

3 Výzkum a ověření reálnosti, stanovení finančních potřeb a identifikace rizik u modelu navrhovaného zadavatelem (SWOT analýza)

Při sledování projektů, standardů a doporučení uvedených v kapitole 2 a dále při komunikaci se subjekty, které úschovu ke své činnosti potřebují nebo plánují se ukazuje, že je velmi důležité se nejprve domluvit na **vhodné terminologii** a především zjistit přesný **účel prováděné či plánované úschovy**.

Ukazuje se, že v této souvislosti běžně používaný **obecný termín archiv resp. digitální archiv**, je pro okruh odborníků z oblasti archivnictví, mimo jejich obor, nepřijatelný. Tato skupina poukazuje na to, že slovo archiv má již svůj přesný terminologický význam, který se vztahuje pouze k uchovávání, ochraňování, evidování a zpřístupňování historicky cenných pramenů, vzniklých v rámci činnosti státních, městských a obecních orgánů či jiných právnických a fyzických osob. Tato uchovávaná data v elektronické podobě jsou pak *archiválie v digitální podobě*. Tím je také přesně dán účel úschovy. Nelze tedy podle tohoto přístupu používat termín archiv jako název „datového skladu“, který slouží pro jiné účely, zejména při zpracování a úschově dat u libovolného subjektu v době jejich zúřadování (jako příklad uveďme úschovu dat zpracovávaných v elektronické podatelně, spisové službě, úložišti elektronických dokumentů určených pro přímé použití) resp. uchovávaných po dobu, která je dána obecnými předpisy a během níž mohou uložená data sloužit jako důkaz, že byla zúřadována správně či, že vůbec byla zúřadována (např. úschova elektronických faktur / daňových dokladů v digitální podobě).

V České republice se předpokládá zřízení Národního digitálního archivu v roce 2010. Harmonogram výstavby byl upřesněn usnesením vlády č. 500/2006 Sb. V souvislosti s tím by měla být stanovena metodika pro realizaci dlouhodobého ukládání a archivaci elektronických dokumentů (archiválií) a to včetně strategie výběru formátů. Ministerstvo spravedlnosti a Ministerstvo vnitra by rovněž mělo provést analýzu právního prostředí pro dlouhodobé ukládání a archivaci elektronických dokumentů a navrhnout kroky k zajištění právní validity těchto dokumentů. Doufáme, že by vhodným podkladem mohla být i tato zpráva, která se touto problematikou zabývá.

Při této příležitosti je vhodné připomenout, že někteří odborníci na archivnictví (včetně českých expertů Ministerstva vnitra) tvrdí, že není prakticky možné zajistit autenticitu dokumentu v elektronické podobě a že tam, kde je nutné zachovat dlouhodobě autenticitu dokumentu (např. jako důkaz o dřívějším zúřadování) je nutné mimo digitální podoby uchovávat i podobu papírovou. Poukazují na to, že zaručený elektronický podpis neřeší problém autenticity, protože dlouhodobě uchovávat elektronický podpis není možné a v žádném případě jej proto nelze použít při dlouhodobém uložení. S tímto názorem lze polemizovat (viz metody a připravované standardy pro dlouhodobé ověření elektronického podpisu např. metody DVCS, TAP, TAA, atd. – podrobnější rozbor této problematiky je v kapitole 2.3.5.). Je však pravdou, že při použití těchto metod jsou potřebné další významné náklady na zachování a udržování možnosti ověřit podpis (zvláště v „masovém měřítku“). Tyto

náklady mohou být z dlouhodobého hlediska velmi vysoké a pravděpodobně budou přesahovat význam podpisu, který se u valné většiny archiválií vyskytuje.

Při budování důvěryhodného dlouhodobého úložiště elektronických dokumentů tedy ve většině „běžných“ případů nepůjde o digitální archiv ve výše uvedeném smyslu, i když se pojem archiv vžil i pro tyto ostatní případy, kdy se dokumenty v elektronické podobě uchovávají z různých příčin po delší dobu.

Nejjednodušší formou datových skladů jsou úložiště, která slouží pro obnovu dat v případě jejich ztráty nebo poškození v hlavním IS. Tedy klasické zálohovací systémy. Požadavky na tyto systémy jsou však odlišné od požadavků, které byly popsány v prvních dvou kapitolách. Plyne to zejména z toho, že tato data se neuchovávají po dlouhou dobu a předpokládá se jejich použití v původním formátu a odpadá tak problém s konverzí datových formátů, otázka původu a integrity se sice u těchto dat částečně řeší, ale není prioritní. Těmito systémy jsme se proto nezabývali a nejsou uvedeny v seznamu projektů, standardů a norem.

Vyčlenit systémy pro úschovu dat za účelem zálohování dat a jejich obnovu od jiných systémů, kdy se data z důvodu zúřadování (zpracování) musí uschovávat po dlouhou dobu, zpravidla není problém a k záměně těchto systémů nedochází. Tyto „ostatní“ systémy se často zjednodušeně a ne zcela přesně nazývají elektronický nebo digitální archiv, ačkoliv s archivem ve smyslu úložiště elektronických archiválií nemají nic společného. To často vede k určitým nedorozuměním a je proto nutné tyto pojmy (hlavně účel úschovy a práce s uloženými daty) přesně definovat.

