Vysoká škola ekonomická v Praze

Fakulta informatiky a statistiky

Katedra informačních technologií

Studijní program: Aplikovaná informatika

Studijní obor: Podniková informatika

Metodika klasifikace dokumentů ve firmě v rámci ECM

Autor diplomové práce: Mgr. Václava Kučerová Zrálíková

Vedoucí diplomové práce: doc. Ing. Vlasta Střížová, CSc.

Oponent diplomové práce: Ing. Vojtěch Braun

Rok odevzdání: 2017

Prohlášení

Prohlašuji, že jsem diplomovou práci zpracovala samostatně a že jsem uvedla všechny použité prameny a literaturu, ze kterých jsem čerpala.

V Praze dne 20. dubna 2017 …….

podpis

Poděkování

Ráda bych touto cestou poděkovala doc. Ing. Vlastě Střížové, Csc., za trpělivost při vedení mé diplomové práce.

Také bych chtěla poděkovat účastníkům připomínkového řízení za jejich cenné podněty na zlepšení.

A v neposlední řadě děkuji své rodině, která mi poskytla nejen časový prostor, aby tato práce mohla vzniknout.

Abstrakt

Podniky v dnešní době musí nakládat s velkým množstvím podnikového obsahu. Aby bylo možné ho během jeho celého životního cyklu kvalitně řídit, je nutné, aby dokumenty byly opatřeny správnými metadaty, která to zaručí.

Cílem této diplomové práce je návrh metodiky klasifikace zákaznických dokumentů, která zachycuje proces návrhu třídění přes testování až po údržbu v běžném provozu a případný drobný vývoj.

Práce je zpracováním tématu jedinečná, neboť používá průnik knihovnické vědy s informačními technologiemi. Metodika mimo jiné obsahuje i návrh věcného selekčního jazyka a návrh IT služeb, které mají zajistit kvalitu poskytované klasifikace. Celá práce je rozdělena na dvě části, teoretickou, ve které jsou představeny základní pojmy, teorie kategorizace a normy, standardy a zákony, které mají na oblast vliv. Druhá část je praktická, obsahuje samotnou metodiku.

Klíčová slova

ECM, správa podnikového obsahu, klasifikace, metodika, předmětový selekční jazyk

Abstract

At present companies are facing to a problem of treating huge amount of company`s content. There is needed to provide correct metadata to the documents to enable its qualified classifying during a whole life cycle of the document.

The target of this diploma thesis puts emphasis on a proposal of the classifying methodology of the customer`s documents, while intercepting a proposal of the process of classifying, testing and support or even enhancement.

Topic of this thesis is unique while it is based on the combination of the IT and librarian`s technologies. Methodology also consists a proposal of a subject retrieval language and a proposal of IT services, which are going to ensure the quality of the provided classification. The thesis is devided into two parts – the theoretical one explains the basic terms, theory of classifying and standards, legislature, which all together have the important influence on the field. The other – practical – part includes itself the methodology.

Key words

ECM, taxonomy, classification, methodology, content services, subject retrieval language

Obsah

1 Úvod 1

1.1 Cíle práce 3

1.2 Použité metody 3

1.3 Komentovaná rešerše 4

TEORETICKÁ ČÁST 7

2 Vymezení základních pojmů 7

2.1 Data, informace, znalosti 7

2.2 Metadata, atribut 9

2.3 Dokument, obsah, záznam 10

2.4 Zákazník 11

2.5 Firma, podnik 12

2.6 ROT obsah 12

2.7 Prohlížení, analytické vyhledávání, mobilní vyhledávání 13

2.8 Podnikové vyhledávání (enterprise search) 14

2.9 Spisový a skartační řád 14

2.10 ECM 15

3 Katalogizace dokumentů 18

3.1 Prekoordionované selekční jazyky 19

3.1.1 Deweyho desetinné třídění 20

3.1.2 Mezinárodní desetinné třídění 20

3.1.3 Předmětová hesla 21

3.2 Postkoordinované selekční jazyky 21

3.2.1 Randhanatanovo dvojtečkové třídění 21

3.2.2 Klíčová slova 22

3.2.3 Deskriptorové selekční jazyky 23

3.3 Klasifikace, indexace a související pojmy 24

3.3.1 Taxonomie 26

3.3.2 Ontologie 29

3.3.3 Nástroje umožňující správu předmětového věcného třídění 30

4 Standardy, normy a legislativní dokumenty týkající se klasifikace dokumentů 33

4.1 Zákon č. 499/2004 Sb., o archivnictví a spisové službě a o změně některých zákonů 33

4.2 Zákon č. 563/1991 Sb., o účetnictví 34

4.3 Zákon č. 235/2004 Sb., o dani z přidané hodnoty 34

4.4 Zákon č. 412/2005 Sb., o ochraně utajovaných informací a o bezpečnostní způsobilosti 34

4.5 Zákon č. 101/2000, o ochraně osobních údajů 35

4.6 Vymáhací proces pohledávek 35

4.7 eIDAS 35

4.8 Normy 36

4.8.1 ISO norma 15489 36
4.8.2 ISO norma 23081 37
4.8.3 ANSI/AIIM TR40-1995 37
4.8.4 ANSI/AIIM/ARMA TR48-2004 38
4.8.5 ANSI/AIIM MS60-1996 38
4.9 Standardy a jiné projekty 38
4.9.1 MoReq2 38
4.9.2 Dublin Core 38
4.9.3 CMIS 40
4.9.4 AGRkMS 41
4.9.5 Schema.org 41
4.9.6 DBpedia 42
4.9.7 OWL 42
4.9.8 SKOS 43
PRAKTICKÁ ČÁST 45
5 Analýza současného stavu ECM z pohledu klasifikace 45
6 Metodika klasifikace zákaznických dokumentů ve firmě 49
6.1 Proces klasifikace zákaznických dokumentů 51
6.2 Fáze plánování 56
6.2.1 Stanovení rozsahu aktivity a cílů 56
6.2.2 Výpočet návratnosti investic a stanovení metrik 58
6.2.3 Vydefinování pracovního týmu 59
6.2.4 Analýza současného stavu 60
6.3 Tvorba obecného návrhu klasifikace 65
6.3.1 Procesní mapa dokumentů 68
6.3.2 Popis postupu tvorby firemního třídění 72
6.3.3 Tvorba předmětového selekčního jazyka pro zákaznické dokumenty 88
6.4 Implementace 104
6.4.1 Tvorba technického provedení 105
6.4.2 Kontrola navrženého třídění 122
6.5 Provoz 127
6.5.1 Služby zajišťující běžný provoz 128
6.5.2 Implementace změn 134
7 Závěr 137
8 Terminologický slovník 139
9 Seznam literatury a použitých zdrojů 141
10 Seznam obrázků, tabulek a grafů 150
10.1 Seznam obrázků 150
10.2 Seznam tabulek 151
11 Přílohy 152

1 Úvod

„Topíme se v informacích a žízníme po znalostech.“

John Naisbitt

Žijeme v době, kdy jsme denně konfrontováni s velkým množstvím dat, která už není v lidských silách zpracovat. Část těchto informací je uložena ve strukturovaných či nestrukturovaných dokumentech. Podniky dnes více než dříve musí hledat a nalézt cestu, jak s těmito informacemi zacházet. Kromě toho, že v dokumentech jsou důležité informace potřebné pro chod samotného podniku, jsou na ně kladeny různé nároky z dodržování oborových standardů či zákonných norem. V odborné literatuře se nyní setkáváme nejen s dolováním dat (data mining), ale i s dolováním dokumentů (document mining, text mining), jehož cílem je získání užitečných, dříve neznámých, poznatků z nestrukturovaného textu, které jsou v písemnostech skryty.

Mimo to díky snahám týkajícím se ochrany životního prostředí podniky preferují chovat se šetrně, což přináší například tlak na snížení spotřeby papíru. V poslední době se více než dřív tedy setkáváme s tím, že již některé typy dokumentů vznikají pouze v elektronické podobě (faktury, zákaznické smlouvy,…). Vzhledem k neexistenci papírového originálu je více než kdy jindy nutné dokument opatřit kvalitními metaúdaji, které umožní nejen jeho bezproblémové nalezení, ale i práci s ním a podpoří procesy, které jsou s dokumentem spjaty.

Při implementaci technologií ECM se setkáváme s tím, že jde spíše o technickou dodávku nového softwaru. Konzultanti dodavatele jsou vybaveni dobrými technickými znalostmi týkajícími se jejich produktu a dobrými komunikačními schopnostmi, nicméně jsou patrné jisté nedostatky z neznalosti oblasti, ve které se podnik pohybuje (interní prostředí, věcná znalost oblasti, externí okolí podniku), v pracovních návycích, procesy, postupy a podobně. V ideálním případě jim toto pomůže překonat interní zaměstnanec, který je v implementaci zainteresován. Nicméně většinou se tak neděje, což je způsobeno několika faktory. K nimž patří zejména vytíženost interního zaměstnance, nezájmem o danou aktivitu, špatná komunikace způsobena vzájemným neporozuměním si (zaměstnankyně z personálního oddělení má jiný styl komunikace a práce než odborník z oddělení IT), anebo nedostatek zkušeností, aby si práci v novém nástroji představil, proto není schopen kompetentně zodpovědět dotazy. Proto bývají tyto implementace spíše záležitostí IT. Koncoví uživatelé či vlastníci dat, se této aktivitě vyhýbají nebo se o ni zajímají pouze okrajově, nový systém

zpravidla odmítají. Nespolupráce na vývoji vede k tomu, že nástroj třeba jen částečně vyhovuje původním očekáváním, špatně se v něm pracuje. Jednou ze součástí aktivity je i návrh, následná správa a vývoj klasifikace dokumentů tak, aby byla umožněna kvalitní práce s dokumenty. Tato část bývá v rámci celé aktivity upozaděna. Domnívám se, že je to většinou dáno tím, že specialisté na daný produkt a samotní uživatelé si navzájem nedokážou pomoci (specialisté žádají o definici metadat svého klienta, a klient, pokud nemá zkušenost s podobným typem technologie, neví, jak data specifikovat, neboť doteď pracoval pouze v hierarchii složek) a vzhledem k složitosti a problematice celé aktivity se řeší na poslední chvíli stylem, nějaká metadata k dokumentům navrhneme, časem se to opraví. Málokdo si v této fázi uvědomí, že časem se to nenapraví, neboť k tomuto stavu nemusí již dojít (aplikace nebude používána a její provoz bude předčasně ukončen). Statistická data ukazující, jaký dopad může mít nekvalitní třídění na práci s podnikovým obsahem, jsou uvedena v kapitole zabývající se analýzou současného stavu.

Diplomová práce se snaží tuto mezeru vyplnit. Kromě základního teoretického vhledu do problematiky klasifikace poskytuje také metodiku, jakým způsobem postupovat při tvorbě zákaznického třídění od návrhu až po jeho správu. Toto téma jsem si vybrala, protože se již několik let pohybuji v této oblasti a mohla jsem tedy využít znalosti, které jsem získala při svém minulém studiu. Při čtení odborné literatury jsem zjistila, že jsou články psány buď z pohledu informančně-knihovnického, nebo jsou technologicky zaměřené na určitý prvek v rámci celého problému. Z toho důvodu je tato práce průnikem těchto dvou světů a dívá se na problematiku jako na celek s přesahem na jiné aktivity v rámci ECM domény.

Práce je určena jak informačním specialistům, kteří se zabývají tvorbou předmětových selekčních jazyků, tak technologickým odborníkům, kteří s doménou ECM přichází do styku (vývojáři, konzultanti, projektoví manažeři, analytici,…).

