

Projekt LIFE

Návrh metodiky pro sledování nákladů životního cyklu v rámci Národní knihovny v Praze za využití softwaru vyvinutého v rámci projektu LIFE.

25. Října

2011

Verze dokumentu 3.1.2, 2011

Koncepce rozvoje Národní knihovny České republiky jako výzkumné organizace na léta od 2010 - 2015

Oblast 2: Dlouhodobá ochrana digitálních dokumentů (LTP – Long-Term Preservation)

Priorita 1: Plánování ochrany (2010-2013)

Vypracovali:

Bc. Zuzana Kvašová

Zaměstnanec: Oddělení digitalizace

Městská knihovna v Praze

<http://www.mlp.cz/cz/>

zuzana.kvasova@mlp.cz | Tel: +420 224 242 792

PhDr. Bedřich Vychodil

Vedoucí: Oddělení pro standardy

Národní knihovna v Praze

www.nkp.cz, www.ndk.cz

bedrich.vychodil@nkp.cz | Tel: +420 221 663 408

Obsah

Obsah.....	2
Úvod	3
1 Metodika.....	5
2 Zachycení/stanovení nákladů	6
2.1 Model LIFE verze 2	6
2.1.1 Náklady mimo životní cyklus dokumentů	6
2.1.2 Náklady životního cyklu dokumentů.....	7
Obecný model uchovávání (Generic preservation model)	14
2.2 Hlavní prvky určování celkových nákladů	14
3 Nástroje pro vypočítání nákladů.....	16
3.1 Finanční model jako excelovská tabulka	16
3.2 Webový nástroj	18
4 Závěr.....	20
Byly kontaktovány vedlejší osoby spolupracující na projektu LIFE.	20

Projekt LIFE (Life Cycle Information for E-Litarature) vznikl ve spolupráci Britské knihovny a University College London (dále UCL). V současné době se na dalším jeho rozvoji podílí také Univerzita v Glasgow (konkrétně Humanities Advanced Technology and Information Institute). Projekt je financován britskou organizací JISC (Joint Information Systems Committee)¹. Cílem je vyvinout metodiku pro vypočítání nákladů na celkový životní cyklus digitálních dokumentů. Nejedná se tedy pouze o běžně využívané odhady, ale o model poskytující exaktní metodiku k výpočtu celkových nákladů. Metodika počítá s požadavky na dlouhodobé uchování a ochranu těchto dat. Celý cyklus je rozdělen do několika etap a procesů, a tím umožňuje efektivně plánovat budoucí náklady.²

První fáze projektu LIFE byla ukončena v dubnu roku 2006. Zaměřena byla na zkoumání životních cyklů klíčových digitálních sbírek vlastněných UCL a Britskou knihovnou. Jednalo se konkrétně o sbírku Voluntarily Deposited Electronic Publications Britské knihovny, data z Web Archiving Britské knihovny a sbírku elektronických časopisů (e-journals) UCL [McLeod, 2006b]. Během této fáze byla vyvinuta metodika analýzy životního cyklu dokumentů, za využití reálných dat těchto kolekcí. Životní cyklus dokumentů definuje jednotlivé prvky existence dokumentu v průběhu doby, od akvizice, přes zpracování, až po zachování [Shenton, 2003]. Projekt LIFE cyklus rozdělil na dílčí etapy (viz kapitola 2.1.2) a vznikla první verze modelu LIFE. Na základě jednotlivých dílčích etap byl vyvinut model pro vypočítání nákladů. Náklady na jednotlivé etapy dávají dohromady plný finanční závazek pro sběr digitálních dokumentů a jejich uchování. V závěru projektu byly nastíněny úkoly a doporučení pro budoucnost. V zájmu efektivního rozvoje metod a nástrojů pro vypočítávání nákladů by měla být podporována spolupráce akademických institucí s knihovnami. Za tento brzký rozvoj by měla převzít odpovědnost komunita zabývající se digitální archivací (digital preservation). Měla by vzniknout dlouhodobá partnerství mezi institucemi, aby projektové týmy mohly řešit společné problémy. Další výhodou by bylo i navázat partnerství s odpovídajícím průmyslem [McLeod, 2006a].

Druhá fáze projektu byla ukončena v srpnu 2008 a přinesla novou verzi modelu LIFE a tři nové případové studie, které celý model představily v jiných podmínkách. Projekt se tak snažil o větší univerzálnost, především v oblasti odhadu nákladů. Začalo být počítáno s tím, že by model LIFE mohl být využit i v dalších institucích, kromě knihoven například muzea

¹ <http://www.jisc.ac.uk/aboutus.aspx>

² <http://www.life.ac.uk/about/>

a archivy. Spolupráci nabídly Dánská královská knihovna³, dánské státní archivy⁴ a Státní a univerzitní knihovna.⁵ Na základě této spolupráce v rámci projektu LIFE vznikly další případové studie.