Zvláštní a významnou kategorii tvoří datová úložiště elektronických faktur, které se z důvodů, které vymezuje zákon o DPH (Zákona o DPH č. 235/2004 Sb. v aktuálním znění) musí uschovávat po dobu deseti let (údaj platí pro ČR, v jiných členských zemích EU je tato lhůta odlišná a většinou kratší). Po celou dobu úschovy je potřeba zachovat původ dat a jejich integritu. Otázka formátu a konverze není v těchto projektech podstatná, neboť je dána formátem výměny elektronických faktur (EDIFACT, PDF, XML) a předpokládá se, že po dobu požadované úschovy bude tento formát stále prakticky použitelný a nebude tedy použita konverze. U úschovy faktur, které existují pouze v elektronické podobě a slouží jako daňový podklad, je však významným atributem zaručený elektronický podpis nebo elektronická značka. Po celou dobu úschovy je proto potřeba zajistit jeho ověřitelnost. Tím se tyto úložiště odlišují od projektů, které se soustřeďují na pouhou úschovu dokumentů se zachováním jejich integrity. Výjimkou v této kategorii jsou datová úložiště elektronických faktur v EDI formátech, kde se pro důkaz původu využívají jiné metody než je elektronický podpis a to zejména archivace logů transakčního systému v kombinaci se zachováním integrity a využitím nějaké vhodné metody zakotvení informace v čase (časové razítko nebo časová značka). Vzhledem k prosazování elektronické fakturace význam těchto „archivů“ elektronických faktur roste.

Dalšími významnými systémy pro úschovu dat jsou systémy elektronické správy dokumentů EDMS (Electronic Document Management System) a systémy elektronické spisové služby ERMS (Electronic Records Management Systems). Vzhledem k provázání těchto dvou systémů se o nich zde zmiňujeme jako o jedné z dalších možných kategorií řešení, které využívají úschovu (často i dlouhodobou) elektronických dokumentů.

EDMS se v organizacích široce používají pro správu a kontrolu elektronických dokumentů. Mnoho funkcí a služeb EDMS se překrývá s ERMS. EDMS obvykle

zahrnuje indexování záznamu, organizaci paměti, správu verzí, úzkou integraci s publikačními aplikacemi a vyhledávacími nástroji pro přístup k dokumentům. Některé systémy ERMS zajišťují úplnou funkcionalitu EDMS, jiné jen některou část. Některé EDMS naopak obsahují základní funkce spisové služby.

Následující tabulka ukazuje typické odlišnosti.

EDMS	ERMS
umožňuje dokumenty upravovat a/nebo vést v několika verzích;	chrání dokumenty před upravováním;
může umožnit, aby dokumenty jejich vlastníci vymazali;	chrání dokumenty před vymazáním, s výjimkou přísně regulovaných okolností;
může obsahovat některé kontroly uchovávání;	musí obsahovat přísnou kontrolu uchovávání;
může obsahovat strukturu uložení dokumentů, kterou mohou kontrolovat uživatelé;	musí obsahovat přísnou strukturu uspořádání dokumentů (schéma třídění), které udržuje správce;
je především určen k podpoře každodenního používání dokumentů pro probíhající činnost.	může podporovat každodenní práci, ale je rovněž určen k zajištění bezpečného uložení pro trvalé provozní dokumenty.

Na tyto systémy lze uplatnit řadu požadavků, které byly rozpracovávány a studovány v předešlých kapitolách. ERMS systémy a příslušná úložiště dokumentů jsou pravděpodobně nejrozšířenějšími systémy úschovy dat, které se v současné době s různým úspěchem budují.

Vzhledem k rozšíření a významu těchto systémů vznikl přirozený požadavek vyvinout vzorovou specifikaci funkčních požadavků na systémy ERMS. Potřebu takovéto specifikace konstatovalo jako první Fórum DLM, organizace úzce spolupracující s Evropskou komisí, která se zabývá všemi aspekty nakládání s dokumenty v digitální podobě. Model těchto funkčních požadavků (MoReq) proto vznikl v rámci programu komisariátu Evropské komise IDA (Interchange of Data between Administrations – vzájemná výměna dat mezi exekutivami). Zdá se, že právě toto evropské doporučení by mohlo ve své právě připravované druhé verzi (MoReq2) sehrát významnou úlohu při unifikaci jedné konkrétní třídy systémů na zpracování dat, kde jedním z hlavních požadavků je právě důvěryhodná úschova elektronických dokumentů (resp. i příprava těchto dokumentů pro předání do digitálního archivu, tzv. předarchivní péče).

Systém navržený zadavatelem, který má být v této kapitole diskutován, není zcela přesně definován. Tím, že se na několika místech zmiňuje zaručený elektronický podpis si lze dovodit, že se nejedná o klasický digitální archiv ve smyslu úschovy elektronických archiválií. Na druhé straně byl v zadání zdůrazněn aspekt dlouhodobé úschovy dat a tedy nutnost studovat takové aspekty jako je formát dat, jejich transformace, prezentace apod. Zadavatel tedy spíše požaduje analýzu systému, který z výše uvedených systémů je nejbližší ERMS systému s úkolem dlouhodobé úschovy

dokumentů v systému a případně jistou přípravou předarchivní péče. Přístup řešitelského týmu tedy vychází především z těchto požadavků i když v naší studii se snažíme o obecnější přístup, který by především definoval klíčové vlastnosti společné pro různé přístupy k dlouhodobému uchování elektronických dokumentů, ale i vhodně zdůraznil specifika jednotlivých systémů, kde je to potřebné.