Práce je rozdělena na dvě části, první je teoretická. V ní jsou definovány základní používané termíny, teorie kategorizace dokumentů a zákonné a normativní požadavky kladené na tuto oblast. V druhé, praktické části je předložena metodika klasifikace dokumentů a to včetně navržených věcných selekčních jazyků. Tato práce se zabývá pouze klasifikací zákaznických dokumentů v podniku, neboť se domnívám, že jsou pro správný a bezproblémový chod organizace důležitější než jiné typy podnikového obsahu. Mimo jiné definuje, jaká mají tyto dokumenty specifika, která se následně promítají do práce s nimi. Některé části a principy z metodiky mohou být aplikovány i na jiný podnikový obsah.

1.1 Cíle práce

Hlavním cílem této diplomové práce je navření metodiky klasifikace zákaznických dokumentů v podniku. Dílčím cílem je vytvoření předmětového selekčního jazyka zákaznických dokumentů ve firmě, který je nedílnou součástí navržené metodiky.

Důležitou částí pro naplnění cílů práce je popis zákonů, standardů a norem týkajících se kategorizace dokumentů v rámci ECM a vymezení termínů, které s danou tématikou souvisejí.

1.2 Použité metody

V teoretické části diplomové práce je použita vědecká metoda **analýzy**:

- Které zákony, standardy a normy a jakým způsobem mají dopad na třídění (kapitola 4).
- Jaké jsou základní pojmy s jejich vymezením s ohledem na práci se tříděním (kapitola 2).
- Představila jsem oblast katalogizace a způsoby jejího dělení (kapitola 3).

V kapitole Analýza současného stavu ECM byl použit přístup **výzkum od stolu** (desk research).

V praktické části jsem na analýzu z teoretické části vytvořila metodiku klasifikace dokumentů, při které byla využita převážně metoda **syntézy**. V kapitole Popis postupu firemního třídění (kapitola 6.3.2) byly použity i metody **analogie** a **indukce**.

Pro ověření, že je správná, byla mnou navržená metodika předložena ke kritickému hodnocení odborníkům z praxe a na základě jejich připomínek revidována. Pro výběr odborníků bylo použito kritérium, že se setkávají s techonologiií ECM v rámci své pracovní náplně na různých pracovních pozicích. Procesu připomínkování se účastnilo 10 respondentů s těmito rolemi: business analytik ECM, běžný koncový uživatel (tzn. specialista v jiném oboru), metodik, specialista na skartační a archivační proces, projektový manažer, servisní manažer, manažer zodpovídající za provoz ECM, konzultant. Jejich připomínky byly v rámci připomínkového řízení do diplomové práce zapracovány. Připomínky byly dvojího typu,a to spíše formálního nebo věcného rázu. Převažovaly připomínky formálního rázu. Neobdržela jsem žádnou zásadní věcnou připomínku, která by s navrženou metodikou byla v rozporu, šlo spíše o doplňující nebo rozšiřující připomínky.

Tento postup je běžný například při řízení projektů, kde se připomínkují jednotlivé návrhy vytvořené pro dané systémy, aby se tím zajistilo, že došlo ke správnému porozumění mezi vývojovými týmy. Každá oblast je v návrhu pokryta. Mnou byl zvolen, protože se jedná o specifickou oblast. Zvažovala jsem i dotazníkové šetření, ale domnívám se, že není vhodné pro tento typ práce, neboť autor si dotazníkem ověřuje svoji hypotézu, kdežto mým cílem bylo získat zpětnou vazbu (názory, postřehy, podněty).

Při navrhování metodiky byla použita i metoda **pozorování**, neboť jsem při navrhování metodiky využila své mnoholeté zkušenosti z praxe. Některé části metodiky byly v praxi využity při migraci a upgrade ECM platformy v telekomunikační firmě (např. tabulka s datovým modelem, tabulka definující vyhledávácí šablony, tabulka definující vkládací šablony, …).

1.3 Komentovaná rešerše

Při vyhledávání odborné literatury jsem narážela na problém s průnikem tohoto tématu mezi dvěma vědními obory. Buď jsem nalézala odborné články, které se zabývaly klasifikací z informačně-knihovnického pohledu, nebo na odborné články, které byly zaměřeny technicky, například detailně rozebíraly metody pro automatické indexování. I přes to se mi podařilo nalézt několik dokumentů, které se na oblast dívaly více analyticky a procesně a kombinovaly poznatky z obou vědních oborů.

V první řadě bych ráda uvedla organizace, které se podnikovým obsahem a technologií zabývají a které publikují odborné materiály k tématu. Organizace **AIIM** (Association for Information and Image Management) je mezinárodně uznávanou autoritou, která se oblastí podnikového obsahu zabývá dlouhodobě. Sdružuje informační profesionály, podílí se na tvorbě standardů, norem, oborných publikací a zajišťuje odborné certifikace. Organizace se informacemi v rámci technologií ECM zabývá z analytického hlediska. Hlediskem procesním, informační strategií a informačním managementem se zabývá nezisková organizace **ARMA** (Association of Information Management Professional International), která stejně jako AIIM sdružuje profesionály, vydává odborné publikace, podílí se na tvorbě standardů. Dalším zajímavým zdrojem je konzultantská společnost **Delphi Group**, která se v oboru pohybuje již 25 let. Mezi hlavní organizace, které vytvářejí normy a standardy pro tuto oblast, patří **Mezinárodní organizace pro standardizaci** (ISO), **Americká standardizační organizace** (ANSI) a mezinárodní konsorcium **W3C.**

Z pohledu monografií je zajímavá kniha od *A. Göker a J. Daviese Information retrieval: seraching in the 21st century* (2009), která se podrobně zabývá analytickým vyhledáváním, metodami používanými pro vyhledávání a automatickou indexaci a fenoménem mobilního vyhledávání. Monografie *The Taxobook: Applications, implementation, and integration in search* od M. M. K. Hlavatá (2014), popisuje proces zavádění třídění v podniku s provázaností na vyhledávání podnikových dokumentů. Kniha uznávané expertky na taxonomie Heather Hedden *The Accidental Taxonomist* (2016a), seznamuje čtenáře se základními standardy a výměnnými formáty, termínem taxonomie a představuje jeho zástupce nejen z knihovnického oboru, dále se zabývá nástroji umožňujícími návrh a správu taxonomie, automatického indexování a v neposlední řadě i procesem plánování a návrhu taxonomie. Kniha od Patricka *Lamba Organising Knowledge: Taxonomiew, Knowledge and Organizastional Effectiveness* (L2007), dopodrobna představuje termín taxonomie, jeho návaznosti na podnikové prostředí a technologie, jako je správa obsahu a správy znalostí, neopomíjí přípravu na projekt a zavádění třídění v podniku.

Z odborných článků zabývajících se tříděním na obecné úrovni, tzn. seznamují čtenáře s odbornými termíny dotýkajícími se třídění, podrobně rozebírají druhy třídění jsou zajímavé toolkit od společnosti AIIM *How to Develop Search Strategy* (2017), *How to Determine Right Taxonomy for Your Business* (2016), vědecký článek od Haniefa Bhata Mohammada Shafia *Taxonomies in knowledge organisation – need, description and benefits* (2014) a odborný text od J. Connella *Shedding Light on Building and Implementing Successful Taxonomies* (2011). Velmi často je v článcích používán pojem taxonomie (taxonomy) jakožto třídění, které je používáno uvnitř podniku naproti třídění (classification), které je spíše chápáno jako odborný termín používající se na poli informačně-knihovnickém, tzn. pro třídění monografií, seriálů apod. Někdy je tento termín chápán v širším slova smyslu, jindy v užším – blíže je tento termín představen ve 2. kapitole této práce. Při analýze těchto odborných textů mě překvapilo, že
i přes svůj informačně-knihovnický původ je pro klasifikaci dokumentů v podniku vhodný i předmětový selekční jazyk – tezaurus, který je v těchto odborných článcích též rozebírán.

Odborné články, které obsahují proces plánování, návrh, vývoj, testování a správu podnikového třídění, jsou například R. Reamy *Taxonomy Development Advice* (2007), od M.H. Bhata *Development and Deployment of Taxonomies* (2014), od S. Bayona-oré *Critical success factors taxonomy for software proces deployment* (2014) a S. Cisca *Creating*

Order out of Chaos with Taxonmies (2005). Je zajímavé, že se shodují v základních stádiích vývoje třídění popsaných výše.

Standardy a normy, které jsou pro tuto oblast zásadní spolu s komentářem, jaký mají na podnikový obsah dopad, jsou představeny ve 4. kapitole této práce.

Specifikem této oblasti je, že kvalitní a relevantní odborné texty nejsou volně dohledatelné na internetu (pouze jejich omezená část, spíše je na webu dostupný jednoduchý základní vhled do problematiky). Při získávání jsem byla nucena využívat přístupné elektronické informační zdroje VŠE a profesionální přístup do zdrojů organizace AIIM. Nejčastěji jsem využívala placené databáze ProQuest Central a EBSCO. Při prvních rešeršních dotazech jsem zjistila, že rešeršní strategie osekávání není vhodná, neboť jsem byla zahlcena velkým množstvím záznamů, proto doporučuji pro toto téma používat rešeršní strategii rostoucí perly. Při své rešerši jsem se snažila využít předmětové selekční jazyky dané databáze, například u EBSCO je vhodné využít jejich termíny s pojmy: taxonomy, knowledge management, information retrieval. Seznam neobsahuje českou odbornou literaturu, neb žádnou literaturu k tématu nelze dohledat. Praktická část staví na odborné cizojazyčné literatuře, česká literatura byla využita pouze pro teoretickou část.

Tato kapitola představila téma, kterým se diplomová práce zabývá a důvody, proč jsem toto téma zvolila. Dále formuluje hlavní a vedlejší cíle a metody, jakými budou cíle dosaženy. V neposlední řadě byla uvedena struktura, předpoklady a omezení této práce. V poslední podkapitole se nachází komentovaná rešerše, která obsahuje základní literaturu, která byla použita při psaní praktické části této práce.

TEORETICKÁ ČÁST

2 Vymezení základních pojmů

V následujících podkapitolách si osvětlíme základní pojmy, které budou v této práci používány, a jejichž význam je nutné znát pro správné pochopení praktické části textu.

V této kapitole se nenacházejí pojmy dotýkající se třídění dokumentů, neboť jsou součástí kapitoly, která se touto oblastí zabývá detailněji.

Kromě toho, že zde budou objasněny základní pojmy, považuji za nutné zde uvést také termíny, se kterými se může čtenář setkat v odborné literatuře zabývající se tématem klasifikace obsahů. Nejsou-li tato slova správně chápána, může dojít k zmatení čtenáře.

2.1 Data, informace, znalosti

Důvody, proč jsou tyto pojmy uvedeny v jedné podkapitole, jsou dva: běžně jsou pojmy „data, informace, znalosti“ používány jako synonyma, což je nejspíše způsobeno tím, že je těžké oddělit, kdy se data stávají informacemi a informace znalostmi. Druhým důvodem je fakt, že mezi pojmy existují vztahy, které jsou vyjádřené níže.

Na pojem **data** se můžeme dívat z různých úhlů, společnost ARMA je vymezuje jako: *„symboly nebo znaky, které představují nezpracovaná fakta nebo čísla, tvoří základ informací.“*(ARMA International 2007).

Výkladový slovník výpočetní techniky je naopak specifikuje jako: „*obecně jakékoli prvky s informační hodnotou, které jsou zpracovávány počítačem. Data jsou uchovávána v ucelených souborech a mohou být velmi rozmanitého typu – např. textová, numerická, binární, obrazová.“* (Hlavenka 1997).