Třetí fáze projektu začala v březnu 2009 a původně měla být ukončena již v roce 2010. Financování zajišťuje Joint Information Systems Committee (JISC) a Research information network⁶. Cílem bylo rozšíření stávajícího Generic Preservation modelu (viz kapitola 0) a vytvoření komplexních sad modelů, které by pokrývaly všechny etapy v životních fázích dokumentů včetně digitalizace a zvýšily tak přesnost odhadu nákladů. Prediktivní nástroj pro výpočet nákladů měl být k dispozici pro testování koncem roku 2010 jako online webová aplikace a též offline verze ve formě Excelovské tabulky. V současné době (září 2011) není k dispozici v žádné z uvedených podob. Tento nástroj je založen na modelu LIFE verze 2.1 (viz kapitola 2.1.2), upravené verzi původního modelu. Kromě klasických materiálů zahrnuje i scénáře uchovávání zvukových médií, webových dokumentů, elektronických časopisů (e-journals) apod. V potaz byly vzaty faktory jako např. zálohování, náklady na hardware a jeho předpokládaná životnost, podpora, infrastruktura a v neposlední řadě náklady na elektrickou energii.

Původním záměrem dohody o pracovní činnosti bylo vytvořit návrh metodiky sledování nákladů životního cyklu digitálních dokumentů pro Národní knihovnu ČR. Práce měla vycházet především ze zkušeností s testováním finančních nástrojů, vzniklých v průběhu projektu LIFE III. K dispozici měl být, na základě dohody s projektovým manažerem Brianem Holem z Britské knihovny, webový nástroj pro predikci a výpočet nákladů. Pracovní tým projektu LIFE vybízí další instituce k jeho testování prostřednictvím oficiálních webových stránek projektu⁷. Po obtížné a zdlouhavé komunikaci se zástupci projektu se však nepodařilo tento webový nástroj do dnešní doby získat (září 2011). Práce tedy vychází zejména z dostupné dokumentace o projektu. Obsahuje metodiku vyčíslování nákladů na životní cyklus dokumentů a podrobný popis nákladů podle modelu LIFE verze 2. Nástroje pro vypočítání nákladů jsou představeny bez praktického otestování, pouze z dostupných zdrojů.

³ <http://www.kb.dk/en/>

⁴ <http://www.sa.dk/content/us/>

⁵ <http://www.statsbiblioteket.dk/>

⁶ <http://www.rin.ac.uk/>

⁷ <http://www.life.ac.uk/>

1 Metodika

Na úplném začátku vyčíslování nákladů na životní cyklus dokumentů si každá instituce musí stanovit několik základních postupů:

- Volba určité sbírky dokumentů, jež mají být uchovávány, a stanovení časového rámce pro trvání jejich uchovávání.
- Přizpůsobení modelu LIFE potřebám instituce.
- Definování lidských zdrojů v jednotlivých fázích projektu.
- Provést pohovory se zaměstnanci za účelem porovnání, zda se terminologie LIFE projektu shoduje s terminologií využívanou v instituci.
- Poskytnutí zpětné vazby a řešení případných problémů v případech, kdy model LIFE neodpovídá potřebám instituce.
- Sestavení workflow.
- Analýza nákladů za využití modelu LIFE.
- Vyplnění tabulek (spreadsheet) modelu LIFE k vypočítání nákladů.
- Vyvození závěrů.

2 Zachycení/stanovení nákladů

Projekt LIFE vytvořil metodologii k zachycení, vyčíslení a popsání nákladů na životní cyklus dokumentů. Mezi tyto náklady patří náklady na vybavení, průběžné náklady na zaměstnance, mzdy apod. (viz následující podkapitoly). Náklady jsou projektovány na základě hodnoty v době vzniku odhadu bez ohledu na inflaci. Odhad udává náklady v průběhu jednoho, pěti a deseti let. Zahrnuje náklady, které mohou být buď vynaloženy pouze jednou, nebo mohou narůstat v průběhu doby, nebo se pravidelně či nepravidelně obnovovat. Tato metodika umožňuje odhad nákladů na jeden titul nebo položku v daném časovém období.

