

Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra informačních technologií



Bakalářská práce

Archivace dat v závislosti na retenčních periodách

Radoslav Nagy

© 2015 ČZU v Praze

ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE

Katedra informačních technologií

Provozně ekonomická fakulta

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Radoslav Nagy

Informatika

Název práce

Archivace dat v závislosti na retenčních periodách

Název anglicky

Data archiving based on retention periods

Cíle práce

Bakalářská práce je zaměřena na problematiku archivování dat. Práce bude směřována i na otázku nutnosti retenčních period z legislativních důvodů.

Cílem práce bude navržení vhodného řešení pro různé situace, které v současné době nastávají, zdůvodnění navržených řešení a podrobný popis využitých technik, nástrojů, postupů, hardware, nebo i informací, které jsou k navrženému postupu klíčové.

Metodika

Metodika řešené problematiky bakalářské práce je založena na studiu a analýze odborných informačních zdrojů, praktické sledování a vyhodnocování řešení v některých společnostech. Zjišťování současného stavu a návrh optimalizace řešení. Návrh nutných podmínek a možných řešení pro korektní archivaci z hlediska legislativy bude základem pro formulaci závěru bakalářské práce.

Doporučený rozsah práce

30 – 40 stran

Doporučené zdroje informací

- (ED), Dorian Cougias). The Backup book: disaster recovery from desktop to data center. 3. ed. Lecanto, FL: Schaser-Vartan Books, 2003. ISBN 09-729-0390-9.
- HUMPHRIES, Mark. Data warehousing návrh a implementace Přel. M. Kocan. 1.vyd. Praha: Computer Press, 2001, 257 s. CD. ISBN 80-722-6560-1.
- MOSLEY, Mark, et al. DAMA guide to the data management body of knowledge. Technics Publications, 2010.
- POKORNY, Jaroslav. Databases in the 3rd millennium: trends and research directions. Journal of Systems Integration, 2010, 1.1-2: 3-15.
- VANÍČEK, Jiří; PAPÍK, Martin; PERGL, Robert; VANÍČEK, Tomáš. Teoretické základy informatiky. Kernberg, 2007.

Předběžný termín obhajoby

2015/06 (červen)

Vedoucí práce

Ing. Jan Tyrychtr, Ph.D.

Elektronicky schváleno dne 10. 3. 2015

Ing. Jiří Vaněk, Ph.D.

Vedoucí katedry

Elektronicky schváleno dne 11. 3. 2015

Ing. Martin Pelikán, Ph.D.

Děkan

V Praze dne 11. 03. 2015

Čestné prohlášení

Prohlašuji, že svou bakalářskou práci "Archivace dat v závislosti na retenčních periodách" jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu literatury na konci práce. Jako autor uvedené bakalářské práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

V Praze dne 13.3.2015

Poděkování

Rád bych touto cestou poděkoval Ing. Janu Tyrychtrovi, Ph.D. jakožto vedoucímu této bakalářské práce, slečně Mgr. Dáši Semjanové za právní poradenství a Zdeňku Bohdaneckému za poskytnutí technických informací a možnost konzultování problematiky digitální archivace.

Archivace dat v závislosti na retenčních periodách

Data archiving based on retention periods

Souhrn

Bakalářská práce se zabývá elektronickou komunikací a uchováváním dat. Vysvětluje a kategorizuje problematiku elektronické archivace. Představuje jednotlivé řešení, které plynou z nástrah krátkodobého a dlouhodobého uchovávání dat. Zaměřuje se na část legislativy Evropské unie, především ale České republiky, která se archivací dat dotýká. Vysvětluje problematiku retence a její souvis s legislativou, kdy jako referenční model je vybrána archivace účetních dokladů. Pomocí vícekriteriální analýzy variant, která je v práci krátce vysvětlena, je pak na základě požadovaných kritérií vybráno nejvhodnější řešení pro archivaci dat v závislosti na retenčních periodách.

Summary

The bachelor thesis is about electronic communication and data archivation. It explains and categorizes the issue of electronic archivation. It presents solutions, which stem from the problematics of short-term and long-term data retention. It focuses on European Union legislation, but especially legislation of the Czech Republic, which affects data archiving. It explains the issue of the retention and its relation to legislation. As the reference model for that is chosen the archivation of an accounting documents. Using multi-criteria analysis of alternatives, which is also briefly explained in the work, is selected the best solution for data archiving, based on the required criteria and on retention periods.

Klíčová slova: archivace; data; retenční periody; zákon; datové úložiště; vícekriteriální; skartace; worm

Keywords: archivation; data; retention periods; law; datastore; multi-criteria; data shredding; worm

1	Úvod.....	4
2	Cíl práce.....	5
3	Metodika	6
3.1	Saatyho metoda.....	7
3.2	Metoda pořadí	8
3.3	Bodovací metoda	8
3.4	Metoda AHP	9
4	Digitální obsah.....	11
4.1	Digitalizace v podnikové sféře	12
4.1.1	Datová zpráva	12
4.1.2	E-government.....	12
4.1.3	E-business	12
4.2	Legislativa Evropské unie.....	13
4.3	Legislativa České republiky.....	14
5	Archivace	16
5.1	Základní dělení archivace	16
5.1.1	Krátkodobá archivace	17
5.1.2	Střednědobá archivace	17
5.1.3	Dlouhodobá a trvalá archivace	18
5.2	Objekt archivace	18
5.3	Použitelnost archivovaných dat	19
5.3.1	Virtualizace.....	19
5.3.2	Emulace	19
5.3.3	Migrace	20
5.4	Referenční model OAIS.....	21
6	Retenční periody	23
6.1	Definice.....	23
6.2	Důkazní pohled	23
6.3	Minimální a maximální doby uchování	24
6.4	Retenční periody v České republice	25
6.4.1	Zákon o účetnictví (zákon č. 563/1991 Sb.).....	25
6.4.2	Zákon o dani z přidané hodnoty (zákon č. 235/2004 Sb.).....	26

6.4.3	Zákon č. 582/1991 Sb.	26
6.5	Řešení rozdílné retence	27
6.5.1	Úložiště	27
6.6	Skartace.....	28
7	Archivní řešení.....	29
7.1	Služby v archivnictví	29
7.2	Open Text Archive Server	30
7.3	IBM FileNet.....	30
7.4	Setce eKeeper.....	31
7.5	ELO.....	31
8	Výběr vhodného archivního softwaru.....	33
8.1	Metoda pořadí	33
8.2	Bodovací metoda	34
8.3	Metoda AHP/Saatyho	34
8.4	Porovnání výsledků.....	36
9	Závěr	38
10	Bibliografie	40
11	Přílohy.....	42

1 Úvod

Ubírání se cestou digitalizace je celosvětovou doménou několika posledních desetiletí. Obecně je digitalizace součástí téměř všeho, co je v lidských silách vykonávat, měnit anebo jenom sledovat. Zážitky a vzpomínky jsou uchovávány ve fotografiích a videích v digitálním formátu, sportovní utkání jsou sledované a vyhodnocované senzory, léčba je prováděná různými autonomními roboty, hudba anebo umění už v digitálním světě své místo nesporně mají taky, různé výrobní procesy, výzkum, ekonomická činnost nebo sociální sféra je s digitálním světem už neoddělitelně spjata. Činnost, kterou do digitálního prostředí je možné zařadit, vyžaduje data. Bez dat o digitalizaci není možné jakkoliv uvažovat a nezáleží, jaká ta data jsou. Můžou být krátkodobá, kterých životní cyklus je jen po dobu jejich zpracování a vygenerování dat dalších, nebo konkrétní data, které chceme anebo musíme uchovat napořád.

Zpracování dat je v současné době závažné téma. Jenom jejich samotná existence otvírá otázky a nutí k diskuzi. Stejně tak při zpracovávání dat je odhaleno množství problémů, se kterými je nutné si poradit ve všech úrovních, jak technické, technologické tak i morální. Jelikož data a technologie, které jsou na nich závislé, ovlivňují dění všeho kolem, není možné tyto problémy podceňovat. Nezodpovězené dotazy jenom prohlubují současný stav, na který je nahlíženo jako na výzvu volající pro okamžité řešení. O to se snaží množství společností a výzkumníků, a i když v tuto chvíli není problém jak obecně pracovat s daty zcela vyřešen, praxe a technologické dovednosti ustanovily několik trendů, na kterých je už teď možné stavět a řešit tak situaci do budoucna.

I když je nejvíce známý a diskutovaný problém současných dat jejich růst, není jediný a možná ani nejpodstatnější. Důležitou součástí práce s daty je i jejich archivace. I když je problematika archivace rozvíjena a ustálená už po staletí, není ji možné zcela stejně aplikovat na práci s daty elektronickými.

2 Cíl práce

V podnikové sféře existuje několik důležitých faktorů, které výrazně ovlivňují anebo definují práci s daty. Cílem bakalářské práce je především nalezení vhodného řešení archivace dat, které bude splňovat požadavky legislativní a technologické. Řešení bude určeno na základě předem definovaných kritérií, která budou vyhodnocena pomocí metod vícekritériální analýzy variant. Kritéria budou zvolena na základě zkušeností z praxe, teoretických požadavků pro archivaci dat a pomocí informací, které jsou poskytnuté a prokonzultované s odborníky z daného oboru. Nedílnou součástí archivace je retence, která ovlivňuje možnosti skartace. Dílčím cílem je proto navržení archivace s ohledem na retenční periody u dat, které tvoří podklady pro finanční úřady (faktury, objednávky, jiné) a kterých definice a povinnosti pro archivaci vycházejí z platné legislativy České republiky.