Informace podle Malé československé encyklopedie je: *„zpráva podávaná ústně, písemně nebo jiným způsobem, s pomocí signálů nebo technických prostředků.“* Od poloviny 20. století se stává obecným vědním pojmem, užívaným v řadě oborů, který znamená (Malá československá encyklopedie 1989, s. 35):

- Zprávu o nějaké věci nebo situaci, jež se někomu předává,
- Zmenšení, snížení neurčitosti v důsledku získaného sdělení,
- Sdělení těsně spojené s řízením, signály spojené se syntaktickými, sémantickými a pragmatickými charakteristikami,

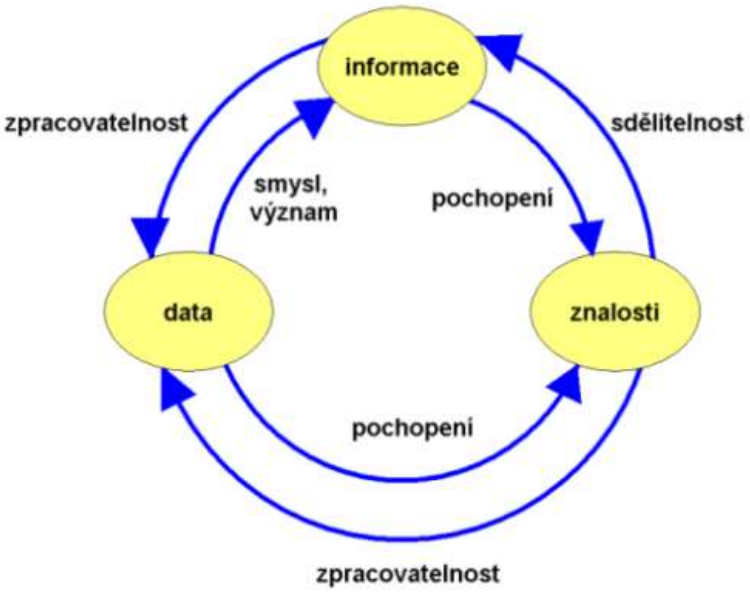
- Odraz různotvárnosti mezi libovolnými objekty a procesy v živé i neživé přírodě

Zajímavé vysvětlení termínu podává Výkladový slovník výpočetní techniky, který říká: *„data, která nesou význam pochopitelný lidmi“* (Hlavenka 1997).

Dle TDKI[1]V můžeme **znalost** definovat jako: *„schopnost člověka nebo jakéhokoli jiného inteligentního systému uchovávat, komunikovat a zpracovávat informace do systematicky a hierarchicky uspořádaných znalostních struktur. Znalost je charakterizována schopností abstrakce a generalizace dat a informací“* (Jonák 2003).

Na obrázku níže jsou vyznačeny závislosti, které se mezi termíny nacházejí. Data pohlížíme jako základní surovinu, ze které se následnými procesy tvoří informace, jež se poté přetvářejí na znalosti. Data se mění na informaci ve chvíli, kdy dávají příjemci smysl. Pokud není příjemce schopen je dekódovat, nedochází k jejich přeměně. Pro uživatele zůstane dokument plný čínských znaků do chvíle, dokud nebude schopen dané znaky přeložit do svého mateřského jazyka.

Informace se mění na znalost, pokud je člověk či inteligentní software dokáže uspořádat do souvislostí.



Obrázek 1: Závislost mezi daty, informacemi a znalostmi (Kučerová 2011)

Z pohledu informačních technologií není problém se správou dat a prací s nimi. Potíže můžeme zaznamenat již při interpretaci data na informace a následné operaci nad nimi.

[1] Pod akronymem TDKIV se nachází terminologická databáze knihovnictví a informační vědy.

2.2 Metadata, atribut

Pojem **metadata** se často definuje jako data o datech. Dle ISO normy 15489 je tento pojem specifikován jako: *„údaje popisující souvislosti, obsah a strukturu záznamů a jejich správu v čase.“*(ISO 15489-1, 2001). Tato definice je následně přejímána do jiných norem a standardů.

Podobnou definici můžeme nalézt v ISO normě 23081 – 2, která říká, že metadata jsou: *„strukturované nebo částečně strukturované informace, které umožňují tvorbu, správu a používání záznamů v čase v rámci domény nebo i mezi doménami.“* (ISO 23081-2, 2009).

Tento pojem se používá pro popis, specifikaci dokumentů a obsahů v rámci technologie ECM[2]. V běžné praxi se můžeme setkat i s termínem **atribut**, který můžeme vymezit jako: *„podstatný znak, neodlučitelná vlastnost“* (Atribut, 2005-2017), většinou je používán uživateli či IT profesionály, kteří mají zkušenosti s relačními databázemi.

Vzhledem k tomu, že pojem metadata existuje v českém jazyce pouze v množném čísle, bude pro jednotné číslo v práci využíván termín **metaúdaj**, neboť údaj lze za určitých okolností vnímat jako synonymum k termínu data.

Jak již definice v ISO normě 23081 – 2 naznačuje, metadata můžeme dělit na strukturované, polostruktrované či nestrukturované.

Metadata můžeme dělit také podle účelu použití na metadata pro obchod, pro ochranu, popisná metadata, metadata pro vyhledávání, pro správu oprávnění.

Jsou používána ze čtyř hlavních důvodů (AIIM Training Programs):

- Identifikace obsahu – jde o jedinečnou identifikaci, která odliší obsah od jiných, mohou zde být použita metadata typu URI, ISBN, referenční číslo, interní číslo v systému.
- Správa obsahu – jedná se o metadata, která pomáhají systému při správě obsahu, kdy mají být jaké funkce/procesy nad obsahem spuštěny, například umožňují identifikovat, které aplikace mohou být použity při zobrazení, kdy má dojít k výmazu dokumentu ze systému. Jde například o metaúdaje typu číslo verze, datum archivace, formát dokumentu.
- Vyhledávání obsahu – metadata, která pomáhají vyhledat relevantní dokument. Zde jsou metaúdaje typu tezaurus, klíčová slova, řízený slovník, heslář.

[2] Termín ECM je v textu následně blíže vysvětlen a rozebrán.

- Sledování využívání obsahu – metadata, která sledují využívání obsahu a na základě toho ho připojují k jinému obsahu. Velmi často se s touto funkcí setkáme u elektronických obchodů, kde se uživateli při prohlížení produktu nabízejí i jiné relevantní produkty, které by ho mohly zaujmout, např. z toho důvodu, že je tento produkt často zakoupen s jiným produktem nebo službou.

Robert F. Smallwood (2004) ve své knize *Information Governance: Concepts, Strategies and Best Practices* dělí metadata na:

- administrativní – prostřednictvím těchto metadat dochází k řízení obsahu uvnitř podniku, správu oprávnění a vlastnictví;
- popisná – popisují obsah dokumentu (předmět, autor, identifikátory,…);
- konzervační – metadata, která umožňují trvalé uchování digitálního obsahu a jeho zpřístupnění v čase, mimo jiné zajišťuji jeho pravost a integritu;
- konstrukční – metadata, která se používají pro zobrazení a procházení souborů, popisují vnitřní uspořádání souboru (např. členění dokumentu, pořadí stránek,…);
- technická – popisují funkce digitálního objektu jako je jeho rozlišení, potřebný hardware. Tyto informace jsou důležité zejména z pohledu migrace na jiný formát.

Pro správu dokumentu v čase je v normách chápáno podpoření celého životního cyklu dokumentu. Životní cyklus je sled jednotlivých fází, kterými prochází dokument od svého vzniku, přes jeho publikování, následnou archivaci až po svůj zánik.

2.3 Dokument, obsah, záznam

Slovo **dokument** pochází z latinského slova documentum a dle slovníku naučného je specifikován jako: *„důkaz, zvláště pak listina, která něco dokazuje.“*(Slovník naučný 1862, s. 239)

Naproti tomu ISO norma 15489 vymezuje tento termín jako: *„zaznamenaná informace nebo objekt, se kterým je možné zacházet jako s celkem.“*(ISO 15489-1, 2001)

V rámci této diplomové práce bude dokument používán ve významu, jakým ho definuje ISO norma. V definici je důležité, že je možné s informací či objektem zacházet jako s celkem, nezáleží, kolik částí má, nebo je-li fixována na nějakém nosiči.

Dokumenty je možné dělit dle standardu **ISBD**, mezinárodního standardu pro bibliografický popis, na: monografie (ISBD(M)), seriály (ISBD(S)), tištěné hudebniny (ISBD(PM)), počítačové soubory (ISBD(CF)), neknižní dokumenty (ISBD(NBM)), kartografické dokumenty (ISBD(CM)), staré tisky a prvotisky (ISBD(A)) a elektronické zdroje (ISBD(ER)). V tomto standardu mimo jiné nalezneme i dělení dle formy dokumentu, a to na: soubor dat, obrázek, pohyblivý obrázek, hudbu, objekt, program, zvuk, mluvené slovo, text, multimediální dokument a ostatní typy dokumentů.

Dle Komenského slovníku naučného můžeme **obsah** vymezit jako: *„souhrn všech vlastností, které vyjadřují charakter předmětu nebo pojmu.“* (Komenského slovník naučný 1938, s. 350)

Tato definice z pohledu využití v rámci technologie ECM není vyhovující, proto budeme obsah chápat spíše jako informace a zkušenosti, které mohou v konkrétním kontextu poskytnout konečnému uživateli přidanou hodnotu. Je jedno, je-li ve formě elektronického dokumentu, obrázku, videa apod.

V rámci této práce budou termíny dokument a obsah používány jako synonymní výrazy.

S pojmem **záznam** se v praxi nejčastěji setkáváme v rámci ERMS[3] aplikace. ISO norma 15489-1:2001 tento pojem specifikuje takto: *„informace, které společnost nebo jednotlivec vytváří, přijímá a uchovává jako důkazy a informace, a to na základě právních povinností nebo během obchodních transakcí.“*(ISO 15489-1, 2001)

2.4 Zákazník

Protože cílem této práce je klasifikace zákaznických dokumentů, domnívám se, že je vhodné vymezit pojem zákazník, i když je tento pojem znám a běžně se s ním setkáváme.

Na pojem se můžeme dívat ze dvou hledisek (Customer 2007):

- Buď jde o stranu, která přijímá nebo spotřebovává výrobky (zboží či služby) a má schopnost volit mezi různými produkty a dodavateli.

[3] Electronic Records Management System (ERMS) vymezuje norma ANSI/AIIM/ARMA TR48-2006 jako *„elektronická správa elektronických a neelektronických záznamů využívající počítačové vybavení a software v souladu s uznávanými zásadami a postupy správy dokumentů.*“(AIIM a ARMA, 2004)

- Nebo oddělení v rámci firmy, která stanovují požadavek na proces (účetnictví) a přijímají výstup procesu (finanční výkaz) od jednoho nebo více interních nebo externích dodavatelů.

V této práci použijeme slovo zákazník jen v prvním ze zmiňovaných významů.

2.5 Firma, podnik

Podnik je obchodním zákoníkem v § 5 vymezen jako: *„soubor hmotných, jakož i osobních a nehmotných složek podnikání.“* (Česká republikaObchodní zákoník, 1991), ten samý zákon vymezuje v § 8 i pojem firma, je to: *„název, pod kterým je podnikatel zapsán do obchodního rejstříku.“* (Česká republikaObchodní zákoník, 1991) Jinými slovy firma je, jako nehmotný majetek, součást podniku. Z čehož například plyne, že je právně převoditelná pouze společně s podnikem (Tietzová 2008).

Pro potřeby této práce bude pojem firma používán synonymně a to z důvodů zajištění čitelnosti textu.

2.6 ROT obsah

V odborných anglických textech týkajících se technologií ECM se setkáváme se zkratkou **ROT**. Jde o počáteční písmena anglických slov redundant, obsolete a trivial. V doslovném překladu je to obsah, který je redundantní, zastaralý a triviální. Toto označení se využívá pro typ dokumentu, který nemá pro podniky žádnou přidanou hodnou, neboť se jedná buď o zastaralou informaci, je to kopie jiného obsahu, nebo obsah, který nikdy nebyl pro firmu potřebný, jedná se o soukromý obsah uživatele (písničky, videa,…). S technologií ECM bývá spojován z důvodu, že by měla být nápomocná při jeho odstraňování nebo vzniku.