Jak je vidět z předchozích dvou kapitol, je činnost a organizace digitálního archivu velmi komplexní záležitostí. Rešerše a analýzy stávajících projektů ve světě ukazují, že různé skupiny k této problematice přistupují z různých stran. Některá řešení jsou budována na velmi pragmatickém základě (viz příloha 6.2), jinde je zvolen spíše holistický přístup. Studium těchto různých projektů a řešení nicméně jasně ukazuje minimálně tři věci:

- 1) Vzhledem k šíři přístupů a přijatých řešení je zřejmé, že v danou chvíli neexistuje jednoznačně a objektivně nejlepší univerzální řešení, které by bylo možné vybrat a doporučit jako základ pro implementaci.
- 2) Většina moderních řešení vychází při definování koncepce OAIS modelu. (viz kapitola 2.3.6.).
- 3) Žádné ze současných řešení není zatím otestováno vůči jednomu z nejpodstatnějších kritérií, tj. jestli opravdu bude naplňovat funkčnost dlouhodobého (> 30 let, > 50 let) důvěryhodného archivu, až se doba fungování archivu bude blížit k této horní hranici.

Nicméně výše uvedené analýzy a závěry přesto poskytují řadu vodítek a kritérií, které je nutné při vytváření jakéhokoliv modelu brát v úvahu.

Tyto kritéria jsou zejména rozebírána v druhé kapitole, kde názvy jednotlivých podkapitol odrážejí nejdůležitější skupiny kritérií:

- Legislativní a normativní předpoklady;
- Technologické podmínky pro vybudování důvěryhodného archivu;
- Organizační zabezpečení archivu;
- Plánování uchování

3.1 Finanční potřeby

Vzhledem k tomu, že detailní finanční analýze modelového pracoviště je věnována celá pátá kapitola, omezíme se zde jen na několik postřehů.

Celková finanční náročnost vybudování a provozování archivu a jeho celkové poslání a zaměření spolu úzce souvisí. Od poslání a zaměření se bude odvíjet charakter uchovávaných dokumentů, jejich objem a předpokládané nárůsty, potřeba dostupnosti a potřeby zabezpečení, redundance, doba úschovy, různé požadavky na prokazatelnost původu atd..

Jak je patrné zejména z druhé kapitoly je důvěryhodný archiv, který by měl fungovat na úrovni státní správy, národní/universitní knihovny, národního archivu velmi komplexní systém. Jádrem tohoto systému může být velmi propracovaný ERMS, nicméně řadu funkcí souvisejících s dlouhodobostí správy uložených záznamů bude nutné doplnit. Pro hrubou představu a ve značném zjednodušení lze říci, že komplexnost IT systému takového důvěryhodného archivu bude odpovídat komplexnosti výkonného ERP systému. Při uvedení tohoto ilustrativního příkladu je nutné upozornit na fakt, že ERP systémy mají potenciální klientelu o několik řádů

větší než systémy pro digitální archivaci a že jednotlivá ERP řešení se konkrétním zaměřením od sebe zdaleka tak neodlišují jako řešení pro dlouhodobou digitální archivaci. Nicméně toto velmi hrubé srovnání poskytuje určité cenové vodítko pro velikost počáteční investice.

Z charakteru provozu archivu vyplývá, že důraz je nutné klást především na bezpečné uchování nepoškozených informací (vlastních dokumentů, ale i metadat) a méně na nepřetržitý provoz a rychlou obnovu provozu. Z tohoto hlediska řešení navrhované zadavatelem v bodě 3 c) se jeví jako postačující. V rámci celkového řešení by ho bylo samozřejmě nutné doplnit o vhodné smlouvy s dodavateli HW, SW a infrastruktury tak, aby v případě kompletního zničení primárního pracoviště bylo možné činnost archivu obnovit na novém místě s použitím záložních kopií, vytvořit nové zálohy, atd..

Při plánování provozu budou největší položkou rozpočtu nákladů lidské zdroje.

Ceny médií soustavně klesají, nicméně množství informací pro digitální uschování roste rychleji. Lze tudíž spíše očekávat nárůst finančních prostředků potřebných pro vhodná média a infrastrukturu. Tyto potřeby je vhodné a nutné usměrnit hned na počátku vhodným vymezením archivní politiky.

Rozpočet archivu mimoto musí počítat s nárazovými náklady, které budou vznikat v souvislosti se správou uložených elektronických dokumentů. Takovéto mimořádné náklady mohou vzniknout např. v souvislosti s migrací formátů, zestárnutím technologií nebo změnou legislativy a je nutné na ně pamatovat při vytváření rezervních fondů.

