pro mě v té době méně známé oblasti digitálních archivů a knihoven.

Jelikož se UK zapojila do projektů Odborné komise pro otázky elektronického zpřístupňování vysokoškolských kvalifikačních prací v rámci Asociace knihoven vysokých škol České republiky, bylo mým počátečním úkolem zprovoznit Digitální univerzitní repozitář (v systému DigiTool od izraelské společnosti ExLibris) a umožnit jednotlivým fakultním knihovnám

vkládání elektronických verzí kvalifikačních prací. Začátky, kdy jsem repozitář vedla sama, nebyly snadné. Vyžadovalo to plné pracovní nasazení (10-12 hod. denně) a rychlé učení se. Přes všechna úskalí byl repozitář zprovozněn a dnes Digitální univerzitní repozitář Univerzity Karlovy zpřístupňuje vysokoškolské kvalifikační práce (do roku 2006), mapy a atlasy Přírodovědecké fakulty UK, digitalizované listiny a rukopisy Archivu UK, historické prameny Filozofické fakulty a materiály českého práva uložené v knihovně Právnické fakulty UK, jako i výroční zprávy, výukové a studijní materiály.

Měnicí se technologie, rychlý rozvoj nových hardwarových a softwarových možností a multidisciplinární záběr mého pracovního zaměření si vyžaduje neustálé vzdělávání, aktivní účast na domácích i mezinárodních konferencích a seminářích (viz 1.1 Zkušenostní základna pro vypracování disertační práce - přehled). I z tohoto důvodu jsem si v dubnu 2009 přidala poloviční úvazek v Národní knihovně ČR jako odborná knihovnice a stala se členkou přípravného týmu NK pro výběr a aplikaci Informačního systému pro projekt NDK ČR. Právě práce v tomto pracovním týmu mě inspirovala k sepsání disertační práce. V České a Slovenské republice totiž doposud neexistuje metodika, která by se podrobněji zabývala strategiemi, návrhem, řízením a administrací větších digitálních knihoven, ale i repozitářů a archivů.

Úvod

Jak ostatní informační systémy, tak i digitální repozitáře (knihovny) prošly během své poměrně krátké existence radikálním vývojem. První pokusy knihoven v oblasti zpřístupňování elektronických či zdigitalizovaných dokumentů byly prezentovány jako tzv. hybridní knihovny (*Hybrid Library*). Ty měly představovat prostředí, v němž jsou propojeny tradiční knihovnické postupy a fondy s určitými prvky knihovny elektronických verzí dokumentů. Sloužily jako doplněk, nikoliv jako náhrada tradiční knihovny. Další v pořadí bylo označení virtuální knihovna, která zpravidla obsahovala materiál z různých samostatných knihoven. S masovou digitalizací zejména vzácných tisků, rukopisů, starých výtisků novin apod. se začal užívat termín digitální knihovny. Původně sloužil termín virtuální knihovna jako synonymum k digitální knihovně, ale nově se používá spíše jako nástroj umožňující přístup k strukturované sbírce materiálů z distribuovaných zdrojů. Digitální a elektronické knihovny mohou být i virtuální knihovny, pokud existují jen virtuálně - to znamená, že knihovna neexistuje „v reálném životě“ (i když se s takovým označením setkáváme už málokdy).

Digitální knihovny nabízejí pomocí internetu přímý přístup k digitálním nebo zdigitalizovaným dokumentům, kdy jejich tradiční předloha byla převedena do digitální podoby. Díky digitálním knihovnám mají uživatelé v řadě případů přístup k ohroženým, historickým nebo unikátním dokumentům.

Dnes se můžeme setkat i s názvem digitální repozitář (*Digital Repository*) či institucionální repozitář (*Institutional Repository*), kdy kromě sběru, organizace a správy také nově dlouhodobě ochraňuje digitální obsah pro příští generace.

Design systémů digitálních knihoven představuje komplexní záležitost, která vyžaduje integraci metodologií, poznatků a praktických zkušeností z různých disciplín. Z těch nejdůležitějších to jsou správa obsahu (*Content Management*), správa metadat (*Metadata Management*), vyhledávání informací, rozšiřování informací, správa databází a HCI (Human-Computer Interaction, tedy interakce člověk – počítač) v kontextu přívětivých uživatelských rozhraní.

Implementace moderních digitálních knihoven je v současné době více náročná než v minulosti. Informační konzumenti vyžadují co nejnadanější přístup k neustále narůstajícím a heterogenním informačním zdrojům. Kvůli své základní roli informační produkce a zpřístupňování informací se od digitálních knihoven očekává vysoká kvalita služeb, poskytovaná nepřetržitě - 7 dní v týdnu, 24 hodin denně.

V průběhu let se zvyšovalo množství rozsáhlých digitálních knihoven, a to nejen co do počtu zpřístupňovaných digitálních objektů ale i „obsloužených“ uživatelů. Uvědomila si to i

Evropská unie, která ve snaze zpřístupňování informací a znalostí obyvatelům EU hledala řešení. Evropská komise dokonce již v roce 2005 definovala 3 klíčové oblasti evropských digitálních knihoven¹ - digitalizace, online dostupnost a dlouhodobá ochrana.

První vlaštkou velkých digitálních knihoven byl projekt „The European Library“. Projekt přístupové brány ke katalogům evropských knihoven pod zastřešujícím názvem „The European Library (dále jen TEL)“ byl financován z 5. výzkumného rámcového programu. Formálně byl ukončen závěrečnou zprávou v lednu 2004, avšak i v současné době poskytuje přístup k obsahu 48 evropských národních knihoven a počet dohledatelných informačních zdrojů neustále narůstá.

TEL ve svých počátcích fungoval na distribuované architektuře, tedy vyhledávání i následné zobrazování relevantních výsledků probíhalo přímo na serverech. Až začátkem roku 2010 (v rámci snah o posílení konkurenceschopnosti vůči Europeana) bylo rozhodnuto, že TEL začne budovat vlastní centrální index, nezávislý na úrovni vyhledávacích mechanismů participujících stran. To samozřejmě přináší řadu výzev, např. mnohojazyčné vyhledávání. Jak teoretické poznatky, tak praktické zkušenosti z TEL později daly základ Europeana, představované jako „evropské online muzeum, knihovna a archiv“ v jednom [Purday, 2010]. Podstatný rozdíl mezi TEL a Europeana je na straně poskytovatelů obsahu. Zatímco do Europeana může přispět dokumenty jakýkoliv typ kulturní instituce členských zemí EU, TEL se zaměřuje pouze na národní knihovny bez geograficko-politického omezení na členské státy EU.

I když Evropská komise (EK) započala digitalizační programy již v roce 2000, myšlenka jednotné evropské digitální knihovny se zrodila až v roce 2005². V listopadu 2006 byla idea Evropské digitální knihovny schválena ministry kultury všech členských států EU a posléze i Evropským parlamentem (v září 2007). Europeana poukazuje na širší problémů a výzev se kterými se knihovny takové velikosti musí potýkat. Podstatně složitější je zde modelování dat od různých institucí do vzájemně kompatibilní struktury, agregování a přijímání obsahu (*Ingest*); samozřejmostí je i jejich dlouhodobá ochrana. Přístup k digitalizovaným sbírkám musí být ošetřen právem na ochranu duševního vlastnictví. Europeana nabízí přístup ke kulturnímu bohatství a vědeckým poznatkům z fondů národních knihoven, institucí uchovávajících audiovizuální materiály či muzeí a galerií. Její poslání, resp. strategie možno obecně shrnout do 4 bodů:

¹ http://europa.eu/rapid/press-release_IP-05-1202_en.htm?locale=en

² Představitelé šesti států EU (Francie, Polsko, Německo, Itálie, Španělsko a Maďarsko) v dopise předsedovi EK ze dne 28. dubna 2002 vyslovili požadavek na vytvoření virtuální evropské knihovny. Prezident EK - José Manuel Barroso - na zmíněný dopis reagoval kladně a snahu podpořil. EK přijala 30. září 2005 doporučení „i2010: Digitální knihovny“, které dalo vzniknout *Digital Library Initiative*. Společně s touto iniciativou byla započata jednání o evropském digitálním kulturním dědictví v rámci iniciativy *i2010 Information society (IP/05/643)*.

1. zpřístupňování evropského vědeckého a kulturního dědictví prostřednictvím multioborového (*Cross-domain*) portálu
2. spolupráce na zpřístupňování dokumentů a udržitelnosti výsledného společného portálu
3. motivace dalších kulturních institucí ke zpřístupňování svého digitálního obsahu prostřednictvím Europeana portálu
4. podpora digitalizace evropského kulturního dědictví a vědeckého odkazu.

Europeana portál nabízí jak knihy, periodika, mapy, fotografie, výtvarná díla, tak muzejní objekty, archivní materiály, hudební nahrávky a mluvené slovo z fonografických válečků, magnetofonových pásek, rozhlasové a televizní přenosy či filmy [VORLÍČKOVÁ, 2009].

Dalším známým projektem rozsáhlé digitální knihovny je *World Digital Library*³. Jde o projekt mezinárodní organizace UNESCO a Kongresové knihovny USA. Cílem je prohlubovat mezinárodní a mezikulturní porozumění, rozšířit rozmanitost digitálního obsahu, zvětšit objem zdrojů i jiných než západních kultur. Současně také zpřístupnit historické kulturní památky pro potřeby vědecké a akademické obce i široké veřejnosti a v neposlední řadě zmírnit digitální propast (*Digital Divide*) mezi jednotlivými státy [VORLÍČKOVÁ, 1999].

Zatímco v minulosti byly rozsáhlé digitální knihovny, repozitáře a archivy vnímány pouze jako „knihovny“ na národní či mezinárodní úrovni, dnes bývá rozsah dán především obsahem než jejich geografickým působením. Implementace moderních digitálních knihoven, repozitářů a archivů je v současné době více náročná než kdykoliv předtím a proto cílem této práce je zanalyzovat současný stav v oblasti problematiky (přehledový a vysvětlující profil) a v neposlední řadě sestavit obecně použitelnou strategii pro návrh, administraci a řízení rozsáhlých digitálních knihoven, repozitářů a archivů.

³ <http://www.wdl.org/>

1. Předmět a cíl práce

O problematice digitálních knihoven již bylo popsáno mnoho stránek, ale jen velmi málo z nich se věnuje nastupujícímu trendu větších až rozsáhlých (byť i konsorciálních) digitálních knihoven. Doposud v této oblasti nebyla vydaná žádná ucelená publikace. To byl jeden z důvodů pro sepsání této disertační práce na téma „Strategie, návrh, řízení a administrace rozsáhlých digitálních knihoven a archivů.“ Autorka se dlouhodobě zabývá problematikou digitálních repozitářů a archivů jako i dlouhodobé ochrany digitálních dokumentů v teoretické i v praktické rovině. Disertační práce má za cíl nezabíhat do technických detailů na hardwarové a softwarové úrovni a k politicko-právním aspektům. Jde primárně o návodný text, který má dát představu o postupech v oblasti návrhu, řízení a administrace digitálních knihoven či archivů tak, aby uložení, správa a dlouhodobá ochrana digitálních dat byla zabezpečena jak pro současné, tak i budoucí uživatele. V úvodní části autorka vysvětluje, co všechno musí zodpovědné osoby vzít v úvahu, pokud jde o návrh, řízení, administraci větších digitálních knihoven a podchycení životního cyklu digitálních objektů (tvorba, zpracování, zpřístupňování a dlouhodobé uchovávání).

HYPOTÉZA, STRUKTURA A PŘÍNOSY PRÁCE

Před vlastním výběrem tématu bylo důležité, aby výzkum s ním spojený byl užitečný a prospěšný pro ostatní, aby měl poznatelný, resp. identifikovatelný předmět a aby poskytl předpoklady pro potvrzení nebo vyvrácení předpokladů, z nichž vychází. Výzkum se musí dopracovat k tomu, aby o předmětu sdělil skutečnosti, které ještě nebyly řečeny, anebo aby se podíval novým pohledem na věci, které již řečeny byly. I proto kompilační dílo může mít určitý vědecký smysl, pokud ještě nic podobného v daném oboru neexistuje. Shromažďuje a spojuje tak názory, které byly vysloveny různými badateli na dané téma [ECO, 1997].

Na počátku výzkumu byla stanovena tzv. pracovní hypotéza, která tvrdí, že: „Neexistuje ucelený přehled problematiky v českém jazyce, ideálně pro specifické české podmínky“, která byla dále přetavena do reálné hypotézy: „V ČR není kladen velký důraz na dodržování knihovních standardů při sestavování digitálních objektů (dokumentů), při návrzích a tvorbě digitálních knihoven, repozitářů a archivů. Možnou příčinou se zdá být i multidisciplinární záběr problematiky, rapidní technologický vývoj a příliš mnoho informací. Avšak standardy hrají vitální roli v efektivní a bezpečné správě, administraci, zpřístupňování a dlouhodobé ochraně obsahu digitálních knihoven, repozitářů a archivů.“ Tato hypotéza společně s definovanými tezemi (jež

jsou součástí samostatného dokumentu), stanovily záběr, strukturu a samotný obsah předložené disertační práce.

Struktura práce vychází z tématu práce, která má především přehledový charakter. Kapitoly 1-6 nabízí přehled multidisciplinární problematiky, která vychází především z amerických a anglosaských informačních pramenů. Pro posouzení stavu v České republice (kapitola 7), bylo nutné provést dotazníková šetření a interview u relevantních institucí. Ty ukázaly, že ne všechny instituce dodržují standardy nebo dokonce ještě nemají svůj digitální repozitář či archiv. Proto se jako nevyhnutné jevílo vypracovat pomůcku, návrh či doporučení v podobě „Best Practices“ pro obecné strategie, plány, management a administraci v českých (případně slovenských) digitálních knihovnách, repozitářích a archivech. Aby návrh, řízení a administrace digitálních knihoven, repozitářů a archivů byly co nejvíce efektivní, je nutné porozumět zaužívané terminologii (viz kapitola 1.2 Vymezení základních používaných termínů).

Součástí jsou přínosy, zmapování problematiky pro digitální knihovny, repozitáře a archivy, kompilace a výběr nejdůležitějších nejnovějších poznatků v této multidisciplinární problematice s důrazem na důležitost výběru signifikantních vlastností digitálních objektů pro efektivní dlouhodobou ochranu. Jak uvádí jedna ze stanovených tezí: „Není možné dlouhodobě uchovávat vše, je nutný výběr dokumentů, včetně jejich signifikantních vlastností.“ Praktická ukázka (sestavena autorkou) možných stanovených klíčových vlastností objektů je dostupná formou myšlenkové mapy (viz Příloha č. 1 – Příklad myšlenkové mapy (formát JP2)). Neméně důležitý je i souhrn nejdůležitějších standardů a de facto standardů podle problematických okruhů (viz Příloha č. 7 – Seznam nejdůležitějších standardů, metodik a dokumentů pro efektivní správu repozitářů, knihoven a archivů). V neposlední řadě je to sestavení „Best Practices“ pro komplexní, přesto snadnou a rychlou orientaci v problematice. Dále jde o nasměrování pozornosti na uživatelská rozhraní z pohledu uživatelsky orientovaného designu a na požadavky uživatelů na digitální knihovny, repozitáře a archivy (včetně získávání zpětné vazby).

UKOTVENÍ PRÁCE V KNIHOVNÍ A INFORMAČNÍ VĚDĚ

Práce se věnuje principům, modelům, standardům, funkcím, procesům a nástrojům při budování (větších) digitálních repozitářů, knihoven a archivů z pohledu knihovní a informační vědy. Je zdrojem informací především pro netechnicky orientované pracovníky digitálních repozitářů (vyšší a nižší management apod.).

METODOLOGIE

Úvod práce vymezuje základní termíny tak, jak je autorka v práci používá. Definice vycházejí z příručkových pramenů i další sekundární literatury a jsou doplněny vlastními interpretacemi.

Metody použité při psaní práce jsou: studium publikovaných primárních a sekundárních pramenů, vlastní výzkum, zkoumání podobných řešení ve světě (vycházející z aktivní účasti na domácích i zahraničních konferencích a participace na českých, slovenských i evropských projektech). Metodologicky práce vychází z komparativní a kritické analýzy především zahraničních materiálů (v anglickém a německé jazyce).

V průběhu vytváření této disertační práce autorka uplatnila zejména tyto myšlenkové operace: analýzu, dedukci, indukci, syntézu, komparaci a v neposlední řadě i kompilaci, která má v první části práce vysvětlující charakter.

Dotazníkem, jako jednou z výzkumných metod, byla zjišťována zpětná vazba od uživatelů ohledně rozhraní digitálních repozitářů na obecné úrovni s konkrétním příkladem v podobě Digitálního univerzitního repozitáře UK v Praze. Zpracování současného stavu v ČR u rozsáhlejších digitálních knihoven, repozitářů a archivů probíhalo kombinací metod dotazníkového šetření a interview. V neposlední řadě jsou výstupem hodnotové analýzy „Best Practices“ - nevhodnější postupy, nástroje a strategie při návrhu a administraci (zejména rozsáhlých) digitálních knihoven, repozitářů a archivů, které autorka uvádí v praktické části disertační práce. Větší digitální knihovny, repozitáře a knihovny jsou pro účely této práce chápány a definovány jako instituce, které mohou se svou digitální knihovnou, repozitářem či archivem absolvovat alespoň jeden stupeň certifikace podle Evropského rámce pro certifikaci a audit digitálních repozitářů⁴.

FORMÁLNÍ ASPEKTY DISERTACE

Pro citování použitých informačních zdrojů je v kvalifikační práci využit tzv. Harvardský systém citování, tedy citace pomocí prvního údaje záznamu s datem vydání dokumentu. Pro větší přehlednost jsou použity pro citace v textu hranaté závorky a záhlaví (příjmení autorů a názvy korporací jsou uvedeny velkými písmeny). Citace, které mají jako záhlaví název zdroje jsou malými písmeny. Terminologie v českém jazyce je doplněná o původní termíny v anglickém a/nebo německém jazyce (dle zpracovaného zdroje), které jsou pro vyšší přehlednost uvedeny velkými počátečními písmeny a kurzívou v závorkách. Je to hlavně z důvodu neustálené terminologie v této oblasti a snaze zabránit nejednoznačnosti vyjádření.

Pod čarou jsou uvedeny vysvětlivky, eventuelně doplňující informace k textu. Údaje z dotazníků (Příloha č. 9, 11) jsou uváděny jako citace (kromě interview v Knihovně Akademie věd

⁴ Podrobnější informace k Evropskému rámci pro certifikaci a audit digitálních repozitářů lze nalézt v kapitole 5 Aktivity na organizační úrovni.

ČR) a nejsou nijak korigovány či doplňovány. Velká počáteční písmena v českém jazyce pod jednotlivými obrázky jsou uváděny v případech, ve kterých by v českém jazyce mohlo dojít k nepřesnému pochopení či významovému zaměnění termínů a názvů za běžný text.

V práci je řádkování „1,5“, font „11“, typ písma Calibri, okraje – 2,5 cm (3,5cm). V části Best Practices jsou praktické ukázky uvedeny písmem Times New Roman.

1.1 Zkušební základna pro vypracování disertační práce – přehled

Podklady pro text předkládané disertační práce vznikaly v průběhu celého doktorandského studia. Získávání teoretických poznatků a výměna praktických zkušeností při budování a provozu Digitálního univerzitního repozitáře v Ústavu výpočetní techniky Univerzity Karlovy v Praze (ÚVT UK) byla možná díky časté účasti na prestižních zahraničních a domácích konferencích a téměř každodenním kontaktu s předními experty a odborníky z celého světa. K tématu digitálních knihoven, komerčních a open-source systémů pro správu digitálního obsahu autorka publikovala řadu článků v českém, slovenském a anglickém jazyce, citované i v časopisech našeho oboru⁵. Další odborný posun nastal díky zapojení se do evropského projektu DigitalPreservationEurope (DPE) a účasti na školení auditorů DRAMBORA (*Digital Repository Audit Method Based On Risk Assessment*). Zaměření autorky na dlouhodobou ochranu, audit a certifikace digitálních knihoven/repozitářů⁶ bylo určující pro odbornou revizi českého vydání publikace „Průvodce plánem důvěryhodného digitálního repozitáře – PLATTER“ [ROSENTHAL, 2009] a také nespornou výhodou při přípravě podkladů pro projekt Národní digitální knihovna (2007-2011) v Národní knihovně ČR, Oddělení digitální ochrany. V současné době autorka pracuje jako vedoucí Digitálního univerzitního repozitáře a vedoucí Oddělení knihovnických aplikací na ÚVT UK, vede tým 20 lidí a aplikuje tak do praxe nejen odborné znalosti z informační vědy, ale i managementu.

⁵ Např. VERMAATEN, S. LAVOIE, B., CAPLAN, P. Identifying Threats to Successful Digital Preservation: the SPOT Model for Risk Assessment. *D-Lib Magazine* [online] 2012, roč. 18, č. 9/10. Dostupné na World Wide Web: <<http://www.dlib.org/dlib/september12/vermaaten/09vermaaten.html>>.

⁶ Používání termínů „digitální knihovna“ a „digitální repozitář“ je podrobněji vysvětleno v kapitole 2.1 Vymezení používaných termínů.

1.1.1 Vzdělání a zaměstnání

září 2000 – červen 2005 - Filozofická fakulta Univerzity Komenského, katedra Knižničnej a informačnej vedy, Bratislava - absolventka magisterského studijního programu Informační věda (titul Mgr.)

leden 2005 – září 2006 IBM Integrated Supply Chain, Ltd., Customer Fulfillment Centre & European System Integrator Sales Centre, Bratislava

říjen 2006 – současnost Univerzita Karlova v Praze, Ústav výpočetní techniky, Praha - vedoucí Digitálního univerzitního repozitáře

- únor 2012 – současnost - vedoucí Oddělení knihovních aplikací
- duben 2009 - prosinec 2012 - Národní knihovna v Praze, Oddělení digitální ochrany, Praha, odborná knihovnice Referátu Centrálního digitálního úložiště

1.1.2 Přehled odborných aktivit

2006 - 2010 - účast v pracovní skupině „Práca so znevýhodnenými používateľmi pri Slovenskej národnej knižnici“ (<http://www.infolib.sk/index/podstranka.php?id=1416>)

2007 - současnost - porotce v soutěži Biblioweb (<http://skip.nkp.cz/akcBweb10.htm>)

2006, 2010 - porotce v soutěži TOP WebLib 2005, 2009
(<http://www.infolib.sk/index/podstranka.php?id=1287>)

září 2008 - příprava materiálů pro prezentaci o informační gramotnosti (setkání na Albeři UK, seminář pro nové studenty univerzity)

2009 - odborná revize českého vydání publikace: ROSENTHAL, C., et al. Průvodce plánem důvěryhodného digitálního repozitáře (PLATTER). 1. vyd. Praha: Národní knihovna České republiky, 2009. 51 s. ISBN 78-80-7050-569-4.

2008 - současnost - oponentura bakalářských a diplomových prací na ÚISK FF UK

1.1.3 Účast na projektech

Účast na národních projektech výzkumu a vývoje:

2007-2011 - projekt Národní digitální knihovna (2007-2011)

- člen přípravného týmu NK pro výběr a aplikaci Informačního systému - projekt NDK ČR

2007 - spolupráce na grantu GA AV ČR A701010606 Interakce člověk - počítač v humanitních vědách

Účast na projektech EU:

2006-2009 - DigitalPreservationEurope – DPE

- účastník projektu za ÚVT Univerzity Karlovy v Praze
- překlady 'Briefing Papers' do slovenštiny „PASQUI, V. Ochrana digitálních dat a otevřené archívy“ a „WRIGHT, R. Dlhodobá ochrana digitálního audiovizuálního obsahu“

1.1.4 Výuka

LS 2007/08 - výběrový seminář Základy uživatelského designu (i pro studenty programu Erasmus)

LS 2008/09 - výběrový seminář pro studenty kombinovaného studia Základy uživatelského designu

LS 2008/09 - výběrový seminář Základy práce s informačními zdroji (spoluvyučující Jana Matějková, Doc. PhDr. Richard Papík, Ph.D.)

ZS 2009/10 - výběrový seminář pro studenty Erasmus Preservation of the Library Collections from the 19th and 20th century (spoluvyučující Mgr. Jan Hutař, Ph.D.)

ZS 2009/10 - výběrový seminář Úvod do uživatelského výzkumu (spoluvyučující Mgr. Lenka Němečková)

1.1.5 Účast na konferencích (výběr)

2007

DELOS Conference, prosinec 2007, Pisa, Itálie

European Conference on Digital Libraries, září 2007, Budapešť, Maďarsko

Konference Digitálna knižnica, říjen 2007, Jasná, Slovensko

2008

DELOS Summer School 2008, červen 2008, Tirrenia, Itálie

9. konference Archivy, knihovny, muzea v digitálním světě 2008, prosinec 2008

2009

DRAMBORA Auditors' Training, leden 2009, London, United Kingdom

INFORUM 2009, květen 2009, Praha

CASLIN 2009, červen 2009, Klášter Teplá

ECDL – 13th European Conference on Digital Libraries – září 2009, Corfu, Greece

LiWA - Living Web Archive - General Assembly and Work Package meeting, červenec, 2010, Praha

ECDL - 14th European Conference on Digital Libraries - září 2010, Glasgow, United Kingdom

2011

Seminář: Představuje se Alma od ExLibris, Praha - listopad 2011

iPres – 8th International on Preservation of Digital Objects, listopad 2011, Singapore

12. konference Archivy, knihovny, muzea v digitálním světě 2011, Praha – prosinec 2011

ANADP - Aligning National Approaches to Digital Preservation 2011 – květen 2011, Talinn, Estonsko

35. seminář knihovníků muzeí a galerií, září 2011, Český Těšín

2012

Seminář AKVŠ, listopad 2012, Brno

2013

Seminář Otevřené repozitáře, květen 2013, Brno

Knihovny budoucnosti, červen 2013, Brno

Digital Preservation Advanced Practitioner Training, červenc 2013, Glasgow, United Kingdom

10th International Conference on Preservation of Digital Libraries, září 2013, Lisbon, Portugal

Meeting of Open Access Week Organizers, září 2013, Pardubice

Aligning National Approaches to Digital Preservation: An Action Assembly, listopad 2013, Barcelona, Spain

2014

Seminář k digitalizaci pro KK III., duben 2014, Brno

Otevřené repozitáře, duben 2014, Brno

INFORUM 2014, květen 2014, Praha

1.1.6 Publikační činnost (výběr)

2004

FOJTŮ, A., GREŠKOVÁ a M., LÁNYIOVÁ, I. Spracovanie a využívanie informácií z hľadiska kognitívnych vied. *Knižničná a informačná veda 20*. Library and Information Science, sv. 20. Bratislava : Univerzita Komenského, 2004, s. 153-179. ISBN 80-223-1993-7.

2005

FOJTŮ, A. Prístupnosť webu alebo prístup k informáciám pre každého. *Infos 2005 : Zborník z 33. medzinárodného informatického sympózia*. Bratislava : SSK; ÚK SAV, 2005. - s. 46-53.

2007

FOJTŮ, A. a KRBEC, P. Digitální zdroje a jejich budoucnost. *Automatizace knihovnických procesů – 11 : sborník z 11. ročníku semináře pořádaného ve dnech 16.–17. května 2007 v Liberci*. Praha : ČVUT, 2007. 136 s. ISBN 978-80-01-03691-4.

FOJTŮ, A. Digitálny obsah a jeho ochrana. *ITlib. Informačné technológie a knižnice* [online], 2007, č. 02 [cit. 2007-15-06]. Dostupné na: <http://www.cvtisr.sk/itlib/itlib072/fojtu.htm>. ISSN 1336-0779.

2008

FOJTŮ, A. Co Web 2.0 dal a vzal. *Sborník prezentací a příspěvků z konference IKI 2008 – Informace, konkurenceschopnost, inovace*. Praha : Česká informační společnost, 2008. Dostupné na: http://cisvts.cz/wp-content/uploads/2012/11/Fojtu_IKI.pdf

HUTAŘ, J., FOJTŮ, A. a PAVLÁSKOVÁ, E. DRAMBORA - nástroj na interní audit digitálních úložišť v nové online verzi a postřehy z provedených auditů [online]. In *Inforum 2008*, Praha 28.-30. května 2008 [online]. Praha : Albertina Icome, 2008 [cit. 2008-11-18]. ISSN 1801-2213. ISSN 1801-2213. Dostupný na: <http://www.inforum.cz/archiv/inforum2008/sbornik/45>

FOJTŮ, A., et al. *Objevujeme planetu Internet : e-learning, multikulturní výchova, modelové hodiny, metodické postupy*. Praha : Erudis, 2008. 88 s. ISBN 978-80-254-1416-3.

FOJTŮ, A. a PAVLÁSKOVÁ, E. Plán dlouhodobé ochrany digitálních dokumentů v praxi. *Příspěvek na konferenci: 9. konference Archivy, knihovny, muzea v digitálním světě 2008.*

2009

FOJTU, A. Open source versus commercial solutions for a long-term preservation in digital repositories. *Systémy pro zpřístupňování VŠKP : zkušenosti, možnosti, nabídky, potřeby : 4. ročník semináře konaného 21. 10. 2009 na VUT v Brně* [online]. Praha : Odborná komise pro otázky elektronického zpřístupňování VŠKP AKVŠ ČR, 2009-11-02 [cit. 2009-11-02]. Dostupné na: <http://www.evskp.cz/seminar/2009/>. ISSN 1803-7003.

FOJTŮ, A. Kognitivně znevýhodnění v kontextu přístupnosti a použitelnosti webu. *Uživatelsky přívětivá rozhraní*. Brno : Hořava & Associates, 2009. 177 s. ISBN 978-80-254-5295-0.

FOJTU, A. Czech National Digital Library and long-term preservation issues. *Cultural Heritage Online . Empowering Users : An active role for user communities*. 2009. Dostupné na: <http://www.rinascimento-digitale.it/eventi/conference2009/slides16-2/fotju.pdf>

FOJTŮ, A. Systém DigiTool a budoucnost dlouhodobé ochrany digitálních dat. Příspěvek na konferenci: 2. Setkání uživatelů DSpace, duben 2009. Dostupné na: http://dspace.vsb.cz/dspace/bitstream/10084/71330/1/fojtu_dsugcz09.pdf

NĚMEČKOVÁ, L. a FOJTŮ, A. Příspěvek na téma 'Výuka Uživatelského výzkumu na ÚISK', Infokon 2009, Masarykova Univerzita v Brně

2010

Studium informační vědy a znalostního managementu v evropském kontextu

- spolupráce na skriptech 'Modul č. 3 - Informační věda', Kapitola 11 – Human-Computer Interaction

2011

FOJTŮ, A. Open source nástroje pro dlouhodobou ochranu digitálních dokumentů – Konference Archivy, knihovny a muzea v digitálním světě 2010.

HUTAŘ, J., MELICHAR, M. a FOJTŮ, A. Dlouhodobá ochrana digitálních dokumentů a projekt NDK. *Konference Knihovny současnosti 2011.*

2014

FOJTŮ, A. Dôvera ako kľúčový koncept digitálnych repozitárov. *Knižničná a informačná veda 25. Library and Information Science*. Bratislava : Univerzita Komenského, 2014.

1.2 Vymezení základních používaných termínů

Data, jako základní jednotky digitálních repozitářů, knihoven a archivů, lze definovat jako hodnoty, čísla, znaky nebo symboly reprezentující informaci, která může být přijata, uchována, zpracována počítačem [KEENAN, 2000, s. 68]. Data mohou být tvořena libovolnými řetězci znaků (čísel, příkazů, vět) uloženými na informačním nosiči. Data nemají zpravidla význam sama o sobě, ale teprve jsou-li pochopena, interpretována, komunikována a využita člověkem nebo počítačem stávají se smysluplnými informacemi [RESSLER, 2006, s. 26].

Dokument (*Document*) jako jakákoliv informace nesoucí obsah v elektronicky zaznamenané podobě je základní jednotkou, ze které jsou tvořeny sbírky informací [REITZ, 2004, s. 192]. Rozlišujeme "**born-digital**" **dokument** (*Born-digital Document*), jako dokument, který vznikl v digitální podobě a nemá tedy svůj digitální protějšek⁷ a digitalizovaný dokument. **Elektronický dokument** (*Electronic Document*) je často pojímán jako synonymum digitálních dokumentů. Dictionary for Library and Information Science [REITZ, 2004, s. 215] jej definuje jako digitální objekt, který je obecnějšího charakteru než papírové dokumenty, přičemž elektronické dokumenty mohou být nejrůznějšího charakteru včetně textových, tabulkových, grafických, audio či video souborů. **Digitální objekt**, v technickém smyslu, představuje typ datové struktury pozůstávající z digitálního obsahu, unikátního identifikátoru obsahu (označovaného jako „handle“) a dalších dat o obsahu, např. popisná, administrativní metadata [REITZ, 2004, s. 217; Kahn, 1995]. **Metadata** či **metainformace** (*Metadata, Metainformation*) v nejobecnějším smyslu chápeme jako „data o datech“ vytvořená za účelem umožnění specifické funkce nebo funkcí. Jsou nástrojem popisu a prostředkem pro uchování, vyhledání a zprostředkování obsahu [KATUŠČÁK, 1998, s. 228; REITZ, 2004, s. 438]. Z technického hlediska se též jedná o datový formát, který může obsahovat informace nepřímo získané z obrazu, jako i informace vztahující se k aktuálnímu popisu obsahu obrazu.

Někdy se v souvislosti s pojmem digitální objekt setkáváme i s označením **intelektuální entita** (*Intellectual Entity*). Jde o termín, jež byl převzat z PREMIS (*Preservation Metadata Maintenance Activity*)⁸ a představuje souhrn veškerých informací, jež tvoří intelektuální jednotku

⁷ Původní digitální dokumenty (tedy nikoliv digitalizované dokumenty vzniklé převodem z analogové do digitální podoby) u kterých nebývá analogový protějšek. I když se to nepředpokládá, jeden z možných (i když ne zcela obvyklých a častých) přístupů k dlouhodobé ochraně představuje tisk na papír. Tento přístup se častěji vyskytuje v soukromé sféře (komerčních společností), než v tradičním prostředí MLA (*Museum, Libraries and Archives*), tedy muzeí, knihoven a archivů.

⁸ PREMIS (*Preservation Metadata Maintenance Activity*) zastupuje skupinu expertů, kteří sestavili mezinárodní standard *The PREMIS Data Dictionary for Preservation Metadata*, definující tzv. preservační, ochranná metadata pro dlouhodobou ochranu digitálních objektů. Tento standard, v podobě datového

za účelem jejich popisu a následné správy; například mapa nebo fotografie. Intelektuální entita může obsahovat další intelektuální (pod)entity. Příkladem může být webová stránka, která může obsahovat text, obraz či video [OCLC, 2008, s. 6].

Digitalizovaný dokument představuje dokument, jenž v procesu digitalizace vznikl pomocí převodu z analogové (v kontextu této práce nejčastěji papírové) podoby do digitálního tvaru. Například digitalizace mapy představuje převod grafického obsahu papírové mapy na digitální a jeho uložení na elektronický nosič (médiu), čehož výsledkem je digitální mapa. **Digitalizace** (*Digitization*⁹) představuje proces konverze analogových dat do digitální podoby, kdy jsou obvykle použité binární systémy [REITZ, 2004, s. 181].

Aby digitální/digitalizované dokumenty mohly naplnit svou podstatu „nosiče informací“ musí být zachován jejich původ, konzistentnost, originalita, a to jak po obsahové (autenticita), tak technické stránce (integrita). **Autenticita digitálních objektů** (*Authenticity Of Digital Objects*) představuje důkaz původnosti dat, jejich pravost, originalitu, pravdivost či reálnost [REITZ, 2004, s. 40]. Naproti tomu **integrita digitálních objektů** (*Integrity of Digital Objects*) umožňuje systému verifikovat, jestli byly provedeny nějaké změny. Je to podmínka, která existuje pokud jsou data nezměněná, neupravená, nepoškozená či nezničená nějakou operací [REITZ, 2004, s. 349]. Představuje též míru, do jaké jsou data imunní vůči poškození, obzvláště při přenosu a uchování [KEENAN, 2000, s. 69]. Tedy **integrita dat** (*Data Integrity*) garantuje, že data při přenosu nebudou nepovoleně vytvořena, zachycena, modifikována či smazána. Datová integrita tak představuje kvalitu nebo stav celistvosti a neměnnosti a poukazuje na konzistentnost, přesnost a správnost dat [REITZ, 2004, s. 155].

Data, resp. digitální obsah lze formou digitálních objektů zpracovávat, uchovávat a zpřístupňovat v **digitálních knihovnách** (*Digital Library*), dále jen DK, které nejčastěji chápeme jako online ekvivalent klasických „kamenných“ knihoven. Tyto objekty mohou mít „born-digital“ charakter (vznikly jako digitální dokumenty) nebo jako digitalizované protějšky fyzických dokumentů. V knihovnách započal proces digitalizace katalogy, dále se posunul k indexům časopisů a abstraktním službám, periodikům, příručkám a nakonec ke knižní produkci [REITZ, 2004, s. 217].

Witten [2010] definuje DK jako „specializovanou (organizovanou) sbírku digitálních objektů, reprezentovanou textem, videem a audiem se stanovenými metodami přístupu a vyhledávání,

slovníku, najdeme jako XML schéma v mnohých komerčních či open-source systémech a nástrojích pro dlouhodobou ochranu.

⁹ V anglické terminologii se nejčastěji setkáme s termínem *Digitization* - digitalizace. Dictionary For Library and Information Science [REITZ, 2004] nacházíme termín *Digitation*, který je definován jako automatický „proces konverze analogových dat do digitální podoby“, zatímco *Digitization* měří tyto automatické procesy.

výběru, organizace a správy.“

Přes značnou heterogenitu vnímání DK, nesou DK tyto společné znaky [BARTOŠEK, 2004; PINKAS, 1999]:

- digitalizace fyzických dokumentů není klíčovým prvkem, organizace obsahu za účelem lepšího přístupu
- interoperabilita - propojení různorodého obsahu do jednoho celku, DK není uzavřenou entitou
- zpřístupňovaný obsah je heterogenní (způsobem uložení, organizací a správou obsahu, různorodostí sbírek, dokumentů a služeb po obsahové i typové stránce a použitými platformami), dynamický (začleňováním a vyřazováním komponentů do/ze struktury DK) a multimediální (povahou dat)
- architektura, tedy realizace DK vyžaduje propojení různých, často autonomně zpracovaných informačních komponentů, které musí být pro uživatele transparentní. Jedná se tedy o distribuovanost - rozdělení sbírek a služeb do prostorově oddělených míst
- cílem je poskytnout uživatelům jednotný přístup k relevantním informacím bez ohledu na jejich formu, formát, způsob a místo uchování
- přizpůsobitelnost (škálovatelnost) - schopnost systému reagovat na vzrůstající počet dokumentů, uživatelů a jejich požadavků bez nutnosti zásadních změn.

Z pohledu uživatelů představují digitální knihovny jeden logický celek s vymezenými částmi, se kterým komunikuje v jednotném informačním jazyce a z něhož dostává informační odpovědi na své požadavky. Jedná se o knihovnu, která virtuálně zpřístupňuje všechny své zdroje uživatelům elektronickými kanály [REITZ, 2004, s. 179]. Setkáváme se však i s definicí, že digitální knihovna představuje knihovnu, která virtuálně zpřístupňuje všechny své zdroje uživatelům elektronickými kanály.

Pro větší typy digitálních knihoven, jako třeba **rozsáhlou digitální knihovnu** (*Very Large Digital Library – VLDL*) neexistuje doposud jednoznačně akceptovaná definice. Nejčastěji je tento typ knihoven vnímán z pohledu množství či typů objektů (např. velké formáty dokumentů – kartografické dokumenty), počtu uživatelů, kterým jsou dokumenty zpřístupňovány či například kvantu heterogenit, jež je potřebné sesouladit [VLDL, 2011].

V kontextu virtuálního zpřístupňování jsme se hlavně v období překotné digitalizace katalogů setkávali nejčastěji s označením **virtuální knihovna** (*Virtual Library*), tedy jakousi „knihovnou bez stěn“, ve které sbírky neexistují v papírové, mikrofišové či jiné hmotné podobě na fyzickém místě, ale jsou elektronicky dostupné v digitálních formátech prostřednictvím počítačové

sítě. Zajímavé je, že některé knihovny a knihovní systémy si i dnes říkají „virtuální“, jako například Colorado Virtual Library¹⁰.

Nutno podotknout, že termín digitální knihovna se zdá být více vhodný, jelikož termín virtuální (vypůjčený z virtuální reality) naznačuje, že zkušenost se čtením a vyhledáváním v takové knihovně není stejně „reálná“ (není zde posuzován rozdíl ve vnímání informací v elektronické či papírové podobě), i když informace jako taková zůstává stejná, zachovaná bez ohledu na formát [REITZ, 2004, s. 760]. Za synonyma virtuální knihovny považujeme i **elektronickou** (*Electronic Library*) a **hybridní knihovnu** (*Hybrid Library*). V nejnovějších informačních zdrojích a výzkumných pracech, i s ohledem na jejich odkazování na ISO standard Otevřeného archivního informačního systému - OAIS, se setkáváme s pojmem **digitální repozitář** (*Digital Repository*) či **digitální archiv**¹¹ (*Digital Archive*). Bývají často významově zaměňovány za termín „digitální knihovna“ (*Digital Library*) či "institucionální repozitář" (*Institutional Repository*), avšak repozitář a archiv (někdy označován též jako úložiště) se používají i pro několik různých vzájemně se překrývajících systémů, aplikací či souborů aplikací nebo organizací, které umožňují uchování digitálního obsahu nejrůznějšího charakteru podle principů dlouhodobé ochrany. Podle amerického modelu „Kahn-Willensky“ [KAHN, 2006; BRATKOVÁ, 2009] je digitální repozitář vedle uživatelského rozhraní (*User Interface*), identifikačního systému (*Handle System*) a vyhledávacího systému (*Search System*) „pouze“ jedním z komponentů architektury digitální knihovny.

Z důvodu nejednoznačných rozdílů v definicích jsou pro účely této disertační práce termíny digitální knihovna, digitální repozitář, digitální archiv, virtuální knihovna, elektronická knihovna a hybridní knihovna významově zaměnitelné a označují tak „organizaci, která může být virtuálního charakteru a jejíž účelem je komplexní sběr, správa a dlouhodobá ochrana bohatého digitálního obsahu“ [AMERICAN, 2007]. Pro snadnější přehled v problematice jsou zde digitální kurátorství a dlouhodobá ochrana chápány jako synonyma. Větší digitální knihovny, repozitáře a knihovny jsou v této práci chápány a definovány jako instituce, které mohou se svou digitální knihovnou, repozitářem či archivem absolvovat alespoň jeden stupeň certifikace podle Evropského rámce pro certifikaci a audit digitálních repozitářů.

Tento obsah by měl být však uchovávan ideálně pouze v **důvěryhodných digitálních repozitářích** (*Trusted/Trustworthy Digital Repository*). Repozitář je považován za důvěryhodný, pokud lze prokázat jeho schopnost plnit určité funkce, a pokud tyto funkce splňují minimální dohodnutá kritéria, jež by měly splňovat všechny důvěryhodné repozitáře. Neexistují jednotná

¹⁰ Dostupné na <http://www.coloradovirtuallibrary.org/>.

¹¹ Digitální archiv (*Digital Archive*) definujeme jako systém navržený pro dlouhodobou lokalizaci, uchování a zpřístupňování digitálních materiálů. Digitální archiv může využívat nejrůznější ochranné metody pro uchování materiálů i přes změnu technologií, včetně emulace a migrace [REITZ, 2004, s. 216; RESSLER, 2006, s. 29].

kritéria napříč celým spektrem typů a druhů digitálních knihoven/repozitářů a archivů. Základem je stanovení podmínek, za jakých má být obsah (resp. jejich podoba ve formě digitálních objektů) uchován a jakými prostředky (technickými, finančními, personálními a administrativními) má být toho docíleno. Jestliže repositář (ideálně externím auditem) dostatečně prokáže, že svým závazkům je schopný dostát, je repositář považován za důvěryhodný. **Audit** zde chápeme jako přezkoumání nebo monitorování výkonu systému [REITZ, 2004, s. 40].

Základem budování důvěryhodných repositářů je respektování referenčního modelu o **Otevřených archivních informačních systémech - OAIS** (*Open Archival Information System - OAIS*). Jedná se o model pro systémy digitální archivace původně sestavený Consultative Committee for Space Data Systems a přijatý ISO jako mezinárodní normu ISO 14721:2002. **Referenční model** (*Reference Model*) je abstraktní konstrukce pro pochopení signifikantních vztahů mezi entitami v určitém prostředí a sestavení sjednocujících standardů nebo specifikací podporujících toto prostředí¹². Není omezený konkrétním standardem, technologiemi či konkrétní implementací, ale společná sémantika určuje jednoznačnost mezi a v rámci jednotlivých provedení. Může též sloužit jako pomůcka pro vysvětlení daného prostředí nespecialistům. OAIS model je hodně využíván knihovnami jako rámec pro sestavení repositářů, knihoven, archivů, resp. informačních systémů obecně dlouhodobé ochrany [REITZ, 2004, s. 502].

Dlouhodobá ochrana (*Digital Preservation*) reprezentuje proces uchovávání materiálů v digitálních formátech, včetně ochrany „bit stream“¹³ a schopnosti zobrazení obsahu reprezentovaného danými „bit stream“. Důvodem je skutečnost, že některá digitální média rychle zastarávají („bit-rot“¹⁴) a jelikož je digitální objekt neoddělitelně provázaný se svým přístupovým prostředím (SW, HW), dochází i k zastarávání samotných informací. Taktéž odkazuje na aktivity digitalizace materiálů originálně vytvořených v nedigitální podobě (tisk, film, pod.), čímž zabraňuje nezvratným ztrátám kvůli degradaci fyzických nosičů [REITZ, 2004, s. 218]. V anglo-saských zemích má anglický termín *Digital Preservation*. **Digitální kurátorství** (*Digital Curation*)¹⁵, které často bývá synonymem dlouhodobé ochrany, ve skutečnosti představuje podsložku dlouhodobé ochrany

¹² Při návrzích informačního systému nesmíme opomíjet ani datový model (*Data Model*), který představuje jeho základní stavební kámen. Popisuje (nejčastěji v grafické podobě) objekty reprezentované počítačovým systémem společně s jejich vlastnostmi a vztahy a taktéž definuje, jaké informace mají být uchovány v databázi a způsob jejich organizace [REITZ, 2004, s. 157].

¹³ V počítačové terminologii představuje „bit-stream“ sekvenci bitů.

¹⁴ V počítačové oblasti označuje „bit-rot“ změnu elektrického náboje v bitech, čímž může pozměnit uchovaná data. Zejména v oblasti dlouhodobé ochrany představuje „bit-rot“ proces postupného zastarávání médií/nosičů.

¹⁵ Podle DCC (*Digital Curation Centre*) je *Digital Curation* „proces správy, ochrany a přidávání hodnoty digitálním datům v průběhu jejich životního cyklu“ [DCC, cca2012]. Podrobněji na <http://www.dcc.ac.uk/digital-curation/what-digital-curation>.

(*Digital Preservation*). Zahrnuje správu, ochranu s přidáním hodnoty digitálním výzkumným datům během jejich životního cyklu, přičemž dlouhodobá ochrana pokrývá „pouze“ ochrannou složku [ŠVÁSTOVÁ, 2014]. Pro snadnější porozumění jsou v této práci termíny dlouhodobá ochrana a digitální kurátorství chápány jako synonyma.

V roce 2007 sestavila pracovní skupina odborníků z Americké asociace knihoven (ALA) a Asociace pro knihovní sbírky a technické služby (ALCTS) tři základní definice dlouhodobé ochrany. Ta nejobecnější a nejstručnější říká, že „dlouhodobá ochrana kombinuje politiky, strategie a aktivity, které dlouhodobě zajišťují přístup k digitálnímu obsahu“ [ALCTS, 2007]. Ty další se pak zaměřují i na podstatu dlouhodobého přístupu, kdy veškeré preservační aktivity a strategie musí brát do úvahy **vytváření**, **integritu** a vhodné metody **správy obsahu**. Při **vytváření obsahu** je nutné dbát na přehledné a úplné technické specifikace, vytváření spolehlivých "master"¹⁶ souborů, dostatečná popisná, administrativní a strukturální metadata pro zabezpečení budoucího přístupu a podrobné kvalitativní kontroly procesů. **Integrita obsahu** zahrnuje dokumentaci veškerých politik, strategií a procedur v rámci repozitáře jako systému i jako organizace, využívání persistentních (trvalých) identifikátorů, zaznamenávání původu (provenience) a dokumentace jakýkoliv změn provedených na objektech, implementaci bezpečnostních pravidel a verifikačních mechanismů a v neposlední řadě provádění rutinních auditů. **Správa obsahu** věnuje pozornost robustní výpočetní a síťové infrastruktuře, uchovávání a synchronizaci souborů v několika lokalitách, neustálému monitorování a správě souborů. Dále též využívání programů pro migraci a emulaci formátů, software a hardware, včetně obnovování (*Refreshing*) souborů, vytváření a testování plánu pro prevenci a řešení následků havárií a jiných pohrom (*Disaster Recovery*). **Emulací** (*Emulation*) rozumíme proces, kterým počítačový program nebo zařízení umožňuje jednomu systému imitovat jiný provedením daného úkolu. V oblasti digitální archivace jde o ochrannou aktivitu, která využívá speciální software, který překládá instrukce z originálně uchovaného programu do novější platformy, čímž obchází nutnost uchovávat zastaralý hardware nebo systémový software [REITZ, 2004, s. 247]. **Migrace** (*Migration*) nejčastěji¹⁷ transformuje,

¹⁶ Master file, volně překládán jako archivní formát, v kontextu problematiky dlouhodobé ochrany odkazuje na formát, který je uchováván v permanentním úložišti digitálních repozitářů. Z těchto typů souborů se často vytváří uživatelské kopie (*User-copy*). Podle referenčního modelu OAIS odpovídá objektům, jež jsou součástí AIP balíčků. Podrobněji viz kapitola 2 Modely digitálních repozitářů, knihoven a archivů. Například v Digitálním univerzitním repozitáři Univerzity Karlovy v Praze jsou uchovávány digitalizované mapové listy ve formátu TIF, uživatelům jsou zpřístupňovány pouze materiály v JP2000 (viz <http://digitool.cuni.cz/DIGITool-33.html>).

¹⁷ Migrace formátů souborů, jako jeden z přístupů k dlouhodobé ochraně digitálních dokumentů, je realizována nejčastěji v případě zastaralých (*Obsolete format*) formátů. Nicméně lze ji využít i případech, kdy např. instituci/organizaci uchovávající digitální dokumenty vypršela softwarová licence nebo uchování konkrétních formátů je finančně náročné.

resp. převádí zastaralé formáty souborů do nových, ideálně neproprietárních formátů. Například převod PDF dokumentů do standardizovaného PDF/A (dle ISO norem 19005-1:2005 a ISO 19005-2:2011).

To vše je nutné mít pod opakovanou kontrolou a nejlépe pravidelnou obnovou či úpravou definovaných politik, strategií a procedur dle potřeb [ALA, 2007]. Těmto strategiím a politikám říkáme **ochranné (preservační) aktivity** (*Preservation Actions*). Ty vycházejí z principu plánování dlouhodobé ochrany (*Digital Preservation Planning*) a její konkrétní implementace ve formě plánu dlouhodobé ochrany (*Digital Preservation Plan*) digitálních dat/objektů.

2. Digitální repozitáře, knihovny a archivy z pohledu informační architektury

Informační architektura je pro digitální repozitáře, knihovny a archivy velmi důležitým oborem, jelikož sdílí stejný cíl – zaměřují se na vyhledatelnost informací, resp. digitálních objektů. Koncepty a principy informační architektury potřebují částečně odlišný přístup v aplikaci, jelikož digitální knihovny disponují nejen uživatelským rozhraním, ale i sbírkami, které mají efektivně uspořádat (i proto, že obsah většinou není generovaný interně), zpřístupňovat a uchovávat [HOWLETT, 2011].

Pojem informační architektura (*Information Architecture - IA*) byl poprvé použit Richardem Saul Wurmanem v roce 1976 na konferenci *The Architecture of Information*. Trvalo více než 20 let než se IA ujala jako vědní disciplína. IA můžeme definovat jako vědeckou disciplínu zabývající se organizací a rozvržením (*Layout*) obsahu systému, rozhraní či webu. Ale také jako umění organizace a modelování informací způsobem, který úspěšně naplňuje potřeby uživatelů obzvláště v komplexních systémech. Jedná se ku příkladu o knihovní systémy, systémy pro správu obsahu, kde je nevyhnutně vhodně podchytit vývoj databází a webu, programování či technické psaní. Institut informační architektury¹⁸ definuje informační architekturu pomocí tří obecných prohlášení [INFORMATION, cca2007]:

1. Strukturální design sdílených informačních prostředí.
2. Umění a věda organizace a popisu webových sídel, intranetů, online komunit a softwaru za účelem následného vyhledání a využití.
3. Vznikající zájmová skupina přenášející principy designu a architektury do digitální sféry.

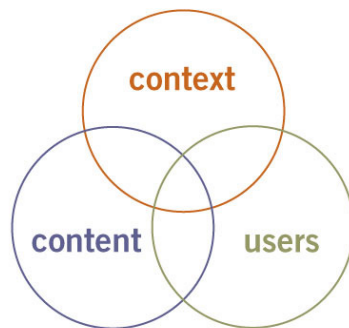
Informační architekturu tvoří čtyři základní prvky - **organizační systém**, **systém popisu** (*Labeling System*), **navigační systém** a **vyhledávací systém**. IA se tedy zaměřuje na informační, nikoli grafický design.

1. **Organizační systém** - informace jsou organizovány za účelem jejich lepšího řízení, vysvětlení, pochopení a v neposlední řadě i využití. Proto základem by měla být náležitá klasifikace. Cílem je podpora navigace, vyhledávání a popis, který bude vyhovovat právě uživatelům systému/ rozhraní.
2. **Systém popisu** - popisný systém sloužící jako reprezentace informací. Úkolem je popis informací sémanticky jednoznačným (a tím pádem i efektivním) způsobem pro daného koncového uživatele či uživatelské komunity.

¹⁸ Institut informační architektury (*The Information Architecture Institute*) představuje organizaci sdružující profesionály. Má za cíl šířit povědomí o problematice, povolání a vzdělávání v oblasti informační architektury. Podrobněji přímo na stránkách Institutu - <http://www.iainstitute.org/en/about/>.

3. **Navigační systém** - jedná se o prvky, které usnadňují navigaci v systému/rozhraní podle předem stanovených kroků. Typickými nástroji jsou globální navigační systémy, lokální navigační systémy, mapy a indexy systému/rozhraní, průvodci či doplňková navigační schémata.
4. **Vyhledávací systém** - patří zde nástroje umožňující formulaci uživatelského požadavku - vyhledávací rozhraní, vyhledávací algoritmus, dotazovací jazyk se svou syntaxí a sémantikou, zobrazené výsledky vyhledávání atd. [ROSENFELD, 2002; MAKULOVÁ, 2006].

Výše uvedené prvky však nevznikají ve vakuu, je nutné brát v potaz uživatele, obsah a kontext (viz Obr. 1), kterých vzájemný průnik představuje základ designu podle IA [MAKULOVÁ, 2006]. IA by měla být vždy navrhována dle potřeb cílové skupiny – uživatelů, měli bychom znát jejich preference, informační potřeby, informační chování atd. Aby uživatelé snáze dokázali definovat své informační potřeby, informační požadavky a rychleji dohledat relevantní/pertinentní informace a znalosti v čase a místě, které potřebují, fungují v digitálních knihovných systémech pořádní znalostí (viz dále v textu, kapitola 2.1 Pořádání znalostí). **Obsah** představuje nejen samotná data, nýbrž i jejich formáty, strukturu, metadata apod. Marchionini [2007] v této souvislosti hovoří o **balíčku** (*Package*), který tvoří základní jednotku informace z pohledu IA. Může jít o text, audio, video, přičemž IA plánuje, co dané balíčky budou obsahovat, jak s nimi bude nakládáno, včetně toho jak a komu budou v informačním systému zpřístupňovány. **Kontext** - zde zastupuje konkrétní prostředí rozhraní či informační systém (digitálního repozitáře knihovny, archivu apod.), které je nutné se všemi jeho specifiky dostatečně podchytit (viz dále v textu). Všechny tyto prvky a pilíře IA dokáže informační architekt zachytit a jasně definovat.



Obr. 1 Základní pilíře designu podle informační architektury
(<http://semanticstudios.com/publications/semantics/images/threecirclesbig.jpg>)

Obrázek popisuje vzájemnou propojenost tří základních konceptů, resp. pilířů designu podle informační architektury. Jedná se o obsah (*Content*) v určitém kontextu (*Context*) pro danou skupinu uživatelů (*Users*).

Informačním architektem je tedy osoba, která organizuje data, transformuje je z komplexní do přehledné podoby, vytváří strukturu nebo mapu informací umožňující snadnou orientaci a cestu k poznatkům. Jde o profesi 21. století s úkolem řešit problémy organizování informací tak, aby byly snadno pochopitelné a srozumitelné [MAKULOVÁ, 2006].

Systémy pro organizaci znalostí mají svou dlouhou historii v klasickém knihovním prostředí, jsou běžnou součástí knihovních systémů a mají své neméně důležité postavení i v digitálních repozitářích. Jejich existence je důležitá pro uživatele, kteří díky nim dokážou snáze definovat své informační potřeby, informační požadavky a rychleji dohledat relevantní/pertinentní informace a znalosti v čase a místě, které potřebují.

Z funkčního hlediska mezi hlavními cíli, které digitální repozitáře, knihovny a archivy v oblasti organizace znalostí naplňují, patří:

- Zlepšení efektivity vyhledávacích mechanismů
- Poskytování jednotného přístupu k materiálům v různých formátech (např. přístup k netextovým dokumentům)
- Poskytování podpory koncovým uživatelům, kteří přistupují k dokumentům - to především znamená pomoc uživatelům při definování informačních potřeb v procesu vyhledávání informací (např. kontextuálními nápovědami apod.) a podporu preciznějšího vyhledávání a zpracovávání již během vyhledávání nebo bezprostředně

po dokončení vyhledávacího procesu.

- Podpora informačního vyhledávání jako integrální součást intelektuální činnosti, řešení problémů, učení se. Důležité je umožnit uživatelům možnost zpětné vazby či jinou formu účasti a vytvořit propojení mezi uživateli.
- Podpora kolaborativní práce - vědecká komunikace jako počítačově podporovaná skupinová konverzace (výměna zkušeností a poznatků) patří mezi nejnovější výzvy. Uvědomělým výběrem KOS (*Knowledge Organization System*) zabezpečíme naplnění výše zmíněných cílů a výzev.

Abychom snáze pochopili organizační hledisko KOS, je vhodné se obeznámit s OAIS referenčním modelem (ISO 14721:2003) – viz kapitola 3.3. Tento model nabízí kompletní funkční a informační popis digitálního repozitáře.

2.1 Pořádání znalostí

Pořádání či organizace znalostí (*Knowledge Organization*) je oblast lidského poznání, která se zabývá uspořádáním jednotek znalostí a objektů všech typů. Ty by měly odpovídat pojmům (teoriím) nebo pojmovým třídám tak, aby byly zachyceny znalosti o poznaném světě v uspořádané podobě spolu s umožněním rozšiřování těchto znalostí za účelem jejich využití [JONÁK, 1999].

Kde ale leží hranice mezi pořádáním a řízením znalostí (*Knowledge Management*) - pojmem který je v současné době hojně (ne-li nadbytečně) užíván? S pojmem řízení znalostí se nejčastěji setkáváme v obchodních kruzích a bývá definován jako systém pro zachycení/získávání, využití a zhodnocení znalostí organizace [KASTEN, 2007]. Řízení znalostí, někdy také i jako management znalostí je hybridní disciplínou; nejde ani čistě o vědu, ani o určitou dovednost. Z hlediska funkce může propojovat oblasti učení a rozvoje organizace, řízení lidských zdrojů, informačních systémů a technologií [COLLISON, 2005]. Na to, aby mohlo být považováno za úspěšné, musí svést dohromady vhodné organizační i technologické procesy a nástroje, které jsou přetaveny do strategických cílů organizace.

Jsou přinejmenším 3 oblasti, ve kterých výzkum pořádání znalostí čelí podobným problémům jako vývoj/design systémů řízení znalostí [KASTEN, 2007]:

- vytváření užitečných struktur znalostí před vlastním sběrem znalostí
- klasifikace (dynamické ne nestrukturované) organizační znalosti; klasifikační schémata musí být navržena tak, aby byla lehce přizpůsobitelná a vyhledatelná nejružnějšími jazykovými přístupovými metodami

- kodifikace tacitní (tj. implicitní, nevyslovené) znalosti (narativní a textovou analýzou).

2.1.1 Implementace systémů pro řízení znalostí do digitálních repozitářů

Prvním krokem, který předchází implementaci jakéhokoliv systému, je analýza současného stavu, organizačního rámce a uživatelských potřeb. Zhodnocení stávající situace dává předpoklad k dobrému rozhodování se pro konkrétní systém(y). V popředí takového rozhodování však musí být znalost prostředí a potřeb uživatelů, nikoliv pouze jejich funkční požadavky a technologická náročnost.

Organizační rámec může být zaměřen pouze na jednu skupinu uživatelů a tím pádem i oborově specifické KOS (*Knowledge Organization System*)¹⁹. Široký záběr digitální knihovny (knihovna o mnoha odděleních, několika online komunitách s přesně vymezeným účelem) podmiňuje vyšší návratnosti investic (*ROI - Return Of Investments*) a sémantickou interoperabilitu. Je však nutné dbát na to, aby existoval vyvážený poměr mezi oborově specifickými a obecnými KOS.

V druhé fázi je nutné zvážit, jestli potřeby uživatelů můžou být naplněny některým ze stávajících dostupných (komerčních či open source) systémů nebo zda-li systémy pořádání znalostí budeme sami vyvíjet. Je nevyhnutné si uvědomit, že hodnota takovýchto systémů je postavena i na tzv. „prověření časem“ a jejich (ne)akceptovatelností v příslušné uživatelské komunitě. Neméně důležitým aspektem v procesu volby je i samotný charakter KOS (offline, online, outsourcing - služba atd.). Třetí krok spočívá v plánování infrastruktury. V tomto stadiu by se o KOS mělo rozhodovat v kontextu digitálních repozitářů. Fyzická lokace KOS může nabývat jedné z dvou podob - interního nástroje anebo externí služby. Oba přístupy mají své výhody i nevýhody.

Pokud je KOS dostupný na webu, je vhodné zvážit odkazování na externí službu. Jelikož v tomto případě nemáme úplnou kontrolu nad systémem, výhody tohoto přístupu nejsou jednoznačné. Problém nastává zejména v případě změny organizace a obsahu či persistence odkazu. Tradičním identifikátorem ve webovém prostředí je hyperlink URL (*Uniform Resource Locator*). Nicméně mnoho odkazů se časem přeruší nebo přestane platit, čímž znepřístupňují původní zdroj. Řešením (kromě organizačních opatření) je užití systému persistentních identifikátorů (např. Handle, URI, URN, DOI, OAI, PURL, NBN). U lokálních systémů (implementovaných jako interní nástroje) mají digitální repozitáře mnohem větší kontrolu nad KOS. Avšak v tomto případě bude úspěšný provoz závislý na dostatečném technologickém

¹⁹ Např. v případě lékařských digitálních knihoven nejčastěji postačí KOS typu Medical Subject Headings.

vybavení a organizační připravenosti.

Implementace systémů pořádání znalostí do digitálních repozitářů může nabývat dvou základních podob:

1. KOS jako skrytá přidaná hodnota pro uživatele, která je součástí samotného vyhledávacího mechanismu;
2. KOS jako interaktivní nástroj pro uživatele, který vede uživatele k definování relevantního vyhledávacího dotazu.

Při implementaci většího počtu KOS je nutné se zamyslet nad škálovatelností celé infrastruktury [HODGE, 2000; SOERGEL, 2009].

3. Modely digitálních knihoven

Pochopit komplexnost a záběr problematiky digitálních repozitářů, knihoven a archivů není snadné. Situaci však ulehčují nejrůznější typy modelů, od těch referenčních, přes objektové a procesní modely až po ty funkční. Na obecné úrovni se jedná o tzv. **konceptuální modely** (*Conceptual Model*), jež představují mentální model, zachycující koncepty (entity) a vztahy mezi nimi s cílem snadnějšího pochopení problému, resp. oblasti, kterou reprezentují. Typickým představitelem modelu digitálních knihoven tohoto typu je model Kahn-Wilenskeho architektury.

Referenční modely (*Reference Model*) digitálních knihoven, repozitářů a archivů zachycují vždy budoucí, chtěný stav a představují abstraktní model reálného otevřeného systému. Příkladem je referenční model Otevřeného archivního informačního systému (*OAIS - Open Archival Information System*) a DELOS Referenční model digitálních knihoven. **Funkční model** se primárně zabývá způsoby interakce se systémem a jakým způsobem funguje [REITZ, 2004, s. 274]. **Procesní model** (*Process Model*) modeluje aktivity reprezentující procesy za účelem jejich zlepšení, analýzy a lepšího pochopení. Nejčastěji mají popisný a vysvětlující charakter. Při jeho sestavování se obvykle začíná zhodnocením dosavadního fungování procesů. Poté se za účelem zvýšení efektivity hledají možná zlepšení, na základě kterých se stanoví nová pravidla a směrnice. Například dřívější přístupy v oblasti digitálních knihoven a dlouhodobé ochrany dat se zaměřovaly zejména na technologické aspekty, zatímco důležité otázky organizační, personální, finanční a další zůstávaly nedotčené. Až po určité době se bral v potaz i celkový kontext, ve kterém vznikají digitální objekty a existují digitální knihovny [REITZ, 2004, s. 542].

Obecně možno říct, že modely digitálních repozitářů, knihoven a archivů nejčastěji vnímají digitální knihovny jako spojení repozitáře (*Repository*), který uchovává obsah digitální knihovny formou strukturovaných objektů, včetně jejich popisu a různých verzí a dále katalogu nebo databáze, která uchovává popis obsahu repozitáře a management či administraci, jež zabezpečuje chod repozitáře.

Mezi nejnámější modely digitálních knihoven (pořadí dle jejich důležitosti a míry implementace v praxi) patří Kahn-Wilenskeho architektura (*Framework for Distributed Digital Object - 1997*), DELOS Referenční model digitálních knihoven (*The DELOS Reference Model of Digital Libraries - 2007*) a Referenční model Otevřeného archivního informačního systému (*Open Archival Information System Reference Model - 2003*).

Kahn-Wilenskeho architektura staví do popředí zájmu digitálních knihoven digitální objekt (*Digital Object*) a za základní prvky systému digitální knihovny považuje uživatelské rozhraní (*User interface*), repozitář (*Repository*), vyhledávací systém (*Search System*) a identifikační systém (*Handle System*).

Refereční model DELOS vidí digitální knihovnu jako systém pozůstávající ze tří základních částí: Digitální knihovna (*Digital Library – DL*), Systém digitální knihovny (*Digital Library System - DLS*) a Systém správy digitální knihovny (*Digital Library Management System - DLMS*).

OAIS model, jakožto představitel nejkompexnějšího modelu, definuje a objasňuje klíčové komponenty OAIS systému: prostředí, informace, informační balíčky a podrobně modeluje funkční entity. Funkční model rozděluje architekturu „ideálního“ repozitáře do pěti hlavních modulů a vztahů mezi nimi: přístup (*Access*), administrace (*Administration*), permanentní úložiště (*Archival Storage*), správa dat (*Data Management*), plánování dlouhodobé ochrany (*Preservation Planning*).

3.1 Kahn-Wilenského architektura

Kahn-Wilenskeho architektura (*Kahn-Wilensky Framework for Distributed Digital Object Services*²⁰) představuje jeden z prvních, ale velmi propracovaných modelů digitální knihoven (z roku 1995), který uvádí základní komponenty, klíčové oblasti a v neposlední řadě upozorňuje i na okolí (ekonomika, duševní vlastnictví) digitálních knihoven. V roce 1997 byl experimentálně realizován v Kongresové knihovně v rámci projektu pilotního systému Národní digitální knihovny (*National Digital Library Project - NLDP*)²¹.

Základní prvky podle Kahn-Wilenskeho modelu tvoří [BARTOŠEK, 2006; BRATKOVÁ, 2009], viz Obr. 2:

- **Uživatelské rozhraní** (*User Interface*) - uživatel specifikuje svůj požadavek přes inteligentní koncové uživatelské rozhraní, které jej přeformuluje na dotaz pro globální vyhledávací systém a vrátí seznam informačních zdrojů zodpovídající požadavku uživatele.
- **Repozitář** (*Repository*) - digitální objekty jsou uloženy v repozitáři a přístup k nim umožňují ostatní nadstavbové prvky (Uživatelské rozhraní, Vyhledávací systém a Handle systém). Pro každý digitální objekt udržuje dva typy metadatových záznamů. Prvním z nich je záznam vlastností objektů (*Properties Records*), který definuje

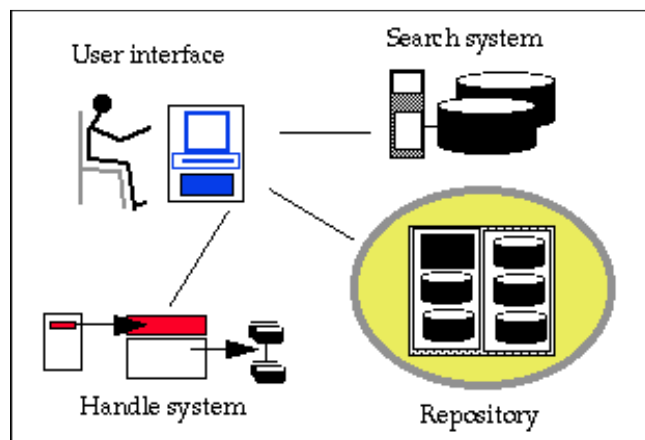
²⁰ KAHN, Robert and WILENSKY, Robert. "A Framework for Distributed Digital Object Services". International Journal on Digital Libraries, Springer, Volume 6, Number 2, April 2006. [doi:10.1007/s00799-005-0128-x]. Dostupný na:<http://www.doi.org/topics/2006_05_02_Kahn_Framework.pdf>.