3 Metodika

Řešení této bakalářské práce využívá metodiku, která je založena na analýze dostupných zdrojů, analýze informací a studiu současných trendů. Výsledky jsou interpretované pomocí metod vícekriteriální analýzy variant.

Teorie a model vícekriteriální analýzy variant se zabývá problémy, jak vybrat jednu nebo více variant z množiny přípustných variant a doporučit je k realizaci. V modelech vícekriteriální analýzy (či zhodnocení) variant je dána konečná (diskrétní) množina m variant, které jsou hodnoceny podle n kritérií. Cílem je najít variantu, která je podle všech kritérií celkově hodnocena co nejlépe, variantu kompromisní, případně seřadit varianty od nejlepší po nejhorší nebo vyloučit neefektivní varianty. [1]

Matice představující kritéria (sloupce) a varianty (řádky) v obecném tvaru:

$$\begin{array}{c} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_p \end{array} \begin{pmatrix} \mathbf{f}_1 & \mathbf{f}_2 & \dots & \mathbf{f}_k \\ \mathbf{y}_{11} & \mathbf{y}_{12} & \dots & \mathbf{y}_{1k} \\ \mathbf{y}_{21} & \mathbf{y}_{22} & \dots & \mathbf{y}_{2k} \\ \dots & \dots & \dots & \dots \\ \mathbf{y}_{p1} & \mathbf{y}_{p2} & \dots & \mathbf{y}_{pk} \end{pmatrix}$$

Celkové hodnocení variant závisí na důležitosti jednotlivých kritérií (interkriteriální preference) a na hodnocení variant (intrakriteriální preference). Důležité jsou právě typy informací o důležitosti jednotlivých kritérií a o hodnocení variant podle každého kritéria. [2]

Výchozím krokem analýzy modelu vícekriteriální analýzy variant je stanovení vah. Definovaných metod pro stanovení vah je několik. Nelze obecně určit, která metoda je nejpřesnější, nebo nejvhodnější, vždy záleží od konkrétní situace. Na základě subjektivních nebo objektivních preferencí se určí vhodná metoda, ale nezřídka je na jeden případ těchto metod použitých i několik a to především z důvodu potvrzení si správnosti výsledků. Pro toto ověření správnosti je zvolena metodika, která využívá několik stanovení vah kritérií, konkrétně:

- Saatyho metoda

- Metoda pořadí
- Bodovací metoda
- Metoda AHP

3.1 Saatyho metoda

Jde o metodu kvantitativního párového porovnání kritérií. Po ohodnocení párových porovnání kritérií se používá devítibodová stupnice a je možné používat i mezistupně (hodnoty 2, 4, 6, 8) [1]:

- 1 – rovnocenná kritéria i a j
- 3 – slabě preferované kritérium i před j
- 5 – silně preferované kritérium i před j
- 7 – velmi silně preferované kritérium i před j
- 9 – absolutně preferované kritérium i před j

Na základě tohoto ohodnocení jednotlivých kritérií je možné porovnat každou dvojici kritérií, rovněž i velikost nastavených preferencí i -tého kritéria k j -tému kritériu a zapsat do matice definované následovně:

$$S = \begin{pmatrix} 1 & s_{12} & \cdots & s_{1n} \\ 1/s_{12} & 1 & \cdots & s_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 1/s_{1k} & 1/s_{12} & \cdots & 1 \end{pmatrix}$$

Na diagonále jsou jedničky (každé kritérium je samo sobě rovnocenné), a pro každé kritérium se vypočte, jako geometrický průměr čísel s_{ij} (k -tá odmocnina jejich součtu). [2]

$$b_i = \sqrt[k]{\prod_{j=1}^k s_{ij}}$$

Dále je nutné vypočíst váhy. Ty jsou podílem geometrických podílů a sumy geometrických podílů, resp. normalizací hodnot b_i .

$$v_i = \frac{b_i}{\sum_{i=1}^k b_i}$$

3.2 Metoda pořadí

Pokud je známa ordinální informace při hodnocení dle jednoho kritéria, je možné tuto informaci vyjádřit pořadím variant a to celými čísly mezi 1 a p . P představuje počet variant. Varianta, která je nejdůležitější (nejhodnotnější), bude ohodnocena číslem p (počtem variant), druhou nejlepší číslem $p-1$, atd. až do poslední, té která je nejhorší. Variantám, které jsou rovnocenné, je přiřazeno číslo průměrného pořadí.

K určení důležitosti vah kritérií se metoda pořadí používá při hodnocení několika experty. Kritéria seřadíme nejprve podle pořadí od nejdůležitějšího po nejméně důležité. Nejdůležitější kritérium ohodnotíme k body ($b_i = k$), druhé nejdůležitější $k-1$ body ($b_i = k-1$), atd. až poslední (nejméně důležité) jedním bodem ($b_i = 1$). V případě, že by některá kritéria byla stejně důležitá, obodujeme je příslušným průměrem. [2]

$$yy = \sum_{j=1}^k b_{ij}$$

V dalším kroku pak budou varianty uspořádané sestupně podle hodnot b_i , a nejlepší variantu lze vypočítat ze vztahu:

$$a_1 : b_1 = \max_{i=1, \dots, s} (b_i)$$

3.3 Bodovací metoda

Při kvantifikaci hodnocení variant podle kritéria je nutné nejdříve stanovit bodovou stupnici. Ta může být v rozmezí, které je nejvíce vhodné pro určování pořadí, častokrát se volí rozmezí od 0 do 10. Podle tohoto kritéria se pak hodnotí každé z variant a je tak vyjádřeno určitým počtem bodů. Při maximalizačním typu ohodnocení se nejlepší varianta ohodnotí počtem bodů, který je nejvyšší. V případě minimalizačního ohodnocení je nejlepší varianta ohodnocena nejnižším počtem bodů.

Také tato metoda se používá pro výpočet vah kritérií, hodnotí-li je více expertů. Stejně jako při hodnocení variant se zvolí bodová stupnice a každé kritérium se ohodnotí určitým počtem bodů podle důležitosti, tj. Čím je kritérium důležitější, tím více bodů dostane. Výpočet vah se z bodového hodnocení provede stejně jako u metody pořadí. [2]

3.4 Metoda AHP

Metoda AHP (Analytic Hierarchy Process) byla navržena prof. Saatyem v roce 1980. Tento hierarchický systém je zobecněním – rozšířením možností vícekritériálního rozhodovacího systému. Na každé úrovni hierarchické struktury se použije Saatyho metoda kvantitativního párového porovnání. Pomocí subjektivních hodnocení párového porovnání pak tato metoda přiřazuje jednotlivým komponentám kvantitativní charakteristiky vyjadřující jejich důležitost. Syntézou těchto hodnocení se pak stanoví komponenta s nejvyšší prioritou, na níž se rozhodovatel zaměří s cílem získat řešení rozhodovacího problému. [1]

Metodu AHP je možné použít pro jakýkoliv typ informace o preferenčních vztazích mezi komponentami modelu. Jedinou podmínkou je, aby uživatel uměl z této informace určit směr a intenzitu preference mezi všemi páry porovnávaných komponent. [1]

Základní prvky a kroky metody AHP jsou:

- Konstrukce hierarchie problému
- Párové porovnávání prvků v jednotlivých hierarchických úrovních.
- Syntéza získaných preferencí a volba nejvýhodnější alternativy. [1]

Úloha vícekritériální analýzy variant je složena s následujícími úrovní:

- Úroveň 1 – cíl vyhodnocování, kterým může být uspořádání variant,
- Úroveň 2 – experti, kteří se na hodnocení podílejí
- Úroveň 3 – kritéria vyhodnocování
- Úroveň 4 – posuzované varianty

Obdobným způsobem, jako mezi kritérii při určování vah kritérií Saatyho metodou, lze určit vztahy mezi všemi komponenty na každé úrovni hierarchie. Pokud máme čtyř

úrovňovou hierarchii, tzn. jeden cíl, h expertů, n kritérií a m variant, bude na druhé úrovni hierarchie jedna matice párového srovnávání o rozměrech $h \times h$. Na třetí úrovni bude h matic o rozměrech $n \times n$ a na čtvrté úrovni n matic o rozměrech $m \times m$. Pomocí propočtů (viz Saatyho metoda) v těchto maticích si varianty „rozdělují“ hodnotu váhy příslušného kritéria (kritéria si pak „rozdělují“ váhy příslušného experta). Hodnoty, které získáme, se nazývají preferenční indexy variant z hlediska všech kritérií. Pokud tedy sečteme tyto preferenční indexy z hlediska všech kritérií, získáme hodnocení varianty z pohledu všech expertů a z hlediska všech kritérií. [2]