Dle studie (Mancini 2011), může jít až o 1/3 dokumentů na neřízených sdílených discích. Důvodů, proč by měly podniky tento druh obsahu vyhledávat a mazat, je několik:

- Díky ROT může docházet ke snížení produktivity, dobré pověsti firmy;
- Když lidé opouštějí podnik, jejich ROT po nich zůstává a často je přidělován zaměstnanci, který po nich jejich práci vykonává;
- Navyšuje náklady na pořízení disků a jejich záloh.

2.7 Prohlížení, analytické vyhledávání, mobilní vyhledávání

Uživatelé při vyhledávání informací na webu, ve specializovaných databázích či aplikacích, používají dvě základní metody vyhledávání.

První z nich je **prohlížení** nebo také listování (browsing), které DTKIV vymezuje jako: *„základní způsob pohybu v informačním prostoru služby WWW prohlížením obsahu webových stránek s využitím hypertextových odkazů. Je protikladem a doplňkem k přímému vyhledávání informací“* (Sklenák 2003). Tato metoda je pro uživatele intuitivní, zajistí přehled v rámci hledané oblasti. Je to přirozené chování pro malé děti, uživatele bez informační gramotnosti a těch, kteří chtějí hledané téma vidět v souvislostech.

Druhou metodou je **analytické vyhledávání** (searching, retrieval), kdy již uživatel musí znát základní principy, jak formulovat rešeršní dotaz v rámci daného systému.

I když se na první pohled může zdát, že jde o zcela rozdílné metody a stojí proti sobě, není tomu tak. Metody se vzájemně doplňují a uživatel i v rámci řešení stejného problému může použít, střídat obě metody (Papík 2014).

Použitá metoda ovlivňuje typ a způsob používání metadat, blíže je to popsáno v praktické části této práce.

Rozšíření mobilních zařízení[4] má dopad i na potřeby a způsoby vyhledávání, pro tento typ se používá termín **mobilní vyhledávání**. Toto vyhledávání má svá specifika (Göker, 2009). Jeho hlavním rysem je to, že uživatelé většinou vyhledávají informace, které jsou relevantní k jejich aktuální činnosti či poloze, popř. potřebují něco naplánovat (hledají restauraci, zajímavosti v okolí apod.), z toho důvodu dochází k velkému množství rychlých dotazů oproti vyhledávání na počítači (uživatel provede rychlé vyhledání a mobilní zařízení odloží). Uživatelé mají málo času a místa na obrazovce, z toho důvodu by výsledky vyhledávání nebo kroků, které musí uživatel udělat, mělo být co nejméně. Uživatelé při tomto typu vyhledávání nejsou na vyhledávání plně soustředěni, neboť zůstávají v kontaktu se svým okolím (konverzují, musí hlídat dopravní situaci, apod.), proto potřebují, aby bylo vyhledání rychlé a cílené na přesnost než úplnost výsledků. A proto potřebují relevatní a aktuální informace, oproti pouhé relevantnosti v rámci hledané oblasti, která je při práci s počítačem postačující. Aby se toho docílilo, používají se při mobilním vyhledávání funkce filtru, který může být nastaven automaticky nebo manuálně.

[4] V tomto případě jsou mobilní zařízením chápány všechna zařízení, které jsou malé a uživatelé je nosí běžně sebou, jsou to mobilní telefon, PDA, tablety apod.

2.8 Podnikové vyhledávání (enterprise search)

Podnikové vyhledávání (v ang. používaný termín enterprise search), dle organizace AIIM[5], můžeme definovat jako: *„identifikace a zpřístupnění konkrétního obsahu v rámci celého podniku k indexaci, vyhledání a zobrazení oprávněným uživatelem"*(AIIM 2017b).

Obsah bez přístupu je bezcenný. Podnikové vyhledávání je o tom, jak organizace pomáhá uživatelům, aby mohli hledat potřebné informace v jakékoliv formátu, odkudkoliv uvnitř daného podniku – v databázi, v DMS, na papíře apod. Získání správné informace ve správný čas.

V rámci této oblasti se v odborné literatuře můžeme setkat s otázkami, zda používat nějaký typ klasifikace nebo využít fultextové vyhledávání. V době kdy jsou výkonné stroje se může zdát, že masovému využívání fultextového vyhledávání nic nebrání. Nicméně právě zde se naráží na to, jak data správně interpretovat, že může být slovo v různých kontextech vnímáno rozdílně. Uživatel je pak často zahlcen velkým množstvím nerelevantních dokumentů. Proto i dnes má klasické tříděné své místo, i když ho existují inciativy spojené například se sémantickým webem, které se snaží o zvýšení kvality v rámci fultextového vyhledávání či automatické indexace, upozaďují (AIIM 2017b).

Schopnost najít relevantní informaci je to co uživatelé potřebují, aby mohli dobře vykonávat svoji práce. Nejde o snadný úkol, jak je již z výše popsaného pochopitelné. Není to jen o vepsání hledaných termínu do vyhledávacího pole, ale jedná se spíše o objevování věcí dotýkající se hledaného tématu.

Síla podnikového vyhledávání je v tom, že uživatel používá jednotou bránu, která prohledává jednotlivé dílčí podsystémy. Způsob jakým může podnikové vyhledávání ovlivňovat tvorbu, využívání klasifikace v rámci podniku je popsáno v praktické části této práce.

2.9 Spisový a skartační řád

Spisový řád je souhrn předpisů, jehož účelem je zabezpečení správy dokumentů (zajistit jednotnou evidenci a manipulaci).

[5]Společnost AIIM byla založena v roce 1943 jako National Micrographics Association a následně přejmenována na Association for Information and Image Management – odtud pochází její akronym. Dlouhodobě se zabývá práci s informacemi v rámci technologie ECM.

Součástí řádu je **spisový a skartační plán**, který definuje přehled typů dokumentů se spisovými znaky, s datem, do kdy budou pro podnik potřebné (skartační lhůta) a jak se s nimi po uplynutí doby naloží (skartační znak).

V angličtině se setkáme s termínem **Business Classification Scheme** (BSC). Je to klasifikační schéma v hierarchické struktuře, které má na hlavní úrovni obchodní funkce a na nižších se rozpadá na aktivity podniku. Bývá často doplněn o skartační a archivní plán. Výhodou dělení na funkce, aktivity a úkoly oproti organizačnímu dělení je, že při každé reorganizaci podniku není nutné promítat změny i do schématu. Příklad třídění z univerzity v Plymouthu je na obrázku 2.

Ref.	Section	Function	Activity	Sub-activity	Notes
A	GOVERNANCE & STRATEGY				
A.1		GOVERNANCE	Legal Framework Development		Instrument and Articles of Government, and other regulations
			Governance Structure Development		
			Governing Body	Development	Board of Governors
				Meeting papers	Agenda, minutes & papers
			Executive Committee	Development	UEG
				Committee papers	Agenda, minutes & papers
			Senior Managers Appointments		Vice-Chancellor, PVC, etc.
			Committees, Boards and Working Groups	Development	Separate folder for each one
				Committee papers	Agenda, minutes & papers
A.2		STRATEGY, POLICY AND PROCEDURE	Strategy	Development	This section is common to all functions and can be used across all Faculties and Directorates
				Planning	
				Performance	
			Policy		
			Procedures		

Obrázek 2: Příklad BSC Univerzity v Plymouthnu (Plymouth University 2017)

Tyto termíny zmiňuji, neboť i zákaznických dokumentů se týká archivace a skartace. Je tedy nutné neopomenout tuto oblast při návrhu povinných metadat k dokumentům. Obsahuje je i srovnávací tabulka v příloze číslo 2.

2.10 ECM

Protože už v názvu práce používám zkratku, považuji za nutné ji čtenáři osvětlit. Půjde o pouhý základní vhled, neboť není obsahem této práce zabývat se technologiemi ECM do detailu.

Pod zkratkou ECM se skrývá **Enterprise Content Management**, který vzhledem k obrovskému množství strukturovaného a nestrukturováno obsahu v podnicích nabývá v dnešní době velkého významu. Jde o relativně nový termín, proto není jeho obsah zcela

ustálen. I přesto se společnost Gartner v prosinci 2016 rozhodla, že používání této zkratky je zastaralé a snaží se pro tuto technologii prosadit termín Content Services.

Společnost AIIM, které byla již výše představena, tento termín vysvětluje jako *„strategii, metody a nástroje používané k zachycení, správě, uložení, uchovávání a dodávání obsahu a dokumentů vztahujících se k organizačním procesům. Nástroje a strategie ECM umožňují správu nestrukturovaných informací dané organizace bez ohledu na jejich umístění.“* (AIIM 2011)

Důvodů, proč jsou technologie ECM v podnicích implementovány, je hned několik: umožňují řídit životní cyklus dokumentů, zřizovat k nim oprávnění pro jednotlivé uživatele, zabezpečit obsah před neoprávněným zneužitím a poškozením, sdílení obsahu napříč jednotlivými odděleními, což má za cíl snížit oběh papírových dokumentů.

Podobně jako v oblasti marketingu, tak i zde existuje **5C**: obsah (content), spolupráce (collaboration), vyhovění zákonným předpisům (compliance), kontinuita (continuity), náklady (cost) (Kučerová Zrálíková 2011).

Trochu překvapivé je, že pod termínem obsah jsou krom samotného digitálního objektu zahrnuta i metadata, záznamy, data, vše, co do systému ECM vstupuje a je jím řízeno. Obsah zde může být statický (neprochází změnami) nebo dynamický (prochází v čase různými změnami), obsah, který podléhá zákonným normám (archivace a skartace), obsah, podobně jako metadata, může nabývat strukturované (formuláře, dotazníky,…), semi-struktrované (část dokumentu je strukturovaná a v jiné části je volný text) a nestrukturované (kniha, článek, příspěvek,…) podoby.

Díky ECM je možná spolupráce a výměna informací nezávisle na čase (dostupnost systému 24 hodin denně 7 dní v týdnu) a místě, nalezneme zde chaty, diskuzní fóra, blogy, wiki stránky, video konference.

Velmi často je využívána pro zákonné splnění podmínek, vnitropodnikových nařízení (v angličtině používán termín compliance), což s sebou kromě jiného nese i bezpečné uložení dokumentů a řízení bezpečného přístupu k nim (logování činnosti, ochrana dat před neoprávněným odmazáním, zneužitím). Ke známým předpisům, zvláště v angloamerickém světě, patří Saranes-Oxley zákon (SOX), Mezinárodní účetní standardy (IAS), HIPPA, Basel II. nebo normy dotýkající se IT, jako jsou ITIL, CoBIT, Sig Sigma, CMMI, PMI (Kučerová Zrálíková 2011).

V poslední době je v praxi patrný trend snižování nákladů, což se společnostem díky ECM nástrojům může podařit (je možné snížit tiskové náklady, náklady na jednotlivá distribuovaná uložiště, zrychlení jednotlivých procesů a zvýšení dostupnosti dokumentů).

V rámci platformy ECM se setkáváme s různými technologiemi, uveďme některé z nich: správa záznamů, správa dokumentů, správa webu, znalostní báze, správa elektronické pošty, archivace dokumentů, správa multimediálních datových zdrojů a jiné.

V této kapitole byly představeny základní termíny, které mají souvislost s tématem diplomové práce. Uvědomuji si, že výčet termínů není úplný, vzhledem k šíři tématu to ani není možné. Vynechané termíny nejsou pro práci zásadní a práce se jich dotýká pouze okrajově, k takovým termínům patří například informační management, informační politika apod.

3 Katalogizace dokumentů

„Taxonomie je někdy popisována jako věda
a někdy jako umění, ale ve skutečnosti je to bojiště.“

Bill Bryson

V této kapitole jsou popsány termíny, které se používají pro oblast třídění dokumentů. Jsou zde vyjasněny základní termíny z knihovnické vědy, ale také znovu objevené termíny, které byly vyvolány s příchodem nových technologií. Na konci se krátce zmíníme o softwarech, které se specializují na správu a tvorbu věcných selekčních jazyků.