²¹ Pilotní systém obsahoval Handle systém pro identifikaci digitálních objektů, repozitář pro jejich uchování a dvě uživatelská rozhraní: jedno navrženo speciálně pro personál spravující digitální objekty v repozitáři a druhý pro uživatele, kteří chtějí k daným objektům přistupovat. Repozitář NLDP již v té době obsahoval rozsáhlé sbírky digitalizovaných textů a fotografií. Podrobnější informace lze najít na: <<http://www.dlib.org/dlib/february97/cnri/02arms1.html>>.

podmínky přístupu k objektu a autorská práva, tzv. přístupová a autorsko-právní metadata (*Administrative And Copyright Metadata*), technické vlastnosti typu formát, tzv. technická metadata (*Technical Metadata*) a bibliografické údaje, tzv. popisná metadata (*Descriptive Metadata*). Druhým je transakční záznam (*Transaction Log*) zaznamenávající veškeré akce a aktivity, resp. transakce repozitáře, které se týkají konkrétního digitálního objektu. Repozitář komunikuje s okolím (zpřístupňuje a ukládá digitální objekty) pomocí repozitářového přístupového protokolu (*RAP - Repository Access Protocol*).

- **Vyhledávací systém** (*Search System*) - určen pro vyhledávání digitálních objektů v repozitáři, resp. v digitální knihovně jako celku. Systém je tvořen indexy (s popisnými metadaty a/nebo plných textů). Z pohledu informační architektury představuje organizační systém. Jeho součástí mohou být nejrůznější typy systémů pořádání znalostí (viz kapitola 2 Digitální repozitáře, knihovny a archivy z pohledu informační architektury).
- **Identifikační systém** (*Handle System*) - představuje resoluční, resp. směřovací systém a mechanismus, který je schopen určit podle jednoznačného identifikátoru - handle, kde se digitální objekt nachází. V případě, že funguje nad několika repozitáři, je schopen zjistit i vícenásobné výskyty. Celý „Handle systém“²² pak velmi zjednodušeně funguje následujícím způsobem: uživatel přes uživatelské rozhraní (*User Interface*) pomocí vyhledávacího mechanismu (*Search System*) vyhledá, kde se jím požadovaný digitální objekt nachází. Poté si vybere konkrétní instanci objektu, který je zaslán do globálního resolučního mechanismu a vyžádá si z konkrétního repozitáře daný dokument pro zobrazení uživateli.

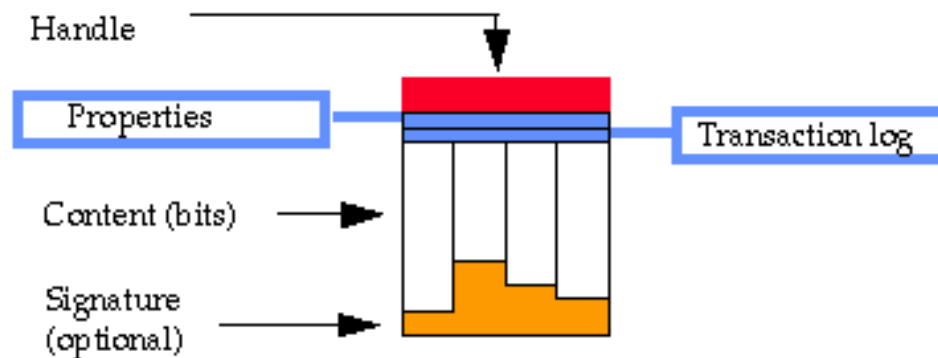
²² Handle systém (*Handles System*) byl úspěšně implementován v *Corporation for National Research Initiatives* - CNRI, ve spolupráci s DARPA (*Defense Advance Research Projects Agency*). Mezinárodní nadace DOI (*The International DOI Foundation*) implementace systému Handle v podobě DOI má v současné době (prosinec 2013) více než 89 mil. registrovaných „handles“. „Handle systém“ je využíván i na Univerzitě Karlově v Praze v rámci Digitálního univerzitního repozitáře (<http://repozitar.cuni.cz/>).



Obr. 2 Hlavní komponenty digitální knihovny podle Kahn-Wilensky
 (<http://knihovna.nkp.cz/NKKR0404/images/0404237.gif>)

Obrázek popisuje základní komponenty digitální knihovny podle rámce sestaveného Robertem Kahnem a Robertem Wilenskym - Uživatelské rozhraní (*User Interface*), Vyhledávací systém (*Search System*), Handle systém (*Handle System*) a Repozitář (*Repository*). Základ zde tvoří repozitář s digitálními objekty, které jsou opatřeny jedinečnými identifikátory handle pro jejich snadné vyhledání (pomocí vyhledávacího mechanismu nad repozitářem) a zpřístupnění uživateli přes uživatelské rozhraní. Slouží tedy jako tzv. "vstupní brána" pro zadávání uživatelských dotazů a výstupní bod pro zobrazení dohledaných resp. uchovaných digitálních objektů.

Základní stavební prvek v tomto modelu architektury informačního systému tvoří digitální objekt jako samostatně použitelná informační jednotka. Představuje specifickou strukturu dat, jejíž hlavní součástí jsou digitální materiály (taktéž označovány i jako data), které kromě svého inherenčního obsahu mají i globální jednoznačný identifikátor, tzv. „handle“ (viz Obr.3).



Obr. 3 Digitální objekt jako základní prvek architektury digitální knihovny v modelu Kahn-Wilensky (<http://knihovna.nkp.cz/NKKR0404/images/0404236a.gif>)

Obrázek popisuje schéma části digitálního objektu, který je součástí základní stavební jednotky architektury Kahn-Wilensky. Je definován obsahem ve formě bitů (*Content - Bits*), jež určují jeho základní vlastnosti (*Properties*) a dále transakčním logem (*Transaction Log*), který zaznamenává veškeré provedené akce a aktivity nad digitálním objektem.

Ten v sobě nese vlastnosti (*Properties*), obsah samotný (*Content*) a transakční log (soupis veškerých operací, které proběhly digitálním objektem). Digitální objekty mohou být proměnlivé, pokud je možné jejich obsah upravovat, nebo neproměnlivé. Mohou existovat v jednoduché formě nebo jako skupina (agregát) několika digitálních objektů [BRATKOVÁ, 2009].

Seskupený objekt tak může vypadat jako složený objekt²³ (*Composite Object*), jejichž obsah zahrnuje dva a více objektů, nebo jako meta-objekty²⁴ (*Meta-objects*), jejichž obsah zahrnuje identifikátory sady objektů. Takovéto objekty již musí obsahovat i tzv. strukturální metadata (*Structure Metadata*), která definují vztahy mezi jednotlivými objekty.

Architektura digitální knihovny podle Kahn-Wilensky se zaměřuje spíše na jednoznačnou identifikaci digitálních objektů, neřeší ještě problematiku dlouhodobé ochrany. Otázky

²³ Seskupený objekt (*Composite Object*) může být ku příkladu elektronická verze vysokoškolské kvalifikační práce, jež tvoří jeden digitální objekt. V objektu je zastoupen samotný text obhajované práce (např. PDF), oponentský posudek (např. PDF, DOC), školitelský posudek (např. PDF, DOC) a naskenovaný protokol průběhu obhajoby (např. JPG nebo PDF).

²⁴ Meta-objekt (*Meta-object*) se vyskytuje např. v případech, kdy se skenují mapové listy, přičemž se jeden mapový list skenuje do vícero formátů a/nebo v různých rozlišeních - jako náhled (*Thumbnail*), archivní kopie (*Master File*), se kterou se nebude aktivně pracovat a uživatelská kopie (*User-copy*), která naopak bude zpřístupňována koncovým uživatelům. Náhled může být v JPG, archivní kopie v nekomprimovaném TIFu a uživatelská kopie jako JP2. Pro popis obrazových souborů a jejich digitálních verzí je sestaven meta-objekt, který obsahuje metadata, jež jsou společná pro všechny objekty.

bezpečného uchovávání a dlouhodobé ochrany se začaly objevovat o něco později a našly své místo v Referenčním modelu Otevřeného archivního informačního systému (viz kapitola 3.3 Open Archival Information System Reference Model - OAIS). Jakýsi pomyslný mezistupeň obecného modelu zastupuje DELOS Referenční model digitálních knihoven. Ten reprezentuje jednotný teoretický rámec životního cyklu digitálních informací a nabízí integrovanou správu obsahu díky definici obecných technologií pro digitální knihovny.

3.2 DELOS Digital Library Reference Model

Projekty skupiny DELOS Network of Excellence byly financovány Evropskou komisí v rámci 6. rámcového programu a její výzkumná činnost trvala do roku konce 2007. Metou byla integrace a koordinace probíhající výzkumné činnosti největších evropských týmů v oblasti digitálních knihoven, jako i šíření povědomí o implementaci technologií digitálních knihoven do mnoha různých aplikačních domén. Hlavním cílem bylo definovat jednotný, ucelený teoretický rámec životního cyklu digitálních informací a nabídnout interoperabilní multimodální/multilingvální služby pro integrovanou správu obsahu vyvinutím obecné technologie pro digitální knihovny.

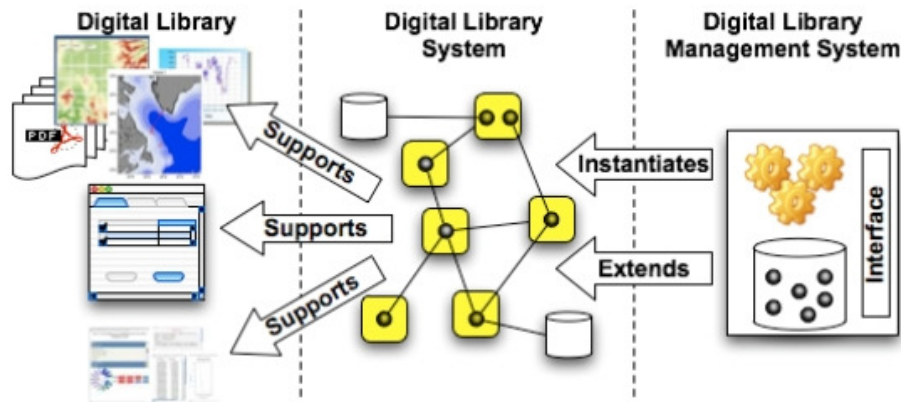
Referenční model digitální knihovny DELOS (*The Delos Reference Model for Digital Libraries – Foundation for Digital Libraries*) představuje jeden z prvních komplexních konceptuálních rámců digitální knihovny. Je tedy jakýmsi základním kamenem pro lepší návrh a tvorbu vhodných systémů digitálních knihoven. Sestaven byl týmem výzkumníků a expertů z 57 institucí - knihoven, muzeí, archivů a univerzit (za Českou republiku to byla Masarykova univerzita v Brně) sdružených v DELOS Network of Excellence.

Z pohledu struktury dokumentu pozůstává Referenční model digitální knihovny ze třech stěžejních částí:

- **Manifest digitální knihovny** (*Digital Library Manifesto*)
- **Referenční model digitální knihovny v kostce** (*Digital Library Reference Model in a Nutshell*)
- **Koncepty & vztahy Referenčního modelu digitální knihovny** (*Digital Library Reference Model Concepts & Relations*).

3.2.1 Třívrstvá konstrukce univerza digitální knihovny

Digitální knihovna představuje vyvíjející se organizaci, která vzniká během řady vývojových kroků, které společně vytvářejí všechny nezbytné prvky. Součástí tohoto procesu jsou právě tyto tři „systémy“ vyvíjející se za pochodu (viz Obr. 4) **Digitální knihovna** (Digital Library – DL), **Systém digitální knihovny** (Digital Library System - DLS) a **Systém správy digitální knihovny** (Digital Library Management System - DLMS)²⁵ [CANDELA, 2007]:



Obr. 4 Konceptuální model digitální knihovny jako třívrstvá konstrukce [CANDELA, 2007a]

Obrázek popisuje digitální knihovnu jako organizaci tvořenou třemi klíčovými systémy. Základní pilíř tvoří Systém správy digitální knihovny (*Digital Library Management System*) se svým specifickým rozhraním (*Interface*), díky kterému funguje (*Instantiates*) a je rozšířen softwarem Systému digitální knihovny (*Digital Library System*), který podporuje (*Supports*) Digitální knihovnu jako celek.

- **Digitální knihovna (DL)** představuje virtuální organizaci, která shromažďuje, spravuje a dlouhodobě uchovává bohatý digitální obsah a svým uživatelským komunitám nabízí pro daný obsah specializované funkce měřitelné kvality podle stanovených pravidel a politik.
- **Systém digitální knihovny (DLS)** reprezentuje software, který podle definované (případně distribuované) architektury nabízí funkce, které daná Digitální knihovna požaduje. Uživatelé interagují s Digitální knihovnou jako organizací právě prostřednictvím Systému digitální knihovny.

²⁵ DELOS Digital Library Reference Model označuje digitální knihovny jako organizační celek/univerzum termínem „Digital Libraries“, které pro potřeby této kapitoly bude použit český ekvivalent „digitální knihovny“. Prvek „Digital Library“ v rámci třívrstvé konstrukce univerza digitálních knihoven je zaznamenán jako „Digitální knihovna“.

- **Systém správy digitální knihovny** (DLMS) zastupuje obecný softwarový systém, který poskytuje vhodnou softwarovou infrastrukturu pro sestavení a administraci Systému digitální knihovny včetně vhodných prvků/funkcí v Digitální knihovně. Integruje doplňkový software, který umožňuje sofistikované a pokročilé funkce.

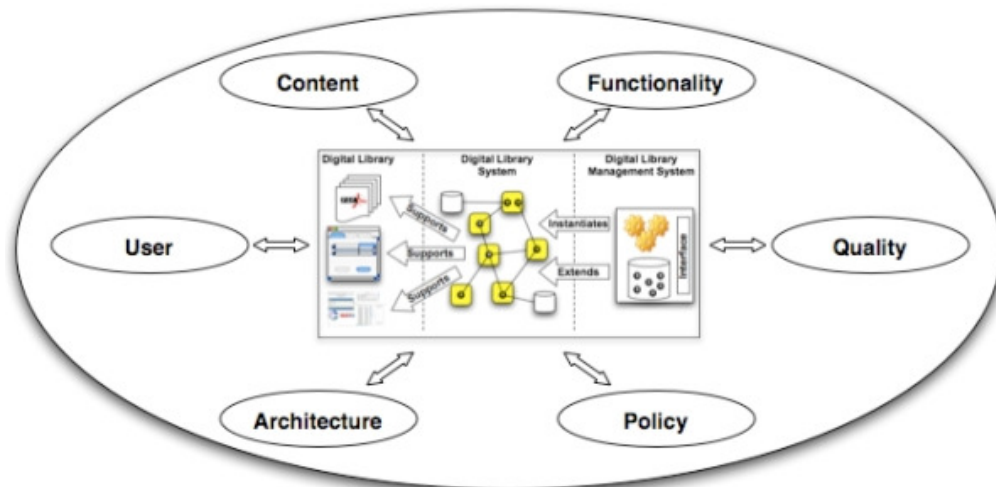
3.2.1.1 Klíčové koncepty konstrukce

Konkrétní naplnění a praktickou realizaci této třívrstvé konstrukce ovlivňuje šest klíčových konceptů (viz Obr. 5) - **Obsah** (*Content*), **Uživatel** (*User*), **Funkce** (*Functionality*), **Kvalita** (*Quality*), **Politiky** (*Policy*) a **Architektura** (*Architecture*) [Candela, 2007a; CANDELA, 2007b]:

- **Obsah** zahrnuje data i informace, které DL spravuje a zpřístupňuje uživatelům. Tyto dále tvoří digitální objekty²⁶ organizované do sbírek.
- **Uživatel** zastupuje v této souvislosti nejrůznější aktéry (actors), lidské i strojové, kteří interagují s digitálními knihovnami.
- **Funkce** představují služby (zejména registrace, vyhledávání, prohledávání a zobrazení), které Digitální knihovna nabízí uživatelům.
- **Kvalita** zosobňuje parametry, které je možné použít pro popis a hodnocení obsahu a chování Digitální knihovny. Kvalitu některých kritérií je možné měřit automaticky, u jiných je nevyhnutná zpětná vazba - vyhodnocení od uživatelů.
- **Politiky**²⁷ definují podmínky, pravidla, omezení a regulace pramenící z interakce uživatelů s Digitální knihovnou.
- **Architekturu** tvoří Systém správy digitální knihovny a mapuje funkce a obsah Digitální knihovny na příslušný software a hardware.

²⁶ V kontextu tohoto modelu představují digitální objekty veškerý obsah, který digitální knihovna jako organizace přijímá, spravuje, uchovává a zpřístupňuje uživatelům. Jedná se o nejrůznější informační objekty, včetně jejich popisných (syntaktických, sémantických, konceptuálních informací) a dalších metadat.

²⁷ Jedná se ku příkladu o pravidla akceptovatelného chování uživatelů, management digitálních práv, podmínky utajení, ochrany soukromí, poplatky za služby apod.



Obr. 5 Hlavní koncepty univerza digitální knihovny
 (<http://www.dlib.org/dlib/march07/castelli/castelli-fig2.jpg>)

Obrázek znázorňuje okolní prostředí digitální knihovny (třívrstvé konstrukce - viz Obr. 4), které ovlivňuje finální podobu, konkrétní naplnění a praktickou realizaci digitální knihovny. Jde o šest klíčových prvků: obsah (*Content*), funkce (*Functionality*), kvalitu (*Quality*), politiky (*Policy*), architekturu (*Architecture*) a v neposlední řadě uživatele (*User*).

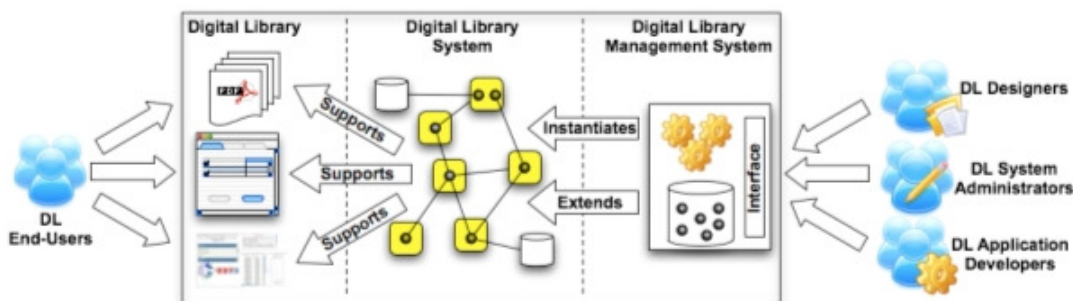
3.2.2 Aktéři digitální knihovny

V tomto komplexním univerzu digitální knihovny interagují čtyři různé skupiny aktérů (*Actors*) (viz Obr. 6): **Koncoví uživatelé** (*DL End Users*), **Designéři** (*DL Designers*), **Systémoví administrátoři** (*DL System Administrators*) a **Vývojáři aplikací** (*DL Application Developers*):

- **Koncové uživatele** je možné zařadit do kategorie Tvůrce obsahu, Konzumenty obsahu nebo knihovníky. Ti všichni těží z příslušných funkcí Digitální knihovny.
- **Designéři** definují, „customizují“ a udržují Digitální knihovnu tak, aby dostala funkčních a informačních potřeb Koncových uživatelů. Jejich úkolem je poskytovat Systému správy digitální knihovny konfiguraci a parametry zjištěných potřeb (např. dotazovací jazyk, datové modely, formáty, ontologie, klasifikační schémy atd.).
- **Systémoví administrátoři** vybírají softwarové komponenty pro Systém digitální knihovny. Výběr je ovlivněn požadavky Koncových uživatelů i Designérů. Administrátoři pracují se

Systémem správy digitální knihovny tak, aby konfigurace zabezpečila nejvyšší možný stupeň kvality poskytovaných služeb.

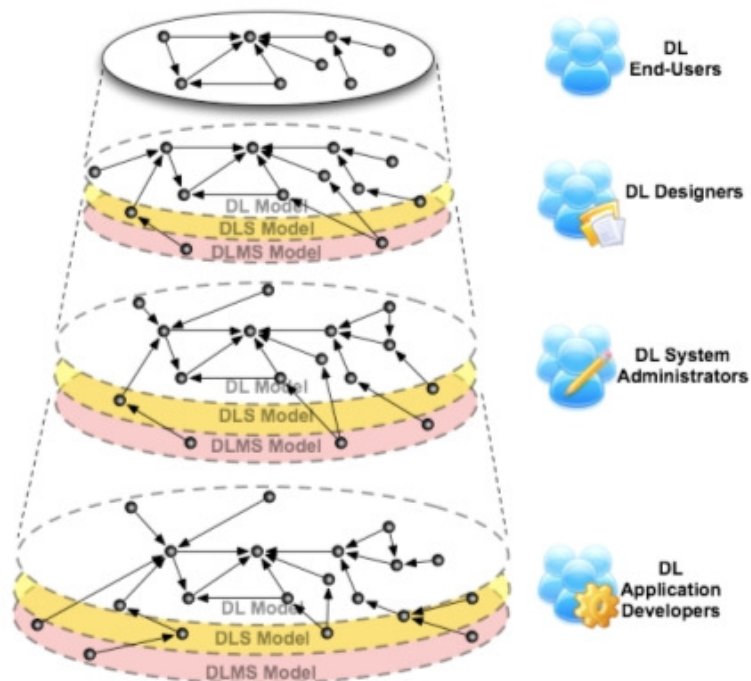
- **Vývojáři aplikací** vyvíjí softwarové komponenty, které budou tvořit potřebné složky Systému digitální knihovny.



Obr. 6 Hlavní role aktérů ve třívrstvé konstrukci univerza digitální knihovny
(<http://jakoblog.de/wp-content/uploads/2008/09/dl-dls-dlms.png>)

Obrázek ukazuje, jakou roli hrají jednotliví aktéři, jež zásadním způsobem ovlivňují fungování digitální knihovny. Systém správy digitální knihovny formují designéři (*DL Designers*), Systémoví administrátoři (*DL System Administrators*) a Vývojáři aplikací (*DL Application Developers*). Digitální knihovnu jako celek, resp. organizaci nejzásadnějším způsobem ovlivňují Koncoví uživatelé (*DL End-Users*), kteří v kontextu tohoto modelu zosobňují tvůrce obsahu, knihovníky a konzumenty.

Všichni výše uvedení aktéři v univerzu digitálních knihoven jsou vzájemně (hierarchicky) propojení (viz Obr. 7). Koncoví uživatelé pracují v Digitální knihovně, zatímco Designéři, Systémoví administrátoři a Vývojáři aplikací využívají Systém digitální knihovny (zprostředkovaně přes Systém správy digitální knihovny) a následně i Digitální knihovnu. Spolupracující aktéři sdílejí společný slovník a znalosti. Například Koncový uživatel definuje požadavky z hlediska modelu digitální knihovny a Designéři, kteří těmto požadavkům rozumí je vhodně zdefinují.



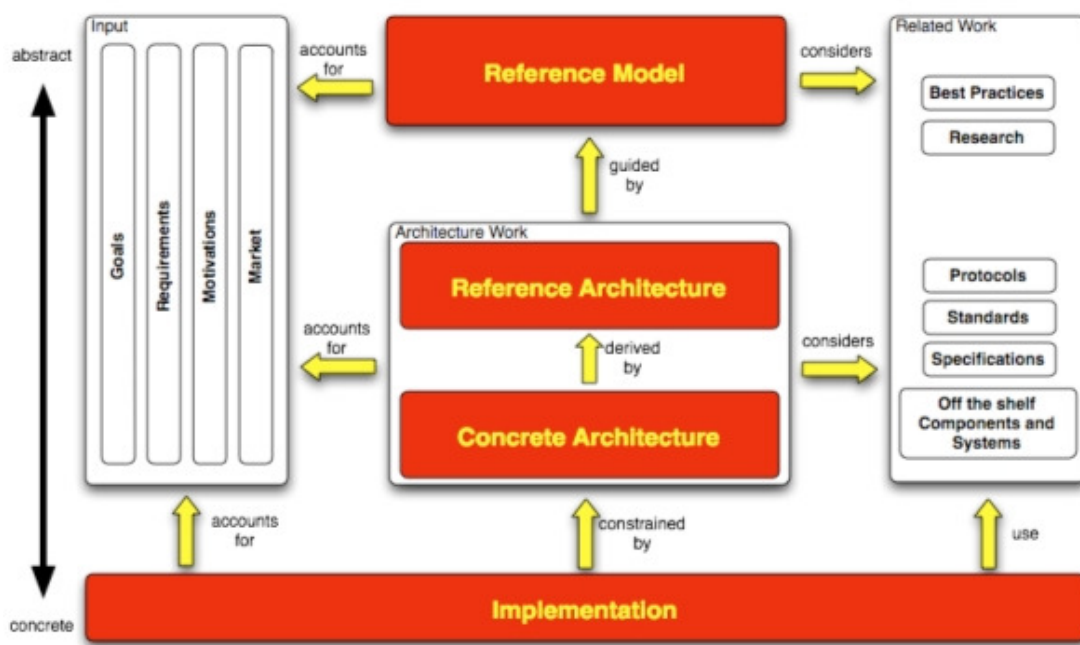
Obr. 7 Hierarchie pohledů uživatelů
 (<http://www.dlib.org/dlib/march07/castelli/castelli-fig4-rev.jpg>)

Obrázek ukazuje čtyři výše popsané typy aktérů (hierarchicky uspořádané), kteří interagují v digitálních knihovnách. Koncoví uživatelé (*DL End-Users*) působí na Digitální knihovny (*DL Model* - bílý ovál, nejvýše položený na obrázku), zatímco Designéři (*DL Designers*), Systémoví administrátoři (*DL System Administrators*) a Vývojáři (*DL Designers*) ovlivňují Systém digitální knihovny (*DLS Model* - žlutý ovál na obrázku) prostřednictvím Systému správy digitální knihovny (*DLMS Model* - růžový ovál na obrázku), a tím pádem nepřímo i samotnou Digitální knihovnu. Takto definované vazby dávají jistotu, že spolupracující aktéři sdílejí společnou slovní zásobu a znalosti. Například, Koncoví uživatelé vyjadřují své připomínky a požadavky na Digitální knihovnu, následně je Designér zpracuje a působí tak na Digitální knihovnu.

Jak ukazuje Obr. 8:

- **Referenční model** (*Reference Model*) – podle MacKenzie et al. [2006] "*Referenční model se skládá z minimální množiny sjednocujících konceptů, axiomů a vztahů v rámci určité problémové domény a je nezávislý na konkrétních standardech, technologiích, implementacích či dalších specifických detailech*". Digitální knihovny potřebují odpovídající referenční model, aby dokázaly zkonsolidovat diverzitu existujících přístupů do jednotného a konzistentního celku.

- **Referenční architektura** (*Reference Architecture*) představuje abstraktní návrh implementace obsahu a vztahů zaznamenaných v referenčním modelu. Těchto referenčních architektur může existovat několik. Například jedna referenční architektura pro Systém digitální knihovny, který pokrývá „federování“ lokálních zdrojů a další pro speciální aplikace digitální knihovny.
- **Konkrétní architektura** (*Concrete Architecture*) v této fázi je architektura v referenčním modelu nahrazená skutečnou implementací. Například „run-time“ prostředí bude *CORBA*²⁸ nebo *Web Services Application Framework*²⁹.



Obr. 8 Univerzum digitální knihovny
(<http://www.dlib.org/dlib/march07/castelli/castelli-fig5.jpg>)

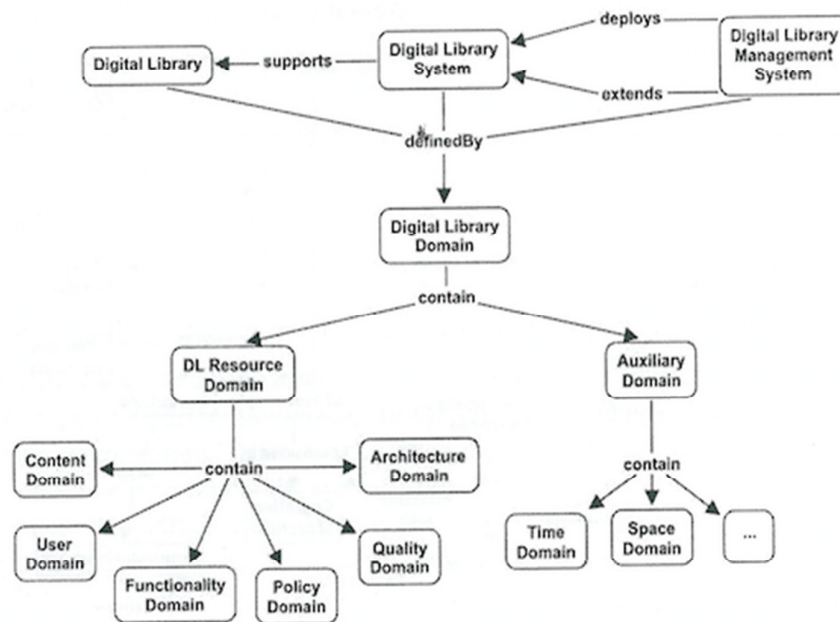
Obrazek podrobně popisuje prostředí digitální knihovny. Konkrétní architektura (*Concrete Architecture*) vychází (*Derived By*) z obecné Referenční architektury (*Reference Architecture*) definované v Referenčním modelu (*Reference Model*). Při konkrétní realizaci (*Implementation*) je nutné vzít v potaz i související Best Practices, výsledky výzkumu (*Research*), u architektury je nutné dbát na vhodné

²⁸ CORBA (*Common Object Request Broker Architecture*) představuje standard umožňující sjednocení různých softwarových komponentů sepsaných v různých jazycích do jednoho prostředí. Podrobněji např. na Wikipedii: <http://en.wikipedia.org/wiki/Common_Object_Request_Broker_Architecture>.

²⁹ Webové aplikační rámce (*Web Services Application Frameworks*) představují do značné míry ucelené a svébytné prostředí, do kterého tvůrci vkládají svůj programový kód, šablony, konfigurace a popisovače, čímž vytváří webovou aplikaci. Podrobněji např. na Wikipedii: <http://en.wikipedia.org/wiki/List_of_web_service_frameworks>.

protokoly (*Protocols*), standardy (*Standards*), specifikace (*Specifications*) a pokud je to vhodné zabudovat již hotové komponenty a systémy (*Off The shelf Components and Systems*). To vše ale musí vycházet (*Accounts For*) z předem stanovených cílů (*Goals*), motivace k provozu digitální knihovny (*Motivations*), konkrétních požadavků na digitální knihovnu (*Requirements*) a v neposlední řadě z toho, co nabízí "trh" pro digitální knihovny (*Market*).

První popisná část modelu se věnuje základním konceptům, složkám a prostředí, které tvoří digitální knihovnu. Cílem je objasnit univerzum digitálních knihoven všem aktérům dle jejich specifických potřeb. Výše (viz Obr. 4) rozebrané tři základní systémy (Digitální knihovna, Systém digitální knihovny a Systém správy digitální knihovny) tvoří třívrstvou konstrukci světa digitálních knihoven, přičemž každá z vrstev je přístupná pro různé aktéry (Koncoví uživatelé, Designéři, Systémoví administrátoři a Vývojáři aplikací). Digitální knihovna je tvořena z prvků, které reprezentují její tři hlavní systémy a je rozdělena do dvou hlavních tříd, v modelu uvedených jako domény: **Doména zdrojů** (*DL Resource Domain*) a **Pomocná doména** (*DL Auxiliary Domain*).



Obr. 9 Konceptuální mapa hierarchie domén digitální knihovny
[NETWORK, 2007]

Obrázek znázorňuje konceptuální mapu základních domén, resp. prvků reprezentující univerzum Digitální knihovny (v modelu uváděné i jako třívrstvá konstrukce digitální knihovny). Základ tvoří dvě domény: Doména zdrojů (*DL Resource Domain*) a Pomocná doména (*Auxiliary Domain*). Doména zdrojů

sestává ze šesti podpodmín: Obsah (*Content Domain*), Uživatelé (*User Domain*), Funkce (*Functionality Domain*), Politiky (*Policy Domain*), Kvalita (*Quality Domain*) a Architektura (*Architecture Domain*). Pomocnou doménu tvoří poddomény Čas (*Time Domain*), Prostor (*Space Domain*) a Jazyk (*DL Language Domain*).

DELOS Referenční model digitální knihovny již počítá i s potřebou dlouhodobé ochrany digitálních objektů (na rozdíl od např. "Kahn-Wilensky architektury"). Tvůrci modelu si uvědomovali jeho důležitost, proto je integrální součástí referenčního modelu. Referenční model v této souvislosti definuje Zdroje (*Resources*), Informační objekt (*Information Object*) a Funkce (*Functions*). Mezi funkce kritické pro dlouhodobou ochranu se řadí:

- **Transformace** (*Transform*) - zobrazuje objekty napříč časem a na různých zařízeních
- **Vizualizace** (*Visualize*) - doplňující funkce k Transformaci pro zobrazování objektů napříč časem a na různých zařízeních
- **Odstranění** (*Withdraw*) - umožňuje odstranění objektů ze Systému Digitální knihovny. Z pohledu dlouhodobé ochrany by tato funkce měla poskytovat mechanismy pro rozhodování - zda objekt k odstranění z digitální knihovny je nutné přesunout na jiné úložiště či do jiného systému nebo jej zcela odstranit
- **Export** (*Export*) - podporuje export částí digitální knihovny nebo jejího celku. Může však mít "jen" formu zálohové kopie nebo zrcadleného sídla (*Mirror Site*³⁰).
- **Porovnání** (*Compare*) - umožňuje osobě nebo počítačovému programu zjistit identitu nebo podobnost mezi dvěma instancemi objektu.

Aby však dlouhodobá ochrana byla opravdu efektivní, je nutné se zaměřit i na **kvalitativní parametry obsahu** (*Content Quality Parameters*) jako jsou **integrita** (*Integrity*), **autenticita** (*Authenticity*), **autoritativnost** (*Authoritativeness*), **interpretace** (*Performance*), **přesnost** (*Fidelity*), **spolehlivost** (*Dependability*) a **původ** (*Provenance*) digitálních objektů. Z pohledu **architektury** (*Architecture Quality Parameters*) je především nutné dodržovat **standardy** (*Compliance To Standards*). **Funkční kvalitativní parametry** (*Functionality Quality Parameters*) zastupuje **kontrola chyb v řízení** (*Fault Management Performance*). Vše je nutné podchytit **obecnými kvalitativními parametry** (*Generic Quality Parameters*) - **prosazováním bezpečnosti** (*Security Enforcement*), **podporou interoperability** (*Interoperability Support*) a **důslednou dokumentací** (*Documentation Coverage*).

³⁰ Zrcadlené sídlo (*Mirror Site*) představuje sídlo, které je replikou již existujícího sídla. Důvodem je snížení síťové zátěže nebo úprava originálního sídla bez toho, aby bylo nutné jej na nějaký čas znepřístupnit.

3.3 Open Archival Information System Reference Model - OAIS

V roce 1982 byla založena *The Consultative Committee for Space Data Systems - CCSDS* s cílem nabídnout mezinárodní fórum pro vesmírné agentury, zajímající se o kolaborativní vývoj standardů v oblasti správy dat na podporu vesmírného výzkumu. Na požádání ISO, CCSDS převzalo koordinaci vývoje archivních standardů, resp. standardů pro potřeby dlouhodobé ochrany digitálních dat. Na základě stanoveného cíle byl sestaven Referenční model Otevřeného archivního informačního systému³¹ (*Open Archival Information System Reference Model*), který zavádí jednotný koncept a terminologii (definice signifikantních vlastností a vztahů mezi entitami v archivním prostředí³²) a zároveň slouží jako podklad pro budoucí standardy v archivní oblasti³³.

Poprvé byl implementován v projektu NEDLIB (*Networked European Deposit Library*)³⁴, pro modelování procesů v systému pro dlouhodobou archivaci a zpřístupnění elektronických dokumentů (*Deposit Systems for Electronic Publications - DSEP*). Díky této iniciativě byl OAIS model obohacen o procesy plánování dlouhodobé ochrany.

V roce 2003 byl schválen jako mezinárodní ISO standard ISO 14721:2003 Space data and information transfer systems – Open archival information system - Reference model. Představuje zcela obecný rámec s cílem pochopit vzájemné vztahy mezi definovanými entitami i vztahy k vnějšímu okolí. Avšak nabízí také kompletní funkční a informační popis digitálního repozitáře a archivu (společně s terminologií a popisem různých typů architektury, resp. souvisejících operací). Cílem tohoto referenčního modelu tedy je:

- zvýšení povědomí a porozumění konceptů relevantních pro archivaci digitálních objektů, zejména pro nearchivní instituce
- sjednocení terminologie a konceptů pro popis a porovnávání datových modelů a archivní architektury
- rozšíření konsenzu prvků a procesů v oblasti ochrany digitálních informací a jejich zpřístupnění (dává základ pro správnou implementaci problematiky dlouhodobé ochrany digitálních dokumentů do procesů digitálních knihoven)

³¹ V názvu uvedený termín „open“, tedy „otevřený“, v kontextu modelu neodkazuje na stupeň zpřístupnění informací v archivu („open access“), ale na skutečnost, že model je otevřený změnám a doporučením v otevřených fórech [DRYDEN, 2009].

³² Archivní prostředí v kontextu tohoto modelu lze chápat jako jakýkoliv informační systém, který uchovává a zpřístupňuje fyzické či digitální objekty/dokumenty.

³³ CCSDS vydal doplňkový standard, který specifikuje první fázi procesu vstupu dat (*Ingest*) – *Producer-Archive Interface Methodology Abstract Standard* – PAIMAS.

³⁴ NEDLIB byl tříletý projekt (od ledna 1998 do konce roku 2000) financovaný Evropskou komisí. Podílelo se na něm 8 národních knihoven Evropy, jeden národní archiv a dvě ICT organizace a tři významná nakladatelství.

- vytvoření rámce pro identifikaci a vývoj standardů.

Rozsáhlý dokument, čítající na 140 stran je přehledně rozdělený do 6 částí. I když to nebývá běžnou praxí, text standardu je poměrně dobře srozumitelný i pro ty, kteří nejsou odborníky v daném oboru.

V úvodních částech definuje a objasňuje klíčové komponenty OAIS systému: **prostředí, informace, informačné balíčky**. Třetí část stanovuje povinné odpovědnosti OAIS. Jádro standardu lze nalézt v části 4, která se věnuje podrobnému modelování funkčních entit. Tvoří jej archivní informační systém a informace, které spravuje. Funkční model rozděluje architekturu „ideálního“ repozitáře do pěti hlavních modulů a vztahů mezi nimi³⁵:

- **přístup** (*Access*)
- **administrace** (*Administration*)
- **permanентní úložiště** (*Archival Storage*)
- **správa dat** (*Data Management*)
- **plánování dlouhodobé ochrany** (*Preservation Planning*).

Poslední dvě části pokrývají perspektivy dlouhodobé ochrany: dlouhodobá ochrana samotných digitálních informací, jako i uchování přístupu k nim a interoperabilitu archivu.

V oblasti terminologie nebyly použity nadužívané pojmy, aby nedocházelo k neúmyslné záměně termínů napříč obory a zároveň jsou podány definice základních termínů. Jednou z největších výhod OAIS je, že nabízí společný odborný slovník. Pro organizace plánující nový systém nabízí inspirativní podkladový materiál, pro systémy ve vývoji kontrolní seznam a pro spolupracující organizace umožňuje srovnávat různé systémy dle stejných kritérií a zjednodušit tak interoperabilitu. V tomto případě by přijetí OAIS rámce mohlo mít i potenciálně ekonomické výhody snížení nákladů skrze sdílené systémové komponenty.

OAIS představuje konceptuální rámec či „šablonu“ pro návrh systémů (včetně archivů a digitálních knihoven). Slouží pro podchycení všech zásadních aktivit archivu uchovávajícího informace, nikoliv „blueprint“, podrobně specifikující konkrétní přístupy. Formu abstraktního a konceptuálního modelu je možné považovat za výhodu i nevýhodu zároveň [DRYDEN, 2009]. Je poměrně složité přeložit model do podoby implementovatelné do praxe. Další nevýhodu je možné vidět v početných rozdílných a některých zcela protichůdných interpretacích. Z pohledu knihovní praxe se největší nevýhodou jeví „oborová nevyhraněnost“, kdy například definování cílové skupiny (pro velké digitální knihovny na národní úrovni) se jeví spíše jako komplikace než pomůcka [WIJNGAARDEN, 2004; LAVOIE, 2000].

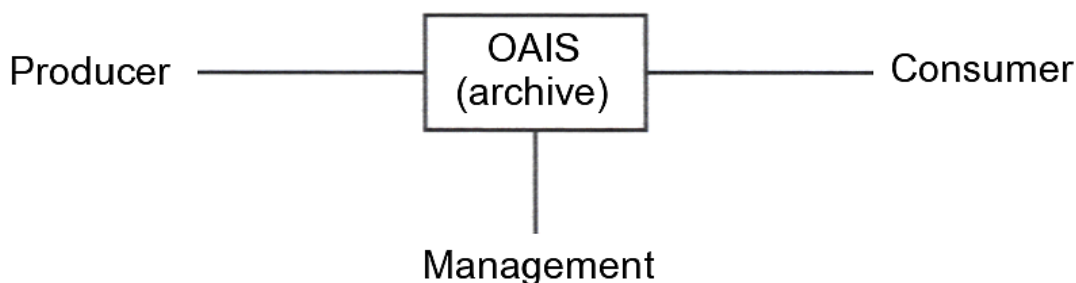
³⁵ Funkce systému OAIS jsou přehledně uvedeny ve formě tabulky v příloze Přehledný popis funkčních entit a rolí OAIS modelu.

Od archivu, resp. digitální knihovny typu OAIS se očekává, že dostojí alespoň těmto minimálním požadavkům [LAVOIE, 2000]:

- vyjednání a přijetí vhodných informací od producentů dat
- získání potřebné kontroly nad procesy a informačním systémem
- určení cílové skupiny (*Designated Community*) a zabezpečení, že informace je dostatečně srozumitelná cílové skupině
- dodržování politik a postupů k zabezpečení informací proti nepředvídatelným událostem a umožnění, aby šířené informace představovaly autentické kopie originálu nebo tak, aby byl originál dohledatelný
- zpřístupnění informací cílové skupině.

3.3.1 OAIS prostředí

OAIS prostředí tvoří **Producenti**, dodávající informace, které archiv uchovává. **Uživatelé** využívají uchovávané informace. Zvláštní podskupinou uživatelů je **cílová skupina** (*Designated Community*), která očekává, že informacím v archivu snadno porozumí. **Management** představuje entitu, která zodpovídá za sestavení politik, které naplní cíle archivu (např. výběr typu informací, které mají být uchovány, rozpočtové zdroje apod.). Nezahrňuje však každodenní správu archivu – ta je v rukách samotné digitální knihovny, resp. instituce, která ji spravuje.



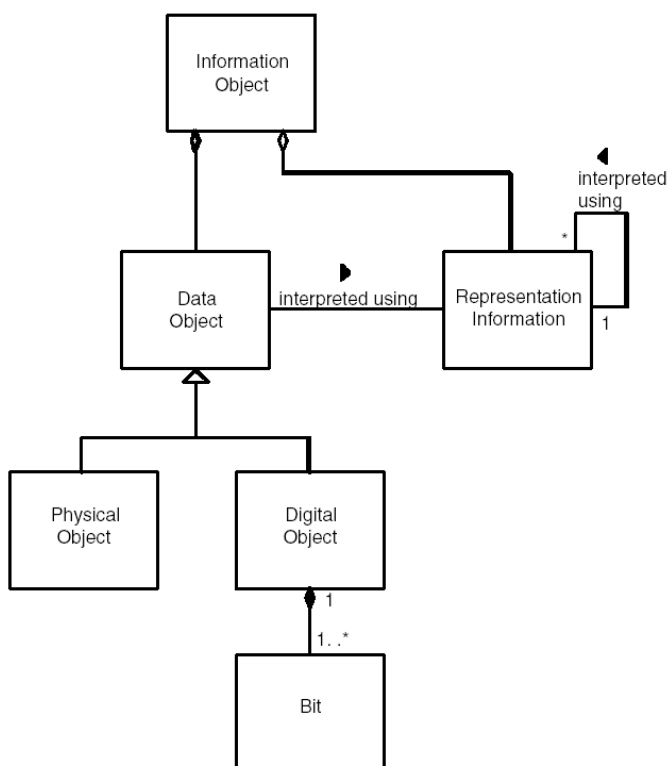
Obr. 10 OAIS v kontextu producentů, managementu a uživatelů
(<http://www.payer.de/digitalebibliothek/digibib0202.gif>)

Obrázek přehledně znázorňuje externí entity v prostředí OAIS, ve kterém na straně vstupu do systému hrají zásadní roli producent (*Producer*), na straně výstupu uživatel (*Consumer*). Mezi oběma skupinami je důležitý management, který zodpovídá za tvorbu politik v rámci repozitáře, resp. digitální knihovny.

Uvedené externí entity (producenti, uživatelé a management) přitom mohou být tvořeny na úrovni jedné instituce nebo mohou být distribuovány na více institucí a to vlastními silami nebo „outsourcovaním“.

3.3.2 Informační model

Základním prvkem modelu OAIS je **informační objekt** (*Information Object*), který sestává ze tří hlavních částí [CUBR, 2010]: **digitální datový objekt** (*Digital Data Object*), **interpretační informace** (*Representation Information*) a **znalostní základna** (*Knowledge Base*). Z důvodu komplexnosti modelu OAIS a definovaných rolí a funkcí bylo nutné přesně zadefinovat pojem informační objekt ve vztahu k informaci (Information). **Informace** je definována jako jakýkoliv typ znalosti, který je komunikován a v rámci komunikace vyjádřen formou dat. **Informační objekt** představuje **kombinaci datového objektu a reprezentace informace** [BRATKOVÁ, 2009] - viz Obr. 11.

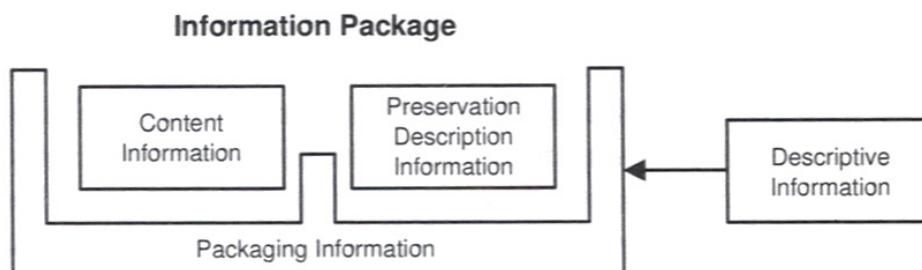


Obr. 11 Schéma získání informace (informačního objektu) z dat, resp. datového objektu (<http://twiki.dcc.rl.ac.uk/pub/OLD/DCCApproachToCuration/InformationObject.gif>)

Obrázek přesně ukazuje, jakým způsobem vzniká informační objekt jako základní prvek modelu OAIS. Bity (*Bit*) tvoří datový objekt (*Data Object*), který je reprezentován fyzickým objektem (*Physical Object*) a digitálním objektem (*Digital Object*), přičemž mezi digitálním objektem a bity existuje vztah 1:n (mnoho bitů tvoří jeden digitální objekt). Datový objekt je dále vyjádřen (*Interpreted Using*) pomocí reprezentace informace (*Representation Information*) a dává vzniknout (*Interpreted Using*) nové informaci - informačnímu objektu (*Information Object*).

Intepretace datového objektu jako smysluplné informace je v rukách cílové skupiny a jejich znalostní báze (příklad programátoři a Java programátoři), která nemusí být srozumitelná pro cílovou skupinu. V takovém případě musí mít datový objekt interpretační informaci. Právě kombinací datového objektu a informací o datovém objektu vznikne „smysluplná informace“ pro cílovou skupinu. Tu je možné strukturovat do balíčků a konceptualizovat tok informací napříč digitálním archivem.

Zavedení **informačních balíčků**: **SIP** (*Submission Information Package*) – **dodavatelský balíček**, **DIP** (*Dissemination Information Package*) – **uživatelský balíček** a **AIP** (*Archival Information Package*) – **archivační balíček** je klíčovou podmínkou pro interoperabilitu mezi digitálními knihovnami, repozitáři a archivy navzájem nebo jinými informačními systémy. Informační balíček představuje logickou jednotku, která seskupuje informační obsah spolu s archivačními informacemi. Je nezbytné, aby informační balíčky měly i tzv. balíčkovací informace (*Packaging Information*), což jsou informace o tom, jak je daný balíček seskupen (v adresářích), komprimován apod. [CUBR, 2010] – viz. Obr. 12:



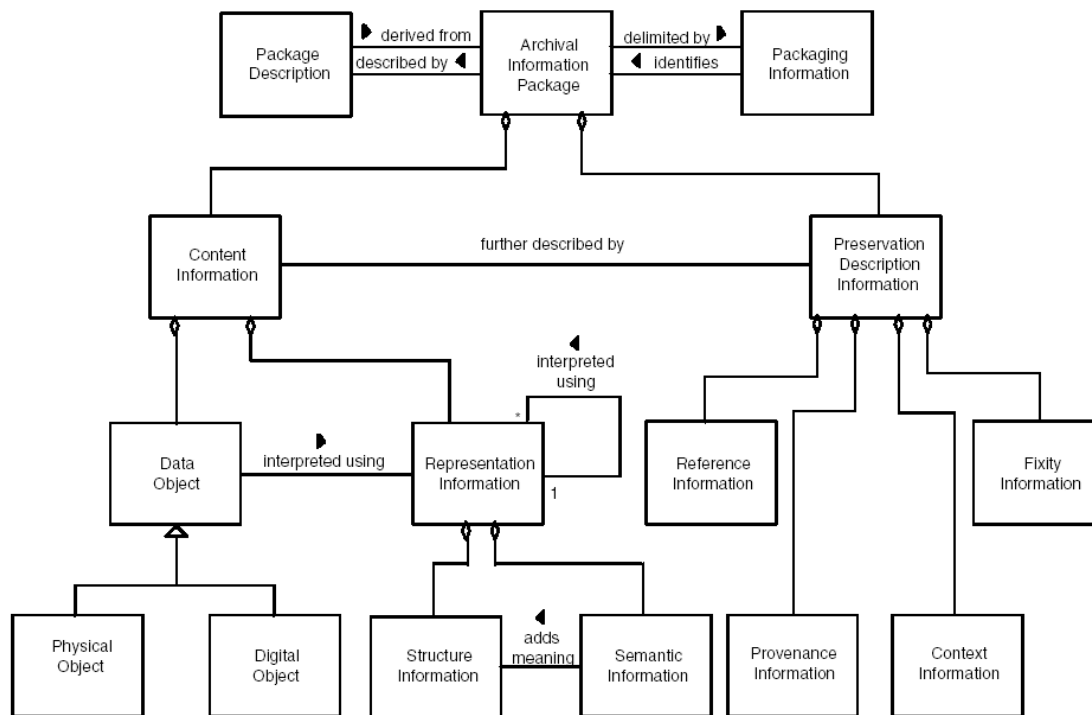
Obr. 12 Informační balíček – koncept a vztahy
(<http://www.payer.de/digitalebibliothek/digibib0203.gif>)

Obrázek popisuje složení a vztahy informačních balíčků v modelu OAIS. Balíček představuje logický kontejner (princip podobný METS standardu), jež obsahuje vlastní informaci (*Content Information*), která má být dlouhodobé uchovávána a

preservační informaci (*Preservation Description Informaton*), která formou metadat podporuje dlouhodobou ochranu. Obsah informace a preservační informace je spojen do logického celku, který se nazývá archivační informace (*Packaging Information*). Informace o balíčku (*Descriptive Information*) zprostředkovávají přístup k informačnímu balíčku přes vyhledávací nástroje archivu.

- **Dodavatelský balíček** – SIP se v kontextu OAIS modelu dostává do digitálního archivu od producentů/tvůrců dat. Jeho forma je obvykle předem dojednána mezi producenty dat a přijímající stranou za archiv, nejčastěji formou smluv a dohod, které podrobně specifikují, jaké dodatečné (např. popisné) informace budou spolu s digitálními daty v balíčcích dodávat. SIP nemusí nutně představovat rovnici jeden digitální objekt = jeden SIP.
- **Archivační balíček** – AIP je vytvořen systémem po úspěšném přijetí a zpracování SIP. Systém ale nemění přijaté SIP, pouze k nim přidává doplňující archivní informace za účelem jejich dlouhodobé archivace.
- **Uživatelský balíček** – DIP reprezentuje uživateli požadované digitální dokumenty (na základě dotazu/ů). Nemusí nutně obsahovat veškeré archivační informace, které obsahuje AIP. Uživatele zpravidla nezajímají záznamy ochranných aktivit, provedených během jejich životního cyklu v digitálním repozitáři.

SIP, AIP a DIP nemusí být nezbytně rozdílné, to závisí na konkrétní implementaci modelu.



Obr. 13 Detailní schéma archivačního balíčku (AIP)

<http://twiki.dcc.rl.ac.uk/pub/OLD/DCCApproachToCuration/AIP-detailed.gif>

Obrázek podrobně rozebírá strukturu archivačního balíčku (AIP) modelu OAIS. Balíček (*Archival Information Package*) je definován (*Defined By*) balíčkovací informací (*Packaging Information*) a informací o balíčku (*Package Description*), jež určují jakou formu a popis bude AIP mít. Inherentní součástí AIP představují obsahové informace (*Content Information*) a archivační informace (*Preservation Description Information*). Součástí obsahové informace je datový objekt (*Data Object*) složený z fyzického objektu (*Physical Object*) a digitálního objektu (*Digital Object*). Datový objekt je interpretován (*Interpreted By*) reprezentací informace (*Representation Information*). Ta je dána strukturálními (*Structure Information*) a semantickými (*Semantic Information*) informacemi. Preservační informace (*Preservation Description Information*) formou metadat podporuje dlouhodobou ochranu a obsahuje důležité informace jako jsou referenční informace (*Reference Information*), informace o původu (*Provenance Information*), kontextové informace (*Context Information*) a informace o stabilitě (*Fixity Information*).

Informační balíček je dále tvořen čtyřmi typy informačních objektů (viz Obr. 13):

- **Obsahová informace** (*Content Information*) je primárním zdrojem zájmu dlouhodobé ochrany – datový objekt a připojená informace v podobě reprezentace (např. jako odkaz mezi rozdílnými verzemi stejného objektu)

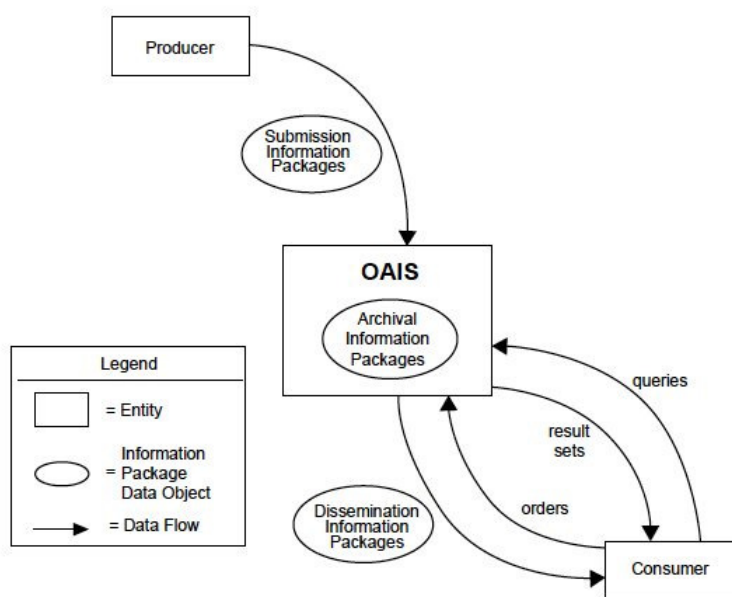
- **Archivační informace** (*Preservation Description Information – PDI*) je potřebná pro podporu a dokumentaci ochranných aktivit. Obsahuje informace nezbytné k adekvátní ochraně informací o objektu (provenience³⁶, jednoznačné identifikátory pro obsahové informace³⁷ a informace validující autenticitu daného obsahu, jako např. kontrolní součet nebo digitální podpis). V digitálním knihovně musí být každý informační balíček jednoznačně identifikovatelný.

PDI se tedy skládá ze čtyř typů metadatových informací [BRATKOVÁ, 2009]:

- **Referenční informace** (*Reference Information*) - představuje unikátní identifikátor pro samotné informace (uvnitř i mimo repozitář). Např. DOI, URN apod.
- **Kontextová informace** (*Context Information*) - popisuje vztahy uvnitř informačního objektu
- **Informace o původu** (*Provenance Information*) - obsahuje historii (vytvoření, změny v obsahu, formátu či lokace) jakýchkoliv provedených akcí na informačním objektu
- **Informace o stabilitě** (*Fixity Information*) - zabezpečuje integritu a autenticitu obsahu (digitální podpis, kontrolní součty apod.)
- **Informace o balíčku** (*Packaging Descriptive Information*) spojuje jednotlivé složky informačního balíčku do identifikovatelných entit
- **Popisné informace** (*Descriptive Information*) podporuje vyhledávání a zpřístupnění obsahové informace. Zprostředkovávají přístup k informačnímu balíčku přes vyhledávací nástroje archivu.

³⁶ Informace o provenienci představují historii obsahové informace (vytvoření, pozměnění, ochranné akce – migrace a konverze).

³⁷ Pod jednoznačným identifikátorem rozumíme identifikátor typu ISBN, DOI apod.

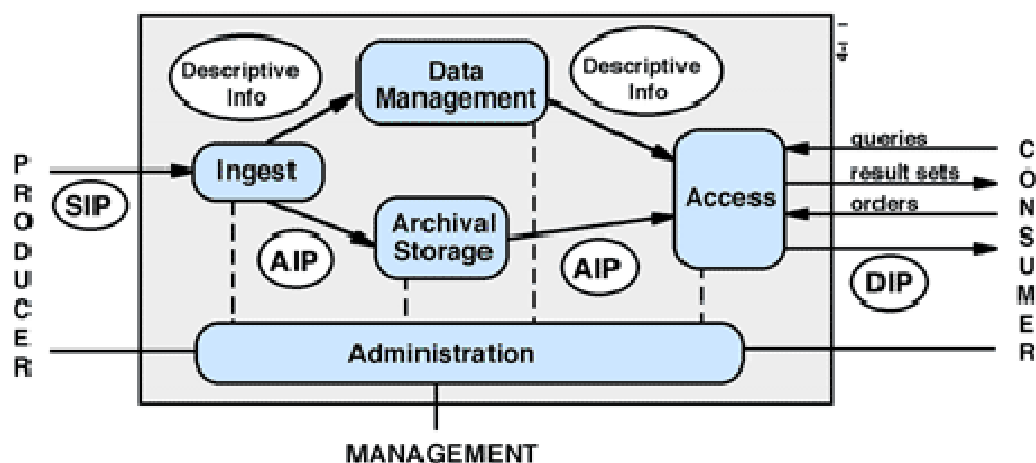


Obr. 14 Interakce a datový tok externích entit modelu OAIS
[ISO, 2003]

Schéma ukazuje datový tok (*Data Flow*; na obrázku znázorněn šipkou) a interakce externích entit (*Entity*; na obrázku je znázorněna obdélníkem) archivu podle OAIS modelu. Producent dat (*Producer*) odevzdává Dodavatelské balíčky SIP (*Submission Information Package*), které jsou v OAIS archivu transformovány a uchovány jako Archivní balíčky AIP (*Archival Information Package*). Na druhé straně uživatel (*Consumer*) žádá (*Queries, Orders*) OAIS o AIP, která se mu navrací ve výsledcích (*Results Set*) ve formě (původně přetransformovaného AIP) Uživatelského balíčku DIP (*Dissemination Information Package*).

3.3.3 Funkční model a entity

Hlavními vlastnostmi OAIS jsou informační a funkční entity. OAIS model rozděluje systém do 5 základních funkčních skupin: **Příjem** (*Ingest*), **Správa dat** (*Data Management*), **Administrace** (*Administration*), **Permanentní úložiště** (*Archival Storage*) a **Přístup** (*Access*) [ISO, 2003]:



Obr. 15 Grafické vyjádření funkčních skupin OAIS modelu
http://nssdc.gsfc.nasa.gov/nssdc_news/dec00/oais_fig3.gif

Obrázek popisuje dílčí funkční oblasti, jejich vztahy a vazbu na okolní prostředí modelu otevřeného archivního informačního systému - Příjem (*Ingest*), Správa dat (*Data Management*), Administrace (*Administration*), Permanentní úložiště (*Archival Storage*) a Přístup (*Access*). Dodavatelské balíčky (SIP) do systému vstupují od producentů dat (*Producer*) a jsou před vstupem přetřansformovány do podoby archivních balíčků (AIP). V této podobě je spravuje management jako entita (*Management*). V případě, že uživatel požaduje (*Queries, Orders*) konkrétní informace obsažené v balíčcích přes modul přístup (*Access*), dostávají (*Results Sets*) se mu AIP přetavené do podoby uživatelských balíčků (DIP). Balíčky v sobě obsahují popisné informace potřebné pro jejich identifikaci (*Descriptive Info*).

1. **PŘÍJEM** (*Ingest*) – zodpovídá za příjem informací od producentů a přípravu pro uchování a správu v archivu; přijímá informace od producentů ve formě SIP, provádí kvalitativní kontroly na SIP, generuje z nich AIP (z jedné nebo několika SIP) a extrahuje popisné informace pro AIP (metadata pro vyhledávání, náhledy pro prohledávání apod.); přenáší nově vytvořené AIP do Permanentního úložiště a přidružené DIP do modulu Správy dat
2. **SPRÁVA DAT** (*Data Management*) – spravuje popisné informace vztahující se k archivním AIP jako doplněk k systémovým informacím, které podporují operace archivu. Modul spravuje databázi obsahující tyto informace; odpovídá na dotazy přicházející z Access modulu a generuje výsledky na výstup pro uživatele. Dále také vytváří reporty pro podporu funkcí modulu Příjem, Přístup a Administrace. Vytváří updaty v databázi Správy dat, včetně doplnění nových popisných informací přijatých od Příjmu nebo nových systémových dat od Administrace.

3. **ADMINISTRACE** (*Administration*) – řídí každodenní operace archivu; to zahrnuje smlouvy s informačními producenty a provádí systémové inženýrství, kontrolu přístupu a zákaznické služby. Funkce Administrace též provádí pravidelné audity na SIP, aby kontrolovala jejich shodu se smlouvou a vyvíjí politiky a standardy, které se vztahují k datovým standardům systému (standardy formátů, požadavky na dokumenty, úložiště, migrace a bezpečnostní politiky). Tyto funkce slouží i jako rozhraní mezi archivem a těmi s komponenty OAIS prostředím: správa a cílová skupina. Jde tedy nejen o fyzickou ale i organizační úroveň administrace archivu.
4. **PERMANENTNÍ ÚLOŽIŠTĚ** (*Archival Storage*) – uchovává, řídí a vybírá AIP v archivu. Přijímá AIP z modulu Příjem, ukládá do permanentního úložiště podle různých kritérií (požadavky informačních medií/nosičů, očekávaná doba zpracování apod.). Dále zabezpečuje migraci AIP na nová média, kontroluje chyby a poskytuje požadované kopie modulu Přístup. Zde se také implementují strategie „disaster recovery“³⁸.
5. **PŘÍSTUP** (*Access*) – pomáhá uživatelům identifikovat a získat popis relevantních informací v archivu a doručuje informace z archivu uživatelům. Tato funkce předpokládá jednotné uživatelské rozhraní pro vyhledávání v obsahu archivu. Umožňuje též vyhledání relevantní AIP z archivu, která je na výstupu přetransformována do DIP. Popisná metadata získává z modulu Správa dat.

Součástí modelu jsou návrhy na uchovávání metadat ve formě ochranné popisné informace (*Preservation Description Information - PID*). Konkrétní implementace modelu , resp. jednotlivých funkčních entit je závislá na samotném repozitáři či archivu.

3.3.3.1 Příjem (Ingest)

Příjem odevzdaných dat (*Receive Submission*) nabízí vhodné úložní kapacity nebo nástroje pro příjem SIP od producentů/tvůrců dat nebo z modulu Administrace³⁹. SIP je možné přijmout přes např. FTP, externí média nebo lokální souborový systém. Producenti nebo modul Administrace poté dostávají zprávu o úspěšném příjmu SIP nebo naopak chybové hlášení společně s žádostí o opětovné zaslání dat.

³⁸ „Disaster recovery“ představují politiky, procesy a postupy, které se vážou na obnovení technologické infrastruktury v případě možného nesprávného či neoprávněného zásahu člověkem nebo přirozeným selháním technologií (např. přírodní katastrofy).

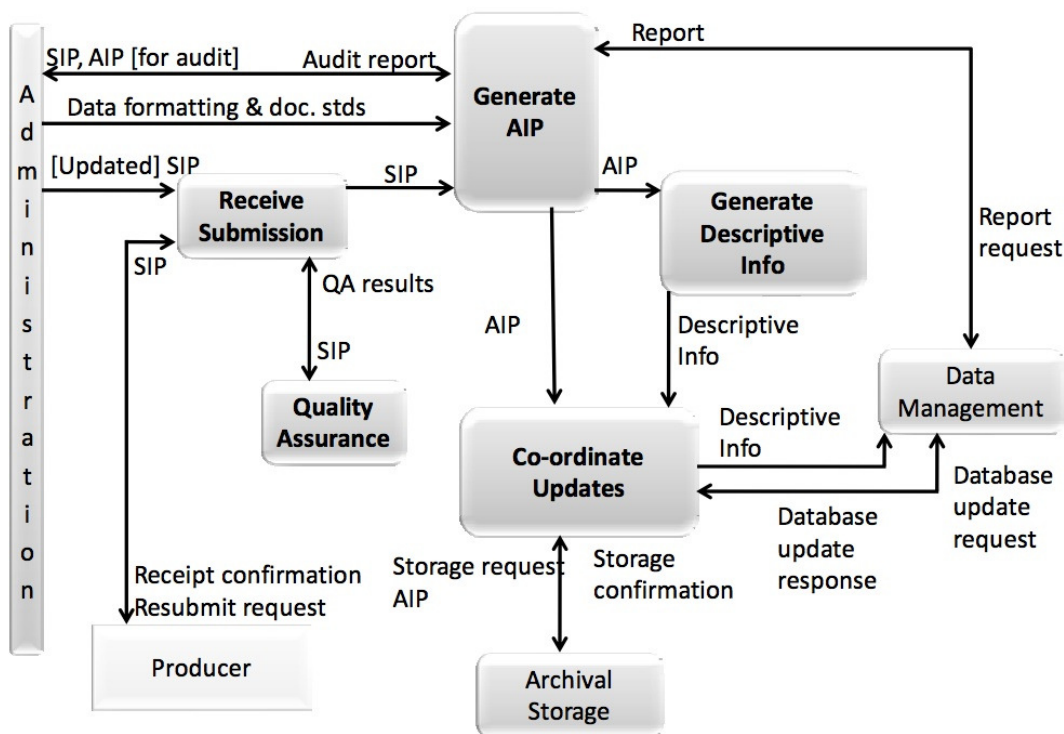
³⁹ To se týká případů „updateů“, kdy data z Archivního úložiště přichází opětovně do modulu Příjem jako SIP (v tomto případě AIP = SIP).

Zabezpečení kvality (*Quality Assurance*) validuje přenos SIP do pracovní oblasti (*Staging Area*). Validace probíhá např. formou CRC (*Cyclic Redundancy Checks*) či různých kontrolních součtů, systémových logů, které zaznamenávají přenos souborů a případné chyby při zápisu a čtení.

Generování AIP (*Generate AIP*) – transformace jedné nebo vícero SIP do AIP, které odpovídají dokumentovým a formátovým standardům archivu. Může zahrnovat konverzi souborových formátů nebo přeskupení obsahové informace v SIP.

Generování popisných informací (*Generate Descriptive Info*) extrahuje popisné informace z AIP a dalších zdrojů za účelem podpory funkce Koordinace updatů (*Co-ordinate Updates*) a modul Správa dat. Jedná se o metadata pro vyhledávání AIP a pomůcky k prohledávání (např. náhledy).

Koordinace updatů (*Co-ordinate Updates*) následně zodpovídá za přenos AIP do Permanentního úložiště a popisných informací do Správy dat. Po ukončení a verifikaci přenosu potvrdí Permanentní úložiště příjem (s informací pro další identifikaci AIP) a Koordinace updatů zašle do Správy dat žádost o update databázy.