2.1 Model LIFE verze 2

Tento model zachycuje procesy, kterými během svého životního cyklu digitální objekty v naprosté většině případů procházejí. Procesy jsou zde uspořádány v chronologickém pořadí, od jejich vytvoření až po případné zpřístupnění. V praxi se však jednotlivé procesy mohou překrývat, nebo být prováděny v jiném pořadí. Životní cykly některých druhů digitálních objektů mohou obsahovat jen některé z uvedených procesů.

2.1.1 Náklady mimo životní cyklus dokumentů

Náklady mimo životní cyklus dokumentů jsou náklady, které by instituce měla vzít v potaz, i přesto, že nejsou přímo spjaty s životním cyklem. Patří sem například oblast personálního managementu a administrativy (finanční a lidské zdroje), nebo zřízení kancelářských prostor a jejich správa. V oblasti systémů a infrastruktury mohou vzniknout náklady na software repozitáře. V oblasti ekonomické přípravy je například třeba mít na zřeteli inflaci, či případná úsporná opatření. Každá instituce by si měla zvážit, zda je pro ni relevantní takové náklady zařadit do celkového modelu či ne.

Fáze mimo životní cyklus	Management a administrativa	Systémy a infrastruktura	Ekonomická příprava
Prvky mimo životní cyklus	Management	Software repozitáře	Inflace
	Administrativa		Úsporná opatření

Obrázek 1: Náklady mimo oblast životního cyklu dokumentů vztahující se ke kapitole 2.1.1

2.1.2 Náklady životního cyklu dokumentů

Náklady životního cyklu dokumentů jsou náklady, které jsou přímo spjaty s procesy nutnými k ochraně digitálních objektů. Celý proces lze rozložit do několika dílčích fází. V nejnovější verzi modelu LIFE jsou to následující procesy:

$$L_T = C + Aq_T + I_T + BP_T + CP_T + Ac_T$$

Obrázek 2: Etapy životního cyklu dokumentů model LIFE verze 2

L_T (Lifecycle) = Životní cyklus, v tomto modelu se jedná o kompletní shrnutí fází a procesů dokumentu v průběhu času T. Podrobněji v tabulce Obrázek 3: Model LIFE verze 2 [Hole, 2010].

C (Creation or Purchase) – Pořízení viz kapitola Pořízení (Creation or Purchase)

Aq (Acquisition) – Akvizice viz kapitola Akvizice (Acquisition)

I (Ingest) – Deponování viz kapitola Deponování (Ingest)

BP (Bit-stream Preservation) – Ochrana Bit-streamu viz kapitola Ochrana bit-streamu (Bitstream preservation)

CP (Content Preservation) – Ochrana obsahu viz kapitola Ochrana obsahu (Content Preservation)

Ac (Access) – Přístup viz kapitola Zpřístupnění (Access)

Fáze životního cyklu	Pořízení	Akvizice	Deponování	Ochrana Bit-stramu	Ochrana obsahu	Přístup
	Prvky životního cyklu	Digitalizace	Výběr	Zajišťování kvality	Správa uložiště	Sledování
....		Dohoda s vydavatelem	Metadata	Zajištění uložiště	Plánování ochrany	Řízení přístupu
....		Práva duševního vlastnictví a licence	Uložení	Obnovování, renovační migrace	Ochranná činnost	Podporad uživatelů
....		Objednávání a fakturování	Update záznamů	Zálohování	Opětní deponování	
		Získávání	Referenční odkazy	Inspekce	Odstranění	
		Kontrola				

Obrázek 3: Model LIFE verze 2 [Hole, 2010].

Softwarový nástroj, který měl vzniknout ve třetí etapě projektu LIFE₁, je založený na tomto konceptuálním schématu. Umožňuje zahrnout větší počet modelových situací pro vypočítání nákladů na životní cyklus různých typů nosičů, například zvukových, webových, e-časopisů, atd.

2.1.2.1 Pořízení (Creation or Purchase)

Existují tři hlavní zdroje digitálních objektů určených k ochraně:

- **Vytvoření** – objekty jsou vytvořené v rámci organizace.
- **Nákup** – objekty jsou nakoupeny, nebo jejich využití podléhá licenci.
- **Darování** – objekty získala organizace darem, bez jakýchkoliv nákladů.

Tato fáze je v modelu LIFE oddělená od ostatních a je doplňková, vzhledem k velkým rozdílům v nákladech. Pokud však organizace digitální objekty vytváří (v modelu se hovoří o tzv. prvku Digitisation), pak je nutné tyto procesy identifikovat a definovat.

2.1.2.2 Akvizice (Acquisition)

Akvizice představuje počáteční fázi získávání a zpracování digitálních objektů před deponováním do digitálního úložiště. Akviziční procesy se týkají managementu sbírek, administrace a získávání samotných objektů.