Oblastí třídění dokumentů se zabývá informační věda a knihovnictví, počátky třídění nacházíme v antickém Řecku, kdy Aristoteles třídil vědy na praktické, teoretické a poetické. Třídění a s tím spojené vyhledávání informací jsou následně vystopovatelné od vzniku prvních knihoven (to je od starověku), kdy si knihovny uspořádávaly dokumenty dle obsahu (dle vědních oborů či oblastí lidské činnosti), aby je bylo možné snáze vyhledat. Pro vyhledávání byly používány katalogy, které obsahovaly záznamy opět tříděné dle vědních oborů či oblastí lidské činnosti. Hlavním cílem katalogů bylo vyhledat relevantní dokument – buď pomocí identifikačních údajů (autor, název dokument, …) nebo z hlediska věcného (obsahového).

Činnost, která se zabývá analýzou a popisem dokumentů, je **katalogizace**, kterou TDKIV vymezuje (Vodičková, 2003):

- *Jmenný a věcný popis dokumentů, někdy chápaný úžeji pouze jako jmenný popis dokumentů, jindy obecněji jako tvorba katalogizačních záznamů pro různé druhy katalogů a jejich organizace.*
- *V době budování lístkových katalogů nauka o knihovních katalozích a o metodách jejich budování včetně jejich organizace.*

Věcný popis nebo také věcná klasifikace patří do dílčího procesu vstupního zpracovávání dokumentů a dle TDKIV ho můžeme definovat jako (Balíková 2003f):

- *Proces tvorby věcných selekčních údajů a jejich přidělování dokumentům.*
- *Nauka o věcných katalozích a o metodách jejich tvorby.*

Pod věcnými selekčními údaji je možné si představit sestavu pojmů, které vypovídají o obsahu, tématu či předmětu dokumentu jako celku a bývají výstupem obsahové analýzy.

Klasifikace má za cíl vytvořit selekční obraz dokumentu, který na základě dostupných technologií umožní koncovému uživateli dosáhnout optimální přesnosti a úplnosti při efektivním vyhledávání dokumentů.

Z obsahové analýzy vychází věcné pořádání informací, které se definuje jako (Balíková 2003g): *„specifický proces pořádání informací, při kterém jsou informace získané obsahovou analýzou dokumentu vyjádřeny pořádacími znaky věcných selekčních jazyků. Podle typu pořádacích znaků a principu organizace informací se vyčleňuje předmětové a systematické pořádání informací.“*

V rámci systematického pořádání informací dochází k vřazování informací o obsahu dokumentu do systému vědních oborů. Slovní formulace je nahrazena znaky umělého jazyka.

Při předmětovém pořádání dochází k převádění výsledků obsahové analýzy z přirozeného jazyka do podoby hesel.

Pro pořádání, ukládání a vyhledávání informace se používá **věcný selekční jazyk,** který je charakterizován množinou termínů, jež dělíme na systematické a předmětové. Dle koordinace jednotlivých znaků můžeme dále věcný selekční jazyk dělit na prekoordinovaný (znaky jsou syntakticky uspořádány v celek a není možné měnit jejich pořadí) a postkoordinovaný (znaky nejsou vytvářeny do celku, lze je využívat samostatně a vytvářet různé kombinace). Následně se setkáme:

- s prekoordinovaným systematickým selekčním jazykem (MDT, DDT);
- s postkoordinovaným systematickým selekčním jazykem (Ranghanatanovo dvojtečkové třídění);
- s prekoordinovaným předmětovým selekčním jazykem (předmětová hesla);
- s postkoordinovaným předmětovým selekčním jazykem (klíčová slova, deskriptorové selekční jazyky).

3.1 Prekoordionované selekční jazyky

Mezi neznámější prekoordinované systematické selekční jazyky patří Deweyho desetinné třídění (DDT) a Mezinárodní desetinné třídění (MDT).

K prekoordinovaným pědmětovým selekčním jazykům patří předmětová hesla.

3.1.1 Deweyho desetinné třídění

Melvin Dewey, který se narodil 10. prosince 1851 v Americe, je považován za otce moderního knihovnictví. Jeho desetinné třídění je jedním z úspěšných klasifikačních systémů po celém světě, zvláště v angloamerické oblasti.

Již od raného dětství se projevoval jeho matematický talent a i přes chudé poměry se mu podařilo dostat se na Amherst College v Massatchusses, kde si jako knihovník vydělával na školné. Deweye zděsil stav, jakým způsobem jsou knihy v knihovně uspořádány, každá knihovna používala vlastní systém a uživatelé museli hledat na různých místech, než se jim podařilo nalézt knihu k hledanému tématu (praxí bylo řazení knih do regálu dle přírůstkového čísla, což způsobilo, že nebyly u sebe knihy tématicky či obsahově podobné).

Mezi vlastnosti **Deweyho desetinného třídění** patří alfanumerická notace, přesně definované kategorie, rozvinutá hierarchie a bohatá síť vztahů mezi tématy.

K přednostem DDC patří to, že je možné si ho rychle osvojit. Poskytuje skvělý způsob, jak řadit knihy do regálu, nicméně vzhledem k tomu, že byl vytvořen pro malé americké univerzitní knihovny, jeho použití pro veřejné knihovny se jeví jako méně vyhovující a velmi obtížně se přizpůsobuje potřebám dnešní informační společnosti.

Podrobněji se Deweyho desetinným tříděním zabývá literatura (Soukupová 2000).

3.1.2 Mezinárodní desetinné třídění

Mezinárodní desetinné třídění je univerzální hierarchická klasifikační soustava využívající desetinné notace pro indexaci dokumentů, které mají vztah k oblastem lidského vědění.

Podobnost MDT s Deweyho desetinným tříděním není náhodná. Vznik MDT je spjat s Mezinárodní bibliografickou konferencí, která se uskutečnila v roce 1895 v Bruselu, kde se mimo jiné rozhodlo o založení mezinárodního bibliografického ústavu, který bude mít rozvoj tohoto třídění na starosti. U zrodu třídění byli Paul Otlet a Henri-Marie Lafontain, kteří se nejdříve pokusili využít DDM beze změn, ale protože jim třídění pro jejich potřeby nedostačovalo, rozšířili ho o celý systém znaků. V roce 1905 bylo publikováno první úplné vydání MDT.

MDT patří mezi nejrozšířenější a nejpoužívanější klasifikaci na světě. V dnešní době se o její vývoj stará the UDC Consortium.

Podrobněji se Mezinárodním desetinným tříděním zabývá literatura (Filoušová 2002).

3.1.3 Předmětová hesla

Předmětová hesla jsou jednoslovné či víceslovné formalizované výrazy, které vystihují obsah dokumentu. Při kategorizaci používáme heslář (seznam hesel – slouží k odstranění subjektivního zpracování), pro bližší specifikaci bývá někdy heslo doplněno podheslem, jež pojem zpřesňuje. Jak heslo, tak i podheslo může být ještě rozšířeno o doplněk, který je uveden v závorce. Dle obsahu je možné rozlišovat hesla a podhesla geografická, historická, tematická, chronologická, korporativní, názvová, personální a formální.

I přesto, že je obsah dokumentu vyjádřen v přirozeném jazyce, je nutné při zápisu dodržovat předepsanou strukturu. Pravidla předmětového seznamu vyšla v roce 1952.

Jako příklad můžeme uvést knihu, která se zabývá komunikací koček (Marek, 1999).

Heslo – podheslo: Kočky – komunikace

Příklad doplňku: *Halas, František* ***(1901 - 1949, český básník)****

Nevytváří se zde hierarchická struktura. Známý předmětový slovník je LCSH (Library of Congress Subject Headings) nebo polytematický strukturovaný heslář (PSH), který je využíván mimo jiné i Národní technickou knihovnou v Praze.

Podrobněji se předmětovými hesly zabývá literatura (Předmětová hesla 2010 a Veličková 1993).

3.2 Postkoordinované selekční jazyky

Postkoordinovaný systematický selekční jazyk zastupuje dvojtečkové třídění Ranghanatana.

K postkoordinovaným předmětovým selekčním jazykům patří deskriptorové selekční jazyky a klíčová slova.

3.2.1 Randhanatanovo dvojtečkové třídění

S nárůstem informací přestávají stávající klasifikační schémata na začátku 20. století dostačovat. Na situaci reaguje indický knihovník Shiyali Ramamrita Ranghanaton (1892-1972) se svým dvojtečkovým tříděním (anglicky Colon Classification). Ranghanatan patří mezi významné postavy knihovnických dějin, působil v různých mezinárodních orgánech, které se věnovaly selekčním jazykům. V roce 1924 se stal knihovníkem univerzitní knihovny v Madrasu a odjíždí na studijní pobyt do Anglie, kde se formuje jeho názor na věcné pořádání informací – nelíbí se mu DDT ani MDT. Proto ve své práci odmítl enumerativní princip

stávajících klasifikačních systémů (vyjmenovávání veškerých procesů, předmětů,… jenž mají být obsahem dokumentu). Podle něj je tento postup nevhodný, protože se jím velmi těžko vyjadřují složitá témata (například řízení základního školství ve Velké Británii v 50. letech). Nicméně i on musel rozdělit klasifikaci do základních tříd. Hlavní třídy již nerozkládá na další podtřídy.

Při obsahové analýze dokumentu dochází k vydefinování charakteristických stránek dokumentu, které se nazývají fasety a k syntéze třídících znaků připravených v tabulkách pro rozmanité fasety. Protože pro správné vyhledání je nutné dát fasety do správného pořadí, je toto třídění též nazýváno analyticko-syntetickým. Fasety jsou ve třídě označeny arabskými číslicemi.

Předností toho třídění je jednoduchost a flexibilita při vyjadřování složitých termínů (v rámci MDT je nutné vybrat aspoň jedno téma, což vede k tomu, že se použijí témata, která jsou obsahu nejblíže, ale nemusí ho zcela vystihovat). Další výhodou je možnost kombinovat jakýkoliv termín s dalším pojmem z tabulek, aby se dosáhlo nejlepšího vyjádření obsahu dokumentu, což může vést k nekonečnému množství číselných variací. Dále je možná rychlá aktualizace tabulek, což podporuje rychlý rozvoj věd a informací.

K nevýhodám patří délka a konstrukce klasifikačního znaku, obtížné zapamatování pro uživatele.

Dvojtečkové třídění se nikdy plošně neujalo díky svým nevýhodám, ale dalo základ fasetové analýze, což vedlo v 60. letech minulého století k částečnému přepracování hlavních klasifikačních schémat na fasetový princip. Mezi ovlivněné klasifikace patří Broad System of Odering nebo The British Catalogue of Music Classification, vliv mělo i na oblast deskriptorových selekčních jazyků a tezaurů (Art and Architecture Thesaurus).

S rozvojem Word Wide Web v 90. letech minulého století byly podniknuty pokusy o použití fasetového principu, analýza by mohla spojit dva základní přístupy, vyhledávání pomocí vyhledávačů a předmětového hesláře tvořeného lidmi (např. Yahoo!).

Podrobněji se dvojtečkovým tříděním zabývá literatura (Dofková 2004).

3.2.2 Klíčová slova

Dle TDKIV můžeme klíčové slovo charakterizovat jako (Balíková 2003d): *„významové slovo nebo slovní spojení, obvykle ve formalizované podobě, vybrané z názvu nebo textu dokumentu a sloužící jako věcný selekční údaj.“* Jejich základní vlastností je stručnost, která umožňuje rychlou orientaci v obsahu.

Klíčová slova se většinou vybírají z názvu dokumentu, ale mohou být vybrána i z jiných části jako jsou anotace, extrakty, referáty, někdy připojuje k dokumentu klíčová slova sám autor.