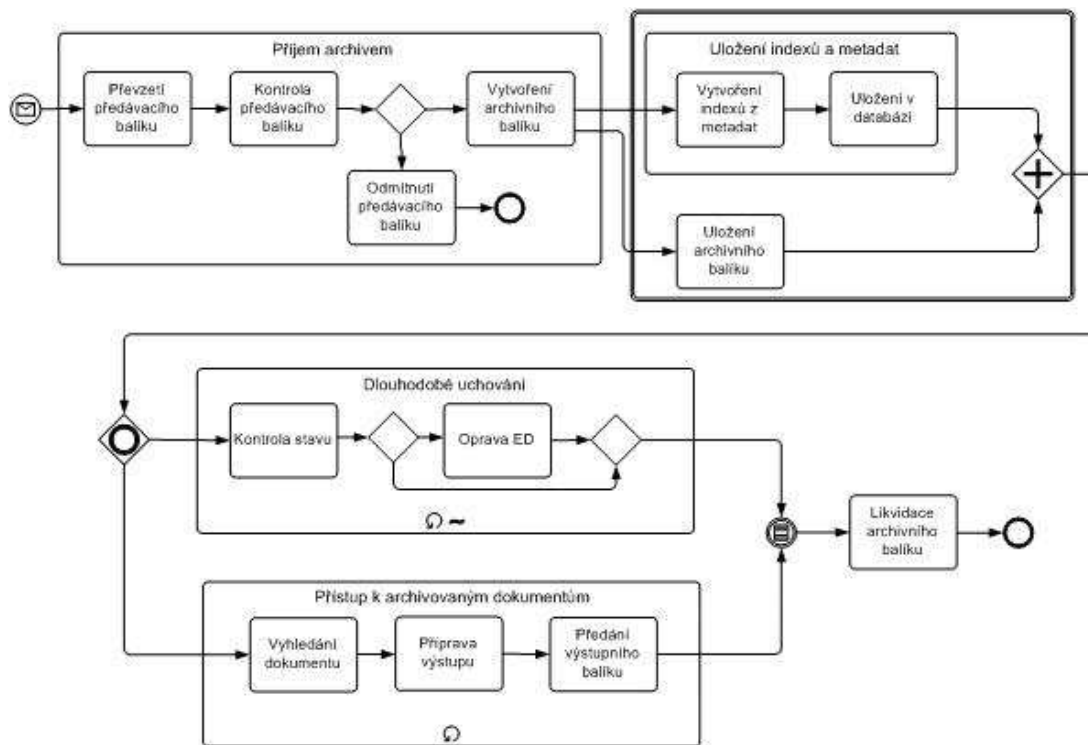
Ke kvantifikaci těchto úvah je zapotřebí mít přesnější zadání a údaje o cílovém zaměření digitálního archivu, než kterými řešitelé v tuto chvíli disponují. Jak již bylo řečeno na začátku tohoto úseku, detailněji se touto problematikou bude zpráva zabývat v páté kapitole.

3.2 Modelové řešení

3.2.1 Modelové workflow/process

Pro potřeby dalších analýz zde navážeme na obecně definované workflow ED z kapitoly 1.4 a definujeme zde úkoly z hlediska dlouhodobého uchovávání ED, které jsou určující pro definování finanční, technické a organizační náročnosti důvěryhodného digitálního archivu (zobecněného podle diskuze v úvodu této kapitoly). Přestože příprava a zpracování elektronického dokumentu u původce zásadně ovlivňuje jeho dlouhodobé uchování, nelze všechny tyto faktory vhodně zobecnit, proto se omezíme pouze na klíčové úkoly archivu s tím, že jsou dány určité předpoklady o předchozím zpracování ED u původce.

Následující obrázek ukazuje modelové workflow pro ED v důvěryhodném archivu:



V následujících bodech rozebereme stručně jednotlivé úkoly (viz též 1.4, kde je další podrobnější diskuze k jednotlivým úkolům):

- **Příjem archivem**
 - **Převzetí předávacího balíku**
Archiv přijímá elektronické dokumenty od původců ve formě předávacího balíku (dávka, která obsahuje dokumenty spolu s metadaty v předepsaném formátu). Původce může předat balík buď na fyzických médiích nebo pomocí elektronické komunikace.
 - **Kontrola předávacího balíku**
Po převzetí provádí archiv kontrolu předávacího balíku, která zahrnuje především formální kontrolu formátu elektronického dokumentu a metadat.
Dále, pokud je předávací balík opatřen zabezpečovacími prvky, jsou tyto ověřeny.
Systém archivu by měl provádět kontroly automatizovaně. Podle možností je také vhodné provádět vybrané namátkové detailní manuální kontroly formátu dokumentů, tak aby bylo zaručeno, že jsou splněny všechny požadavky na kvalitu vstupních dokumentů.
 - **Odmítnutí předávacího balíku**
Pokud jsou zjištěny nedostatky v předávacím balíku, archiv tento odmítne a informuje původce o důvodu odmítnutí. Původce musí opravit zjištěné nedostatky a iniciovat následně nový proces předání dokumentů archivu.
 - **Vytvoření archivního balíku**
Pokud je předávací balík v pořádku, systém archivu jej transformuje do vnitřního formátu ve kterém bude uložen v archivu – archivního balíku. Tento obsahuje dokument ve vstupním formátu, případně v dalších formátech vhodných pro dlouhodobé uchování (kdy konverze

do těchto formátů je součástí vytvoření archivního balíku), všechna příslušná metadata (která jsou jednak získána z metadat dodaných původcem, jednak doplněna v archivu, kdy může být požadováno i manuální doplnění metadat) a archivní balík může být opatřen vhodnými zabezpečovacími prvky.