- **Výběr** (Selection) – klíčová etapa, ve které se rozhoduje, jaké materiály by měly být získány. Výběr by měl záviset na poslání organizace, účelu, ke kterému mají digitalizované materiály sloužit a hodnotě materiálu. Tento proces realizuje tým složený z managementu, specialistů na dané dokumenty, či obsah, zaměstnanců provozu a zaměstnanců z oblasti ochrany dokumentů.
- **Dohoda s vydavatelem** (Submission agreement) – v této fázi se uzavírá dohoda s případnými dodavateli digitálních objektů. Jsou dojednány podmínky z obou stran a detaily, jako např. očekávané formáty digitálních objektů, způsob přenosu těchto materiálů, a procedury nutné pro migraci.
- **Práva duševního vlastnictví a licence** (IPR & Licensing) – tento proces zahrnuje zkoumání a vyjednávání práv duševního vlastnictví, přístupu a ochrany digitalizovaných objektů. V některých případech mohou být práva sjednána prostřednictvím licenční smlouvy. Jednání s majiteli duševních práv je nezbytné a pro nakládání s digitalizovanými objekty je jedním ze základních předpokladů.
- **Objednávání a fakturování** (Ordering & Invoicing) – administrativní proces zahrnující objednávání, fakturování a placení digitálních objektů.
- **Získávání** (Obtaining) – proces přesunu digitálních objektů od zdroje k instituci zajišťující dlouhodobé uložení.
- **Kontrola** (Check-in) – při tomto procesu dochází ke kontrole přesunovaných materiálů. Nejedná se o kontrolu kvality dat, ale o ujištění se, že objednané materiály skutečně dorazily v uvedeném množství.

2.1.2.3 Deponování (Ingest)

Deponování zahrnuje procesy hodnocení a analýzy digitálních objektů a poté jejich deponování do digitálního repozitáře.

- **Zajišťování kvality** (Quality Assurance) – v této fázi se digitální objekty zkoumají z hlediska kvality. Pokud není dosaženo předpokládané kvality, pak je nutné zvážit změnu strategie. Objekty mohou být opraveny, znovu získány, nebo mohou být připojena metadata popisující kvalitu. Tento proces také obvykle zahrnuje kontrolu digitálních objektů na přítomnost virů.
- **Metadata** (Metadata) – proces identifikace, extrahování a nahrávání metadat popisujících obsah. Charakteristické je zkoumání obsahu za účelem porozumět jeho technickým vlastnostem. První fází tohoto procesu je obvykle charakterizace formátu digitálního obsahu. Extrahování metadat je proces identifikování metadat z obsahu a z existujících metadat spojených s obsahem.
- **Uložení** (Deposit) – proces vkládání digitálních objektů do úložiště a další související činnosti.
- **Update záznamů** (Holdings Update) – update záznamů při zpřístupnění nového digitálního objektu.
- **Referenční odkazy** (Reference linking) – proces přidávání nebo updatování informací použitých v systémech pro vyhledávání v digitalizovaných objektech

2.1.2.4 Ochrana bit-streamu (*Bitstream preservation*)

Tento proces se týká ukládání a údržby digitálních objektů v průběhu času. Zajišťují, aby nedocházelo ke ztrátě nebo poškození bitů tvořících tyto objekty. Samotné úložiště neposkytuje dostatečnou ochranu bit-streamu. Ochranu bit-streamu v úložišti podporuje efektivní řízení, zálohování a pravidelné kontroly.

- **Správa úložiště** (Repository Administration) – správa úložiště s dalšími úkoly spojenými s poskytováním bit-streamové ochrany.
- **Zajištění úložiště** (Storage Provision) – proces ukládání digitálních objektů s možností jejich vyhledání na základě dotazu. Tento proces zahrnuje podporu a údržbu úložiště.
- **Obnovování, renovační migrace** [Cubr, 2007] (Refreshment) – proces přesunu uložených složek do nového repozitáře pokud současnému repozitáři skončí životnost.
- **Zálohování** (Backup) – proces, při kterém jsou vytvářeny kopie uložených předmětů, aby byla zamezena případná ztráta, poškození, nebo smazání dat.
- **Inspekce** (Inspection) – tato fáze kontroluje, zda digitální objekty lze získat beze ztrát. Obsahuje jak automatickou, tak manuální kontrolu vyhledávání.