Obr. 16 Funkce modulu Příjem (*Ingest*)
[ISO, 2003]

Obrázek popisuje funkce modulu Příjem (*Ingest*). Příjem odevzdaných dat (*Receive Submission*) přijímá odevzdaná data přímo od producentů (*Producer*) nebo z modulu Administrace (např. v případě migrací či jiných akcí, které je na

objektech potřebné provést ([*Updated*] *SIP*). Zabezpečení kvality (*Quality Assurance*), potvrdí (*QA results*) úspěšný přenos *SIP* do pracovní oblasti. Funkce Generování AIP (*Generate AIP*) transformuje jednu nebo více *SIP* do jedné nebo více AIP, které odpovídají dokumentovým a formátovým standardům archivu (*Data formatting & doc. stds*). Generování popisných informací (*Generate Descriptive Info*) extrahuje popisné informace z AIP a sbírá popisné informace za účelem podpory funkce Koordinace updatů (*Co-ordinate Updates*) a modulu Správa dat (*Data Management*).

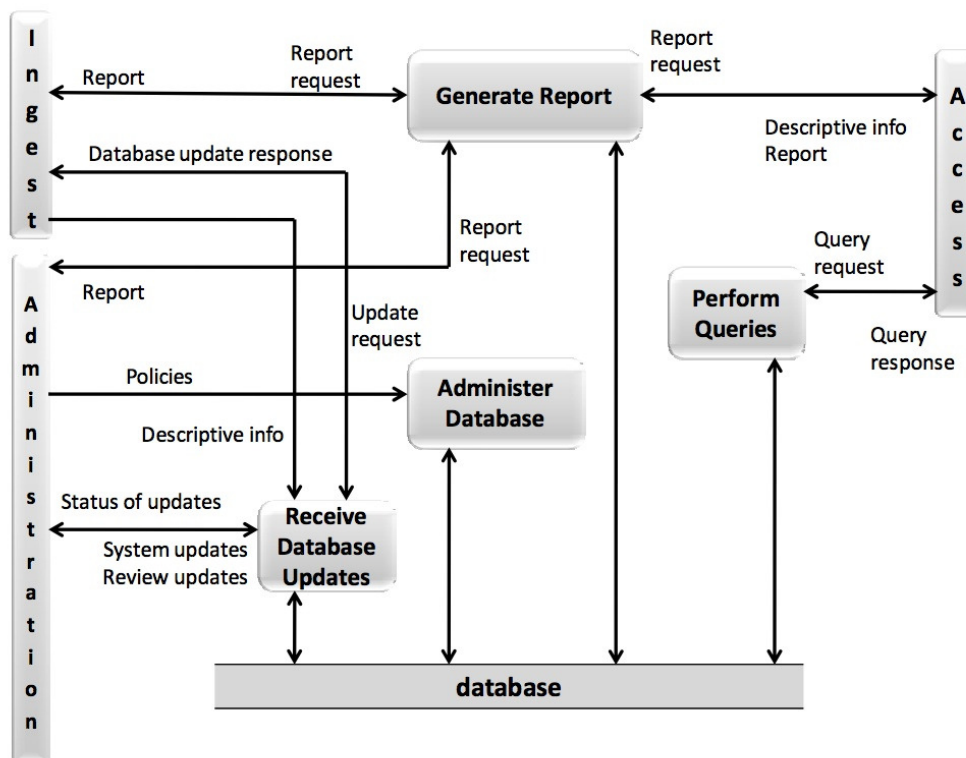
3.3.3.2 Správa dat

Administrace databáze (*Administer Database*) zabezpečuje integritu databáze entity Správa dat, která obsahuje popisné a systémové informace. Cílem funkce je vytvářet definice schémat dostatečně podporující ostatní funkce modulu Správa dat. Důležitá je i interní kontrola a validace obsahu databáze podle politik stanovených v modulu Administrace.

Dotazování (*Perform Queries*) přijímá dotazy z modulu Přístup a pro žadatele sestavuje relevantní výsledky.

Generování reportů (*Generate Report*) přijímá dotazy z modulů Příjem, Přístup a Administrace a vykonává všechny patřičné operace pro zpřístupnění dotazů funkci Dotazování.

Příjem updatů databáze (*Receive Database Updates*) přidává, upravuje nebo odstraňuje informace z úložiště Správy dat. Hlavním zdrojem úprav jsou funkční entity Příjem (poskytuje popisné informace pro nové AIP) a Administrace (vytváří systémové „updaty“ a „review“ updaty).



Obr. 17 Funkce modulu Správa dat (*Data Management*)
[ISO, 2003]

Obrázek popisuje funkce modulu Správa dat (*Data Management*). Administrace databáze (*Administer Database*) má za hlavní úkol udržovat integritu databáze Správa dat. Dotazování (*Perform Queries*) přijímá dotazy (*Query Requests*) z modulu Přístup (*Access*) a vrací zpátky požadované výsledky (*Result Set*). Generování reportů (*Generate Report*) přijímají žádosti (*Report Request*) z modulů Příjem, Přístup a Administrace (*Administration*) a zároveň se vykonávají všechny patřičné akce pro zpřístupnění dotazů funkcí Dotazování. Příjem updatů databáze (*Receive Database Updates*) přidává, upravuje nebo odstraňuje informace z úložiště Správy dat.

3.3.3.3 Permanentní úložiště

Příjem dat (*Receive Data*) přijímá žádosti o uložení AIP od modulu Příjem a přesouvá vyhovující AIP do Permanentního úložiště. Odpovídá za výběr typu média, přípravu zařízení nebo jednotek a fyzický transfer do Permanentního úložiště.

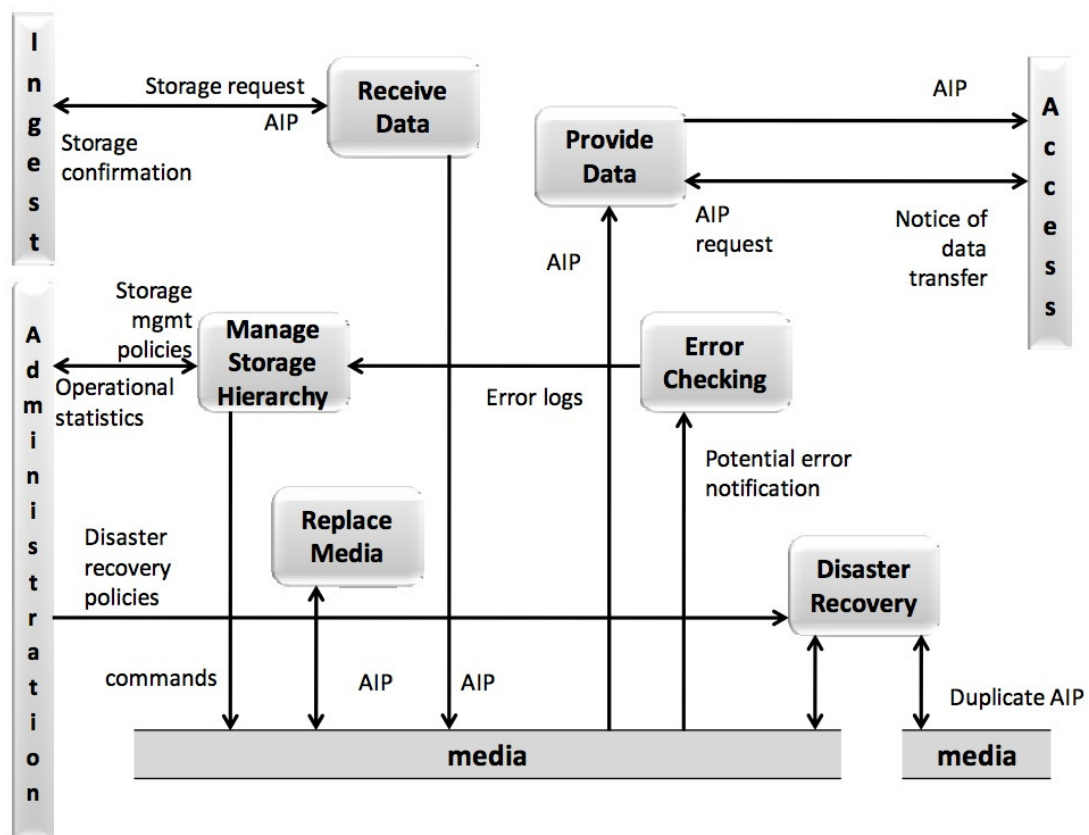
Správa hierarchie úložiště (*Manage Storage Hierarchy*) selektuje zápis obsahu AIP na vhodná média podle politik archivu, operačních statistik atd.

Výměna médií (*Replace Media*) kopíruje AIP bez pozměnění obsahové informace a popisné informace pro dlouhodobou ochranu (PDI). Pokud jsou součástí obsahové informace s atributy závislými na médiích, zabezpečuje, aby byly zachovány i po migraci.

Kontrola chyb (*Error Checking*) zajišťuje, aby žádné součásti AIP nebyly poškozeny během interních datových přesunů v Permanentním úložišti. Upozorňuje přitom na potenciální chyby, vypisuje chybové logy, které následně kontrolují osoby zodpovědné za Permanentní úložiště.

Obnova škod (*Disaster Recovery*) duplikuje digitální obsah archivních sbírek a uchovává duplikáty na fyzicky odděleném místě zkopírováním archivního digitálního obsahu na obnovitelná média.

Poskytování dat (*Provide Data*) poskytuje kopie AIP do modulu Přístup. Přijímá také žádosti o AIP, identifikuje vyžádané AIP a poskytuje je na požadovaných typech médií nebo je přesune do pracovního prostoru. Po splnění žádosti posílá oznámení o provedeném datovém přenosu do modulu Přístup.



Obr. 18 Funkce modulu Permanentní úložiště (*Archival Storage*)
 [ISO, 2003]

Obrázek popisuje funkce a procesy v rámci modulu Permanentní úložiště (*Archival Storage*). Z modulu Přijem (*Access*) se přijímají data v podobě AIP (*Receive Data*). V této fázi se žádá o přijetí a kontroluje dostupnost úložiště pro danou AIP

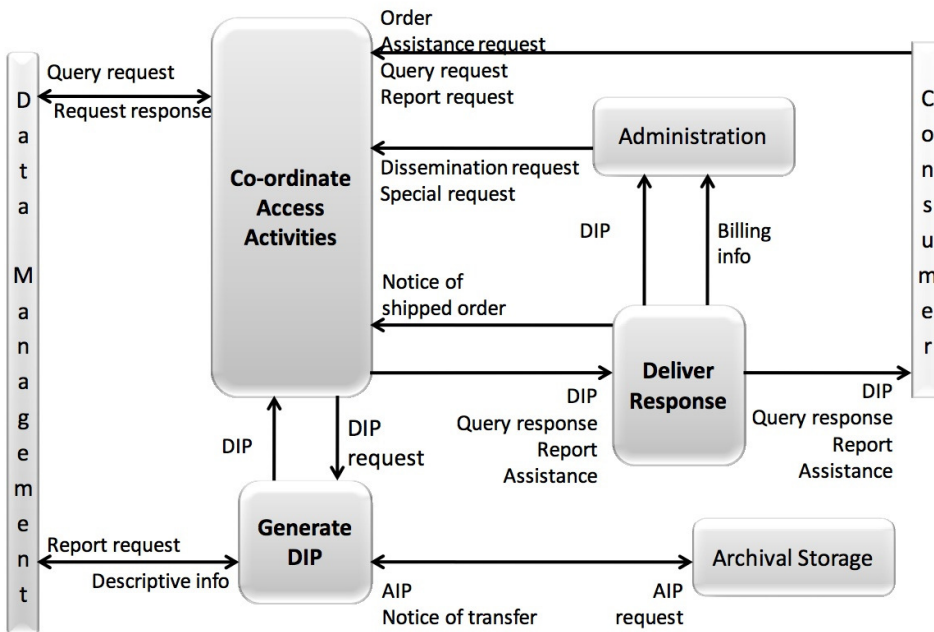
(*Storage Request*). V případě kladně vyřízené žádosti se do modulu Příjem vrací potvrzení o přijetí (*Storage Confirmation*). Správa hierarchie úložiště (*Manage Storage Hierarchy*) vybírá vhodná média pro zápis obsahu AIP podle stanovených politik archivu, operačních statistik atd. Výměna médií (*Replace Media*) zabezpečuje kopírování AIP bez pozměnění jejich obsahové a popisné informace. Funkce Kontrola chyb (*Error Checking*) má na starost integritu AIP během interních datových přesunů v Permanentním úložišti. Poskytování dat (*Provide Data*) poskytuje kopie AIP do modulu Přístup a Přijímá také žádosti o AIP.

3.3.3.4 Přístup

Koordinace přístupových aktivit (*Coordinate Access Activities*) nabízí jednoduché uživatelské rozhraní k informačním objektům (Information Holdings) v archivu. Rozlišuje tři typy uživatelských požadavků: **dotazy** (*Query Request*), které zpracuje Správa dat a navrací soubor výsledků pro prezentaci uživateli, **žádosti o reporty** (*Report Queries*) a **objednávky-požadavky** (*Orders*), které osloví buď Správu dat a/nebo Permanentní úložiště, aby připravily DIP.

Generování DIP (*Generate DIP*) přijímá žádosti o vyvolání AIP, vybírá příslušnou AIP z Permanentního úložiště a kopii dat přesouvá do pracovního prostoru (*Staging Area*) k dalšímu zpracování. Následně, za účelem získání popisné informace k DIP, posílá zprávu se žádostí do modulu Správa dat. Kompletní DIP poté přesouvá do pracovního prostoru a tuto skutečnost oznámí funkci Koordinace přístupových aktivit.

Doručení odpovědi (*Deliver Response*) (pro online doručování žádostí uživatelům) přijímá z Koordinace přístupových aktivit žádost a připravuje ji k zobrazení v reálném čase přes příslušné komunikační linky. Po úspěšném doručení navrací oznámení o doručení odpovědi do Koordinace přístupových aktivit a případné fakturační informace do modulu Administrace.



Obr. 19 Funkce modulu Přístup (*Access*)
[ISO, 2003]

Obrázek popisuje funkce modulu Přístup (*Access*). Koordinace přístupových aktivit (*Coordinate Access Activities*) poskytuje jednoduché uživatelské rozhraní k archivu a rozlišuje tři typy uživatelských požadavků: dotazy (*Query Request*), které zpracuje Správa dat (*Data Management*) a navrací soubor výsledků pro prezentaci uživateli (*Result Set*); žádosti o reporty (*Report Queries*) a objednávky-požadavky (*Orders*), které osloví buď Správu dat a/nebo Permanentní úložiště, aby připravily DIP. Generování DIP (*Generate DIP*) přijímá žádosti o vyvolání AIP, vybírá příslušnou AIP z Permanentního úložiště a kopii dat přesouvá do pracovního prostoru (*Staging Area*) k dalšímu zpracování. Doručení odpovědi (*Deliver Response*) připravuje k zobrazení v reálném čase žádosti uživatelů přes příslušné komunikační linky.

3.3.3.5 Administrace

Vyjednání smluv pro odevzdávání dat (*Negotiate Submission Agreement*) – vyjednání smluv s producenty, které stanoví harmonogram odevzdávání dat. Tato funkce také přijímá doporučené šablony a úpravy od modulu Plánování dlouhodobé ochrany a posílá návrhy AIP a SIP k vyhodnocení do „Žádosti o audit“. Ty jsou vyhodnoceny na základě již kvalitně zdokumentovaných politik preferovaných formátů a požadovaných postupů.

Správa systémové konfigurace (*Manage System Configuration*) – poskytuje archivnímu systému prostředky systémového inženýrství, kterými monitoruje archiv a systematicky kontroluje případné změny v konfiguraci. Tato funkce odpovídá za snadnou ovladatelnost a integritu konfigurace ve všech fázích životního cyklu systému. Kontroluje též operace, výkon a využití systému. Přijímá a sumarizuje reporty ze Správy dat a operační statistiky z Permanentního úložiště, které dále zasílá do modulu Plánování dlouhodobé ochrany. Zasílá žádosti o změny, postupy a nástroje do pracovního prostoru Update archivní informace.

Update archivní informace (*Archival Information Update*) nabízí mechanismy pro update obsahu archivu. Přijímá žádosti o změny, postupy a nástroje ze Správy systémové konfigurace. Přeposílá tyto žádosti do modulu Přístup. Po updatu obsahu se změní DIP a vrací se jako SIP do modulu Příjem.

Kontrola fyzického přístupu (*Physical Access Control*) umožňuje nebo také zabraňuje fyzickému i vzdálenému přístupu do archivu podle stanovených archivních politik.

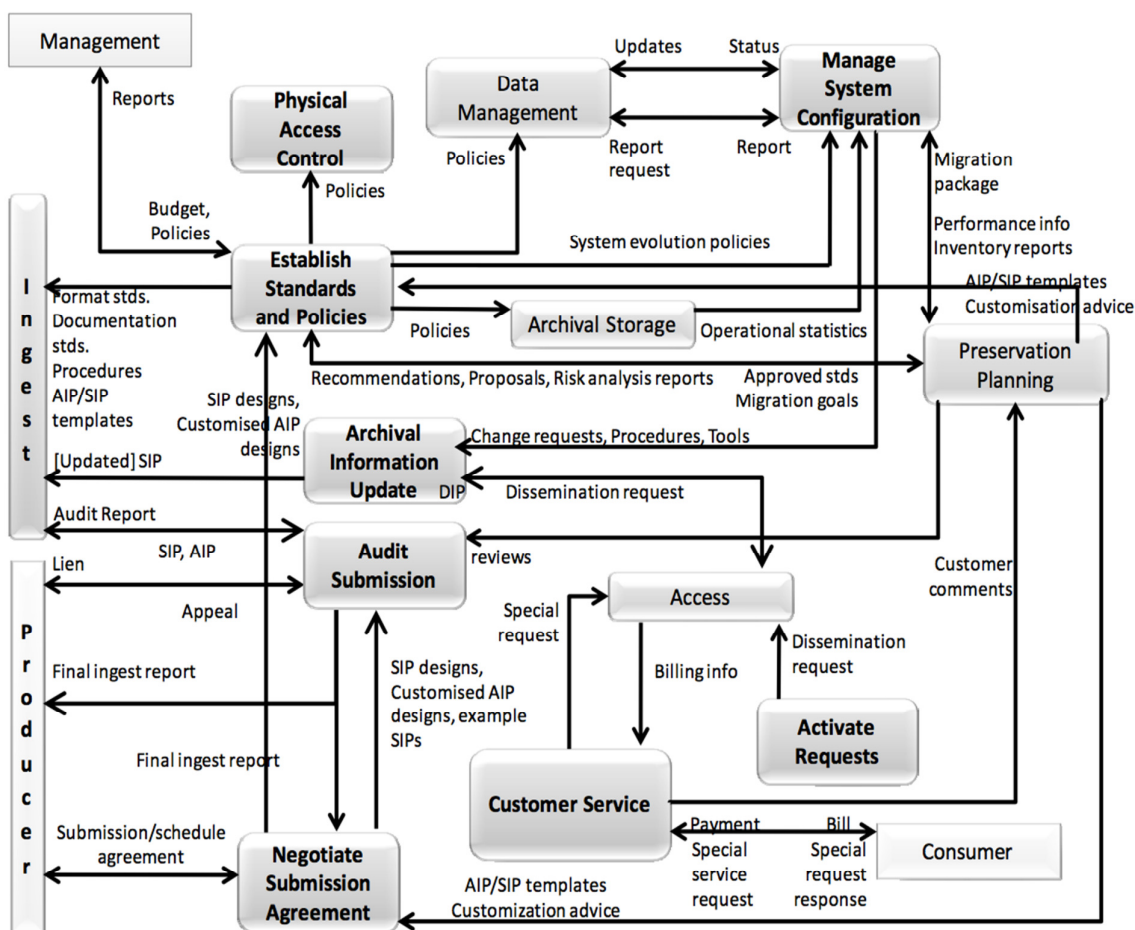
Zavedení standardů a politik (*Establish Standards and Policies*) akceptuje informace a politiky k rozpočtu a vytváří pravidelné reporty pro management archivu. Přijímá doporučení a návrhy na nové archivní datové standardy od modulu Plánování dlouhodobé ochrany. Na základě přijatých

informací a reportů ze Správy systémové konfigurace vytváří nebo pozměňuje politiky a standardy pro ostatní funkce modulu Administrace. Funkce vytváří i politiky pro správu úložiště a politiky obnovy dat.

Aktivace žádostí (*Activate Requests*) aktivuje požadavky na zpřístupnění přicházející z modulu Přístup.

Žádosti o audit (*Audit Submission*) verifikuje přijímané SIP (nebo AIP) podle vyjednaných podmínek. Samotný proces auditu může rozhodnout, které části SIP nejsou vhodné pro přijetí do archivu a musí být zaslány v jiné podobě nebo jsou zcela vyloučeny. Poté je vypracovaný report auditu zaslán do modulu Přijím a případné nesrovnalosti i producentovi dat.

Zákaznický servis (*Customer Service*) vytváří, spravuje a maže uživatelské účty, přijímá fakturační údaje od modulu Přístup, zasílá faktury a přijímá platby od uživatelů.

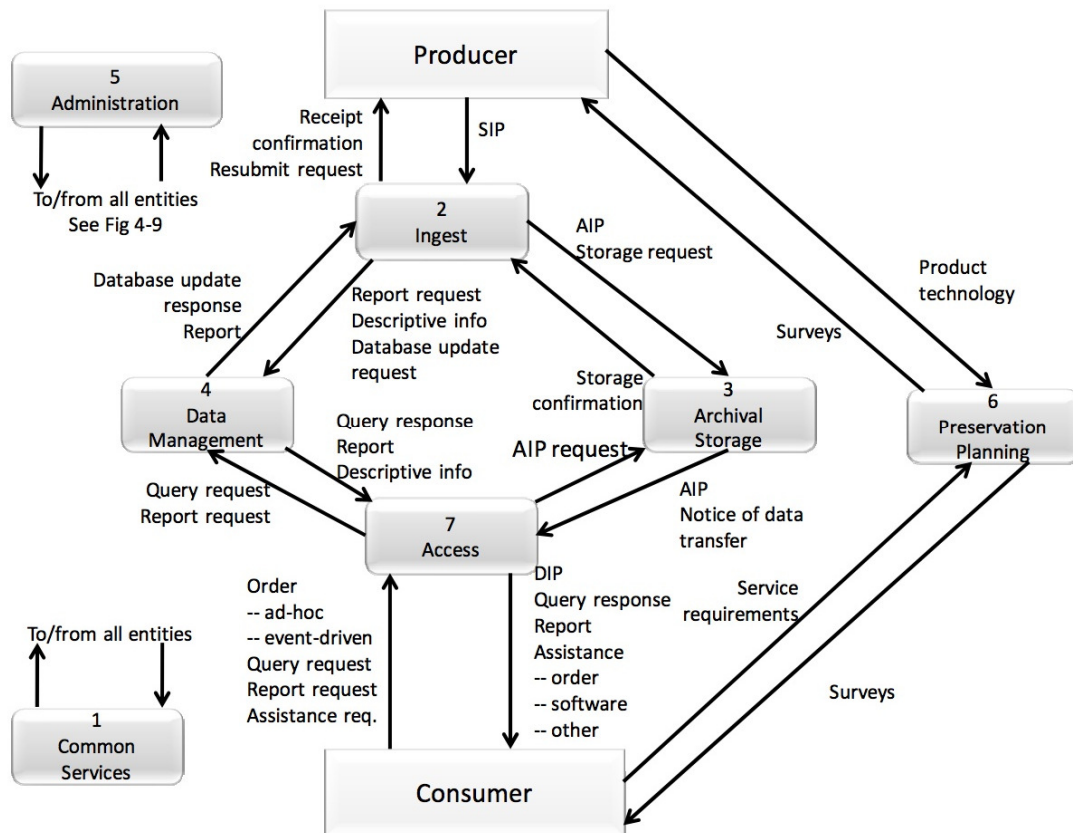


Obr. 20 Funkce modulu Administrace (*Administration*)
[ISO, 2003]

Obrázek podrobně objasňuje funkce modelu Administrace (*Administration*).
Funkce Vyjednání smluv pro odevzdávání dat (*Negotiate Submission Agreement*)

ujednává smluvní podmínky a stanovuje harmonogram odevzdávání dat (*Submission/Schedule Agreement*). Správa systémové konfigurace (*Manage System Configuration*) monitoruje archiv, systematicky kontroluje případné změny v konfiguraci a zodpovídá za snadnou ovladatelnost a integritu konfigurace ve všech fázích životního cyklu systému. Update archivní informace (*Archival Information Update*) přijímá žádosti o změny, postupy a nástroje ze Správy systémové konfigurace. Přeposílá tyto žádosti do modulu Přístup (*Access*). S updatem obsahu se změní DIP a vrací se jako SIP do modulu Příjem (*Ingest*). Kontrola fyzického přístupu (*Physical Access Control*) kontroluje fyzický i vzdálený přístup do archivu (ve smyslu systému) dle stanovených politik ve funkci Zavedení standardů a politik (*Establish Standards And Policies*). Aktivace žádostí (*Activate Requests*) aktivuje požadavky na zpřístupnění přicházející z modulu Přístup. Funkce Žádosti o audit (*Audit Submission*) ověřuje přijímané SIP (nebo AIP) podle vyjednaných podmínek a *Zákaznický servis (Customer Service)* vytváří, spravuje a maže uživatelské účty. Tam, kde je to požadováno, přijímá fakturační údaje (*Billing Info*) od modulu Přístup, zasílá faktury (*Bill*) a přijímá platby (*Payment*) od uživatelů.

Zjednodušeně možno říct, že modul Administrace představuje jakési "jádro" archivu podle OAIS (viz následující Obr. 21).



Obr. 21 Datové toky z/do modulu Administrace v archivu a externích entit modelu OAIS [ISO, 2003]

Obrázek ukazuje významné postavení modulu Administrace (5. *Administration*) v archivu OAIS. Naznačuje nejdůležitější datové toky z/do Administrace. Společné služby (1. *Common Services*) představují podporující služby systému (jako např. bezpečnost, zálohování řešení výjimek apod.) a jejich centrální správu má na starost Administrace. Z modulu Příjem (2. *Ingest*) se do modulu Administrace dostávají SIP nebo AIP (pro auditní účely). Příjem dostává z modulu Administrace požadavky na formáty dat a stanovené standardy (*Data Formating & Documentation Stds.*), auditní reporty (*Audit Report*) a v neposlední řadě upravené SIP ([*Updated*] SIP). Modul Permanentního úložiště (3. *Archival Storage*) odesílá do Administrace operativní statistiky (*Operational Statistics*) a Administrace informuje Permanentní úložiště o politikách uchování dat (*Storage Mgmt. Policies*). Správa dat (4. *Data Management*) poskytuje modulu Administrace nejrůznější typy reportů (*Report*) a informací o updatech (*Status Of Updates*). Administrace žádá Správu dat o reporty (*Report Request*), systémové a „review“ updaty (*System Updates a Rewiev Reports*) a politiky (*Polciies*). Plánování dlouhodobé ochrany (6. *Preservation Planning*) přijímá z Administrace šablony a přehled AIP/SIP (*AIP/SIP Templates a AIP/SIP Review*), návrhy pro „customizaci“ (*Customization Advice*), balíčky dat pro migraci (*Migration Package*) a další doporučení (*Recommendation*). Modul Plánování dlouhodobé ochrany na druhou stranu odesílá do Administrace schválené standardy (*Approved Stds.*), migrační plány (*Migration Goals*), inventarizační reporty (*Inventory Reports*), reporty o výkonu systému (*Performance Reports*) a komentáře a připomínky od uživatelů (*Consumer Comment*). Modul Přístup (7. *Access*) poskytuje Administraci DIP a fakturační údaje (*Billing Information*). Administrace žádá Přístup o data (*Dissemination Request*). Externí entita Producenti dat (*Producer*) vyjednávají přes Administraci podmínky smlouvy o dodávce dat (*Submission Agreements Negotiation, In*), jejich rozvrhu dodávání (*Submission Schedule Negotiation, In*) a rozklady (*Appeal*). Modul Administrace navrácí schválené smlouvy o dodávání dat (*Submission Agreement Negotiation, Out*), rozvrhu (*Submission Schedule Negotiation, Out*), retenčním právu (*Lien*) a taktéž finální podobu správy o dodávce dat (*Final Ingest Report*) do systému. Uživatelé (*Consumer*) poskytují archivu informace (*Information Request*) a plní své finanční závazky formou plateb (*Payment*). Administrace žádá od uživatelů informace (*Info Requests*) a vystavuje platební "příkazy" (*Bill*). Management ovlivňuje archiv stanoveným rozpočtem (*Budget*) a závaznými politikami (*Policies*). Ty často zohledňují reporty (*Reports*) zasílané z Administrace.

3.3.3.6 Plánování dlouhodobé ochrany

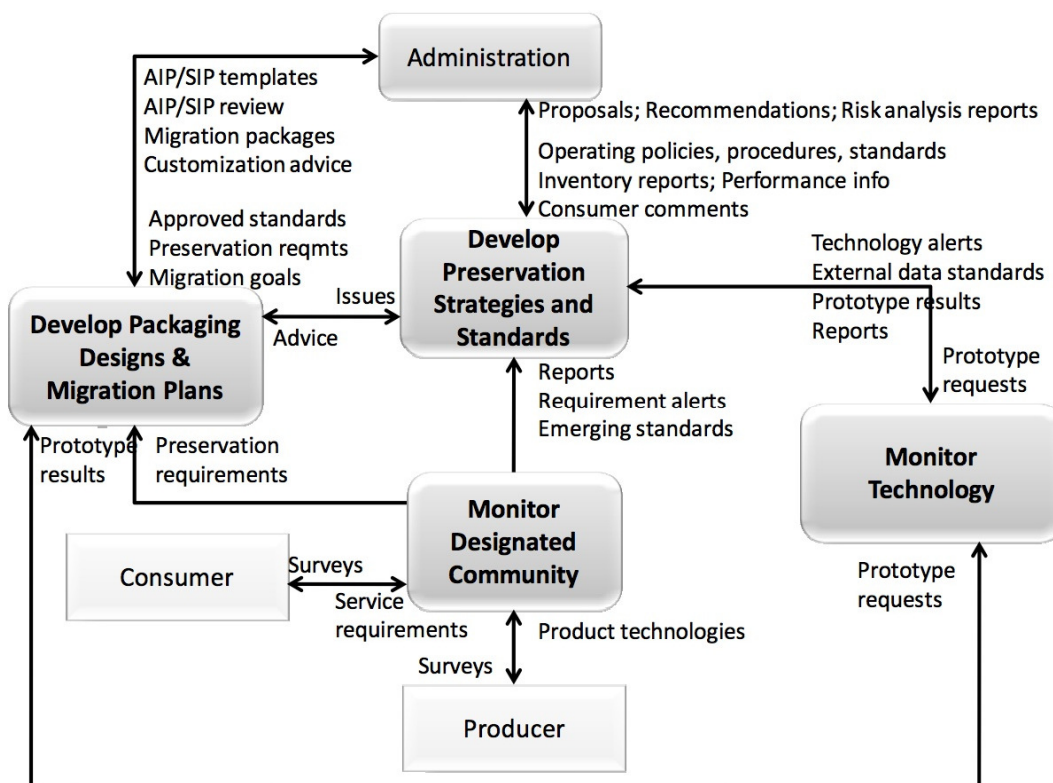
Monitorování cílové skupiny (*Monitor Designated Community*) komunikuje s uživateli archivu a producenty/tvárci dat, aby dostatečně sledovali změny požadavků na služby a technologie. Může jít např. o datové formáty, výběr médií, preference formy softwarových balíčků, nové výpočetní platformy a mechanismy pro komunikaci s archivem. Požadavky je možné zjistit pomocí průzkumů, pravidelných formálních hodnotících procesů, workshopů, kdy se vyžádá zpětná vazba

nebo se získá při individuálních interakcích. Výsledek je ve formě reportů, upozornění, standardů i požadavků na ochranu.

Sledování technologií (*Monitor Technology*) monitoruje vyvíjející se digitální technologie, informační standardy a výpočetní platformy za účelem zjišťování potenciálně zastaralých technologií/produktů v archivu. Přijímá žádosti na prototypy pro Vytvoření strategií a standardů pro dlouhodobou ochranu (*Develop Preservation Strategies And Standards*). Tato funkce též zasílá reporty, externí datové standardy, výsledky prototypování funkci Sestavení podoby balíčků a migračních plánů (*Develop Packaging Design & Migration Plans*).

Vytvoření strategií a standardů pro dlouhodobou ochranu (*Develop Preservation Strategies And Standards*) zodpovídá za vývoj a doporučení strategií a plánů pro lepší podchycení změn v požadavcích cílové skupiny nebo technologického vývoje, které si vyžádají migraci archivních jednotek nebo nový způsob odevzdávání dat.

Sestavení podoby balíčků a migračních plánů (*Develop Packaging Design & Migration Plans*) modul přijímá archivem schválené standardy (formátů, metadat a dokumentové standardy) a migrační plány z modulu Administrace. Následně tyto standardy implementuje do formy šablon pro AIP a DIP. Migrační plány přijaté do této funkce představují transformaci AIP včetně obsahové informace tak, aby nedošlo ke ztrátě zpřístupnění v důsledku zastaralých technologií. Po otestování migračního plánu zašle tato funkce migrační balíček do modulu Administrace, který naplánuje a provede vlastní migrace.



Obr. 22 Funkce modulu Plánování dlouhodobé ochrany (*Preservation Planning*) [ISO, 2003]

Obrázek popisuje funkce modulu Plánování dlouhodobé ochrany (*Preservation Planning*). Funkce Monitorování cílové skupiny (*Monitor Designated Community*) sleduje u cílové skupiny (*Consumer* a *Producer*) změny požadavků na služby (*Service Requirements*) a technologie (*Product Technologies*) a to nejčastěji formou průzkumů (*Surveys*). Monitorování vyvíjejících se digitálních technologií, informačních standardů a výpočetní platformy zabezpečuje funkce Sledování technologií (*Monitor Technology*). Pro lepší podchycení změn v požadavcích cílové skupiny nebo technologického vývoje existuje funkce Vytvoření strategií a standardů pro dlouhodobou ochranu (*Develop Preservation Strategies And Standards*). Sestavení podoby balíčků a migračních plánů (*Develop Packaging Design & Migration Plans*) přijímá archivem (ve smyslu organizace) schválené standardy (formátů, metadat a dokumentové standardy) a migrační plány z modulu Administrace. Následně tyto standardy implementuje do formy šablon pro AIP a DIP.

Kromě funkčních entit jsou pro fungování digitálních archivů nezbytné i tzv. **společné služby** (*Common Services*).

3.3.4 Společné služby

Společné služby představují podporující služby systému (jako např. bezpečnost, zálohování, řešení výjimek apod.) a je možné je zatřídit do 3 základních podtříd:

- **Služby operačního systému** (*Operating System Services*) jsou klíčové služby potřebné pro tvorbu a správu aplikační platformy a poskytují rozhraní mezi aplikačním softwarem a platformami. Jde o následující služby:
 - operace jádra (*Kernel Operations*) nabízí „low-level services“ důležité pro vytváření a správu procesů, spouštění programů, definování a komunikace signálů, definice a zpracování „system clock“ operací, správu souborů a adresářů, kontrolu I/O procesů do a z externích programů
 - příkazy a utility včetně mechanismů pro operace na operační úrovni, například porovnávání, tištění a zobrazení obsahu souborů, editování souborů, vyhledávání vzorů, vyhodnocování výrazů, logování zpráv, přesun souborů mezi adresáři, třídění dat, spuštění příkazových skriptů a správa procesů
 - systémová správa představuje schopnosti definice a správy zdrojů a přístupu (např. které zdroje jsou spravovány a třídy přístupu definovány), konfiguraci a řízení výkonnosti zařízení, souborových systémů, administrativních procesů, front, strojové profily a profily platformy, autorizaci využívání zdrojů a systémové zálohování
- **Síťové služby** (*Network Services*) nabízí mechanismy pro podporu distribuovaných aplikací poskytujících přístup k datům a zabezpečují interoperabilitu aplikací napříč síťovým prostředím. Jedná se o služby:
 - datové komunikace, které zahrnují API a specifikace protokolu pro spolehlivý, transparentní přenos dat napříč komunikačními sítěmi
 - transparentního přístupu k souborům, které se nacházejí kdekoli v heterogenní síti
 - podpory interoperability s jinými operačními systémy, obzvláště mikropočítačových operačních systémů, které nemusí být formálně specifikované v národních nebo mezinárodních standardech
 - vzdáleného volání procedury (*Remote Procedure Call*) se specifikací pro šíření lokální procedury do distribuovaného prostředí. To znamená, že procedura může být uložena na jiném místě než je umístěn sám volající program

- síťové bezpečnosti podporující přístup, autentikaci, důvěrnost, integritu kontroly a správu komunikace mezi odesílatelem a příjemcem informace v síti
- **Bezpečnostní služby** (*Security Services*) pro ochranu citlivých informací v informačním systému. Stupeň ochrany je určen hodnotou informace a hrozbou její ztráty. Definují kontrolu přístupu k systémovým datům, funkcím, hardwarovým a softwarovým zdrojům uživatele a uživatelskými procesy. Jde o služby:
 - **identifikační/autentizační**, které potvrzují identitu žadatele. Autentizační služby se mohou vyskytnout na začátku nebo v průběhu relace (*Session*)
 - **kontroly přístupu**, které zabraňují neautorizovaným uživatelům k přístupu a využívání informačních zdrojů a využívání zdrojů nepovoleným způsobem
 - **datové integrity**, jež zabezpečují, že data nejsou pozměněna nebo zničena nesprávnou manipulací. Vztahuje se zejména na data v permanentním úložišti a data v podobě komunikačních zpráv (*Communication Messages*)
 - **utajení dat**, která zajišťují, že data nebudou dostupná nebo zpřístupněná nepovolaným osobám nebo počítačovým procesům
 - **nepopíratelnosti** zabezpečují, že entity zapojené do výměny informací nemůžou odmítnout účast. Existuje nejčastěji v jedné nebo dvou formách, kdy příjemci dat je poskytnut důkaz původnosti dat (ochrana proti dodatečnému pokusu odesílatele popřít (ne)zaslání dat nebo jejich obsah) neboť odesílatel dat má důkaz přijetí dat.

Technology) a žádá o průzkum (*Surveys*) uživatele i producenty dat. Modul Přístup (7. *Access*) přijímá od uživatelů ad-hoc nebo událostmi řízené objednávky (*Orders - Ad Hoc* a *Event-driven*), dotazy (*Query Request*), požadavky na reporty (*Report Request*) a případnou žádost a pomoc (*Assistance Req.*). Přístup navrácí uživatelům DIP, výsledky zasláných dotazů (*Result Set*), reporty (*Report*) a nabízí pomoc při objednávkách, práci se softwarem a j. (*Assistance - Orders Status, Software, Other*).

Jak je vidět z výše uvedeného, všechny tyto modely přistupují k digitálním knihovnám, repozitářům a archivům odlišným způsobem. V popředí Kahn-Wilenskeho architektury, stojí digitální objekt s persistentními identifikátory, jež tvoří základ digitálních knihoven. Zajímavostí Kahn-Wilenskeho architektury je rozlišování mezi tzv. „master“ souborem (*Master File*), který můžeme považovat za archivní verzi a souborem pro zpřístupňování (*Access File*). Master soubory jsou obvykle (nikoliv nutně) výsledkem digitalizace, ve vysoké kvalitě a velkého objemu (např. TIFF skeny). Pro uživatele se zobrazuje soubor pro zpřístupňování (např. JP2 nebo JPG)⁴⁰. **Digitální objekty však nejsou jedinou podstatnou složkou digitálních knihoven, repozitářů či archivů. Významnou roli zde hrají lidé (specializovaný personál, uživatelé), zařízení, softwarové systémy, postupy, komunity apod.** DELOS Digital Library Reference Model proto definuje třívrstvou konstrukci digitálních knihoven, kterou tvoří digitální knihovna na obecné úrovni, vybraný systém digitální knihovny a systém pro správu digitální knihovny (na technické úrovni). Tento model je průlomový především definováním prostředí digitálních knihoven, ve kterých působí a dlouhodobou ochranu (*Long-term Preservation*), která zásadním způsobem ovlivňuje dostupnost (*Availability*) digitálních objektů. Tuto myšlenku dále rozvíjí Referenční model Otevřeného archivního informačního systému (OAIS), který v rámci svého modelu definuje (zejména pro repozitáře a archivy) modul pro dlouhodobou ochranu (Plánování dlouhodobé ochrany – *Preservation Planning*). Velkou předností tohoto modelu je též vymezení terminologického základu, stanovení tzv. informačních balíčků a popis funkčních entit. **V současné době se v odborných kruzích se téměř výlučnou normou pro digitální repozitáře a archivy stal právě OAIS model**

⁴⁰ Tento přístup respektuje i Digitální univerzitní repozitář UK v Praze.

3.4 Od referenčních, konceptuálních a procesních modelů k důvěryhodným digitálním knihovnám, repozitářům a archivům

Důvěra uživatelů je jedním z klíčových aspektů repozitáře jakéhokoliv typu a obsahu. Nejčastěji ji definujeme jako spolehlivost, víru či předpoklad, že osoba nebo organizace bude dodržovat rámec společných hodnot a představ. Za důvěryhodný digitální repozitář možno označit repozitář, kterého posláním je poskytovat spolehlivý a dlouhodobý přístup k organizovaným digitálním zdrojům v cílové skupině - dnes i do budoucna⁴¹. Takový repozitář porozumí hrozbám i rizikům správy digitálních informací. V dalších částech práce jsou představeny i kritéria certifikace digitálních repozitářů (podle Evropského rámce pro audit a certifikaci repozitářů), jež jsou založeny na posouzení schopnosti dlouhodobě a spolehlivě uchovávat a zpřístupňovat autentické digitální dokumenty. Certifikát je v tomto ohledu „jistotou“ cílových skupin, které s důvěrou odevzdávají svá data do digitálních repozitářů. Slouží ale i jako objektivní (auditní) nástroj a výchozí bod pro hodnocení veškerých procesů a činností vztahujících se k digitálním objektům v repozitáři.

Důležité je též prozkoumat vlastní životní cyklus digitálních dokumentů. Ten totiž zasahuje do organizační infrastruktury, administrace systému repozitáře a managementu rizik. Postupy správy digitálních objektů, technologický rámec a zabezpečení dat musí být na rozumné úrovni a adekvátní úkolům a cílům repozitáře. Identifikovaná kritéria vychází především z referenčního modelu OAIS.

Takový repozitář může být jak lokálního charakteru, spravovaný interními zdroji dané instituce, či pomocí externích zdrojů (*Outsourcing*). Avšak bez ohledu na typ architektury a způsob uchování a zpřístupňování musí tyto repozitáře [HUTAŘ, 2008; RESEARCH, 2009]:

1. zavázat se k trvalé správě/ochraně digitálních objektů pro definovanou komunitu/-y
2. prokázat organizační způsobilost pro tento úkol (tj. vhodné financování, personální zajištění, vhodné procesy)
3. dostatečně demonstrovat fiskální zodpovědnost a udržitelnost
4. dostát smluvním a právním požadavkům a splnit své povinnosti v této oblasti

⁴¹ V roce 1996 RLG/CPA publikovali společné prohlášení o důvěře v digitálních archivech: „Pro zabezpečení dlouhé životnosti informací, možná tu nejdůležitější roli při operacích digitálních archivů hraje udržování identity, integrity a kvality samotného archivu jako důvěryhodného zdroje kulturních záznamů. Uživatelé archivovaných informací v elektronické podobě a archivních služeb na ně se vážících musí mít jistotu, že digitální archivy jsou tím, čím prohlašují, že jsou a že informace uchovávané v nich jsou v bezpečné formě i do budoucna“ (Trusted, 2000).

5. mít vypracován účelný a účinný strategický rámec
6. získávat a zpracovávat digitální objekty podle stanovených kritérií, která odpovídají jeho cílům a schopnostem
7. udržovat a zajišťovat dlouhodobou integritu, autenticitu a použitelnost spravovaných digitálních objektů
8. archivovat potřebná metadata o všech akcích, které byly s digitálními objekty v průběhu skladování provedeny; související informace o vzniku, podpoře dostupnosti a využívání objektů před jejich vstupem do repozitáře
9. splnit potřebná kritéria pro zpřístupňování
10. mít strategický program pro plánování ochrany
11. mít odpovídající technickou infrastrukturu, potřebnou k trvalému udržování a zabezpečení spravovaných digitálních objektů.

Tato kritéria představují klíčové charakteristiky důvěryhodnosti repozitářů, jež je možné shrnout do sedmi základních atributů důvěryhodných repozitářů.

3.4.1 Atributy důvěryhodných repozitářů

Konformita s referenčním rámcem OAIS (*OAIS Compliance*) - tento atribut je implicitní modelu důvěryhodných digitálních repozitářů (viz Obr. 24 na str. 82). Z organizačního hlediska musí repozitáře dodržovat zavedené politiky a procedury, které vychází z OAIS⁴². Po technologické stránce je nutné přetavit funkční entity modelu OAIS do plně fungujícího repozitáře. Finanční aspekt shody s OAIS vyžaduje důkladné plánování nákladů a tím pádem zabezpečení adekvátních udržitelných finančních prostředků.

Administrativní odpovědnost (*Administrative Responsibility*) - dodržovat národní a mezinárodní standardy, vyhledat rady expertů a širší komunity pro pravidelnou validaci a certifikaci zvolených a zavedených procesů a standardů. Spolehlivý repozitář je zavazuje ke transparentnosti a zodpovědnosti při veškerých svých aktivitách.

Organizační životaschopnost (*Organizational Viability*) - atribut zdůrazňuje potřebu prokázání životaschopnosti a důvěryhodnosti. Důležitý je i závazek k dlouhodobému uchování v prohlášeních o poslání a odpovídající právní status. Neméně důležité jsou i transparentní

⁴² Referenční model OAIS nabízí společný rámec (včetně terminologie a konceptů) pro popis architektury a operací digitálních archivů/repozitářů. Je to funkční model (popisující jednotlivé specifické úkoly, které repozitář provádí) a k němu se vztahující informační model (popisující tvorbu metadat pro dlouhodobou ochranu a přístup).

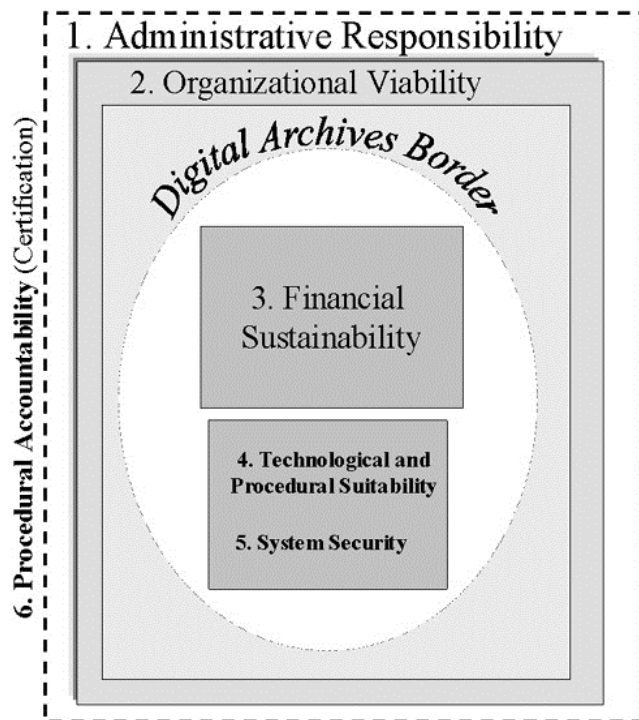
„business“ praktiky, efektivní manažerské politiky, přezkoumání a dodržování politik a procedur, správa rizik (*Risk Management*), podchycení nepředvídatelných událostí (*Contingency Planning*) a plánování případného nástupnictví (*Succession Planning*) důvěryhodným „dědicům“. Nutné je myslet i na zaměstnance, zabezpečení jejich profesního vývoje a obsažný písemný souhlas s vkladateli dat o možnosti spracování a/nebo dlouhodobém uchovávání dokumentů. Top-level podpora a závazek v administrativní zodpovědnosti je stěžejním faktorem pro kvalitní program dlouhodobé ochrany.

Finanční udržitelnost (*Financial Sustainability*) - zavedení a udržení vhodných „business“ praktik a kontrolovatelného „business“ plánu, vyvažování rizik, výhod, investic i výdajů, demonstrace finančního zdraví a trvajících finančních závazek. Tomu musí odpovídat adekvátní rozpočet a rezervy. Nesmí se podceňovat aktivní vyhledávání finančních zdrojů. Finanční udržitelnost byla až do nedávna nejvíce podceňovaný atribut, ale díky nejrůznějším projektům se dostává do pozornosti i po skončení projektů. Ty jsou důležitým začátkem v oblasti dlouhodobé ochrany.

Technická a procedurální vhodnost (*Technological and Procedural Suitability*) - výběr a zavedení nejvhodnější strategie pro dlouhodobou ochranu digitálních dokumentů, zajištění vhodné infrastruktury (hardware, software a zařízení) pro získávání, uchovávání a zpřístupňování. Dále sestavení politiky správy technologií repozitáře (nahrazení, posílení, financování), shoda s relevantními standardy a „best practices“ (podpořených adekvátní expertízou), jako i podstoupení externího auditu pro vyhodnocení systémových komponentů a jejich výkonu. Jde o atribut, kde se přesouvají organizační a technologické zodpovědnosti. Organizační zodpovědnost je poskytnout informace o možných technických problémech a manažerským úkolem je zabezpečit, aby tyto byly vyřešeny dle zavedených procedur a obecně platných standardů.

Systémová bezpečnost (*System Security*) - repozitář má v písemné podobě politiky a plány postupů v případě náhlých pohrom nejrůznějšího typu (od finančních potíží až po přírodní katastrofy). Rozpracovaný je i plán obnovy dat a zaměstnanci jsou patřičně proškolení. Zvláštní pozornost je v tomto ohledu zaměřena na integritu dat tak, aby nedošlo k jejich ztrátě, dále na zachycení změn v datech a na schopnost opravy poškozených dat či obnovy (ze záloh) těch ztracených.

Procedurální zodpovědnost (*Procedural Accountability*) - zodpovědnost za množství vzájemně souvisejících úkolů a funkcí. Praktiky repozitáře jsou dokumentovány a mohou být zpřístupněny na požádání. Mechanismy pro zpětnou vazbu jsou k dispozici pro řešení problémů a vyjednávání budoucích požadavků mezi repozitářem, poskytovatelem služeb třetích stran a cílovou komunitou.



Obr. 24 Model důvěryhodného digitálního repozitáře
 (<http://quod.lib.umich.edu/s/spobooks/images/bbv9812.0001.001-00000010.jpg>)

Obrázek přehledně ilustruje model důvěryhodného digitálního repozitáře, jeho atributy a jejich vzájemnou provázanost. Výchozí podmínkou důvěryhodného repozitáře je jeho shoda s OAIIS referenčním modelem. Dalšími důležitými atributy jsou: administrativní odpovědnost (1. *Administrative Responsibility*), organizační životaschopnost (2. *Organizational Viability*), finanční udržitelnost (3. *Financial Sustainability*), technická a procedurální vhodnost (4. *Technological and Procedural Suitability*), systémová bezpečnost (5. *System Security*) a procedurální zodpovědnost (6. *Procedural Accountability Certification*). Bezproblémové fungování digitálních archivů (*Digital Archives Borders*; ekvivalent repozitáře prezentovaný v jiných modelech) bezprostředně ovlivňuje finanční udržitelnost, technickou a procedurální vhodnost a systémovou bezpečnost. Digitální archiv může být spravován více než jednou institucí a jedna instituce může spravovat jeden a víc repozitářů. Na digitální repozitář jako organizaci působí administrativní odpovědnost, kdy se instituce zavazuje k vytvoření důvěryhodného repozitáře. Organizační životaschopností demonstruje instituce/organizace své poslání, právní status a aktivity směřující k důvěryhodnému repozitáři. Procedurální zodpovědnost znamená, že organizace kompletně dokumentuje veškeré preservační aktivity a dodržuje schválené politiky, praktiky a procedury.

3.4.2 Faktory zodpovědných důvěryhodných repozitářů

Organizace spravující repozitář potřebuje porozumět vlastním, lokálním požadavkům a musí znát své zodpovědnosti. Zodpovědný digitální repozitář, mimo jiné, vyjednává data v požadované

podobě, akceptuje je od producentů dat a držitelů práv. Musí mít dostatečnou kontrolu nad informacemi podporujícími dlouhodobou ochranu. Repozitář sám nebo ve spolupráci se zúčastněnými stranami (cílové skupiny uživatelů, producenti dat apod.) zabezpečuje, že informace, které uchovává budou „nezávisle srozumitelné“ (*Independently Understandable*) cílové skupině, tj. že daná komunita uživatelů porozumí daným informacím bez asistence expertů. Samozřejmě je dodržování implementovaných politik a jim odpovídajících procedur tak, aby obsah repozitáře byl chráněn před případnými nepředvídatelnými událostmi. Ochrana má svůj protipól v podobě zpřístupňování uchovávaných informací cílové skupině a repozitář úzce spolupracuje s cílovou skupinou pro potřebnou zpětnou vazbu. Výslednou podobu důvěryhodných a zodpovědných repozitářů však ovlivňují zejména tyto faktory [RESEARCH, 2000]:

- **rozsah sbírek** (*Scope Of Collections*) - pro materiály, které mají své fyzické protějšky, rozhodnutí o dlouhodobé ochraně ovlivňuje stav originálních materiálů a jejich důvod pro digitalizaci. U materiálů, jež vznikly v digitální podobě (*Born-digital*) existuje vyšší riziko ztráty a je pak kromě samotného obsahu nutné vzít v potaz i standardy, formáty, kontrolu kvality, integrity a dostupnou dokumentaci
- **management dlouhodobé ochrany a životního cyklu dat** (*Preservation And Lifecycle Management*) - dlouhodobá ochrana vyžaduje aktivní správu, která začíná při tvorbě materiálů a závisí na proaktivním přístupu digitálních repozitářů a za spolupráce všech zúčastněných stran, včetně producentů, resp. poskytovatelů dat
- **různorodost zúčastněných stran** (*Wide Range Of Stakeholders*) - tvůrci obsahu, vývojáři systémů, správci sbírek či budoucí uživatelé představují potenciální zúčastněné strany při dlouhodobé ochraně digitálních materiálů, čímž se značně komplikuje rozdělení zodpovědností - kdo, kdy, na jak dlouho? Např. u komerčních vydavatelů a nakladatelství není přílišný zájem o dlouhodobou ochranu kvůli obavám ze snížení tržeb a na druhé straně knihovníci a cílové skupiny uživatelů požadují dlouhodobý přístup. Proto rozhodnutí o tom, jaké materiály a jak budou uchovávány v repozitářích je nutné řešit již na začátku životního cyklu dat, v době jejich vytvoření.
- **vlastnictví materiálů a jiné právní otázky** (*Ownership Of Material And Other Legal Issues*) - u digitálních materiálů nejsou odpovědi na otázky ohledně vlastnictví snadno řešitelné. To se týká např. elektronických zdrojů a přístupu k nim i po vypršení licence. Často se práva vztahují na software a systémy, kterými dané materiály byly vytvořeny⁴³.

⁴³ V tomto ohledu existuje jen velmi nízké povědomí, jak na straně institucí, které dané materiály uchovávají, tak na straně tvůrců dat, kteří v daných systémech příslušné materiály tvoří. V neposlední řadě podstatnou roli sehrávají i samotní dodavatelé softwaru. O tom, že se jedná o dost neprobádanou, ale zároveň důležitou oblast, svědčil i workshop na mezinárodní konferenci iPRES 2013 - *10th International*

Dlouhodobá ochrana má však širší právní důsledky. Do jaké míry dlouhodobá ochrana překračuje autorské právo není zcela známo. Například, tvůrce obsahu obvykle nemá práva na software a systémy, kterými dané digitální soubory vytváří. V případě nutných změn pak vyvstává otázka, jak takové případy co nejvhodněji, přitom legálně a při zachování práv, vyřešit⁴⁴.

- **dopady nákladů** (*Cost Implications*) - podle [RESEARCH, 2000] si ochrana digitálních materiálů vyžádá mnohem více prostředků a zdrojů než v případě tradičních, tištěných materiálů. Zatímco pro tištěné dokumenty jsou dostupné prostředky při stabilních podmínkách k jejím uchování, u digitálních dokumentů je nutná periodická analýza a aplikace nových strategií a technologií pro zachování přístupu i do budoucna. Existují 4 vzájemně související faktory, jež významně ovlivňují náklady na dlouhodobou ochranu [DIGITAL, 2008]:
- potřeba aktivně řešit přicházející změny v technologiích v pravidelných intervalech a po neomezený časový rámec
- nedostatek standardizace u samotných zdrojů a licenčních ujednáních s vydavateli a jinými producenty dat, čímž je těžké splnit „úspory z rozsahu“⁴⁵ (*Economies of Scale*)
- nevyřešené prostředky spolehlivé a přesné interpretace (*Rendering*) některých digitálních publikací bez ztráty jejich významných informací (v případech změny technologií)
- digitální ochrana bude vyžadovat vyšší náklady než u tradičních materiálů.

Conference on Preservation of Digital Objects (Lisabon, Portugalsko). Workshop *Open Research Challenges in Digital Preservation* nastínil 4 okruhy problémů dlouhodobé ochrany, které nabízí možný prostor pro další bádání: emulační strategie, dlouhodobá ochrana jako věda, (polo)automatizované hodnocení kvality digitalizovaných obrazových materiálů a v neposlední řadě právě udržitelné výpočetní systémy [FOJTŮ, 2013].

⁴⁴ U nás tato situace není zatím kritická, nicméně zkušenosti zahraničních institucí ukazují, že dodatečné získávání práv může ohrozit jak zpřístupňování, tak uchovávání dokumentů. Nezřídka se pak stává, že materiály musí být z repozitářů, resp. systémů pro dlouhodobou ochranu odstraněny.

⁴⁵ V případě úspor z rozsahu výroby klesají průměrné náklady výrobků s růstem výroby.

4. Dlouhodobá ochrana digitálních objektů

Není tomu tak dávno, kdy ochrana digitálních dokumentů byla pojmána čistě z technokratického hlediska, tedy jako změna dokumentů pomocí kopírování datových toků, migrace, obnovení (*Refreshing*), přesunu dat na média (CD, DVD a pod.), analógového zálohování (např. vytištěním informačních zdrojů na papír), emulace (napodobení systému zpracování údajů nebo jeho části jiným systémem nebo technickými prostředky) a v neposlední řadě UVC (*Universal Virtual Computer*)⁴⁶.

S narůstajícím objemem digitálních a digitalizovaných dat vznikla potřeba definice nového oboru dlouhodobé ochrany digitálních dokumentů (*Long-Term Preservation Of Digital Documents*). Nemělo by však jít jen o samoučelnou archivaci, ale o skutečnou ochranu, zpřístupnění a šíření. Správcové digitálních sbírek a systémů (rozsáhlých) digitálních knihoven proto postupně obracejí pozornost na skutečnost, že dlouhodobé uchovávání a ochrana není jen o zachovávání fyzického datového toku, ale i o schopnosti prezentovat obsah na všech úrovních. Cílem je tedy zabezpečit (pravidelnými kontrolami) jak fyzickou, tak i logickou integritu dat.

Repozitář je důvěryhodný pokud splňuje předem definované funkce (alespoň s minimálním počtem stanovených kritérií). Plnění těchto funkcí, resp. kritérií musí být průkazné. To v praxi znamená, že dosažení důvěryhodnosti je do značné míry závislé na auditu a certifikaci.

Často publikovaným názorem je, že dlouhodobou ochranu dokumentů je možné vyřešit pouze metadatovým popisem a uchováním souborů na (optická či pevná) média. Není tomu však tak. V dnešní době již problematika dlouhodobé ochrany digitálních dat není pouze otázkou správně zvolených technologií. Zahrnuje i celou řadu dalších problematických okruhů - od organizace a řízení, kvalifikovaného personálu až po finanční stránku. Pod ochranou digitálních dokumentů tedy rozumíme soubor činností a zásahů, které jsou nevyhnutné pro zabezpečení trvalého a spolehlivého přístupu k digitálním dokumentům bez ohledu na čas [KNIGHT, 2009].

Obecně existují 4 základní přístupy k dlouhodobé ochraně digitálních dokumentů [WILSON, 2008]:

- **technokratický** (*Techno-centric*) - ochrana HW a SW po co nejdéle možnou dobu:
- **datový** (*Data-centric*) - ochrana dat na úkor originální aplikace, např. konverze dokumentů do PDF,

⁴⁶ Myšlenka UVC, tedy jakéhosi univerzálního virtuálního počítače se zrodila počátkem roku 2000 a byla otestována pro potřeby dlouhodobé ochrany JPEG v Národní knihovně Nizozemska. UVC je postaven na principu dekodérů formátů souborů a programů. Podrobněji na: <http://www-935.ibm.com/services/nl/dias/cs/uvc.html>.

- **procesní** (*Process-centric*) - ochrana originální aplikace a procesů, např. emulace starého HW nebo SW
- **“post hoc”** (*Post Hoc*) - bez aktivní ochrany; v případě potřeby data „archeologickými“ metodami vyfiltrovat⁴⁷.

Obr. 25 Životní cyklus/politika digitální ochrany
[FOJTŮ, 2011]



Obrázek popisuje, jakým způsobem má fungovat životní cyklus dlouhodobé ochrany v institucích spravujících repozitáře. Na začátku sestavování interního plánu, resp. politiky dlouhodobé ochrany je nutné vyhodnotit současnou situaci (Zhodnocení). Pokud repozitář ještě neexistuje, je možné vybrat vhodný nástroj (Výběr) a otestovat jej (Testování). Poté objektivně vyhodnotit auditem, externím nebo interním samo-auditem (Audit) a sepsat patřičnou dokumentaci (Dokumentace). Pokud se objeví případné nedostatky v politice, je nutné ji upravit (Úpravy/doplnění). Po určitém čase či zásadnější změně začíná hodnocení od začátku (Zhodnocení).

Je dobré si uvědomit, že **cílem dlouhodobé ochrany není uchovávat všechno, nýbrž stanovit pravidla a podmínky za jakých se vybrané dokumenty budou uchovávat.**

Při vytváření plánu, resp. interní politiky dlouhodobé ochrany digitálních dokumentů je vhodné využít postup dle Obr. 25 (na str. 88). Na začátku je potřebné zmapovat výchozí situaci repozitáře, pokračovat výběrem nejvhodnějších formátů, postupů, SW, HW, testováním

⁴⁷ Datová archeologie (*Data Archeology*) představuje disciplínu zabývající se obnovou počítačových dat, jež byla vytvořena v zastaralých formátech nebo uchována na přežitých médiích. Poprvé se termín objevil v roce 1993 v souvislosti s obnovou počítačových záznamů o klimatických podmínkách. Ty byly klíčové pro vyhodnocování teorií klimatických změn. Více na: http://en.wikipedia.org/wiki/Global_Oceanographic_Data_Archaeology_and_Rescue_Project.

vybraných řešení až po audit a eventuelně certifikaci. S tím tedy souvisí i porozumění životního cyklu digitálních dokumentů.

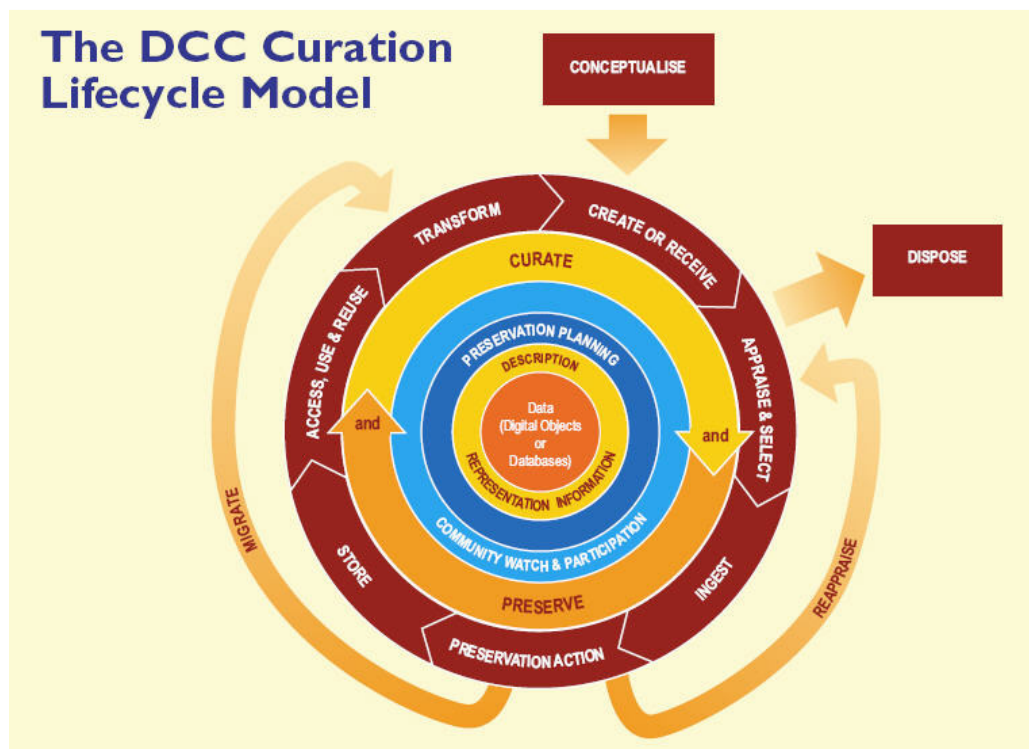
4.1 Životní cyklus digitálních dokumentů

Jednotlivé fáze životního cyklu digitálních dokumentů (objektů) definujeme podle klíčových oblastí problematiky. První krok představuje tvorbu digitálních dokumentů (primární – původní digitální dokument a sekundární – digitalizovaný dokument). V současné době se posouváme od problémů souvisejících s technologickým řešením (retrospektivní) digitalizace ke kvalitativním problémům digitalizace (součástí kterých je výběr materiálů pro transformaci do digitální podoby). Následně k dalším fázím životního cyklu digitálních dokumentů - zpracování (výběr nejvhodnějšího metadatového popisu, evidence a katalogizace), zpřístupňování (dostupnost a přístupnost), šíření a dlouhodobá ochrana (*Long-term Preservation*) digitálních dokumentů.

4.1.1 DCC model životního cyklu dlouhodobé ochrany

DCC model životního cyklu uchování (*DCC Curation Lifecycle Model*) zastupuje kategorii grafických modelů dlouhodobé ochrany. Představuje ideální model organizace či konsorcia se všemi fázemi a kroky, které je nutné při dlouhodobé ochraně prosadit a implementovat. Model je dostatečně granulární a je možné začít s aktivitami dlouhodobé ochrany v kterékoli fázi jejího cyklu⁴⁸.

⁴⁸ Podrobnější a velmi přehledný popis DCC modelu lze najít v online prezentaci Pavly Švástové [ŠVÁSTOVÁ, 2014].



Obr. 26 Životní cyklus dlouhodobé ochrany podle DCC
 (<http://www.dcc.ac.uk/sites/default/files/documents/publications/DCCLifecycle.pdf>)

Jak je vidět na Obr. 26 klíčovým prvkem životního cyklu podle DCC modelu jsou data, resp. digitální objekty či databáze ((Data Digital Objects Or Databases) - oranžový kruh). Data, jako libovolné informace v binární podobě tvoří jednoduché až komplexní digitální objekty (textové, obrazové soubory, společně s jejich identifikátory a metadaty). Pro tyto digitální objekty musí existovat dostatečný popis (informace formou reprezentace) pomocí popisných, technických, strukturálních, administrativních či ochranných metadat (žlutý kruh). Díky plánování a specifickému plánu dlouhodobé ochrany mají takto popsané objekty větší šanci k permanentní archivaci (tmavě modrý kruh). Konkrétní preservační aktivity a nástroje se díky spolupráci v komunitě mohou vylepšovat a snadněji obměňovat (středně světlý modrý kruh). Tyto výše uvedené aktivity se prolínají celým životním cyklem a podporují je následující sekvenční aktivity: tvorba či příjem (*Create Or Receive*), výběr (*Appraise And Select*), import (*Ingest*), ochranné aktivity (*Preservation Action*), uložení (*Store*), zpřístupnění či využití (*Access, Use And Reuse*) a případná transformace (*Transformation*), která pomyslně uzavírá cyklus (červený kruh) a eventuálně začíná celý proces nanovo.

4.2 Platter – Plán důvěryhodného digitálního repozitáře

Problematika dlouhodobé ochrany digitálních dokumentů by se měla objevit již v prvotních fázích plánování repozitáře. Pro instituce, které ještě nemají repozitář, je ideálním řešením Plán důvěryhodného digitálního repozitáře – PLATTER⁴⁹. Vznikl v rámci projektu DPE (*Digital Preservation Europe*). Ten má za cíl popularizovat problematiku dlouhodobé ochrany digitálních dat v širší odborné veřejnosti a koordinovat spolupráci jednotlivých výzkumných záměrů v této oblasti.