4 Digitální obsah

Zásadním z klíčových faktorů pro vznik elektronických archivů je rozvoj technologií v oblasti digitalizace dokumentů. Potíže, které přináší práce s papírovými dokumenty, jsou natolik rozsáhlé, že je vhodné zaobírat se jejich digitalizací. Zákon v některých případech jasně stanovuje, jak je možné nakládat s evidencí a správou listinných dokumentů. Proto je v rámci jakéhokoliv subjektu, kterého se zákon přímo dotýká, nutné zabezpečit procesy, které vymezují činnosti archivace a skartace. Tyto procesy jsou pro každý subjekt jiné a nad rámec zákona jsou jejich procesy rozšiřovány pro interní potřeby. V takovém případě představuje digitalizace možnost využití informačních technologií pro snižování nákladů a zvyšování efektivity pro činnost subjektu. V praxi to znamená, že elektronická archivace dokumentů pro subjekty státní správy, neziskové organizace, podniky nebo i jednotlivé živnostníky přináší mnohonásobní úsporu prostorů pro skladování, času a prostředků, které jsou na tuto činnost vyžadované. Pro představu, tuto chvíli znamená elektronická archivace obrovské zhuštění dat do malého prostoru, v průměru je to 300 stran znaků na megabajt na disku [3]. Další užitečnou vlastností elektronických archivů je vyhledávání uložených elektronických dokumentů pomocí vyhledávacích metadat, například dle vystavovatele, datu, částky nebo zboží, které bylo například fakturováno. Velmi podstatný argument je možnost vícenásobného přístupu k dokumentu v jedné chvíli, v případě papírového dokumentu by to znamenalo zdoluhavé vytváření kopií, nebo stanovení čekací lhůty, po kterou by byl dokument rezervován osobou nebo celým pracovištěm. Dalším argumentem jsou oprávnění. Přístup k takovému obsahu je možné hlídat líp, pro každou instanci v archivu je možné specifikovat konkrétní přístupová práva, které v případě papírového dokumentu lze jenom stěží. Oprávnění pro smazání, úpravu, nebo výhradně čtení dokumentu představuje dimenzi, která je pro papír nepředstavitelná. Dokumenty ve formátu dat ani nepodléhají povětrnostním vplyvem a živelným pohromám.

Digitální obsah ale má i několik nevýhod, které v dnešní době brzdí jeho přebírání nadvlády nad papírovým obsahem. Zákony, které vymezují aspekty digitalizace, stále nejsou dostatečně definované. Pokrok v oblasti informačních technologií je natolik rychlý, že v některých případech znemožňuje definování jasných pravidel pro uchovávání obsahu a pro takovéto případy jsou stále vhodnější archivy, kterých pravidla se formovaly po staletí.

4.1 Digitalizace v podnikové sféře

Digitální obsah a způsoby jeho archivace vycházejí primárně z pravidel pro uchovávání papírového obsahu. Evropská unie proto vydala směrnici *1999/93/EC*, která umožňuje využití elektronicky podepsané dokumentů v oblastech státní správy (E-government) nebo při elektronickém obchodování (E-business). Tyto směrnice definují především krátkodobou archivace elektronických dokumentů.

4.1.1 Datová zpráva

Pro zjednodušení obchodního styku byla zavedena definice datové zprávy. Tu definuje zákon „§ 2 zákona č. 227/2000 Sb., o elektronickém podpisu“ a jeho znění je:

- Elektronická data, která lze přenášet prostředky pro elektronickou komunikaci a uchovávat na záznamových médiích, používaných při zpracování a přenosu dat elektronickou formou. [4]

4.1.2 E-government

Elektronická forma výkonu veřejné správy může být nazývána jako E-government. Jelikož definicí pro E-government (nazývaný taktéž eGovernment) je nepřehledné množství a pořad přibývají, jako nejhodnější pro Českou republiku se jeví následující definice:

- eGovernment je využívání informačních technologií veřejnými institucemi pro zajištění výměny informací s občany, soukromými organizacemi a jinými veřejnými institucemi za účelem zvyšování efektivity vnitřního fungování a poskytování rychlých, dostupných a kvalitních informačních služeb. [5]

Pro styk orgánů státní správy, občany a dalšími institucemi může E-government využít elektronický podpis jako adekvátní substituci za ručně psaný podpis. Využití je daleko širší a lze jej využít i jako prostředek pro komunikaci mezi orgány moci a státní správou. Zákonní ustanovení je definované *vyhláškou Ministerstva informatiky č. 496/2004 Sb.*

4.1.3 E-business

Obchodování prostřednictvím elektronického styku je trendem, který se rozvíjí a jeho rozšíření je nejvíce očekáván a považován za největší přínos v soukromé sféře. Podobně jako datová zpráva a e-government je e-business popsán (anebo definován) směrnicí Evropské unie 2000/31/ES, které cílem ne přispění k řádnému fungování trhu v rámci členských států

Evropské unie. Myšlenku „elektronického obchodování“ podpořil rozvoj v oblasti elektronických podpisů, bez kterého nelze obchodování prostřednictvím internetu vykonat. [6]

4.2 Legislativa Evropské unie

Základní pilíř pro vývoj technologií v oblasti elektronického podpisu, elektronické archivace dokumentů a s nimi související digitalizace papírových dokumentů bylo přijetí směrnice Evropské unie 1999/93/EC. Směrnici tvoří čtyři základní body [7]:

- dobrovolný akreditační systém – každý poskytovatel certifikačních služeb získá po splnění legislativních podmínek akreditaci
- každý členský stát uplatňuje svoje vnitrostátní právní předpisy, přijaté na základě tyto směrnice, pro poskytovatele certifikačních služeb,
- členské státy dbají na plnění právních požadavků na elektronický podpis a stejně tak umožňují použít elektronický podpis při soudním sporu,
- členské státy dbají na neodopíratelnost právních účinků elektronických podpisů, i když mají elektronickou podobu anebo se nezakládají na kvalifikovaném certifikátu.

Snaha Evropské unie je evidentně k rozšiřování používání elektronických podpisů, technologiemi s nimi souvisejících, především ve veřejné správě, neméně i v komerční sféře napříč všemi jejími členy. Stavební kámen pro tuto možnost je definování a vyspecifikování toho, co je ještě považované za dokument a co nikoliv. S každým takovým dokumentem ale vzniká problém jak uchovat digitální data a jak následně potvrdit jejich pravost, proto je řešení problematiky uzákoněno a to ve smyslu [6]:

- typy dokumentů, které jsou zákonem určené, se musí archivovat v určitém časovém rozpětí s možností elektronické archivace,
- při elektronické formě archivace je potřeba zaručit neměnnost dokumentu po celou dobu archivace a zajistit čitelnost obsahu

Z pohledu elektronické archivace dokumentů má zavedení legislativních opatření své opodstatnění a k jejich vzniku vedly především tyto důvody:

- v rámci rozvoje v oblasti elektronické archivace bylo nutné přistoupit k ochraně uživatelů a jejich dat,
- Evropská unie určila standardy, legislativní požadavky a návrhy zákonů pro členské státy s cílem ochrany osobních údajů, podpořit a usměrnit elektronickou komunikaci a uchovávání datových záznamů.

Přínosem těchto předpisů je především to, že zřizovatelé elektronických archivů mají v ruce legální a právně podepřený nástroj, bez kterého by vytváření elektronických archivů nemělo požadovaný rozsah přínosů pro jeho zřízení. Z toho proto vyplývá, že dokumenty uložené v takovém elektronickém archívu mají stejnou právní váhu jako v případě fyzického (papírového) nosiče. Podmínkou je však splnění veškerých požadavek, které elektronickou archivaci definují. Definicí jsou pak určeny i standardy, na základě kterých je komunikace mezi subjekty snazší.

4.3 Legislativa České republiky

Česká republika má na základě směrnic Evropské unie zákon, který definuje elektronickou komunikaci a elektronickou archivaci elektronicky podepsaných dokumentů. Zákony doplňují i ustanovení a vyhlášky, demonstrativní výčet dokumentů je následující:

- zákon č. 227/2000 Sb., o elektronickém podpisu a o změně některých dalších zákonů, v znění pozdějších předpisů,
- zákon č. 300/2008 Sb., o elektronických úkonech a autorizované konverzi dokumentů, v znění pozdějších předpisů,
- zákon č. 500/2004 Sb., správní řád, v znění pozdějších předpisů,
- zákon č. 499/2004 Sb., o archivnictví a spisové službě a o změně některých v znění pozdějších předpisů,
- vyhláška č. 193/2009 Sb., o stanovení podrobností provádění autorizované konverze dokumentů,
- vyhláška č. 259/2012 Sb., o podrobnostech výkonu spisové služby,
- vyhláška č. 194/2009 Sb., o užívání a provozování informačního systému datových schránek,
- vyhláška č. 212/2012 Sb., o struktuře údajů, na základě kterých je možné jednoznačně identifikovat podepisující osobu, a postupech pro ověřování

platnosti zaručeného elektronického podpisu, elektronické značky, kvalifikovaného certifikátu, kvalifikovaného systémového certifikátu a kvalifikovaného časového razítka.