- Uložení indexů a metadat
 - Vytvoření/extrakce indexů z metadat
Z metadat archivního balíku jsou automaticky vybrány ty informace, které budou potřebné pro vyhledávání a další manipulaci s elektronickými dokumenty v archivu.
 - Uložení v databázi
Informace získané v předchozím kroku jsou uloženy v relační databázi a jsou aktualizovány indexy sloužící pro vyhledávání.
- Uložení archivního balíku
Archivní balík je uložen na média v několika kopiích. V tomto modelu počítáme s magnetickými páskami, kdy jsou vytvořeny 3 kopie, 2 jsou potom skladovány mimo hlavní lokalitu.
- Dlouhodobé uchování
 - Kontrola stavu
Pro dokumenty, které jsou uchovány v digitálním archivu musí probíhat jejich kontrola. Jedná se o nepřetržitou činnost, kdy podle předem definovaného plánu (který je přizpůsoben jednak možnostem archivu, jednak požadavkům na uchování daných dokumentů) jsou jednotlivé uložené archivní balíky kontrolovány na čitelnost a použitelnost obsažených dokumentů.
 - Oprava ED(Archivního balíku)
Pokud jsou při kontrole zjištěny nedostatky, nebo nastane nějaká asynchronní událost (např. předpokládané vypršení použitelnosti určitého formátu dokumentů), je nutné zajistit opravné akce, tyto mohou zahrnovat:
 - Migrace medií (pokud byly zjištěny chyby při čtení z daného media)
 - Migrace technologie (pokud se daná technologie ukládání dat blíží ke konci své životnosti)
 - Migrace nebo emulace formátu (pokud daný formát přestává být využíván a stane se zastaralý)
 - Aktualizace zabezpečení (pokud vyprší platnost nebo účinnost použitých zabezpečovacích prvků)
 - Aktualizace metadat a indexů (pokud se ukáže jako potřebné)
- Přístup k archivovaným dokumentům
Model archivu, se kterým zde pracujeme, předpokládá pouze příležitostný a omezený přístup k uloženým dokumentům. Přístup k dokumentu zahrnuje následující aktivity:
 - Vyhledání dokumentu(archivního balíku) v databázových indexech podle parametrů zadaných uživatelem
 - Příprava výstupu – archivní balík je stažen do dočasného úložiště, kde je konvertován do formy vhodné pro uživatele – výstupního balíku, kdy jednak jsou vybrána potřebná metadata, která budou předána s dokumentem, jednak může dojít ke konverzi dokumentu do formátu

přijatelného uživatelem, nebo k jiným úpravám, které budou potřebné (např. odstranění osobních údajů) .

- Předání výstupního balíku
- Likvidace archivního balíku
Podle našich dosavadních zjištění nepovažuje většina archivů likvidaci archivního balíku za potřebný krok. Předpokládá se, že většinou budou dokumenty uchovávány stále. Nicméně, mohou existovat situace, kdy je potřeba provést likvidaci některých balíčků, proto je vhodné aby existovali pro toto definované procesy a vhodná podpora v rámci archivního systému. Pro určité typy archivů je pak pravidelná likvidace naopak potřebná (např. pro archiv elektronických faktur, kdy nemá smysl uchovávat faktury, které přesáhnou určený časový limit).