PLATTER není technickým návodem pro programátory, kteří by chtěli budovat důvěryhodný digitální repozitář. Je spíše prakticky orientovaným průvodcem pro manažery či pracovníky zodpovědné za digitální data. Ukazuje rozsah problematiky a upozorňuje na možná rizika a problémy důvěryhodnosti digitálního repozitáře⁵⁰ (*Trust*, či v gramaticky chybných tvarech *Trustedness* a *Trustworthiness*). Největší výhodou PLATTER je identifikace silných a podchycení slabých stránek (v ideálním případě) budoucího, eventuálně současného repozitáře. Za slabé stránky PLATTER možno považovat především neschopnost aplikace veškerých požadavků na jakýkoliv typ repozitáře a v některých ohledech až příliš všeobecná doporučení. To by ale nemělo odpovědné osoby odradit od hledání vlastních řešení a dalších postupů při plánování dlouhodobé ochrany digitálních dokumentů.

PLATTER procesy vychází ze skupiny Strategických objektivních plánů (*Strategic Objective Plans* - SOPs), pomocí kterých definuje repozitář své cíle a klíčové indikátory výkonu, jež posléze vytváří důvěru uživatelů (viz Obr. 27 na str. 91). Záměrem je, aby tyto strategické plány fungovaly jako „živé“ dokumenty, které se mění společně se změnami v repozitáři.

V první fázi PLATTER představuje klasifikaci vlastního repozitáře pomocí 4 popisných tříd (*Strategic Planning*), které mají širší organizační záběr, cíle i funkce a vytváří podklad pro detailnější plánování [ROSENTHAL, 2009]:

- **účel a funkce** - typ repozitáře z funkčního hlediska, mandát, pověření, provozní vyzrállost, právní podmínky získávání obsahu apod.
- **velikost** - ve smyslu lidských zdrojů, financí, dat apod.
- **provoz** - typ a zdroj dat, citlivost dat, metody akvizice, oprávnění přístupu apod.
- **technické řešení a implementace** - zdroj metadat, standardy interoperability, strategie ukládání, strategie softwarové podoby (*Software Management*) apod.

⁴⁹ <http://www.digitalpreservationeurope.eu/platter.pdf>

⁵⁰ Repozitář z pohledu PLATTER představuje organizaci odpovědnou za uchování digitálních dat. Nejde tedy jen o technologii samotnou. Uchování představuje časový rámec přesahující možnosti současných technologií [HUTAŘ, 2008].

V dalších fázích, po zodpovězení těchto důležitých strategických otázek, je možné přistoupit k vytvoření strategických cílů (*Goal Expression*). Ty vychází z 9 základních principů důvěryhodného repositáře⁵¹:

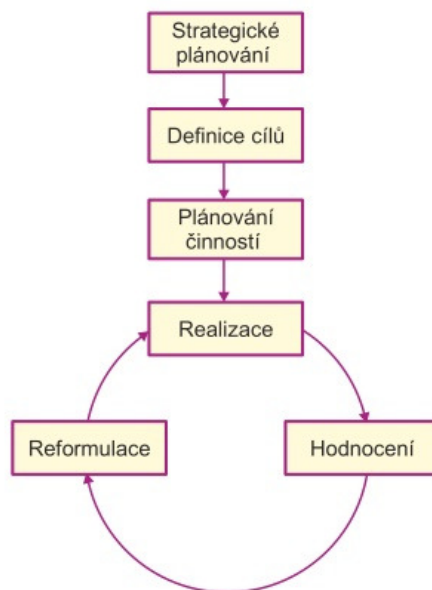
1. Finanční plán
2. Plán řízení lidských zdrojů
3. Datový plán
4. Akviziční plán
5. Plán zpřístupňování
6. Plán ochrany
7. Technický plán
8. Plán zajištění kontinuity
9. Krizový plán.

Všechny tyto cíle odpovídají tzv. SMART požadavkům⁵². Na základě SMART cílů lze vytvořit základ pro realizaci hodnocení cílů (*Realization*), kontrolu průběhu (*Review*) a průběžnou úpravu (*Reformulation*), tedy reformulaci cílů a následné vyhodnocení výsledků. Jde o cyklický proces, který do značné míry odpovídá analýze rizik DRAMBORA⁵³ [ROSENTHAL, 2009].

⁵¹ Viz Kapitola 4 – Dlouhodobá ochrana digitálních objektů

⁵² SMART požadavky představují akronym prvních písmen anglických termínů Specific, Measurable, Assignable, Realistic, Time-Related. Požadavky tedy musí být dostatečně specifikované, měřitelné, realistické, časově vymezené a s konkrétní odpovědností.

⁵³ viz Kapitola 5.1.3.1 DRAMBORA - Digital Repository Audit Method Based on Risk Assessment.



Obr. 27 Plánovací cyklus PLATTER
[ROSENTHAL, 2009]

4.3 PLATO Preservation Planning Tool

Dalším užitečným nástrojem pro digitální repozitáře, knihovny a archivy může být PLATO Preservation Planning Tool⁵⁴. Ten staví svoji funkcionalitu na analýze užitek. Podporuje rozhodování v oblasti plánování dlouhodobé ochrany digitálních dat a výběr nejvhodnějšího formátu pro budoucí migrace. Hlavním principem webového nástroje Plato je vyhodnocování ochranných a archivačních aktivit podle stanovených cílů, signifikantních vlastností a požadavků na jejich zachování. Využívá webového rozhraní a integruje online služby poskytované třetími stranami (např. online formátové registry). Vychází z referenčního modelu OAIS, konkrétně jeho funkční entity Plánování dlouhodobé ochrany⁵⁵. Jde o online softwarovou aplikaci (licencovanou jako CC-GNU LGPL) pro evaluaci potenciálních ochranných řešení a strategií. Samotnému testování musí předcházet **stanovení signifikantních vlastností** u testovaných typů objektů [FOJTŮ, 2011].

⁵⁴ Garant: Department of Software Technology and Interactive Systems, Vienna University of Technology (<http://www.ifs.tuwien.ac.at/>); Instalace: online nástroj, otevřené porty 8080 a 8443; Poslední verze: 4.2 (červen 2013).

⁵⁵ Mezi její základní funkce patří hodnocení obsahu repozitáře, sledování změn ve vývoji technologií, pravidelná doporučení pro migraci, vytváření a navrhování ochranných opatření, sledování požadavků cílové skupiny [CUBR, 2010].

4.3.1 Signifikantní vlastnosti

Termín signifikantní vlastnosti v kontextu dlouhodobé ochrany digitálních dat (significant properties, significant characteristics, transformational information properties nebo také essence) byl poprvé zaveden v projektu CEDARS⁵⁶. Vyjadřuje důležité charakteristiky (komponenty) informačního objektu, které mají být uchovány tak, aby jeho použitelnost, přístupnost a srozumitelnost byla zabezpečena i do budoucna [WILSON, 2007]. Správci sbírek mohou využít i formální vyjádření signifikantních vlastností pro výběr ochranných technik a načasování různých preservačních akcí.

Signifikantnost má několik **základních aspektů** [KNIGHT, 2009]:

- **relativita** (*Relativity*), jež není univerzální a neměnná
- **závislost** (*Dependency*) na cílech a aplikovaných kritériích
- **interpretace** (*Interpretation*) dle cílové skupiny a její znalostní báze apod.

Formální vyjádření významných vlastností je rekurzivní, tj. mnoho z vlastností lze dále dělit na dílčí vlastnosti. Vlastnost písmo, například, může být dále rozdělena na dílčí vlastnosti, jako je typ písma, styl, velikosti nebo barva. Pokud některé z těchto dílčích vlastností pozmění vzhled nebo význam digitálního objektu, mohou být považovány za důležité a stojí za to je zachovávat v přesné shodě jako u originálního objektu. Je dobré si uvědomit, že například jen formát konverze může změnit strukturu dokumentu, jeho stylistické rysy, navigaci či vzhled.

⁵⁶ CEDARS - CURL Exemplars in Digital Archives (<http://www.ukoln.ac.uk/services/elib/projects/cedars/>). Jelikož šlo o iniciativu v oblasti vyššího vzdělávání (*Higher Education*) a financování probíhalo z *Joint Information Systems Committee* (JISC) v programu *Electronic Libraries (eLib) Programme*, byl projekt zaměřen na otázky vztahující se k dlouhodobé ochraně a přístupu k digitálním materiálům zejména ve vědeckých a univerzitních knihovnách. Nejznámější výstup projektu představuje metadatová specifikace *Cedars Preservation Metadata*.

Kategorie	Požadavek	Signifikantnost
Kontext	organizační kontext (např. název organizace, datum, vazba na jiné dokumenty apod.)	uchování logových souborů s informacemi o preservačních akcích, o originálu a současném formátu souboru, názvu a verzi HW, SW, OS
Obsah	veškerý obsah musí být uchován, včetně záhlaví, zápatí, poznámek pod čarou, obsahu (souhrn) a vlastností dokumentu	prostý text musí být vždy čitelný
Struktura	struktura dokumentu musí být uchována tak, aby reprezentovala logické vazby mezi jednotlivými částmi dokumentu	pořadí kapitol, odstavců, ale též správná pozice vložených poznámek, poznámek pod čarou a obrázků
Chování	aktivní chování, „update“, záznam o původních formách chování musí být zaznamenán	popis aktivních odkazů musí být uchován
Vzhled	vzhled originálu a uchovávané verze nemusí být identické, nicméně „nový“ vzhled nesmí ovlivnit pochopení originálního významu	k subjektivnímu vyhodnocení

Tab. 1 Požadavky a signifikantní vlastnosti textového dokumentu
[SLATS, 2004]

Stanovování signifikantních vlastností však není podchyceno jednotnou metodikou⁵⁷. Koncepty signifikantních informací se do určité míry překrývají s reprezentacemi informací. Signifikantní vlastnosti představují charakteristiky abstraktního informačního objektu, naproti tomu reprezentace informací (strukturní a sémantické informace) indikují charakteristiky objektu z pohledu dat, jakou je např. kódování [BROWN, 2008].

Základní požadavky na dlouhodobou ochranu je možné rozdělit do pěti základních okruhů, resp. kategorií: **kontext** (*Context*), **obsah** (*Content*), **chování** (*Behavior*), **vzhled** (*Appearance*) a **struktura** (*Structure*). Z nich se poté „extrahují“ signifikantní vlastnosti. Ve výsledku pak například požadavky na dlouhodobou ochranu textového dokumentu a (slovní) vyjádření jeho

⁵⁷ Dollar navrhuje 8 výchozích kritérií: čitelnost, srozumitelnost, identifikovatelnost, zapouzdřenost, vyhledatelnost, transformovatelnost a autenticitu. Jiný přístup představuje tzv. Model důvěry (z angl. Model of Trust) postavený na autenticitě, aktuálnosti, přesnosti, pokrytí, kredibilitě, objektivitě, validitě, předvídatelnosti a věrohodnosti [KORENKOVÁ, 2011]. Knight [2009] identifikuje složení digitálního objektu (formu jakou je idea vyjádřena), účel (zamýšlené funkce, typ uživatele), organizační investice (strategické, finanční), potenciál (právní, finanční).

signifikantních vlastností mohou vypadat jako v Tab. 1.

Tyto je dále možné “přetavit” do podoby myšlenkové mapy (viz Příloha č.1 – Příklad myšlenkové mapy (formát JP2)) nebo inventář signifikantních vlastností pro jednotlivé formáty⁵⁸.

Výběru signifikantních vlastností by však měl předcházet výběr samotných zdrojů (nebo jejich typů) k uchování, pro který se v anglo-amerických kruzích ustálil termín *Appraisal*. Do češtiny by se dal volně přeložit jako (kontrolovaný) **výběr**. Ten představuje proces evaluace aktivit, kterým se určuje jaké dokumenty a na jak dlouho mají být uchovány (při respektování organizačních pravidel a očekávání komunity). Základní stavební kameny kontrolovaného výběru představují kompletní analýzu dokumentů, hodnocení aktivit a jejich právních omezení, určení reálnosti uchování dokumentů a konečné rozhodnutí (viz Obr. 28) [FOJTŮ, 2011].

Podle Neumeyera a Raubera [2007] však kontrolovaný výběr obvykle nahrává hodnotám hlavního proudu, zatímco ty ostatní jsou často potlačovány. Důležitou roli by podle nich mělo mít tzv. “řízené opomenutí” (*Planned Omission*), kdy dokumenty hlavního proudu jsou přijímány, ale neuchovájí se pro příští generace a poté následuje výběr materiálů kulturního dědictví pro dlouhodobou ochranu na bázi náhodného výběru. Představuje totiž nejméně nákladné a zároveň racionální řešení. Hlavní výhodou náhodného výběru před kontrolovaným výběrem (*Appraisal*) je nestranný pohled na současný život, jednoduchost a finanční přívětivost. Nicméně nejlepší řešení se jeví jako kombinace náhodného výběru a kontrolovaného výběru (společně s řízeným opomenutím) pomocí tzv. 3-stupňové strategie⁵⁹. Tato se částečně opírá i o kontrolovaný výběr jakožto jeden z prostředků poskytující autentický a realistický pohled na minulost. Jeho součástí je [NEUMEYER, 2007]:

- náhodný výběr
- manuální, tj. tradiční kontrolovaný výběr (*Appraisal*) za účelem zachování kompletnosti
- přístupová omezení z důvodu ochrany autorských práv.

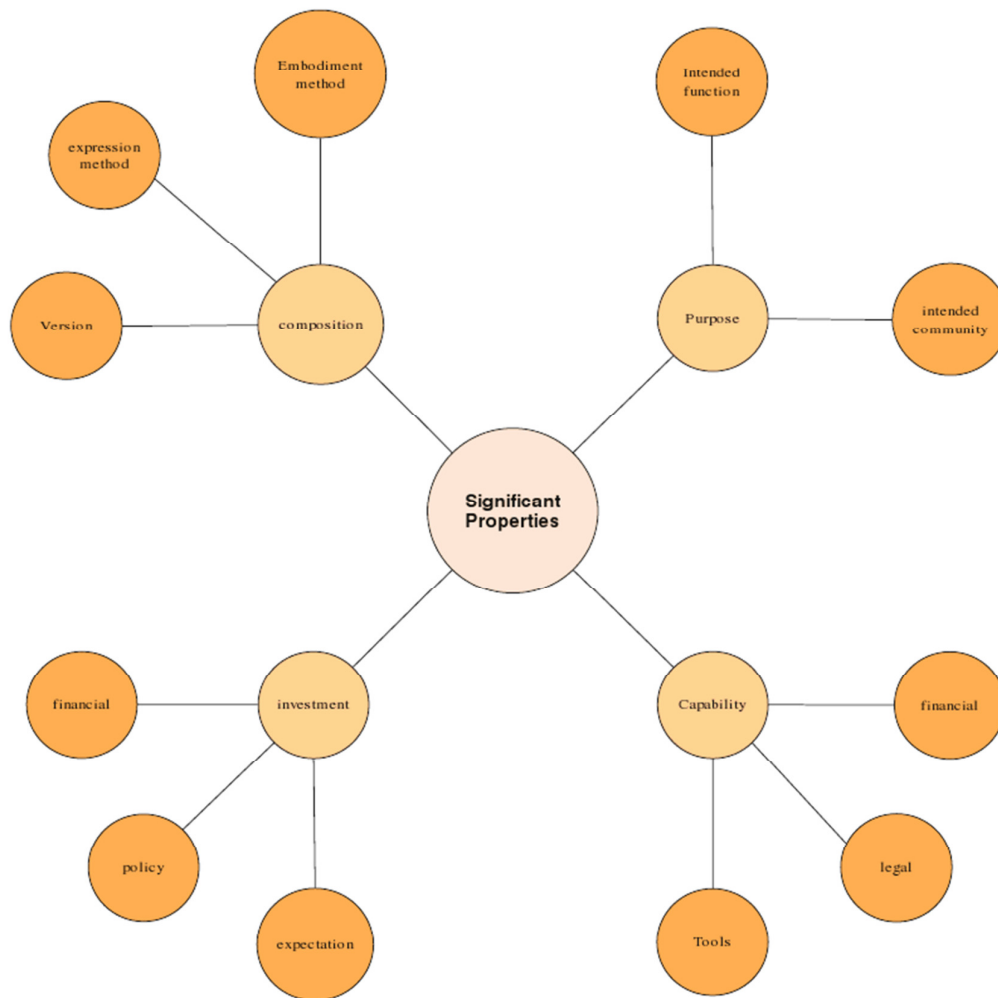
Analýza užitků (*Utility Analysis*) poskytuje managerům informace pro vyhodnocení finančního dopadu (včetně ROI) v případě určité intervence [NETWORK, 2007]. Tato analýza byla původně vyvinuta pro ekonomickou oblast. V ekonomické praxi je často velmi obtížné stanovit a kvantifikovat přínosy a užitky projektu a následně je vyjádřit v peněžních jednotkách, neboť celá řada efektů plynoucích z investice je nefinanční a někdy dokonce nehmotné povahy. Pro tyto

⁵⁸ Výbornou ukázkou je např. *Florida Digital Archive Format Information* -

<http://web.archive.org/web/20100527210441/http://www.fcla.edu/digitalArchive/formatInfo.htm>.

⁵⁹ Tuto metodu je dále možné kombinovat např. s rozhodovacím stromem od *Digital Preservation Coalition - Decision Tree for Selection of Digital Materials for Long-term Retention*. Podrobněji na <http://www.dpconline.org/advice/preservationhandbook/decision-tree>.

případy se nesnadno ocenitelné náklady a přínosy převádí na finanční toky pomocí různých metod, které umožňují tyto nefinanční náklady a přínosy ocenit [RRA, cca2010].



Obr. 28 Příklad signifikantních vlastností
<http://www.significantproperties.org.uk/images/atomic-model.gif>

Obrázek ukazuje, na jaké signifikantní vlastnosti se lze u objektů zaměřit. Základní 4 oblasti v této ukázce tvoří: účel (*Purpose*), kompozice (*Composition*), investice (*Investment*), schopnosti (*Capability*), k nimž jsou přiřazeny konkrétní vlastnosti či aspekty jakými jsou např. finanční, právní, politické apod.

Tento koncept (viz Obr. 28) je však možné aplikovat i na oblast digitální ochrany, kdy obecné strategie (emulace a migrace) je možné rozdělit na klíčové charakteristiky (modifikace originálníhoho

souboru, náklady při aplikování modifikovaného souboru do praxe apod.) a poté je (ke zvolení optimálního řešení) převést na kvantitativní rozhodovací kritéria.



Obr. 29 Postup analýzy užiteků (*Utility Analysis*)
[NETWORK, 2005]

Obrázek popisuje postup při analýze užiteků pro vyhodnocení finančního dopadu. Koncept je možné aplikovat i pro oblast dlouhodobé digitální ochrany, kdy obecné strategie, resp. přístupy k ochraně (např. emulace nebo migrace) začínají výběrem ukázkového příkladu pro analýzu, detailním sestavením jejich struktury. Poté je nutné identifikovat technické vlastnosti vybraného příkladu a určit výstupní očekávané chování. To je vhodné klasifikovat podle funkcí, kterým se přiřadí vlastnosti. Vše se následně vyhodnotí a sestaví konečná analýza, včetně finančních požadavků a možností.

Rozhodování v oblasti plánování dlouhodobé ochrany digitálních dat a výběru nevhodnějšího formátu pro budoucí migrace velmi dobře podporuje právě nástroj PLATO.

4.3.2 Testování pomocí nástroje PLATO

Celý proces evaluace PLATO je rozdělen na 4 fáze o 14 posloupných krocích: (podrobněji v části 4.3.2.2 Práce s nástrojem Plato):

1. **Definice požadavků** (*Define Requirements*) – pozůstává ze specifikace a podrobnějšího popisu sbírky (obsahu dat), která je vybrána pro naplánování ochranných akcí. Tento krok je obdoba první fáze auditu v nástroji DRAMBORA.

Výběr vzorových dat - vybrána jsou konkrétní vzorová data (optimálně 5-10 objektů), u kterých se podrobně stanoví požadavky na dlouhodobou ochranu. Autoři nástroje k tomuto účelu doporučují zejména open source nástroj FreeMind⁶⁰ pro vytváření myšlenkových map. Základním hodnotícím kritériem je obvykle soubor nebo záznam sbírky.

2. **Hodnocení alternativ** (*Evaluate Alternatives*) – hodnotící kritéria z předchozí fáze jsou podkladem pro provádění experimentálních migračních aktivit.
3. **Posouzení výsledků** (*Consider Alternatives*) - experimentální migrace jsou podkladem pro následnou analýzu a vyhodnocení nejlepšího formátu dané vzorové sbírky digitálních dat.
4. **Sestavení plánu dlouhodobé ochrany** (*Build Preservation Plan*) - výstupem nástroje PLATO je komplexní dokumentace s objektivním posouzením nejrůznějších migračních alternativ dle specifických požadavků dané instituce. Proto implementace nástroje do workflow dlouhodobé ochrany digitálních dat lze doporučit v plném rozsahu.

4.3.2.1 Hlavní implementované prvky nástroje PLATO ve SW verzi 4.2

K dispozici je několik drobných vylepšení implementovaných napříč nástrojem PLATO. Často jsou výsledkem zpětné vazby od uživatelů. Jde o:

1. **stahování nových „stromů“ požadavků** (*Requirements Tree*), které je možné dále upravovat v nástroji Freemind (viz výše 4.3.2., 1. Definice požadavků). Tato možnost je

⁶⁰ Nástroj je dostupný z: http://freemind.sourceforge.net/wiki/index.php/Main_Page. Jde o aplikaci napsanou v Java jazyce, dostupnou pro všechny nejrozšířenější operační systémy (platformy MS Windows, Mac OS X, Linux, eComStation a j.).

nejjednodušší, kdy se vlastní požadavky a znalosti zkombinují se zkušenostmi jiných uživatelů a odborníků. Nástroj (pro inspiraci) nabízí prohlížení již testovaných objektů jiných uživatelů/institucí, v případě, že je testující neoznačí za „neveřejný projekt“.

2. **automatizované měření**, jež představuje rozšiřitelný rámec pro připojení rozhodujících kritérií pro měřitelné vlastnosti a metriky v šesti kategoriích:
 - **výstupní objekt** (*Object Outcome*) a jeho požadované vlastnosti (např. zachování možnosti editace nebo šířky obrazu)
 - **výstupní formát** (*Outcome Format*) spolu s kritérii na formát objektu, ve kterém má být uchovávan (např. standardizace)
 - **důsledky** (*Outcome Effects*) vyvolané danou ochrannou akcí (např. náklady)
 - **runtime** (*Action Runtime*) potřebný k provedení akce (např. spotřebovaný čas a paměť)
 - **vlastnosti** (*Action Static*) náklady na licence a kvalita dokumentace konkrétního nástroje
 - **posouzení** (*Action Judgement*) vlastností, jakou je např. použitelnost.
3. **rychlé vyhodnocení** - workflow rychlého vyhodnocení, které umožňuje z několika základních předpokladů rychlé vyhodnocení možných řešení (ve třech krocích)
4. **vzdálená emulace** je integrována do PLATO (v rámci *Planet service Registry*⁶¹) a zprostředkována přes službu GRATE, jež běží na samostatném serveru
5. **integrace P2**⁶² - pro zvýšení počtu alternativních akcí a umožnění automatické evaluace formátů pomocí měřicího rozhraní bylo přidáno 44 tisíc údajů o formátech, které jsou popsány v registru P2. Umožňuje dotazovacím nástrojům (*Querying Tool*) testovací objekty převádět do daného formátu i přesto, že nejsou obsaženy v „Service Registry“.
6. **přednastavení pro ePrints** (<http://www.eprints.org>)⁶³ - repozitář je schopen realizovat ochranný plán vytvořený v Plato.

⁶¹ „Service Registry“ je poměrně nekompletní; kdokoliv, kdo hledá nástroj pro konverzi objektů též požaduje jiné zdroje potenciálních nástrojů. Jedním z řešení by bylo přímé prohledávání *Planets Service Registry*. Ten ale bohužel není dostupný na žádné veřejné adrese [PROM, 2010].

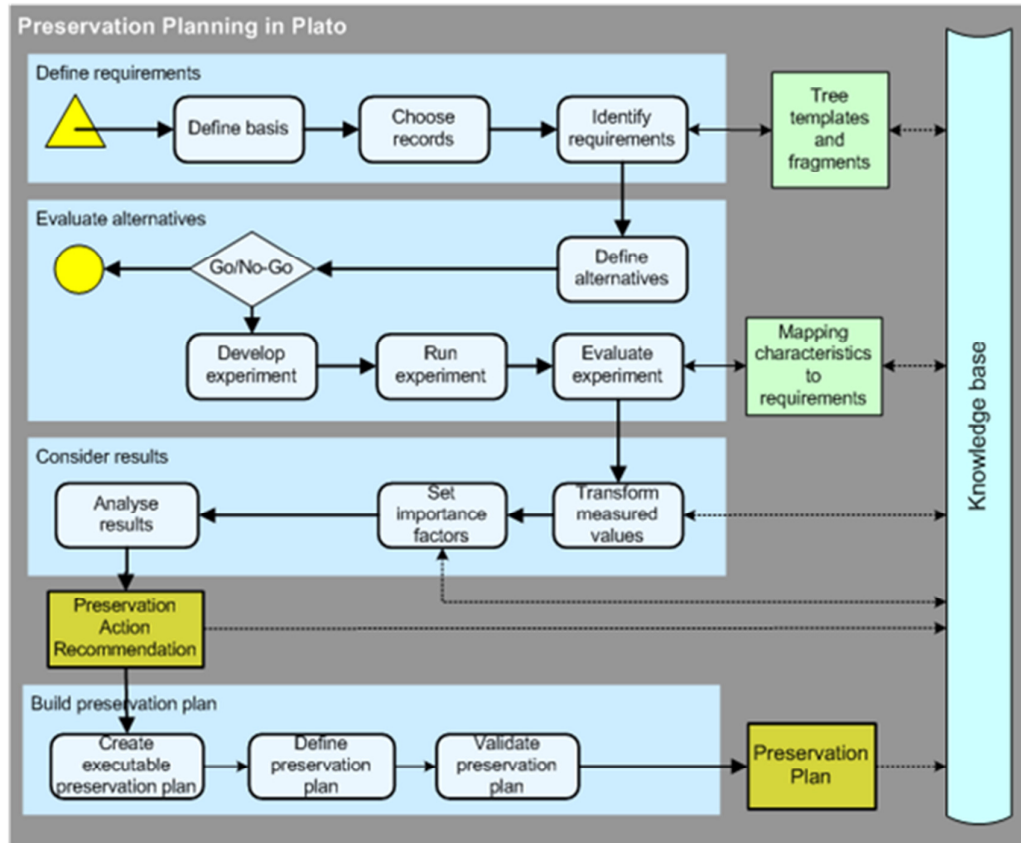
⁶² Podrobněji v TARRANT, D., HITCHCOCK, S., CARR, L. Where the Semantic Web and Web 2.0 Meet Format Risk Management:P2 Registry. *The International Journal of Digital Curation* [online]. Issue 1, Volume 2, 2011 [cit. 2011-05-22]. Dostupné z: <http://www.ijdc.net/index.php/ijdc/article/viewFile/171/239>. ISSN: 1746-8256.

⁶³ P2 představuje registr formátů, který je sémanticky obohacený a obsahuje informace pro podporu procesů dlouhodobé ochrany. Podrobněji v prezentaci: TARRANT, D., HITCHCOCK, S., CARR, L. A complete preservation workflow in EPrints (+ Plato) - 10 Minute Summary. *Preservation and Archiving Special Interest Group (PASIG)* [online]. Madrid : 5th July 2010 [cit. 2011-05-22]. Dostupné z: <http://eprints.ecs.soton.ac.uk/21279/>.

7. integrovaná FITS⁶⁴
8. vylepšená znalostní báze.

4.3.2.2 Práce s nástrojem Plato

Při přípravě nového plánu dlouhodobé ochrany digitálních dokumentů pomocí nástroje PLATO je nutná posloupnost kroků (viz Obr. 30):



Obr. 30 Postup plánování dlouhodobé ochrany v nástroji PLATO
(<http://blog.soton.ac.uk/keepit/files/2010/05/Preservation-planning-image.png>)

Obrázek popisuje proces plánování dlouhodobé ochrany pomocí nástroje PLATO. Workflow je možné rozdělit do čtyř základních fází: definování požadavků a výběr vzorových dat (*Define Requirements*), hodnocení alternativ (*Evaluate Alternatives*), posouzení výsledků (*Consider Results*) a sestavení plánu dlouhodobé ochrany (*Build Preservation Plan*). V první fázi je nutné definovat bázi (*Define Basis*), popsat akce a procedury pro vybrané, testované dokumenty či sbírky. Dále vybrat reprezentativní vzorky dat pro experimenty (*Choose Records*),

⁶⁴ FITS - *File Information ToolSet* (<http://code.google.com/p/fits>) zahrnuje a sjednocuje výstupy charakterizačních nástrojů (DROID a JHOVE), včetně extraktoru ExifTool (<http://www.sno.phy.queensu.ca/~phil/exiftool>) a dalších.

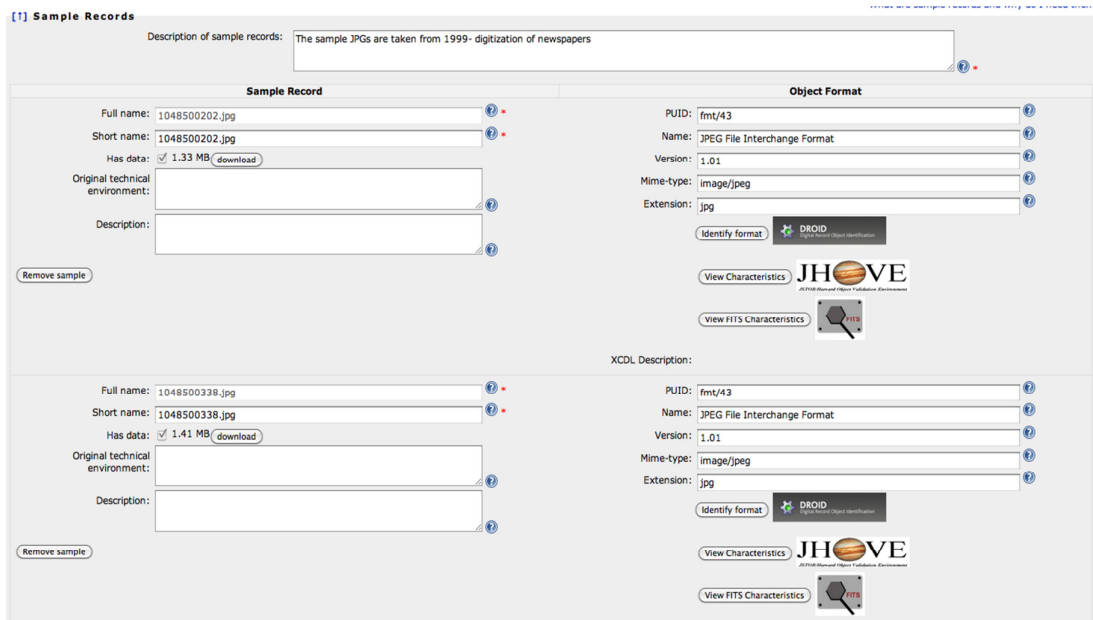
identifikovat jejich klíčové charakteristiky (*Identify Requirements*). Pro určité typy objektů nabízí PLATO vzory stromů cílů a signifikantních vlastností (*Tree Templates And Fragments*), jež jsou dostupné ve znalostní bázi PLATO (*Knowledge Base*). V druhé fázi je nutné z klíčových charakteristik stanovit alternativy (*Define alternatives*) a rozhodnout se pro nebo proti jejich testování (*Go/No-Go*). Z nich totiž vychází konkrétní experiment (*Develop Experiment* a *Run Experiment*), který je poté nutné vyhodnotit (*Evaluate Experiment*). Díky znalostní bázi je možné namapovat charakteristiky pro jednotlivé požadavky (*Mapping Characteristics To Requirements*). Po vyhodnocení experimentu je možné přetransformovat jednotlivé výsledky na měřitelné hodnoty (*Transform Measured Values*) a přiřadit jim potřebnou váhu (*Set Importance Factors*). Po analýze měřitelných hodnot a váh (*Analyze Results*) je možné doporučit ochranné akce pro testované objekty/sbírký. Na jejich základě se sestaví automatizovatelný plán dlouhodobé ochrany (*Create Executable Preservation*). Plán se zadefinuje (*Define Preservation Plan*) a ověří se jeho efektivnost (*Validate Preservation Plan*). Výsledkem má být písemná forma plánu dlouhodobé ochrany (*Preservation Plan*).

1. Definování báze (*Define Basis*) - tj. sesbírání veškeré dostupné dokumentace popisující akce a procedury (včetně nejrůznějších omezení) pro vybrané dokumenty/sbírký.

- **identifikace** je důležitá pro pozdější dohledání konkrétního plánu; všechny ochranné plány je potřebné náležitě označit a popsat. Definují se institucionální politiky, právní regule a uživatelská kritéria, která mohou ovlivnit plánovaná rozhodnutí pro ochranu dokumentů. Například politiky definující povolené formáty pro „ingest“, zákon o ochraně duševního vlastnictví apod.
- **status, popis** představují další krok v plánování - tedy jestli je konkrétní plán definovaný, čeká na schválení nebo je již aktivní a co vedlo k tomu, že jej bylo nutné nově definovat či pozměnit.
- **politiky** organizace zodpovědné za dlouhodobou ochranu svých dokumentů se v různých ohledech liší. Proto nejen duševní vlastnictví, technické požadavky, ale i strategie, cíle a politiky představují integrální součást plánu dlouhodobé ochrany.

2. Definování příkladů (*Define Sample Objects, Choose Records*) - jde o obecný popis objektů ve sbírce a sbírký jako celku (velikost sbírký, formáty ve sbírce apod.).

- **profil sbírký** - kromě obecného popisu výběru dokumentů, které nejlépe vystihují sbírký je možné uvést i stupeň rizika pro každý profilový objekt.
- **upload ukázkových souborů** – nahrání konkrétních reprezentačních souborů dané sbírký.



Obr. 31 „Upload“ ukázkových souborů do nástroje PLATO Preservation Planning Tool

„Screenshot“ ukazuje jakým způsobem probíhá nahrání testovaných objektů do nástroje PLATO. Na začátku je potřebné popsat společné charakteristiky vybraného vzorku dat a důvod pro jejich testování (*Description Of Sample Record*). Popis může být i poměrně stručný, jako na tomto obrázku: „Ukázkové JPG soubory jsou z roku 1999 - digitalizace novin“. U importovaných souborů (*Sample Record*) jsou automaticky rozpoznány charakteristiky jako jsou PUID formátu, název formátu (*Name*), jeho verze (*Version*), „mime-type“ a odpovídající přípona (*Extension*). Pokud identifikace není přesná, je možné využít další identifikační služby jako jsou DROID, JHOVE nebo FITS Characteristics.

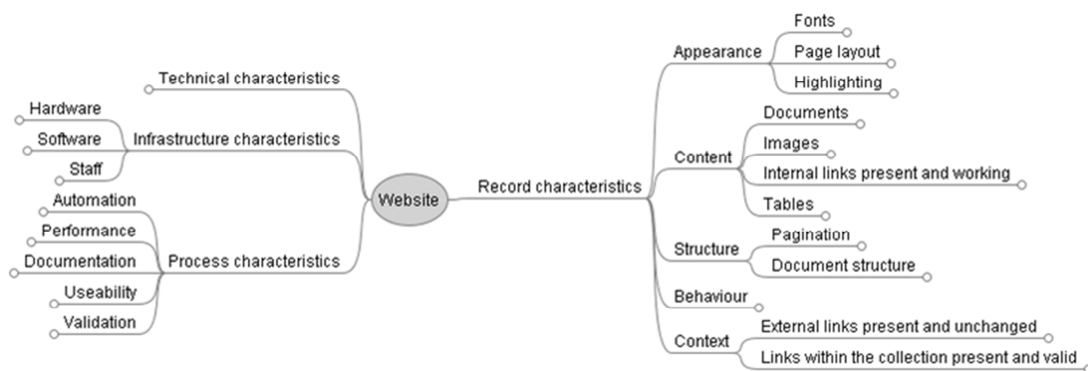
3. Identifikace požadavků (*Identify Requirements*) - představuje nejdůležitější krok při plánování dlouhodobé ochrany. Požadavky jsou definovány ve formě tzv. stromu cílů, myšlenkové mapě kritérií popisujících současné charakteristiky daných formátů včetně jejich budoucí podoby. Konkrétní podoba vychází ze samotných digitálních objektů a potřeb instituce, přesto mají společný základ, který je možné rozdělit do čtyř základních skupin:

- **vlastnosti objektu (*Object Characteristics*)** – popisuje vizuální a kontextovou zkušenost, kterou uživatel získává při práci s digitálním objektem. Jde o **vzhled (*Appearance*)**, **obsah (*Content*)**, **strukturu (*Structure*)** a **chování (*Behaviour*)**, s charakteristikami nejnižší úrovně typu barevná hloubka (*Colour Depth*), rozlišení (*Image Resolution*), způsob interakce (*Forms Of Interactivity*),

podpora maker (Macro Support) nebo vložená metadata (Embedded Metadata).

- **charakteristiky záznamu** (*Record Characteristics*) – určují technický základ digitálního objektu, kontext, vztahy mezi objekty a metadaty.
- **charakteristiky procesů** (*Process Characteristics*) – poukazují na procesy dlouhodobé ochrany, včetně použitelnosti, komplexnosti nebo škálovatelnosti.
- **náklady** (*Costs*) – mají podstatný vliv na výběr konkrétní strategie dlouhodobé ochrany. Obvykle jsou rozděleny na technické a personální nebo na počáteční a provozní.

Pro **Identifikaci požadavků** je možné využít i výsledky jiných institucí ze **znalostní báze PLATO**.



Obr. 32 Příklad „stromu“ cílů pro webovou stránku

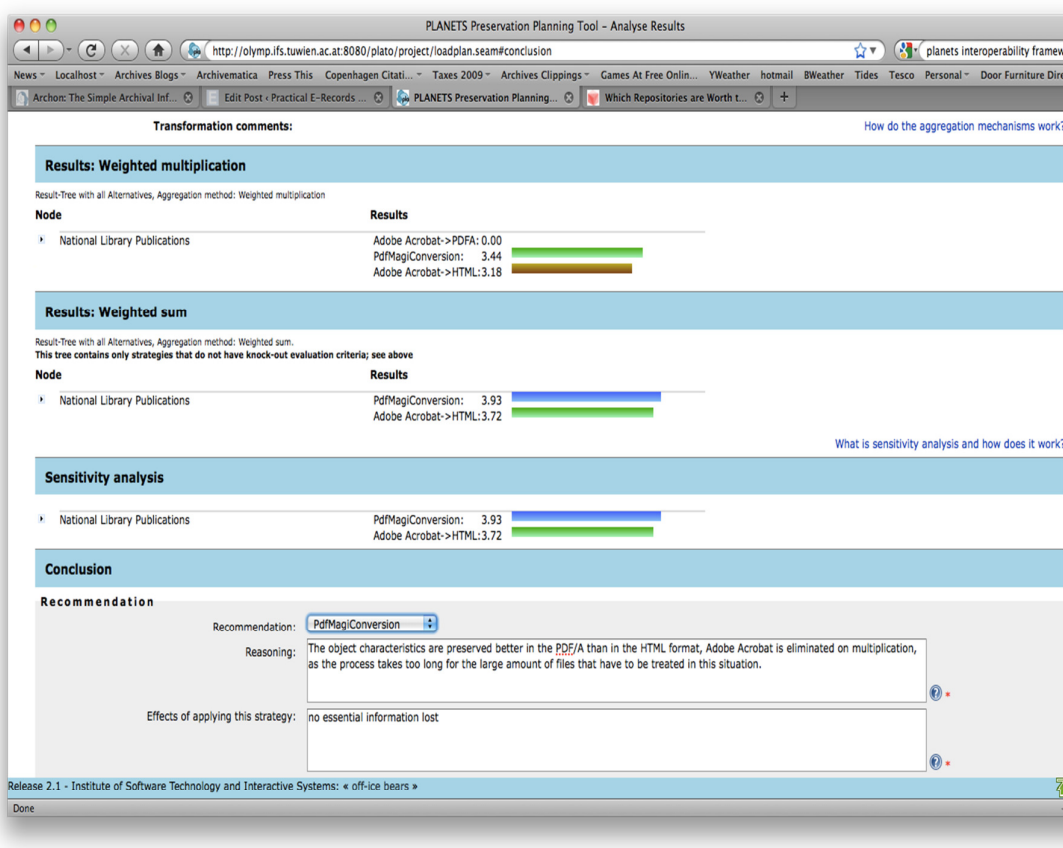
Obrázek přehlednou a výpočetní formou definuje signifikantní vlastnosti digitálního objektu (v příkladu webová stránka). Klíčové charakteristiky jsou zobrazeny formou myšlenkové mapy. Pro potřeby nástroje PLATO jsou myšlenkové mapy (nazývány jako „strom požadavků“) vytvářeny online, ručně přímo na webu PLATO, nebo je možné je nejdříve vytvořit v aplikaci FreeMind a uploadovat do PLATO jako celek. Myšlenková mapa má kořenový uzel (*Website*) ze kterého vedou jednotlivé uzly a poduzly. Pro webovou stránku je nutné podchytit technické vlastnosti (*Technical Infrastructure*), vlastnosti infrastruktury (*Infrastructure Characteristics*) a vlastnosti vlastního záznamu (*Records Characteristics*), pro které jsou podstatné poduzly vzhled (*Appearance*), obsah (*Content*), struktura (*Structure*), chování (*Behaviour*) a kontext (*Context*). Tyto charakteristiky ale nejsou univerzální napříč všemi repozitáři, každá instituce si sama, podle své cílové skupiny a svého poslání, stanovuje signifikantní vlastnosti, které se zavazuje do budoucna uchovávat.

4. **Definice alternativ** (*Define Alternatives*) – jsou vybrány rozdílné strategie dlouhodobé ochrany (např. migrace vs. emulace). Pro každou alternativu se vyhodnocuje množství práce, času, finančních prostředků.
5. **Rozhodnutí k akci** (*Take Go Decision*) – některé testy strategií dlouhodobé ochrany pro dané sbírky digitálních objektů počítají s mnohými alternativami nebo vysokými náklady na hardware a software. V tomto kroku je tedy nutné zvážit jejich proveditelnost.
6. **Navržení experimentu** (*Develop Experiment*) – aby bylo možné provádět opakovatelné a tím pádem i objektivní testy, je nutné vše dobře zdokumentovat (včetně experimentálního prostředí, workflow, software a hardware a mechanismy získávání výsledků).
7. **Provedení experimentu** (*Run Experiment*) – experimenty mají být sestaveny tak, aby bylo možné otestovat jednu či více alternativ dlouhodobé ochrany pro předem definovaný soubor digitálních objektů. Experiment poté např. ukáže zkonvertované souborové formáty, změní hardwarovou zátěž konverze apod.
8. **Vyhodnocení experimentu** (*Evaluate Experiment*) – výsledky experimentů je nutné zkontrolovat a vyhodnotit do jaké míry definované požadavky byly naplněny.
9. **Převedení naměřených hodnot** (*Transform Measured Values*) - hodnoty ve „stromu“ požadavků jsou uvedeny v různých jednotkách (vteřiny, koruny, bity apod.). Aby bylo možné tyto rozdílné hodnoty agregovat, je nutné je přetransformovat do jednotné škály. Vývojáři Plato uvádějí, že k nejrelevantnějším výsledkům se hodnotitelé dopracují při škále celých hodnot od 0 do 5, kdy 0 představuje zcela neakceptovatelný výsledek a 5 naopak nejlepší možný rezultat. Podle vývojářů Plato umožňuje tento nástroj vyhodnocování signifikantních vlastností u téměř všech typů objektů, a to automaticky nebo ručně. Automaticky hodnocená kritéria pro obrazové dokumenty typu JPG jsou: “imagequality:AE”, “imagequality:MAE”, “imagequality:MEPP”, “imagequality:MSE”, “imagequality:PAE”, “imagequality:PSNR”, “imagequality:RMSE”, “input:filesize”. Všechny ostatní možné signifikantní vlastnosti je nutné vyhodnotit ručně; což v případě cca 25 kritérií a 3 testovacích souborů přes min. 3 externí služby představuje ruční hodnocení cca 225 bodů. Ruční kontrola musí proběhnout i u takových signifikantních kritérií, u kterých by strojová kontrola neměla být překážkou (např. “image size”, “image height”, “image width” apod.). Jde totiž o porovnávání číselné shody vstupu a výstupu, nikoliv o vyhodnocení kvality (např. vizuální čitelnost). V případě TXT souborů jsou testovací migrace ještě komplikovanější. Externí služby (*External Services*) nenabízejí žádnou vhodnou alternativu pro migrace TXT souborů.

10. Stanovení váhy (Set Importance Factors) – ne všechna kritéria jsou stejně důležitá. Pro každou přetransformovanou jednotku je též nutné stanovit její váhu .

11. Vyhodnocení alternativ a výsledků (Analyse Results) - jde o proces empirické evaluace výsledků, který tvoří základ pro další fáze (zejména transformaci).

- **suma (Sum)** – naměřené hodnoty výkonu jsou znásobeny váhou. Tyto hodnoty jsou sečteny do jedné porovnatelné váhy pro danou alternativu. Výsledky kritérií u koncových uzlů (*Leaf Values*)⁶⁵, kterých hodnota dosáhne 0 nemají rozhodující vliv na výslednou hodnotu.
- **násobení (Multiplication)** – zde je váha kritéria koncového uzlu vynásobená celkovou váhou uzlu.
- **souhrn výhod (Sum Of Advantages)** – umožňuje vzájemné srovnání dvou alternativ.



Obr. 33 Ukázka analýzy výsledků v nástroji PLATO

⁶⁵ Označení „koncového uzlu“ (*Leaf Node*) vychází z konceptu myšlenkových map. Např. u Obr 32. „Strom“ cílů pro webové stránky představují „koncové uzly“ – *Technical characteristics, Infrastructure characteristics, Process characteristics, Appearance, Content, Structure, Behaviour, Context*.

(<http://e-records.chrisprom.com/wp-content/uploads/2010/04/Screen-shot-2010-04-23-at-1.16.23-PM.png>)

Screenshot ukazuje výsledky zkušební migrace souborů, ke kterým se ručně přidává komentář s hodnocením výsledné transformace (v tomto konkrétním případě se jedná o PDF nebo HTML). V příkladu je porovnána konverze Adobe Acrobat do PDF/A, PdfMagiConversion a Adobe Acrobat do HTML. Jak je uvedeno na obrázku, zdůvodnění výběru konverzního nástroje PdfMagiConversion (pro budoucí aplikaci do praxe) může mít následující formu: „Vlastnosti objektu jsou lépe uchovány v PDF/A než v HTML formátu, Adobe Acrobat je vyloučen multiplikací, jelikož proces pro tak velké množství ošetřených souborů trvá neúměrně dlouho.“

Některé z hodnocených výsledků je možné provést automaticky, stále však ještě většinu z nich je nutné ohodnotit ručně. To znamená procházet jednotlivá neautomatizovatelná kritéria, doplňovat hodnotu a poté je finálně odsouhlasit. Následuje transformace, kdy je možné Plato použít pouze jako migrační nástroj, nikoliv jako hodnotící nástroj. Vybrané vzorky formátů se vloží do Plato, přemigrují do jiných formátů, stáhnou se z nástroje a mimo něj vyhodnotí. Při vyhodnocování se kritéria, resp. signifikantní vlastnosti rozdělí na kvantifikovatelné (tedy strojově kontrolovatelné) a kvalitativní (ručně kontrolovatelné) a přiřadí se jim potřebná váha. Všechny výsledky je následně možné sepsat např. ve formě tabulky vytvořené v některém z dostupných tabulkových procesorů. Odrazovým můstkem k efektivní dlouhodobé ochraně digitálních dokumentů, které daná instituce spravuje, je pochopení, že cílem nemá být uchovávání všeho. Předpokládá se výběr dle stanovených pravidel a podmínek, za jakých se budou dokumenty chránit. Zde se jako nejvhodnější metodika pro selekci objektů a jejich typů jeví tzv. 3-stupňová strategie, která kombinuje náhodný výběr a kontrolovaný výběr společně s řízeným opomenutím. Jakmile víme co uchovávat, je možné se zaměřit na to jakým způsobem. To předpokládá znalost klíčových prvků (objektů), které je nutné dlouhodobě uchovávat. Správci sbírek tedy potřebují znát signifikantní vlastnosti pro výběr preservačních technik a načasování různých preservačních akcí u objektů určených k dlouhodobé ochraně. Ne všechny však byly vytvořeny v otevřených specifikacích. Jelikož jsou proprietární formáty navázány na komerční subjekt, přetrvává u nich velké riziko ztráty informací (autentický, srozumitelný obsah). Proto je žádoucí takovéto formáty souborů migrovat do otevřenějších protějšků při zachování jejich signifikantních vlastností. Ulehčují to nejrůznější nástroje (většinou „open source“), ale jen některé jsou vhodné k integraci do systémů rozsáhlých digitálních knihoven a archivů. **Nejpoužitelnějším nástrojem pro plánování dlouhodobé ochrany se zdá být Plato, který v současné době nemá žádný vhodnější ekvivalent.** Hlavním principem webového nástroje Plato je vyhodnocování preservačních aktivit

podle stanovených cílů, signifikantních vlastností a požadavků na jejich zachování. Reálná práce s nástrojem Plato ukazuje, že většinu z předem definovaných signifikantních vlastností u JPG, JP2 a TXT není možné automaticky otestovat. U JPG a JP2 je nutné většinu kritérií ohodnotit ručně. U TXT souborů není možné provést ani samotné migrace. Nadějí na zlepšení je v tomto ohledu další vývoj Plato v rámci evropského projektu SCAPE: *Scalable Preservation Environments*.

12. Sestavení proveditelného plánu (*Create Executable Plan*) – představuje definici jakýchsi „spouštěčů“, tedy podmínek, kdy budou uskutečněny dané akce dlouhodobé ochrany (např. migrace/emulace pro jaké soubory a signifikantní vlastnosti).

13. Definování plánu dlouhodobé ochrany (*Define Preservation Plan*) – na rozdíl od předchozích bodů, které se zaměřují spíše na technickou otázku a definování plánu dlouhodobé ochrany představuje spíše záležitost na úrovni organizace. Finanční náklady do značné míry ovlivňují výběr konkrétních alternativ. V tomto kroku lze tedy provést detailní kalkulaci pomocí modelu nákladů LIFE2 nebo Total Cost of Ownership (TCO)⁶⁶. Zároveň se v této fázi zdokumentují i přiřazené odpovědnosti za monitorování a provádění činností spojených s dlouhodobou ochranou.

14. Validace plánu (*Validate Plan*) – finální fáze dokumentace konkrétního plánu dlouhodobé ochrany pro předem stanovené digitální objekty. Kontrolují se proveditelné plány (*Executable Plan* – viz bod 12.) a poté se zvalidují. Po validaci plánu zodpovědnou osobou již není možné provádět částečné změny bez celkové revize plánu.

Instituce má mít ve výsledku k dispozici konkrétní (na základě měřitelných kritérií) plán dlouhodobé ochrany pro konkrétní sbírku digitálních objektů.

4.4 Komerční systémy pro dlouhodobou ochranu

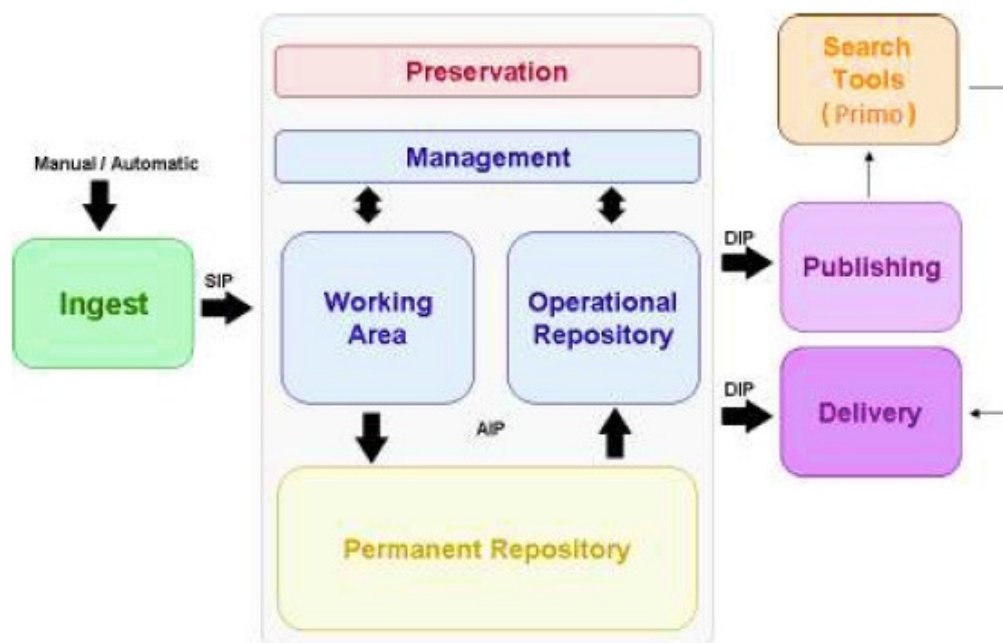
Za překotným vývojem „open source“ systémů a nástrojů dlouhodobé ochrany digitálních dokumentů do značné míry stály komerční systémy a jejich první „pokusy“ implementace ISO normy 14723:2003 do reálně fungujících informačních systémů. Prvním takovým komerčním systémem byla Rosetta (dříve známa jako DPS - *Digital Preservation System*).

4.4.1 Rosetta

Rosetta, jako první komerční řešení pro digitální uchování (dříve nazýváno *Digital Preservation System*), bylo navrženo ve spolupráci ExLibris s Národní knihovnou Nového Zélandu

⁶⁶ Podrobněji viz kapitola 5.2 Ekonomické modely.

(NLNZ) a připomínkováno mezinárodně uznávanými odborníky a inovátory. Systém pro NLNZ byl uveden do běžného provozu v říjnu 2008. Mezi klíčové vlastnosti systému lze považovat plánování dlouhodobé ochrany a příslušných aktivit (identifikace formátů, identifikace rizik pro jednotlivé formáty apod.), podpora metadat pro dlouhodobou ochranu (PREMIS), přístupný archiv (online přístup k informacím v permanentním úložišti) a webové uživatelské rozhraní. **Rosetta díky své škálovatelnosti⁶⁷ o otevřenosti systému umožňuje uchovávat digitální obsah nejrůznějšího typu i pro budoucí generace a obohacovat systém o „plug-in“ nástroje či jiné aplikace dle potřeb konkrétní instituce.**



Obr. 34 Architektura systému Rosetta
(<http://www.exlibrisgroup.com/ru/files/Products/Preservation/RosettaDescription.pdf>)

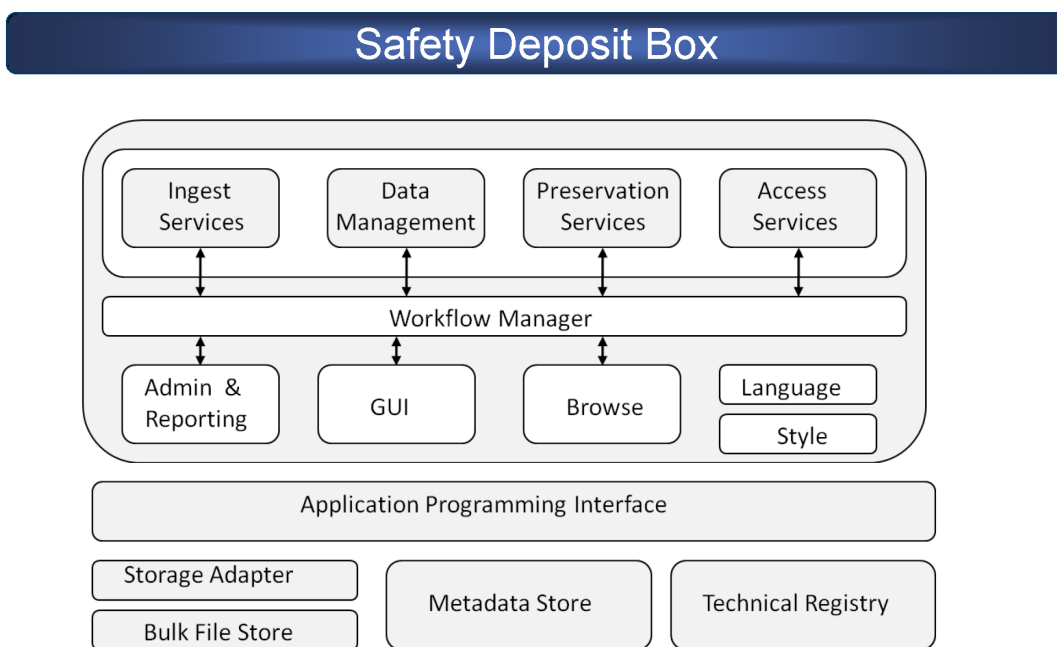
Obrázek popisuje obecnou architekturu systému Rosetta. Data vstupují do systému automaticky nebo manuálně (*Manual /Automatic*) formou vstupních informačních balíčků (SIP) přes modul Import (*Ingest*) do pracovní oblasti (*Working Area*), odkud můžou být zasílány pro obohacení do modulu Správy (*Management*) nebo po kompletním zpracování do modulu Permanentního úložiště (*Permanent Repository*). Operační repositář řeší spojení mezi Permanentním úložištěm, modulem Správy a vyhledávacími nástroji. Vyhledávací rozhraní (*Search Tools*) může fungovat jako nadstavba systému, do kterého jsou uživatelské informační balíčky zasílány (DIP) pomocí nástrojů publikování

⁶⁷ Dokument ke škálovatelnosti v podobě „Proof of Concept“ je dostupný na:
<http://www.exlibrisgroup.com/files/Products/Preservation/RosettaScalingProofofConcept.pdf>.

(Publishing) a zobrazovány pomocí modulu pro zpřístupňování (Delivery).

4.4.2 SDB - Safety Deposit Box

Safety Deposit Box je produktem britské firmy Tessella⁶⁸, jehož architektura (stejně jako v případě Rosetta) vychází z referenčního modelu OAIS. Tomu odpovídají i jednotlivé moduly systému (*Ingest, Access, Preservation, Storage, Data Management a Admin & Reporting*). Původně byl systém navržen zejména pro větší archivy a knihovny, dnes oslovuje i zákazníky z komerčního sektoru (farmacie, zdravotnictví apod). **Základní předností systému je jeho flexibilita a škálovatelnost. Zde jde především o schopnost pracovat s obsahem v nejrůznějších formátech (jak strukturovaná data (databáze), tak i nestrukturované dokumenty) a schopnost zpracování obrovského objemu dat v co nejkratším možném čase⁶⁹.**



Obr. 35 Architektura systému SDB
[EVANS, 2011]

⁶⁸ <http://www.digital-preservation.com/solution/>

⁶⁹ Škálovatelnost byla testována na projektu Family Search (<https://familysearch.org>). Jednalo se ambiciózní projekt digitalizace rozsáhlé sbírky genealogických informací. Testoval se import SIP balíčků o velikosti 4.4GB. Potřebné bylo naimportovat cca 20 TB za den. Řešením byl update diskového pole na 168 disků paralelně a přidán byl další "job queue server". Podrobněji ve videu: <http://www.digital-preservation.com/lessons-from-a-real-large-scale-digital-preservation-system/>.

Obrázek zachycuje obecnou architekturu systému SDB. Základ tvoří služby importu (*Ingest Services*), správy dat (*Data Management*), dlouhodobé ochrany (*Preservation services*) a přístupu (*Access Services*). Nad nimi funguje systém pro komunikaci s nadstavbovými prvky (*Workflow Manager*), jež představují správu a reporting (*Admin & Reporting*), grafické rozhraní (*GUI*), nástroj pro vyhledávání (*Browse*). Přes API (*Application Programming Interface*) jsou napojeny další nástroje: připojení k úložišti (*Storage Adapter*), úložiště hromadných souborů (*Bulk File Store*), úložiště metadat (*Metadata Store*) a technický registr (*Technical Registry*).

4.4.2.1 Preservica

V současné době přichází Tessella i s možností cloudového řešení dlouhodobé ochrany pomocí systému Preservica⁷⁰, který vychází z SDB. Výhodou tohoto řešení jsou nižší pořizovací náklady řešení dlouhodobé ochrany (není potřebné nakupovat úložiště, zaměstnávat specializované odborníky pro správu systému). Instituce platí pouze za skutečně využitý prostor, který lze kdykoliv navýšit a jež je neustále zálohovaný. V neposlední řadě jsou formáty objektů průběžně kontrolovány pro jejich případné zastarání. **Nevýhodou tohoto, resp. jakéhokoliv komerčně spravovaného cloudového řešení je skutečnost, že data se dostávají mimo původní instituce, není stoprocentní jistota jejich ochrany, zabezpečení a případného navrácení v případě bankrotu či neshod se společností. Vlastní data instituce pak mohou být předmětem vydírání. Proto je vždy nutné zvážit jak výhody i nevýhody a vše dostatečně právně ošetřit.**

4.5 Open source systémy a nástroje pro dlouhodobou ochranu

V současné době vznikají „open source“ nástroje pro dlouhodobou ochranu digitálních dat doslova jako „houby po dešti“. Některé jsou vyvíjeny za účelem jejich integrace do stávajících systémů institucí či cílových skupin (např. CRI-B). Jiné fungují jako první prototypy „open source“ digitálních repozitářů odpovídajících referenčnímu modelu OAIS a plně podporujících životní cyklus dlouhodobé ochrany digitálních dat (např. Mopseus, Archivematica, RODA). Samostatnou skupinu nástrojů tvoří metodologie (online, offline) sledování, evaluace a hodnocení činnosti repozitáře včetně plánování dlouhodobé ochrany dat [FOJTŮ, 2010].

⁷⁰ <http://preservica.com>

4.5.1 Archivematica

Archivematica⁷¹ je open source systém vyvíjený jako sada 24 „mikroslužeb“ (*Microservices*) postavených kolem souborového systému, rozdělených do 9 základních kategorií:

1. Příjem SIP (*ReceiveSIP*)
2. Kontrola SIP (*ReviewSIP*)
3. Karanténa SIP (*QuarantineSIP*)
4. Ocenění AIP (*AppraiseSIP*)
5. Příprava AIP (*PrepareAIP*)
6. Kontrola AIP (*ReviewAIP*)
7. Uchování AIP (*StoreAIP*)
8. Zpřístupnění DIP (*ProvideDIP*)
9. Monitorování ochrany (*MonitorPreservation*).

Celá sada pak vyhovuje referenčnímu modelu OAIS

Beta verze pro implementaci do partnerských institucí byla vydána začátkem roku 2011, technicky zdatnější uživatelé (jde o „virtual appliance“ - upravený Ubuntu Linux s „open source“ nástroji) mohou systém implementovat do svých infrastruktur.

Cílem systému postavených na „mikroslužbách“ je redukce technické komplexnosti a údržby, zjednodušení vývoje a možnost snadného nahrazení zastaralých prvků systému.

Zatím nejvyužitelnější součástí se zdá být *MonitorPreservation* - Archivematica uchovává originální formát všech importovaných dokumentů pro migraci a emulaci, přičemž normalizace souborů pro dlouhodobou ochranu je v otevřených formátech.

4.5.2 RODA & CRIB

RODA – *Repositório de Objectos Digitais Autênticos*⁷² nebo-li Repozitář autentických digitálních objektů představuje technické řešení na úrovni národního archivu. **Na rozdíl od ostatních hodnocených systémů, které se zaměřují pouze na některou z oblastí problematiky dlouhodobé ochrany, RODA je komplexní systém správy digitálních objektů - od vkládání až po**

⁷¹ Autor: Artefactual Systems (<http://artefactual.com/>), finanční garant - UNESCO Memory of the World Subcommittee on Technology, Instalace: Windows XP, Linux, Mac; dostupné na: <http://archivematica.org/>, Požadavky: Java Runtime Environment 1.6, doporučená instalace OpenOffice, Poslední verze: 0.6-alpha nebo 0.6.2-alpha (pouze pro vývojáře a technicky zdatné uživatele)

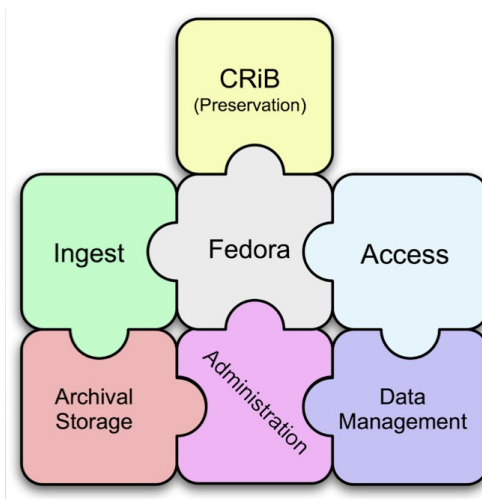
⁷² Autor: Portuguese National Archives (Directorate General of the Portuguese Archives), University of Minho (<http://antt.dgarq.gov.pt/>, <http://www.uminho.pt/>), Instalace: GNU/Linux systém - ověřeno na Ubuntu 8.04 a Ubuntu 8.10, Požadavky: doporučená instalace OpenOffice, Poslední verze: 29.07.2009, Plány pro další verze: projekt byl ukončen (1.4.2006 - 31.1.2008)

jejich zpřístupnění.

Jádro systému formuje Fedora Commons. Servisně orientovaná architektura, tedy jakýsi „balíčkový systém“ interoperabilních služeb je možné využít v různých systémech. RODA vychází z referenčního modelu OAIS, pro popisná metadata využívá schéma EAD. Digitální objekty jako soubor binárního obsahu, vztahů a reprezentací, jsou strukturovány v METS⁷³.

CRiB (*Conversion And Recommendation Of Digital Object Formats*), který původně vznikl jako samostatný migrační nástroj, je integrován do RODA v rámci preservačního modulu. Zabezpečuje identifikaci formátů, doporučuje optimální verze pro migraci, provádí (distribuované) migrace objektů do nových, vhodnějších formátů a následně kontroluje jejich integritu. Po úspěšném vyhodnocení migračního procesu generuje ochranná metadata v PREMIS⁷⁴.

I když se jedná o technické řešení na úrovni národního archivu, díky servisně orientované architektuře je možné jej (především modul CRiB) integrovat do různých systémů digitálních knihoven (např. do Transformačního modulu či Modulu plánování dlouhodobé ochrany) [FOJTŮ, 2010].



Obr. 36 Architektura systému RODA
(http://www.bl.uk/ipres2008/presentations_day2/37_Ramalho.pdf)

Obrázek přehledně zobrazuje architekturu Repozitáře autentických digitálních objektů (RODA). Jádro systému tvoří Fedora, kolem kterého jsou vybudovány jednotlivé moduly systému v souladu s ISO normou OAIS 14721:2003: vkládání

⁷³ METS - Metadata Encoding and Transmission Standard pro zakódování popisných, administrativních a strukturálních metadat pro objekty v digitálních knihovnách.

⁷⁴ PREMIS - *PREservation Metadata: Implementation Strategies*, představuje pracovní skupinu, která definuje vhodná metadata pro dlouhodobou ochranu. Akronym PREMIS se však používá i pro označení datového slovníku PREMIS Data Dictionary.

dat (*Ingest*), permanentní úložiště (*Archival Storage*), administrace (*Administration*), přístup (*Access*) a správa dat (*Data Management*). Modul Plánování dlouhodobé ochrany zde zastupuje systém CRiB (*CriB (Preservation)*).

4.5.3 AIDA

AIDA (*The Assessing Institutional Digital Assets*⁷⁵) představuje metodologii pro hodnocení činností souvisejících s repozitářem, především kapacity, stav připravenosti a celkovou schopnost správy digitálních sbírek. Toto hodnocení však neprobíhá formou auditu (jak je tomu např. v případě nástroje DRAMBORA⁷⁶). Jde o evaluaci (*Measurement*), proto „důkazy“ pro jednotlivé aktivity či naopak nečinnost nejsou nutné. V podstatě tedy špatný výsledek neexistuje.

Na evaluaci činnosti repozitáře a procesů dlouhodobé ochrany se nepodílí jedna osoba, nýbrž všechny zodpovědné osoby (na úrovni oddělení i institucionální úrovni), které postupně vyplňují hodnotící tabulky. Pomocí nástroje se hodnotí 3 klíčové oblasti:

1. **organizace** (*Organizational Leg*)
2. **technické zabezpečení** (*Technical Leg*)
3. **zdroje** (*Resource Leg*).

V každé této oblasti je nutné zhodnotit 11 aspektů správy digitálních dat/sbírek (pro zdroje pouze 9) na 5 úrovních připravenosti⁷⁷. Ty mají svůj stručný popis s příkladem, který vysvětluje a usnadňuje evaluaci. Ku příkladu prvek č. 6 (*Sustainability Of Funding*), 2. stupeň na institucionální úrovni hodnotí kontinuitu finančních prostředků⁷⁸. Jelikož jde o metodologii, integrace AIDA je možná v případě její implementace do stávajících systémů či budoucích systémů [FOJTŮ, 2010].