2.1.2.5 Ochrana obsahu (Content Preservation)

- **Sledování** (Preservation Watch) – monitoruje kontext, ve kterém ochrana životního cyklu dokumentu probíhá pro budoucí plánování těchto aktivit.
- **Plánování ochrany** (Preservation Planning) – v této fázi jsou zvažovány profily digitálních objektů určených k dlouhodobé ochraně z hlediska využití a dalších požadavků. Posuzuje se dostupné řešení ochrany a rozvíjí se plán pro dlouhodobou ochranu.
- **Ochranná činnost** (Preservation Action) – proces zajišťování dlouhodobé přístupnosti digitálních objektů. Zahrnuje i hodnocení a implementaci softwaru usnadňujícího tyto procesy.
- **Opětné deponování** (Re-ingest) – tato fáze představuje proces deponování migrovaných objektů zpět do úložiště. Může opakovat předchozí fáze:
 - Získávání
 - Kontrola
 - Zajišťování kvality
 - Metadata
 - Uložení
 - Update záznamů
- **Odstranění** (Disposal) – odstranění digitálních objektů z repozitáře, pokud již dále není nutná jejich dlouhodobá ochrana. K tomuto odstranění by mělo dojít v souladu s tzv. *odstraňovací procedurou*.

2.1.2.6 Zpřístupnění (Access)

- **Poskytování přístupu** (Access Provision) – proces poskytování přístupu pro uživatele.
- **Řízení přístupu** (Access Control) – představuje aplikaci akcí nebo technických opatření k zajištění přístupu příslušným uživatelům dle předem sjednaného přístupového práva.
- **Podpora uživatelů** (User Support) – podpora poskytování přístupu uživatelům, kteří již mají přístup k digitálním objektům. Např. referenční služby, obecná uživatelská podpora a korespondence.

Obecný model uchovávání (Generic preservation model)

Projekt LIFE má, kromě identifikace nákladů pro uchovávání digitálních sbírek, za cíl odhadnout jejich výši. Proto vznikl tzv. Obecný model uchovávání (Generic preservation model), doprovázený excelovskou tabulkou (ve třetí fázi projektu LIFE také webovým nástrojem), která slouží jako nástroj pro odhad nákladů. Tabulka obsahuje vzorová data - formáty, počty objektů atd.

2.2 Hlavní prvky určování celkových nákladů

- **Sledování technologie** (Technology Watch) – náklady monitorovacích nástrojů, standardů a dalších technologií souvisejících s konkrétním formátem (TEW) za určité časové období (y). To je sníženo o POC, náklady na technologie, není nutno v jednotlivých institucích duplikovat.
 - $\text{Technology Watch} = \text{TEW} * y * \text{POC}$
- **Náklady na nastavení nástrojů** (Tool Setup Cost) – náklady na přípravu konkrétního nástroje, které nezahrnují náklady na jeho pořízení. Tyto náklady jsou založeny na konstantě COA. Opět jsou náklady sníženy o POC.
 - $\text{Tool Setup Cost} = \text{COA} * \text{ENP} * \text{POC}$
- **Plánování uchovávání** (Preservation Planning) – náklady na plánovací procesy, které nastanou v období vyčíslování nákladů. Tyto náklady budou vysoké v období, kdy budou tyto činnosti probíhat (PLA), oproti tomu v dalších obdobích budou velmi nízké (PLN)
 - $\text{Preservation Planning} = (\text{PLA} * \text{ENP} + \text{PLN} * (y - \text{ENP})) * \text{POC} * \text{FCX}$
- **Náklady na provedení uchovávacích procesů** (Execute Preservation Action) – jedná se o náklady na provedení uchovávajících procesů během tohoto období. Základem je ENP, očekávaný počet uchovávacích procesů. PON reprezentuje procento fondu, které bude uchováno použitím migračních procesů.
 - $\text{Execute Preservation Action} = \text{ENP} * \text{PON} * n * (\text{FCM} / n + \text{PUM})$
- **Kontrola kvality** (Quality Assurance) – náklady akce zajišťující kontrolu kvality v období vyčíslování nákladů. Základní náklady (BCT) jsou násobeny počtem objektů v bázi dat.
 - $\text{Quality Assurance} = \text{BCT} * n * \text{ENP} * \text{FCX}$

- **Očekávaný počet uchovacích akcí** (Expected Number of Preservation Actions) – náklady na uchovací akce. PON ovlivňuje frekvenci těchto akcí (např. za 10 let proběhne 1,5 akcí).