Klíčová slova vystihují téma dokumentu, zároveň jsou předmětem informačního dotazu (budou použita při vyhledávání). Podobají se přirozeným heslům, ale na rozdíl od nich je jejich spojení vzájemně nezávislé a neuspořádané.

Při použití klíčových slov dochází k tomu, že systém porovná jednotlivá slova či slovní spojení se slovy a slovními spojeními uvedených u dokumentů. To s sebou většinou nese seznam dokumentů, které zcela neodpovídají uživatelovým potřebám.

3.2.3 Deskriptorové selekční jazyky

Patří mezi postkoordinované selekční jazyky, které jsou definovány jako (Balíková 2003e): *„selekční jazyk určený pro indexaci a vyhledávání prostřednictvím řízeného slovníku označovaného jako tezaurus, jehož lexikum je složeno z deskriptorů a nedeskriptorů.“*

Tyto jazyky dělíme na dva základní druhy, unitermi a selekční jazyky založené na deskriptorech. Též je možné dělit je dle rozsahu tematického zaměření na univerzální a speciální (zahrnují pouze část lidského poznání).

Tezaurus je deskriptorový selekční jazyk, jenž má formu standardizovaného slovníku se strukturou (jsou zde zachyceny vztahy nadřazenosti, podřazenosti). Vznik tezaurů se datuje po 2. světové válce, kdy došlo k poznání, že stávající selekční jazyky nejsou dostačující a je nutné začít efektivně řešit informační explozi.

Tezaurus obsahuje deskriptor – preferovaný termín, a nedesktriptor – nepreferovaný termín. Ten slouží jako informace pro uživatele, že jde o nepreferovnaý termín, a odkazuje na ekvivalentní preferovaný termín.

Postup při indexaci dokumentů (Strenková, 2002):

- Prostřednictvím obsahové analýzy určíme obsah dokumentu (základní téma, doplňková témata);
- Formulujeme obsah pomocí klíčových slov;
- Klíčová slova převedeme na deskriptory daného tezauru.

Práce s tezaurem specifikují normy ISO 2788, ISO 5964 a ISO 25964.

3.3 Klasifikace, indexace a související pojmy

Při čtení odborné literatury se nepoučený čtenář může ztrácet, neboť pro přiřazování metadat k dokumentu se používá několik odborných termínů s rozdílným významem, ale v určitém kondextu mohou být využívány synonymně.

Termín **klasifikace** vysvětluje TDKIV jako (Balíková 2003c): *„uspořádání pojmů nebo jiných entit do tříd a dalších skupin (podtříd apod.) za účelem vyjádření jejich sémantických vztahů. Třídy a další skupiny jsou obvykle vyjádřeny notací. Nebo přidělování notací (znaků tříd) klasifikačního systému za účelem vyjádření obsahu dokumentu.“*

Indexaci TDKIV definuje jako (Balíková 2003b): „*proces vyjádření obsahu dokumentu pomocí prvků selekčního jazyka, obvykle s cílem umožnit zpětné vyhledávání. Podle použitých metod se rozlišuje pojmová a slovní indexace, podle použitých postupů se rozlišuje intelektuální, automatická a poloautomatická indexace. Z hlediska použitých selekčních jazyků se rozlišuje prekoordinovaná indexace a postkoordinovaná indexace.“*

I pojem **kategorizace** znamená třídění dokumentů do jednotlivých kategorií na základě společných vlastností. Často se s tímto termínem setkáme u automatického zatříďování obsahu nebo při přidělování štítků (tagů) k obsahu.

Pojmy třídění, klasifikace, indexace jsou v práci používány jako synonyma. Termín kategorizace je použit u automatického třídění.

DTKIV definuje i pojem **folksonomie** (Havlová 2003): *„schéma kategorizace obsahu internetu, kdy uživatelé označují obsah WWW stránek popisky (tagy) podle svého výběru za účelem utřídění těchto stránek a jejich rychlejšího vyhledávání. Tagy a hierarchie pořádacího systému jsou plně řízeny uživatelem v závislosti na jeho požadavcích.“* S touto možností se setkáváme už i v rámci technologie ECM.

Způsob, jaký můžeme použít pro přiřazování selekčního jazyka dokumentům, dělíme dle použitých postupů na intelektuální, automatický a poloautomatický (Schwarz 2003).

Intelektuální indexace je náročný proces, jehož kvalita je ovlivněna řadou faktorů, například znalostí selekčního jazyka, oboru, schopností porozumět textu, systematickým myšlením, ale i únavou, náladou či motivací. Proto bývá upřednostňována automatická klasifikace, při níž tyto faktory nehrají roli, a je cenově dostupnější. Navíc enormní nárůst elektronického obsahu není možné kategorizovat pouze lidskými. Z pohledu kvality výstupu je však stále manuální indexace nejlepší.

Poloautomatická indexace spočívá v kombinaci automatických a intelektuálních postupů. Během procesu je část provedena automaticky (z textu jsou předvybrány vhodné termíny) a výstup slouží jako podklad pro intelektuální indexaci (je na indexátorovi[6], aby navrhované termíny potvrdil či zamítl). I tato metoda má svá omezení: obsah indexovaného dokumentu může být vyjádřen i jinými výrazy přirozeného jazyka, než které jsou v dokumentu obsaženy. Nastávají tři typy situací (Strenková 2002):

- dva různé výrazy mají stejný význam (synonyma) – klisna = kobyla;
- dva různé výrazy jsou ve vztahu obecné – konkrétní (nadřazené – podřazené) – zvíře > hospodářské zvíře > kůň > hřebec;
- dva různé výrazy mají pouze v širším smyslu něco společného – kůň a dostih.

Posledním typem je automatická indexace, kterou můžeme rozdělit na dva základní typy (Schwarz 2003):

- Automatická extrakce (někdy nazývána jako slovní indexace) – založená na extrakci indexačních termínů přímo z plného textu.
- Automatické přiřazování (někdy nazýváno jako pojmová indexace) – určování indexačních termínů z řízeného slovníku či znalostní databáze, na základě porovnání s výrazy z plného textu dokumentu.

Automatická extrakce je jednodušší a snadno programovatelná metoda, kdežto automatické přiřazování je pro praktickou aplikaci náročnější. V procesu automatického přiřazování mohou být použity i metody automatické extrakce, jenž slouží pouze jako dílčí podklad pro další postup.

Postup při exrakci termínů (Schwarz 2003):

- Lexikální analýza - identifikace jednotlivých slov a sousloví v textu dokumentu. Někdy se identifikace sousloví uvádí vzhledem ke své složitosti jako samostatná procedura;
- Odstranění nevýznamových a nespecifických slov – provádí se pomocí negativního slovníku;

[6] Dle TDKIV můžeme roli indexátora definovat jako (Plankova, 2003): *„odborný pracovník zajišťující věcné zpracování dokumentů prostřednictvím obsahové analýzy a indexace dokumentů.“*

- Lematizace – redukce slov na jejich základní tvar (kmen a kořen). Opakem procesu je derivace, kdy se k základnímu tvaru slova generují jeho inflexní tvary (využívá se při vyhledávání);
- Srovnání slov – jde o porovnání termínů z vybraného textu s termíny řízeného slovníku;
- Vážení nebo stanovení vah termínů – nejčastěji na základě frekvenční metody.

3.3.1 Taxonomie

Pojem **taxonomie** slovník TDKIV nezná. Pročítáme-li obornou literaturu, zjistíme, že tento termín má několik významů. Široké pojetí termínu je synonymní s termínem třídění. Je možné se setkat s významem, že jde o třídění dokumentů v rámci podnikové sféry. Úzké pojetí chápe taxonomii jako typ věcného selekčního jazyka. Objasnění, jak se tentto termín používá, předkládá práce Mohammada Hanniefa Bhata a Sheikha Mohammada Shafia (Bhat 2004), ze které jsou použity definice níže:

Boer termín definuje: *„taxonomie je logickým pořádáním informací do kategorií“*. Podle Heddena je: *„druh řízeného slovníku, který má hierarchii (nadřazený termín/podřazený termín), ale ne nutně související vazby, jako je tomu u standardního tezauru.“* Dle Keheta je to: *„praxe a věda, formální a institucionalizovaná klasifikace se nazývá klasifikace. Jako klasifikační systém taxonomie obvykle organizuje poznání světa v hierarchické stromové struktuře, kde jsou zastoupeny vazby nadřazenosti-podřazenosti, včetně-zahrnuto, nadtřída-podtřída nebo dominantní-dominující vztahy mezi koncepty.“*

Společnost AIIM tvrdí, že klíčovou funkcí taxonomie je dávat související věci do skupin tak, aby bylo snazší tyto věci nalézt.

Taxonomie, stejně jako klasifikace, má za cíl splnit čtyři základní cíle: identifikovat obsah, jeho správu, pomoc při vyhledávání obsahu a sledovat jeho využívání. Aby bylo možné tyto cíle realizovat, je nutné splnit devět pravidel (9 Ingredients to a Successful Taxonomy, s. 03-04):

- Intuitivnost – uživatelé se snadno orientují a používají navržené třídění. Bez problémů předpovídají, ve které kategorii pravděpodobně naleznout hledaný dokument. Třídění je koncipováno způsobem, který odráží pracovní zvyky a používané struktury uvnitř podniku.

- Jednoznačnost – třídění nenabízí alternativní místa, kde by mohl být dokument k nalezení. Uživatelé nemají více možností, kam dokument uložit, respektive kde ho vyhledat.
- Pohostinnost – taxonomie pojme nový typ dokumentu bez nutnosti provést její revizi. Nepohostinné třídění podporuje tendenci uživatele dávat obsah do nevhodných kategorií. Naopak pohostinnost umožňuje bez problémů přijmout nový typ dokumentu bez významného rozšiřování či zásadní restrukturalizace.
- Konzistentnost a předvídatelnost – uživatelům poskytuje komfort při navigaci uvnitř struktury. Podporují intuitivnost, protože kategorie jsou organizovány abecedně, hierarchicky… To zajišťuje rychlé a úspěšné procházení tříděním. Na základě toho se s tříděním uživatel rychle seznamuje. Pokud je použita stromová struktura, pak princip podřízenosti je shodný pro celou úroveň. Toto pravidlo bývá někdy ohroženo díky soutěživosti různých skupin uživatelů.
- Relevantnost – třídění odráží uživatelský jazyk, uživatelský pohled na to, jak by měl být obsah organizován. Třídění má odrážet způsoby uspořádání informací v podniku.
- Šetrnost – v taxonomii není nic navíc, než je pro typy ukládaných obsahů požadováno, a ani nic nechybí. Neexistuje zde redundance. Někdy může omezovat pravidlo pohostinnosti, ale podporuje důležitý princip jednoznačnosti.
- Smysluplnost – výsledky vyhledávání jsou přesné a relevantní.
- Trvanlivost – třídění nepotřebuje časté změny nebo rozšíření a už vůbec ne změny radikální. Robustní třídění většinou potřebuje změny jen jedenkrát za rok, nejsou-li vyžádány radiální změnou typu ukládaného obsahu.
- Vyváženost – struktura používaného třídění by měla velké množství informací rozdělit do vyvážených klastrů. Sleduje, zda nějaká část struktury není nadměrně vytěžována (ukládá se zde velké množství obsahů) oproti jiné části.

V odborné literatuře nalezneme rozdělení taxonomie (How to Develop Taxonomie 2012, s. 03-05). Nejjednodušší formou je **řízený seznam**. Je to seznam termínů, které vznikly pro nějaký účel nebo jde o kroky v procesu, role v projektu a podobně. Při prvním pohledu by měl uživatel být schopen porozumět principu podobnosti, která je spojuje. Je však obtížné ho

na první pohled rychle projít a seznámit se s ním. Doporučuje se, aby seznam nepřekračoval 12 - 15 položek.

Ve chvíli, kdy je těžké se v seznamu orientovat a začíná být dlouhý, začne se seznam členit do kategorií, vytvoří se **stromová struktura**. Právě ta bývá s taxonomií nejčastěji spojována. Velké množství stromových struktur není pro uživatele předvídatelné, důvodem bývá špatně uplatněný princip dělitelnosti mezi jednotlivými úrovněmi.