3.2.2 IT infrastruktura

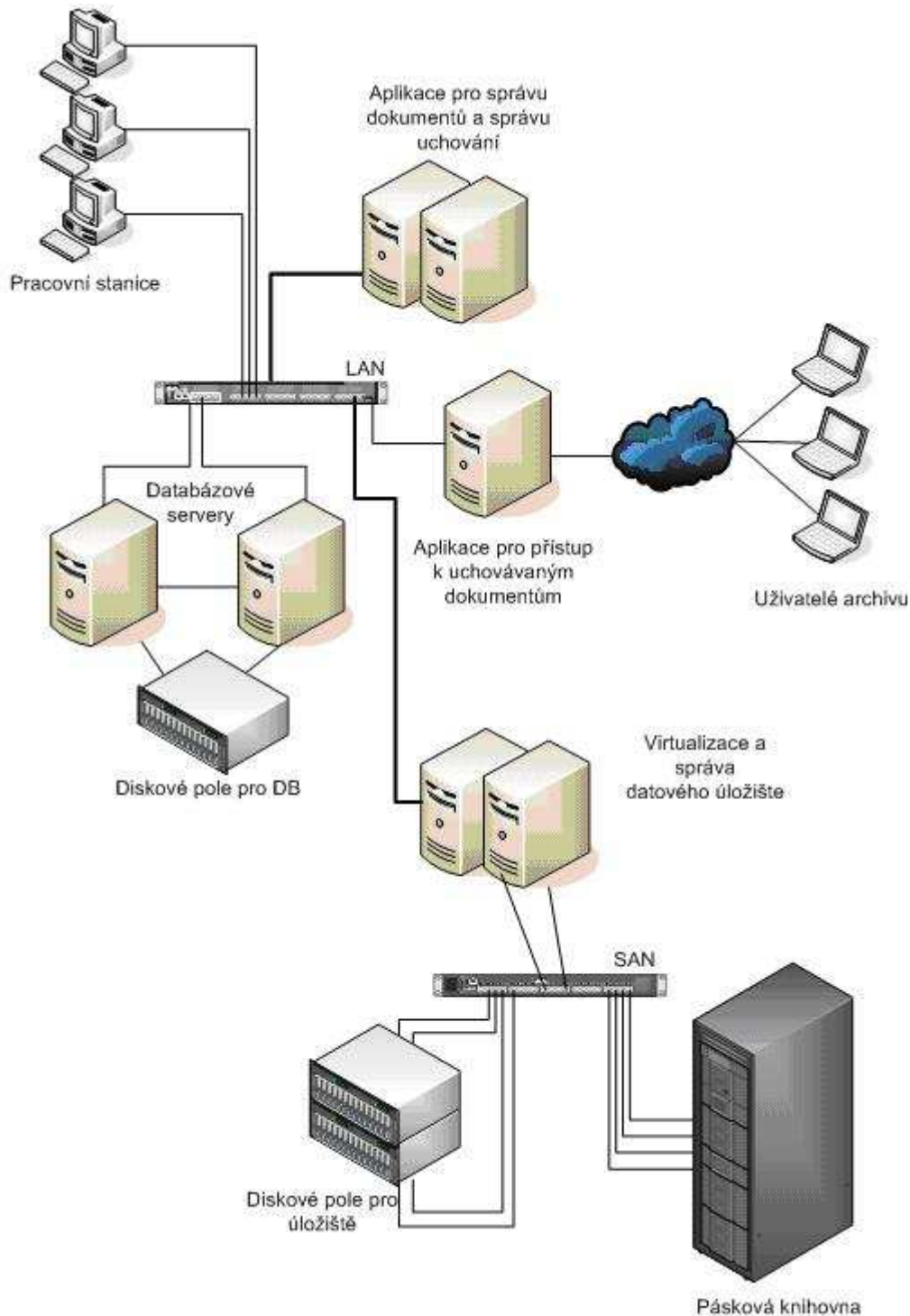
Vlastní IT infrastruktura důvěryhodného archivu bude záviset od řady detailních faktorů vlastní implementace, takže tento pracovní model je možné chápat pouze jako velmi přibližný.

Mezi hlavní faktory, které zásadním způsobem ovlivňují IT infrastrukturu patří:

- objem a roční nárůst uchovávaných dokumentů,
Pro níže uvedený model předpokládáme zhruba roční nárůsty v desítkách TB (s postupným zvyšováním ročních nárůstů) a celková kapacita takového řešení by mohla postačit do cca 500 TB.
- typy uchovávaných dokumentů
Pro níže uvedený model předpokládáme že se bude jednat o textové dokumenty, obrázky, tabulky a databáze.
- uchovávání dokumentů ve více formátech
Pokud budou dokumenty uchovávány ve více formátech (původní formát a standardní formáty, do kterých bude dokument převeden při příjmu) může být kapacita uchovávaných dat až několikanásobně vyšší, než je celkový objem přijatých elektronických dokumentů.
Pro níže uvedený model předpokládáme ukládání elektronických dokumentů pouze v jednom formátu (ve formátu v jakém byly přijaty do archivu nebo dohodnutém standardním formátu, na který je dokument převeden při příjmu).
- hodnota uchovávaných dokumentů (s tím spojené riziko z jejich ztráty a adekvátnost dodatečných opatření pro minimalizaci tohoto rizika)
Pro níže uvedený model předpokládáme že se bude jednat o dlouhodobé uchování běžných dokumentů, které pochází z běžné agendy státních nebo soukromých organizací. Nejedná se o dokumenty, které by měly výjimečnou hodnotu, která by ospravedlňovala mimořádný přístup k jejich uchování.
- požadavky na dostupnost dokumentů
Pro níže uvedený model předpokládáme, že nebude nutná nepřetržitá dostupnost, bude tedy možné odstavit řešení v případě rozsáhlé údržby na několik hodin a v případě mimořádných událostí i na několik dní.
- frekvence přístupu k uchovávaným dokumentům
Pro níže uvedený model předpokládáme, že přístup k uchovávaným dokumentům bude spíše výjimečný. Nepředpokládá se masové zpřístupnění pro širokou veřejnost.

- další požadavky na speciální funkcionalitu archivu, týkající se např. automatizované konverze dokumentů, komplexních schémat zabezpečení dokumentů aj.

Pro níže uvedený model předpokládáme, že kromě potřebných migrací dokumentů pro jejich uchování, bude nutné také provádět konverze dokumentů v případě zpřístupnění, kdy kromě formátu, budou nutné i např. úpravy obsahu (jako je třeba anonymizace dokumentu odstraněním osobních údajů).



Uvedený obrázek ukazuje rámcovou hardwarovou architekturu obecného systému digitálního archivu. Vycházíme z logického modelu, který by popsán v kapitole 2.3.6 a který vychází z modelu OAIS.