- $\text{Expected Number of Preservation Actions} = y / (\text{BLE} + 0.1 * y) + \text{PON}$

3 Nástroje pro vypočítání nákladů

3.1 Finanční model jako excelovská tabulka

Finanční model byl vytvořen kombinací excelovské tabulky a Visual Basic. Excelovský sešit obsahuje:

- Vstupní a výstupní list, obsahující vypočítané náklady na všechny etapy.
- Šest listů pro úpravu dat (data refinement), které umožňují uživateli měnit odhady v jednotlivých fázích modelu.
- Šest modelových listů (model sheets), obsahujících finanční modely používané pro výpočet nákladů během životního cyklu dokumentů.

	A	B	C	D	E	F	G	H
1								
2	Burney Digital Collection- Lifecycle Pr							
3	v 2.0							
4								
5	Lifecycle Stages	Practical Explanation	Cost notes	% of Staff time	Year 1 (individual costs)	Year 1	Year 2	Year 3
6								
7	Creation or Purchase							
8						0.00	0.00	0.00
9		Cost to digitise the total archive plus create associated project information	creation of digital archive			189,710.00	129,373.00	129,373.00
10								
11		Sub Total				189,710.00	129,373.00	129,373.00
12								
13	Acquisition							
14	Selection							
15	Selection Policy (policy/procedure)		10% Curatorial Grade B	10.0%	5,556.32	5,556.32	0.00	0.00
16	Selection (action)	Sort Microfilm into correct order	50% Curatorial Grade C, 6 months	25.0%	10,713.15	10,713.15	0.00	0.00
17	Selection Metadata (metadata)							
18	Submission Agreement							
19	Submission Agreement (policy/procedure)	Setting up of contract with third	10% project Grade A	10.0%	6,647.62	6,647.62	6,647.62	6,647.62
20	Negotiation of Submission (action)	Communication and negotiation with producers/depositors.	5% Legal support Grade A	5.0%	3,323.81	3,323.81	0.00	0.00
21	Submission Metadata (documentation)							
22	IPR & Licensing							
23	IPR & Licensing (policy/procedure)	Management of contract agreement with third party	5% Product development Grade B	5.0%	2,778.16	2,778.16	0.00	0.00
24	Negotiation of Rights (action)		5% Product development Grade B	5.0%	2,778.16	2,778.16	0.00	0.00

Obrázek 4: Náhled nástroje pro vypočítání nákladů z fáze LIFE 2 [Hole, 2010a]

Model je vytvořen pro odhad reálných nákladů životního cyklu dokumentu. Je však možné ho využít pro jakési cvičné odhady nákladů, například za účelem zjistit, jak jednotlivé změny v nákladech změni celkové náklady. Uživatel tak zadává data pouze do pěti polí základního vstupního listu s cílem získat přibližný prvotní odhad. Těmito poli jsou:

- Časový rámec projektu.
- Typ uchovávaného materiálu (tištěný, webový, zvukový atd.).
- Zdroje (nakoupení, dary atd.).
- Počet položek, které mají být ročně zpracovány.
- Velikost instituce.
- V případě digitalizace také požadovaná kvalita.

První věc, kterou by měl uživatel udělat, je přizpůsobit model LIFE instituci, pro kterou bude použit. Do listu „Refine Organizational Profile“ upřesní výchozí data pro konkrétní instituci. Uživatel se může zvolit typ úložišť a jejich počet a technologie v nich využitě (pevné disky, enterprisové pásky, flashové paměti, and pay per use (e.g., cloud storage)). Vysokou bezpečnost zajišťuje využití různých úložných metod (storage methods) a dodavatelů (vendors). Zahrnuty by měly být rovněž personální náklady (roční, hodinové, měsíční atd.) a to všude tam, kde se s pracovníky počítá.

Fáze „Creation or purchase“ slouží k vypočítání nákladů na zajištění dokumentů určených k ochraně. V případě digitalizace je zobrazeno 23 položek ve 3 sloupcích pro zachycení malých, středních a velkých projektů. Jednotlivé prvky jsou založeny na pracovních nákladech (počet dní kdy projektový manažer vytváří projekt) a nákladech na samotnou digitalizaci. Tato čísla by měla být zkontrolována a případně nastavena na nulu, pokud žádné náklady nejsou potřeba.

V etapách „Akvizice“ (Aquisition) a „Deponování“ (Ingest) uživatelé vyplňují 35 položek. Náklady zde se týkají hodin, dnů, nebo procenta času personálu v určité roli. V této části je možné experimentovat a zkoušet různé možnosti, neboť je to jedna z etap, kde se dá poměrně významně ušetřit.

Fáze „Bitstream preservation“ umožňuje upravit náklady na správu úložiště, obnovování, zálohování a spravování. Měly by být brány v potaz případné změny související s časem. Je důležité si uvědomit, že dnes užívané technologie nejsou trvalým řešením a nelze předpovědět, jaké budou k dispozici v následujících 20 letech. Instituce by zároveň neměly spoléhat na klesající náklady na úložiště, provozní náklady na uchování mohou v některých případech překročit hodnotu sbírky.