Pokud při dělení důsledně uplatňujeme stejný princip na všech úrovních, pak se ze stromové struktury stává **hierarchická struktura,** která už je velmi snadno předvídatelná. Hierarchie by měla být vyčerpávající (obsahovat všechna možná témata, i ta, která se v daném podniku nepoužívají) a kategorie by se měly vzájemně vylučovat, což podpoří pravidlo jednoznačnosti. Téma v dané struktuře existuje pouze jednou na jednom místě.

Trochu specifický typ je **polyhierarchie**, která se snaží řešit problém s tím, že tématům přiřadí více nadřazených objektů. Tady se využívají křížové vazby, které zajistí provázanost na všechny nadřazené objekty. To ale snižuje předvídatelnost.

Fasetová struktura naopak bere v úvahu pouze jednu vlastnost. Pro popis objektu potřebuje několik faset, které daný objet specifikují, například typ dokumentu, typ služby. Každá faseta může mít formu řízeného seznamu, stromu nebo hierarchie. Použitím kombinace jednotlivých faset se docílí bohatého popisu obsahu a zároveň se zajistí odlišný způsob organizace a hledání. Uspokojí to různé skupiny uživatelů a překoná konkurující způsoby organizování dokumentů u různých skupin uživatelů. Fasetová struktura je oblíbená při navigaci ve velkých kolekcích. Specifické použití faset je v **maticové struktuře**, která má obvykle dvě fasety ve formě tabulky. Umožňuje to prohlédnutí všech možných kombinací. Tato struktura je funkční, existuje-li omezený počet subjektů, se kterými se pracuje. Díky tabulkovému zobrazení se snadno odhalují mezery (některá buňka není vyplněná). K omezením patří to, že v případě velkého počtu objektů se uživatelům obtížně používá.

Posledním typem jsou **systémové mapy**, které dělíme na popisné a koncepční. Jde o vizuální reprezentaci znalostí, kde propojení vyjadřují vztah mezi entitami. Popisnou mapou je například schéma metra, mapy tepen v lidském těle. Ke koncepční naopak patří procesní mapy.

V rámci této práce bude termín taxonomie používán v užším významu v hierarchické podobě.

3.3.2 Ontologie

S pojmem ontologie se setkáme ve filozofii antického Řecka. Všeobecná encyklopedie tento pojem definuje jako *„základní část filozofie, věnována otázce bytí jako takového a nejobecnějším kategoriím jsoucna.“(*Anon. 1999, s. 464) Pojem byl do filozofie zaveden v 18. století německým osvícenským filozofem Christianem Wolfem.

Potřeba ontologie v prostředí webových stránek byla vyvolána nutností vyměňování informací, které by bylo srozumitelné i pro stroj. Toto je realizováno prostřednictvím takzvaného sémantického webu.

Organizace W3C[7], která se dlouhodobě zabývá oblastí sémantického webu a vývojem standardů, definuje ontologii takto*: „vymezuje pojmy a vztahy, které jsou použity k popisu a prezentaci oblasti zájmu.“* (Semantic web c2017).

T. Gruber ontologii definuje jako: *„explicitní specifikaci konceptualizace.“* (Svátek c2009, s. 2). Neboli pojmy, které modelují určitou část statě, je nutné explicitně vyjádřit.

Ontologie podporuje výměnu informací mezi počítačovými systémy, porozumění mezi různými typy lidí (například znalostními inženýry a experty).

Terminologie týkající se struktur ontologie se různí, nicméně v základu je možné ji rozdělit na:

- Třídy – množina konkrétních objektů;
- Vztahy – definované vztahy, které jsou mezi konkrétními objekty;
- Instance – konkrétní objekt reálného světa;
- Pravidla – kromě vztahů definujících příslušnost ke třídě je možné zařazovat i jiné logické formule, které například vyjadřují vztah ekvivalence, disjunktnost tříd apod.

Ontologie je možné členit podle míry formalizace, podle předmětu formalizace a podle paradigmatu. Podle paradigmatu je dělíme na (Svátek, c2009, s. 3):

- Terminologické – obsahují termíny, které se dále nedefinují a vztahy vymezují obecnějšího a speciálnějšího termínu, dále zde nalezneme i synonymní vztah a vztahy obecného charakteru.

[7]Mezinárodní organizace W3C se dlouhodobě zabývá tvorbou webových standardů.

- Informační – jedná se o nadstavbu databázových schémat, zabezpečuje vyšší úroveň kontroly integrity než běžně dostupné nástroje, a abstrakci, která je potřebná pro pojmové dotazování.
- Znalostní – reprezentace znalostí v rámci umělé inteligence.

V praktické části diplomové práce je blíže představeno terminologické paradigma. Z něj vyplývá, že ontologie je podobná tezaurům. Oba vymezují vztahy mezi termíny a věcně pokrývají nějakou oblast lidského poznání. V některých případech je možné je zobrazit v hierarchické struktuře.

Vztahy mezi termíny v tezauru jsou předem definovány (nadřazený, podřazený, asociativní, nepreferovaný). Ontologie umožňuje definovat vlastní vztahy a tím zajistit jejich přesnější vyjádření, vytvořit vazby mezi třídami a jednotlivými instancemi, vztahy mezi částmi a celkem.

3.3.3 Nástroje umožňující správu předmětového věcného třídění

Vzhledem k tomu, že se diplomová práce zabývá tříděním dokumentů, domnívám se, že je vhodné aspoň krátce představit software, který se specializuje na tvorbu a správu předmětových selekčních jazyků, neboť je možné říci, že technologie ECM jsou už pouze konzumenty navržené a používané klasifikace.

Software umožňující práci se tříděním by měl umět podpořit tyto fáze (Abdus 2010):

- Vývoj třídění – měl by umět podpořit nejen vývoj uvnitř nástroje, ale i automatické generování třídění na základě dostupných dat, import a export třídění a jeho následnou úpravu. Je vhodné, když nástroj umožňuje pro import a export více druhů formátů například: XML, CSV, XML2, OWL, SKOS, HTML, RDF/XML, MARC.
- Kategorizace dokumentů – měl by umět podpořit nejen ruční přiřazování metaúdajů k dokumentům, ale i automatickou klasifikaci nebo kombinaci obou (hybridní přístup).
- Vyhledávací a navigační funkce – software by měl být přehledný a intuitivní, umožnit kvalitní zobrazení výsledků hledání na základě zadaného rešeršního dotazu.

- Změny v rámci třídění – měl by umět podpořit vývojové změny v rámci třídění, tak aby bylo možné včlenit nový typ obsahu do již používaného třídění.

Při výběru softwaru, který se bude při vývoji používat, bychom kromě ceny, podpory, architektury a podporovaných platforem měli mít na zřeteli tato kritéria (Abdus, 2010):

- Přesnost – nástroj musí být schopen kategorizovat dokument přesně;
- Podpora různých formátů – software by měl podporovat různé typy používaných formátů pro obsah;
- Podpora vícejazyčnosti – měl by umět podpořit různé typy jazyků, zvláště jde-li o podnik působící v mezinárodním prostředí;
- Údržba – správce softwaru by měl mít možnost ručně upravovat třídění (přidávat, odebírat).

Na trhu existují jak velké robustní nástroje, které umožňují automatickou tvorbu třídění nebo automatickou klasifikaci dokumentů, tak i menší nástroje, které umožňují pouze návrh třídění. Mají zde zastoupení komerční programy i volně šířený otevřený software (open-source).

Pro potřeby této diplomové práce byly použity programy **Protégé**, který je vyvíjen na půdě Standfordské univerzity a **ThManager**, který byl vyvinut na Zaragotské univerzitě.

Mezi další otevřené softwary patří například **TemaTres**, který má svůj původ v knihovně a univerzitě v Buenos Aires, nebo **VocBench** původně užitý pro vývoj tezauru AGROVOC, který má sloužit potřebám Organizace pro výživu a zemědělství (FAO) v rámci Spojených národů (OSN), a nebo **SWOOP**, vytvořen na půdě univerzity v Marylandu.

Ke komerčním nástrojům patří například **Autonomy IDOL,** jehož výrobcem je společnost HP, dále **Intelligent Topic Manager** od společnosti Mondeca, **ConceptTaxonomyManager** od firmy Concept Searching nebo **Thesaurus Master** od Access Innovations.

V rámci této kapitoly byly představeny základní termíny týkající se třídění dokumentů, byly rozebrány prekoordinované a postkoordinované selekční jazyky i s příklady. Kapitola mimo jiné obsahuje význam nových termínů taxonomie, ontologie. Kapitola je zakončena

informacemi o nástrojích, které umožňují tvorbu a následnou správu předmětových selekčních jazyků.

4 Standardy, normy a legislativní dokumenty týkající se klasifikace dokumentů

Vzhledem k tomu, že organizace je otevřený systém, je ovlivňována svým okolím a musí reagovat na změny, požadavky, které jsou na ni kladeny. Některé z dokumentů, které generuje či přijímá od externích subjektů, mohou podléhat různým zákonným či normativním standardům, anebo mohou být použity jako důkazní prostředek při soudních a mimosoudních sporech. Aby bylo možné s těmito dokumenty nakládat dle potřeb organizace, je nutné je správně katalogizovat.

Na klasifikaci dokumentů v širším slova smyslu mohou mít dopad některé zákony či normy. Jsou představeny v následujících podkapitolách, vždy je dokument popsán obecně a následně je charakterizováno, z jakého hlediska je pro klasifikaci dokumentů důležitý.

4.1 Zákon č. 499/2004 Sb., o archivnictví a spisové službě a o změně některých zákonů

Zákon upravuje archiválie (Česká republikaZákon o archivnictví a spisové službě a o změně některých zákonů 2004), specifikuje způsob výběru, evidence a ochrany. Dále řeší práva a povinnosti vlastníků, správce archiválií. V neposlední řadě se zabývá spisovou službou.

Za archiválii je považován ten dokument, který má trvalou hodnotu z pohledu obsahu, původu, doby vzniku nebo vnějších znaků.

Tento zákon je pro podnikový obsah důležitý, protože definuje procesy související se skartačním řízením. V rámci skartačního řízení dochází k oddělování archiválií, které je nutné následně postoupit příslušnému archivu, od bezcenných dokumentů, které nejsou již pro podnikání důležité a je možné je po skončení řízení bez problému zničit (např. odmazat ze systému a záloh). Dokumenty jsou během řízení označovány skartačním znakem A (archiválie), S (skart – možné zničit) nebo V (dokumenty, u kterých ještě není možné určit, zda jde o archiválie či stoupu). Tento proces generuje požadavky na specifický typ metadat, jako jsou například skartační doba nebo skartační znak.

Zákon ve své příloze č. 1 jasně vymezuje, které dokumenty vzniklé z činností obchodních společností jsou archiválie. Jsou to například zakladatelské dokumenty, stanovy, jednací řád, organizační řády a schémata, výroční zprávy, zprávy o auditu, účetní uzávěrky, podnikatelské záměry, vývojové studie a jiné.

4.2 Zákon č. 563/1991 Sb., o účetnictví

Zákon na základě předpisů Evropské unie vymezuje účetnictví (způsob a rozsah), dále požadavky týkající se průkaznosti účetnictví, způsob a rozsah zveřejňování informací týkajících se účetnictví a v neposlední řadě podmínky, za kterých se účetní záznamy předávají státu (Česká republikaZákon o účetnictví 1991).