5 Archivace

V řečtině slovo archeion, někdy též archeia, označovalo v antice původně úřední instituci nebo její budovu, jakož i samotné písemnosti vzniklé z úřední činnosti této instituce a místo, kde se tyto písemnosti ukládaly. V soudobém českém archivnictví se archiv z hlediska jeho funkcí vymezuje jako informační instituce, která shromažďuje archiválie podle územní nebo věcné příslušnosti za účelem jejich trvalého a bezpečného uchování, jejich odborného zpracování a jejich zpřístupňování pro vědecké, správní, hospodářské a kulturní využití. Archiválie jsou písemné, obrazové, zvukové a jiné záznamy - dokumenty (bez ohledu na způsob záznamu informace), které vzešly z činnosti státních a samosprávných orgánů a jiných právnických osob nebo fyzických osob (obecně původců), a které byly vzhledem ke svému dokumentárnímu významu (trvalé hodnotě) vybrány k trvalému uložení (§2 zákona o archivnictví). [8]

Na základě výše uvedené definice je možné odvodit archivaci elektronickou. Podnik nebo instituce anebo jakýkoliv jiný subjekt může narazit na problematiku elektronické archivace, pokud svou činností jakkoliv pracuje s elektronickými informacemi. Povinnost archivace je v mnoha případech daná zákonem. Například vymezená zákonem č. 499/2000 o archivnictví a spisové službě, především ale i další legislativou, která ukládá povinnost uchovávat data anebo informace. Příkladem může být zákon č. 563/1991 o účetnictví, který definuje povinnost uschovávat účetní záznamy po minimálně stanovenou dobu pro každý typ záznamu. Pokud subjekt eviduje tyto data výhradně v elektronické formě, je povinen zabezpečit náležitou archivaci. Pochopitelně, podobných zákonů je v České republice několik, od oblasti telekomunikací, přes bankovníctví až po zdravotnictví. Závazná jsou i směrnice Evropské unie, která pro archivaci vydává množství doporučení a nařízení.

5.1 Základní dělení archivace

Rozdělit archivaci můžeme z několika pohledů. Základním rozdělením je to, které archivaci definuje už ze své podstaty. Tím je myšleno rozdělení z hlediska času a lze ji definovat na:

- Krátkodobou,
- střednědobou,
- a dlouhodobou až trvalou archivaci.

Toto rozdělení je nezbytné pro určení strategie archivování dat a z ní vyplývajících investic. Jelikož je pro každé zařazení dokumentu doporučený jiný způsob archivace, není možné bez tohoto základního rozdělení určit, jak bude infrastruktura podniku sestavena. Podobně je na tom i jakýkoliv jiný subjekt a může se jednat o množství dat zcela zanedbatelné, které čítá třeba i jen několik jednotek dokumentů anebo i obrovské korporace, které pro ukládání dat využívají samostatné „datawarehouses“.

5.1.1 Krátkodobá archivace

Archivace pro krátkodobé účely může vyžadovat stejné prvky archivace jako při archivaci na delší dobu. Krátkodobé archivace jsou tvořeny převážně daty, které čekají na další zpracování anebo daty, které z hlediska časové potřeby jsou vyžadována jenom v daném okamžiku, nebo řádově dnů až týdnů. Archivace takových dokumentů má důvod hlavně kvůli indexaci anebo jako způsob bezpečného, dočasného, úložiště dat. Problémy s krátkodobou archivací nejsou tak závažné jako při střednědobé anebo dlouhodobé archivaci, zřídka je vyžadováno ukládat tyto data se zabezpečením. I takové situace ale mohou nastat, v praxi lze najít archivaci datových zpráv z datových schránek, které se z objemového hlediska do schránek nevejdou, mají však důležitou funkci při zpracování a je nezbytné zabezpečit to, aby jejich obsah nebyl smazán anebo upraven.

5.1.2 Střednědobá archivace

Předpokládaný čas archivace v střednědobém horizontu se počítá na jednotky let. Pokud se jedná o uložení dat v bezpečném archivu, problém může nastat s procesem zabezpečení těchto dat z pohledu důvěryhodnosti. Certifikáty, které důvěryhodnost zabezpečují, mají limitovanou platnost. Mechanizmy, které s certifikáty pracují, zastarávají, je proto nezbytné data archivované tímto způsobem postupně a průběžně ošetřovat. Pro bezpečnou archivaci je množství teorií a způsobů, jednoznačnou odpověď na správnost bezpečné archivace ale stanovit nelze. Pro každou situaci, typ dokumentu, subjekt a právní normy musí být vykonaná analýza a specifikovaný postup, jakým tato archivace bude vykonaná.

5.1.3 Dlouhodobá a trvalá archivace

Za dlouhodobou archivaci se považuje ukládání a uchovávání dat po dobu desítek let. V některých případech definovaných zákonem č. 499/2004 Sb., o archivnictví a spisové službě není povoleno informace, dokumenty, data, nebo cokoli smazat, nebo zničit vůbec. V případě zabezpečeného archivu to představuje ohromnou výzvu, kterou je nutné řešit. Problémy, které vstupují, jako aspekty do takového řešení nejsou jenom ty, které se týkají střednědobé archivace, nad rámec toho totiž vyvstává otázka, jakým způsobem a v jakém formátu budou data přístupna o 50 nebo 100 let. Dlouhodobou archivaci dat se proto zabývají specializované společnosti, které na archivaci vytvářejí sofistikované postupy a využívají nejmodernějších technologií a poznatků vědy. Tyto archivy se pak označují jako „důvěryhodné archivační autority“ (TAA).

5.2 Objekt archivace

Elektronické dokumenty, data, informace, jsou po vykonání archivačního procesu součástí archivačního objektu. Význam tohoto uskupení je sjednocení možných dat, které s daným dokumentem souvisejí. Příkladem budiž faktura, která je uložena v archivu jako obrázek. Z obrazového zdroje není triviální strojové čtení informací, jako jsou jednotlivé položky faktury, k obrázku jsou proto přiloženy tzv. metadata, které tyto informace obsahují a které byly vytvořeny při zpracování. Patří tam i poznámky, grafické úpravy, informace o průběhu zpracování a množství dalších informací. Do archivačního objektu patří i časové razítko, které dokládá informaci o čase zpracování a archivaci dokumentu. Stejně tak i elektronický podpis, který je zárukou toho, jaký subjekt, případně osoba s daným dokumentem pracovala. Důležitou součástí objektu archivace je případná informace o retenci daného objektu, na základě které archivační systém dokáže s dokumentem pracovat. To všechno může být součástí objektu archivace.

Archivační objekty jsou v archivu udržované tak, že ho po jeho vytvoření udržují kontinuální integritu. Znamená to, že udržují aktuální jeho součásti, jako je časové razítko nebo podpis. Důvod je zabezpečení platnosti důkazových záznamů a je tak učiněno v průběhu celé archivační doby.

5.3 Použitelnost archivovaných dat

Dlouhodobá archivace není jenom o uložení dat. Data, která projdou archivačním procesem, musí být přístupná ale především čitelná po celou dobu archivačního cyklu. Jelikož tempo vývoje informačních technologií je enormně vysoké, může nastat taková situace, že platforma, která k archivaci slouží, nebude podporovaná výrobcem hardware nebo software a po určitém čase nebude možné zarchivované data přečíst. Problémem je taky podpora formátů, ta musí být zabezpečená primárně softwarově, ale navázán je na to i hardware. Pokud nebude splněná tato podmínka, archivace jako taková nebude plnit svůj účel.

Z toho hlediska jsou vytvořené postupy a mechanismy, které umožňují to, že obsah, který byl elektronicky zarchivován, bude možné požadovaným způsobem číst, nebo jej dál zpracovávat. Z praktického hlediska však není moc řešení, které by bylo možné plnohodnotně využít, použitelnost se omezuje na tři technologie, a to:

- virtualizace
- emulace
- migrace

5.3.1 Virtualizace

Pokud je uchován program, který umožňuje interpretaci obsahu v strojovém jazyku „univerzálního virtuálního počítače“, jedná se o virtualizaci. Stejným způsobem ale musí být uchovány i data, které tímto způsobem budou čtené, nebo zobrazované. Tento způsob vyžaduje jednoduchost a otevřenost a všeobecnost. Při přechodu na novou platformu už jenom stačí, aby obsahovala interpret tohoto jazyka. [9]

5.3.2 Emulace

Emulace je simulace původního hardwarového prostředí (architektury, operačního systému) na aktuální platformě. Prostřednictvím emulace je možné zobrazit elektronický dokument v originálním software a to je jeden z možných způsobů, jak s ním pracovat v původní formě. Spolu s emulací je potřeba uchování instalačních souborů původní platformy. Výhodou emulace je možnost uchovávání taky vektorové grafiky a multimediální obsah. Rizikem je ztráta schopnosti emulovat některou vlastnost původní platformy. [9]

5.3.3 Migrace

Nejvhodnějším způsobem pro uchování čitelnosti a možnosti s daty dále pracovat je jednoznačně migrace. Rozumí se tím především převod elektronických dokumentů do jiných formátů, případně na jinou konfiguraci hardwaru. Migrací je možné vykonat i převod mezi softwarovými platformami. Na rozdíl od emulace a virtualizace přináší migrace i řadu problémů. Týkají se toho, jak zabezpečit autenticitu migrovaných dokumentů, pokud je vyžadována. Zároveň může být migrací změněn obsah archivovaných dat natolik, že bude považovaný za znehodnocený a tudíž bude dál zbytečné jej archivovat.