Fáze „Ochrany obsahu“ (Content preservation) byla zjednodušena a přepracována, oproti situaci v době ukončení projektu LIFE2. Každému typu obsahu je přiřazen stupeň heterogenity (vysoký - webové stránky, nízký - tištěné materiály) a komplexnosti (vysoký - MS Word, PDF, nízký - TIFF). Kombinace těchto faktorů se pak použije k určení nákladů migraci obsahu. Uživatelé si mohou vybrat ze tří strategií – nedělat nic, migrovat při deponování (migrate on ingest), migrovat periodicky. V případě zvolení „nedělat nic“ mohou být zadány náklady na emulaci. Tuto strategii zvolila například nizozemská Královská knihovna⁸, která vsadila na emulaci v dlouhodobém horizontu. The Wellcome Library⁹ oproti tomu počítá s tím, že přijímat pouze omezený počet formátů, za účelem stability, bude fungovat i bez emulace. V případě „migrovat při deponování“ jsou náklady na migraci počítány každý rok. Při „migrovat periodicky“ lze zvolit procento položek, které mají být migrovány a dobu mezi migrací. Model umožňuje specifikovat i procento položek, jejichž migrace vyžaduje manuální zásah, spolu s dalšími náklady.

Fáze „Zpřístupnění“ (Access) poskytuje výchozí odhady nákladů na vytvoření, údržbu a správu digitální knihovny. Lze určit, zda se některé náklady budou pravidelně opakovat při obnově systému.

3.2 Webový nástroj

Cílem je usnadnit přístup k modelu LIFE a vylepšit ho. Uživatel například vidí pouze údaje, které jsou pro něj v daném okamžiku relevantní. Tento nástroj byl vyvinut za použití open-source nástroje Symfony a MySQL databáze. Pro snadnou udržitelnost nástroje v budoucnosti bylo nutné zajistit, aby ho mohl spravovat administrátorem bez potřeby zásahu developera. Projekt je posouzen z různých hledisek (kategorie, zdroje, typ organizace) a podle počtu položek zpracovaných každý rok. (viz Obrázek 5: Náhled webového nástroje pro vypočítání nákladů). Tento model může být pojat jako soubor vlastností, které lze využít k popisu projektu [Hole, 2010].

⁸ <http://www.kb.nl/index-en.html>

⁹ <http://library.wellcome.ac.uk/>

Mozilla Firefox

LIFE3 Demo

Google

LIFE Predictive costing tool for digital preservation

Introduction Basic Input **Output** Refine Logout Changelog About

Name: Manuscripts digitisation
 Start year: 2011
 End year: 2020

Creation or Purchase	Acquisition	Ingest	Bit-stream Preservation	Content Preservation	Access
Conceive Digitisation Activity	Selection	Quality Assurance	Repository Administration	Preservation Watch	Access Provision
£51,000	£630	£1,000	£11,000	£18,000	£0
Digitisation Selection and Preparation	Submission Agreement	Metadata	Storage Provision	Preservation Planning	Access Control
£17,000	£390	£0	£2,100	£18,000	£0
Transport	IPR & Licensing	Deposit	Refreshment	Preservation Action	User Support
£0	£210	£500	£0	£14,000	£0
Digitisation	Ordering & Invoicing	Holdings Update	Backup	Re-ingest	
£40,000	£580	£670	£45	£0	
Digitisation Quality Assurance	Obtaining	Reference Linking	Inspection	Disposal	
£890	£600	£0	£0	£0	
IPR	Check-In				
£0	£680				
£110,000	£3,000	£2,200	£14,000	£51,000	£0
Total: £180,000.00					

UCL British Library Humanities Advanced Technology & Information Institute JISC research information network

Obrázek 5: Náhled webového nástroje pro vypočítání nákladů [Hole, 2010a]

4 Závěr

Cíle vytyčené na začátku projektu nebyly zcela splněny. Hlavními důvody byla nedostatečná komunikace ze strany zainteresovaných lidí v projektu LIFE. A i přes oficiální oznámení na webových stránkách projektu, že aplikace je přístupná a v případě zájmu je možné si tuto aplikaci ve formě online aplikace nebo offline excelovského sešitu stáhnout nebo o ní požádat, se nám nepodařilo i přes mnohanásobné kontaktování se k těmto aplikacím dostat.