Hardwarové komponenty:

- **Servery pro aplikaci správa dokumentů a správa uchování**
Jedná se o klíčovou aplikaci, která zajistí řízení životního cyklu elektronického dokumentu v archivu a poskytuje funkcionalitu potřebnou pro práci s dokumentem a archivním balíkem. Lze předpokládat, že tato aplikace bude implementována na několika serverech, jednak z důvodu architektury aplikace, jednak pro zlepšení dostupnosti a výkonnosti aplikace.
- **Pracovní stanice**
Umožní pracovníkům archivu provádět potřebné operace s dokumenty a archivními balíky a spravovat systémy archivu.
- **Databázové servery**
Tyto servery provozují relační databázi, obsahující doplňující informace o uložených dokumentech potřebné jak pro vyhledávání, tak pro správu archivovaných dokumentů. Dostupnost a integrita této databáze bude klíčová pro celý systém archivu, proto by zde bylo vhodné využít databázový cluster.
- **Diskové pole pro DB**
Diskové pole obsahuje databáze. Předpokládáme, že toto diskové pole bude zaměřené především na výkon. Jeho kapacita bude daná především počtem uchovávaných archivních balíků. Samozřejmostí je pravidelné zálohování této databáze.
- **Server(y) pro přístup k uchovávaným dokumentům**
Zde poběží aplikace/modules, které umožní uživatelům přístup k uchovávaným dokumentům. Uživatelé budou moci pomocí vhodného rozhraní (předpokládáme využití web technologií) vyhledat požadovaný dokument, zadat požadavek na prohlédnutí a případné získání elektronické kopie daného dokumentu.
- **Pásková knihovna**
Slouží jako hlavní zařízení pro dlouhodobé ukládání dat. Jako vhodná technologie se nyní jeví LTO4. Pro každý archivní balík je potřebné vytvořit 3 kopie, každou na nezávislém mediu, jedna kopie je uchovávána lokálně, další dvě v různých oddělených lokalitách. LTO4 umožňuje vytvářet pásky typu WORM (data tedy nelze mazat ani přepisovat), které se jeví velmi vhodné pro tyto účely.
- **Diskové pole pro úložiště**
Toto diskové pole slouží jako dočasné úložiště pro elektronické dokumenty a archivní balíky, které jsou používány nebo budou používány v krátké době. Jedná se tedy především o příjem dokumentů do archivu a vytváření archivních balíků, přístup k archivním balíkům a další činnosti spojené s uchováním. Toto diskové pole bude zaměřené na kapacitu, bude tedy výhodné použít SATA disky v konfiguraci RAID5 nebo RAID6. Kapacitu tohoto pole lze zhruba odhadnout jako 10% kapacity celkově uchovávaných dat.

Předpokládáme, že na popsané HW platformě bude vybudován komplexní IS, který se bude skládat z různých komponent, kdy některé budou standardní SW balíky, další budou moduly vyvinuté přímo pro potřeby daného archivu.

Mezi standardní SW balíky, které jsou zahrnuty v našem modelu, patří:

- **Operační systémy**
Předpokládáme, že serverové aplikace budou preferovat operační systémy UNIX.
- **Databáze**
Nároky na databázový systém budou značné, bude tedy vhodné zvolit ověřený databázový systém, který zvládá bez problémů větší množství záznamu, je maximálně spolehlivý a poskytuje vysokou výkonnost (z komerčních systémů mohou daná kritéria splňovat např. databáze Oracle, Sybase, DB2 ...). Na databázovém systému se rozhodně nevyplatí šetřit.
- **Systémy virtualizace datových úložišť a správa datových úložišť**
Bude se jednat o několik vhodně se doplňujících komerčních balíčků, které poskytnou, jak vhodné zobecnění heterogenního datového úložiště, tak prostředí pro jeho kompletní správu. Podrobnější popis vlastností těchto systémů lze nalézt v kapitole 2.3.6.
- **Systém správy dokumentů**
Pro centrální aplikaci digitálního archivu, která se bude starat o celý životní cyklus elektronického dokumentu, bude výhodné jako základ použít existující komerční aplikaci správy digitálních dokumentů (EDMS/ERMS), která poskytne základní funkcionalitu. Pro potřeby konkrétního archivu, bude nutné provést přizpůsobení a úpravy, proto je nutné zvážit flexibilitu a otevřenost této aplikace jako jedno z nejdůležitějších kritérií při jejím výběru.
- **Web a komunikační servery**
Pro potřeby uživatelského rozhraní a přenosu dokumentů budou součástí řešení digitálního archivu i web a komunikační servery. Pro tyto SW produkty lze využít „open source“ projekty, kde existují kvalitní řešení.
- **Middleware - Aplikační server a případně další SOA aplikace**
Tyto řešení budou sloužit pro vývoj a provozování zakázkových modulů. Také systém správy dokumentů bude velmi pravděpodobně vybudován na nějakém middleware.
- **Vývojové nástroje**
Budou sloužit pro návrh, úpravu a vývoj potřebných zakázkových komponent archivu.