Existují ale časově stálé formáty, které mají svou definici doporučenou normou ISO a stanovené vyhláškou č. 191/2009 §20 Sb., o podrobnostech výkonu spisové služby ze dne 23. června 2009 [10]:

- statické textové a statické textově-obrazové dokumenty – formát PDF/A (norma ISO 19005),
- statické obrazové dokumenty – formáty PNG (norma ISO/IEC 15948), TIFF (revize 6, nekomprimovaný), JPEG (norma ISO/IEC 10918),
- dynamické datové dokumenty – formáty MPEG-2 (norma ISO/IEC 13818), MPEG-1 (norma ISO/IEC 11172), GIF,
- zvukové dokumenty – formáty MP2 (MPEG-1 Audio Layer 2), MP3 (MPEG-1 Audio Layer 3), WAV (PCM modulace),
- metadata dokumentů – formát XML (dle schémat přílohy národního standardu pro elektronické systémy spisové služby)

Použití konkrétních datových formátů souborů je specifikováno v některých dalších právních ustanoveních České republiky, které jsou vhodné pro elektronickou archivaci dat z dlouhodobého hlediska [11]:

- vyhláška č. 259/2012 Sb., o podrobnostech výkonu spisové služby,
- vyhláška č. 193/2009 Sb., o stanovení podrobností provádění autorizované konverze dokumentů,
- vyhláška č. 194/2009 Sb., o stanovení podrobností užívání a provozování informačního systému datových schránek, ve znění vyhlášky č. 422/2010 Sb.

5.4 Referenční model OAIS

Standardizace se dotýká i samotné filozofie archivace. Vytvořen byl proto model, který určuje optimální skladbu archivačního bloku. OAIS je definovaný jako model pro archivní systémy (systém elektronických dat). Pokud je systém navržen v modelu OAIS, má subjekt vlastnící tento systém jistotu nebo záruku, že systém splňuje požadavky na dlouhodobou archivaci.

OAIS je chápán jako procesy vytvářené systémy a osobami, které přebírají odpovědnost za dlouhodobé uchování informací a jejich zprostředkování pro zobrazení. Součástí řešení na základě OAIS je „producent digitálních záznamů“ (původce informací), „management“ (správa dat, archivace a uložení, plánování ochrany během uchovávání) a „koncový uživatel“ (spotřebitel, čtenář) digitálních záznamů. Každá z těchto tří stran má oprávnění vykonávat činnosti a úkony za účelem využití archivu podle OAIS. Tyto činnosti jsou částečně definovány referenčním modelem OAIS, částečně jsou upravené na základě potřeb subjektu, které vyplývají z analýz pro konkrétní situaci, ale vždy s ohledem na dlouhodobé zachování digitálních záznamů.

Data, které jsou pomocí modelu OAIS archivované, jsou chápány jako informační balíčky, definované jako „datové objekty“.

Model OAIS definuje pojem „informační balíčky“. Ty jsou definované ve třech kategoriích. Každá kategorie má vlastní sadu popisných (nepovinných a povinných) informací a každá slouží k účelu zcela jinému (vyhledávání, přijímání a ukládání):

- SIP - soubor informací pro dodávání (soubor podávaných informací) – od producenta /původce
- AIP - soubor informací pro archivaci (soubor archivovaných informací) – v digitálním archivu
- DIP – soubor informací pro šíření (soubor informací pro rozšiřování) – směrem k uživateli /spotřebiteli / badateli

Model je rozdělen do samostatných funkčních jednotek [13]:

- Ingest (recepce)

- Archivní úložiště (storage) – fyzická ochrana bitů tvořících data
- Plánování ochrany (zachování) – monitoring prostředí v OAIS, zaručuje logickou ochranu s cílem zachování funkčnosti souborů
- Správa dat (data management) – správa dat a metadat ke všem akcím u všech digitálních záznamů
- Přístup – umožňuje přístup koncovým uživatelům k informacím v digitálním archivu
- Administrace (správa) – správa a řízení souborového systému (HW, SW), sledování systému, aktualizace obsahu archivu, dodržování standardů, postupů a poskytování nezbytné podpory.

6 Retenční periody

6.1 Definice

Slovem retence je myšleno následující [14]:

- zadržení, zadržování
- schopnost zapamatovat si něco, udržet v paměti
- zachování, uchování

Pojem „Retenční perioda“ proto představuje dobu, přesněji konkrétní časovou délku, po kterou je nutnost držet obsah v nezměněném stavu. Pro dokumenty v elektronickém archivu to znamená, že nesmí být po určitou dobu smazané, nijak pozměněné a v případě, že budou přesouvány, musí být zaručeno to, že se migrací nijak nepoškodí, neznehodnotí anebo úplně neztratí.

6.2 Důkazní pohled

V občanskoprávním řízení, může elektronická verze naskenovaných papírových dokumentů sloužit jako důkaz o obsahu a existenci originálu, pokud je tento zlikvidován. Nicméně, ve všech jurisdikcích Evropské unie, soudy připisují tomuto formátu menší důkazní hodnotu, než by byla hodnota originálu. Důvodem je podezření, že původní záznam mohl být pozměněn ještě před tím, než byl zdigitalizován, nebo by byly provedeny změny v elektronickém záznamu. V obou případech je možnost vykonat změny relativně snadno, ale není snadné tyto změny zpětně detekovat. Další problémy mohou nastat, pokud se elektronický dokument převede na jiný typ souboru, nebo je převeden na jiné datové nosiče. To činí záznam ještě méně čitelným nebo nepřístupným úplně a jeho důkazní hodnota se tak sníží na minimum. Na druhé straně, důkazní hodnotu elektronických záznamů je možné zvýšit tím, že bude u soudu prokázáno, že byla přijata odpovídající opatření k zajištění pravosti a integrity záznamu. V tomto ohledu jsou zákony České republiky postavené tak, aby odpovídaly mezinárodní normě je ISO 15489. Mezi hlavní body obsažené v této normě týkající se uchovávání záznamů v archivu, jsou [15]:

- Retenční podmínky a postupy zacházení by měly být nastavené tak, aby ochraňovali záznamy před neoprávněným přístupem, ztrátou nebo zničením, jakož i proti krádeži a přírodním katastrofám

- Organizace by měly mít politiku a pokyny týkající se konverze nebo migrace záznamů z jednoho archivačního systému na jiný
- Systémy pro uchovávání elektronických dokumentů by měly být navrženy takovým způsobem, že záznamy budou přístupné, autentické, spolehlivé a použitelné při jejich uchovávání v průběhu celého období, bez ohledu na změny systémů

Jsou-li přijata tyto opatření, soudy ve většině případů berou v úvahu digitální verzi jako počáteční důkaz o existenci a obsahu původního záznamu. Je tedy na druhé straně, aby prokázala opak.

6.3 Minimální a maximální doby uchovávání

Minimální lhůty pro uchovávání údajů vyžadují uchovávání účetních knih a záznamů subjektu. Minimální doba uchovávání údajů se liší v každém členském státu Evropské unie. Pro ilustraci je povinnost uchovávat tato data pět let v Polsku, šest let v Německu, Finsku a Španělsku, sedm let v Rakousku, Švédsku a Nizozemí, osm let v Maďarsku a deset let v Belgii, Francii, Itálii a Rumunsku (a pro některé záznamy taky v Německu a ve Finsku). Pokud jsou záznamy relevantní pro statistiky subjektu, nebo jsou pro práci subjektu nezbytně nutné, je možné tato data uchovávat v podstatě pořád. Například, v Polsku je doporučeno zachovat schválené účetní uzávěrky natrvalo.

Kromě minimálních retenčních period existují rovněž povolené maximální doby uchovávání pro některé druhy záznamů. Účelem omezení uchovávání délky záznamů je především ochrana soukromí osob, jejichž osobní údaje jsou v datech obsaženy. Evropské zákony na ochranu údajů, které jsou realizovány ve všech výše uvedených státech, zakazují uchovávání záznamů obsahujících osobní údaje po dobu déle, než je nezbytně nutné pro dosažení cílů, pro které osobní údaje byly shromážděny nebo následně zpracovány. Co ale znamená "nezbytně nutná doba" v tomto ohledu není přesně definovaná. Maximální doba uchovávání ale platí, a liší se podle kategorie záznamu a nesmí být v žádném případě kratší, než je platná minimální zákonná doba uchovávání údajů. Poté, co tyto maximální lhůty pro uchovávání údajů uplynou, tyto dokumenty by měly být:

- bezpečně odstraněny,

- de-identifikované - všechny odkazy na subjekty údajů by měly být učiněny nečitelné,
- nebo v některých státech uložené v uzavřeném a nepřístupném archivu.

Je třeba brát v úvahu i to, že de-identifikaci nebo smazání dat případě právních sporů, daňových kontrol nebo vyšetřování je nutné pozdržet až do jeho konce. V takovém případě, jsou retenční periody prodlužovány případně probíhajících soudních sporů až do soudního rozhodnutí v poslední instanci.

6.4 Retenční periody v České republice

Každá firma musí povinně ze zákona archivovat některé dokumenty. Délky nutnosti uschovávat účetní závěrku, účetní a daňové doklady anebo mzdové listy popisuje několik zákonů a to následujícím způsobem.

6.4.1 Zákon o účetnictví (zákon č. 563/1991 Sb.)

Nejsnadnější popis toho, jak dlouho mají být dokumenty archivované je krátká citace samotného zákona:

§ 31 a § 32

Účetní jednotky jsou povinny uschovávat účetní záznamy pro účely vedení účetnictví po dobu, kterou stanovuje zákon.