Pokud by si člověk měl z problematiky dlouhodobé ochrany odnést pouze jednu myšlenku, byla by to ta, která říká o neschopnosti uchovávat vše. Jednak to z hlediska personálních, finančních a technických možností není pravděpodobné a zůstává otázkou, jestli je to vůbec vhodné. I u samotných objektů nelze uchovávat vše, proto je nutné vybrat klíčové/signifikanční vlastnosti,

⁷⁵ Autor: University of London Computer Centre, financováno z 1/07 JISC Capital Programme Call, Instalace: nejde o softwarovou aplikaci, Požadavky: nejsou známy žádné specifické požadavky, nicméně AIDA je cílen především na anglo-saské akademické instituce, Poslední verze: květen 2009, Plány pro další verze: projekt byl ukončen (1.10.2007 - 31.3.2009); v dalším samostatném projektu je snaha o vytvoření praktického nástroje, který pro jednotlivá kritéria bude navrhopvat praktická řešení

⁷⁶ <http://repositoryaudit.eu/>

⁷⁷ Stupeň připravenosti představuje úroveň splnění jednotlivých kritérií, přičemž úroveň 1 (stage) představuje nejnižší možný stupeň.

⁷⁸ „Finance přichází krátkodobě, formou jednorázových grantů a „ad hoc“ ocenění“.

kteřé se knihovna, repozitář či archiv zaváže do budoucna zachovat nejrůznějšími prostředky a pomocí dostupných komerčních či "open source" řešení. Koneckonců nejde ani tak o konkrétní řešení jako o koncepční, systematický přístup a dodržování standardních postupů a příslušných norem.

5. Aktivity na organizační úrovni

Posledním krokem před vypracováním konečné verze plánu dlouhodobé ochrany digitálních dokumentů je kontrola (audit) předcházejících postupů a výsledků. Výběr konkrétní metodiky je podmíněný personálním obsazením, materiálovým vybavením a v neposlední řadě finančními zdroji.

5.1 Metriky, certifikační a auditní nástroje

Nejznámější metodologie hodnocení a certifikace důvěryhodnosti jsou (viz Obr. 37 na str. 116):

- **Preserving Digital Information** – 1996
- **Trustworthy Digital Repositories** - 2002
- **Catalogue of Criteria for Trusted Digital Repositories** (nestor Catalogue) – 2006, 2009 (později schváleno jako německý standard DIN 31644 Criteria for trustworthy digital archives)
- **Digital Repository Audit Method Based on Risk Assessment** (DRAMBORA) - DCC and DPE - 2007
- **Trustworthy Repositories Audit & Certification: Criteria and Checklist** (TRAC) – 2007
- **Data Seal of Approval** (DSA) - 2008

Nejčastěji probíhá hodnocení kritérií pomocí metodik DSA, TRAC nebo online nástroje DRAMBORA.

V červenci 2010 byl signatáři *Memorandum of Understanding* odsouhlasen Evropský rámec pro audit a certifikaci digitálních repozitářů (*European Framework for Audit and Certification of Digital Repositories*). Dohodu podepsali zástupci *CCSDS/ISO Repository and Audit Certification Working Group /RAC*, *Data Seal of Approval Board* a *DIN Working Group Trustworthy Archives – Certification*. Tento rámec stanovil tři základní úrovně auditů a certifikace digitálních repozitářů:

- **Základní certifikace** (*Basic Certification*) – udělena repozitářům, které dostojí směrnicím *Data Seal of Approval* (DSA).
- **Rozšířená certifikace** (*Extended Certification*) - přiřazena repozitářům, které již splnily Základní certifikaci a navíc provedly za externího prozkoumání strukturovaný a veřejně dostupný „samo-audit“ podle ISO 16363 nebo DIN 31644.
- **Formální certifikace** (*Formal Certification*) - po dosažení Základní certifikace repozitář nechal provést plně externí, nezávislý audit a certifikaci dle ISO 16363 nebo DIN 31644.

Repozitáře poté mohou na svých web stránkách a v dokumentech uvádět jeden ze tří symbolů certifikace.

Problémová oblast	Ochranná opatření
důvěryhodnost repozitáře	certifikační a auditní nástroje
ekonomická rizika	ekonomické modely
organizační a procedurální překážky, ekonomická rizika	motivační mechanismy
ekonomická rizika, organizační a procedurální překážky	obchodní modely
institucionální rovina obecně	komplexní plánování
Informační, systémová a institucionální rovina obecně	řízení digitálního životního cyklu

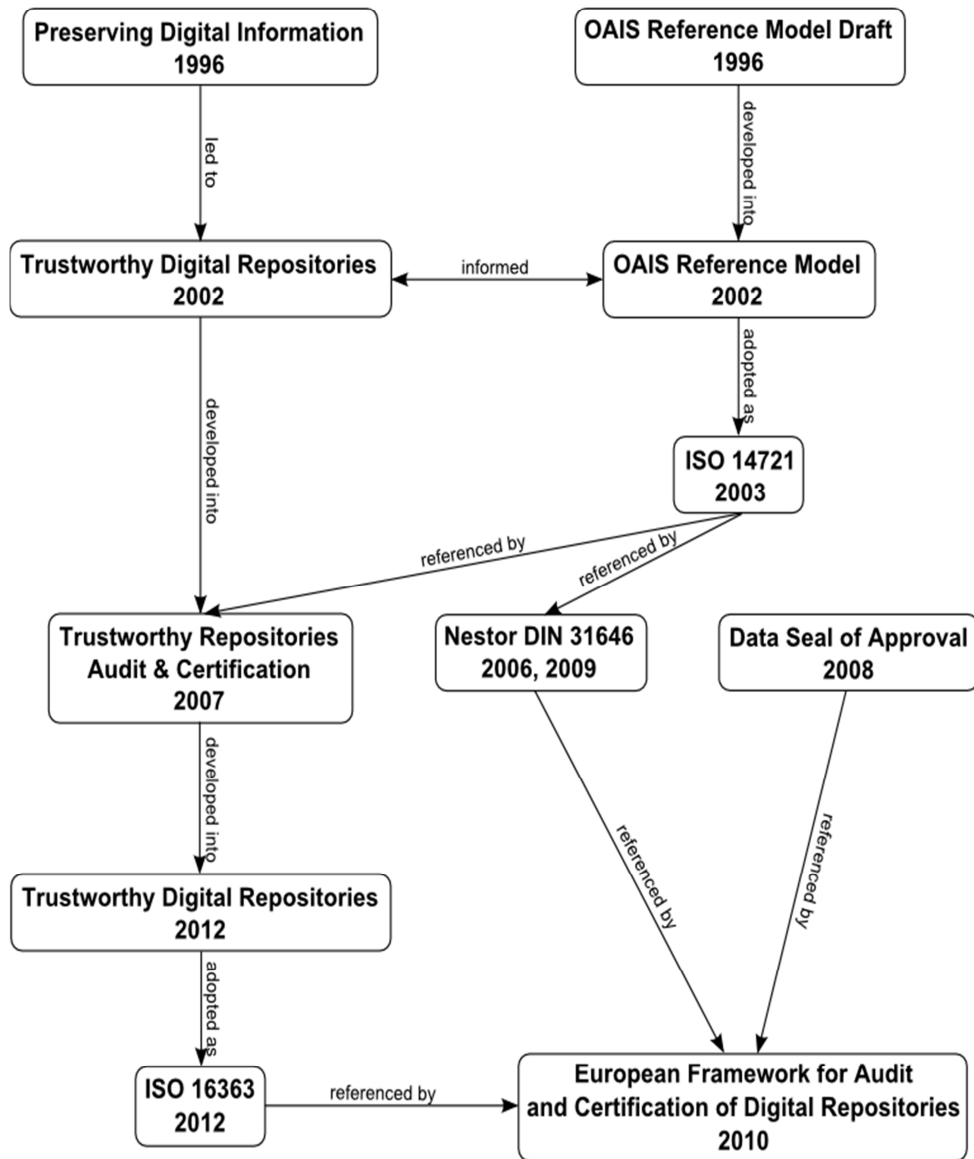
Tab. 2 Ochranná opatření institucionální roviny
[CUBR, 2010]

Tabulka popisuje problémové oblasti dlouhodobé ochrany, kterým repozitáře nejčastěji čelí a jim odpovídající obecná ochranná opatření (dlouhodobé ochrany). Repozitáře musí mít k dispozici motivační mechanismy, komplexní plánování, musí vhodně řídit životní cyklus digitálních dokumentů v rámci stanovených ekonomických modelů, aby dostaly certifikaci či případnému auditu.

Všechny tyto metodiky, auditní a certifikační nástroje slouží především pro instituce uchováající kulturní dědictví – knihovny, muzea a archivy – a hodí se jak v přípravných (pro kontrolu návrhů), tak post-implementačních fázích běžného chodu repozitáře (pro kontrolu implementace). Poslouží však i komerčním a nekomerčním poskytovatelům služeb, vývojářům software a dodavatelům třetích stran pro případnou spolupráci s kulturními institucemi při návrhu, vývoji a provozu technologií pro digitální repozitáře. V neposlední řadě prokáže službu všem zájmovým skupinám důvěryhodnosti repozitáře:

- uživatelům repozitáře, kteří potřebují důvěryhodný zdroj informací dnes i do budoucna
- producentům dat a poskytovatelům obsahu, pro které certifikace repozitáře představuje záruku správně zvoleného partnera
- finančním institucím a grantovým agenturám či jiným organizacím, pro které důvěryhodnost repozitáře znamená garanci správně investovaných prostředků
- samotným digitálním repozitářům, jež auditem a certifikací prokáží své odpodstatnění.

Digital Repository Standards Development



Obr. 37 Vývoj standardů pro digitální repozitáře
 (<http://en.wikipedia.org/wiki/File:Digitalrepositorystandards.png>)

Obrázek popisuje vývoj jednotlivých norem a de facto standardů pro digitální repozitáře. V roce 1994 *Research Libraries Group* (RLG) a *Commission on Preservation And Access* (CPA) založili pracovní skupinu *Task Force on Archiving of Digital Information*, které cílem bylo prozkoumat způsoby, jakými je možné zabezpečit dlouhodobou ochranu a trvalý přístup k digitálním objektům. Na základě výzkumu pracovní skupiny vznikl v roce 1996 dokument *Preserving Digital Information: Report of the task force on archiving of digital documents*. V roce 2000 OCLC a RLG započali spolupráci v oblasti definice atributů pro důvěryhodné digitální repozitáře (*Trustworthy Digital Repositories - 2002*). Ve stejném roce, kdy byl představen dokument *Preserving Digital Information*, vznikl návrh

referenčního modelu OAIS (*OAIS Reference Model Draft* - 1996). Ten se rozvinul do komplexního modelu (*OAIS Reference Model*) a prošel náročným procesem standardizace podle ISO (ISO 14721 - 2003). Z této mezinárodní normy poté vycházely projekty pro audit a certifikaci repozitářů - *Trustworthy Repositories Audit & Certification* (2007), Nestor DIN 31646 (2006, 2009) a *Data Seal of Approval* (2008). Poslední dvě poskytly základ Evropskému rámci pro audit a certifikaci repozitářů (*European Framework For Audit And Certification of Digital Repositories* - 2010). Vrchol „pyramidy“ tohoto rámce pro audit a certifikaci tvoří ISO standard 16363:2012 *Audit and Certification Of Trustworthy Digital Repositories*, jež mimo jiné vycházel i z propracovaných definic atributů důvěryhodných digitálních repozitářů z let 2002 i 2012 (*Trustworthy Digital Repositories* - 2012).

5.1.1 Základní certifikace - Data Seal of Approval

Data Seal of Approval Guidelines (DSA) version 2, July 19, 2013 nabízí nástroj, který má především manažerům repozitářů a digitálních knihoven poskytnout dobrý základ pro dlouhodobou ochranu digitálních dokumentů. V Evropě je DSA součástí integrovaného rámce pro audit a certifikaci repozitářů - *European Framework For Audit And Certification Of Digital Repositories* a slouží jako první základní certifikace (*Basic Certification*). Hlavním centrem pozornosti při hodnocení důvěryhodnosti pomocí DSA jsou digitální objekty a databáze a jejich životní cyklus v repozitářích, knihovnách a archivech. Pro dlouhodobou udržitelnost archivovaných dat musí data splňovat 5 základních principů [DATA, 2013]:

1. lze je najít na internetu
2. jsou přístupná a respektují příslušnou legislativu s ohledem na ochranu autorských práv a osobních údajů
3. jsou dostupná v použitelném formátu
4. jsou spolehlivá
5. lze na ně odkazovat.

Nepříliš rozsáhlý dokument obsahuje celkem 16 zásad/kritérií (viz Příloha č. 5 - Příklad kritéria č. 14 v *Data Seal of Approval*) ve třech základních kategoriích pro aplikaci a ověření kvalitativních aspektů při vytváření, uchování a využití výzkumných dat společenských a humanitních věd v digitální podobě. Aby repozitář dostal označení důvěryhodného repozitáře (*Trusted Digital Repository* - TDR) musí vyhovět minimálně zásadám 4 až 13. Pro získání *Data Seal of Approval* musí navíc umožnit producentům dat a uživatelům naplnění v zásadách 1 až 3 a 14 až 16 [DATA, 2010]:

Kategorie **Producenti dat** (*Data Producers*) – odpovídá za kvalitu dat, kterou ovlivňuje samotná kvalita vědeckého a akademického výzkumu i formát, ve kterém jsou data a doplňkové informace pro jejich uchování a dokumentace v podobě metadat a dalších kontextuálních informací:

1. Producent dat vkládá data do digitálního repozitáře společně s informacemi, které ostatním subjektům umožňují zhodnotit jejich kvalitu a to, nakolik odpovídají etickým a jiným normám platným pro danou disciplínu.
2. Producent dat odevzdává data ve formátech doporučovaných digitálním repozitářem.
3. Producent dat odevzdává data společně s metadaty vyžadovanými digitálním repozitářem.

Kategorie **Datový repozitář** (*Data Repository*) – odpovídá za kvalitu uchování a dostupnost dat v dlouhodobém měřítku. Na kvalitu dat působí dva důležité faktory – procesy a kvalita prostředí organizace, do které je repozitář zasazen + kvalita technické infrastruktury repozitáře:

4. Digitální repozitář má jasně stanovené poslání (*Mission*) v oblasti digitální archivace a uplatňuje jej.
5. Digitální repozitář věnuje dostatečnou péči dodržování právních předpisů a smluv, a to případně včetně těch, které se vztahují k ochraně osob.
6. Digitální repozitář aplikuje zdokumentované procesy a postupy pro správu ukládání dat.
7. Digitální repozitář má plán dlouhodobé ochrany digitálního obsahu v něm uloženého.
8. Archivace probíhá v průběhu celého životního cyklu dat a podle jasně stanovených postupů
9. Digitální repozitář přebírá od producentů dat odpovědnost za zpřístupnění digitálních objektů.
10. Digitální repozitář umožňuje uživatelům najít a použít data a trvale na ně odkazovat.
11. Digitální repozitář zajišťuje integritu digitálních objektů a metadat.
12. Digitální repozitář zajišťuje autenticitu digitálních objektů a metadat.

Kategorie **Příjemci dat** (*Data Consumers*) – odpovídá za kvalitní využití dat. To ovlivňuje zejména míra, do jaké jsou výstupy z vědy a výzkumu dostupné (co nejsnáze a co nejširší cílové skupině) v rámci stanoveného „etického kodexu“ (*Codes Of Conduct*). Jde o národní legislativní rámec a politiky přístupů (např. dostupnost dat podle autorského práva, omezení přístupu pro uživatele dané instituce či země, ve které se repozitář (nejen fyzicky) nachází):

13. Technická infrastruktura výslovně podporuje úkoly a funkce popsané v mezinárodně uznávaných archivních standardech jako je např. OAIS.
14. Uživatel dat dodržuje přístupová pravidla stanovená digitálním repozitářem.
15. Uživatel souhlasí s pravidly pro sdílení a správné využívání znalostí a informací obecně uznávanými v dané oblasti a řídí se jimi.
16. Uživatel respektuje digitálním repozitářem stanovená licenční omezení týkající se užití dat.

Všechny směrnice by měly být podpořeny veřejně dostupným prohlášením (URL odkaz na text, nejlépe v anglickém jazyce). Pokud daný dokument není k dispozici v anglickém jazyce, musí být v tomto jazyce alespoň jeho stručné shrnutí. Hodnocení v první fázi probíhá "samo-auditem", poté vyhodnocené směrnice slouží jako podklad pro udělení 'Data Seal of Approval' od Data Seal of Approval Board. Tato hodnotící komise v průběhu přibližně 2-3 měsíců projde dokumentaci ke všem 16 směrnicím a výsloví verdikt k (ne)udělení certifikace DSA⁷⁹. Hodnotí se stupeň shody (viz Tab 3):

Stupeň hodnocení	Minimální požadavek pro směrnice	Zdůvodnění (při "samo-auditu")	Poznámka a/nebo URL odkaz (pro hodnotící komisi DSA)
0	x	N/A: Není aplikovatelné	Doplňte vysvětlení
1	x	Ne: Ještě jsme nad tím neuvažovali	Doplňte vysvětlení
2	x	Teoreticky: Máme teoretický návrh.	Dodejte URL výchozího dokumentu
3	1,2,7,8,10,11,12,13	Rozpracováno: Jsme v implementační fázi	Dodejte URL podpůrného dokumentu
4	3,4,5,6,9,14,15,16	Implementováno: Tuto zásadu jsme plně implementovali pro potřeby našeho repozitáře.	Dodejte URL podpůrného dokumentu

Tab. 3 Hodnocení shody se směrnicemi Data Seal of Approval [DATA, 2013]

Tabulka popisuje hodnocení shody, resp. splnění podmínek pro udělení pečeti Data Seal of Approval. Existují 4 základní stupně shody (0-4), přičemž je nevyhnutné splnit zejména 3. a 4. stupeň a jako důkaz je nutné odkázat na URL s průkazným materiálem (např. politiky, plány apod).

⁷⁹ Seznam digitálních repozitářů, kterým hodnotící komise DSA přidělila pečeti důvěryhodného repozitáře - <https://assessment.datasealofapproval.org>.

Zásady DSA se brzy dočkají i českého překladu. Ústav výpočetní techniky UK v Praze (Mgr. Andrea Fojtů, Mgr. Eliška Pavlásková), ve spolupráci s Archives New Zealand (Mgr. Jan Hutař, Ph.D.) a firmou Multidata Praha, s.r.o. (Marek Melichar) po dohodě s tvůrci DSA kritérií zabezpečují oficiální překlad do českého jazyka⁸⁰.

5.1.2 Rozšířená a formální certifikace

Rozšířená a formální certifikace představují druhý a třetí (nejvyšší možný) stupeň certifikace repozitářů podle Evropského rámce pro audit a certifikaci digitálních repozitářů (viz výše kapitola 5.1 Metriky, certifikační a auditní nástroje). Jedná se o německý standard DIN 31644, který vychází z pravidel Nestor a ISO standard 16363.

5.1.2.1 NESTOR - Network of Expertise in Long-Term Storage and Long-Term availability of Digital Resources in Germany a DIN 31644 Criteria for trustworthy digital archives

Nestor – *Network of Expertise in Long-Term Storage and Long-Term availability of Digital Resources in Germany* představuje katalog 14 kritérií s podrobným vysvětlením a konkrétními příklady. **Klíčovým konceptem nestor katalogu kritérií pro dlouhodobé uchovávání a dlouhodobé zpřístupňování je důvěryhodnost repozitáře** (*Vertrauenswürdigkeit des Langzeitarchivs*). Celou problematiku vyhodnocení repozitářů člení do tří skupin [NESTOR, 2008]:

- A. **organizační rámec** (*Organisatorischer Rahmen*) – 5 kritérií; digitální repozitář působí v určitém organizačním rámci, který stanovuje jeho cíle, právní podmínky, personální a finanční zabezpečení.
- B. **správa objektů** (*Umgang Mit Objekten*) – 7 kritérií; repozitář analyzuje stanovené cíle a strategie, na základě kterých specifikuje veškeré nutné požadavky a akce na dlouhodobé uchovávání či jiné nutné zacházení s digitálními objekty během jejich životního cyklu. Tento cyklus zodpovídá hlavním fázím referenčního modelu OAIS (funkční entity) – Příjem, Permanentní úložiště a Přístup. Při práci s objekty je nutné dodržet [FRUCHT, cca2009]:
 - **integritu** a autenticitu získaných informací, resp. digitálních objektů. Ty totiž představují klíčový aspekt důvěryhodnosti
 - **dlouhodobé plánování** technických ochranných opatření pro trvalé uchování informací, resp. digitálních objektů

⁸⁰ Viz tisková zpráva: <http://www.datasealofapproval.org/en/news-and-events/news/2013/10/2/dsa-translated-czech/>.

- **standardy** přenosu, uchovávání a využívání informací, resp. digitálních objektů
- **správu dat**, pomocí které je možné dostat z digitálních objektů užitečné informace pro vymezené cílové skupiny.

C. **infrastruktura a zabezpečení** (*Infrastruktur Und Sicherheit*) – 2 kritéria; představují souhrn technických a bezpečnostních záležitostí, včetně komunikačních a síťových služeb, hardware a programového vybavení (software). Chrání tak digitální objekty před systémovými a vnějšími hrozbami.

Hodnocení repozitářů podle abstraktních kritérií⁸¹ vychází ze čtyř základních principů [DOBRATZ, 2007; NESTOR, 2008]:

- **dokumentace** (*Dokumentation, Documentation*) – základní cíle, koncepty, pracovní postupy a specifikace pro digitální repozitář by měly být patřičně zdokumentovány. Včasné prozkoumání dokumentace při plánování může zabránit pozdějším chybám v implementaci navrhovaného řešení. Ukazuje též, že plánovaná architektura a design jsou kompletní a odpovídají standardům kvality a bezpečnosti.
- **transparentnost** (*Transparenz, Transparency*) – týká se zveřejňování relevantních částí dokumentace partnerům a uživatelům a možnosti jejich připomínkování. Partneři (např. producenti dat) tak mají přehled o tom, komu a za jakých podmínek poskytují svá data.
- **adekvátnost** (*Angemessenheit, Adequacy*) – hodnocení musí proběhnout ve vhodném kontextu v souladu s cíly a posláním digitálního repozitáře, jelikož absolutní standard neexistuje. Některá kritéria a požadavky stanovené metodikou tak nemusí být pro daný digitální repozitář relevantní a tím pádem splnitelné.
- **měřitelnost** (*Bewertbarkeit, Measurability*) – dlouhodobá ochrana digitálních materiálů není vždy jednoznačně měřitelnou záležitostí, přesto je možné nalézt jisté indikátory, které poukazují na stupeň důvěryhodnosti repozitáře.

První verze nestor katalogu byla vydána v německém i anglickém jazyce v roce 2006. O tři roky později byla zveřejněná druhá, přepracovanější verze, tato však pouze v německém jazyce.

5.1.2.2 TRAC - Trustworthy Repositories Audit & Certification: Criteria and Checklist, ISO 16363

Trustworthy Repositories Audit & Certification : Criteria and Checklist, známý pod

⁸¹ Podrobný přehled nestor kritérií je v přílohách: Příloha č.6 – Seznam kritérií katalogu nestor

akronymem TRAC:CC⁸² představuje OAIS-kompatibilní auditní a certifikační nástroj digitálních repozitářů. Vznikl z iniciativy konsorcia knihoven RLG a Národního archivu USA - NARA (*National Archives And Records Administration*). V začátcích existoval TRAC jako kontrolní seznam RLG/NARA Audit Check-list, který se v roce 2012 stal součástí normy ISO/IEC 16363 Space data and information transfer systems -- Audit and certification of trustworthy digital repositories⁸³.

Vzhledem ke své organizační, časové a finanční náročnosti má smysl zejména pro rozsáhlejší digitální knihovny, repozitáře a archivy. Jako auditní nástroj slouží pro vyhodnocení spolehlivosti, závaznosti a připravenosti institucí převzít na sebe zodpovědnost za dlouhodobé uchování obsahu. Je určen pro ty, kteří pracují pro nebo jsou zodpovědní za digitální repozitáře, knihovny nebo archivy a hledají objektivní prostředek pro vyhodnocení důvěryhodnosti jejich repozitáře.

Je nutné si však uvědomit, že repozitář jako organizační celek se dotýká pracovníků na nejrozličnějších úrovních. Vedení a nižší management bude muset mít povědomí alespoň o požadavcích z části A. Organizační struktura. Systémoví a síťoví správci, kteří zodpovídají za mnohé části infrastruktury budou pracovat s částí C. Technologie, technická infrastruktura a bezpečnost. Producenti a příjemci dat naleznou relevantní informace především v dokumentaci pro část A. a B. Správa digitálních objektů:

A. Organizační infrastruktura

1. Řízení a životaschopnost organizace
2. Organizační struktura a personální zabezpečení

⁸² Dokument dostupný na: <http://www.crl.edu/PDF/trac.pdf>.

⁸³ ISO 16363:2003 je součástí souboru několika dalších norem, které jsou nezbytným podkladem při návrhu, plánování, administraci digitálních repozitářů a archivů. Jde o:

ISO 9000 – soubor norem managementu jakosti, jež přináší organizacím zvýšení důvěryhodnosti prostřednictvím tohoto mezinárodně uznávaného certifikátu.

ČSN ISO/IEC 17799:2006 Informační technologie - Bezpečnostní techniky - Soubor postupů pro management bezpečnosti informací poskytuje doporučení a obecné principy pro vymezení, zavedení, udržování a zlepšování systému managementu bezpečnosti informací v organizaci, a to na základě analýzy rizik. Může proto sloužit i jako praktický průvodce při sestavování bezpečnostních politik, standardů a postupů v organizacích.

ISO 15489-1:2001 Information and documentation -- Records management -- Part 1: General a ISO/NP TR 15489-2 Information and documentation -- Records management -- Part 2: Guidelines definují systematický a na procesy orientovaný přístup k managementu dat a záznamů.

ISO 14721:2002 Space data and information transfer systems -- Open archival information system (OAIS) -- Reference model poskytuje referenční model pro identifikaci účastníků, jejich role a zodpovědnosti při dlouhodobé ochraně digitálního obsahu v digitálních repozitářích a archivech.

3. Procedurální zodpovědnost a strategický rámec
4. Finanční udržitelnost
5. Smlouvy, licence a závazky

B. Správa digitálních objektů

1. Ingest: akvizice obsahu
2. Ingest: tvorba archivních balíčků
3. Plánování dlouhodobé ochrany
4. Archivní úložiště & ochrana/správa AIP balíčků
5. Informační management
6. Správa přístupu

C. Technologie, technická infrastruktura a bezpečnost

1. Systémová infrastruktura
2. Vhodné technologie
3. Bezpečnost

Proces hodnocení důvěryhodnosti a certifikace repozitáře podle TRAC je rozdělený do 84 kritérií ve třech základních kategoriích a 14 podkategoriích. Jelikož **TRAC představuje velmi nákladnou záležitost, mnoho institucí upřednostňuje „samo-audit“ (self-audit) pomocí nástroje DRAMBORA.**

5.1.3 Další auditní a certifikační nástroje

5.1.3.1 DRAMBORA - Digital Repository Audit Method Based on Risk Assessment

Auditní nástroj *Digital Repository Audit Method Based On Risk Assessment*⁸⁴ (DRAMBORA) je výsledkem úsilí *Digital Curation Centre* - DCC a *Digital Preservation Europe* – DPE. Projekt, probíhající v letech 2006-2007 byl financován Evropskou komisí a měl za cíl zlepšit koordinaci a kooperaci mezi členskými státy v oblasti dlouhodobé ochrany. **V současné době tvůrci DRAMBORA usilují o uznání nástroje za mezinárodní ISO standard** (v rámci technické komise ISO/TC 46/SC 11 – Archive/records management).

Zatímco TRAC a Nestor pro získání certifikátu a „nálepky“ důvěryhodného repozitáře sestavili dotazníky k ověření definovaných kritérií, DRAMBORA interactive představuje metodologii „sebe-hodnocení“ (*Self-Assessment*) repozitáře, které je v kontextu projektu DRAMBORA vnímáno jako interní audit. Na rozdíl od jiných auditních a certifikačních metod je více

⁸⁴ <http://www.repositoryaudit.eu>

zaměřen na kvantitativní než kvalitativní hodnocení. Zastupuje **metodologický přístup k hodnocení důvěryhodnosti repozitáře**.

DRAMBORA chápe digitální ochranu (*Digital Curation*) jako aktivitu managementu rizik, v rámci kterého je nutné podchytit nejistoty a hrozby, které ohrožují autenticitu, integritu, pochopení digitálních objektů a jejich proměnu na rizika, které je možné zachytit a do budoucna odstranit. Jde o on-line nástroj, který napomáhá auditorům k hodnocení současného stavu digitálního repozitáře. Cílem není samoučelné hledání chyb, nýbrž kritické vyhodnocení jak pozitivních, tak negativních aspektů funkčnosti repozitáře. Postup auditu je stanovený v 6 fázích (fáze 4 se skládá ze čtyř kroků, vid' obr. 39). Jednotlivé fáze jsou:

1. Identifikace organizačního kontextu

- Specifikace mandátu repozitáře nebo organizace, ve kterém repozitář existuje
- Soupis cílů repozitáře

1. Dokumentové politiky a regulační rámec

- Soupis strategických plánovacích dokumentů repozitáře
- Soupis legislativních, regulačních a smluvních rámců a dohod, ze kterých repozitář vychází
- Soupis dobrovolných závazků, které se repozitář rozhodl plnit
- Soupis jiných dokumentů a principů, na kterých repozitář stojí

2. Identifikace aktivit, prostředků a jejich vlastníků

3. Identifikace rizik - Identifikace rizik spojených s aktivitami a prostředky repozitáře

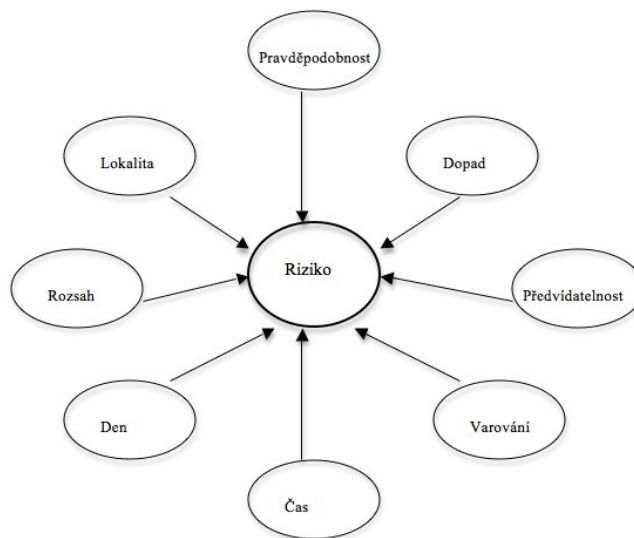
4. Hodnocení rizik - Vyhodnocení identifikovaných rizik

5. Management rizik - Řešení rizik

Prvním krokem by měla být registrace repozitáře, resp. organizace, která daný repozitář spravuje. Auditori (interní, případně externí) při ní vytvoří profil organizace, popíše a zdokumentují mandát, cíle, aktivity instituce, která má auditem projít. Neméně důležitým krokem je následná identifikace rizik a jejich atributů (viz Obr. 38), které můžou v souvislosti se stanovenými aktivitami nastat a vyhodnocení jejich pravděpodobnosti výskytu a potenciálního dopadu (důsledků) na repozitář. Aby nešlo jen o bezúčelný audit je nevyhnutné promyslet správná, resp. nejvhodnější řešení pro identifikovaná rizika.

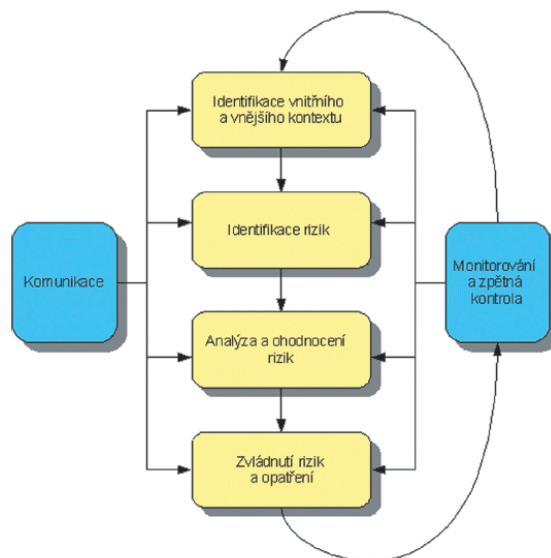
Značnou výhodou auditu DRAMBORA je možnost specifikace dlouhodobého plánu spolu se střednědobými cíly a krátkodobými úkoly. DRAMBORA audit poukazuje na rizika pramenící z neexistence krizových plánů. Přínosem auditu tedy není až tak zjištění, že je potřebné „něco“

udělat. Podstatnější je spíše stanovení priorit řešených úkolů. Zejména ve fázi definování rizik se ukáže, že absence řešení některých otázek by se snadno mohla stát příčinou vzniku celé řady problémů, kterých kombinace by v krajním případě mohla být pre repositář kritická. **Je nutné však zdůraznit, že zejména začátky práce s online nástrojem DRAMBORA nejsou snadné.**



Obr. 38 Atributy rizik
[WALLACE, 2004]

Obrázek ukazuje 8 základních atributů, které ovlivňují riziko. Jde o čas, den, rozsah, lokalitu, pravděpodobnost výskytu, dopad, možnost předvídatelnosti a způsob varování.



Obr. 39 Struktura hodnocení repozitáře pomocí nástroje DRAMBORA
[HUTAŘ, 2008]

Obrázek popisuje proces hodnocení repozitáře pomocí auditního nástroje DRAMBORA. Celý postup začíná identifikací vnitřního a vnějšího kontextu repozitáře (např. specifikace mandátu repozitáře nebo organizace, ve kterém repozitář existuje, soupis cílů repozitáře.

DRAMBORA nenavádí k jednoznačnému hodnocení „úspěš/neúspěš“. Vychází se z předpokladu, že důvěryhodnost ohrožují rizika, které je nutné dostatečnými prostředky a formou podchytit a potírat. **Zajímavé je, že čím více rizik se během auditního procesu objeví, tím má repozitář více nakročeno k důvěryhodnosti repozitáře. To z toho důvodu, že si je vědom svých problémů a rizik, které má snahu vyřešit.**

5.2 Ekonomické modely

Jakákoliv instituce spravující repozitář musí mít dostatečné povědomí o svém „finačním zdraví“, tedy o příjmech i nevyhnutných výdajích. Odhadování nákladů či dokonce přesná kalkulace nákladů specifických aktivit je pro mnohé zodpovědné osoby a organizace velmi těžký úkol. Je to navíc komplikováno i tím, že je nutné počítat i s IT náklady. Výsledkem pak bývá, že organizace často nejsou schopné odůvodnit své výdaje. Proto modely nákladů (ekonomické modely) mohou dané výdaje jednodušeji a spolehlivěji prokázat. Porozumění nákladů totiž umožní repozitářům provádět informovaná rozhodnutí o investicích na základě zvážení investic, jež mají největší dopad. Ekonomický model, resp. model nákladů, definuje APARSEN [KAUR, 2013] jako

„reprezentaci aktivit dlouhodobé ochrany digitálních dat, které je možné sdílet, prozkoumat a kriticky ohodnotit a jejichž cílem je objasnit náklady spojené s aktivitami v oblasti dlouhodobé ochrany.“

Výběr vhodného ekonomického modelu je podmíněn především záběrem, resp. sektorem, který daný model pokrývá, pro jaký typ organizace je určen a jaký je jeho hlavní účel (stanovení rozpočtu, účetnictví, alokace zdrojů/nákladů či získávání financí pro specifické projekty nebo aktivity). Dále je nutné zvážit, jaká data bude potřebné získat/dodat, jaké jsou silné a slabé stránky a jak jsou parametry nákladů v modelu stanoveny. V neposlední řadě je potřeba vědět, jestli daný model nabízí i určité nástroje (tabulkový procesor či webové formuláře) pro snadnější práci. Za poslední desetiletí vzniklo nemálo ekonomických modelů. Organizace, resp. instituce spravující repozitáře mají snahu upravovat stávající modely pro své potřeby nebo vytvářet další, vlastní, pokud nejsou pro ně ty dostupné zcela vyhovující. Z důvodu lepší orientace v ekonomických modelech, vydal APARSEN⁸⁵ [KAUR, 2013] podrobnou zprávu o nejznámějších ekonomických modelech pro digitální repozitáře. Aby jejich hodnocení bylo co nejvíc objektivní, jednotlivé hodnotící parametry modelů nákladů byly mapovány vůči atributům důvěryhodných repozitářů, definovaných v ISO standardu ISO 16363⁸⁶. Ten totiž slouží jako jakýsi výchozí bod a společné měřítko pro danou problematiku důvěryhodnosti.

Nejčastěji tyto ekonomické modely vznikají jako součást národních či mezinárodních projektů⁸⁷ (podrobněji o nejznámějších modelech viz Tab 4 Stručný přehled modelů nákladů, str. 128-130). Mezi nejznámější modely nákladů počítáme Lifecycle Information for E-Literature, LIFE3, Cost Model for Digital Preservation, CMDP, Keeping Research Data Safe 1 and 2, PrestoPRIME cost model for digital storage, Cost Estimation Toolkit, Digital Preservation for libraries, DANS, ENSURE (analyzující cloudové řešení úložišť).

⁸⁵ APARSEN - *Alliance for Permanent Access to the Records of Science in Europe Network* (<http://www.alliancepermanentaccess.org>) seskupuje organizace a výzkumníky z nejrůznějších oborů za účelem koherence, koheze a spojení výzkumu v oblasti dlouhodobé ochrany a zpřístupňování digitálních dat.

⁸⁶ Standard ISO 16363 je rozdělen do 5 částí. První dvě se věnují struktuře dokumentu a přehledu certifikačních a auditních kritérií. Části 3 až 5 nabízí metriky pro audit a certifikaci důvěryhodného digitálního repozitáře, tj. Organizační infrastruktura, Správa digitálních objektů, Infrastruktura a Bezpečnostní správa rizik.

⁸⁷ Například projekty ENSURE (2007-2013) a PrestoPRIME (2009-2012) byly financovány Evropskou komisí (v rámci 7. rámcového projektu). Oba modely byly vyvíjeny jako integrální součást konkrétního řešení dlouhodobé ochrany. Zatímco ENSURE se zaměřuje na zdravotnictví, klinické testy a finanční sektor. PrestoPRIME (s několika dostupnými nástroji) se orientuje na praktická řešení pro dlouhodobou ochranu digitálních mediálních objektů a programů. DANS Cost Model při svém vzniku vycházel z konkrétních případových studií v sektoru kulturního dědictví.

CET (Cost estimation toolkit)⁸⁸	<ul style="list-style-type: none"> - nástroj vyvinutý NASA a SGT⁸⁹ - odhad nákladů aktivit u vědeckých dat; aplikovatelné pro systémy dlouhodobé ochrany - dvouúrovňový nástroj; softwarový balíček dostupný na: http://opensource.gsfc.nasa.gov/projects/CET/#software - forma: tabulkový procesor
CMDP (Cost Model for Digital Preservation)⁹⁰	<ul style="list-style-type: none"> - model vyvinutý Royal Library of Denmark a Danish National Archives - odhad nákladů dlouhodobé ochrany (vkládání dat, plánování dlouhodobé ochrany a migrace, archivní uchování) - zaměřený na organizace uchovávající kulturní dědictví - dostupný na: http://www.costmodelfordigitalpreservation.dk - forma: tabulkový procesor dostupný na http://www.costmodelfordigitalpreservation.dk/contact/archival-storage/doc
DANS cost model⁹¹	<ul style="list-style-type: none"> - nástroj vyvinutý DANS⁹² - výpočet nákladů archivních dat na základě nákladů spojených s aktivitami (activity based costing), které jsou pro dlouhodobou ochranu dat nevyhnutná
DP4lib - Digital Preservation for libraries⁹³	<ul style="list-style-type: none"> - model sestaven Deutsche Nationalbibliothek - výpočet nákladů na dlouhodobou ochranu formou služeb od třetích stran; není omezen na konkrétní sektor - forma: tabulkový procesor
ENSURE project⁹⁴	- model a nástroj vyvíjen v rámci projektu ENSURE ⁹⁵ , financovaný

⁸⁸ Podrobnější informace v příspěvku z konference: FONTAINE, K., HUNOLT, G., BOOTH, A., BANKS, M. Observations on Cost Modeling and Performance Measurement of Long Term Archives in PV2007. Dostupné na: http://www.pv2007.dlr.de/Papers/Fontaine_CostModelObservations.pdf.

⁸⁹ *Stinger Ghaffarian Technologies, Inc.* (SGT) je aeronautická společnost založená v roce 1994 a zaměřena na oblast podpory projektů v oblasti systémového inženýrství, vědecké analýzy a modelování, projektového managementu a informačních technologií.

⁹⁰ Další doporučená literatura: KEJSER, U. B., NIELSEN, A. B., THIRIFAYS, A. Cost Model for Digital Preservation: Cost of Digital Migration, *International Journal of Digital Curation*, 6(1), 2011. doi:10.2218/ijdc.v6i1.186. Dostupné na: <http://www.ijdc.net/index.php/ijdc/article/view/177>.

⁹¹ Podrobnější informace v článku: PALAIOLOGK, A. S., ECONOMIDES, A. A., TJALSMA, H. D., SESINK, L. B. An activity-based costing model for long-term preservation and dissemination of digital research data: the case of DANS. *International Journal on Digital Libraries*, 12(4), 195–214. doi:10.1007/s00799-012-0092-1. Dostupné z: <http://link.springer.com/article/10.1007%2Fs00799-012-0092-1>.

⁹² *Data Archiving and Networked Services* (DANS) je institut založen *Royal Netherlands Academy Of Arts And Sciences* (KNAW) a *Netherlands Organisation For Scientific Research* (NWO). Jeho cílem je propagovat udržitelný přístup k digitálním výzkumným datům. DANS sestavili certifikační rámec *Data Seal of Approval* (<http://www.datasealofapproval.org/>).

⁹³ Podrobnější informace ke stažení na: <http://aparsen.digitalpreservation.eu/pub/Main/CostModels/DP4lib-Cost-By-Service-CostModel.docx>.

⁹⁴ Doplňkové čtení: BADAWY, M., SHEHAB, E., BAGULEY, P., WILSON, M.D. Towards a cost model for long term digital preservation. Proc. 2012 ISPA/SCEA Joint International Conference: Assuring cost efficiency: global solution, Brussels, Belgium, 14-16 May 2012. Dostupné z:

	<p>EK v 7. rámcovém programu (únor 2011 - leden 2014) - http://www.ensure-project.eu</p> <ul style="list-style-type: none"> - odhaduje náklady aktivit spojených s dlouhodobou ochranou dat v oblasti zdravotnictví, klinických testů, finančnictví; může být rozšířen na výrobní sektor - zaměřuje se na odhad financí dat v cloudových řešeních - pracuje s nejistotami, riziky a otázkami zastarávání formátů
ISIS facility model ⁹⁶	<ul style="list-style-type: none"> - vyvinutý Cranfield University ve spolupráci s STFC⁹⁷ - speciálně pro účely zařízení ISIS (ISIS facility) na STFC, není proto využitelný v jiných oblastech
KRDS - Keeping research data safe (KRDS + KRDS2) ⁹⁸	<ul style="list-style-type: none"> - nástroje financovány a vyvíjeny partnery JISC - Charles Beagrie Ltd., UKOLN, DCC, UKDA - UK Data Archive, ADS - Archeology Data Service, OCLC - model zaměřen na náklady spojené s aktivitami, nabízí seznam užitek a teoretické metriky pro výzkumná data - nástroje pro analýzu užitek, hodnot a dopadu, podrobněji na http://www.beagrie.com/krds.php - má širší využití
LIFE - Life Cycle Information for E-literature	<ul style="list-style-type: none"> - model vyvinutý University College London a British Library (projekt LIFE3 je financován JISC a Research Information Network) - forma: model přetaven do podoby komplexního webového nástroje pro odhad nákladů dlouhodobé ochrany; nástroj dostupný na http://www.life.ac.uk/tool/
OECD - International Standard Cost Model Manual	<ul style="list-style-type: none"> - model vyvinutý Standard Cost Model Network - stanovuje administrativní náklady, nabízí transparentní postupy - manuál v písemné formě dostupný na http://www.oecd.org/regreform/regulatorypolicy/34227698.pdf
PrestoPRIME ⁹⁹	<ul style="list-style-type: none"> - model vyvíjený jako součást projektu EK (7. rámcový program) - http://www.PrestoPRIME.eu/ - nabízí výpočet nákladů a dlouhodobý odhad finančních potřeb pro masovou digitalizaci audiovizuálních dokumentů

<http://epubs.stfc.ac.uk/bitstream/7711/Towards%20a%20Cost%20Model%20for%20Long%20Term%20Digital%20Preservation.pdf>.

⁹⁵ Projekt ENSURE (*Enabling Knowledge Sustainability Usability And Recovery For Economic Value*) je zaměřen na oblast dlouhodobé použitelnosti (*Long-term Usability*) dat vyprodukovaných v zdravotní péči i klinických testech, v oboru finančnictví. Podrobněji na: <http://ensure-fp7-plone.fe.up.pt/site/about>.

⁹⁶ Podrobnější informace na: <http://ensure-fp7-plone.fe.up.pt/site/Poster.pdf>.

⁹⁷ *Science & Technology Facilities Council* je ve Spojeném Království Velké Británie a Irska jednou ze sedmi veřejně financovaných výzkumných rad (*Research Council*) zodpovídající za podporu a koordinaci výzkumu, inovací a vývoje pro oblast technologií.

⁹⁸ Podrobnější informace: BEAGRIE, N., LAVOIE, B., AND WOOLLARD, M. Keeping Research Data Safe 2. 2010. Dostupné z:

<http://www.jisc.ac.uk/media/documents/publications/reports/2010/keepingresearchdatasafe2.pdf>.

⁹⁹ Další informace: WESTERHOF, H., UBOIS, J., SNYDERS, M. Financial Models and Calculation Mechanisms. PrestoPRIME Consortium: 2011. Dostupné z:

https://prestoprime.ina.fr/public/deliverables/PP_WP6_D6.3.1_FM_calculation_R0_v1.01.pdf.

4C - Collaboration to Clarify the Costs of Curation	-projekt s cílem pomoci organizacím objasnit náklady na dlouhodobou ochranu - financovaný Evropskou komisí, své aktivity začal v únoru 2013 a výsledky mají být prezentovány v lednu 2015.
--	---

Tab. 4 Stručný přehled modelů nákladů
[KAUR 2013]

Je nutné si uvědomit, že náklady pro chod (digitálních) repozitářů, knihoven a archivů a dlouhodobou ochranu zasahují do celého životního cyklu digitálních dokumentů a mohou, kromě jiných, obsahovat [KAUR, 2013]: náklady na import dokumentů (validace při přijetí), migraci formátů či emulaci dokumentů, periodický upgrade a obnovy hardware, doplňování sekundárních informací, resp. reprezentací (*Representation Information*), vyhledávání a vytváření DIP¹⁰⁰, až po samotnou likvidaci dat.

Je tedy jasné, že nejde o jednoduchou záležitost. Z toho důvodu vznikají nejrůznější nástroje, metodiky či modely odhadů nákladů, z nichž nejznámější je právě LIFE - *Life Cycle Information for E-Literature*.

5.2.1 LIFE (Life Cycle Information for E-Literature)

Projekt LIFE (2005-2006) a jeho navazující projekty LIFE2 (2007-2008) a LIFE3 (2009-2010) vznikly ze spolupráce Britské národní knihovny (*The British Library*) a *University College Of London* z grantových prostředků JISC¹⁰¹. Cílem bylo zmapovat náklady na činnosti spojené zejména s dlouhodobou ochranou (z pohledu životního cyklu dokumentů). **Finanční možnosti totiž představují důležitý faktor při plánování a rozhodování se pro konkrétní strategii dlouhodobé ochrany.** LIFE vypracoval metodiku pro kalkulaci nákladů dlouhodobé ochrany digitálních dat na pět, deset a sto let. Je určena pro středně velké až rozsáhlé (digitální) repozitáře, knihovny a archivy či paměťové instituce (archivy a muzea) a také akademické a výzkumné instituce obecně. Vývoj metodiky LIFE probíhal od sestavení vlastního modelu, přes obecné kalkulace, kalkulace s proměnnými v příkladech až po možnost zadávání zdrojových dat do webového nástroje LIFE.

Přesnost modelu vychází z kvality a kvantity vstupních dat. Přidělením nákladů co nejvíce relevantním částem a aplikováním obecného modelu *Generic LIFE Preservation Model* je možné

¹⁰⁰ DIP (*Dissemination Information Package*) jako informační balíček pro zpřístupňování uživatelů podle ISO 14721:2003 Space data and information transfer systems -- Open archival information system -- Reference model.

¹⁰¹ JISC (původně jako *Joint Information Systems Committee*, dnes pouze jako JISC; <http://www.jisc.ac.uk>) představuje britskou grantovou agenturu, která podporuje projekty z oblasti osvojování si a využívání technologií při výuce, učení a výzkumu.

podle autorů dosáhnout poměrně přesného odhadu nákladů. Zatímco projekt LIFE1 se zaměřil výhradně na ekonomické aspekty digitálních dokumentů, projekt LIFE2 počítal již i s digitalizovanými dokumenty a nabídl komparační mechanismy nákladů dlouhodobé archivace digitálních a tištěných verzí dokumentů [CUBR, 2010]. LIFE1 k sestavení výsledního modelu využil poznatky z případových studií Britské národní knihovny a *University College Of London*. LIFE2 nechal přezkoumat aplikaci modelu i externě - na různých životních cyklech v nejrůznějších typech institucí¹⁰².

¹⁰² Případové studie *SHERPA DP Case study*, *SHERPA –LEAP Case study*, *British Library Newspaper Case study* - *Burney Digital Newspapers*, *British Library Newspapers Case Study* - *Legal Deposit of Newspapers*, lze najít na webových stránkách projektu LIFE - <http://www.life.ac.uk/2/documentation.shtml>.

5.2.1.1 Metodologie LIFE

V původním modelu LIFE1 byl testován výpočetní vzorec [LIFE, 2014], ve kterém jsou náklady rozděleny do šesti různých časových období digitálního životního cyklu uchovávaných dokumentů a na jejich základě poté vypočítal celkové náklady:

$$L_T = Aq + I_T + M_T + Ac_T + S_T + P_T$$

kde **T** zastupuje čas (*Time*), **L** představuje celkové náklady v čase od 0 do **T** (*Complete Lifecycle Cost*), **Aq** = **akvizice** (*Acquisition*), **I** = **příjem dat** (*Ingest*), **M** = **metadata** (*Metadata*), **Ac** = **přístup** (*Access*), **S** = **uchování** (*Storage*) a **P** = **dlouhodobá ochrana** (*Preservation*). Tento model byl na základě zpětné vazby a reálných případových studiích z LIFE2 ve své verzi LIFE3 pozměněn na:

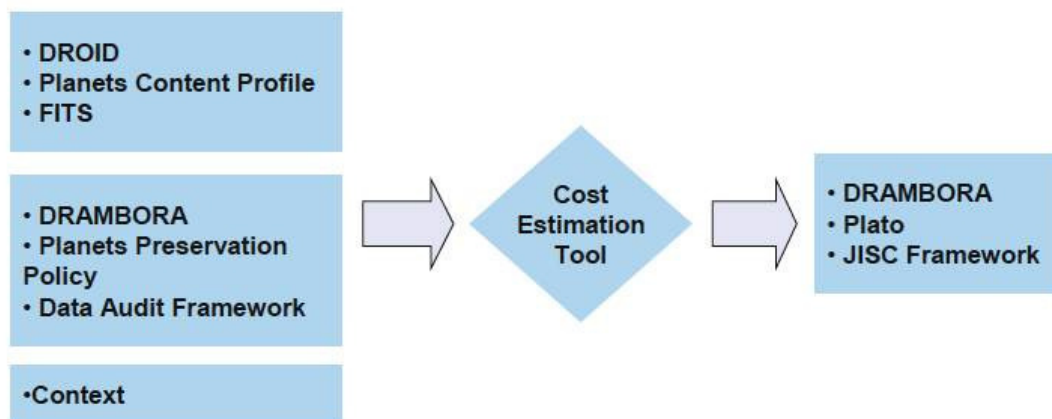
$$L_T = C + Aq + I_T + M_T + BP + CP + Ac_T$$

kde **T** zastupuje jednotku času (*Time*), **L** představuje celkové náklady v čase od 0 do **T** (*Complete Lifecycle Cost*), **C** = **vytvoření** (*Creation*), **Aq** = **akvizice** (*Acquisition*), **I** = **příjem dat** (*Ingest*), **M** = **metadata** (*Metadata*), **BP** = **bitová ochrana** (*Bit-stream Preservation*), **CP** = **ochrana obsahu** (*Content Preservation*) a **Ac** = **přístup** (*Access*).

LIFE3 nabídl sérii modelů nákladů pro každou fázi a složku digitálního životního cyklu (viz Tab. 5), dále jednoduchý webový nástroj pro odhad nákladů (včetně podpory jednoduchého vkládání dat) a integraci do dalších nástrojů pro komplexnější vyhodnocování nákladů a aktivit spojených s dlouhodobou ochranou (viz Obr. 38 Integrace LIFE3 do nástrojů).

Komponent životního cyklu	Vysvětlení
Životní cyklus	Proces od vytvoření až po přístup a dlouhodobou ochranu určitého digitálního objektu. Tento proces může být dále rozdělen na řadu dalších samostatných procesů.
Fáze životního cyklu	Vysoká úroveň procesu v rámci životního cyklu. Nabízí možnost seskupení souvisejících složek životního cyklu. Procesy v rámci Fáze životního cyklu se obvykle (opětovně) vyskytují ve stejném časovém okamžiku.
Složka životního cyklu	Jasný a signifikantní proces životního cyklu, který nabízí užitečné informace o nákladech pro organizace, které vykonávají plánování, hodnotící a komparační testy.
Podsložka životního cyklu	Navrhovaný klíčový prvek Složky životního cyklu není dostatečně signifikantní, aby si zasloužil začlenění jako v případě Složky životního cyklu.

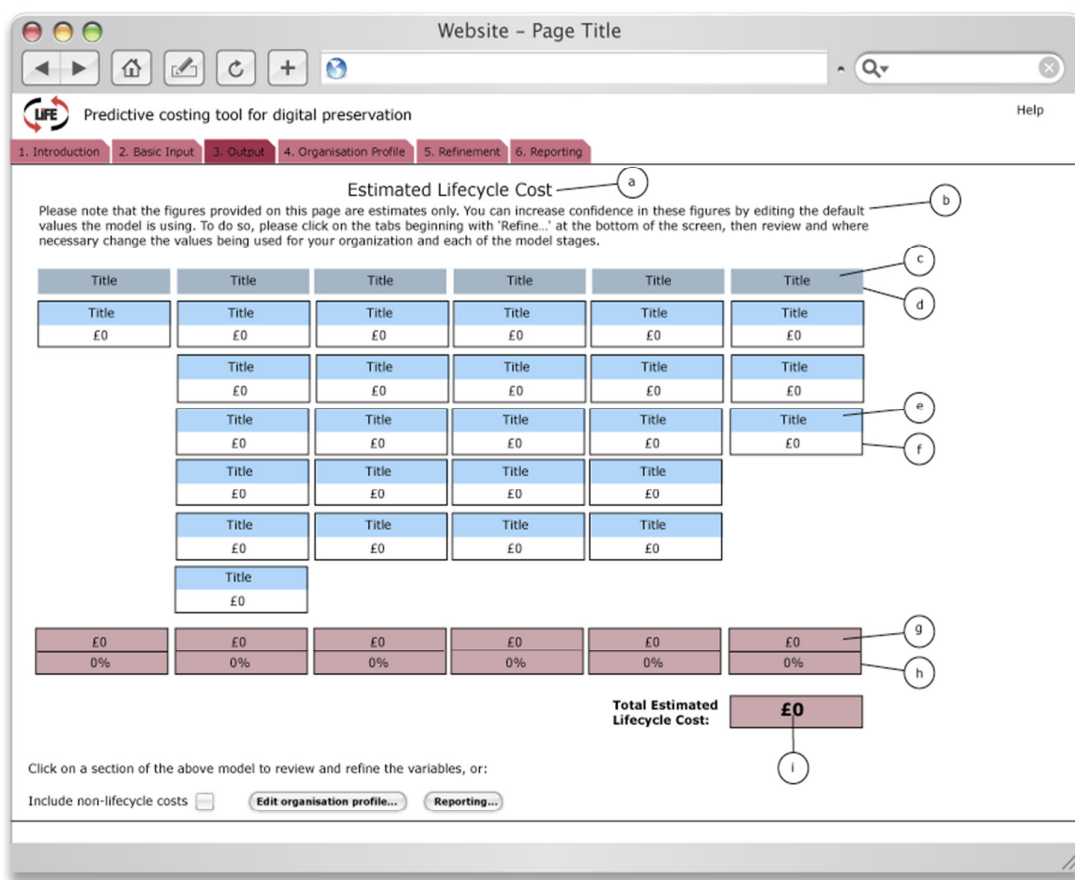
Tab. 5 Komponenty životního cyklu modelu LIFE
[LIFE, 2013]



Obr. 40 Integrace LIFE3 do dalších nástrojů
[HOLE, 2010]

Obrázek popisuje jak nástroje nejrůznějšího typu – od nástrojů na identifikaci formátu (DROID, *Planets Content Profile*, FITS), nástrojů pro identifikaci a podchycení rizik ochrany dat (DRAMBORA, *Planets Preservation Policy*, *Data Audit Framework*) a jejich kontext (*Context*) – ovlivňují samotný nástroj pro vyhodnocování nákladů a aktivit spojených s dlouhodobou ochranou. V té podobě je poté integrován do nejrůznějších nástrojů a pomůcek (DRAMBORA, Plato, JISC Framework).

I proto verze 3 nabízí šablony pro nejčastější typ obsahu a organizační profil při respektování detailních vstupů a specifických výstupů. Doplňkovou vlastností je automatické vyplňování specifických dat podle výstupů z případových studií s možností úpravy profilu. Modely jsou k dispozici v podobě excelovské tabulky, umožňující technicky zdatnějším uživatelům vyšší „customizaci“. Webový nástroj *Predictive Costing Tool For Digital Preservation*, je vytvořen ve spolupráci s HATII (*The Humanities Advanced Technology And Information Institute At The University Of Glasgow*), představuje přístupnější a uživatelsky přívětivější verzi LIFE modelu.



Obr. 41 Screenshot z webového nástroje LIFE
[Hole, 2010]

Screenshot ukazuje jak vypadá základní rozhraní webového nástroje LIFE pro mapování nákladů pro dlouhodobou ochranu. Práce v něm je rozdělena do 6 základních částí (červené záložky v levém horním rohu): 1. Úvod (*Introduction*), 2. Základní vstupy (*Basic Input*), 3. Výstupy (*Output*), 4. Organizační profil (*Organization Profile*), 5. Upřesnění (*Refinement*), 6. Vykazování (*Reporting*). Na základě vyplněných údajů lze poté získat konečný finanční odhad (*Total Estimated Lifecycle Cost*).

Na rozdíl od excelovské tabulky, webový nástroj¹⁰³ zobrazuje pouze data, které jsou relevantní pro uživatele. Pro udržitelnost aplikace do budoucna byla sestavena tak, aby dodatečné úpravy nevyžadovaly vývojáře, ale pouze úpravy administrátora.

5.2.2 4C - Collaboration to Clarify the Costs of Curation

Cílem projektu 4C - *Collaboration to Clarify the Costs of Curation*, jak již samotný název napovídá, je pomoci organizacím napříč Evropou stanovit náklady na dlouhodobou ochranu. Projekt je financovaný Evropskou komisí, své aktivity začal 1. února 2013 a výsledky by měly být prezentovány ke konci ledna 2015. Konsorcium organizací, mající poměrně rozsáhlé zkušenosti s modelováním a odhadem nákladů na dlouhodobou ochranu (viz Tab 6 Seznam partnerských institucí projektu 4C, str. 135-136). Zahrnuje národní knihovny a archivy, členské organizace, poskytovatele služeb, výzkumné instituce a SME¹⁰⁴.

Dánsko	
	Danish National Archives
	Royal Library of Denmark
Estonsko	
	National Library of Estonia
Nizozemí	
	DANS - Data Archiving and Network Service (KNAW/NWO)
Německo	
	Deutsche National Bibliothek, for the network of expertise in digital preservation 'nестor'
Portugalsko	
	Institute for Information Systems and Computing Research (INESC-ID)
	Keep Solutions
Rakousko	
	Secure Business Austria
Velká Británie	

¹⁰³ Webová aplikace byla vytvořena pomocí objektově-orientovaného PHP open source rámce Symfony a MySQL database [HOLE, 2010].

¹⁰⁴ SME (*Small And Medium Businesses*) - malé a střední podniky. Jde vesměs o podniky do 250 zaměstnanců.

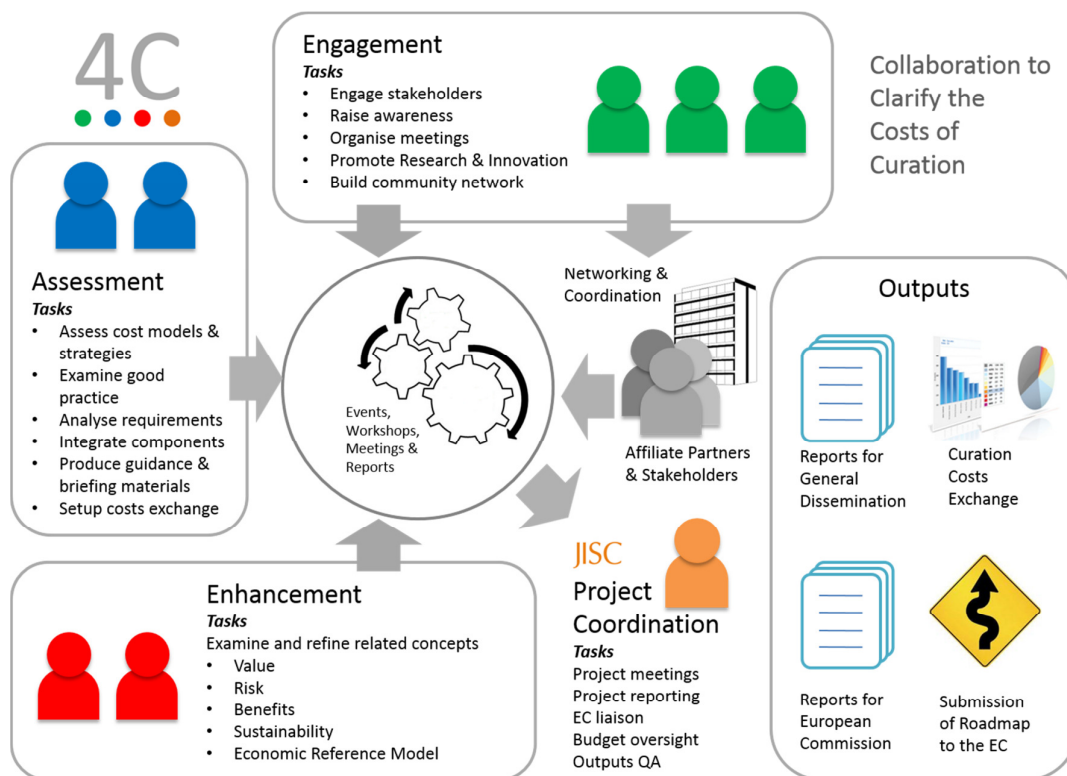
	Digital Curation Centre, University of Edinburgh
	Digital Preservation Coalition
	JISC
	Humanities Advanced Technology and Information Institute, University of Glasgow
	UK Data Archive, University of Essex

Tab. 6 Seznam partnerských institucí projektu 4C
[PARTNERS, 2013]

Aby bylo možné přinést výsledky v poměrně krátké době trvání projektu, je činnost participujících stran rozdělena do pracovních skupin (*Work Packages*), dále jen jako PS:

- **PS Hodnocení** (*Assessment*) má sestavit co možná nejefektivnější modely nákladů pro soukromý i veřejný sektor. Tomu musí předcházet řádné vyhodnocení stávajících modelů, strategií a praktik. Analýza bude mít poměrně široký záběr kritérií: od cen, kvality, hodnot, rizik, užiteků až po udržitelnost. Na základě jejich analýzy a integrace jednotlivých komponentů poskytne rady a návody, jež povedou ke konkrétním výsledkům při odhadování nákladů. Kritickou oblastí této skupiny bude především vhodné propojení výsledků PS Závazek a jejich výsledná použitelnost pro PS Plán. Tuto skupinu řídí Dánský národní archiv (*Danish National Archives*).
- **PS Závazek** (*Engagement*) - jelikož je 4C navržen jako otevřený a sociální projekt, úkolem této PS je navázat spolupráci s co nejširší skupinou zainteresovaných osob a jejich příslušných organizací - paměťové instituce, výzkumné skupiny a nejrůznější organizace průmyslu a obchodu. Dalším úkolem je zvýšit povědomí o problematice, podněcovat diskusi a organizovat setkání, podporovat výzkum a inovace a vytvořit tak rozsáhlou komunitní síť. Tuto skupinu koordinuje Národní knihovna SRN (*Deutsche Nationalbibliothek*).
- **PS Zlepšení** (*Enhancement*) se zaměří na veškeré nepřímé faktory, které ovlivňují náklady dlouhodobé archivace, tj. na hodnotu, rizika, výhody, udržitelnost, ze kterých vznikne Ekonomický referenční model (*Economic Reference Model*). Výsledky svých aktivit poskytne PS Hodnocení a PS Plán. Tuto skupinu řídí Národní knihovna Estonska (*National Library of Estonia*).
- **PS Plán** (*Roadmap*) musí sesbírat výsledky činností předchozích skupin, které poté zpracuje do podoby reportů k obecnému využití, zveřejnit odhady finančních prostředků a přístupy/postupy pro dlouhodobou ochranu u oslovených institucí, resp. Repoziťářů. Následně odevzdat Evropské komisi souhrnný report a odpovídající plán pro stanovování nákladů dlouhodobé ochrany digitálních dokumentů (leden 2015). Tuto skupinu řídí JISC.

- **PS Projektový management (Project Management)** zabezpečí kvalitní řízení projektu a součinnost jednotlivých PS. Tuto roli převzal JISC. Jeho úkolem bude organizování pracovních setkání, vypracovávání reportů, kontakt s Evropskou komisí, kontrola rozpočtu a obhajitelných výstupů.



Obr. 42 Pracovní skupiny (Work Packages) modelu nákladů 4C
(<http://www.4cproject.eu/work-packages>)

Obrázek popisuje 5 klíčových pracovních skupin, které se na projektu 4C podílí. Skupina oblasti "Závazku" (*Engagement*) má za úkol (*Tasks*) navázat spolupráci s nejužnějšími zainteresovanými skupinami (*Engage Stakeholders*), zvýšit povědomí o problematice (*Raise Awareness*), podněcovat diskusi a organizovat setkání (*Organise Meetings*), podporovat výzkum a inovace (*Promote Research & Innovation*) a vytvořit tak rozsáhlou komunitní síť (*Build Community Network*). Skupina oblasti Hodnocení (*Assessment*) se snaží sestavit nejefektivnější metody modelování nákladů pro soukromý i veřejný sektor, kterým musí předcházet řádné vyhodnocení stávajících modelů a strategií (*Assess Cost Models & Strategies*) a praktik (*Examine Good Practice*). Na základě jejich analýzy (*Analyse Requirements*) a integrace jednotlivých komponentů (*Integrate Components*) poskytnout rady a návody (*Produce Guidance & Briefing Materials*), jež povedou k sestavení dotazů o konkrétních přístupech organizací k odhadu nákladů (*Setup Costs Exchange*). Skupina Zlepšení (*Enhancement*) se zaměřuje na veškeré nepřímé faktory, které ovlivňují náklady dlouhodobé archivace, tj. na hodnotu (*Value*), rizika (*Risk*), výhody (*Benefits*), udržitelnost (*Sustainability*), ze které vznikne Ekonomický referenční model (*Economic Reference Model*). Skupina Plánu (*Roadmap*), na obrázku jako výstupy (*Outputs*) sesbírá výsledky činností předchozích skupin, které nabídne v podobě reportů k obecnému využití (*Reports*

For General Dissemination), zveřejní odhady finančních prostředků a přístupy/postupy pro dlouhodobou ochranu u oslovených institucí, resp. repozitářů (*Curation Costs Exchange*), odevzdá Evropské komisi (EK) souhrnný report (*Report For European Commission*) a odpovídající plán (*Submission Of Roadmap To The EC*) pro stanovování nákladů dlouhodobé ochrany digitálních dokumentů. Koordinaci projektu (*Project Management*) zabezpečí jeden z partnerů - JISC. Úkolem bude organizace pracovních setkání (*Project Meetings*), vypracovávání reportů (*Project Reporting*), kontakt s EK (*EC Liaison*), kontrola rozpočtu (*Budget Oversight*) a výstupů (*Output QA*).

Z prvních výstupů projektu 4C¹⁰⁵ je zřejmé, že modely¹⁰⁶ by měly podporovat výstupy, u kterých jsou „na první pohled“ zřejmé užítky a hodnoty dlouhodobé ochrany. Neméně důležitá je i jasná definice nákladů u modelování variabilních položek, jednoduchá práce s nástroji, ideálně při standardních nastaveních a hodnotách, implementace účetnictví, návrhů rozpočtů, zvládnutí nejrůznějších typů a objemů aktiv a případové studie (*Use Cases*) [KEJSER, 2011].