Bylo kontaktováno zákaznické centrum Britské knihovny:

Digital preservation department

<http://www.bl.uk/aboutus/contact/index.html>

Customer-Services@bl.uk

+44 (0)1937 546254

Byl kontaktován LIFE Team

LIFE Team

life@bl.uk

Byl využit automatický formulář „Feedback on LIFE project“

Feedback form

<http://www.life.ac.uk/feedback/>

Byly kontaktovány hlavní osoby v tomto projektu. Opětovně emailem a opětovně telefonicky

Brian Hole

LIFE3 Project Manager, The British Library

http://www.life.ac.uk/team/bio_hole.shtml

<http://brianhole.com>

Tel: +442074127152

Tom Pollard

Datasets Outreach Officer (Kolega Biana)

T: +44 (0) 207 412 7767

M: +44 (0) 750 012 6200

E: tom.pollard@bl.uk

www.twitter.com/DatasetsBL

Paul Wheatley

Digital Preservation Manager, The British Library

http://www.life.ac.uk/team/bio_wheatley.shtml

paul.wheatley@bl.uk

Byly kontaktovány vedlejší osoby spolupracující na projektu LIFE.

Li Lin

LIFE3 Financial Analyst and Modeller,

The British Library

Li2518@126.com

Li2518@Hnu.cn (nefunkční email)

Tel: 0731-88821540, 13973188887

I přes neúspěch získat reálnou aplikaci k testování, vznikla přehledná studie vycházející z veřejně dostupných dokumentů, která překládá a popisuje jednotlivé kroky a postupy v rámci metodiky pro sledování nákladů životního cyklu dokumentu. I přesto, že nebylo z výše uvedených důvodů testovat ani jednu ze dvou aplikací je tato studie dobrým podkladem pro hlubší studium a napomáhá k hlubšímu pochopení metodiky pro sledování nákladů životního cyklu dokumentů v prostředí paměťových institucí, kam Národní knihovna v Praze bezesporu patří.

Budeme nadále sledovat situaci okolo projektu LIFE, protože se jedná v současné době o nejrozvinutější a největší projekt svého druhu v této oblasti výzkumu. V případě, že se nám podaří získat alespoň jednu, ze zmiňovaných aplikací, bude tento dokument sloužit jako základ pro hlubší analýzu aktuální implementace aplikace.

Literatura

- AYRIS, P. ; DAVIES R. ; MCLEOD R. ; MIAO, R. ; SHENTON, H. ; WHEATLEY, P. 2008. [PDF soubor]. *The LIFE2 final project report*. London : LIFE Project, 2008 [cit. 2011-08-12]. Dostupný z WWW: <<http://discovery.ucl.ac.uk/11758/>>.
- CUBR, Ladislav. *Dlouhodobá ochrana digitálních dokumentů*. 1.vyd. Praha : Národní knihovna České republiky, 2010. 154s. ISBN 978-80-7050-588-5 (brož.).
- HOLE, Brian. 2010a [PDF soubor]. *LIFE3: Predictive Costing of Digital Preservation*. London : LIFE Project, 2010. [cit. 2011-08-12]. Dostupný z WWW: <http://www.life.ac.uk/3/docs/Hole_pasig_v1.pdf>.
- HOLE, Brian ; WHEATLEY, Paul ; LIN, Li ; MCCANN, Patrick ; AITKEN, Brian. The Life3 Predictive Costing Tool for Digital Collections. In *New Review of Information Networking* [online]. Volume 15, Issue 2, October 2010b, p. 81-93 [cit. 2011-09-15]. Dostupný z WWW: <http://brianhole.com/download/hole2010_life.pdf>. ISSN 1361-4576>.
- MCLEOD, R. ; WHEATLEY, P. ; AYRIS, P. 2006a. [PDF soubor] *LIFECycle information for e-literature: a summary from the LIFE project*. London : LIFE Project, 2006. [cit. 2011-08-12]. Dostupný z WWW: <<http://discovery.ucl.ac.uk/1855/1/LIFEProjSummary.pdf>>.
- MCLEOD, R. ; WHEATLEY, P. ; AYRIS, P. 2006b. [PDF soubor]. *LIFECycle information for e-literature: full report from the LIFE project*. London : LIFE Project, 2010. [cit. 2011-08-12]. Dostupný z WWW: <<http://discovery.ucl.ac.uk/1854/1/LIFEProjMaster.pdf>>.
- SHENTON, H. Life Cycle Collection Management. In *LIBER Quarterly*. Volume 13, Issue 3-4, 2003 , 2003, p. 254-272 [cit. 2011-10-20]. Dostupný z WWW: <<http://liber.library.uu.nl/publish/articles/000033/article.pdf>>.