Mezi zakázkové SW komponenty patří:

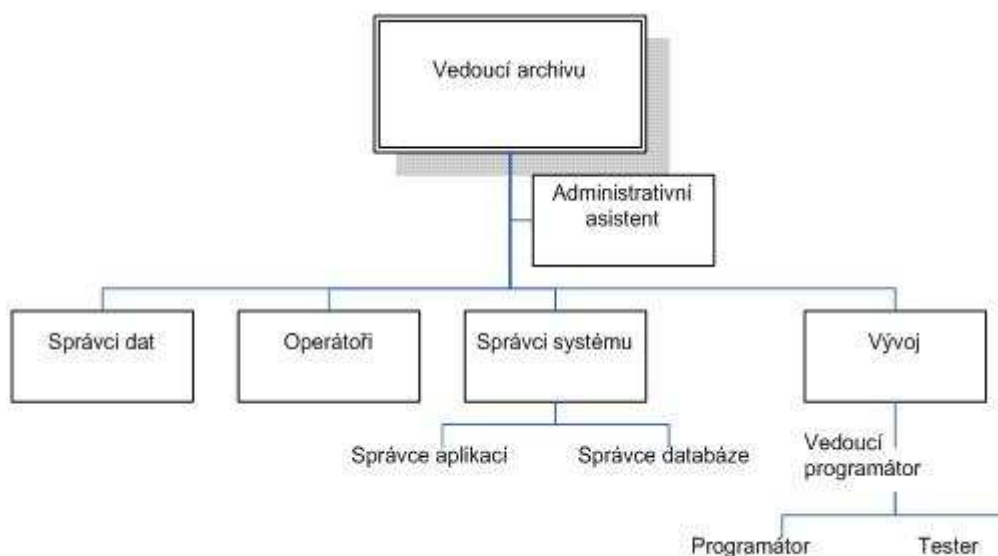
- **Moduly pro systém správy dokumentů a správu archivu rozšiřující správu dlouhodobého uchování**
Dodatečné moduly, navazující na standardní funkcionalitu správy dokumentů a poskytující funkce potřebné pro dlouhodobé uchování elektronických dokumentů.
- **Modul pro příjem a kontrolu archivních balíčků**
Moduly pro příjem a kontrolu elektronických dokumentů budou přizpůsobeny potřebám daného archivu.
- **Moduly pro přístup do archivu**

Tyto moduly opět budou vyžadovat individuální přizpůsobení potřebám archivu.

- Drobné úpravy
Předpokládáme, že kromě výše uvedených modulů budou potřebné další drobné úpravy, jako je např. lokalizace uživatelských rozhraní.

3.2.3 Organizační a personální zabezpečení

Modelová organizační struktura digitálního archivu vychází z výše popsaných rámcových modelů procesů a IT infrastruktury. Následující obrázek ukazuje předpokládanou organizační strukturu:



V rámci této organizační struktury definujeme tyto role:

- Vedoucí archivu
- Administrativní asistent
- Správce dat
- Operátor
- Správce systému
 - Správce aplikací
 - Správce databáze
- Vývoj
 - Vedoucí programátor
 - Programátor
 - Tester

3.3 SWOT analýza oproti původnímu modelu zadavatele

Strengths

Navrhovaný model v sobě obsahuje předpoklady pro velmi dobré zabezpečení archivu a autentičnost uchovávaných digitálních objektů.

Weaknesses

Model představuje pouze menší část celkového řešení problému. Z toho plyne, že zůstává řada otázek a problémů, na které model neodpovídá a tudíž celková podoba výsledného řešení není a ani nemůže být z uvedeného modelu jasná. Tato skutečnost samozřejmě ovlivňuje možnost odpovídajícího celkového zhodnocení. Podrobný a faktický výčet problémů, ke kterým musí model zaujmout stanovisko je v kapitole 2.4 Organizační zabezpečení archivu. Pokud vybereme jen ty nejdůležitější, tak jistě jsou to tyto:

- Model je příliš hrubý, celkově je malá pozornost věnovaná „detailům“. Nicméně právě správné vyřešení celé řady „detailů“ je to, co odliší certifikovaný důvěryhodný archiv od ostatních archivů.
- Návrh se nezabývá charakterem ukládaných informací.
- Model nevěnuje patřičnou pozornost formátu ED a uchování použitelnosti a jeho vztahu k případnému použití elektronického podpisu
- Model nebere v úvahu „stárnutí“ elektronického podpisu
- Tvorba metadat k uloženým informacím není vůbec řešena.
- Chybí vymezení cílové komunity a analýzy jejich potřeb.
- Není řešeno vyhledávání dokumentů v archivu v závislosti na potřebách uživatelů a smysluplná prezentace uložených dokumentů a uchovávaných informací.
- Chybí ošetření všech nestandardních situací.
- Chybí celkové smluvní a právní zázemí:
- Vztah původce – archiv není definován;
- Ošetření práv archivu k nakládání se spravovanými digitálními objekty, otázky duševního vlastnictví, atd.. není řešeno
- Model není založen na OAIS konceptu.

Opportunities

Dopracování modelu do šíře a detailu na základě druhé kapitoly.

Významný je pokus o vyřešení otázky důvěryhodnosti archivu a poskytovaných informací jinými způsoby než je zaručený elektronický podpis. Bližší informace lze nalézt v úvodu k druhé kapitole.

Threads

Nejsou řešeny základní podmínky dlouhodobého uchování

Přílišné spoléhání na zaručený elektronický podpis jako jediný prostředek pro zaručení důvěryhodnosti.

Malá pozornost věnovaná organizačně-právním otázkám.

Volba proprietárního řešení neodpovídajícího obecně přijatým standardům a normám.

Nedostatečné ošetření všech výjimečných a nestandardních situací. Nedostatky v této oblasti by mohly podstatným způsobem ovlivnit vnímání důvěryhodnosti archivu u cílové komunity.