Mnohé z analyzovaných modelů chtějí být obecné a použitelné pro jakoukoliv instituci, která uchovává objekty v digitální podobě. Některé naopak pokrývají pouze specifickou oblast – dle instituce, typu dokumentu apod. Například PrestoPRIME model pokrývá pouze náklady na dlouhodobou ochranu audio-vizuálních materiálů. CET model může být použit jinou institucí uchovávající vědecká data a ISIS model je specificky zaměřen pouze na ISIS STFC formáty, software atd. Je jasné, že různé repozitáře a archivy se zabývají jen určitou podčástí celkových nákladů, v závislosti na mandátu, cílech a povaze materiálů, které uchovávají. LIFE model spoléhá na metodologii životního cyklu a rozděluje jej na aktivity (*Activity-based Models*), které jsou součástí „procesního modelu“. Všechny modely se liší záběrem, stupněm vývoje, oblastí využitelnosti (všeobecný nebo doménově-specifický model) a uživatelskou zodpovědností pro dodávání dat. **Hlavními nedostatky modelů je však poměrně nízká uživatelská přívětivost, nedostatek konsenzu při definování, kvalifikování a strukturování dat nákladů a nedostatek důkazů důležitosti investování do dlouhodobé ochrany dat** [KAUR, 2013; KEJSER, 2014].

¹⁰⁵ Podrobná zpráva je dostupná na: <http://www.4cproject.eu/community-resources/outputs-and-deliverables/d3-1-evaluation-of-cost-models-and-needs-gaps-analysis-ms12-draft>.

¹⁰⁶ Hodnocené byly tyto ekonomické modely: *Test Bed Cost Model For Digital Preservation* (T-CMDP), *NASA Cost Estimating Tool* (NASA-CET), *LIFE³ Costing Model* (LIFE3), *Keeping Research Data Safe* (KRDS), *Cost Model For Digital Archiving* (CMDA), *Cost Model For Digital Preservation* (CMDP), *DP4LIB Cost Model* (DP4Lib), *PrestoPrime Cost Model For Digital Storage* (PP-CMDS), *Total Cost Of Preservation* (CDL-TCP) a *Economic Model For Long-Term Storage* (EMLTS).

6. Požadavky uživatelů na digitální knihovny, repozitáře a archivy

S nástupem osobních počítačů si mnozí uvědomili, jak důležité je věnovat pozornost interakci člověk-počítač (*Human-Computer Interaction* - HCI), jelikož nedostatečně navržená rozhraní mezi člověkem a počítačem mohou vést k nepředvídatelným problémům. Změnil se tak přístup k návrhu a designu rozhraní informačních systémů.

HCI představuje spleť vzájemně propojených oborů - počítačové grafiky, kognitivní psychologie (včetně lidských faktorů), ergonomiky, inženýrských oborů a systémů, jenž jsou součástí počítačové vědy, operační systémy nevyjímaje. HCI je ale orientována i na design, tvorbu a testování přístupnosti a použitelnosti informačních systémů a jejich rozhraní tak, aby byly co nejjednodušší a nejintuitivnější pro konkrétně definovanou skupinu uživatelů. Dodržování nejdůležitějších zásad použitelnosti systému či rozhraní zvyšuje srozumitelnost a přehlednost. Interakce s rozhraním by tedy měla probíhat bez nutnosti složitého přemýšlení. Rozhraní by mělo být uživatelsky co nejpřívětivější. **Současný trend je zcela zaměřen na uživatele a jeho potřeby. Uživatel je totiž centrálním a určujícím činitelem celého informačního systému,** bez respektování potřeb uživatelů funguje jakýkoliv systém (tedy i digitální repozitář, knihovna či archiv) víceméně naprázdno.

Uživatelsky orientovaný design (UCD - *User-centred design*) je design zaměřen na uživatele. Představuje moderní, rychle se rozvíjející filosofii designu, která má za to, že dobrý design musí sloužit především zájmu uživatelů, přičemž uživatelské požadavky jsou od začátku (životního cyklu tvorby uživatelského rozhraní) zaznamenávány a následně doladovány prostřednictvím různorodých metod (např. etnografických studií, testování prototypů apod.). Aktivní participaci koncových uživatelů na designu rozhraní prostřednictvím různých workshopů a sezení představuje tzv. participační design. Nabízí vhodné společné komunikační prostředí pro technické i netechnické účastníky (vývojáře, designéry i uživatele). To vše s cílem nejvyšší možné použitelnosti rozhraní či systémů jako celku.

Mezinárodní norma ISO 9241 (v části 11) [ISO, 2013] definuje použitelnost jako využívání produktu uživateli za účelem dosažení specifických cílů, a to efektivně, s vysokým stupněm výkonnosti a míry uspokojení. Půžitelnost lze v obecnějším chápání brát jako takový přístup k vývoji výrobku, který předpokládá přímou zpětnou vazbu uživatele během výrobního cyklu, s cílem snížení nákladů a naplnění potřeb uživatelů [MAKULOVÁ, 2007]. Výsledkem může být i samotný informační systém, potažmo digitální knihovna, repozitář nebo archiv. V tomto kontextu se objevuje i termín systémová použitelnost (*System Usability*), definující co nejlepší využití nabízených funkcionalit systému uživateli [KLING, 1994].

Pro vytvoření digitální knihovny splňující kritéria použitelnosti, je nutné se zaměřit zejména na tyto základní složky, vycházející mimo jiné i z informační architektury: obsah, funkčnost/funkce (navigace, vyhledávání) a rozhraní. Kling [1994] však zdůrazňuje nejen použitelnost rozhraní (*Interface Usability*), ale i tzv. organizační použitelnost (*Organizational Usability*). Organizační použitelnost určuje možnosti efektivního začlenění systémů do celkové struktury organizace. To vše ale není jednoduché podchytit, jelikož lidské, resp. uživatelské chování je složité a technologie pro digitální knihovny prochází neustálým a rychlým vývojem.

Mezi nejčastější metody testování použitelnosti patří např. uživatelské testování, „eye-tracking“, testování prototypů (na papíře, nebo ve formě elektronické prezentace). Dále je nutné se zaměřit i na způsob vytváření obsahu, navigaci, vyhledávání včetně grafiky a animace, grafického designu a prvků uživatelského rozhraní. Vlastnosti určující kvalitu rozhraní [USER, 2014]:

- **interaktivita** - rozhraní by mělo komunikovat s uživatelem, informovat jej o tom, co se děje
- **průhlednost** - ovládání by mělo být jednoduché a reakce systému předvídatelné
- **přehlednost** - uživatel se musí dobře orientovat nejen na jedné obrazovce, ale v celém systému
- **empatie** - uživatel by měl mít pocit, že rozhraní je vytvořeno přesně pro něj
- **standardizace** - zachování klasických a navykých pracovních postupů uživatele i s ohledem na zvyklosti a standardy operačního systému, pro který je program navržen
- **srozumitelnost** - pojmy, které rozhraní používá, by měly být obecně známy, popřípadě by měly být známy skupině uživatelů, pro které je program určen
- **návodnost** - rozhraní by mělo provázet uživatele celým systémem tak, aby se uživatel navigaci v systému nemusel „učit“
- **ochrana** - před nebezpečnými operacemi, které by mohly připravit uživatele o data
- **pomoc** - vybavení přehlednými helpy a pomůckami

Pokud totiž většinu výše uvedených kritérií systémy/rozhraní mají splňovat tyto čtyři charakteristické znaky [USABILITY, 2014]: **snadnou naučitelnost a zapamatovatelnost, vysokou výkonnost uživatelů při provádění a dokončování stanovených úkolů, nízkou chybovost a spokojenost**. Příkladem zpětné vazby od uživatelů a jedním ze způsobů měření spokojenosti systémů/rozhraní může být dotazník (viz příklad str. 141 – odezva na Digitální univerzitní repozitář UK).

6.1 Zpětná vazba – dotazníkové šetření

Pro získání rychlé zpětné vazby k obecnému uživatelsky přívětivému rozhraní a konkrétním připomínkám k Digitálnímu univerzitnímu repozitáři se celkově vyjádřilo pomocí dotazníkového šetření 21 studentů magisterského studia na Univerzitě Karlově v Praze, v Ústavu informačních studií a knihovnictví. Dotazník¹⁰⁷ byl zpracován v březnu 2014. Obsahoval celkem 14 otázek; od základních identifikačních údajů (otázky č. 1-5), následoval prostor pro vlastní názor ohledně uživatelsky přívětivých informačních systémů (otázky č. 6-11), až po zpětnou vazbu k univerzitnímu repozitáři (otázky č. 12-14).

Před samotným dotazníkovým šetřením byly stanoveny tři základní hypotézy, které měl dotazník potvrdit nebo případně vyvrátit:

1. Uživatelé od přívětivého systému či rozhraní očekávají především přehlednost **(potvrzeno)**.
2. Uživatelům Digitálního univerzitního repozitáře UK v Praze (<http://repozitar.cuni.cz/>) činí významné problémy vyhledávání v repozitáři a vyhledané výsledky jsou nerelevantní **(nepotvrzeno)**.
3. Grafické rozhraní Digitálního univerzitního repozitáře je zastaralé a uživatelé jej vnímají záporně **(potvrzeno)**.

Prvním předpoklad tedy byl, že uživatelé od přívětivého systému či rozhraní očekávají především přehlednost (Hypotéza č. 1). Další dvě hypotézy se přímo dotýkaly Digitálního univerzitního repozitáře a vycházely ze získané zpětné vazby několika fakultních knihoven. Vyplývá z nich, že uživatelům činí významné problémy vyhledávání v daném repozitáři a vyhledané výsledky nejsou relevantní (Hypotéza č. 2). Naproti tomu se potvrdilo, že uživatelské rozhraní je z grafického hlediska zastaralé a nutně potřebuje obměnu (Hypotéza č. 3).

Většina respondentů – necelých 86% (18 studentů) patří do věkové skupiny 20-29 let, 2 studenti jsou ve věku mezi 30-39 let. Pouze jeden dotazník informací o věkové skupině neobsahoval. Průzkumu se zúčastnilo 2x více žen (téměř 67%) než mužů (14 žen a 7 mužů), což je

¹⁰⁷ Viz nevyplněný dotazník v Příloze č. 10 – Dotazník pro studenty výsledky a seznam výsledků v tabulce Přílohy č. 11 – Tabulka souhrnných údajů dotazníku pro studenty.

dáno i samotným oborem Informační studia a knihovnictví, který tradičně studovalo a studuje víc žen než mužů.

Naprostá většina odpovědí na otázku č. 5 (téměř 62% - 13 dotazovaných) a otázku č. 8 (více než 90% - 19 dotazovaných) **potvrdily předpoklad stanovený hypotézou č. 1**, tedy že uživatelské rozhraní by mělo být především přehledné. Pouze jedna odpověď se výrazně odlišovala. Dotazovaný zastával názor, že *„data by měla být dostupná přes Google, pak není co řešit“*. To koneckonců vlastně potvrzují i odpovědi na otázku č. 10 - jestli uživatelé hledají informace přímo na stránkách digitálních repozitářů a knihoven nebo využívají nejprve možnosti vyhledávacích služeb typu Google. Pouze 3 uživatelé uvedli, že hledají přímo na stránkách knihoven a archivů. Tedy téměř 85% (18 respondentů) nejdříve hledá informace ve vyhledávacích a po neúspěšném hledání navštíví stránky knihoven. Zajímavostí je, že kritérium empatie u dobrého uživatelského rozhraní (tedy uživatel má pocit, že rozhraní bylo vytvořeno pro něj) uvedly pouze ženy. U otázky č. 8 se pouze 6 studentů shodlo na digitální knihovně, která podle nich odpovídá principům uživatelsky přívětivého rozhraní. Jednalo se 2x o digitální knihovnu Europeana (<http://europeana.eu/>), 2x World Digital Library (<http://www.wdl.org/>) a Manuscriptorium (<http://www.manuscriptorium.com/>). Dodržování skupiny standardů ČSN EN ISO 9241 při návrzích uživatelských rozhraní považuje za důležité víc jak 85% (18 respondentů).

U otázky č. 12 (Hypotézy č. 2) pouze 7 dotazovaných nebylo spokojeno s vyhledáváním a výsledky vyhledávání, což činí něco málo přes 33%. **Hypotéza č. 2 se tedy nepotvrdila.**

Jelikož otázky 13 a 14 nabízely volnou odpověď, nebylo jednoduché stanovit jednoznačnou kladnou nebo zápornou odpověď. Nejčastěji byl vytýkán zastaralý, málo interaktivní design, malé písmo, ikonka „Go“ pro zahájení vyhledávání či neexistence návodů. Asi nejlépe většinové názory na repozitář shrnula studentka slovy: *„rozhraní dle mého názoru nepatří rozhodně mezi ty nejhorší, které lze objevit – setkala jsem se s mnohem horšími, ale jistě je tu možnost vylepšení. Část uživatelů jistě trochu odradí nepříliš moderní grafika (já osobně se také moc nemůžu přenést přes tu světle modrou – té je až dost ve Windows)“* **Hypotéza č. 3 tedy byla více méně potvrzena**, kdy „pouze“ 9-10 uživatelů (přibližně 43- 48%) celkově hodnotilo rozhraní (otázka č. 14) kladně a přívětivě.

Ze získaných výsledků je tedy nejdůležitější zaměřit se na inovaci grafického rozhraní, zmodernizovat jej, odstranit přebytečné informace z úvodní stránky, zvětšit písmo, přidat návody a přejmenovat vyhledávací tlačítko.

7. Současný stav problematiky v ČR

V České republice funguje několik digitálních knihoven, repozitářů a archivů většího rozsahu, postavených zejména na „open-source“ technologiích a vlastních (zejména upravených „open-source“) řešeních.

Aby byl poskytnut pokud možno objektivní přehled (napříč spektrem knihoven) o typech systémů, strategiích, administrativě a managementu digitálních knihoven a repozitářů, byl sestaven dotazník a e-mailem byly osloveny tyto instituce: Národní filmový archiv (NFA), Národní technická knihovna (NTK), Národní knihovna ČR (NK ČR), Národní archiv (NA), Městská knihovna v Praze (MKP), Moravská zemská knihovna (MZK), Národní knihovna ČR (NK ČR) a Knihovna Akademie věd ČR (KNAV). Poslední zmiňovaná knihovna upřednostnila formu interview. Zaznamenané odpovědi autorky byly poté e-mailem autorizovány. I přes opakovanou prosbu o vyplnění dotazníků, odpovědi z MZK a NK ČR nebyly zaslány.

Dotazník měl potvrdit či vyvrátit **tyto stanovené hypotézy**¹⁰⁸:

1. Knihovny nerozlišují mezi termíny digitální knihovna a digitální repozitář **(nepotvrzeno)**
2. Knihovny a archivy své digitální knihovny, repozitáře a archivy necertifikují **(potvrzeno)**
3. Standardy pro dlouhodobou ochranu nejsou téměř vůbec dodržovány **(potvrzeno)**.

Cílem dotazníkového šetření v institucích¹⁰⁹ bylo v devíti okruzích otázek zjistit, jakým způsobem vnímají instituce rozdíly mezi digitálními knihovnami a repozitáři (pokud vůbec), jaká byla motivace pro jejich digitální repozitář či knihovnu (Otázka č.1), jakým způsobem byl výsledný systém vybírán a proč bylo zvoleno komerční či „open source“ řešení (Otázka č.2). Otázka č. 3 směřovala k finančním záležitostem – nakolik byl výběr ovlivněn finančními prostředky z grantů, jak tyto prostředky byly investovány (jestli pouze hardware nebo software apod.).

U KNAV a MKP bylo hlavní motivací pro digitální knihovnu zpřístupňovat uživatelům digitalizované materiály vlastního fondu (zejména monografie, noviny a časopisy). Digitalizace byly financovány z grantů (VISK7¹¹⁰ – KNAV a Norské fondy¹¹¹ – MKP). Vznik digitálního repozitáře

¹⁰⁸ Hypotézy vychází ze studované literatury, komunikace s odborníky doma i v zahraničí a z účasti na nejrůznějších knihovnických akcích (semináře, workshopy, konference) – viz 1.1 Zkušenostní základna pro vypracování disertační práce – přehled, Seznam použité literatury a Seznam doporučené literatury.

¹⁰⁹ Viz nevyplněný dotazník v Příloze č. 8 - Dotazník pro instituce a seznam výsledků v tabulce Přílohy č. 9 – Tabulka souhrnných údajů dotazníku pro instituce.

¹¹⁰ Podmínky Programu Veřejné informační služby knihoven 7 lze najít na v dokumentu na: <<http://goo.gl/kzM52d>>.

u KNAV byl motivován zejména potřebou evidování a hodnocení vědeckého publikování a zavedením zlaté cesty „open access“¹¹² (*Gold Open Access*) politiky. Zatímco v digitální knihovně uchovávají zdigitalizované materiály, v digitálním repozitáři ASEP „born-digital“ dokumenty¹¹³. Tím se nepotvrdila hypotéza č. 1, že knihovny nerozlišují mezi digitální knihovnou a digitálním repozitářem.

U NA, NFA a NTK je hlavním podnětem k vytvoření repozitářů samotné poslání institucí (národního charakteru) a s tím úzce spojená potřeba řešit problematiku dlouhodobého uchovávání digitálních dat. Projekt Národního digitálního archivu (NDA) byl budován již v letech 2007-2008. Po vleklých problémech s výběrovými řízeními se instituce nakonec rozhodla pro implementaci systému dlouhodobé ochrany formou „open source“ řešení Archivematica¹¹⁴. Jak poznamenal NFA: **„Archivematicu používá nebo o ní uvažuje řada dalších paměťových institucí v ČR a spolupráce na implementaci by mohla být užitečná. Open source řešení považujeme principiálně vhodné pro instituce našeho typu (nižší riziko ztráty kontinuity produktu, než u komerčního produktu).“**

Náklady spojené s pořízením, resp. úpravou „open source“ repozitářového systému byly v případě NA a NTK pokryty z grantových prostředků (u NA ze strukturálních fondů EU v rámci Integrovaného operačního programu). NTK před výběrem konkrétního řešení vypracovala podrobnou analýzu¹¹⁵, ve které srovnávala systémy s předem zadanými požadavky. Je možné tedy usoudit, že instituce nemají odpovídající personální zabezpečení, kterými by pokryly nedostatečné finanční prostředky pro pořízení a následnou implementaci systémů či případný vlastní vývoj. I když NFA nemá doposud vybudovaný odpovídající repozitář, uvědomují si, že se jedná o nevyhnutnost. Jak uvedl vedoucí digitální laboratoře Jan Zahradníček: *„Přijímáme do sbírek nativně digitální audiovizuální data a chápeme, že je třeba je prezervovat nejen na bitové, ale i logické úrovni. Jsme však na obou úrovních zcela na začátku: bitovou ochranu zajišťujeme jedním online diskovým polem zabezpečeným technologií DDP (obdoba RAID6) a dvěma redundantními řadami LTO-6 pásek (LTFs). Doplňkový a spíše záložní systém (protože neobsahuje plnou kvalitu děl) je postaven na optických médiích (nyní dvouvrstvé BD). Výhradní prezervační*

¹¹¹ Podrobnější informace viz portál Ministerstva financí ČR: <<http://www.eeagrants.cz/cs/programy>>.

¹¹² Zlatá cesta publikování v režimu „open access“ představuje publikování v otevřených časopisech, kde přístup poskytují vydavatelé.

¹¹³ KNAV chápe digitální knihovnu jako digitální protějšek knihovny, která zpřístupňuje zdigitalizované dokumenty a „digital-born“ dokumenty fondu knihovny. Naproti tomu digitální repozitář může uchovávat i digitální materiály mimo knihovní fond (ve smyslu institucionálního repozitáře) i bez prvku dlouhodobé ochrany (dle modelu OAIS – viz část 3.3 Open Archival Information System Reference Model – OAIS).

¹¹⁴ Podrobnější informace lze najít v části 4.5 Open source systémy a nástroje pro dlouhodobou ochranu nebo přímo na stránkách projektu Archivematica: <https://www.archivematica.org/wiki/Main_Page>.

¹¹⁵ Analýza je dostupná na: <http://invenio.nusl.cz/record/111525/files/idr-270_1.pdf>.

formát (který zároveň může sloužit i k prezentaci) je nekryptované DCP (Digital Cinema Package). Další formáty jsou jen pomocné, slouží k prezentaci. Na logické úrovni nyní vedeme základní evidenci o sbírkách v excelové tabulce, do budoucna uvažujeme o implementaci zřejmě Archivematicy (s extrakcí technických metadat zřejmě pomocí MediaInfo), příp. P4.”

Na druhé straně NTK má digitální repozitáře dva. Ten první, Národní úložiště šedé literatury, vznikl původně jako pilotní projekt sběru, dlouhodobé archivace a zpřístupnění šedé literatury v České republice¹¹⁶. V současné době plní funkci Národního úložiště pro šedou literaturu. Druhý repozitář představuje institucionální repozitář NTK. Zde se uchovávají výsledky činnosti zaměstnanců NTK a materiály z akcí konaných NTK.

Další okruh (Otázka č. 4) zjišťoval jaký typ dat - audio, video, text, obraz – repozitář, či archiv knihovna má, jestli jde převážně o digitalizované či „digital-born“ dokumenty a jakým způsobem je zdefinován samotný objekt v repozitáři (co objekt to jedna stránka?, METS objekty? apod.) a metody vytváření metadat.

Oslovené knihovny a archivy zpřístupňují zdigitalizované i „digital-born“ dokumenty převážně obrazového a textového charakteru (KNAV, MKP, NA). NFA primárně audiovizuální data (s obrazovými a textovými doplňky v podobě fotografií z natáčení, scénáři atd.). NA má v plánu zpracovávat a uchovávat i tabulky a databáze. Jak dále poznamenává: „*opomenout nelze ani vektorovou grafiku a GIS*¹¹⁷.“ **Žádná oslovená instituce ale nerespektuje pravidla typu METS. Pro objekty vesměs platí, že 1 objekt = 1 bibliografický záznam.** Zajímavostí je, že zatímco některé oslovené instituce nejdříve digitalizovaly a poté uplatňovaly zkušenosti při zpracování a archivaci „digital-born“ dokumentů (KNAV, MKP), jiné své zkušenosti získávaly nejdříve z nativně digitálních dokumentů a posléze na digitalizovaných materiálech. Původní překotné digitalizační období se jich netýkalo. Jak napsali z NA: „*[Máme*¹¹⁸*] Born digital dokumenty. Zkušenosti s jejich uložením však budou využity rovněž při ukládání digitálních reprodukcí.*“ Záznamy jsou pořizovány jak ručně, tak přebírány z katalogů knihoven.

Otázka č. 5 zkoumala nakolik architektury repozitářů splňují mezinárodní ISO standard 14721:2003 Space data and information transfer systems -- Open archival information system -- Reference model, příp. jeho revidované verzi ISO 14721:2012. **Až na NTK, žádné z úložišť neodpovídá standardu ISO 14721:2003 ani jeho revidované verzi, i když by většina z nich takovou normu ráda splnila.** NTK pro potvrzení souladu s tímto ISO standardem vypracovala certifikovanou (a odborníky oponovanou) metodiku Metodika zpracování, dlouhodobého

¹¹⁶ Národní úložiště šedé literatury vzniklo za podpory Ministerstva kultury v rámci projektu Digitální knihovna pro šedou literaturu.

¹¹⁷ Podrobnější informace na: <http://cs.wikipedia.org/wiki/Geografický_informační_AD_systém>.

¹¹⁸ Pro lepší pochopení kontextu je slovo [Máme] doplněno autorkou.

uchování a zpřístupnění šedé literatury v ČR na příkladu Národního úložiště šedé literatury¹¹⁹. KNAV uvádí, že: „*k dispozici je digitální úložiště, nikoli ale ve smyslu knihovnickém. Z toho pohledu tedy systém neodpovídá OAIS modelu, jelikož chybí LTP systém.*“ MKP repozitář: „*neodpovídá standardu ISO 14721:2003 ani jeho revidované verzi ISO 14721:2012 a zatím pro naše potřeby není nutné zavádět změny, které by s sebou nesly akceptaci všech podmínek.*“

V dotazníku byla věnována pozornost (Otázka č. 6) i způsobům dlouhodobého uchování dat v repozitáři, postupům při fyzické ochraně dat, včetně intervalu a testování obnovitelnosti dat ze záloh. Dále jestli instituce provádí i logickou ochranu dat (např. dodržování standardů, výběr neproprietárních formátů, stanovení signifikantních vlastností objektů pro případné budoucí migrace apod.).

Zálohování probíhá ve všech institucích na bitové úrovni. Částečnou logickou ochranu v NTK supluje pravidelná měsíční kontrola postupů, stanovených v dokumentaci instituce. KNAV pro zálohování nemá písemné směrnice, spoléhá na znalosti IT personálu, čímž, jak si sami uvědomují, se vystavují riziku ztráty „know-how“ odchodem klíčových zaměstnanců IT oddělení. Toto riziko chtějí v nejbližší době podchytit řádným zdokumentováním postupů a procesů. MKP provádí pravidelné zkoušky obnovy dat: „*Dlouhodobé uchování dat je palčivá otázka, v tuto chvíli se snažíme – máme dvě záložní kopie na jiném diskovém poli – zálohy probíhají každou noc a pravidelně testujeme jejich obnovitelnost – minimálně jednou za 3 měsíce.*“ Zálohování je většinou na dvou geograficky oddělených místech, frekvence je většinou denního intervalu. Z dotazníku bylo zjištěno, že NTK kromě denních záloh provádí i tzv. měsíční archivní zálohy: „*máme systém záloh jak na serverech tak na páskách a uchováváme je geograficky odděleně. Přístupy a manipulace je dána jen pověřeným osobám. Systém záloh je pracován od denních inkrement[al]ních až po měsíční archivní.*“ Jelikož NFA nemá repozitář s prvkem dlouhodobé ochrany, spoléhají prozatím na kontrolní součty MD5, přičemž plánují pravidelné kontroly integrity.

Další část byla věnována otázkám kolem správy a administrace systému a uživatelské přívětivosti rozhraní repozitářů a knihoven (Otázka č.6, 7). Důležité bylo vědět, jestli organizace „outsourcují“ některou část repozitáře nebo knihovny (např. zálohování, úpravy systému apod.). **I když je v zahraničí „outsourcování“ poměrně běžnou záležitostí, u nás se instituce v co největší možné míře spoléhají na vlastní zdroje.** Pouze NA sdělil, že „outsourcován“ bude „*vývoj portálu pro přístup k NDA, tzv. Archivního portálu.*“

K otázce důležitosti přívětivých a přístupných rozhraní pro uživatele se vyjádřily pouze NA a NFA. I když NA soulad s normou ISO 9241 v tomto okamžiku neřeší, požadují, aby odpovídal

¹¹⁹ Metodika je dostupná na webu NTK: <<http://nusl.techlib.cz/index.php/Metodika>>.

pravidlům přístupného webu a po technické stránce HTML5. Přestože NFA nemá repozitář, jeho zástupce vtipně poznamenal: „*na uživatele určitě hledíme, tu naši tabulku se snažíme dělat přehlednou*“. Jestli a jakým způsobem knihovny a archivy získávají zpětnou vazbu od uživatelů se dotazníkem nepodařilo zjistit. Aby však systém nefungoval „naprázdno“, komunikace s uživateli je nevyhnutná.

Poslední část dotazníku (Otázky č. 8 a 9) se zaměřila na audity, certifikace podle Evropského rámce pro audit a certifikaci repozitářů (DSA, ISO 16363 or DIN 31644) a politiky finančního plánování, managementu rizik, dlouhodobé ochrany dat, sběru dat, přístupných formátů apod.

Instituce jsou přesvědčené, že certifikaci nepotřebují nebo naopak potřebují, ale je příliš nákladná. Potvrdila se tedy druhá hypotéza, která předpokládala, že knihovny a archivy své digitální knihovny, repozitáře a archivy necertifikují. Certifikaci má předcházet (externí) audit, ve kterém vyhodnotí současný stav, procesy a postupy v rámci jednotlivých repozitářů a archivů.

Certifikace podle ISO 16363 je skutečně nákladnou záležitostí, nicméně v současné době již existují i bezplatné certifikace typu Data Seal of Approval¹²⁰. Jelikož NDA v rámci NA byl financován z evropských fondů, projekt předpokládá ve svém závěru audit a certifikaci. Jak napsal NA: „*výstupem auditu budou zprávy o souladu informačního systému s ISO 16363 a OAIS. Výstupem certifikace bude certifikát o splnění požadavků kladených [na] informační systém veřejné správy a o splnění požadavků kladených na IT bezpečnost.*“ MKP a NTK alespoň pravidelně provádí interní audit pomocí nástroje DRAMBORA. I tím dávají závazek svým uživatelům, že znají rizika a ohrožení v repozitářích a v rámci knihoven, čímž vzbuzují vyšší důvěryhodnost. Audit má za cíl objektivně posoudit pracovní postupy, procesy, technické a personální zabezpečení, finanční možnosti apod.

Pracovní postupy, resp. politiky pokrývající nejrůznější oblasti mají kodifikované pouze NTK a MKP. Nedávno byl v MKP zveřejněn dokument „*Základní pravidla archivace elektronických dat v MKP, který definuje politiku archivace v MKP. V rámci dokumentu byly popsány dokumenty a formáty, které se v MKP v rámci e-publikování uchovávají, možná rizika. V nejbližší době bude ještě doplněn o vlastnosti formátů, které MKP přijímá od uživatelů a jiných producentů dat a které by měla uchovávat.*“

Jak tento dotazník potvrdil, standardy pro dlouhodobou ochranu nejsou téměř vůbec dodržovány (Hypotéza č. 3), což je velmi zneklidňující. Obdobně je to se standardy i v oblasti

¹²⁰ Podrobnější informace v části 5.1 Metriky, certifikační a auditní nástroje.

návrhu architektury digitálních knihoven a repozitářů, sestavování digitálních objektů, jako i při návrzích uživatelského rozhraní a certifikace procesů uvnitř knihoven a archivů. Rozlišují pojmy digitální knihovna a digitální repozitář, kdy druhý zmíněný je buď přesný protějšek kamenné knihovny či archivu nebo má na rozdíl od digitální knihovny nastavené procesy pro dlouhodobou ochranu a jim odpovídající nástroje. Oslovené instituce si tedy uvědomují naléhavost řešení dlouhodobé ochrany a potřebu auditů – dvě knihovny pravidelně provádí interní audit DRAMBORA. Pro knihovny a archivy, jež ještě nemají implementovaný repozitář s prvky dlouhodobé ochrany byly autorkou sestaveny následující „Best Practices“, které prokrývají problematické body z dotazníku.

8. Best Practices

„Best Practices“ představují metodu nebo techniku (nejčastěji v podobě příručky či návodu), jež vychází ze sběru zkušeností od lidí, kteří se danou problematikou systematicky zabývají. Jejich hlavním cílem je rychlé zorientování se v dané oblasti a poskytnutí co nejsnadněji aplikovatelných poznatků do praxe¹²¹. Pro tyto typy dokumentů je typický stručný charakter s maximální obsahovou výtažností se zaměřením na praxi. Vychází z dostupných informačních pramenů, metodik pro důvěryhodné repozitáře (zejména PLATTER a DRAMBORA) a téměř 8leté zkušenosti autorky z administrace Digitálního univerzitního repozitáře Univerzity Karlovy v Praze. Cílem této metodiky, resp. „Best Practices“ příručky je poskytnout zejména managementu repozitářů instrukce při vytváření digitálních knihoven, repozitářů a archivů. Dokument je rozčleněn do 11 základních okruhů, na které je nutné se zaměřit. Nejsou však rozepsány do podrobnů. Hlavním důvodem je možnost rychlé orientace v problematice a v specifických podmínkách v prostředí jednotlivých digitálních knihoven, repozitářů a archivů, které není snadné ani žádoucí příliš generalizovat. Tato stručná příručka může posloužit pro sestavení vhodných plánů, procedur a postupů v rámci organizace repozitáře. Že nejde o snadnou věc dokazuje i multidisciplinární záběr problematiky; od technických poznatků přes krizový management, management rizik až po audit a certifikaci. I proto autorka sestavila přehled nejdůležitějších standardů, metodik a dokumentů pro efektivní správu repozitářů, knihoven archivů, který je součástí přílohy č. 7.

Konkrétní (skutečné či fiktivní) příklady pro jednotlivé podčásti „Best Practices“ jsou uvedeny odlišným typem písma Times New Roman, pro lepší rozlišení v textu orámovány. Doplňující informace a informační zdroje jsou v poznámkách pod čarou¹²².

¹²¹ Dobrým příkladem takového dokumentu jsou např. „Best practice – Pravidla tvorby přístupného webu“ z „dílny“ již neexistujícího Ministerstva informatiky, jež nesloužily pouze pro orientaci v problematice, ale i jako normativní dokument (pro účely novely Zákona č. 365/2000 Sb. o informačních systémech veřejné správy, provedenou zákonem č. 81/2006 Sb.). Dokument je dostupný například z: http://blindfriendly.cz/download/doc/MI_best_practice.pdf.

¹²² Podrobnější informace jsou součástí kapitol 1 až 5 a lze je též dohledat v informačních pramenech uvedených v Seznamu použitých a doporučené literatury této disertační práce.

8.1 Současný stav, definice repozitáře

Prvním krokem, který podmiňuje dobře fungující repozitáře je vyhodnocení současného stavu, identifikace cílové skupiny a zabezpečení mandátu. Stanovení cílové skupiny repozitáře¹²³ vyžaduje analýzu a definování způsobů přístupů budoucích uživatelů. Pro „federované“¹²⁴ (*Federated Digital Repositories*) či spolupracující repozitáře je nutné sdílené porozumění cílové skupiny, kterým mají repozitáře sloužit.

Organizace si musí být vědoma, jaký repozitář má nebo do budoucna chce mít a co je hlavní motivace k jeho vytvoření či provozu. Vše ostatní se odvíjí právě od tohoto prvního kroku (od výběru vhodné technologie až po formu zpřístupňování dat). Proto je důležité definování repozitáře pomocí mandátu (*Mandate*) a poslání (*Mission*). Jde o identifikaci organizačního kontextu, hlavních cílů¹²⁵ a poslání repozitáře, jež nepřímo určují i omezení činností (*Constraints*). Mandát může mít podobu zakládací listiny nebo obdobného oficiálního dokumentu od zřizovatele¹²⁶, resp. organizace odpovídající za digitální repozitář. Přičemž může platit, že **digitální repozitář = celá organizace nebo repozitář je pouhou částí organizace jako celku** (technická realizace – úložiště)¹²⁷. Nejčastěji je možné se s požadavkem na zdokumentované cíle a plány setkat u repozitářů, knihoven či archivů, jež usilují o některý z typů certifikace. Tento dokument není zavazující či jediný platný. Představuje „živý“ dokument, který se upravuje s měnícími technologiemi a požadavky knihoven či uživatelské komunity.

Je možné uvést následující fiktivní příklad mandátu repozitáře UK:

¹²³ Cílová skupina je často definována i samotným mandátem repozitáře (např. v případě firemních nebo institucionálních repozitářů).

¹²⁴ Podrobněji o „federovaných“ digitálních repozitářích například zde: http://eprints.rclis.org/7629/1/ATINERconf_paper_Chumbe.pdf.

¹²⁵ Veškeré stanovené cíle by měly odpovídat tzv. SMART cílům. Poprvé se s tímto pojmem setkáváme v roce 1981. SMART představuje anglický akronym slov konkrétní (*Specific*), měřitelný (*Measurable*), dosažitelný (*Attainable*), odpovídající či realistický (*Relevant/Realistic*), termínovaný (*Time-bound*) a vyhodnotitelný (*Evaluable*). Českým ekvivalentem bývá KARAT (konkrétní, ambiciózní, reálný, akceptovatelný a termínovaný) [SMART, 2014].

¹²⁶ Jeden z pěkných příkladů mandátu pro institucionální repozitář lze najít zde: <http://roarmap.eprints.org/825/>.

¹²⁷ Jako příklad možno uvést velké národní knihovny, jež kromě tradičního knihovního fondu nabízí i digitální knihovnu/repozitář/archiv. Zvláštním typem repozitářů jako celků jsou pak konsorciální digitální knihovny a repozitáře.

Poslání a kontext Digitálního univerzitního repozitáře UK

Univerzita Karlova v Praze (dále jen UK) začala budovat vlastní Digitální repozitář v roce 2006, a to v rámci Transformačních a rozvojových projektů Ministerstva školství, mládeže a tělovýchovy České republiky. Cílem bylo zlepšení dostupnosti digitálních dokumentů (jako jsou např. závěrečné práce, studijní materiály, periodika, vědecké články a příspěvky do sborníků, digitalizované historické dokumenty apod.) vytvořených na UK v souladu s požadavky na jejich dlouhodobou ochranu a archivaci.

Hlavním posláním Digitálního univerzitního repozitáře UK v Praze je tedy shromažďovat, uchovávat a poskytovat přístup k vybraným materiálům, které UK vlastní nebo byly vytvořené studenty, akademickými pracovníky a vědci s afiliací k UK v Praze.

Klíčovými hodnotami jsou spolehlivost, transparentnost a bezpečnost.

Spolehlivost

- Trvalé uchovávání dat a digitálních dokumentů dle stanovených signifikantních vlastností
- Zavázání se k osvědčeným postupům pro důvěryhodné digitální repozitáře

Transparentnost

- Otevřenost v oblasti investičních strategií a nákladů

Bezpečnost

- Zabezpečení obsahu a povolení přístupu pouze vybraným lidem
(dle stanovených politik přístupů)
- Zabezpečených dat a připojení.

Konec příkladu.

8.2 Hardwarové a softwarové zabezpečení, datový plán

Na základě definovaných cílů, plánů akvizice a předpokládaného datového modelu je možné začít sestavovat hardwarovou infrastrukturu a softwarové řešení. Datový model nejčastěji vychází z výběru konkrétního hardware a software pro repozitář, knihovnu či archiv. Jen zřídkakdy je tento model sestaven před vlastním zabezpečením hardwaru a softwaru. Volnější ruku v návrhu mají ty repozitáře, které jsou navrženy a naprogramovány přímo pro potřeby organizace repozitáře. Zde však vzniká riziko nedostatečného podchycení všech důležitých aspektů repozitáře a nedodržování odpovídajících de facto norem a mezinárodních standardů. To platí zejména pro prvky dlouhodobé ochrany, tak jak jsou definovány ISO standardem 14721:2003.

Podkladem pro efektivní IT infrastrukturu je vhodný technický plán. PLATTER zdůrazňuje tři hlavní oblasti technické a technologické infrastruktury – rozsah, bezpečnost a služby [ROSENTHAL, 2009]. **Pokud se repozitář drží standardů, má zabezpečnou dostatečnou podporu a je si vědomý rizik a hrozeb, jež vznikají z daného typu repozitáře, potom není ani tak podstatné, jestli repozitář funguje na systémech otevřeného kódu (*Open-source*) či v některém z komerčních řešení. Oba přístupy mají svoje klady i zápory.** U komerčních řešení může být riziko ztráty podpory v případě zániku subjektu, u „open-source“ pomalejší vývoj a riziko ztráty dostatečného personálního zabezpečení.

Důležitější než volba „open-source“ versus komerční systém je zabezpečení dostatečně škálovatelných hardwarových, softwarových a síťových systémů, garance integrity, bezpečnosti a obnovitelnosti dat. Ideální je nechat repozitář podrobit bezpečnostnímu auditu v pravidelnějších intervalech (např. 1x za 2-3 roky). Kromě offline záloh na páskách je vhodné mít i online zálohy (pro snadnější obnovitelnosti dat v RAIDu).

Postupně se začíná prosazovat i bezpečnost, která se týká ochrany proti neautorizovanému přístupu jak vně systému, tak směrem ven, k cílovým skupinám. Nápomocný může být i ISO 14641-1:2012 Electronic archiving -- Part 1: Specifications concerning the design and the operation of an information system for electronic information preservation či další standardy¹²⁸. Popisuje metody a techniky¹²⁹, které mohou být použity u elektronického informačního systému pro správu dokumentů v archivu.

Aby byl chod repozitáře bezproblémový, musí technický plán vycházet z managementu rizik a být sestaven v úzké návaznosti na krizový plán.

8.2.1 Datový plán

Datový plán definuje informační balíčky SIP, DIP a AIP (podrobněji viz Kapitola 3.3 Open Archival Information System Reference Model - OAIS) a podrobně rozpisuje přijímané a spravované formáty a metadata v nich. Datový plán v tomto ohledu navazuje na akviziční plán. U všech typů informačních balíčků, tak jak jsou definovány v ISO 14721:2003, je důležité specifikovat

¹²⁸ Další důležité standardy pro tuto oblast jsou: ČSN ISO/IEC 27000 - 27002 - Systém řízení bezpečnosti informací, ČSN ISO/IEC 15408 - Kritéria pro hodnocení bezpečnosti IT, ČSN ISO/IEC TR 13335-1 - 13335-4 - Směrnice pro řízení bezpečnosti IT nebo ISO/IEC TR 13335-5 - Guidelines for the management of IT Security, Management guidance on network security.

¹²⁹ V úvahu přichází využití jednoho ze tří možných archivačních médií: fyzické WORM, logické WORM a prepisovatelná média. Archivní integrita je zajištěna na fyzických a logických WORM médiích. Na prepisovatelných médiích je integrita zabezpečena pomocí šifrování a technik typu výpočtu kontrolního součtu nebo „hashovacích“ funkcí, datum a časové razítko či digitální podpis. Ve všech případech je nutné, aby byly v souladu s příslušným workflow v organizaci repozitáře.

zdroj a formát metadat (jak popisných, tak technických, strukturálních a administrativních). Je potřebné vědět, která metadata repozitář získá od producentů dat či jiných třetích stran, která získává automatizovaně (přes protokoly Z39.50 nebo OAI-PMH) či extrahuje (např. pomocí nástrojů JHOVE, DROID apod.) a která doplňuje manuálně. Nesmí se zapomínat na sestavení a dodržování specifikací transformace SIP na AIP (z akvizice do permanentního úložiště) a AIP na DIP (z permanentního úložiště do modulu zpřístupnění uživatelům).

8.3 Akviziční plán repozitáře

Jakmile repozitář zná své poslání a cíle, má zabezpečenou robustní hardwarovou a softwarovou architekturu s jasně definovaným datovým modelem, může se pustit do plánování akvizice dat. Podle Rosenthal et al. [2009] má mít akviziční plán tři hlavní části: specifikaci uchovávaných dokumentů, vyjednávání smluv potřebných k jejich získání a vývoj procedur i metod jejich akvizice. Je důležité zdokumentovat transakce mezi repozitářem, knihovnou či archivem a poskytovatelem obsahu, jež budou odsouhlasené všemi zúčastněnými stranami dohody k odevzdávání dat (případně včetně vhodných formátů). Také je potřebné smluvně ošetřit [RESEARCH, 2002; DOBRATZ, 2009]:

- seznam intelektuálních entit (digitálních objektů), které mají být archivovány (ideálně taktéž stanoveny signifikantní vlastnosti)
- seznam informačních objektů a dat, která představují tyto intelektuální entity
- seznam požadovaných metadat
- stanovení SIP, včetně požadovaných metadat a identifikátorů a přiřazení informačních objektů do SIP balíčků
- ustanovení zástupců pro právní úkony a zodpovědné osoby na obou stranách
- způsob odevzdávání dat (FTP, externí hard disk apod.)
- efektivní procedury a „workflow“ pro získání odbavení autorských práv pro krátkodobý a okamžitý přístup i dlouhodobou ochranu
- procedury a systémy pro zabezpečení autenticity odevzdaných materiálů
- vstupní posouzení kompletních dat
- efektivní evidence veškerých transakcí, včetně probíhajících vztahů s poskytovateli dat
- ověření metadat - metadata je nutné ověřovat již při jejich odeslání do systému společně s objektem, který má být uchován
- unikátní a persistentní identifikace materiálů apod.
- harmonogram činností.

Je jen samozřejmostí, že veškeré smluvní ujednání musí být pravidelně aktualizováno.

Příklad dokumentu určující podmínky příjmu dat do repozitáře [Požadavky, 2013]:

Požadavky na producenty dat vkládaných do repozitáře

Vložení objektů do repozitáře

Do Digitálního univerzitního repozitáře jsou objekty přijímány pouze po dohodě s administrátory repozitáře (digitoool@cuni.cz). Na základě dohody s administrátory bude stanovena přesná podoba importního balíčku. Producent (součást UK poskytující data) v první fázi poskytne testovací vzorek dat, na jehož základě bude vytvořen testovací digitální objekt (případně objekty). V průběhu testování bude optimalizováno importní workflow. Poté budou objekty vkládány dávkově na základě předchozí dohody.

Informace o sbírce

Producent společně s digitálními objekty dodá i stručnou informaci o dané sbírce, rozsáhlejší text popisující sbírku, který bude dostupný v rámci informačních stránek repozitáře a ideálně i ilustrační obrázek, který bude zobrazen na hlavní stránce repozitáře. Texty je vhodné dodat v českém a anglickém jazyce.

Dostupnost objektů

Producent dodá písemně (elektronickou formou) souhlas se zveřejněním digitálních objektů, a to včetně přístupové politiky k nim. Možnosti jsou následující:

1. objekty jsou v režimu open access – jsou dostupné komukoli
2. objekty jsou dostupné komukoli, ale po odsouhlasení licenčních podmínek – producent dodá text licenčních podmínek v českém a anglickém jazyce
3. objekty jsou dostupné pouze studentům a zaměstnancům Univerzity Karlovy
4. objekty jsou dostupné pouze z určitých IP adres (variantně pouze z počítačů v síti univerzity)
5. objekty jsou pouze v archivním režimu (dokumenty nejsou vyhledatelné prostřednictvím standardního uživatelského rozhraní).

Metadata

Jakýkoli objekt vkládaný do repozitáře musí být bezpodmínečně opatřen popisnými metadaty a v případě, že se jedná o komplexní objekt (skládající se z více než jednoho souboru) i strukturálními metadaty.

Vkládaný digitální objekt by měl mít záznam v centrálním katalogu univerzity (CKIS), který bude následně připojen i k objektu v repozitáři. Pokud existují důvody bránící tomu, aby objekt měl záznam v CKIS, je např. možné připojit záznam ve formátech MARCXML, MODS a DC.

Do repozitáře je možné vložit i komplexní strukturovaný objekt – tedy objekt skládající se z více částí - například historický dokument, kde jedné stránce odpovídá jeden obrázek. V tomto případě je však nutné dodat informaci o tom, jaká přesně je struktura objektu. Tuto informaci je možné předat dvěma způsoby:

1. ve formátu strukturálních metadat METS (*Mets Profile*¹³⁰)
2. objekt bude dodán v adresářové struktuře odpovídající hierarchii fyzických souborů a případně s excelovou tabulkou odpovídající jejich logické struktuře (např. rozdělení do kapitol).

Rozsáhlejší sbírky musí být opatřeny základními technickými metadaty, která budou obsahovat minimálně ID objektu, název souboru a kontrolní součet (např. MD5). ID objektu musí být unikátní.

Formáty objektů

Přijímaný formát objektů je závislý na povaze konkrétní sbírky.

Obrazové materiály

- bezztrátový TIFF – bude uložen jako archivní kopie, rozlišení minimálně 300 dpi, nepřijímáme obrázky s LZW kompresí (u této komprese existuje riziko z hlediska dlouhodobé ochrany digitálních objektů)

- JP2000 – generuje se v rámci importního workflow, využít jako kopie pro zobrazení

Textové materiály:

Formát PDF, v ideálním případě jeho archivní variantu PDF/A.

¹³⁰ <http://www.loc.gov/standards/mets/profiles/00000021.html>

Audiovizuální materiály:

Zatím neurčeno

Přenos objektů

Digitální objekty je možné DUR předat na externím záznamovém médiu (disk e-SATA, USB 3, DVD) nebo v případě větších sbírek přes SFTP.

Při přenosu většího množství souborů vyžadujeme doplnění várky předávacím protokolem (formát txt), který bude obsahovat počet souborů a jejich seznam.

Názvové konvence

Předávané soubory musí být pojmenované podle následujících pravidel:

- Název souboru nesmí obsahovat diakritiku a apostrofy, jako oddělovač může být využita pomlčka (-) nebo podtržítka (_)
- Název souboru je zároveň jeho unikátním identifikátorem, ideální je zachovat nějakým způsobem vazbu na digitální objekt (v případě digitálních objektů skládajících se z jediného souboru název souboru odpovídá unikátnímu identifikátoru)
- Kdykoli je to možné, doporučujeme při pojmenování souborů využít číselnou řadu, která zachová správné řazení souborů v rámci objektu.

Konec příkladu.

8.4 Práce s objekty, dlouhodobá ochrana

Získání dostatečné kontroly nad informacemi si vyžaduje analýzu současného či budoucího předpokládaného digitálního obsahu. Zde je vhodné stanovit a uplatňovat politiky o akceptovatelných formátech, podobě SIP (např. co SIP to jeden digitální objekt; jeden digitální objekt zodpovídá METS standardu). **Z praktických zkušeností existujících repozitářů totiž vyplývá, že dlouhodobá ochrana funguje nejlépe, pokud je digitální obsah spolu se všemi metadaty uchováván jako jeden soubor. METS (Metadata Encoding and Transmission Standard) pro kódování popisných, administrativních, strukturálních metadat tento požadavek naplňuje.**

Jakmile je rozhodnuto, co se bude uchovávat, mělo by příslušné rozhodnutí být co možná nejvíc automatizováno a obsah by měl být systematicky analyzován již při odevzdání materiálu do

systému. Repozitář též musí zabezpečit, že dohody s poskytovateli dat se vztahují jak na popisná, tak technická metadata, které musí poskytovatel dodat. Proto je nevyhnutné zabezpečit kontroly autenticity, potvrzení, že dokumenty odevzdané do repozitáře jsou přesně ty, které poskytovatel dat odeslal. Přičemž kontrola integrity by měla zahrnovat ověření nejen integrity datového toku (*Bitstream*), ale též potvrzení funkčnosti a použitelnosti objektu.

Je potřebné určit i signifikantní vlastnosti digitálních objektů, které však nejsou obecně stanovitelné, resp. absolutní a statické. Jejich výběr závisí na typu organizace, služeb, které poskytuje a jejich roli v dlouhodobé ochraně. Avšak **je nutné si uvědomit, že čím víc signifikantních vlastností je stanoveno pro další uchování, tím víc metadat je nutných a dlouhodobá ochrana je náročnější.** Vytvoření a uchování detailních metadat,¹³¹ vztahujících se k signifikantním vlastnostem objektů je kritické pro podporu ochranné funkce repozitáře.

Vyšší počet vybraných klíčových vlastností pro další uchování komplikuje i přístupy k dlouhodobé ochraně. V případě migrací se často zužuje možný výběr formátu a navíc je testování vhodných formátů mnohem náročnější. Repozitáře mohou změnit svůj přístup k objektům i bez pozměnění samotného objektu. Je možnost zpřístupňovat dokumenty tzv. migrací "za pochodu" (on-the-fly migration), bez uchování daných migrovaných verzí v permanentním úložišti. Pokud však objekt spoléhá na komplexní technické prostředí nebo využívá proprietární technologie, emulace prostředí se zdá být vhodnějším řešením pro zabezpečení dlouhodobého přístupu.

Řádně zdokumentované politiky pro uchování, úložiště a údržbu (permanentní úložiště), jež zaručí bezproblémový a spolehlivý chod repozitáře jsou samozřejmostí při dodržení základní premisy repozitáře - informace budou uchovány srozumitelně, nezávisle na cílové skupině tak, aby jim uživatelská komunita porozuměla bez asistence expertů.

8.5 Finanční zabezpečení repozitáře

Jednou z největších obav současného (na neustálý ekonomický růst zaměřeného) světa je ztráta finančních zdrojů. Právě nedostatek finančních prostředků bývá hlavní hrozbou správy repozitáře a dlouhodobé ochrany dokumentů v nich uchovaných. Proto je důležité mít k dispozici finanční plán. I když je nemožné sepsat jednotná pravidla finančního plánu pro všechny typy repozitářů na několik let předem, je možné začít u vyhodnocení finančních zdrojů za posledních několik let a pokusit se o prognózu vývoje, zejména pro oblast dlouhodobé ochrany digitálních

¹³¹ Podrobná metadatová specifikace je kritická i pro konsorciální úložiště či síťové repozitáře.

objektů. Zde nám poměrně dobře poslouží ekonomické modely¹³² (např. LIFE). Ty mohou napomocť při obhajobě vynaložených nákladů či navýšení finančních prostředků u vyššího managementu organizace repozitáře. Pokud to dovolí status repozitáře a organizace, pomohou podnikatelské aktivity. Taktěž může být kužitku hledání sponzorů či nových grantových příležitostí.

Finanční plán je úzce propojen s managementem rizik a odpovídajícím krizovým plánem.

8.6 Personální zajištění repozitáře

Dobře fungující repozitář vyžaduje nejen dobré technické zázemí, finanční zabezpečení či podporu managementu, ale především lidský prvek v podobě kvalifikovaných zaměstnanců nejrůznějších oborů – knihovníky, archiváře, administrátory a další IT personál. Jelikož problematika repozitářů představuje relativně novou oblast pracovní realizace, pro kterou neexistuje klasický „rámec“ kariérního postupu a terciární vzdělávání, není jednoduché sehnat a udržet si vhodné odborníky. Tuto situaci navíc často ztěžují finanční možnosti repozitářů. I pro odůvodnění oprávněnosti personálního zajištění u vyššího managementu repozitáře **je vhodné mít tzv. plán personálního zajištění, jak ideální stav i tzv. minimálně udržitelný stav.** Plán by měl obsahovat název pracovní pozice, podrobný popis pracovní náplně – klíčové úkoly, řešené problémy, spolupráce s jinými součástmi organizace repozitáře, zkušenosti, znalosti, pravomoci, odpovědnosti apod.

8.7 Plán zpřístupňování

Zpřístupňování dokumentů repozitáře cílové skupině je definováno v mandátu, resp. posláni repozitáře. Plán zpřístupňování by měl být součástí vize či programového prohlášení repozitáře. Měl by obsahovat:

- definované cíle repozitáře
- vyjádřené závazky dlouhodobého uchování, správy a zpřístupňování digitálních informací cílové komunity uživatelů a potvrzený závazek např. prostřednictvím auditů a certifikace
- potvrzení důležitosti vztahu repozitáře se svými cílovými skupinami získáváním zpětné vazby (např. pomocí průzkumů, dotazníků apod.)

¹³² Podrobnější informace jsou dostupné v kapitole 3.2 Ekonomické modely.

- krátkodobé, střednědobé a dlouhodobé cíle (pro všechny základní oblasti repozitáře – příjem, správa, dlouhodobá ochrana a zpřístupnění).

Plán by měl zahrnout i autorsko-právní problematiku (formulace a implementace politiky zpřístupňování obsahu repozitáře), použitelnost a přístupnost uživatelských rozhraní (interních pro zaměstnance i externích pro uživatele).

8.8 Krizový plán a management rizik

Krizový plán¹³³ by měl řešit ohrožení, kterým repozitář může čelit a nikoliv pouze na úrovni dat nebo služeb, ale i repozitáře jako celku, resp. organizace (např. v případech, kdy je ohrožena existence samotného repozitáře). Je potřebné začít přehledem prvků kritické infrastruktury, kterých narušení by mělo zásadní dopad na funkčnost celého systému repozitáře. **Jde o seznam možných zdrojů rizik a analýzu ohrožení, tedy výčet konkrétních hrozeb, které mohou způsobit vznik specifické krizové situace a zhodnocení působení konkrétní hrozby.** Ty se zpracovávají do písemných výstupů, přičemž je důležité uvést seznam zaměstnanců, přehled právnických osob a podnikajících fyzických osob, které zajišťují plnění opatření vyplývajících z krizového plánu. V přehledu spojení se uvede zejména seznam telefonních kontaktů a elektronických adres na ustanovený krizový štáb a další subjekty (osoby, instituce, firmy) podílející se na odstraňování krizových situací. Dále je potřebné sestavit i souhrn využitelných právních předpisů, eventuelně další dokumenty související s připraveností na krizové situace a jejich řešení.

Plán obnovy po havárii (*Disaster Recovery Plan*) představuje sadu dokumentů¹³⁴ - postupů řešení konkrétních hrožících krizových situací identifikovaných v analýzách ohrožení, zejména:

- název a charakteristika krizové situace
- plánovaná činnost zodpovědných osob, oddělení apod.
- opatření k řešení krizové situace včetně stanovení odpovědností
- síly a prostředky nezbytné pro řešení krizové situace.

¹³³ Poměrně obsáhlý zdroj informací ke krizovému managementu a zvládání katastrof (crisis and disaster management) lze najít na stránkách: <<http://eulibraryblog.com/crisis-disaster-management-resources/#DisPol>>.

¹³⁴ Dokumenty je možné klasifikovat na politické (na globální i lokální úrovni – např. politické převraty a ořesy nebo i ztráta mandátu či potřeby existence repozitáře), ekonomické (např. citelný úbytek finančních zdrojů), technické (sledování nejnovějšího vývoje technologií nebo selhání technických prostředků), lidské či personální (selhání lidského faktoru, úmyslná nebo neúmyslná sabotáž, ztráta klíčového zaměstnance), externí (např. dodavatelé služeb) a v neposlední řadě ty nejméně ovlivnitelné – přírodní katastrofy (povodně, extrémní teploty, zemětřesení atd.).

Pro sestavení dobrého plánu obnovy je potřebné znát odpovědi na tyto otázky (Wallace, 2004):

- Co jsou naše kritické aktiva?
- Jaká jsou rizika pro tyto aktiva?
- Jak je možné snížit riziko výskytu hrozby?
- Jakým způsobem je možné případné škody minimalizovat?
- Co má dělat tým repozitáře, pokud k událostem dojde?
- Kde můžeme najít informace pro sestavení našeho plánu?

Příklad zjištěných rizik repozitáře NK ČR¹³⁵ [HUTAŘ, 2008]:

Organizace a management

Celková koncepce Národní digitální knihovny není schválena vládou	21
Nedostatečné vyčlenění prostředků na zajištění chodu repozitáře (nejen finančních)	20
Pozice "digitální ochrany" v organizační struktuře Národní knihovny ČR	15
Smlouvy s producenty, dodavateli nejsou naplněny	12
Dokumentace procesů, metod a procedur upravujících různé aspekty všech aktivit	12
Mandát repozitáře	12
Certifikace digitálního repozitáře jakožto důvěryhodného nebude úspěšná	12
Selhání managementu	12
Plán pro následné aktivity digitální ochrany po skončení životnosti repozitáře	10
Odpovědnosti a povinnosti vyplývající z legislativy nebo jiných nařízení nejsou naplněny	9
Ztráta důvěryhodnosti	9
Nereálná očekávání rozsahu využití/úspěchu repozitáře	9

Personální zajištění

Nedostatek kvalitních zaměstnanců	20
Ztráta klíčového zaměstnance/ů	15
Vypovězení nebo změna smluv ze strany externích dodavatelů	9
Nedostatečný počet zaměstnanců v odboru digitální ochrany	9
Zkušenosti a znalosti zaměstnanců zastarají	9
Dobré pracovní prostředí a odměny zaměstnanců	8

Konec příkladu.

¹³⁵ Číselné hodnoty u jednotlivých rizik představují součin pravděpodobnosti výskytu rizika a jeho možného dopadu (podle metodiky DRAMBORA).

U identifikace rizik to však nekončí, pokračuje se jedním ze tří přístupů řešení. Prvním je potlačení (*Mitigation*), kdy je potřebné provést něco, čím by se snížila pravděpodobnost výskytu nebo výše škody způsobené událostí, které nebylo možné se vyhnout. Dále existuje možnost vyvarování se, tedy vyhnoutí se určité události či situaci. Posledním způsobem je delegace rizika třetí straně či stranám.

8.9 Kontinuita repozitáře

Vyjednání nástupnictví v praxi není snadná záležitost, protože odhadnout, kdo za 20-30 či více let bude „vhodným kandidátem“ pro převzetí dat, je téměř nemožné. Na jedné straně může jít o zánik repozitáře, na druhé straně zánik instituce, která daný repozitář spravuje. V tomto ohledu mají konsorciální repozitáře jistou výhodu.

Plán nástupnictví může mít alespoň rámcovou podobu, přičemž v době, kdy repozitáři hrozí reálný zánik, vyjedná konkrétní podmínky nástupnictví s vybranou organizací (např. podobného typu, velikosti, ve stejném regionu apod.). V takovém případě se uvedou podrobné informace o instituci a/nebo repozitáři, jež data převezme a stanoví, v jaké podobě se převezmou (včetně datové a metadatové specifikace) a v jaké podobě se budou do budoucna zpřístupňovat či uchovávat. Nesmí se zapomenout na licenční ujednání (autorsko-právní problematika). V neposlední řadě je vhodné (tam, kde je to možné) nabídnout případné kompenzace za akceptaci dojednaných závazků (např. finanční zdroje původní instituce).

Ilustrativní „Plán nástupnictví Digitálního univerzitního repozitáře“:

Digitální univerzitní repozitář se zavazuje dlouhodobě uchovávat a zpřístupňovat své cílové skupině veškeré informace a dokumenty, jež mu byly svěřeny, a to podle nejlepšího vědomí a svědomí. Jelikož není možné podchytit všechna možná rizika, které mohou činnost repozitáře ohrozit, je nutné zabezpečit obsah a práva pro cílovou skupinu (např. v případě zániku repozitáře).

Principy nástupnictví

- Univerzita Karlova v Praze (dále jen UK) si ponechá vlastnictví materiálů Digitálního univerzitního repozitáře (dále jen DUR).
- UK se zavazuje k pokračování uchování a zpřístupňování materiálů.
- UK může zajistit převedení materiálu třetí osobě za účelem dalšího uchování a/nebo poskytnutí přístupu. To nemusí být v rozporu s právem UK na uchování a zpřístupňování materiálů pro svou cílovou skupinu, na které má tato skupina nárok.

Konec příkladu.

8.10 Projektové řízení

K tomu, aby projekt mohl být úspěšně dokončen (a tím pádem i hlavní cíl naplněn či problém vyřešen) je nutné využít metod projektového řízení. V anglosaské literatuře se nejčastěji setkáme s termínem *Project Management*. Do češtiny jej lze přeložit dvěma způsoby - jako řízení projektu, případně projektové řízení. Projektové řízení je často chápáno jako filosofie přístupu k řízení projektu s jasně stanoveným cílem, který musí být dosažený v požadovaném čase, s určenými náklady a v požadované kvalitě. Projektový cíl se skládá ze tří základních podmínek, proto také někdy bývá označován jako trojpodmínka či trojimperativ. Jednotlivé podmínky - provedení (kvalita výstupu), čas a náklady - jsou na sobě závislé. Tedy změna jedné podmínky nutně vyvolá změnu dvou zbývajících. V ideálním případě jsou všechny tři roviny naplněny současně a nedochází ku příkladu k preferenci času na úkor kvality a nákladů.

Výběr nejvhodnější organizační formy projektového řízení ovlivňují zejména následující faktory: struktura projektu, jeho fáze a rozsah řešených problémů, způsob zapojení organizačních struktur účastníků projektů, míra institucionalizace subjektu projektového řízení, právní, ekonomická a jiná omezení a formy. Dále systém legislativních, ekonomických, ale také "zvykových" norem vytváří specifické podmínky pro organizaci projektového řízení, potenciál spolupracovníků, míru ochoty a schopnost spolupráce účastníků projektu a v neposlední řadě úroveň informačních systémů účastníků projektu.

Řízení projektů je ale především o práci s lidmi a za její podcenění se draze platí! Spousta projektů je dnes svěřena lidem, kteří s řízením takto specifických aktivit nemají zkušenosti. Většinou bývají odborníky v oboru, skrze který bude získán výstup projektu. Ovšem odborník nemusí být vždy zdatným manažerem schopným uřídit tak nestabilní a nejistotou zatížený prvek, jakým je projekt. **Existuje však i opačný extrém, kdy se největší důraz klade nikoliv na lidskou složku, ale na používání nejrůznějších nástrojů a zpracování co nejvíce projektových metodik s cílem eliminovat rizika nesplnění termínů, kvality a nákladů, přičemž uniká samotná podstata řízení a řešení projektu.**

8.11 Audit, certifikace a politiky repozitáře

Posledním krokem při sestavování repozitáře by měl být nezávislý (ideálně externí) audit a certifikace dle některé z mezinárodně uznávaných certifikačních metod. Rozhodně nemá být samoučelný. Repozitář si musí být vědom svých silných, ale i slabých stránek. Repozitář je totiž pouze tak pevný, jaký pevný je jeho nejslabší článek.

Cílem je identifikace a podchycení slabých míst repozitáře, přičemž je nutné najít optimální řešení. **Před externím auditem by měl proběhnout interní audit (např. dle metodiky Platter, DSA; u rozsáhlých repozitářů spíše TRAC).** Začínat by se mělo vyhodnocením současného stavu. Je nutné (kromě jiného) se ptát, jaká je cílová skupina repozitáře (zůstává pořád stejná nebo se naopak mění?). Jaké sbírky repozitář má, kolik objektů jednotlivé sbírky čítají, jaké formáty a v jakých verzích dané digitální objekty sbírky obsahují, jak je to s integritou dat (ověřování kontrolních součtů). Dále jaké je hardwarové a softwarové řešení (je dostatek místa na úložištích, funguje bezproblémové zálohování, lze zálohy bez příslušných komplikací obnovit, je repozitář co do architektury škálovatelný, je dostatek finančních prostředků na nákup dalšího hardware a software atd.). Neméně důležitá je otázka zabezpečení chodu repozitáře: jaká je podpora vyššího managementu, jaké je zabezpečení personálních zdrojů repozitáře pro svou činnost, jak úzká je specializace znalostí a zkušeností zaměstnanců, jaká hrozí rizika v případě jejich odchodu, jaké jsou finanční prostředky pro chod repozitáře pro technické i organizační stránce, jaký je poměr stálých a „náhodných“, resp. nepravidelných prostředků, např. grantových apod.

Pro zabezpečení bezproblémového zpřístupňování dokumentů a jejich dlouhodobou ochranu je též nevyhnutné zdokumentovat politiky a procedury, jež ovlivňují chod repozitáře jako organizace. Pro distribuované repozitáře (konsorciální řešení) to znamená především přesné vyjádření zodpovědnosti jednotlivých participujících stran. Politiky pro dlouhodobou ochranu digitálních souborů musí vhodně "zapadnout" do již existujících politik pro nedigitální obsah. Spojení mezi politikou a výslednými postupy v praxi je velmi důležité [RESEARCH, 2002].

Doporučení [RESEARCH, 2002] pro důvěryhodné a efektivně fungující repozitáře:

1. Vytvořte rámec a sestavte procesy, které podporují certifikaci digitálních repozitářů
2. Prozkoumejte a vytvořte nástroje pro identifikaci atributů digitálních materiálů, které musí být uchovány
3. Prozkoumejte a vytvořte modely pro kooperativní sítě a služby repozitářů
4. Navrhněte a vytvořte systémy pro unikátní, persistentní identifikaci digitálních objektů, které podporují dlouhodobou ochranu

5. Hledejte a šiřte informace o komplexních vztazích mezi dlouhodobou ochranou a právem na duševní vlastnictví.
6. Prozkoumejte a určete, které technické strategie nejlépe slouží dlouhodobému přístupu k digitálním zdrojům
7. Prozkoumejte a definujte minimální úroveň metadat potřebných pro dlouhodobé uchování a správu digitálních informací. Navrhněte nástroje pro automatické generování a/nebo extrakci co největšího množství požadovaných metadat.

Závěr

V předložené disertační práci je zanalyzován současný stav, návrhy, administrace a řízení v oblasti problematiky digitálních knihoven, repozitářů a archivů. **Disertační práce má nejen přehledový a vysvětlující charakter problematiky, ale i praktický přínos v definování obecných strategií, návrhů, řízení a administrace jak pro management digitálních knihoven, tak i repozitářů a archivů.**

Při výběru vlastního zaměření práce bylo důležité, aby s ní spojený výzkum byl užitečný i prospěšný pro praxi a aby měl poznatelný, resp. identifikovatelný předmět. Na počátku výzkumu byla stanovena tzv. pracovní hypotéza, která byla dále přetavena do reálné hypotézy (viz část 1. Předmět a cíl práce). V průběhu zpracování práce byla hlavní hypotéza na základě studia dostupné odborné literatury, komunikace s odborníky (nejen na konferencích) a dotazníkového šetření (mezi uživateli/studenty i institucemi) potvrzena. Tato hypotéza společně s definovanými tezemi (jež jsou součástí samostatného dokumentu), stanovily záběr, strukturu a samotný obsah předložené disertační práce. Navíc byla doplněna o množství obrazových materiálů, přehledných tabulek, příloh a poznámek pod čarou odkazujících na podrobnější informace či další doporučenou literaturu.

Práce má ucelenou strukturu jednotlivých kapitol, kdy kapitoly 1 až 6 nabízí přehled multioborového okruhu otázek, který vychází především z amerických a anglo-saských informačních pramenů. Jsem si vědoma toho, že některé okruhy problematiky nejsou rozebírány téměř vůbec nebo ne zcela do hloubky, což je dáno především přehledovým charakterem práce a skutečností, že k těmto dílčím tématům již bylo napsáno několik kvalitních disertačních či diplomových prací (metadata, dlouhodobá ochrana, open source systémy pro digitální knihovny apod.). Pro posouzení stavu v České republice (kapitola 7), bylo nutné provést dotazníková šetření a interview u relevantních institucí. Kapitola 8 pak vychází z výsledků výzkumu a nabízí příručku pro sestavení vhodných plánů, procedur a postupů v rámci organizace repozitáře. Že nejde o snadnou věc, dokazuje i multidisciplinární záběr problematiky; od technických poznatků přes krizový management, management rizik až po audit a certifikaci. I proto jsem sestavila přehled nejdůležitějších standardů, metodik a dokumentů pro efektivní správu repozitářů, knihoven a archivů, jež jsou součástí přílohy č. 7.