Účetní záznamy se uschovávají [17]:

- účetní závěrka a výroční zpráva po dobu 10 let (počínajících koncem účetního období, kterého se týkají)
- účetní doklady, účetní knihy, odpisové plány, inventurní soupisy, účtový rozvrh, přehledy po dobu 5 let (počínajících koncem účetního období, kterého se týkají)
- účetní záznamy, kterými účetní jednotky dokládají vedení účetnictví, po dobu 5 let (počínajících koncem účetního období, kterého se týkají)

Pokud je záruční lhůta nebo reklamační řízení delší, uschovává účetní jednotka doklady a jiné účetní záznamy po dobu, po kterou tato lhůta běží nebo toto řízení trvá. Jestliže se účetní záznam vztahuje k nezaplacené pohledávce či nesplněnému závazku,

uschovává účetní jednotka tento účetní záznam do konce prvního účetního období následujícího po účetním období, v němž došlo k zaplacení pohledávky nebo ke splnění závazku. [17]

6.4.2 Zákon o dani z přidané hodnoty (zákon č. 235/2004 Sb.)

§ 27

Plátce DPH musí ze zákona uchovávat všechny daňové doklady rozhodné pro stanovení daně nejméně po dobu 10 let od konce zdaňovacího období, ve kterém se uskutečnilo zdanitelné plnění nebo plnění osvobozené od daně s nárokem na odpočet daně, na jím zvoleném místě. Po tuto dobu plátce odpovídá za věrohodnost původu dokladů, neporušitelnost jejich obsahu, jejich čitelnost a na žádost správce daně za umožnění přístupu k nim bez zbytečného odkladu. [20]

6.4.3 Zákon č. 582/1991 Sb.

Přesné znění názvu zákona je „Zákon o organizaci a provádění sociálního zabezpečení“.

§ 35a

Zaměstnavatelé jsou povinni uschovávat [19]:

- stejnopisy evidenčních listů vyhotovených v kalendářním roce, kterého se týkají, nebo v bezprostředně následujícím kalendářním roce po dobu 3 kalendářních roků po roce, kterého se týkají, a stejnopisy ostatních evidenčních listů po dobu 3 kalendářních roků po roce, ve kterém byly vyhotoveny
- mzdové listy nebo účetní záznamy o údajích potřebných pro účely důchodového pojištění po dobu 30 kalendářních roků následujících po roce, kterého se týkají; jde-li o mzdové listy nebo účetní záznamy o údajích potřebných pro účely důchodového pojištění vedené pro poživatele starobního důchodu, po dobu 10 kalendářních let následujících po roce, kterého se týkají
- u obchodní společnosti seznam společníků a členů statutárního orgánu a dozorčí rady této společnosti za jednotlivé kalendářní měsíce a přehled kalendářních měsíců, za které tato společnost neodvedla pojistné na sociální zabezpečení a příspěvek na státní politiku zaměstnanosti, které byla povinna

odvést (platí to obdobně pro družstvo) - po dobu 6 kalendářních let následujících po měsíci, kterého se záznam týká, vždy však po dobu 3 kalendářních let následujících po měsíci, v němž bylo dlužné pojistné za tento měsíc zapláceno

6.5 Řešení rozdílné retence

Protože právní předpisy definují různé doby retence pro různé typy dokumentů, není řešení situace triviální. První podmínkou je nalezení vhodného úložiště, které takovou retenci zabezpečí. Dalším krokem je vypořádání se s rozdílným přístupem k datům s rozdílnou dobou retence, v poslední řadě jsou to legislativní nároky na zabezpečení pravosti dokumentu.

6.5.1 Úložiště

Pro dodržení stanovené retence jsou vhodné úložiště, tzv. „WORM“, které jsou programovatelné. WORM je akronym anglického spojení „Write Once, Read Many“. To znamená, že úložiště dat obsahuje mechanismus, který umožní zapsat data jednou, ale neumožní jejich smazání, nebo modifikaci. Povolené je jenom čtení. [18]

Programovatelné znamená to, že po uplynutí naprogramované retence je zmazání nebo modifikace dat zařízením povolena.

Další možností je vypalování dat na nosiče CD/DVD/BR, které svou podstatou neumožňují změnu obsahu. Fyzická činnost s nosiči je náročná, dostupnost nízká a cena za malé množství uchovaných dat poměrně vysoká. V úvahu připadají i jiné technologie pro ukládání dat na bázi opticko – magnetického záznamu, ty však jsou rozšířené jen minimálně a pro dlouhodobou archivaci nejsou ani tak vhodné, jako archivace na optické nosiče.

Celkovým a nejvhodnějším řešením pro správné uchovávání dat v retenčních periodách je zařízení typu WORM, které dokáže pracovat s vícero retenčními periodami a které dokáže definovat pro jednotlivé skupiny dat rozdílné hodnoty retenčních period. Obrovskou výhodou má takové zařízení, které dokáže určovat retenci pro každý dokument

separátně. Tyto zařízení dodávají společnosti: IBM, HP, EMC², Fujitsu, Oracle, Hitachi a další.

Velmi rozšířeným zařízením typu WORM je produkt Centera společnosti EMC². Poměr cena za MB dat a výkon zařízení je vyvážený. Funkcionalitou je Centera mezi lídry ve své kategorii, její hlavní přednosti jsou:

- Certifikace úložiště podle legislativních norem a norem ISO pro chráněné úložiště
- Možnosti nastavení retenčních period na všech úrovních, retence interní pro objekt anebo jednotlivé soubory, retence řízená aplikací, nekonečná retence, atp.
- Síťové zabezpečené připojení
- Deduplikace dat
- Škálovatelnost, automatické zálohování, podpora výrobce

6.6 Skartace

Po uplynutí archivační povinnosti je vhodné písemnosti zlikvidovat. A jelikož se jedná o citlivá firemní nebo osobní data, nejvhodnějším způsobem je skartace. Zařízení, které pokyn pro skartaci dokumentů dostane, obsah ze svého úložiště smaže, nebo zlikviduje nevratným způsobem. Existují nosiče, na kterých je tento proces nutné vykonat ručně, typicky fyzickou likvidací CD/DVD média. Existují i další metody, které nosiče dat znehodnocují softwarově anebo přímo fyzicky likvidují.

7 Archivní řešení

7.1 Služby v archivnictví

I když v tuto chvíli ještě není elektronická archivace natolik rozšířená, aby bylo možné konstatovat, že dokumenty uchovávané v elektronické podobě převyšují ty papírové, už teď ale existuje velké množství nástrojů, které elektronickou archivaci umožňují. Poskytovatelé těchto služeb se snaží, aby jejich řešení splňovalo legislativní kritéria konkrétní krajiny, reálně je však cílem všech splňovat mezinárodní standardy.

V České republice existuje několik firem, které takové služby poskytují. Ve svém portfoliu nabízejí komplexní řešení, které se zabývá všemi činnostmi, které s archivací souvisí. Dokáží svým řešením sbírat data v podstatě z jakéhokoliv zdroje, dokumenty pak ošetřit časovým razítkem, elektronickým podpisem, nebo jej komplexně šifrovat proti zneužití. Tato data pak náležitě uchovávat na fyzických nosičích a pečovat o ně podle potřeb legislativních anebo preferencích subjektu. Na druhé straně, dokáží zabezpečit dostupnost těchto dat v požadovaném intervalu.

Některé společnosti, které se v České republice archivací zabývají:

- O2 Czech Republic, a.s.
- T-Systems Czech Republic, a.s.
- Autocont, a.s.
- Software602, a.s.
- Ixtent s.r.o.
- Sabris a.s.
- ICZ a.s.
- a další.

Profesionální řešení nabízí i nespočet zahraničních společností. Každá společnost se specializuje na jinou oblast trhu, základní funkcionalitu na archivaci ale splňují všechny. Společnosti dokáží dodat subjektu komplexní řešení. To představuje základní hardwarovou výbavu, softwarový produkt a služby pro rozšiřování řešení a údržbu.

Protože potřeb pro vykonávání samotného aktu archivace je mnoho, lze konstatovat, že je pro tyto služby je neideálnější samostatný software, který může být mezičlánkem mezi fyzickým úložištěm a aplikací, která data generuje, nebo s nimi zachází. Software je pak předmětem výběru na základě funkcí, které jsou požadované legislativou, technologickými možnostmi anebo i splněním požadavků referenčního modelu pro archivační software O AIS.

Řešení, které firmy nabízejí, jsou postavené jenom na několika produktech, čtyři z nich jsou všeobecně známé produkty, které jsou v České republice zastoupené nejvíce. Společnost Software602, a.s. a O2 Czech Republic, a.s. nabízejí vlastní řešení, které má specifickou úlohu a jako univerzální řešení pro široké spektrum subjektů není vhodné. V analýze pro výběr vhodného řešení proto bude uvažováno jenom se čtyřma základními produkty.

7.2 Open Text Archive Server

Kanadská společnost Open Text nabízí produkt pojmenován „Archive Server“, ten je historicky navázán na produkty společnosti SAP a.g. Archivovat tímto produktem ale lze cokoli, především data navázané na produkty z obrovského portfolia, které Open Text nabízí. Archiv této společnosti splňuje nejpřísnější podmínky, které je nutné při zabezpečeném archivu dodržovat z hlediska legislativních a právních norem.