Z důvodu nejednoznačné definice jsou pro účely této disertační práce termíny digitální knihovna, digitální repozitář, digitální archiv, virtuální knihovna, elektronická knihovna a hybridní knihovna významově zaměnitelné a označují tak organizaci, jejímž účelem je komplexní sběr, správa a dlouhodobá ochrana digitálních materiálů. Pro větší typy digitálních knihoven, jako třeba

rozsáhlou digitální knihovnu neexistuje doposud jednoznačně akceptovaná definice. Nejčastěji je tento typ knihoven vnímán z pohledu množství či typů objektů (např. velké formáty dokumentů jako jsou mapy), počtů uživatelů, kterým jsou dokumenty zpřístupňovány apod. **Větší digitální knihovny, repozitáře a knihovny jsou v této práci chápány a definovány jako instituce, které mohou se svou digitální knihovnou, repozitářem či archivem absolvovat alespoň jeden stupeň certifikace podle Evropského rámce pro certifikaci a audit digitálních repozitářů.**

Aby však digitální knihovny, repozitáře či archivy mohly takovou certifikací projít, musí zodpovídat některým ze standardních modelů architektury. Všechny tři nejnámější a v práci rozebírané modely přistupují k digitálním knihovnám, repozitářům a archivům odlišným způsobem.

V popředí Kahn-Wilenskeho architektury stojí digitální objekt s persistentními identifikátory, jež tvoří základ digitálních knihoven. Avšak digitální objekty nejsou jedinou podstatnou složkou digitálních knihoven, repozitářů či archivů. Významnou roli zde hrají i lidé (specializovaný personál, uživatelé), zařízení, softwarové systémy, postupy, komunity apod. Proto DELOS Digital Library Reference Model definuje třívrstvou konstrukci digitálních knihoven, kterou tvoří digitální knihovna na obecné úrovni, vybraný systém digitální knihovny a systém pro správu digitální knihovny (na technické úrovni). Daný model je průlomový především definováním prostředí digitálních knihoven, ve kterých působí a dlouhodobou ochranu, která zásadním způsobem ovlivňuje dostupnost digitálních objektů. Tuto myšlenku dále rozvíjí Referenční model Otevřeného archivního informačního systému (OAIS), který v rámci svého modelu definuje, zejména pro repozitáře a archivy, modul pro dlouhodobou ochranu. Aby návrh, řízení a administrace digitálních knihoven, repozitářů a archivů byly co nejvíce efektivní, je nutné porozumět zaužívané terminologii (viz kapitola 1.2 Vymezení základních používaných termínů). Velkou předností uvedeného OAIS modelu je i vymezení terminologického základu, stanovení tzv. informačních balíčků a popis funkčních entit. V současné době se stal v odborných kruzích téměř výlučnou normou pro digitální repozitáře a repozitáře právě OAIS model.

Pro digitální repozitáře, knihovny či archivy, které se drží standardů, mají zabezpečnou dostatečnou podporu a jsou si vědomé rizik a případných hrozeb, není ani tak podstatné, jestli fungují na systémech otevřeného kódu („open-source“) či v některém z komerčních řešení. Oba přístupy mají svoje klady i zápory. U komerčních řešení může být riziko ztráty podpory v případě zániku subjektu, u „open-source“ pomalejší vývoj a riziko ztráty dostatečného personálního zabezpečení. Důležitější než volba „open-source“ versus komerční systém je zabezpečení dostatečně škálovatelných hardwarových, softwarových a síťových systémů, garance integrity, bezpečnosti a obnovitelnosti dat.

To by mělo vést k optimální dlouhodobé ochraně dat, resp. z nich vycházejících digitálních dokumentů. Pokud by si člověk měl z problematiky dlouhodobé ochrany odnést pouze jednu myšlenku, byla by to ta, která říká o neschopnosti uchovávat vše. Z hlediska personálních, finančních, ale i technických možností to není možné a zůstává otázkou, zda-li je to žádoucí. I u samotných objektů nelze uchovávat vše, proto je nutné stanovit klíčové, tzv. signifikantní vlastnosti, které se knihovna, repozitář či archiv zaváže do budoucna zachovat nejrůznějšími prostředky a pomocí dostupných komerčních či "open source" řešení. Koneckonců nejde ani tak o konkrétní řešení jako o koncepční, systematický přístup a dodržování standardních postupů, zásad a příslušných norem.

To platí i pro uživatelská rozhraní, kde dodržování nejdůležitějších zásad použitelnosti zvyšuje jejich srozumitelnost a přehlednost. Interakce s rozhraním by tedy měla probíhat bez nutnosti složitého přemýšlení. Rozhraní by mělo být uživatelsky co nejpřívětivější. **Současný trend je zcela zaměřen na uživatele a jeho potřeby. Uživatel je totiž centrálním a určujícím činitelem celého informačního systému, bez respektování potřeb uživatelů funguje jakýkoliv systém (tedy i digitální repozitář, knihovna či archiv) víceméně naprázdno.**

Potvrzení, jestli instituce se svými digitálními knihovnami, repozitáři či archivy jdou správným směrem a jejich systémy a procesy nefungují bezúčelně, je dáno několika auditními a certifikačními nástroji. Hodí se jak v přípravných (pro kontrolu návrhů), tak post-implementačních fázích běžného chodu repozitáře (pro kontrolu realizace). Poslouží však i komerčním a nekomerčním poskytovatelům služeb, vývojářům software a dodavatelům třetích stran pro případnou spolupráci s kulturními institucemi při návrhu, vývoji a provozu technologií pro digitální repozitáře.

Jak potvrdil výzkum v této disertační práci (dotazníková šetření a interview), standardy pro dlouhodobou ochranu nejsou téměř vůbec dodržovány, což je velmi zneklidňující. Obdobně je to i se standardy v oblasti návrhu architektury digitálních knihoven a repozitářů, sestavování digitálních objektů, jako u návrhů uživatelského rozhraní a certifikace procesů uvnitř knihoven a archivů. Knihovny rozlišují pojmy digitální knihovna a digitální repozitář, kdy druhý zmíněný je buď přesný protějšek kamenné knihovny či archivu nebo má na rozdíl od digitální knihovny nastavené procesy pro dlouhodobou ochranu a jim odpovídající nástroje. Také si uvědomují naléhavost řešení dlouhodobé ochrany a potřebu auditů – dvě knihovny pravidelně provádí interní audit DRAMBORA.

Protože odpovědi ukázaly, že minimum institucí dodržuje standardy nebo dokonce některé ještě nemají svůj digitální repozitář či archiv, jevílo se jako nevyhnutné vypracovat autorkou pomůcku, návrh či doporučení v podobě „Best Practices“ pro obecné strategie, plány, management a administraci v českých (slovenských) digitálních knihovnách, repozitářích a

archivech. Slouží pro komplexní, přesto snadnou a rychlou orientaci v problematice.

Velký přínos práce spatřuji ve zmapování dané problematiky, kompilace a výběr nejdůležitějších nejnovějších poznatků v této multidisciplinární problematice s důrazem na důležitost výběru signifikantních vlastností digitálních objektů pro efektivní dlouhodobou ochranu. Není možné dlouhodobě uchovávat vše, je nutný výběr dokumentů, včetně jejich signifikantních vlastností. Praktická ukázka (sestavena autorkou) možných stanovených klíčových vlastností objektů je dostupná formou myšlenkové mapy (viz Příloha č. 1). Neméně podstatný je i souhrn nejdůležitějších standardů a de facto standardů podle problematických okruhů (viz Příloha č. 7 – Seznam nejdůležitějších standardů, metodik a dokumentů pro efektivní správu repozitářů, knihoven a archivů). Dále o nasměrování pozornosti na uživatelská rozhraní z pohledu uživatelsky orientovaného designu a na požadavky uživatelů na digitální knihovny, repozitáře a archivy (včetně získávání zpětné vazby). Právě dotazníkem, jako jednou z výzkumných metod, byla zjišťována zpětná vazba od uživatelů ohledně rozhraní digitálních repozitářů na obecné úrovni s konkrétním příkladem v podobě Digitálního univerzitního repozitáře UK v Praze.

I když je zřejmé, že díky rapidnímu vývoji technologií a získávání nových poznatků v multidisciplinární problematice repozitářů může dojít k částečnému posunu aktuálnosti této disertační práce. Přesto **je možné konstatovat, že předložená disertační práce odpovídá na otázky a plní cíle, které stály na počátku jejího vzniku a ukazuje tak možné směřování do budoucna.**

Slovník zkratk a akronymů

AIP	Archival Information Package
ALA	Americal Library Association
ALCTS	Association for Library Collections & Technical Services
API	Application Programming Interface
CCSDS	The Consultative Committee for Space Data Systems
CC-GNU LGPL	Creative Commons-GNU Lesser General Public License
CD	Compact Disk
CEDARS	CURL Exemplars in Digital ARchiveS
CORBA	Common Object Request Broker Architecture
CPA	Commission on Preservation And Access
CRC	Cyclic Redundancy Check
CRiB	Conversion and Recommendation of Digital Object Formats
ČSN	Česká technická norma
DANS	Data Archiving and Network Service
DARPA	Defense Advanced Research Projects Agency
DC	Dublin Core
DCC	Digital Curation Centre
DIN	Deustches Institut für Normung
DIP	Delivery Information Package
DK	Digitální knihovna
DL	Digital Library
DLS	Digital Library System
DLMS	Digital Library Management System
DOC	Document File Format
DOI	Digital Object Identifier
DPE	DigitalPreservationEurope
DPS	Digital Preservation System
DROID	Digital Record Object Identification
DSA	Data Seal of Approval
DSEP	Deposit Systems for Electronic Publications
DRAMBORA	Digital Repository Audit Method Based on Risk Assessment

DUR	Digitální univerzitní repozitář
DVD	Digital Versatile Disk
EAD	Encoded Archival Description
EK	Evropská komise
EU	Evropská unie
eSATA	External Serial ATA
FITS	Flexible Image Transport System
GIS	Geographic information system
GUI	Graphic User Interface
GNU	"GNU's Not Unix!" – rekurzivní akronym
HATII	The Humanities Advanced Technology and Information Institute
HCI	Human Computer Interaction
HTML/HTML5	HyperText Markup Language
HW	Hardware
IA	Information Architecture
ISBN	International Standard Book Number
ISO	International Organization for Standardization
ISO/IEC	International Organization for Standardization/International Electrotechnical Commission
ISO/NP TR	International Organization for Standardization Not Published Technical Report
ISO/TC	International Organization for Standardization/Technical Committee
JHOVE	JSTOR/Harvard Object Validation Environment
JISC	Joint Information Systems Committee
JP2	JPEG 2000
JPEG	Joint Photographic Experts Group
JPG	Přípona formátu JPEG
KNAV	Knihovna Akademie věd ČR
KOS	Knowledge Organization System
LIFE	Lifecycle Information for E-literature
METS	Metadata Encoding and Transmission Standard
MKP	Městská knihovna v Praze
MODS	Metadata Object Description Schema
MPEG-7	The Multimedia Content Description Interface
MZK	Moravská zemská knihovna

NA	Národní archiv ČR
NARA	National Archives and Records Administration
NBN	National Bibliography Numbers
NDK	Národní digitální knihovna
NEDLIB	Networked European Digital Libraries
nestor	Network of Expertise in Long-Term Storage and Long-Term availability of Digital Resources
NDLP	National Digital Library Project
NFA	Národní filmový archiv ČR
NISO MIX	Z39.87 Data dictionary - technical metadata for digital still images (MIX) – NISO Metadata for Images in XML
NK/NK ČR	Národní knihovna ČR
NLNZ	National Library of New Zealand
NTK	Národní technická knihovna
OAI	Open Archives Initiative
OAI-PMH	The Open Archives Initiative Protocol for Metadata Harvesting
OAIS	Open Archival Information System
OCLC	Online Computer Library Center
OS	Operační systém
PAIMAS	Producer-Archive Methodology Abstract Standard
PID	Persistent identifier
PDF	Portable Document Format
PDF/A	Portable Document Format/Archival
PDI	Preservation Description Information
PHP	PHP: Hypertext Preprocessor
PLATO	The Preservation Planning Tool
PREMIS	Preservation Metadata: Implementation Strategies
PS	Pracovní skupina
PURL	Persistent Uniform Resource Locators
RLG	Research Libraries Group
RAP	Repository Access Protocol
RODA	Repositório de Objectos Digitais Autênticos = Repository of Authentic Digital Objects
SCAPE	Scalable Preservation Environments

SDB	Safety Deposit Box
SFTP	SSH File Transfer Protocol
SIP	Submission Information Package
SMART	Specific, Measurable, Attainable, Relevant, Time-bound
SME	Small and Medium Enterprise
SOHO	Small Office Home Office
SOP	Strategic Objective Plan
SW	Software
TCO	Total Cost of Ownership
TDR	Trusted Digital Repository
TEL	The European Library
TIF	Tagged Image File
TRAC	Trustworthy Repositories Audit & Certification
TXT	Souborový textový formát
UCD	User-centred design
UK	Univerzita Karlova v Praze
UNESCO	United Nations Educational, Scientific and Cultural Organization
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
USB	Universal Serial Bus
UVC	Universal Virtual Computer
ÚVT	Ústav výpočetní techniky
4C	Collaboration to Clarify the Costs of Curation

Seznam obrázků

Obr. 1 Základní pilíře designu podle informační architektury.....	35
Obr. 2 Hlavní komponenty digitální knihovny podle Kahn-Wilensky.....	42
Obr. 3 Digitální objekt jako základní prvek architektury digitální knihovny v modelu Kahn-Wilensky.....	43
Obr. 4 Konceptuální model digitální knihovny jako třívrstvá konstrukce	45
Obr. 5 Hlavní koncepty univerza digitální knihovny	47
Obr. 6 Hlavní role aktérů ve třívrstvé konstrukci univerza digitální knihovny	48
Obr. 7 Hierarchie pohledů uživatelů.....	49
Obr. 8 Univerzum digitální knihovny	50
Obr. 9 Konceptuální mapa hierarchie domén digitální knihovny.....	51
Obr. 10 OAIS v kontextu producentů, managementu a uživatelů.....	55
Obr. 11 Schéma získání informace (informačního objektu) z dat, resp. datového objektu.....	56
Obr. 12 Informační balíček – koncept a vztahy.....	57
Obr. 13 Detailní schéma archivačního balíčku (AIP)	59
Obr. 14 Interakce a datový tok externích entit modelu OAIS	61
Obr. 15 Grafické vyjádření funkčních skupin OAIS modelu.....	62
Obr. 16 Funkce modulu Příjem (<i>Ingest</i>)	64
Obr. 17 Funkce modulu Správa dat (<i>Data Management</i>)	66
Obr. 18 Funkce modulu Permanentní úložiště (<i>Archival Storage</i>)	67
Obr. 19 Funkce modulu Přístup (<i>Access</i>)	69
Obr. 20 Funkce modulu Administrace (<i>Administration</i>)	70
Obr. 21 Datové toky z/do modulu Administrace v archivu a externích entit modelu OAIS	71
Obr. 22 Funkce modulu Plánování dlouhodobé ochrany (<i>Preservation Planning</i>).....	74
Obr. 23 Zobrazení datového toku uvnitř modelu OAIS.....	77
Obr. 24 Model důvěryhodného digitálního repozitáře	82
Obr. 25 Životní cyklus/politika digitální ochrany	86
Obr. 26 Životní cyklus dlouhodobé ochrany podle DCC.....	88
Obr. 27 Plánovací cyklus PLATTER	91
Obr. 28 Příklad signifikantních vlastností	95
Obr. 29 Postup analýzy užiteků (<i>Utility Analysis</i>)	96
Obr. 30 Postup plánování dlouhodobé ochrany v nástroji PLATO	99
Obr. 31 „Upload“ ukázkových souborů do nástroje PLATO Preservation Planning Tool.....	101

Obr. 32 Příklad „stromu“ cílů pro webovou stránku	102
Obr. 33 Ukázka analýzy výsledků v nástroji PLATO	104
Obr. 34 Architektura systému Rosetta	107
Obr. 35 Architektura systému SDB	108
Obr. 36 Architektura systému RODA.....	111
Obr. 37 Vývoj standardů pro digitální repozitáře	116
Obr. 38 Atributy rizik.....	125
Obr. 39 Struktura hodnocení repozitáře pomocí nástroje DRAMBORA.....	126
Obr. 40 Integrace LIFE3 do dalších nástrojů	133
Obr. 41 Screenshot z webového nástroje LIFE.....	134
Obr. 42 Pracovní skupiny (<i>Work Packages</i>) modelu nákladů 4C	137

Seznam tabulek

Tab. 1 Požadavky a signifikantní vlastnosti textového dokumentu	93
Tab. 2 Ochranná opatření institucionální roviny	115
Tab. 3 Hodnocení shody se směrnicemi Data Seal of Approval	119
Tab. 4 Stručný přehled modelů nákladů	130
Tab. 5 Komponenty životního cyklu modelu LIFE	133
Tab. 6 Seznam partnerských institucí projektu 4C	136

Seznam citované literatury

Seznam citované literatury je řazen abecedně dle prvního údaje v záznamu (záhlaví), seznam není číslován. Pro lepší přehlednost je v seznamu literatury každé záhlaví zvýrazněno. Jednotlivé záznamy jsou v souladu s pravidly uvedenými v normě ČSN ISO 690 (01 0197) platné od 1. dubna 2011. Výjimkou je uvádění roku vydání, který je uveden oproti normě dvakrát, jednou za záhlavím a jednou za údaji o vydavateli, a to z důvodu rychlejší orientace v seznamu.

ALCTS Preservation and Reformatting Section, Working Group on Defining Digital Preservation. 2007. Definitions of Digital Preservation [online]. Washington, D.C.: ALA Annual Conference, June 2007 [cit. 20013-12-10]. Dostupné z: <http://www.ala.org/alcts/resources/preserv/defdigpres0408>

AMERICAN LIBRARY ASSOCIATION. 2007. Fostering media diversity in libraries strategies and actions [online]. American Library Association, June 2007 [cit. 2009-09-05]. Dostupné z: http://www.ala.org/ala/aboutala/offices/oif/ifissues/fostering_media_dive1.pdf

BARTOŠEK, Miroslav. 2004. Digitální knihovny – teorie a praxe. *Národní knihovna*, 2004, 15(4), s. 233-254. ISSN 0862-7487. Dostupné také z: <http://knihovna.nkp.cz/NKKR0404/0404233.html>

BRATKOVÁ, Eva. 2009. *Digitální knihovny s volným přístupem v oblasti vědy a výzkumu a identifikace a metadatový popis jejich objektů* [online]. Praha: Univerzita Karlova v Praze, 2009. 314 s. [Vedoucí práce Rudolf Vlasák. [cit. 2014-04-17]. Dostupné z: http://is.cuni.cz/webapps/zzp/search/?lang=cs&tab_searchas=basic

BROWN, A. 2008. *White paper: Representation information registries* [online]. 29th January 2008 [cit. 2014-04-18]. Dostupné z: http://www.planets-project.eu/docs/reports/Planets_PC3-D7_ReplInformationRegistries.pdf

CANDELA, L. et al. 2007a. Setting the Foundations of Digital Libraries. *DLib Magazine* [online]. Vol 13, no 3/4. [cit. 2012-11-08]. ISSN 1082-9873. Dostupné z: <http://www.dlib.org/dlib/march07/castelli/03castelli.html>

CANDELA, L. et al. 2007b. Manifest digitálnych knižníc. *Knižnica* [online]. 2007, 8(1) [cit. 2013-07-08]. Dostupné z: http://www.snk.sk/swift_data/source/casopis_kniznica/2007/januar/07.pdf

COLLISON, C. a PARCEL, G. 2005. *Knowledge management*. Brno : Computer Press, 2005. 236 s. ISBN 80-251-0760-4.

CUBR, Ladislav. 2010. *Dlouhodobá ochrana digitálních dokumentů*. 1. vyd. Praha: Národní knihovna České republiky, 2010. 154 s. ISBN 978-80-7050-588-5.

- DATA SEAL OF APPROVAL BOARD. 2010.** *Data Seal of Approval : Guidelines version 1* [online]. June 1, 2010 [cit. 2014-04-15]. Dostupné z: http://www.datasealofapproval.org/media/filer_public/2013/09/27/guidelines_01-june-2010.pdf
- DATA SEAL OF APPROVAL BOARD. 2013.** *Data Seal of Approval : Guidelines version 2* [online]. July 19, 2013 [cit. 2014-04-15]. Dostupné z: http://www.datasealofapproval.org/media/filer_public/2013/09/27/guidelines_2014-2015.pdf
- DCC. cca2012.** *What is digital curation?* [online]. cca2012. [cit. 2014-04-17]. Dostupné z: <http://www.dcc.ac.uk/digital-curation/what-digital-curation>
- DIGITAL PRESERVATION COALITION. 2008.** *Preservation Management of Digital Materials: The Handbook* [online]. November, 2008 [cit. 2014-03-12]. Dostupné z: <http://www.dpconline.org/advice/preservationhandbook>
- DOBRATZ, Susanne, SCHOGER, Astrid a STRATHMANN, Stefan. 2007.** The nestor Catalogue of Criteria for Trusted Digital Repository Evaluation and Certification. *Journal of Digital Information* [online]. Volume 8, Number 1, 2007 [cit. 20012-03-17]. Dostupné z: <http://www.informatik.uni-trier.de/~ley/db/journals/jodi/jodi8.html> - DobratzSS07
- DRYDEN, J. 2009.** The Open Archival Information System Reference Model. *Journal of Archival organization* [online]. Vol. 7, Iss. 4, 2009. [cit. 2014-04-07]. ISSN 1533-2756, DOI: 10.1080/15332740903334116. Dostupné z: <http://www.tandfonline.com/doi/abs/10.1080/15332740903334116?journalCode=wjao20-.U2s4II5E2KU>
- ECO, Umberto. 1997.** *Jak napsat diplomovou práci.* Olomouc: Votobia, 1997. ISBN 80-7198-173-7.
- EVANS, M. 2011.** *SDB: A Cross Industry, Flexible and Scalable Preservation Architecture* [online]. Oracle PASIG, 11th May 2011 [cit. 2013-12-12]. Dostupné z: <http://lib.stanford.edu/files/pasig-may-2011/Evans-PASIG-May-2011-Tessella.pdf>
- FOJTŮ, Andrea. 2010.** *Open source nástroje pro dlouhodobou ochranu digitálních dat: Podklad pro závěrečnou zprávu grantu Výzkum a vývoj.* Praha: Národní knihovna ČR, prosinec 2010.
- FOJTŮ, Andrea. 2011.** *Plánování dlouhodobé ochrany pomocí nástroje PLATO Preservation Planning Tool: Podklad pro závěrečnou zprávu grantu Výzkum a vývoj.* Praha: Národní knihovna ČR, září 2011.
- FOJTŮ, Andrea. 2013.** *Osobní poznámky z konference iPres 2013.* iPRES 2013 - 10th International Conference on Preservation of Digital Objects (Lisabon, Portugalsko).
- FRUCHT, B., et al. cca2009.** *nestor Kriterien Katalog vertrauenswürdige digitale Langzeitarchive* [online]. cca2009 [cit. 2013-05-30]. Dostupné z: <http://www.wirtschaftsarchiv.de/arbeitskreise/fachliche-arbeitskreise/elektronische-archivierung/NestorKriterienkatalog2.pdf>

- HODGE, G. 2000.** *Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files.* (s.n.) : Digital Library Federation, 2000. 43 p. ISBN 1-887334-76-9. Dostupný i z: <http://www.clir.org/pubs/reports/pub91/contents.html>
- HOLE, Brian. 2010.** *LIFE3: Predictive Costing of Digital Preservation* [online prezentace]. July 5th 2010 [cit 2013-08-15]. Dostupné z: http://www.life.ac.uk/3/docs/Hole_pasig_v1.pdf
- HOWLETT, Allison. 2011.** *Information Architecture and Digital Libraries (Part 1)* [online]. June 11, 2011 [cit. 2013-08-26]. Dostupné z: <http://acrystelle.com/2011/06/11/information-architecture-and-digital-libraries-part-1/>
- HUTAŘ, Jan 2008.** Proč jsou české digitální repozitáře nespolehlivé? *Knihovna* [online]. 2008, roč. 19, č. 1, s. 39-53 [cit. 2013-08-12]. ISSN 1801-3252. Dostupné z: <http://knihovna.nkp.cz/knihovna82/82039.htm>
- INFORMATION ARCHITECTURE INSTITUTE. cca2007.** *Our Mission* [online]. cca2007 [cit. 2014-04-17]. Dostupné z: http://iainstitute.org/en/about/our_mission.php
- ISO 14721:2003.** *Space data and information transfer systems -- Open archival information system -- Reference model.* Geneva: International Organization for Standardization, 2003.
- ISO 9241. 2013.** In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 15 October 2013 [cit. 2014-04-21]. Dostupné z: http://en.wikipedia.org/wiki/ISO_9241
- JONÁK, Z., PÁLKA, P. a SCHWARZ, J. 1999.** Potřeby a možnosti mezinárodní spolupráce v oblasti pořádání znalostí. *Ikaros* [online]. 1999, roč. 3, č. 1. [cit. 2009-09-05]. URN-NBN:cz-ik255. ISSN 1212-5075. Dostupné z: <http://www.ikaros.cz/node/255>
- KAHN, Robert a WILENSKY, Robert. 1995.** *A Framework for Distributed Digital Object Services* [online]. May 13, 1995 [cit. 2014-04-17]. Dostupné z: <http://www.cnri.reston.va.us/k-w.html>
- KASTEN, J. 2007.** Thoughts on the Relationship of Knowledge Organization to Knowledge Management. *Knowledge Organization* [online]. January 2007 [cit. 2009-09-06]. Dostupný z Library, Information Science & Technology Abstracts with Full Text, EBSCOhost.
- KATUŠČÁK, Dušan, MATTHAEIDESOVÁ, Marta a NOVÁKOVÁ, Marta. 1998.** *Informačná výchova: terminologický a výkladový slovník: odbor knižničná a informačná veda.* 1. vyd. Bratislava: Slovenské pedagogické nakladateľstvo, 1998. 375 s. Edícia terminologických a výkladových slovníkov, Zv. 6. ISBN 80-08-02818-1.
- KAUR, Kirnn et al. 2013.** *D32.1 Report on cost parameters for Digital Repositories* [online]. Aparsen, 2013-02-28 [cit. 2014-04-15]. Dostupné z: http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2013/03/APARSEN-REP-D32_1-01-1_0.pdf

- KEENAN, Stella a JOHNSTON, Colin. 2000.** *Concise Dictionary of Library and Information Science*. 2nd ed. London: Bowker-Saur, 2000. 265 s. ISBN 1-85739-251-5.
- KEJSER, Ulla Bøgvad, Anders Bo NIELSEN a Alex THIRIFAYS. 2011.** Cost Model for Digital Preservation: Cost of Digital Migration. *The International Journal of Digital Curation*. [online]. Vol. 6, No. 1, pp. 255-267, 2011 [cit. 2014-04-21]. ISSN 1746-8256. DOI: doi:10.2218/ijdc.v6i1.186. Dostupné z: <http://www.ijdc.net/index.php/ijdc/article/view/177>
- KLING, Rob a ELLIOTT, Margaret. 1994.** Digital library design for organizational usability. In: *ACM SIGOIS Bulletin* [online]. 1994 [cit. 2014-04-21]. DOI: 10.1145/192611.192746. Dostupné z: <http://dl.acm.org/citation.cfm?id=192746>
- KNIGHT, G. a PENNOCK, M. 2009.** Data Without Meaning: Establishing the Significant Properties of Digital Research. *The International Journal of Digital Curation* [online]. Vol 4, No 1 (2009) [cit. 2011-05-22]. ISSN: 1746-8256. Dostupné z: <http://www.ijdc.net/index.php/ijdc/article/view/110>
- KORENKOVA, M. a HÄGERFORS, A. 2011.** *Archiving 2011: Quality criteria for digital information in long-term preservation*. Salt Lake City (UT): Society for Imaging Science and Technology, 2011. 232 p. ISBN / ISSN: 978-0-89208-294-0.
- LAVOIE, B. 2000.** Meeting the challenges of digital preservation: The OAIS reference model. In: *OCLC Newsletter* [online]. No. 243:26-30 (January/February 2000) [cit.2013-07-12]. Dostupné z: <http://digitalarchive.oclc.org/da/ViewObject.jsp?objid=0000001747>
- LIFE. 2014.** [online]. 2014 [cit. 2014-04-21]. Dostupné z: <http://www.life.ac.uk/3/>
- MACKENZIE, Matthew C., LASKEY, Ken a MCCABE, Francis, et al. 2006.** *Reference Model for Service Oriented Architecture. Committee Specification* [online]. 1, 2 August 2006 [cit. 2013-07-22]. Dostupné z: <https://www.oasis-open.org/committees/download.php/19679/soa-rm-cs.pdf>
- MAKULOVÁ, Soňa. 2005.** Informačná architektúra. *Ikaros* [online]. 2005, roč. 9, č. 9 [cit. 2013-08-26]. urn:nbn:cz:ik-002007. ISSN 1212-5075. Dostupné z: <http://www.ikaros.cz/node/2007>
- MAKULOVÁ, Soňa. 2007.** Použitelnosť webových sídiel ako základný predpoklad ich úspešnosti. In: *ELET web solutions* [online]. 2007 [cit. 2014-04-21]. Dostupné z: <http://www.elet.sk/?pouzitelnost-pristupnost&sprava=pouzitelnost-webovych-sidiel-ako-zakladny-predpoklad-ich-uspesnosti>
- MARCHIONINI, Gary. 2007.** "Information Architecture." [online video]. YouTube, 13 Dec. 2007 [cit. 2013-02-01]. Dostupné z: http://www.youtube.com/watch?v=W0VYRev7_bQ
- NESTOR. 2008.** *nestor - Materialien 8: nestor-Kriterien, Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, Version 2* [online]. Frankfurt am Main : nestor c/o Deutsche Nationalbibliothek, 2008 [cit. 2013-05-30]. urn:nbn:de:0008-2008021802. Dostupné z: <http://edoc.hu-berlin.de/series/nestor-materialien/8/PDF/8.pdf>

NETWORK OF EXCELLENCE ON DIGITAL LIBRARIES. 2007. *The DELOS Digital Library Reference Model: Foundations for Digital Libraries, [version 0.98, December 2007]*. Pisa: GRUPPO ALI, 2007. ISBN 291233537X, 9782912335371.

NEUMEYER, R. a RAUBER, A. 2007. *Why appraisal is not 'utterly' useless and why it's not the way to go either: A provocative position paper (PPP)* [online]. 12th November 2007 [cit. 2011-05-22]. Dostupné z: http://www.digitalpreservationeurope.eu/publications/position/appraisal_final.pdf

OCLC/RLG PREMIS WORKING GROUP. 2008. *PREMIS Data Dictionary for Preservation Metadata version 2.0* [online]. Dublin (OH): OCLC, March 2012 [cit. 2011-08-21]. 237 s. Dostupné z: <http://www.loc.gov/standards/premis/v2/premis-2-0.pdf>

Partners. 2013. 4C – Collaboration to Clarify the Costs of Curation [online]. 2013 [cit. 2014-04-21]. Dostupné z: <http://www.4cproject.eu/about-us/partners>

PINKAS, Otakar. 1999. Digitální knihovny, jejich organizace a služby. *Knihovny současnosti '99*. Brno: Sdružení knihoven, 1999. s. 100-105. ISBN 80-86249-04-2.

Požadavky na producenty dat vkládaných do repozitáře. 2013. DigiTool [online]. 2013 [cit. 2014-05-08]. Dostupné z: <http://digitool.cuni.cz/DIGITool-73.html>

PROM, Christopher. 2010. *PLATO (Digital Preservation Planning) Software Review* [online]. April 25, 2010 [cit. 2011-05-22]. Dostupné z: <http://e-records.chrisprom.com/?p=108>

PURDAY, J. 2010. *Europeana v1.0 : Annual Report : 1 February 2009 – 31 January 2010 (Grant Agreement Number: 55800)* [online]. 26th February 2010 [cit. 2010-08-01]. Dostupné z: http://version1.europeana.eu/c/document_library/get_file?uuid=df49b382-aa16-497f-8405-948b532dc0d7&groupId=10602

REITZ, Joan M. 2004. *Dictionary for library and information science*. Westport: Libraries Unlimited, 2004. 788 s. ISBN 1-59158-075-7.

RESEARCH LIBRARIES GROUP. 2002. *Trusted Digital Repositories: Attributes and Responsibilities: An RLG-OCLC Report* [online]. Mountain View (CA): RLG, 2002 [cit. 2011-07-25]. Dostupné z: <http://www.oclc.org/research/activities/past/rlg/trustedrep/repositories.pdf>

RESSLER, Miroslav, ed. 2006. *Informační věda a knihovnictví: výkladový slovník české terminologie z oblasti informační vědy a knihovnictví: výběr z hesel v databázi TDKIV*. Vyd. 1. Praha: Vysoká škola chemicko-technologická, 2006. 161 s. ISBN 80-7080-599-4.

ROSENFELD, Louis a MORVILLE, Peter. 2002. *Information architecture for the World Wide Web*. 2nd ed. [Sebastopol]: O'Reilly, 2002. 461 s. ISBN 0-596-00035-9.

ROSENTHAL, Colin a kol. 2009. *Průvodce plánem důvěryhodného digitálního repozitáře (PLATTER)*. 1. vyd. Praha: Národní knihovna České republiky, 2009. 51 s. ISBN 978-80-7050-569-4.

RRA VYSOČINA. cca2010. *Ekonomická analýza* [online]. cca2010 [cit. 2011-05-22]. Dostupné z: http://www.rrav.cz/sluzby/ekonom_analyzy.html

SMART metoda. 2014. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 18.4.2014 [cit. 2014-04-21]. Dostupné z: http://cs.wikipedia.org/wiki/SMART_metoda

SOERGEL, D. 2009. *Osobní poznámky z workshopu KOS in Digital Libraries*. ECDL 2009. September, 2009.

ŠVÁSTOVÁ, Pavla. 2014. *Digital Curation* [online prezentace]. Brno : Moravská zemská knihovna, 2014 [cit. 2014-21-05]. Dostupné z: <http://www.ics.muni.cz/mba/eiz/eiz11c-6.pdf>

Usability. 2014. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2014 [cit. 2014-04-21]. Dostupné z: <http://en.wikipedia.org/wiki/Usability>

User interface design. 2014. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2014 [cit. 2014-04-21]. Dostupné z: http://en.wikipedia.org/wiki/User_interface_design

VLDL 2011: Fourth Workshop on Very Large Digital Libraries 2011. 2011. DELOS Network of Excellence on Digital Libraries [online]. 2011 [cit. 2014-06-07]. Dostupné z: <http://www.delos.info/vldl2011/>

VORLIČKOVÁ, B. 2009. Světová digitální knihovna otevírá pokladnici kulturního bohatství lidstva na internetu. *Ikaros* [online]. 2009, roč. 13, č. 5 [cit. 2010-08-26]. URN-NBN:cz-ik5428. ISSN 1212-5075. Dostupné z : <http://www.ikaros.cz/node/5428>

WALLACE, M.a WEBBER, L. 2004. *Disaster Recovery Handbook*. New York: AMACOM Books, 2004. 416 p. ISBN: 9780814472408.

WILSON, A. 2007. *Significant Properties Report* [online]. InSPECT Work Package 2.2, 10/04/2007 [cit. 2014-04-18]. Dostupné z: http://www.significantproperties.org.uk/wp22_significant_properties.pdf

WILSON, A. 2008. *Significant properties of Digital Objects* [online]. 7 April 2008 [cit. 2011-05-22]. Dostupné z: https://www.google.cz/search?client=safari&rls=en&q=wilson+significant+propersties&ie=UTF-8&oe=UTF-8&gfe_rd=cr&ei=sOxUU57SF-Tb8geAw4CoDg

WIJNGAARDEN, H. 2004. *OAIS and the e-Depot*. [online]. 2004 [cit. 2013-11-22]. Dostupné z: http://www.ndl.go.jp/en/publication/ndl_newsletter/136/5hvw_136_361.pdf

WITTEN, I. H., BAINBRIDGE, D. a NICHOLS, D. M. 2010. *How to build a digital library*. 2nd ed. Amsterdam: Morgan Kaufmann Publ., 2010. xxiii, 629 s. The Morgan Kaufmann series in multimedia information and systems. ISBN 978-0-12-374857-7.

Seznam studované literatury

Seznam studované literatury je řazen abecedně dle prvního údaje v záznamu (záhlaví), seznam není číslován. Pro lepší přehlednost je v seznamu literatury každé záhlaví zvýrazněno. Jednotlivé záznamy jsou v souladu s pravidly uvedenými v normě ČSN ISO 690 (01 0197) platné od 1. dubna 2011.

ABID, Abdelaziz. Safeguarding our digital heritage: a new preservation paradigm. In: LUSENET, Yola de a Vincent WINTERMANS (eds.). *Preserving the digital heritage: principles and policies*. Amsterdam: Netherlands National Commission for UNESCO, European Commission on Preservation and Access, 2007, s. 7-14. ISBN 978-90-6984-523-4. Dostupné také z: <http://www.ica.org/5697/paag-resources/preserving-the-digital-heritage-principles-and-policies.html>

BONIN, Sonja. *Preservation and Long-term Access via NETWORKED Services: keeping digital information alive for the future* [online]. PLANETS Consortium, 2009 [cit. 2011-12-04]. 20 s. Dostupné z: http://www.planets-project.eu/docs/comms/PLANETS_BROCHURE.pdf

BORBINHA, José (ed.), et al. *iPRES 2011: 8th International Conference on Preservation of Digital Objects, 1.-4.11.2011, Singapore*. Singapore: National Library Board Singapore & Nanyang Technology University, 2011. 287 s. ISBN 978-981-07-0441-4. Dostupné také z: <http://getfile3.posterous.com/getfile/files.posterous.com/temp-2012-01-02/dHqmzjcCGoexvmiBzJDCyhrhlgswoffzvsnfpEAXjHFEesarvwahEHrmyvj/iPRES2011.proceedings.pdf>

CAPLAN, Priscilla a Carol C.H. CHOU. DAITSS Grows Up: Migrating to a Second Generation Preservation System. In: *Archiving 2011, May 16-19, 2011, Salt Lake City Utah: Final Program and Proceedings*. Springfield (VA): Society for Imaging Science and Technology, 2011, s. 101-104. ISBN 978-0-89208-294-0.

CARPENTER, Leona. *Repositories in Context: Digital Repositories as components of an integrated infrastructure for education* [online prezentace]. 2005 [cit 2011-04-15]. Dostupné z: <http://www.ukoln.ac.uk/events/delos-rep-workshop/presentations/carpenter.ppt>

COLLISON, C., PARCEL, G. *Knowledge management*. Brno : Computer Press, 2005. 236 s. ISBN 80-251-0760-4.

CONCORDIA, C., GRADMANN, S. Not just another portal, not just another digital library: A portrait of Europeana as an application program interface. *IFLA Journal* [online]. March 2010, vol. 36, no. 1, p. 61-69 [cit. 2010-07-29]. Dostupné v SAGE journals online. DOI 10.1177/0340035209360764.

COUSINS, J., CHAMBERS, S., MEULEN, E. van der. Uncovering cultural heritage through collaboration. *International Journal on Digital Libraries* [online]. November 2008, vol. 9, no. 2 [cit.

2010-08-10]. ISSN 1432-5012. Dostupné z: <http://www.springerlink.com/content/q1067151j02j248n/>

DAPPERT, Angela. *Digital Preservation Metadata* [online prezentace]. WePreserve, March 2009 [cit. 2011-10-20]. 122 snímků (ve formátu PPT). Dostupné z: <http://www.slideshare.net/DigitalPreservationEurope/preservation-metadata-1258027>

DAVENPORT, T.H., PRUSAK, L. *Working Knowledge : How organizations manage what they know.* Boston (Ma.) : Harvard Business School Press, 2000. 199 p. ISBN 1-57851-301-4.

EUROPEAN COMMISSION. Commission unveils plans for European digital libraries (IP/05/1202). In: *Europa Press Releases RAPID* [online]. Brussels: [European Commission], 30 September 2005 [cit. 2011-05-01]. Dostupné z: <http://europa.eu/rapid/pressReleasesAction.do?reference=IP/05/1202&format=HTML&aged=0&language=en&guiLanguage=en>

FOJTŮ, Andrea, Jan HUTAŘ a Marek MELICHAR. Dlouhodobá ochrana digitálních dokumentů a projekt NDK. In: *Knihovny současnosti 2011: Sborník z 19. konference, konané ve dnech 13.-15. září 2011 v Českých Budějovicích.* Ostrava: Sdružení knihoven ČR, 2011, s. 73-79. ISBN 978-80-86249-62-9. Dostupné také z: http://www.svkos.cz/data/xinha/sdruk/ks2011/sbornik_2011.pdf

GARRETT, John a Donald WATERS. *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information* [online]. The Commission on Preservation and Access and The Research Libraries Group, 1996 [cit. 2011-11-17]. 64 s. Dostupné z: <http://www.clir.org/pubs/reports/pub63watersgarrett.pdf>

GAVRILIS, Dimitris, Christos PAPATHEODOROU, Panos CONSTANTOPOULOS a Stavros ANGELIS. Mopseus - A Digital Library Management System Focused on Preservation. In: LALMAS, Mounia, et al. (Eds.). *Research and Advanced Technology for Digital Libraries. 14th European Conference, ECDL 2010 Glasgow, UK, September 6-10, 2010.* Berlin: Springer, 2010, s. 445-448. Lecture Notes in Computer Science, Vol. 6273. ISSN 0302-9743. ISBN 3-642-15463-8.

GLADNEY, Henry M. *Preserving Digital Information.* Berlin: Springer, 2007. xxxiii, 314 s. ISBN 978-3-540-37886-0.

GRADMANN, S. *Knowledge = Information in Context : on the Importance of Semantic Contextualisation in Europeana* [online]. 2010 [cit. 2010-08-10]. Dostupné z: http://version1.europeana.eu/c/document_library/get_file?uuid=cb417911-1ee0-473b-8840-bd7c6e9c93ae&groupId=10602

HATHI TRUST. HathiTrust Certified as Trustworthy Repository. In: *HathiTrust* [online]. 30. March 2011 [cit. 2011-04-16]. Dostupné z: <http://www.hathitrust.org/hathitrust-certified-as-trustworthy-repository>

HILL, L., et al. *Integration of Knowledge Organization Systems into Digital Library Architectures:*

Position Paper for 13th ASIS&T SIG/CR Workshop, "Reconceptualizing Classification research" [online]. 2002 [cit. 2009-09-06]. Dostupné z: <http://www.alexandria.ucsb.edu/~gjanee/archive/2002/kos-dl-paper.pdf>

HJØRLAND, B. Fundamentals of Knowledge Organization. *Knowledge Organization*. ISSN 0943-7444, 2003, vol. 30, no. 2, p. 87-111.

HM TREASURY. *The Orange Book: Management of Risk - Principles and Concepts*. London: HM Treasury, 2004. 50 s. ISBN 1-84532-044-1. Dostupné také z: http://www.hm-treasury.gov.uk/d/orange_book.pdf

HUTAŘ, Jan, Andrea FOJTŮ a Eliška PAVLÁSKOVÁ. DRAMBORA - nástroj na interní audit digitálních úložišť v nové online verzi a postřehy z provedených auditů. In: *Inforum 2008, Praha 28.-30.5.2008* [online]. Praha: Albertina lcome, 2008 [cit. 2011-11-14]. 9 s. ISSN 1801-2213. Dostupné z: <http://www.inforum.cz/pdf/2008/hutar-jan-cze.pdf>

HUTAŘ, Jan a Colin ROSENTHAL. Practical ways to tackle digital preservation using DPE tools and services. In: *WePreserve Forum, Friday 17.10 Prague* [online]. 2008 [cit. 2011-11-03]. Dostupné z: http://www.digitalpreservationeurope.eu/platter/platter_presentation_prague.pdf

KAN, M.Y., LEE, D., LIM, E.P. Scholarly digital libraries at scale: introduction to the special issue on very large digital libraries. 2008. *International Journal on Digital Libraries* [online]. October 2008, vol. 9, no. 1 [cit. 2010-08-10]. ISSN :1432-5012. Dostupný z: <http://www.springerlink.com/content/ r2580kx1jp2t7460/>

KIRCHHOFF, Amy, et al. Becoming a certified trustworthy digital repository: the Portico experience. In: RAUBER, Andreas, et al. (eds.). *IPRES 2010: Proceedings of the 7th International Conference on Preservation of Digital Objects, Vienna, Austria, September 19-24, 2010*. Vienna: Oesterreichische Computer Gesellschaft, 2010, s. 87-93. ISBN 978-3-85403-262-5.

KUNIAVSKY, Mike. *Observing the user experience: a practitioner's guide to user research*. San Francisco, CA : Morgan Kaufmann Publishers, c2003. xvi, 560 s. ISBN 1-55860-923-7.

LAVOIE, Brian a Lorcan DEMPSEY. Thirteen Ways of Looking at... Digital Preservation. *D-Lib Magazine* [online]. July/August 2004, Volume 10, Number 7/8 [cit. 2011-12-29]. ISSN 1082-9873. DOI 10.1045/july2004-lavoie. Dostupné z: <http://www.dlib.org/dlib/july04/lavoie/07lavoie.html>

LAVOIE, Brian a Richard GARTNER. *Preservation metadata: Technology Watch Report* [online]. [Dublin (OH)]: OCLC, DPC, Oxford University Library Services, 2005 [cit. 2011-08-14]. 21 s. DPC Technology Watch Series Report 05-01. Dostupné z: <http://www.dpconline.org/docs/reports/dpctw05-01.pdf>

LAWRENCE, Gregory W., et al. *Risk Management of Digital Information: a File Format Investigation*. Washington (DC): Council on Library and Information Resources, 2000. 75 s. ISBN 1-887334-78-5. Dostupné také z: <http://www.clir.org/pubs/reports/pub93/pub93.pdf>

LEVACK, Kinley. Europe Gets Serious About Cultural Collection. *EContent* [online]. November 2009, 11 [cit. 2010-08-10]. ISSN 1525-2531.

Dostupné z: <http://www.econtentmag.com/Articles/News/News- Feature/Europe-Gets-Serious-About-Cultural- Collection-57724.htm>

LHOTÁK, Martin. Open source pro digitální knihovnu. In: *Automatizace knihovnických procesů – 11: sborník z 11. ročníku semináře pořádaného ve dnech 16.–17. května 2007 v Liberci*. Praha: ČVUT, 2007, s. 65-72. ISBN 978-80-01-03691-4. Dostupné také z: <http://www.akvs.cz/akp-2007/09-lhotak.pdf>

NÁRODNÍ TECHNICKÁ KNIHOVNA. *Audit důvěryhodnosti Digitálního repozitáře NUŠL* [online]. 8. únor 2011 [cit. 2011-05-01]. Dostupné z: <http://nysl.techlib.cz/index.php/Audit>

PALFREY, John G. a Urs GASSER. *Born digital: understanding the first generation of digital natives*. New York: Basic Books, 2008. 288 s. ISBN 9780465005154.

PREECE, Jenny; ROGERS, Yvonne, aj. *Human-computer interaction*. Wokingham, England; Reading, Mass. : Addison-Wesley Pub. Co., 1995. xxxviii, 775 s. ISBN 0-201-62769-8.

PREMIS EDITORIAL COMMITTEE. *PREMIS Data Dictionary for Preservation Metadata* [online]. Version 2.1. Washington: PREMIS Editorial Committee, January 2011 [cit. 2011-10-14]. 226 s. Dostupné z: <http://www.loc.gov/standards/premis/v2/premis-2-1.pdf>

ROSS, Seamus. *Changing Trains at Wigan: Digital Preservation and the Future of Scholarship*. London: National Preservation Office, 2000. 44 s. Dostupné také z: <http://www.bl.uk/blpac/pdf/wigan.pdf>

ROSS, Seamus, et al. *An Introduction to the DRAMBORA toolkit and its Underlying Principles* [online prezentace]. Barcelona, 26. 3. 2009 [cit. 2011-04-17]. 116 snímků (ve formátu PPT). Dostupné z: http://www.digitalpreservationeurope.eu/preservation-training-materials/files/DRAMBORA_barcelona.ppt

ROTHENBERG, Jeff. *Avoiding technological quicksand: finding a viable technical foundation for digital preservation*. Washington (DC): Council on Library and Information Resources, 1999. 35 s. ISBN 1-887334-63-7. Dostupné také z: <http://www.clir.org/pubs/reports/rothenberg/pub77.pdf>

RUSSELL, Kelly. EDARS: Long-term Access and Usability of Digital Resources: the Digital Preservation Conundrum. *Ariadne* [online]. December 1998, no. 18 [cit. 2011-09-16]. ISSN 1361-3200. Dostupné z: <http://www.ariadne.ac.uk/issue18/cedars/>

RUSSELL, Kelly, et al. *Metadata For Digital Preservation: The Cedars Project Outline Specification* [online]. CEDARS, March 2000 [cit. 2011-09-24]. 33 s. Dostupné z: <http://web.archive.org/web/20010613073650/http://www.leeds.ac.uk/cedars/MD-STR~5.pdf>. Zdroj je dostupný pouze v Internet Archive.

SCHREIBMAN, Susan (ed.). *Best Practice Guidelines for Digital Collections* [online]. College Park (MD), May 2007 [cit. 2011-10-08]. 76 s. Dostupné z: http://www.lib.umd.edu/dcr/publications/best_practice.pdf

SCHULTZ, Matt a Emily B. GORE. The importance of trust in distributed digital preservation: a case study from the MetaArchive cooperative. In: RAUBER, Andreas, et al. (eds.). *IPRES 2010: Proceedings of the 7th International Conference on Preservation of Digital Objects, Vienna, Austria, September 19-24, 2010*. Vienna: Oesterreichische Computer Gesellschaft, 2010, s. 105-111. ISBN 978-3-85403-262-5.

SIEBINGA, S., PURDAY, J., WERF, B. Van der. *Guidelines for the use of EuropeanaLabs* [online]. Version 1 of 20 July 2009 [cit. 2010-07-30]. [file://localhost/Dostupné z/
http://version1.europeana.eu:c/document_library:get_file%3Fuuid=9f7ed5a7-fdaf-404a-b1a1-6a55b97b9b6a&groupid=10602](file://localhost/Dostupné%20z/http://version1.europeana.eu:c/document_library:get_file%3Fuuid=9f7ed5a7-fdaf-404a-b1a1-6a55b97b9b6a&groupid=10602)

SINCLAIR, Pauline a Amir BERNSTEIN. *An Emerging Market: Establishing Demand for Digital Preservation Tools and Services* [online]. PLANETS, 2010 [cit. 2011-12-09]. 10 s. Planets White Paper. Dostupné z: <http://www.planets-project.eu/docs/reports/Planets-VENDOR-White-Paperv4.pdf>

SMIRAGLIA, Richard P. Introducing Metadata. In: *Metadata: a cataloger's Primer*. Binghamton (NY): Haworth Information Press, 2005, s. 1-15. ISBN 978-0-7890-2801-3.

STEINEROVÁ, Jela. *Informačné správanie : pohľady informačnej vedy*. 1. vyd. Bratislava : Centrum vedecko-technických informácií SR, 2005. 199 s. ISBN 80-85165-90-2.

STOKLASOVÁ, B. Perspektivy důvěryhodného digitálního úložiště v rámci Národní digitální knihovny. *Knihovna* [online]. 2006, roč. 17, č. 2, s. 49-55 [cit. 2009-09-05]. ISSN 1801-3252. Dostupné z: <http://knihovna.nkp.cz/knihovna62/stoklas.htm>

STONE, Debbie, aj. *User interface design and evaluation*. San Francisco, CA. : Morgan Kaufmann, c2005. xxviii, 669 s. ISBN 0-12-088436-4.

STRODL, Stephan, Petar PETROV a Andreas RAUBER. *Research on Digital Preservation within projects co-funded by the European Union in the ICT programme* [online]. Vienna: Vienna University of Technology, 2011 [cit. 2012-01-30]. 51 s. Dostupné z: http://cordis.europa.eu/fp7/ict/telearn-digicult/report-research-digital-preservation_en.pdf

TONNINGEN, B.R. van. *The Knowledge Organization : Knowledge management = asset management* [online]. June 2000 [cit. 2009-09-06]. Dostupné z: <http://www.managementsite.com/236/The-Knowledge-Organization.aspx>

Trusted Digital Repository. In: *TrustedDigitalRepository.eu* [online]. TrustedDigitalRepository.eu, 2010 [cit. 2011-12-12]. Dostupné z:

<http://www.trusteddigitalrepository.eu/Site/Trusted%20Digital%20Repository.html>

VAN GARDEREN, Peter. Archivematica: using micro-services and open-source software to deliver a comprehensive digital curation solution. In: RAUBER, Andreas, et al. (eds.). *IPRES 2010: Proceedings of the 7th International Conference on Preservation of Digital Objects, Vienna, Austria, September 19-24, 2010*. Vienna: Oesterreichische Computer Gesellschaft, 2010, s. 145-150. ISBN 978-3-85403-262-5.

VERHEUL, Ingeborg. *Networking for Digital Preservation: Current Practice in 15 National Libraries*. München: Saur, 2006. 269 s. ISBN 3-598-21847-7. Dostupné také z:

<http://archive.ifla.org/VI/7/pub/IFLAPublication-No119.pdf>

VOJTÁŠEK, Filip. Dlouhodobá archivace digitálních dokumentů. *Ikaros* [online]. 2000, roč. 4, č. 10 [cit. 2011-09-10]. ISSN 1212-5075. Dostupné z: <http://www.ikaros.cz/node/675>

WALKER, Alison. Preservation. In: BOWMAN, J. H. (ed.). *British Librarianship and Information Work 2001-2005*. Hampshire: Ashgate Publishing Limited, 2007. 18 s. (501-518). ISSN 1752-556X. ISBN 978-0-7546-4778-2. Dostupné také z: <http://www.bl.uk/blpac/pdf/bliwb.pdf>

WATERS, Donald J. What Are Digital Libraries? *CLIR Issues* [online]. July/August 1998, Nr. 4 [cit. 2011-12-29]. Dostupné z: <http://www.clir.org/pubs/issues/issues04.html#dlf>

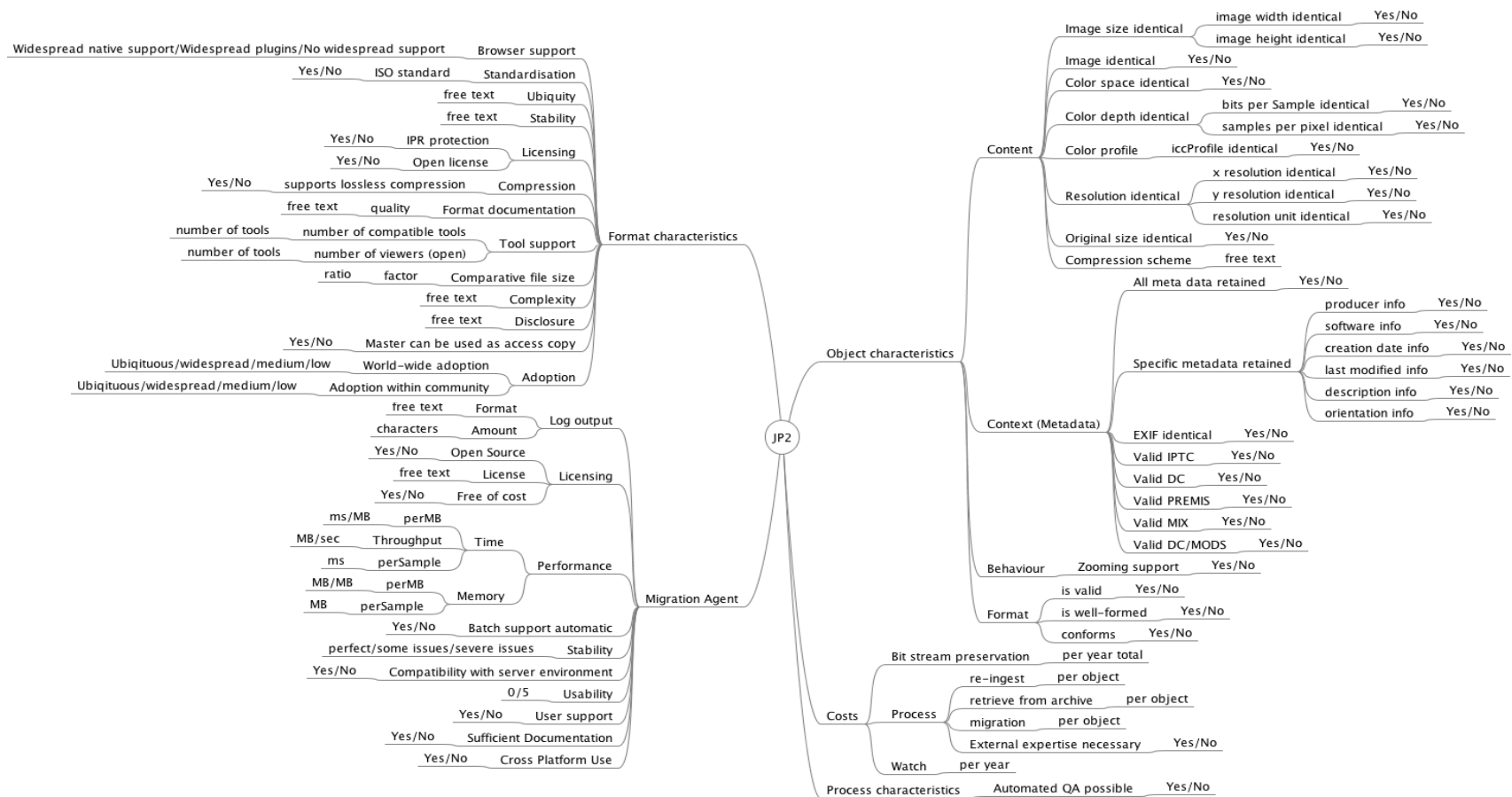
WEBB, Colin. Preservation Metadata for Digital Collections. In: *National Library of Australia* [online]. Canberra: National Library of Australia, 15. October 1999 [cit. 2011-09-23]. Dostupné z: <http://www.nla.gov.au/preserve/pmeta.html>

WEBB, Colin. *Guidelines for the Preservation of Digital Heritage* [online]. Paris: UNESCO, 2003 [cit. 2012-01-01]. 177 s. Dostupné z: <http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>

WILSON, Andrew. *Significant Properties Report: InSPECT Work Package 2.2* [online]. V2. London: AHDS, 10/04/2007 [cit. 2012-02-06]. 10 s. Dostupné z:

http://www.significantproperties.org.uk/wp22_significant_properties.pdf

Příloha č. 1 - Příklad myšlenkové mapy (formát JP2)



Příloha č. 2 - Přehled strategických plánů, odpovědností a cílů PLATTER

Strategický cíl	Odpovědnost	Cíl	Související cíle	
1. Finanční plán	Zabývá se plánováním, monitorováním a vykazováním financí.	1.1 Pravidelně sledovat a revidovat finanční plán	3.2 Získat a udržet zaměstnance pro práci na specifických pozicích	
			4.1 Formulovat, udržovat a aktualizovat programové prohlášení, které bude odpovídat mandátu repozitáře	
			5. 1 IT infrastruktura si musí umět poradit s takovým rozsahem ukládání, zpracování a přenosu dat, který odpovídá potřebám daného repozitáře	
		1.2 Udržet financování na úrovni, kterou vyžaduje běžný provoz repozitáře	1.4 Stanovit si a naplňovat marketingové plány a plány externí komunikace, které odpovídají potřebám daného repozitáře	
			1.3 Vytvořit nouzové plány pro případ finančních omezení nebo krizí, které dostatečně ochrání důležitá data	7.1 Ochrana digitálních dat by měla být zajištěna po dobu, která překračuje existence digitálního repozitáře
				8. 1 Digitální repozitář včas reaguje na podstatné změny v prostředí
1.4 Stanovit si a naplňovat marketingové plány a plány externí komunikace, které odpovídají potřebám daného repozitáře	2.1 Získat relevantní dokumenty			
	4.2 Definovat komunitu/y uživatelů repozitáře, rozumět jejich potřebám a dovednostem			
2. Akviziční plán	Zabývá se vztahy s depository a dalšími poskytovateli dat.	2.1 Získat relevantní dokumenty	4.2 Definovat komunitu/y uživatelů repozitáře, rozumět jejich potřebám a dovednostem	
		2.2 Vyjednat dohody o uložení	2.5 Zajišťovat aktuálnost smluv o uložení	
			1.1 Pravidelně sledovat a revidovat finanční plán	
			6.1 Určit jaké formáty digitálních objektů bude repozitář akceptovat (SIP)	
			6.1.1 Specifikovat zdroje a formát bibliografických a	

			popisných dat pro SIP
			6.1.2 Specifikovat technická metadata pro SIP
		2.3 Získat fyzickou kontrolu nad dokumenty	5. 3 Infrastruktura IT musí garantovat dostupnost daných služeb pro uživatele
		2.4 Monitorování akvizice	6.1 Určit, jaké formáty digitálních objektů bude repozitář akceptovat (SIP)
			6.2 Specifikovat formáty dat a obsah metadat pro archivaci digitálních objektů (AIP)
			6.3 Specifikovat formáty dat pro digitální objekty distribuované uživatelům (DIP)
		2.5 Zajišťovat aktuálnost smluv o uložení	
3. Plán řízení lidských zdrojů	Zabývá se získáváním a udržováním souboru dovedností, které jsou relevantní pro správu repozitáře.	3.1. Definovat zaměstnanecké pozice, odpovědnosti a pravomoci pracovníků repozitáře	
		3.2 Získat a udržet zaměstnance pro práci na specifických pozicích	
		3.3 Rozvíjet kvalifikaci zaměstnanců	
4. Plán zpřístupňování	Zabývá se vztahy s koncovými uživateli a pravidly pro zpřístupňování.	4.1 Formulovat, udržovat a aktualizovat programové prohlášení, které bude odpovídat mandátu repozitáře	1.1 Pravidelně sledovat a revidovat finanční plán
			9.3 Repozitář musí udržovat znalosti o komunitě uživatelů, jejich kompetencích a znalostní bázi
			1.4 Stanovit si a naplňovat marketingové plány a plány externí komunikace, které odpovídají potřebám daného repozitáře
		4.2 Definovat komunitu/y uživatelů repozitáře, rozumět jejich potřebám a dovednostem	5. 1 IT infrastruktura si musí umět poradit s takovým rozsahem ukládání, zpracování a přenosu dat, který odpovídá potřebám daného repozitáře
		5. 3 Infrastruktura IT musí garantovat dostupnost daných služeb pro uživatele	
		4.3 Formulovat a implementovat politiku zpřístupňování	5. 1 IT infrastruktura si musí umět poradit s takovým

		obsahu repozitáře	rozsahem ukládání, zpracování a přenosu dat, který odpovídá potřebám daného repozitáře
			6.2.1 Specifikovat metadata pro AIP
		4.4 Specifikovat a realizovat technologické požadavky distribuce a zpřístupňování	6.3.1 Specifikovat metadata pro DIP
			6.3 Specifikovat formáty dat pro digitální objekty distribuované uživatelům (DIP)
			6.5 Specifikovat transformaci AIP do DIP
			4.2 Definovat komunitu/y uživatelů repozitáře, rozumět jejich potřebám a dovednostem
			5. 3 Infrastruktura IT musí garantovat dostupnost daných služeb pro uživatele
5. Technický plán	Specifikuje požadavky na hardware, software a síťové systémy.	5. 1 IT infrastruktura si musí umět poradit s takovým rozsahem ukládání, zpracování a přenosu dat, který odpovídá potřebám daného repozitáře	1.1 Pravidelně sledovat a revidovat finanční plán
			3.3 Rozvíjet kvalifikaci zaměstnanců
		5.2 Infrastruktura IT musí garantovat integritu a bezpečnost uložených dat	4.3 Formulovat a implementovat politiku zpřístupňování obsahu repozitáře
			3.3 Rozvíjet kvalifikaci zaměstnanců
			9.5 Repozitář musí udržovat, provádět a hodnotit takové strategie dlouhodobé ochrany digitálních dat, které vyhovují konkrétním cílům dlouhodobé ochrany
		5. 3 Infrastruktura IT musí garantovat dostupnost daných služeb pro uživatele	9.1 Repozitář musí mít přehled o současném a vznikajícím hardware, software a technologiích ukládání dat
			3.3 Rozvíjet kvalifikaci zaměstnanců
6. Datový plán	Specifikuje datové a metadatové objekty, formát, struktury pro vkládání, uchovávání a	6.1 Určit jaké formáty digitálních objektů bude repozitář akceptovat (SIP)	2.2 Vyjednat dohody o uložení
		6.1.1 Specifikovat zdroje a formát bibliografických a popisných dat pro SIP	2.2 Vyjednat dohody o uložení
		6.1.2 Specifikovat technická metadata pro SIP	2.2 Vyjednat dohody o uložení
		6.2 Specifikovat formát dat a obsah metadat pro archivaci digitálních objektů (AIP)	9.1 Repozitář musí mít přehled o současném a vznikajícím hardware, software a technologiích ukládání dat

	zpřístupňování dat a související transformace a mapování dat.		9.2 Repozitář by měl udržovat srozumitelné informace o všech strukturálních standardech (tj. například kódování souborů) a formátech
			9.5 Repozitář musí udržovat, provádět a hodnotit takové strategie dlouhodobé ochrany digitálních dat, které vyhovují konkrétním cílům dlouhodobé ochrany
		6.2.1 Specifikovat metadata pro AIP	9.5 Repozitář musí udržovat, provádět a hodnotit takové strategie dlouhodobé ochrany digitálních dat, které vyhovují konkrétním cílům dlouhodobé ochrany
		6.3 Specifikovat formáty dat pro digitální objekty distribuované uživatelům (DIP)	4.2 Definovat komunitu/y uživatelů repozitáře, rozumět jejich potřebám a dovednostem
		6.3.1 Specifikovat metadata pro DIP	4.2 Definovat komunitu/y uživatelů repozitáře, rozumět jejich potřebám a dovednostem
		6.4 Specifikovat transformaci SIP do AIP 6.5 Specifikovat transformaci AIP do DIP	
7. Plán zajištění kontinuity	Zabývá se povinnostmi a zajištěním ochrany dokumentů i po zániku repozitáře.	7.1 Ochrana digitálních dat by měla být zajištěna po dobu, která překračuje existenci digitálního repozitáře	1.3 Vytvořit nouzové plány pro případ finančních omezení nebo krizí, které dostatečně ochrání důležitá data
8. Krizový plán	Reaguje na náhlé změny v prostředí repozitáře.	8.1 Digitální repozitář včas reaguje na podstatné změny v prostředí	5.2 Infrastruktura IT musí garantovat integritu a bezpečnost uložených dat
			1.3 Vytvořit nouzové plány pro případ finančních omezení nebo krizí, které dostatečně ochrání důležitá data
			9.5 Repozitář musí udržovat, provádět a hodnotit takové strategie dlouhodobé ochrany digitálních dat, které vyhovují konkrétním cílům dlouhodobé ochrany
9. Plán ochrany	Zajišťuje zpřístupňování a použitelnost	9.1 Repozitář musí mít přehled o současném a vznikajícím hardware, software a technologiích ukládání dat	3.3 Rozvíjet kvalifikaci zaměstnanců
		9.2 Repozitář by měl udržovat srozumitelné informace o všech	6.2 Specifikovat formát dat a obsah metadat pro archivaci

dokumentů bez ohledu na zastarávání nebo změny technologií.	strukturálních standardech (tj. například kódování souborů) a formátech	digitálních objektů (AIP)
	9.3 Repozitář musí udržovat znalosti o komunitě uživatelů, jejich kompetencích a znalostní bázi	4.2 Definovat komunitu/y uživatelů repozitáře, rozumět jejich potřebám a dovednostem
	9.4 Repozitář by měl vědět, jaké nároky na dlouhodobou ochranu má každý typ uloženého informačního zdroje nebo třída uložených dat	3.3 Rozvíjet kvalifikaci zaměstnanců
	9.5 Repozitář musí udržovat, provádět a hodnotit takové strategie dlouhodobé ochrany digitálních dat, které vyhovují konkrétním cílům dlouhodobé ochrany	6.2 Specifikovat formáty dat a obsah metadat pro archivaci digitálních objektů (AIP)
		6.2.1 Specifikovat metadata pro AIP
		8.1 Digitální repozitář včas reaguje na podstatné změny v prostředí
	5.2 Infrastruktura IT musí garantovat integritu a bezpečnost uložených dat	
9.6 Repozitář by měl na základě vypracované strategie pravidelně hodnotit, které informace mají být uchovávány		

Příloha č. 3 - Přehledný popis funkčních entit a rolí OAIS modelu

Funkční entita a role	Služba	Funkce
Příjem Zabezpečuje služby a funkce pro akceptaci vstupních informačních balíčků (SIPů) od producentů dat a upravuje obsah pro uložení a správu v archivu.	Příjem	Příjem SIPů.
	Zabezpečení kvality	Provádění kontrol na SIPEch.
	Generování AIP	Generování AIPů, které splňují formátování archivních dat a dokumentové standardy.
	Generování popisných informací	Extrakce popisných informací a AIPů pro přidání do archivní databáze.
Permanентní úložiště Poskytuje služby a funkce pro uchovávání, správu a vyhledávání AIPů.	Koordinace updatů	Koordinace updatů Permanentního úložiště a Správy dat.
	Příjem dat	Příjem AIPů do modulu Příjem a přidání AIPů do permanentního úložiště.
	Správa hierarchie úložiště	Správa hierarchie úložiště
	Výměna médií	Obnova médií úložiště
	Kontrola chyb	Vykonávání rutinních a speciálních kontrol chyb.
	Obnova škod	Nabízí nástroje pro obnovu dat.
	Poskytování dat	Poskytování AIPů modulu Přístupu.
	Administrace databáze	Administrace funkcí archivní databáze (zachování schémat a definice zobrazení včetně referenční integrity).
Správa dat Poskytuje služby a funkce pro populaci, správu a přístup k popisným informacím identifikujícím a dokumentujícím archivní jednotky a administrativním datům ke správě archivu.	Příjem updatů databáze	Provedení updatů databáze (nové popisné informace nebo archivní administrativní data).
	Dotazování	Posílání dotazů modulu Správy dat pro získání požadovaných výsledků.
	Generování reportů	Vytváření reportů z výsledků.
Administrace	Vyjednání smluv pro	Vyžádání a vyjednání smluv s producenty pro

Poskytuje služby a funkce pro veškeré operace v permanentním úložišti.	odevzdávání dat	odevzdávání dat.
	Žádosti o audit	Žádosti o audit tak, aby odpovídaly archivním standardům.
	Správa systémové konfigurace	Monitoring a zlepšení archivních operací. Správa konfigurace hardwaru a softwaru systému.
	Update archivní informace	Inventarizace, hlášení a migrace/update obsahu archivu.
	Zavedení standardů a politik	Sestavení a dodržování archivních standardů a politik.
	Zákaznický servis	Poskytování zákaznických služeb.
	Aktivace žádostí	Aktivace uložených žádostí.
Plánování dlouhodobé ochrany Poskytuje služby a funkce pro monitorování systému a nabízí doporučení, které vedou k tomu, aby informace uchované v systému zůstaly přístupné cílové uživatelské skupině i do budoucna, kdy původní/originální prostředí bude překonáno.		Hodnocení obsahu archivu a pravidelné doporučování k updatu/migraci archivních informací.
	Vytvoření strategií a standardů pro dlouhodobou ochranu	Vytvoření strategií a standardů pro dlouhodobou ochranu.
	Sledování technologií Monitorování cílové skupiny	Monitorování změn v technologiích a požadavků cílové skupiny a jejich znalostní báze.
	Sestavení podoby balíčků a migračních plánů	Vytvoření šablon informačních balíčků i podoba SIPů a AIPů pro specifické vstupy. Sestavení podrobných migračních plánů, softwarových prototypů a testovacích plánů pro naplnění migračních cílů.
Přístup Poskytuje služby a funkce uživatelům ke zjištění existence, popisu, lokace a dostupnosti informací uložených v systému. Umožňuje uživatelům zažádat a přijat informační balíčky.	Koordinace přístupových aktivit	Komunikace s uživateli ohledně žádostí. Kontrola přístupu ke chráněným informacím. Vyřízení žádostí.
	Generování DIP	Generování výsledků (Dissemination Information)

		Packages, výsledky, reporty).
	Dodání výsledků	Doručení výsledků uživatelům.

Příloha č. 4 - Příklad jednoduché politiky a příslušných postupů podle RLG Report

Politika sbírky University of Anytown se vztahuje na několik stupňů sběru dat pro tyto oblasti:

- všeobecné/výzkumné
- studijní
- minimální

Její politika je aplikována na následující stupně zodpovědností pro dlouhodobou ochranu:

- archivní (uchováváno navždy)
- servisní (dostupné pro dohlednou budoucnost)
- zrcadlené (zodpovědnost jen po krátkou dobu)
- odkazované, linkované (nepředpokládá se dlouhodobá ochrana)

Každý stupeň vyžaduje jiné odpovídající technické procesy a postupy pro digitální uchování. Pro „archivní“ materiály ve „všeobecné“ sbírce všechny signifikantní vlastnosti budou ochraňovány a míra povinných metadat bude přiřazena (v ideálním případě) již při odevzdání dat do repozitáře. Na druhé straně, "odkazované" materiály ze sbírky „minimální“ si vyžadují nízký či žádný stupeň (dlouhodobé) ochrany.

Splnění této zodpovědnosti vyžaduje:

- politiky pro rozšiřování sbírek (např. výběr a uchování), které odkazují na technické postupy o tom jak a na jaké úrovni budou uchovávané materiály ochraňovány a jakým způsobem k nim bude zabezpečený přístup v krátkodobém i dlouhodobém měřítku
- politiky pro kontrolu přístupů, jež zaručí, že zájmy všech zúčastněných stran jsou dostatečně ochráněny; jde zejména o autentizaci uživatelů a jim zpřístupňovaný rozsah sbírek, resp. materiálů
- politiky pro uchování materiálů - včetně smluv „Service-level Agreements“ s externími dodavateli
- politiky, které definují cílové skupiny repozitáře a popisují jejich znalostní bázi
- důkladný systém updatů politik a postupů v souladu se změnami v technologiích a cílové skupině repozitáře
- explicitní odkazy mezi těmito politikami a postupy, čímž je zabezpečena jednoduchá aplikace mezi heterogeními sbírkami

Příloha č. 5 - Příklad kritéria č. 14 v Data Seal of Approval

14. Uživatel dat dodržuje přístupová pravidla stanovená digitálním repozitářem

Požadovaná úroveň shody:

4. Implementováno: tuto zásadu jsme plně implementovali pro potřeby našeho repozitáře.

Návod pro žadatele

Tato Zásada se týká zodpovědnosti repozitáře za tvorbu pravidel přístupu k datům. Tato pravidla odpovídají příslušné národní a mezinárodní legislativě. Zásada se vztahuje i k tomu, do jaké míry digitální repozitář informuje uživatele o podmínkách přístupu do repozitáře.

- Využívá repozitář licencí pro koncové uživatele (*End User Licence*)?
- Vyžadují některá z dat uložených v digitálním repozitáři specifický režim?
- Existují smlouvy pro přidělení přístupu k datům s omezeným přístupem, případně k důvěrným datům?
- Využívá repozitář speciální licence, např. „Creative Commons“?
- Existují opatření pro případ, že stanovené podmínky nejsou splněny?

Návod pro hodnotitele

- Vztahuje se na vyjádření k sebehodnocení k příslušné Zásadě? (Odkážte na doplňkový text pro žadatele náležející ke každé Zásadě).
- Jsou odkazy na doplňující dokumentaci veřejně dostupné?
- Souhlasíte se stupněm shody uvedeným v sebehodnocení a považujete poskytnuté podklady za dostačující pro udělení DSA?
- Jsou všechny zkratky vysvětleny?
- Tam kde je to možné, proveďte svou odpověď na konkrétní sebehodnocení spíše návodnými komentáři než konkrétními dotazy.

Příloha č. 6 - Seznam kritérií katalogu nestor

A. Organizační rámec
<p>1. Repozitář má jasně definované své cíle.</p> <p>1.1. Výběr kritérií 1.2. Odpovědnost za dlouhodobou ochranu informací v podobě digitálních objektů 1.3. Cílová skupina</p>
<p>2. Repozitář zpřístupňuje své cílové skupině vhodné informace reprezentované digitálními objekty.</p> <p>2.1. Přístup cílové skupiny 2.2. Interpretace digitálních objektů cílovou skupinou</p>
<p>3. Dodržování zákonných a smluvních pravidel</p> <p>3.1. Smluvní ujednání mezi producenty a repozitářem 3.2. Repozitář má na archivaci legální základ 3.3. Repozitář má na využívání informací legální základ</p>
<p>4. Organizační forma pro digitální repozitář je adekvátní</p> <p>4.1. Adekvátní financování 4.2. Vyhovující stav kvalifikovaného personálu 4.3. Organizační struktura 4.4. Repozitář má plány dlouhodobé ochrany 4.5. Pokračování archivačních funkcí i po skončení repozitáře</p>
<p>5. Existuje adekvátní management kvality</p> <p>5.1. Definice procesů a zodpovědností 5.2. Dokumentace prvků a procesů 5.3. Reakce na výrazné změny</p>
B. Správa objektů¹³⁶
<p>6. Repozitář zabezpečuje integritu digitálních objektů během všech procesních fází:</p> <p>6.1. Příjem (Ingest) 6.2. Permanentní úložiště 6.3. Přístup</p>
<p>7. Repozitář zaručuje autenticitu digitálních objektů během všech procesních fází:</p> <p>7.1. Příjem (Ingest) 7.2. Permanentní úložiště 7.3. Přístup</p>
<p>8. Repozitář má strategický plán pro technická archivační opatření</p>
<p>9. Repozitář akceptuje digitální objekty od producentů podle předem stanovených kritérií</p> <p>9.1. Specifikace SIP 9.2. Identifikace signifikantních vlastností digitálních objektů určených k dlouhodobé archivaci 9.3. Technická kontrola digitálních objektů za účelem provedení preservačních/archivačních metod</p>
<p>10. Permanentní úložiště digitálních objektů funguje podle stanovené specifikace</p> <p>10.1. Definice AIP 10.2. Transformace SIP do AIP</p>

¹³⁶ Terminologie vychází z ISO normy ISO 14721:2002 Space data and information transfer systems - Open archival information system (OAIS) -- Reference model. Podrobněji viz kapitola 1.2 Open Archival Information System Reference Model - OAIS.

<p>10.3. Uchovávání a čitelnost AIP</p> <p>10.4. Implementace archivačních strategií pro AIP</p>
<p>11. Repozitář umožňuje využívání dat podle předem stanovených kritérií</p> <p>11.1. Definice DIP</p> <p>11.2. Transformace DIP do AIP</p>
<p>12. Systém pro správu dat je schopen nabídnout důležité funkce digitálního repozitáře</p> <p>12.1. Persistentní identifikaci objektů a jejich vazeb</p> <p>12.2. Metadata pro obsahový a formální popis a identifikaci digitálních objektů</p> <p>12.3. Metadata pro strukturální popis digitálních objektů</p> <p>12.4. Metadata pro dokumentaci změn provedených na digitálních objektech</p> <p>12.5. Metadata pro technický popis digitálních objektů</p> <p>12.6. Metadata pro přístupová práva k digitálním objektům</p> <p>12.7. Přiřazení metadat je zaručeno pro všechny digitální objekty</p>
<p>C. Infrastruktura a bezpečnost</p>
<p>13. IT infrastruktura je adekvátní</p> <p>13.1. IT infrastruktura implementuje požadavky z managementu objektů¹³⁷</p> <p>13.2. IT infrastruktura implementuje bezpečnostní nároky na správu objektů</p>
<p>14. IT infrastruktura implementuje požadavky managementu objektů</p>

¹³⁷ Management objektů je zde chápán na úrovni organizační struktury, na rozdíl od správy objektů, která je pojmána jako proces/aktivita.

Příloha č. 7 - Seznam nejdůležitějších standardů, metodik a dokumentů pro efektivní správu repozitářů, knihoven a archivů

Oblast	Název
Strategie a návrh	<ul style="list-style-type: none"> • Dokumenty (standardy a projekty) technické komise ISO/TC 20/ podskupiny SC 13 - Space data and information transfer systems <ul style="list-style-type: none"> • z toho (pro digitální repozitáře, knihovny a archivy) nejdůležitější standard ISO 14721:2003 Space data and information transfer systems -- Open archival information system (OAIS) -- Reference model • doporučený standard 14641-1:2012 Electronic archiving -- Part 1: Specifications concerning the design and the operation of an information system for electronic information preservation • Referenční model "DELOS Digital Library Reference model. Foundations for digital Libraries" • Důvěra jako klíčový koncept systémů <ul style="list-style-type: none"> • PLATTER - Plán důvěryhodného digitálního repozitáře • Trusted Digital Repositories: Attributes and Responsibilities (An RLG-OCLC Report)
Administrace - Bezpečnost dat (výběr nejdůležitějších)	<ul style="list-style-type: none"> • Dokumenty (standardy a projekty) technické komise ISO/IEC JTC1, SC27 (podskupiny 27) <ul style="list-style-type: none"> • ČSN ISO/IEC 27000 - 27002 - Systém řízení bezpečnosti informací • ČSN ISO/IEC 15408 - Kritéria pro hodnocení bezpečnosti IT • ČSN ISO/IEC TR 13335-1 - 13335-4 - Směrnice pro řízení bezpečnosti IT • ISO/IEC TR 13335-5 - Guidelines for the management of IT Security, Management guidance on network security • Dokumenty v přípravě: <ul style="list-style-type: none"> • ISO/IEC 27017 — Information security management for cloud systems • ISO/IEC 27018 — Data protection for cloud systems • ISO/IEC 27040 — Guideline on storage security

Administrace – Metadata a persistentní identifikátory (výběr nejnámějších)	<ul style="list-style-type: none"> • Popisná (volitelný výběr, podle typu systému): <ul style="list-style-type: none"> • EAD - Encoded Archival Description • MARC - MACHine Readable Cataloging • DC – Dublin Core • MODS - Metadata Object Description Schema • MPEG-7 - The Multimedia Content Description Interface • Strukturální, administrativní, preservační (autorkou velmi doporučeno): <ul style="list-style-type: none"> • METS - Metadata Encoding and Transmission Standard • PREMIS - PREMIS Data Dictionary for Preservation Metadata, version 2.2 • Technická (nezbytnost): <ul style="list-style-type: none"> • NISO MIX - Z39.87 Data dictionary - technical metadata for digital still images (MIX) - NISO Metadata for Images in XML • Persistentní identifikátory (autorkou velmi doporučeno): <ul style="list-style-type: none"> • Handle
Uživatelské rozhraní	<ul style="list-style-type: none"> • Zejména set standardů, které řeší ergonomiku a především interakci člověk-počítač (HCI): standardy technické komise TC 159
Audit a certifikace (výběr nejdůležitějších)	<ul style="list-style-type: none"> • Certifikace podle Evropského rámce pro certifikaci a audit digitalních repozitářů <ul style="list-style-type: none"> • DSA – Data Seal of Approval • Nestor/DIN 31644 - NESTOR - Network of Expertise in Long-Term Storage and Long-Term availability of Digital Resources in Germany/DIN 31644 Criteria for trustworthy digital archives • ISO 16363 - TRAC – Trustworthy Repositories Audit & Certification: Criteria and Checklist • Audit/"samo-audit" <ul style="list-style-type: none"> • DRAMBORA - Digital Repository Audit Method based on Risk Assessment
Finanční odpovědnost (výběr nejnámějších)	<ul style="list-style-type: none"> • DP4lib - Digital Preservation for libraries • LIFE - Life Cycle Information for E-literature

ekonomických modelů)	<ul style="list-style-type: none"> • PrestoPRIME (audio-vizuální dokumenty) • 4C - Collaboration to Clarify the Costs of Curation
Organizace (vyplátí se spíše pro organizce většího typu)	<ul style="list-style-type: none"> • Set standardů ISO 9000 (volitelné, náročné na implementaci)

Příloha č. 8 - Dotazník pro instituce (vzor)

V Praze 22.3.2014

Vážený/á XXX,

jako vedoucího Oddělení standardů v NK ČR bych Vás touto cestou chtěla požádat o spolupráci při vyplnění krátkého dotazníku, který slouží pro získání části dat pro moji disertační práci "Strategie, návrh, řízení a administrace rozsáhlých digitálních knihoven a archivů", kterou zpracovávám pod vedením doc. PhDr. Richarda Papíka, Ph.D. Cílem dotazníku je zjistit, jakým způsobem knihovny většího typu přistupují ke svým digitálním knihovnám a archivům.

Pokud byste chtěl/a být informován o výsledcích dotazníku nebo byste měl/a jakékoli dotazy, napistě mi, prosím, na email xxx@xxx.com.

Vyplnění dotazníku zabere cca 25 min. času. Byla bych Vám velmi vděčná, kdybyste byl/a ochotný/á se tohoto průzkumu zúčastnit! Prosím o jeho vyplnění do 14. dubna 2013.

Předem moc děkuji!

Mgr. Andrea Fojtů

doktorandka Ústavu informačních studií a knihovnictví FF UK v Praze
vedoucí Oddělení knihovnických aplikací ÚVT UK v Praze

Dotazník (9 okruhů otázek)

1. Jaká byla motivace k vytvoření Vašeho digitálního repozitáře? Co Vás ovlivnilo? Jaký máte digitální repozitář?

2. Jakým způsobem jste vybírali Váš systém? Proč jste zvolili open source systém? Zvažovali jste i komerční systém? Ovlivnily vás zkušenosti se systémem a repozitáři obecně u nás nebo v zahraničí?

3. Jaké finanční prostředky byly využity ke zřízení digitálního repozitáře? Čerpali jste finance z grantu či jste využili interní finanční zdroje? Jak byly prostředky využity? Pouze nákup HW, nebo i úpravy systému apod.?

4. Jaký typ dat (audio, video, text, obraz) repozitář obsahuje? Jde převážně o digitalizované či digital-born dokumenty? Jak máte zdefinovaný objekt v repozitáři (jedna stránka, jeden komplexní či METS objekt apod.)? Kolik objektů máte v repozitáři? Jakým způsobem získáváte metadata pro tyto objekty (stahujete je z interního katalogu, souborného katalogu, ručně doplňujete apod.)?

5. Zodpovídá architektura a workflow repozitáře standardu ISO 14721:2003 Space data and information transfer systems -- Open archival information system -- Reference model, příp. jeho revidované verzi ISO 14721:2012? Pokud ne, proč (nepovažujete jej za nutný pro správný chod repozitáře)? Pokud částečně, která část chybí? Plánujete ji doplnit?

6. Jakým způsobem dlouhodobě uchováváte data v repozitáři? Jakým způsobem chráníte fyzicky svá data a na jakém typu nosiče jsou data uchovávána? Pokud provádíte zálohy, tak v jakém intervalu? Testujete obnovitelnost záloh a v jakém intervalu? Provádíte i logickou ochranu dat (např. dodržování standardů, výběr neproprietárních formátů, stanovení signifikantních vlastností objektů pro případné budoucí migrace apod.)?

7. Outsourcujete některou část Vašeho repozitáře (např. zálohování, úpravy systému apod.)? Je pro Vás důležité přívětivé a přístupné (usable and accessible) rozhraní pro uživatele? Zodpovídá rozhraní alespoň částečně okruhu ISO standardů 9241 nebo některého jiného standardu či de-facto standardu? Získáváte nějakým způsobem zpětnou vazbu o rozhraní od uživatelů?

8. Je pro Váš repozitář, resp. organizaci spravující repozitář důležitá certifikace? Podstoupili jste již některý z auditních či certifikačních procesů Evropského rámce pro audit a certifikaci repozitářů (<http://www.trusteddigitalrepository.eu/Site/Trusted%20Digital%20Repository.html>)? Pokud ne, proč? Pokud ano, který (DSA, ISO 16363 or DIN 31644) a kdy?

9. Máte pro repozitář vypracované politiky (finanční plánování, management rizik, dlouhodobá ochrana dat, sběr dat a přístupných formátů apod.)? Pokud ne, plánujete je v nejbližší době?

DĚKUJI ZA VYPLNĚNÍ DOTAZNÍKU!

Příloha č. 9 - Tabulka souhrnných údajů dotazníku pro instituce

Otázka	Instituce	Odpovědi institucí
1. Jaká byla motivace k vytvoření Vašeho digitálního repozitáře? Co Vás ovlivnilo? Jaký máte digitální repozitář?	KNAV (interview – autorizované)	Knihovna AV ČR má k dispozici digitální knihovnu (systém Kramerius - vydání periodik a monografií z 19. a 20. století z produkce Akademie věd ČR a jejích historických předchůdců), která vznikla jako odpověď na zpřístupňování digitalizovaných dokumentů knihovny. Nově je v systému Eprints zpřístupňována digitální produkce ústavů AV ČR. Navíc s potřebou evidování a hodnocení vědeckého publikování a zavedením zlaté OA politiky (1 mil. Kč), byl sestaven digitální repozitář ASEP.
	MKP	MKP musela existenci repozitáře řešit, protože začala v rámci projektu HISPRa z norských fondů digitalizovat a potřebovala uložit digitální kopie dokumentů. MKP nemá klasický digitální repozitář, nepoužívá žádný klasický systém na správu. Repozitář tvoří disková pole o kapacitě 3x10 TB, data jsou zálohována v námi zvolené struktuře a existují dvě plnohodnotné zálohy rozmístěné v jiné budově MKP.
	NA	Motivací pro vytvoření NDA byla potřeba řešit problematiku dlouhodobého uchovávání digitálních archiválií. NDA byl budován na základě technologického projektu vytvořeného v letech 2007-2008. Repozitář je součástí LTP systému Archivematica.
	NFA	Nutnost. Přijímáme do sbírek nativně digitální audiovizuální data a chápeme, že je třeba je prezervovat nejen na bitové, ale i logické úrovni. Jsme však na obou úrovních zcela na začátku:

		<p>bitovou ochranu zajišťujeme jedním online diskovým polem zabezpečeným technologií DDP (obdoba RAID6) a dvěma redundantními řadami LTO-6 pásek (LTFS). Doplnkový a spíše záložní systém (protože neobsahuje plnou kvalitu děl) je postaven na optických médiích (nyní dvouvrstvé BD). Výhradní prezervační formát (který zároveň může sloužit i k prezentaci) je nekryptované DCP (Digital Cinema Package). Další formáty jsou jen pomocné, slouží k prezentaci. Na logické úrovni nyní vedeme základní evidenci o sbírkách v excelové tabulce, do budoucna uvažujeme o implementaci zřejmě Archivematicy (s extrakcí technických metadat zřejmě pomocí MedialInfo), příp. P4.</p>
	<p>NTK</p>	<p>Máme dva digitální repozitáře. První digitální repozitář Národního úložiště šedé literatury (dále jako A) vznikl jako pilotní projekt sběru, dlouhodobé archivace a zpřístupnění šedé literatury v České republice. Druhý Institucionální digitální repozitář: Studijní knihovna pro výzkum vývoje NTK (dále jako b) vznikl pro potřeby NTK. Zde se ukládají veškeré výstupy zaměstnanců NTK a materiály z akcí konaných NTK.</p>
<p>2. Jakým způsobem jste vybírali Váš systém? Proč jste zvolili open source systém? Zvažovali jste i komerční systém? Ovlivnily vás zkušenosti se systémem a repozitáři obecně u nás nebo v zahraničí?</p>	<p>KNAV (interview – autorizované)</p>	<p>Původně byl v KNAV komerční systém. Po nespokojenosti s podporou a nepružným vývojem padlo rozhodnutí přejít na open source systém – Kramérius (Fedora systém).</p>
	<p>MKP</p>	<p>Systém jsme vybírali velmi dlouho, bohužel se nám žádný nejevil jako optimální pro naše účely, vybírali jsme především z open source systémů. Inspiraci jsme hledali jak u nás, tak především v zahraničí.</p>
	<p>NA</p>	<p>Systém byl vybírán v otevřeném výběrovém řízení. Po dvou neúspěšných pokusech jsme se byli nuceni zvolit jiné řešení a</p>

		vybovat NDA vlastními silami.
	NFA	Archivematicu používá nebo o ní uvažuje řada dalších paměťových institucí v ČR a spolupráce na implementaci by mohla být užitečná. Open source řešení považujeme principiálně vhodné pro instituce našeho typu (nižší riziko ztráty kontinuity produktu, než u komerčního produktu).
	NTK	U A jsme provedli analýzu dostupných open source, naše požadavky a vše další je popsáno v dokumentu viz http://invenio.nusl.cz/record/111525/files/idr-270_1.pdf Pro B jsme použili již lokalizované Invenio z A.
3. Jaké finanční prostředky byly využity ke zřízení digitálního repozitáře? Čerpali jste finance z grantu či jste využili interní finanční zdroje? Jak byly prostředky využity? Pouze nákup HW, nebo i úpravy systému apod.?	KNAV (interview – autorizované)	Z programu VISK7 byla financována digitalizace, nákup HW.
	MKP	Finance na nákup potřebného HW byly v první vlně v rámci grantu z norských fondů, nyní jsme v rámci vlastního rozpočtu řešili reorganizaci celého systému a především rozšíření kapacity úložiště o dalších 30 TB.
	NA	NDA je spolufinancován ze strukturálních fondů EU (Integrovaný operační program).
	NFA	Zatím máme pouze HW. Vše šlo z grantu ministerstva kultury, pod níž naše příspěvkovka spadá.
	NTK	A bylo částečně hrazeno z projektu a částečně z rozpočtu NTK (HW plně nTK, SW projekt a spoluúčast NTK). B plně z prostředků NTK.
	MKP	V MKP uchováváme především textové soubory v různých formátech: TXT jako výsledek OCR; HTML, PDF, EPUB, PRC/MOBI, RTF jsou formáty pro e-knihy. Dále jsou uchovávány audio dokumenty – digitalizované nahrávky z gramofonových desek a samozřejmě obrazové soubory ve formátech JPEG,

		<p>JPG2000, TIFF a v případě digitalizace na zakázku i PDF. Dokumenty jsou ze ¼ digitální kopie, z ¼ digital-born (e-knihy). Základní objekt je definován jako 1 digitální kopie dokumentu či 1 e-kniha nebo 1 nahrávka. Tyto jsou pak tvořeny dalšími objekty. V tuto chvíli archivujeme kolem 6600 základních objektů. Pokud to lze, stahujeme metadata z knihovního systému Koniáš, doplňujeme je ručně a částečně automaticky (v případě administrativních metadat a strukturálních metadat).</p>
<p>4. Jaký typ dat (audio, video, text, obraz) repozitář obsahuje? Jde převážně o digitalizované či digital-born dokumenty? Jak máte zdefinovaný objekt v repozitáři (jedna stránka, jeden komplexní či METS objekt apod.)? Kolik objektů máte v repozitáři? Jakým způsobem získáváte metadata pro tyto objekty (stahujete s interního katalogu, souborného katalogu, ručně doplňujete apod.)?</p>	<p>KNAV (interview – autorizované)</p>	<p>Dokumenty jsou obrazového a textového charakteru. Metadata splňují standardy NK. Je to dáno především dobou (2003), kdy vznikl popis k digitalizovaným dokumentům. Tehdy METS nebyl rozšířeným standardem. Administrativní a technická metadata máme v systému, nejsou součástí objektu v METSovém chápání. Provádíme kontroly MD5. Textové dokumenty jsou v PDF (1 článek = 1 objekt = 1 bibliografický záznam). K dispozici je v digitální knihovně cca 5000 dokumentů.</p>
	<p>NA</p>	<p>Born digital dokumenty. Zkušenosti s jejich uložením však budou využity rovněž při ukládání digitálních reprodukcí. V případě digital-born dokumentům půjde zejména o textové dokumenty a obrázky. Dále půjde o tabulky, databáze, video i audio. Opomenout nelze ani vektorovou grafiku a GIS. Digitalizované dokumenty jsou ukládány zejména v podobě rastrových obrázků.</p>
	<p>NFA</p>	<p>Viz ot. 1</p>
	<p>NTK</p>	<p>A i B - video, text, obraz - digital-born – k jednomu záznamu více objektů A – data buď stahujeme přes OAI-PMH nebo je partnerské instituce vkládají ručně</p>

		B - ručně
5. Zodpovídá architektura a workflow repozitáře standardu ISO 14721:2003 Space data and information transfer systems -- Open archival information system -- Reference model, příp. jeho revidované verzi ISO 14721:2012? Pokud ne, proč (nepovažujete jej nutný pro správný chod repozitáře)? Pokud částečně, která část chybí? Plánujete ji doplnit?	KNAV (interview – autorizované)	K dispozici je digitální úložiště (nikoli ve smyslu knihovnickém), z tohoto poledu systémy neodpovídají OAIS modelu, jelikož chybí LTP systém.
	MKP	Naše úložiště neodpovídá standardu ISO 14721:2003 ani jeho revidované verzi ISO 14721:2012 a zatím pro naše potřeby není nutné zavádět změny, které by s sebou nesly akceptaci všech podmínek.
	NA	Budovaný systém vychází z OAIS. Soulad s touto normou bude ověřen v závěru projektu formou auditu.
	NFA	Nyní ne, určitě bychom rádi standardů docílili.
	NTK	A i B ano
6. Jakým způsobem dlouhodobě uchováváte data v repozitáři? Jakým způsobem chráníte fyzicky svá data a na jakém typu nosiče jsou data uchovávána? Pokud provádíte zálohy, tak v jakém intervalu? Testujete obnovitelnost záloh a v jakém intervalu? Provádíte i logickou ochranu dat (např. dodržování standardů, výběr neproprietárních formátů, stanovení signifikantních vlastností objektů pro případné budoucí migrace apod.)?	KNAV (interview – autorizované)	Provádí se bitová ochrana pro cca 50-70TB dat. Nejsou k dispozici žádné obecné směrnice (pouze IT pravidla), je to něco, co zřejmě budeme muset řešit v blízké budoucnosti.
	MKP	Dlouhodobé uchování dat je palčivá otázka, v tuto chvíli se snažíme – máme dvě záložní kopie na jiném diskovém poli – zálohy probíhají každou noc a pravidelně testujeme jejich obnovitelnost – minimálně jednou za 3 měsíce. Snažíme se udržovat si povědomí o trendech, takže dodržujeme vypsání standardy - pokud je to možné, převádíme starší formáty do novějších atd.
	NA	Data uchováváme na diskových polích RAID6 ve dvou lokalitách. V rámci projektu NDA bude záložní úložiště umístěno v nově rekonstruovaných prostorách vzdálených cca 100 km od hlavního úložiště. Rovněž bude upraven způsob uložení dat. Nově budou uložena ve dvou úložištích a v každém na dvou typech médií (diskové pole, Bluray).
	NFA	Viz ot. 1. K datům děláme MD5 checksumy. Plánujeme i nějaké periodické kontroly integrity dat, ale takhle daleko jsme se ještě

		nedostali.
	NTK	Máme systém záloh jak na serverech tak na páskách a uchováváme je geograficky odděleně. Přístupy a manipulace je dána jen pověřeným osobám. Systém záloh je propracován od denních inkrementních až po měsíční archivní. Test obnovitelnosti záloh neděláme. Pravidelně 1x měsíčně provádíme dodržování standardů.
7. Outsourcujete některou část Vašeho repozitáře (např. zálohování, úpravy systému apod.)? Je pro Vás důležité přívětivé a přístupné (usable and accessible) rozhraní pro uživatele? Zodpovídá rozhraní alespoň částečně okruhu ISO standardů 9241 nebo některého jiného standardu či de-facto standardu? Získáváte nějakým způsobem zpětnou vazbu o rozhraní od uživatelů?	KNAV (interview – autorizované)	Ne, systémy máme ve vlastní správě.
	MKP	Veškerá údržba je v rukách MKP, žádnou činnost neoutsourcujeme.
	NA	Outcosorsován bude vývoj portálu pro přístup k NDA, tzv. Archivního portálu). Přívětivost a přístupnost rozhraní pro uživatele portálu jsou pro nás důležité. Soulad s normou ISO 9241 v tomto okamžiku neřešíme. V současnosti požadujeme, aby portál odpovídal pravidlům pro přístupný web a po technické stránce HTML5.
	NFA	Neoutsourcujeme. Na uživatele určitě hledíme, tu naši tabulku se snažíme dělat přehlednou ☺
	NTK	Ne
8. Je pro Váš repozitář, resp. organizaci spravující repozitář důležitá certifikace? Podstoupili jste již některý z auditních či certifikačních procesů Evropského rámce pro audit a certifikaci repozitářů (http://www.trusteddigitalrepository.eu/Site/Trusted%20Digital%20Repository.html)? Pokud ne, proč? Pokud ano, který (DSA, ISO 16363 or DIN 31644) a kdy?	KNAV (interview – autorizované)	Ano, ale v nejbližší době není v plánu. „čekáme“ i na výsledky implementace Archivematica v grantovém projektu CESNET.
	MKP	Pro MKP není certifikace nutná, přesto se snažíme podle našich možností jednou za rok až za dva roky zkontrolovat situaci a položit si například otázky ze selfauditu DRAMBORA.

	NA	Projekt předpokládá ve svém závěru audit a certifikaci. Výstupem auditu budou zprávy o souladu informačního systému s ISO 16363 a OAIS. Výstupem certifikace bude certifikát o splnění požadavků kladených informační systém veřejné správy a o splnění požadavků kladených na IT bezpečnost.
	NFA	Takhle daleko ještě nejsme. Certifikace jistě může být užitečná. Nevím ovšem, zda pro specifickou prezervaci AV dat bude nějaký certifikační model vhodný...
	NTK	Pravidelně 1x ročně provádíme audit v nástroji DRAMBORA. O dalším certifikačním procesu neuvažujeme, především z finančních důvodů.
9. Máte pro repozitář vypracované politiky (finanční plánování, management rizik, dlouhodobá ochrana dat, sběr dat a přístupných formátů apod.)? Pokud ne, plánujete je v nejbližší době?	KNAV (interview – autorizované)	Ne. V nejbližší době něco takového neplánujeme.
	MKP	V nedávné době vznikl v MKP dokument nazvaný Základní pravidla archivace elektronických dat v MKP, který definuje politiku archivace v MKP. V rámci dokumentu byly popsány dokumenty a formáty, které se v MKP v rámci e-publikování uchovávají, možná rizika. V nejbližší době bude ještě doplněn o vlastnosti formátů, které MKP přijímá od uživatelů a jiných producentů dat a které by měla uchovávat.
	NA	Politiky nejsou v tuto chvíli kodifikované. Máme v zásadě nastavena pravidla pro příjem dokumentů, zajištěno financování pro dobu udržitelnosti a určeny formáty pro dlouhodobé uchovávání.
	NFA	Nyní ne. Rádi bychom, ale obávám se, že v nejbližší době to nebude. Konsolidace v této oblasti a nějaké smysluplné plánování očekávám za cca 3-5 let.
	NTK	Ano, vše zmíněné.

Příloha č. 10 - Dotazník pro studenty (vzor)

V Praze 22.3.2013

Vážené kolegyně, vážení kolegové,

dovolte mi požádat Vás o spolupráci při vyplnění krátkého dotazníku, který slouží pro získání částí dat pro moji disertační práci "Strategie, návrh, řízení a administrace rozsáhlých digitálních knihoven a archivů", kterou vypracovávám pod vedením doc. PhDr. Richarda Papíka, Ph.D. Cílem dotazníku je zjistit, jak uživatelé (jako klíčový prvek digitálních, repozitářů, knihoven a archivů), definují uživatelsky přívětivé rozhraní dle standardů – uvedených na příkladu Digitálního univerzitního repozitáře UK v Praze.

Pokud budete chtít být informováni o výsledcích dotazníku nebo budete mít jakékoli dotazy, napište mi, prosím, na email xxx@gmail.com.

Vyplnění dotazníku je zcela anonymní a zodpovězení 14 otázek Vám zabere cca 15 min. času.

Děkuji za Váš čas a ochotu zúčastnit se tohoto průzkumu!

Mgr. Andrea Fojtů

doktorandka Ústavu informačních studií a knihovnictví FF UK v Praze
vedoucí Oddělení knihovnických aplikací ÚVT UK v Praze

U otázek 1.-3. zaškrtněte, prosím, pouze jednu možnost!

1. Věk (v letech)

- méně než 20 30–39 starší než 50
 20–29 40–49

2. Pohlaví

- žena muž

3. Studuji

- bakalářské studium doktorské studium
 magisterské studium jiné

4. Pokud studujete, máte zároveň také zaměstnání?

- ano ne

5. Kterou VŠ studujete?

název univerzity _____ (např. Univerzita Karlova v Praze)

obor _____ (např. Psychologie)

Pokud studujete na dvou a více školách, запиšte je prosím!

6. Popište, prosím, vlastními slovy, co pro Vás znamená uživatelsky přívětivé rozhraní digitální knihovny?

8. Které vlastnosti mají podle Vás nejzásadnější vliv na výsledný dojem uživatelsky přívětivého designu digitálních knihoven a archivů?

Vyberte, prosím, max. 4 možnosti. Zakroužkujte tu nejdůležitější!

- interaktivita (komunikace rozhraní s uživatelem)
- přehlednost (uživatel se dobře orientuje nejen na jedné obrazovce, nýbrž v celém systému knihovny/archivu)
- průhlednost (jednoduché rozhraní, předvídatelné reakce systému)
- empatie (uživatel má pocit, že rozhraní bylo vytvořeno pro něj)
- standardizace (zachování klasických a navykých pracovních postupů uživatele i s ohledem na zvyklosti a standardy operačního systému)
- srozumitelnost (pojmy, které rozhraní používá, by měly být obecně známy)
- návodnost (rozhraní by mělo provázet uživatele celým systémem bez nutnosti učení se systému)
- ochrana (před nebezpečnými operacemi, které by mohly připravit uživatele o data)
- pomoc (vybavení přehlednými návody a pomůckami)
- jiné _____

7. Znáte nějakou digitální knihovnu (domácí nebo zahraniční), která podle Vás odpovídá principům uživatelsky přívětivého rozhraní?

Napište prosím její název a URL adresu.

9. Myslíte si, že by digitální knihovny a archivy měly respektovat sadu ČSN EN ISO standardu 9241, zejména části 210: Human-centred design for interactive systems?

ano stručné zdůvodnění

ne stručné zdůvodnění

10. Hledáte informace přímo na stránkách digitálních knihoven a archivů nebo využíváte některý ze známých vyhledávačů?

Zaškrtněte pouze jednu možnost.

- hledám přímo na stránkách knihoven a archivů
- hledám pouze ve vyhledávačích
- nejdříve hledám informace ve vyhledávačích, po neúspěšném hledání navštívím stránky knihoven

11. Pokud vyhledáváte přímo na stránkách digitálních knihoven a archivů, jaké operátory využíváte?

Je možné zaškrtnout i více možností.

- vyhledání přesné fráze („dotaz“)
- operátor OR (dotaz OR dotaz)
- odstranění výrazu z vyhledávání (dotaz NOT Dotaz)
- nepoužívám žádný z uvedených operátorů

Další tři otázky se týkají Digitálního univerzitního repozitáře UK v Praze (<http://repozitar.cuni.cz/>)

12. Zadejte uživatelský dotaz v Digitálním univerzitním repozitáři. Napište jej, prosím, do tohoto dotazníku spolu s informací, zda-li vyhledané výsledky odpovídaly Vaším potřebám. Jak složité či jednoduché bylo pro Vás potřebné informace dohledat?

13. Napište ve třech bodech co vidíte u Digitálního univerzitního repozitáře UK za největší problém z hlediska líbivosti a funkčnosti rozhraní? Seřadte dle důležitosti od 1. do 3., kdy bod 1 označí nejzávažnější problém.

1.

2.

3.

14. Popište v několika větách, jak celkově hodnotíte rozhraní Digitálního univerzitního repozitáře UK?

DĚKUJI VÁM ZA VYPLNĚNÍ DOTAZNÍKU!

Příloha č. 11 - Tabulka souhrnných údajů dotazníku pro studenty

Otázka	Odpovědi respondentů
1. Věk (v letech) Možné odpovědi : <ul style="list-style-type: none"> • méně než 20 • 20–29 • 30–39 • starší než 50 • 40–49 	18x 20-29 2x 30-39 1x bez odpovědi
2. Pohlaví Možné odpovědi : <ul style="list-style-type: none"> • žena • muž 	14x žena 7x muž
3. Studuji Možné odpovědi : <ul style="list-style-type: none"> • bakalářské studium • magisterské studium • doktorské studium • jiné 	21x magisterské studium
4. Pokud studujete, máte zároveň také zaměstnání?	15x ano, 5x ne

<p>Možné odpovědi :</p> <ul style="list-style-type: none"> • ano • ne 	
<p>5. Kterou VŠ studujete? <i>Pokud studujete na dvou a více školách, запиšte je prosím!</i></p> <p>Možné odpovědi :</p> <ul style="list-style-type: none"> • název univerzity (např. Univerzita Karlova v Praze) • obor (např. Psychologie) 	<p>20 x Univerzita Karlova v Praze, Informační studia a knihovnictví 1x Univerzita Karlova v Praze, Informační studia a knihovnictví a Cestovní ruch</p>
<p>6. Popište prosím vlastními slovy, co pro Vás znamená uživatelsky přívětivé rozhraní digitální knihovny?</p>	<p>Intuitivní, jednoduché, přehledné.</p> <p>Uživatelsky přívětivé rozhraní pro mě znamená pohodlné a rychlé využití všech možností, které digitální knihovna poskytuje.</p> <p>Dostatečně jednoduché, pohodlné a svižné. Bez zbytečných prvků - přehledné. Návodné.</p> <p>Takové, ve kterém se uživatel dobře orientuje, je přehledné, jednoduché a intuitivní.</p> <p>Snadná práce se systémem, vše je přehledné.</p> <p>Jednoduché, přehledné.</p> <p>Takové, které lze použít snadno, bez překážek a k uspokojení svých potřeb.</p> <p>Takové rozhraní, ve kterém se snadno orientuje - vím, kde hledat pole pro hledání, fazety apod. Intuitivní rozhraní</p> <p>Přehledné, čitelné písmo, hierarchie, spíše méně informací na 1 straně než více. Při vyhledávání nabídka příbuzných dokumentů.</p> <p>Google. Víím, je to hrozné a naivní, ale byl bych rád, aby knihovny poskytly svá data předním vyhledávačům, a to zejména Googlu. Zároveň bych byl rád, aby existovalo jedno jediné rozhraní pro profesionální vyhledávání ve všech knihovnách. Obdivuji projekty jako Worldcat, Google Books apod. a myslím si, že je to především politická (ekonomická) otázka, proč ještě data nejsou takto široce dostupná a propojená. Technologie existují a výpočetní</p>

	výkon je dostatečný.
	Jednoduchost ve vyhledávání, discovery systém typu Google.
	Takové rozhraní, které je přehledné a umožní mi dosáhnout co nejpřesnějšího výsledku dotazu, bez potřeby přílišného zadání informací.
	Uživatelsky přívětivé rozhraní pro mě zjednodušeně řečeno znamená, aby určitá prvky (např. vyhledávací pole) byly tam, kde obvykle vyhledávací pole bývá; aby bylo možné filtrovat, řadit, aby bylo dostupné i pokročilé vyhledávání.
	Rozhraní, kde hned vím co a kam napsat, kde si můžu dotaz po vyhledání zpřesnit či případně upravit, kde mi rozhraní našeptává nebo odkáže na možnou pomoc, pokud si nevím rady.
	Rozhraní je přehledné, snadno se v něm orientuje a nachází to, co uživatel hledá.
	Intuitivnost, snadné vyhledávání, snadná orientace, nápověda, barvy, které „neruší“.
	Uživatelsky přívětivé rozhraní v digitální knihovně pro mne znamená intuitivní ovládání, přehledné a dobře zorganizované uspořádání, snadné a pohodlné vyhledávání (oceňují i možnost rozšířeného vyhledávání, našeptavač či fasetové kategorie) vizuálně příjemný design, vlastní prostor (například možnost vlastní schránky), možnost projevit názor či položit dotaz.
	Přehledné, jednoduché, opticky příjemné, návody jak s rozhraním pracovat.
	Snadná orientace, jasné rozvržení, v základním nastavení co nejpodobnější klasickým internetovým vyhledávačům (pokud budu chtít „profesionálnější“ vyhledávání možnost zvolit rozšířené hledání), co nejnávodnější používání
	Prostredie by malo byť jednoduché a prehľadné (systematicky usporiadané), k dispozícií by mali byť návody, prípadne presmerovanie na pomoc s prípadnými otázkami.
	Prehľadné, s ľahkou a intuitívnou orientáciou.
7. Znáte nějakou digitální knihovnu (domácí nebo zahraniční), která podle Vás odpovídá principům uživatelsky přívětivého rozhraní? <i>Napište prosím její název a URL adresu.</i>	ERIC http://eric.ed.gov
	EUROPEANA http://europeana.org
	ROAD Directory of Open Access Scholarly Resources http://road.issn.org
	NE
	Bez odpovědi
	Neznám
	Universitat Autònoma de Barcelona http://ddd.uab.cat
	dspace.k.utb.cz

	<p>k4.techlib.cz Digitální knihovna NTK</p> <p>Přiznávám, že se snažím vyhledávat ve stylu „výraz + jméno website“, kterýžto výraz lepím do Googlu. V podstatě očekávám, že budoucnost patří JEDNOMU rozhraní, kde bude možné vyhledat vše.</p> <p>World Digital Library http://www.wdl.org/en/</p> <p>Nemohu úplně tvrdit, že by byla úplně přívětivá Kramerius http://kramerius.nkp.cz/kramerius/Welcomedo.jsessionid=DE3146902D0D7B324BDFCC1C61DB0BCB nevzpomenu si, digitální knihovny zatím spíš nenavštěvuji</p> <p>ne</p> <p>Manuscriptorium - http://www.manuscriptorium.com</p> <p>Europeana, http://www.europeana.eu/</p> <p>Osobně mi nejvíce vyhovuje Europeana, tj. http://www.europeana.eu/. Docela vysoce hodnotím také World Digital Library, tj. http://www.wdl.org/en/.</p> <p>http://www.manuscriptorium.com/?q=cs</p> <p>Library of Congress - http://www.loc.gov/collections/ Google Books - http://books.google.com/?hl=cs</p> <p>Digitální knihovna Gutenberg - www.gutenberg.org</p> <p>Digitální knihovna Kramerius – Moravská zemská knihovna http://kramerius.mzk.cz/search/</p>
<p>8. Které vlastnosti mají podle Vás nejzásadnější vliv na výsledný dojem uživatelsky přívětivého designu digitálních knihoven a archivů?</p> <p><i>Vyberte prosím max. 4 možnosti. Zakroužkujte tu nejdůležitější!</i></p> <p>Možné odpovědi:</p> <ul style="list-style-type: none"> • interaktivita (komunikace rozhraní s uživatelem) • přehlednost (uživatel se dobře 	<p>interaktivita, průhlednost, empatie, standardizace</p> <p>přehlednost, standardizace, srozumitelnost, ochrana</p> <p>přehlednost, průhlednost, srozumitelnost, návodnost</p> <p>přehlednost, průhlednost, empatie, standardizace</p> <p>interaktivita, přehlednost, srozumitelnost, ochrana</p> <p>přehlednost, standardizace, srozumitelnost, pomoc</p> <p>přehlednost, průhlednost, návodnost</p> <p>přehlednost, průhlednost, návodnost, ochrana</p>

<p>orientuje nejen na jedné obrazovce, nýbrž v celém systému knihovny/archivu)</p> <ul style="list-style-type: none"> • průhlednost (jednoduché rozhraní, předvídatelné reakce systému) • empatie (uživatel má pocit, že rozhraní bylo vytvořeno pro něj) • standardizace (zachování klasických a navykých pracovních postupů uživatele i s ohledem na zvyklosti a standardy operačního systému) • srozumitelnost (pojmy, které rozhraní používá, by měly být obecně známy) • návodnost (rozhraní by mělo provázet uživatele celým systémem bez nutnosti učení se systému) • ochrana (před nebezpečnými operacemi, které by mohly připravit uživatele o data) • pomoc (vybavení přehlednými návody a pomůckami) • jiné 	<p>přehlednost, průhlednost, srozumitelnost přehlednost, průhlednost, jiné (_Data by měla být dostupná přes Google, pak není co řešit__) přehlednost, srozumitelnost, návodnost přehlednost, průhlednost, srozumitelnost, návodnost průhlednost, přehlednost, standardizace, srozumitelnost interaktivita, přehlednost, návodnost, pomoc přehlednost, průhlednost, návodnost interaktivita, přehlednost, empatie, návodnost interaktivita, přehlednost, průhlednost, pomoc průhlednost, empatie, srozumitelnost, pomoc přehlednost, standardizace, návodnost přehlednost, průhlednost, srozumitelnost, pomoc přehlednost, srozumitelnost, návodnost, pomoc</p>
<p>9. Myslíte si, že by digitální knihovny a archivy měly respektovat sadu ČSN EN ISO standardu 9241,</p>	<p>2x bez odpovědi 4x ano (bez stručného zdůvodnění)</p>

<p>zejména části 210: Human-centred design for interactive systems?</p> <p>Možné odpovědi :</p> <ul style="list-style-type: none"> • ano - stručné zdůvodnění • ne - stručné zdůvodnění 	<p>14 x ano se stručným zdůvodněním:</p> <p>ano - standardy jsou důležité jak pro provozovatele, tak i pro uživatele</p> <p>ano - myslím, že dodržovat standardy je vždy výhodné</p> <p>ano - určitě ano, aby byly uživat. přívětivé</p> <p>ano - zvýšení počtu uživatelů</p> <p>ano - protože to, aby se uživatelé se systémem dobře pracovalo je základ</p> <p>ano - aby byl design jednotný - přehledný pro více systémů</p> <p>ano - pokud si mám zvolit z těchto dvou možností, tak dám „ano“, ale přiznám se, že tuto normu neznám. Pokud bych měla odpovědět sama, dala bych „nevím“. Dávám tedy „ano“, norma nebyla určitě vytvořena jen tak bezdůvodně, dá se čekat, že její obsah je opodstatněný a že pokud se někdo bude řídit normou, „nebude to špatně“. Ale jak jsem psala, normu neznám, takže tohle jsou jen dohady...</p> <p>ano - alespoň částečně ano, standardy dávají návody čeho se vyvarovat a co je pro uživatele vhodné – proto mohou ušetřit čas při vývoji a odstraňování nedostatků</p> <p>ano - jedná se o interaktivní systém</p> <p>ano – vždy by měl být brán ohled na uživatele</p> <p>ano - myslím si, že to může zvýšit použití digitálních knihoven, jelikož tato část se zabývá interakcí člověka a různých systémů. Pro uživatele by ovládání vytvořené podle této části normy mohlo být mnohem jednodušší</p> <p>ano - je to myslím, užitečné z více pro více stran</p> <p>ano - pokud dělám službu pro uživatele (mnohdy laika) musím tomu přizpůsobit prostředí, ve kterém se má, pokud možno bez problémů, pohybovat a orientovat</p> <p>ano - pre uľahčenie práce aj bežným používateľom</p> <p>1x ne - zbytečné, ať si trh sám nastaví, co je pro uživatele dobré a co ne</p>
<p>10. Hledáte informace přímo na stránkách digitálních knihoven a archivů nebo využíváte některý ze známých vyhledávačů?</p> <p><i>Zaškrtněte pouze jednu možnost.</i></p>	<p>3x hledám přímo na stránkách knihoven a archivů</p> <p>18x nejdříve hledám informace ve vyhledávačích, po neúspěšném hledání navštívím stránky knihoven</p>

<p>Možné odpovědi :</p> <ul style="list-style-type: none"> • hledám přímo na stránkách knihoven a archivů • hledám pouze ve vyhledávacích • nejdříve hledám informace ve vyhledávacích, po neúspěšném hledání navštívím stránky knihoven 	
<p>11. Pokud vyhledáváte přímo na stránkách digitálních knihoven a archivů, jaké operátory využíváte? <i>Je možné zaškrtnout i více možností.</i></p> <p>Možné odpovědi :</p> <ul style="list-style-type: none"> • vyhledání přesné fráze („dotaz“) • operátor OR (dotaz OR dotaz) • odstranění výrazu z vyhledávání (dotaz NOT Dotaz) • nepoužívám žádný z uvedených operátorů 	<p>4x vyhledávání přesné fráze, operátor OR, odstranění výrazu z vyhledávání 6x vyhledávání přesné fráze, operátor OR 1x operátor OR 3x vyhledání přesné fráze („dotaz“) 1x vyhledání přesné fráze („dotaz“) - K tomu bych napsal poznámku, že nehledám přesný výraz, naopak. Hledám nějaký obecnější výraz, který pak zpřesňuji. Operátory „AND“, „NOT“ atd. zásadně nikdy nepoužívám, to je snad záležitost 90. let?!</p> <p>1x opět – digitální knihovny moc nepoužívám, ale např. u elektronických informačních zdrojů (databáze Proquest, Ebsco)... používám při vyhledávání různé operátory (AND, NOT, OR). Používám také uvozovky a *.</p> <p>1x nepoužívám žádný z uvedených operátorů 1x bez odpovědi</p>
<p>12. Zadejte uživatelský dotaz v Digitálním univerzitním repozitáři. Napište jej prosím do tohoto dotazníku spolu s informací, zda-li vyhledané výsledky odpovídaly Vaším potřebám. Jak složité či jednoduché bylo pro Vás potřebné informace dohledat?</p>	<p>Zadávaný dotaz: čtenářská gramotnost Výsledky neodpovědali mojim potřebám, málo výsledkov. "Školní knihovny" sbírka: Vysokoškolské kvalifikační práce, výsledky odpovídaly Dotaz zněl čtenářská gramotnost, výsledky odpovídaly, snadné "Elektronické informační zdroje". S vyhledáváním jsem spokojen, výsledky se zdají být relevantní dotazu Pokročilé vyhledávání - informační management OR znalostní management. 3 výsledky - ano, celkem odpovídaly mým potřebám</p>

	Informační politika. Většina nalezených záznamů byla relevantní, občas se objevil nerelevantní výsledek
	Fuzzy logika. Výsledky boli OK. Vyhľadávanie jednoduché
	1. Zadáno: mapa Praha Až 9. výsledek je mapu Prahy 2. Při zadání "Praha" a omezení na mapovou sbírku je až 8. výsledek mapa Prahy
	„Knihy v životě" - začíná na/obsahuje - vysokoškol. kvalif. Práce - nenalezeno nic (ani pokročilé vyhledávání)
	"Digitální knihovny" ve VŠKP sbírce. Podle systému měly všechny výsledky stejnou váhu, ale třeba hned výsledek č. 2 mi přijde naprosto zcestný a nerelevantní
	Zadal jsem výraz: „mapa prahy historie“. Výsledky odpovídají tomu, co jsem hledal (tj. něco to najde), ale otázka je přece jiná: PROČ bych měl vyhledávat přes toto rozhraní, když si mohu podobný výraz „vygooglovat“? Je pro mne mnohem lepší tento výraz napsat do vyhledávacího pole Googlu a nechat si přes vyhledávání obrázků přímo zobrazit výsledky odkudkoliv z webu. Teprve v další fázi bych se začal dívat na specializované weby (jako např. na tento), kde bych doufal, že najdu další materiály. Každopádně, na prvním místě je Google, teprve poté jdu do „izolovaných informačních ostrůvků“...
	Dotaz: „marketingová strategie“ V rámci daných možností shledávám vyhledané výsledky jako použitelné, ale pochopitelně bych jen s tímto zdrojem nevystačila.
	Mapová sbírka PŘF UK - Československá republika 1945 – obsahuje Vyhledané informace přesně odpovídali mému požadavku
	Použila jsem pokročilé vyhledávání a zadala toto: knihovn* AND hendikep* OR handicap, typ dokumentu – všechny, formát jsem nevolila. Dostala jsem 9 výsledků, přibližně to, co jsem očekávala (např. publikace o rozvoji čtenářství u sluchově postižených dětí)
	„Informační potřeba“- výsledky odpovídaly
	Patent* AND rodin* Ano, odpovídaly. Spíše jednoduché.
	'Všechna slova= biblioterapie' ve sbírce 'Centrál.repozitář' - Výsledky odpovídaly mým potřebám - Problém mi to nedělalo, vyhledávám tam celkem často
	Zadala jsem dotaz na osobní knihovnu Karla Havlíčka Borovského. Ze semináře Knihoven osobností kultury jsem

	<p>věděla, že by ve webové aplikaci měl být. Zkusila jsem nejprve základní jednoduché vyhledávání (tak postupuje asi běžný a možná trochu líný uživatel) a dala jsem Karel Havlíček Borovský. Poněkud mě překvapilo, že to osobní knihovnu K. H. Borovského nenalezne (za předpokladu vynechání výrazů „osobní knihovna“). Po zadání v pokročilém vyhledávání „Borovský and knihovna“ mne mírně překvapilo, že stále nenajdeme celý odkaz na osobní knihovnu, ovšem v klíčových slovech již nalezneme u některých záznamů „osobní knihovna“. Při vyhledávání „osobní and knihovna and Borovský“ je prakticky stejný výsledek. Úspěšný je vlastně uživatel při zadání plného dotazu, tj. „osobní knihovna Karla Havlíčka Borovského“. Když ale zadáme „osobní knihovna Borovského“, opět dostaneme výsledky jednotlivé dokumenty. Dojem z toho mám takový, že by uživatel neměl čekat žádné složitosti (jako já) a zároveň si přesně ujasnit, co hledá (a v ideálním případě zadat přesný název). Po nalezení jsem již byla plně spokojená, náhled je velmi uspokojivý, přidány jsou rozšiřující funkce (možnost zoomu, rotace, zobrazení metadatových popisků).</p>
	<p>Informační gramotnost u dětí – z 19 výsledků pro mě byly pertinentní pouze dva, k dohledání informací jsem použila webové vyhledávače(pro mě snazší cesta)</p>
	<p>Vyhledávala jsem „Alois Jirásek“</p> <ul style="list-style-type: none"> - hledaná informace mezi výsledky byla, ovšem také zde byly výsledky, které se tématu vůbec netýkaly (u dvou dokonce naprosto netuším, co vedlo systém k tomu, aby je do výsledků zařadil) - jen doplním: z vlastních dřívějších zkušeností s vyhledáváním, že v případě, kdy dotaz není takto „jasný“, a jedná se například o sadu dvou a více pojmů, které netvoří obvyklé sousloví, vylézají mnohdy mezi výsledky i naprosto zcestné dokumenty - takže k složitost x jednoduchost > velmi záleží na hledaném dotazu, někdy je to snadné, někdy tedy rozhodně ne
	<p>Zadala som do vyhľadávača spojenie "Informačná gramotnosť" - výsledky nie celkom zodpovedali mojim potrebám, avšak po využití pokročilého vyhľadávania to korešpondovalo s mojimi požiadavkami.</p>
<p>13. Napište ve třech bodech co vidíte u Digitálního univerzitního repozitáře UK za největší problém z hlediska líbivosti a funkčnosti rozhraní? Seřadte dle důležitosti od 1. do 3., kdy bod 1 označí nejzávažnější problém.</p>	<p>Nenapadlo mě nic negativního</p> <ol style="list-style-type: none"> 1. kvalifikační práce jen do r. 2006, 2. vzhled homepage 1. vyhledávací pole nedisponuje našeptávačem a nenavrhuje opravu při překlepu 2. výsledky - stručné, ani stručné tabulkové rozhraní nezobrazuje rok publikace dok. (pro mě jedna z nejdůležitějších informací) 3. Modré tlačítko "GO" označující funkci spustit vyhledávání - na první pohled není najná jeho funkce (např.

	nahradit lupou - ikonka)
	1. málo interaktivní 2. zastaralý grafický design
	1. malé písmo
	Neviem
	1. Na první pohled nerelevantní vyhledávání 2. Otevření fulltextu klikem na pravou ikonku s popiskem "Objekt" 3. Chybějící fazety
	1. nemožnost vyhledávat v podsbírkách - a už vůbec ne přímo (např. po vstoupení do VŠKP - FF není možno vyhledávat jen v diplomkách. 2. nemožnost řadit výsledky podle roku (třeba u VŠKP to chybí citelně...) 3. možná to byl jen můj problém, ale odkaz na "vstoupení" nebo prosté zobrazení full-textu bych hledala "pod záznamem", prosté pod těmi identifikačními informacemi + nějak jednoznačněji, momentální forma je prostě: klikněte na obrázek full-text a čekejte, zda se něco stane
	1. velký problém s vyhledáváním 2. při rozkliknutí sbírky - malá písmena, tučná, těžko k přečtení, velké mezery 3. Kvalifikační práce - text práce v 1 záložce, nerozděleno na kapitoly
	1. Absolutní nepřehlednost. Když už bych měl využívat, tak chci jedno jediné textové pole, do kterého napíšu hledaný výraz. Podobně jako má Google, BING, Baidoo, apod. Prostě současný standard. 2. Výsledky a jejich (ne)použitelnost. Namátkou jsem zkusil oblast „Historické knihovní fondy FFUK“. Náhodně jsem rozkliknul nějakou knihu – výsledek jsem obdržel v nějakém flashovém či javascriptovém okně. Proč prostě nemohu dostat normální odkaz v podobě PDF, JPEG, apod., kde bych si mohl knihu stáhnout? Podobně jako to má třeba archive.org. Prostý odkaz. Žádné ovládací prvky na zvětšování/ zmenšování dokumentu atd., které stejně slouží jen k tomu, aby uživateli znemožnily download kýženého dokumentu... 3. Nestandardnost. Souvisí s bodem 1. Knihovny by se měly navzájem domluvit na jednom jediném rozhraní.
	1. Příliš mnoho textu na úvodní straně. Informace zde uvedené by bylo možné přesunout jinam. 2. Malé vyhledávací okno a okolní text.

	3. Jinak nic ☺.
	Bez odpovědi
	<p>1. U pokročilého vyhledávání mi přijde nevhodně umístěné tlačítko „GO“, napřed jsem ho vůbec nemohla najít, čekala bych ho rozhodně někde více vlevo.</p> <p>2. U pokročilého vyhledávání mi rozhraní přijde nějak roztažené do šířky. Napřed jsem si dokonce myslela, jestli to není třeba tím, že rozhraní nepodporuje můj prohlížeč, ale po kontrole v jiném prohlížeči to vypadá stejně. Není mi moc jasné, proč to tak je. Rozhodně bych to nějak zúžila, jednotlivá pole mají okolo sebe zbytečně moc prostoru.</p> <p>3. Dvě věci, které mě napadly jako první jsem napsala už výše. Tak ještě 3. věc- tlačítko GO by mohlo být pojmenované jinak – např. „Search“ nebo „Hledej“ (no ale hlavně bych ho prostě umístila jinam)</p>
	<p>1. nelze upravit výsledky (např. vyřadit typ dokumentu)</p> <p>2. další problémy nevidím</p>
	1. Nevidím výrazné problémy
	<p>1. přeplněnost úvodní stránky</p> <p>2. nezajímavý design</p> <p>3. absence helpu</p>
	<p>Nemyslím si, že by něco bylo opravdu palčivé. Design je jednoduchý, ale plně postačující a ne nadarmo se říká „v jednoduchosti je krása“. Ani k funkčnosti nemám větších námitek. Se svým dotazem jsem byla úspěšná až později spíše proto, že jsem čekala nějaké složitosti.</p> <p>1. vyhledávání – nutnost přesnosti</p> <p>2. seřazení dle váhy (asi bych se více přikláběla ke shodě s dotazem, případně neutrální abecední)</p> <p>3. někomu může připadat vizuální stránka příliš jednoduchá a malinko mi vadí tlačítko „GO“ (když zbytek stránek je v českém jazyce, nejspíš bych nahradila „Najdi“, popř. „Hledej“, „Hledat“).</p>
	<p>1. vizuálně nezajímavý</p> <p>2. chybí stručná anotace u každého dokumentu</p> <p>3. chybí řazení výsledků podle roku vydání</p>
	<p>1. při řazení výsledků dle „váhy“ se mnohdy relevantní výsledk dostanou na konec seznamu</p> <p>2. na pohled zastaralé tabulkové zobrazení výsledků</p> <p>3. podle mě zbytečné zaškrťátko při vyhledávání „obsahuje“ a „přesně“ – snad všichni využívající webové</p>

	<p>vyhledávače umí užívat značení přesného dotazu přímo v řádku na vyhledávání, proč tedy přidávat tento krok, když je to v podstatě zbytečné; naopak „začíná na“ by mi například v rozšířeném vyhledávání přišlo jako dobrý prvek</p>
	<ol style="list-style-type: none"> 1. Nedostatočné tematické pokrytie 2. Nezodpovedajúce výsledky vyhľadávaných fráz / pojmov 3. Absencia nápovedy, zvláštna formulácia upozornení (nesprávne vyhľadanie a pod)
	<ol style="list-style-type: none"> 1. trošku nudný design 2. malé písmenká 3. trošku neprehľadný zoznam digitálnych zbierok
<p>14. Popište v niekoľika vetách, jak celkově hodnotíte rozhraní Digitálního univerzitního repozitáře UK?</p>	<p>bez odpovědi</p> <p>Působí příjemným vizuálním dojmem. Vyhledávání je přívětivé</p> <p>Repozitář hodnotím kladně. Zamrzí absence dnes standardních prvků (našeptávač, automatický návrh při překlepu). Velice nešťastné je rovněž rozdělení sbírky VŠKP mezi Digitool a IS. Chybí fazetové omezení výsledků.</p> <p>Rozhraní DUR se mi zdá málo interaktivní, také graf. design se mi jeví zastaralý, ale zase mi připadá velmi funkční, relevantní vyhledávání, přehledný.</p> <p>Celkově rozhraní hodnotím příznivě. Uživatel se v něm poměrně snadno vyzná.</p> <p>Neviem objektívne zhodnotiť DUR, keďže som s ním pracoval pri vyplňaní tohoto dotazníku.</p> <p>Hodnotím jej jako vcelku uživatelsky z hlediska designu přívětivé, ale málo prakticky použitelné.</p> <p>Osobně mi nevádí ta vzhledová strohost, ale fakt je, že by se ten vzhled měl nějak modernizovat a odlehčit. Problémy s funkčností jsem popsala výše.</p> <p>Špatné vyhledávání, nepřehledné! Chválím rychlé načítání, vyhledávání závěr. prací v Centrálním katalogu UK.</p> <p>Funkční, nicméně dle mého názoru odrazující potenciální uživatele, neboť ti již očekávají rozhraní a la Google, kde prostě jen napíšu hledaný výraz a tím to pro ně končí. Z mého pohledu jsme nyní ve fázi, kdy jedn. firmy a organizace zjednodušují uživatelské rozhraní svých systémů, protože pochopili, že lidé chtějí co nejjednodušší UI. V další fázi vývoje dle mého názoru dojde ke „konsolidaci“ uzavřených informačních ostrůvků, kdy budou stále další a další databáze připojovány a indexovány na jednom místě. Budoucnost patří integraci a standardizaci.</p> <p>Z portálu mám spíše pozitivní dojem. Vyhledávání je rychlé a intuitivní, a to v několika bázích najednou. Omezující je pochopitelně fakt, že je vyhledání zdrojů ohraničeno možnostmi univerzitních sbírek.</p> <p>Celkově hodnotím repozitář UK jako kvalitní a přínosný. Vychází uživatelům vstříc několika možnostmi po otevření</p>

	<p>sekcí, přičemž si uchovává přehlednost. V případě potřeby něco vyhledat mi repozitář našel přesně to co jsem potřeboval.</p>
	<p>Celkové hodnocení rozhraní: celé rozhraní nedokážu z jednoho vyhledávání posoudit. Věci, které mě nejvíc „pálily“ jsem zmínila u předchozí otázky. Oceňuji ale, že je možné volit pokročilé vyhledávání, že rozhraní obsahuje operátory (AND, OR, NOT) = věci, na které jsem zvyklá z vyhledávání v Ebscu nebo Proquestu.</p>
	<p>Vyhledávání (jednoduché i pokročilé) vypadá jednoduše, chybí mi jednoduchá a přehledná nápověda (např. fráze se vyhledává v „“ apod. ale to jsem možná jen přehlédla) a chybí mi lepší práce s výsledky.</p>
	<p>Rozhraní je jednoduché, přehledné, konzistentní s jinými univerzitními rozhraními.</p>
	<p>Úvodní stránka je příliš strukturovaná. Nelíbí se mi bílé pozadí. Celkový vzhled mi přijde zastaralý a ne moc interaktivní.</p>
	<p>Rozhraní hodnotím velmi pozitivně. Je přehledné, intuitivní, dobře organizované, s možností jednoduchého i pokročilého vyhledávání, dává uživateli vlastní prostor, nápovědu, řadu informací, možnost změnit jazyk. Dobře odvedená práce!</p>
	<p>Celkově nepůsobí rozhraní špatný dojem, jen by chtělo zapracovat na grafické stránce a některých již zmíněných detailech.</p>
	<p>Rozhraní dle mého názoru nepatří rozhodně mezi ty nejhorší, které lze objevit – setkala jsem se s mnohem horšími, ale jistě je tu možnost vylepšení. Část uživatelů jistě trochu odradí nepřiliš moderní grafika (já osobně se také moc nemůžu přenést přes tu světle modrou – té je až dost ve Windows). Docela se mi líbí nabídka v rozšířeném vyhledávání, ale i s ním se vždy ve výsledcích objeví 2-3 šotci – je mi záhadou, jak systém na některé dotazy vyplivne naprosto zcestné výsledky v kombinaci s těmi dobrými.</p>
	<p>Prostredie hodnotím kladne - jeho jednoduchost, organizovanost a ucelenost. Oceňujem najmä záložku s archívom vyhľadávania, čo uľahčuje nasledovné vyhľadavanie. Uvítala by som však tematické rozšírenie repozitára a bližšiu špecifikáciu obsiahnutých dokumentov (oblasti, témy...).</p>
	<p>Celkovo sa mi rozhranie páči, je prehľadné a jednoduché. Nemalo by robiť problém aj menej skúseným používateľom.</p>