Základní charakteristika:

- Jednoduchost a flexibilita řešení je výhodou produktu
- Funkcionalita je malá, ale produkt je vhodný pro obrovské množství dat, pro kterého správu zvládá dokonale
- Výkon řešení je obrovský vzhledem k požadovanému hardware
- Licence a cena produktu je nastavena převážně na velké korporátní subjekty
- Napojení na portfolio společnosti Open Text je téměř automatické a nabízí tak další dimenzi pro použití

7.3 IBM FileNet

Společnost IBM má ve svém portfoliu produkt s názvem FileNet. Více než archiv je tento produkt robustní řešení, které nahrazuje podnikové aplikace typu DMS, ECM, apod.

FileNet ve své funkcionalitě dokáže taktéž splňovat nejpřísnější kritéria pro archivaci. Není ale natolik flexibilní jako konkurenční produkt od společnosti Open Text. FileNet se nasazuje tam, kde není modularita řešení zásadní, rozšiřování tohoto archivu je z hlediska jeho architektury složité. Podstatná je i cena za řešení, která vzhledem k nabízeným možnostem produktu, je adekvátní jenom pro velké subjekty čítající stovky až tisíce uživatelů.

Základní charakteristika

- Funkcionalita řešení je obrovská, produkt nahrazuje produkty, které jsou v ostatních případech na archiv navázané a zvyšují tak cenu, nebo výkon apod.
- Výkon řešení je malý, pro rychlý běh v korporaci se stovkami uživatelů jsou vyžadované vysoké hardwarové nároky
- Cena za řešení je ze všech produktů nejvyšší

7.4 Setcce eKeeper

Slovinská společnost Setcce nabízí produkt eKeeper. Hlavní zaměření je uchovávání dokumentů ve spolupráci s certifikačními autoritami. Produkt eKeeper splňuje normu OAIS, je tím transparentní a předurčen pro dlouhodobou archivaci. Podpora normy zaručuje to, že čitelnost archivovaných dat je zaručena i po ukončení životního cyklu produktu a případné migraci dat na jiný produkt taktéž podporující normu OAIS.

Základní charakteristika:

- Cenově nejpřístupnější řešení vhodné pro malé ale i nadnárodní subjekty
- Funkcionalita je minimální, automatická podpora množství certifikačních autorit je velkou výhodou

7.5 ELO

Jednoduché řešení pro archivaci dat spojeného s podnikovým ECM řešením. Produkt ELO tak může být nasazený v kancelářích o jednotkách uživatelů až po obrovské společnosti. Funkcionalitou se dotahuje na IBM FileNet, výkonem a robustností ale není předurčen jako řešení pro kritické situace.

Základní charakteristika:

- Pokročilé funkce nabízející jednoduché workflow
- Ukládání dokumentů není primární zaměření produktu
- Jednoduchá a levná údržba a konfigurace

8 Výběr vhodného archivního softwaru

Metodika pro definování nejvhodnějšího řešení pro archivaci vzhledem k retenčním periodám je popsána v kapitole 2 tohoto dokumentu. Jedná se o metody z vícekritériální analýzy variant.

Kritéria stanovená pro produkty, které jsou metodikou ohodnocované, jsou navrženy na základě zkušeností odborníků z mnohaleté praxe. Zároveň jsou navrženy podle modelové situace, která je vyváženým průměrem potřeb, které mají jednotlivé subjekty pro zavedení archivace elektronických dokumentů.

8.1 Metoda pořadí

Seřazení důležitosti jednotlivých parametrů je v stupnici od 1 do 4. Nejlepší ohodnocení znamená číslo 4. Promítnutí ohodnocení je ve výsledné tabulce. Kritéria jsou zaznamenané ve sloupcích. Váhy pak sloupce doplňují. Jednotlivé řešení jsou zaznamenané v řádcích, na konci kterých je výsledek vypočten z hodnot ohodnocení a vah. Řešení s nejvyšší hodnotou je řešení, které je dle metody pořadí to nejlepší.

Tabulka 1 Ohodnocení a výsledky metody pořadí

Řešení	Cena	Funkcionalita	Výkon	Robustnost	Výsledek
Archive Server	2	2	4	3,5	3,2
FileNet	1	4	1	3,5	2,6
eKeeper	3	1	3	1,5	2
ELO	4	3	2	1,5	2,2
váhy	0,1	0,2	0,3	0,4	

Tabulka 2 Váhy metody pořadí

Kritérium	Pořadí	Body	Váha
Cena	4	1	0,1
Funkcionalita	3	2	0,2
Výkon	2	3	0,3
Robustnost	1	4	0,4

8.2 Bodovací metoda

Použití metody vyžaduje ohodnocení parametrů. Stupnice, která byla pro ohodnocení parametrů zvolena má rozpětí od 1 do 10. Nejlepší ohodnocení znamená nejvyšší hodnotu, v tomto případě 10.

V tabulce jsou v řádcích řešení, které jsou porovnávány. Kritéria jsou zaznamenána ve sloupcích. Výpočet se provádí za pomoci vah (v posledním řádku) a výsledek je zaznamenán v posledním sloupci. Nejvyšší hodnota znamená nejlepší řešení.

Tabulka 3 Výsledky bodovací metody

Řešení	Cena	Funkcionalita	Výkon	Robustnost	Výsledek
Archive	4	3	9	8	6,6
FileNet	2	10	4	7	6,28
eKeeper	9	2	7	2	4,24
ELO	10	5	3	3	4,32
Váha	0,12	0,24	0,28	0,36	

Tabulka 4 Váhy bodovací metody

Kritérium	Body	Váha
Cena	3	0,12
Funkcionalita	6	0,24
Výkon	7	0,28
Robustnost	9	0,36
Součet	25	

8.3 Metoda AHP/Saatyho

Způsob vzájemného porovnávání kritérií (AHP) je využíván v Saatyho metodě. Nutnost porovnat všechny kritéria na všech úrovních znamená použití několika tabulek pro

výpočet pořadí řešení. Po součtu všech vypočtených vah lze získat pořadí, ve kterém se jednotlivé řešení umístily. Nejvyšší hodnota znamená první místo v ohodnocení.

Tabulka 5 Stanovení vah

Stanovení vah	Cena	Funkcionalita	Výkon	Robustnost	ri	Váhy
Cena	1,00	1/3	1/5	1/7	0,312394	0,058656
Funkcionalita	3,00	1,00	1/2	1/5	0,782542	0,146933
Výkon	5,00	2,00	1,00	1/2	1,495349	0,280772
Robustnost	7,00	4,00	2,00	1,00	2,735565	0,513639
				součet	5,325850	1,000000

Tabulka 6 Cena

Cena	Archive	FileNet	eKeeper	ELO	ri	vi	Váhy
Archive	1,00	3,00	0,25	0,20	0,622333	0,105365	0,00618
FileNet	0,33	1,00	0,14	0,11	0,269702	0,045662	0,002678
eKeeper	4,00	7,00	1,00	0,50	1,934336	0,327496	0,01921
ELO	5,00	9,00	2,00	1,00	3,080070	0,521476	0,030588
				součet	5,906442	1,000000	0,058656

Tabulka 7 Funkcionalita

Funkcionalita	Archive	FileNet	eKeeper	ELO	ri	vi	Váhy
Archive	1,00	0,17	2,00	0,20	0,508133	0,081662	0,011999
FileNet	6,00	1,00	9,00	3,00	3,567621	0,573354	0,084244
eKeeper	0,50	0,11	1,00	0,14	0,298475	0,047968	0,007048
ELO	5,00	0,33	7,00	1,00	1,848148	0,297016	0,043641
				součet	6,222376	1,000000	0,146933

Tabulka 8 Výkon

Výkon	Archive	FileNet	eKeeper	ELO	ri	vi	Váhy
Archive	1,00	9,00	2,00	6,00	3,223710	0,546065	0,15332
FileNet	0,11	1,00	0,17	2,00	0,438691	0,074310	0,020864
eKeeper	0,50	6,00	1,00	4,00	1,861210	0,315271	0,088519
ELO	0,17	0,50	0,25	1,00	0,379918	0,064354	0,018069
				součet	5,903529	1,000000	0,280772

Tabulka 9 Robustnost

Robustnost	Archive	FileNet	eKeeper	ELO	ri	vi	Váhy
Archive	1,00	2,00	8,00	6,00	3,130169	0,528231	0,27132
FileNet	0,50	1,00	6,00	5,00	1,967990	0,332107	0,170583
eKeeper	0,13	0,17	1,00	0,50	0,319472	0,053912	0,027691
ELO	0,17	0,20	2,00	1,00	0,508133	0,085750	0,044044
				součet	5,925763	1,000000	0,513639

Tabulka 10 Výsledek AHP

Řešení	Součet vah	Pořadí
Archive	0,442819	1,00
FileNet	0,278370	2,00
eKeeper	0,142468	3,00
ELO	0,136343	4,00

8.4 Porovnání výsledků

Tabulka 11 Celkové pořadí

Použité metody	Archive	FileNet	eKeeper	ELO
Metoda pořadí	1	2	4	3
Bodovací metoda	1	2	4	3
Saatyho metoda (AHP)	1	2	3	4
Pořadí	1	2	4	3

Zásadní podíl na určení vhodnosti řešení mají váhy. V případě určení řešení pro archivaci, byly váhy stanovené tak, aby preferovali robustnost řešení a výkon před cenou a funkcionalitou. Pokud by byly váhy stanovené jinak, je pravděpodobné, že se to projeví v celkovém výsledku pořadí.

Všechny tři použité metody pro ohodnocení dopadly s podobným výsledkem. Jako nejvhodnější řešení pro Archivaci a pro stanovené kritéria je produkt od společnosti Open Text a má název Archive Server.

Druhým v pořadí je i navzdory extrémně vysoké ceně a nepružnosti řešení produkt FileNet od společnosti IBM.

O třetí a čtvrtou příčku pak bojují levné řešení eKeeper a ELO, více funkcionality zřejmě převažuje nad výkonem eKeeperu a má mírně navrch ve dvou ze tří porovnání.

9 Závěr

Legislativní ukotvení pro elektronickou archivaci je specifikované v několika navazujících zákonech, které jsou představené v jednotlivých kapitolách. Tyto zákony lze rozdělit na směrnice a nařízení Evropské unie a na právní ustanovení České republiky. Zákony a směrnice Evropské unie nespécifikují konkrétní doby nutné pro uchovávání dat, zvolení retence je na každém členském státu podle vlastních požadavků. Česká republika uzákoňuje retence v několika navazujících zákonech. Na základě analýz legislativy lze konstatovat, že retence účetních dokladů se pohybuje od minimální doby řádově v měsících až po několik desetiletí, pokud za účetní doklady považujeme data o zaměstnancích. Není proto optimální stanovit konkrétní (nejvyšší) hodnotu, náklady pro uchovávání nepotřebných dat by pro subjekt mohly být neúnosné. Pro každý typ dokumentů je proto vhodné stanovit konkrétní retenci a archivačním procesem po vypršení doby tyto dokumenty elektronicky skartovat. Na činnost pro celý archivační proces je vhodný specifický software, který je předmětem výběru pomocí vícekriteriální analýzy variant.

Sestavení řešení na základě výše popsaných podmínek by mělo být z prvků:

- Zabezpečené hardwarové úložiště typu WORM.
- Archivační řídicí aplikace pro nakládání s retencemi.
- Vykonávání zabezpečení dat pomocí elektronického podpisu, označování časovými razítky a volitelně i kryptováním.

Pro porovnání byly vybrány čtyři produkty, které jsou pro archivaci určené a které splňují požadavky na obsluhování retencí, HW úložiště a dokáží zabezpečit data tak, aby je bylo možné použít v jakémkoliv legislativním řízení. Na základě výsledků metody vícekriteriální analýzy variant a zvolených vah podstatných pro subjekt je optimálním řešením aplikování software s názvem Archive Server od společnosti Open Text.

Stejný způsob použité metodiky pro zvolení produktu „Archive Server“ lze aplikovat pro výběr HW úložiště typu WORM a požadované certifikační autority. Problematika je ale natolik komplexní a specifická pro každý subjekt, který o elektronické archivaci dat uvažuje, že nelze určit jednoznačné a nejoptimálnější všeobecné řešení. Na základě konzultací s odborníky a průzkumu trhu lze konstatovat, že populární kombinace mezi subjekty je

implementace Open Text Archive spolu s EMC² Centera a napojením na certifikační autoritu České pošty – PostSignum. Lze tedy konstatovat, že tato kombinace pro mnoho subjektů v České republice představuje ideální řešení problematiky elektronické archivace dat v závislosti na retenčních periodách.

10 Bibliografie

- [1] ŠUBRT, T. a kolektiv. *Ekonomicko – matematické metody*. Vyd. 1. Plzeň: Vydavatelství a nakladatelství Aleš Čeněk, 2011, 351 s. ISBN 978-80-7380-345-2
- [2] BROŽOVÁ, H., ŠUBRT, T. a HOUŠKA, M. *Modely pro vícekritériální rozhodování*. Vyd. 1. Praha: Credit, 2003, 172 s. ISBN 978-80-213-1019-3.
- [3] PODSTRELENEC, J. *DIP: Elektronická archivácia papierových dokumentov*. In: Trend [online]. Trend Holding, 1998. [cit. 31. 12. 2013]. Dostupné z URL: <http://www.etrend.sk/trend-archiv/rok-/cislo-April/dip-elektronicka-archivacia-papierovych-dokumentov.html>.
- [4] ČESKÁ REPUBLIKA, [online], *Zákon č. 227/2000 Sb., ze dne 29. června 2000 o elektronickém podpisu a změně některých dalších zákonů*. In: *Sbírka zákonů České republiky*, 2000. Dostupné z URL: <http://www.mvcr.cz/clanek/zakon-c-227-2000-sb-o-elektronickem-podpisu.aspx> , ISSN 1211-1244
- [5] LIDINSKÝ, V. *eGovernment bezpečně*. Praha: Grada, 2008. ISBN 978-80-247-2462-1.
- [6] DOSTÁLEK, L.; VOHNOUTOVÁ, M.; KNOTEK, M. *Velký průvodce infrastrukturou PKI a technologií elektronického podpisu*. 2001, 534 s. ISBN 80-251-0828-7
- [7] KUNSTOVÁ, R. *Efektivní správa dokumentů*. Praha: Grada, 2009. 204s. ISBN 978-80-247-3257-2.
- [8] NÁRODNÍ ARCHIV, [online]. Národní archiv, r.v.: 2000, [cit. 12. 12. 2014] Dostupné z URL: http://www.nacr.cz/sua/cinnost/o_archiv.htm
- [9] HÖNIG, P. *Důvěryhodný elektronický archiv pro dlouhodobou archivaci, požadavky a jejich řešení*. Liberec: AKP 2007, 16-17. 5. 2007. 11. ročník semináře.
- [10] ČESKÁ REPUBLIKA, [online], *Vyhláška č. 191 ze dne 23. června 2009 o podrobnostech výkonu spisové služby*. In: *Sbírka zákonů České republiky, 2009*. Dostupné z URL: <http://www.mvcr.cz/soubor/archivnictvi-a-spisova-sluzba-dokumenty-vyhlaska-o-spisove-sluzbe-pdf.aspx>
- [12] LECHNER, T. *Elektronické dokumenty v právní praxi*. Praha: Leges, 2013. 256s. ISBN 978-80-87576-41-0.
- [13] AIPSAFE, [online]. Aip Safe s.r.o., 2009, [cit. 12. 12. 2014] Dostupné z URL: <http://www.aipsafe.cz/cs/oais>

- [14] **KLIMEŠ, L.** *Slovník cizích slov*. Praha: SPN, 1983, 791 s., ISBN 14-545-83
- [15] **IRON MOUNTAIN**, [online], *European Document Retention Guide 2013*. In: Iron Mountain, 2013. [cit. 12. 2. 2015]. Dostupné z URL:
<http://www.ironmountain.co.uk/~media/files/iron%20mountain/knowledge%20center/reference%20library/europe/best%20practices/e/european%20document%20retention%20guide%202013.pdf> .
- [16] **SOBOTKOVÁ, M.**, [online], *Jak dlouho musíte archivovat firemní dokumenty?* In: STORMWARE s.r.o., 2012. [cit. 12. 2. 2015]. Dostupné z URL:
<http://portal.pohoda.cz/dane-ucetnictvi-mzdy/ucetnictvi/jak-dlouho-musite-archivovat-firemni-dokumenty/>
- [17] **ČESKÁ REPUBLIKA**, [online], *Zákon č. 563/1991 Sb. ze dne 12. prosince 1991 o účetnictví*. In: *Sbírka zákonů České republiky*, 1992. Dostupné z URL:
<http://portal.gov.cz/app/zakony/zakonPar.jsp?idBiblio=39611&nr=563~2F1991&rpp=15#local-content>
- [18] **JANSEN, C.**, [online], TechoPedia - Janalta Interactive Inc., 2010, [cit. 21.1.2015] Dostupné z URL: <http://www.techopedia.com/definition/27405/write-once-read-many-worm-data-storage>
- [19] **ČESKÁ REPUBLIKA**, [online], *Zákon č. 582/1991 Sb. ze dne 17.12.1991 o organizaci a provádění sociálního zabezpečení* In: *Sbírka zákonů České republiky*, 2000. Dostupné z URL: http://www.mpsv.cz/ppropo.php?ID=z582_1991o
- [20] **ČESKÁ REPUBLIKA**, [online], *Zákon č. 325/2004 Sb. ze dne 1.4.2004, o dani z přidané hodnoty*. In: *Sbírka zákonů České republiky*, 2005. Dostupné z URL:
<http://portal.gov.cz/app/zakony/zakon.jsp?page=0&nr=235~2F2004&rpp=15#seznam>

11 Přílohy

Tabulka 1 Ohodnocení a výsledky metody pořadí	33
Tabulka 2 Váhy metody pořadí	33
Tabulka 3 Výsledky bodovací metody	34
Tabulka 4 Váhy bodovací metody	34
Tabulka 5 Stanovení vah	35
Tabulka 6 Cena.....	35
Tabulka 7 Funkcionalita	35
Tabulka 8 Výkon	35
Tabulka 9 Robustnost	36
Tabulka 10 Výsledek AHP	36
Tabulka 11 Celkové pořadí	36