Z pohledu práce s dokumenty je důležitý § 31 a §32, který definuje délku času, po kterou jsou účetní záznamy uchovávány. Pro výroční zprávu a účetní uzávěrku je to 10 let, doba 5 let platí pro účetní doklady, odpisové plány, účtový rozvrh, inventurní soupisy, účetní knihy a účetní záznamy, kterými dokládají vedení účetnictví. V případě, že běží reklamační řízení nebo záruční doba a je delší než doby stanovené výše, je nutné po dobu běhu procesů lhůtu pro uchování prodloužit. Totéž platí i pro případ nesplněného závazku nebo nezaplacené pohledávky, doklady jsou uchovávány až do konce prvního účetního období následujícího po období, kdy došlo k vyrovnání závazku či pohledávky.

4.3 Zákon č. 235/2004 Sb., o dani z přidané hodnoty

Zákon upravuje daň z přidané hodnoty (Česká republikaZákon o dani z přidané hodnoty 2004).

Z pohledu dokumentů je důležitá skartační doba u plátců DPH, kteří mají povinnost archivovat, dle § 27, daňové doklady nejméně po dobu 10 let od konce zdaňovacího období. Na plátci též leží břemeno zodpovědnosti za věrohodnost, neporušitelnost a čitelnost původního dokladu. Podnik může doklad v listinné podobě digitalizovat, uvedené podmínky však musí být zachovány.

4.4 Zákon č. 412/2005 Sb., o ochraně utajovaných informací a o bezpečnostní způsobilosti

Zákon (Česká republikaZákon o ochraně utajovaných informací a o bezpečnostní způsobilosti 2005) upravuje stupně utajovaných informací, jakým způsobem se s těmito informacemi má pracovat (ochrana, přístup). Dále jmenuje citlivé činnosti, podmínky jejich výkonu.

Z pohledu práce s dokumenty je důležité, že dělí informace na přísně tajné, tajné, důvěrné a vyhrazené. Informace o stupni utajení bývá většinou součástí klasifikace dokumentů.

4.5 Zákon č. 101/2000, o ochraně osobních údajů

Zákon vymezuje právo na ochranu před neoprávněným zasahováním do soukromí, upravuje práva a povinnosti při zpracování osobních údajů (Česká republikaZákon o ochraně osobních údajů 2000). Dále definuje podmínky, za kterých smí tyto údaje být předány do jiných států.

Některé dokumenty obsahují osobní údaje, které je správce povinen zlikvidovat následně poté, co uplynula potřeba uchování pro provoz společnosti. Způsoby, jak se s touto zákonnou podmínkou vyrovnat, jsou dva: začernit osobní údaje na dokumentu, nebo ho celý vymazat.

4.6 Vymáhací proces pohledávek

Na práci s dokumenty, zvláště na dobu skartace, má významný dopad vymáhací proces, který má za úkol zajistit vymožení nezaplacených pohledávek.

Pojem pohledávka vymezuje Občanský zákoník č. 89/2012 v paragrafu 1721, který říká, že věřitel má vůči dlužníku právo na plnění jako na pohledávku a dlužník má povinnost toto právo uspokojit (Česká republikaZákon občanský zákoník 2012). Plnění musí být majetkové povahy.

Způsobů, jak vzniklé pohledávky zajistit a utvrdit, je několik: zástavní právo, zadržovací právo, ručení, smluvní pokuta, finanční záruka a směnka.

Pohledávky mohou být vymáhány soudní nebo mimosoudní cestou. V obou případech je nutné zajistit evidenci pohledávek a veškerých návazných dokumentů, které mohou být následně u soudu označeny jako důkaz dle zákona č. 99/1963 Sb., občanský soudní řád, §120. To znamená mít k dispozici doklad, na jehož základě pohledávka vznikla (smlouva, závazná objednávka,…), účetní doklady, korespondenci týkající se dané věci, upomínky o nezaplacení dluhu apod. (Doleček 2014)

4.7 eIDAS

Při práci s elektronickým dokumentem, zvláště u smluv a faktur, se v rámci technologie ECM setkáváme s elektronickým podpisem. Ještě začátkem roku 2016 u nás platil zákon č. 227/2000 Sb. o elektronickém podpisu, který vycházel ze směrnice Evropské unie 1999/93/ES. Ta byla nahrazena nařízením č. 910/2014, o elektronické identifikaci a službách vytvářejících důvěru na vnitřním trhu, které je obecně známé pod termínem eIDAS,

což je odvozeno od eID and Signature. Nařízení je u nás aplikováno v zákoně č. 297/2016 o službách vytvářejících důvěru pro elektronické transakce (Peterka 2016b).

Tento zákon zcela mění používání elektronických kvalifikovaných podpisů (dříve osobní certifikát, elektronická značka, systémový certifikát), definuje dva druhy podpisů: elektronických podpis a elektronická pečeť (Peterka 2016a).

Evropská unie místo našich elektronických značek zavádí elektronické pečetě, které může na rozdíl od el. značky připojit pouze právnická osoba, a to jen na své původní dokumenty. Zároveň pečeť neřeší potřebu něco strojově podepisovat, jako tomu bylo u elektronických značek.

Zároveň na podepisující (jedno zda elektronickým podpisem či pečetí) přenáší odpovědnost seznámení se s obsahem toho, co podepisuje(Peterka 2016a).

4.8 Normy

4.8.1 ISO norma 15489

Norma 15489 je mezinárodně uznávaná ISO norma týkající se správy informací a záznamů, rozdělena do dvou částí:

- ISO 15489-1:2001 Information and documentation – Records management Part 1 : General;
- ISO 15489-2:2001 Information and documentation – Records management Part 2 : Guidelines.

V první části nalezneme definici základních pojmů, pokyny pro navrhování a implementaci spisové služby, procesní řízení záznamů a určování odpovědností. Druhá část je metodická při zavádění normy do organizace, která chce splnit její podmínky.

V metodické části nalezneme minimální seznam metadat, kterými jsou: přiřazení jedinečného identifikátoru systémem, datum a čas registrace, název nebo zkrácený popis, autor (osoba či organizace), odesílatel nebo příjemce.

Mimo jiné norma upozorňuje na to, že je možné k jednotlivým záznamům přidávat i jiné metaúdaje, které se mohou týkat obsahu, struktury záznamu nebo popisné informace týkající se kontextu. Navrhuje podle typu záznamu připojit tyto metaúdaje: název dokumentu, popis nebo abstrakt, datum vytvoření, datum a čas převzetí, autor, odesílatel, příjemce, klasifikace, odkaz na související záznamy, software a jeho verze, ve které byl záznam

vytvořen nebo ze které přišel, přístupová práva, doba uchovávání. (15489- Information and documentation – Records management 2001)

4.8.2 ISO norma 23081

Norma 23081 doplňuje ISO normu 15489 a má také dvě části:

- ISO 23081-1:2006 Information and documentation – Records management processes – Metadata for records – Part 1: Principes;
- ISO 23081-2:2006 Information and documentation – Records management processes – Metadata for records – Part 2 : Conceptual and implementation issues.

Norma poskytuje rámec pro použití, vytvoření a řízení správy záznamů a jejich metadat. Norma neobsahuje povinný soubor metadat k záznamům, neboť vzhledem k různým typům organizací a rozdílným legislativním požadavkům to není možné.

Norma dělí metadata do šesti typů, a to na (23081-1 Information and documentation – Records management processes 2006):

- metadata o záznamu samotném,
- metadata o obchodních pravidlech nebo postupech a nařízeních,
- metadata o agentech,
- metadata o podnikatelské činnosti nebo procesech,
- metadata týkající se procesu řízení záznamu.

Na tuto normu ve své specifikaci navazuje MoReq2, který je představen v podkapitole standardy.

4.8.3 ANSI/AIIM TR40-1995

Norma s názvem Navrhované metaúdaje pro dokumenty v prostředí pro správu digitalizace dokumentů (Suggested Index Fields for Documents in Electronic Image Management Environments) popisuje metadata, která jsou často využívána v systému pro správu digitalizace dokumentů.

Norma je z pohledu architektury, aplikace, implementace nezávislá. Popisuje metaúdaje z těchto oblastí: vyhledávání, workflow, údržby, směrování, zpracování, dotazování a zpracování. Tato metadata mohou být zároveň využita i pro auditní logování v rámci systému.

4.8.4 ANSI/AIIM/ARMA TR48-2004

Norma obsahuje rámec pro integraci elektronických systémů pro správu dokumentů (EDMS) a systém pro řízení elektronických záznamů (ERMS). Obě technologie jsou součástí ECM.

Obsahem normy jsou integrační požadavky na spolupráci ERMS a EDMS. Popisuje rámec integrace ve třech klíčových oblastech: správa metadat, funkčnost, typické přístupy implementace.

4.8.5 ANSI/AIIM MS60-1996

Norma se nazývá Výměna datového toku v elektronické složce (Electronic folder interchange datastream).

Definuje a popisuje nezbytné a dostatečné konstrukce pro datový tok, které poskytují možnost komplexní výměny obrazových objektů mezi systémy pro digitalizace dokumentů. Norma se zabývá výměnou objektů, jako jsou složky, dokumenty s jejich souvisejícími objekty, daty a metaúdaji.

4.9 Standardy a jiné projekty

4.9.1 MoReq2

Standard MoReq (Model Requirements for the Management of Electronic Records, Modelové požadavky na správu elektronických dokumentů) popisuje modelové požadavky pro správu elektronických záznamů se zaměřením na funkční požadavky systémů elektronické spisové služby (ERMS). Vznikl na základě projektu IDA, který byl v roce 1999 vyhlášen Evropskou komisí a spravuje ho DLM fórum. Dnes se již setkáváme s jeho druhou verzí. **MoReq2**, je rozšířen o metodiku uchování digitálních dokumentů a jejich metadat.

Metadatový model se snaží být konzistentní s ISO normami: ISO 23081 (má širší působnost než MoReq2), ISO 1549 a ISO 15836 (některé prvky pro správu záznamů neobsahuje).

4.9.2 Dublin Core

V roce 1995 se v americkém městě Ohio v Dublinu konala konference (odtud název), kde standard vznikl. Původním cílem bylo zajistit popis zdrojů v rámci Word Wide Web, který bude sestavovat přímo autor, ale během let si získal i přízeň institucí, které se zpracováním zdrojů primárně zabývají – vládní organizace, muzea, knihovny. Standard byl v roce 2001 přijat jako americká norma pod ANSI/NISO Z39.85, následně v roce 2003 jako

ISO norma 15836:2003 a od roku 2008 je součástí formátu Office Open XML, který je ISO normou ISO/IEC 29500:2008. O oblíbenosti svědčí i fakt, že se s ním setkáváme v různých projektech zabývajících se klasifikací elektronických dokumentů (SKOS, jeden z bibliografických formátů v rámci katalogu Kongresové knihovny, v programu ThManager či Protégé). V odborné literatuře se pro něj používá zkratka DC. (DCMI Home 1995-2017)

Asociace DCMI (Dublin Core Metadata Initiative), zastřešuje veškerý vývoj, správu, podporu a propagaci tohoto standardu. Její náplní je například činnost v pracovních skupinách, pořádání výročních konferencí a seminářů, vzdělávací činnosti, podpora šíření normy a postupů.

Základní seznam metadat v anglickém a českém znění je uveden v tabulce 1. Z výčtu je stále patrný původní záměr standardu – klasifikace elektronických zdrojů.

Tabulka 1: Dublin Core Metadata (Dublin Core 2006)

Anglicky	Česky
Title	Název
Creator	Tvůrce
Subject	Předmět
Description	Popis
Publisher	Vydavatel
Contributor	Přispěvatel
Date	Datum
Type	Typ
Format	Formát
Identifier	Identifikátor
Source	Zdroj
Language	Jazyk
Relation	Vztah
Coverage	Pokrytí
Rights	Práva

Další prvky byly přidány následně (viz příloha číslo 1).

Při uvádění jednotlivých metadat není důležité jejich pořadí, neboť každá položka má své jedinečné ID. Příklad zápisu XML pro demonstraci uvedeného (Riley):

```
<dc:subject>children</dc:subject>
<dc:coverage>United States</dc:coverage>
<dc:coverage> New South, Populism, Progressivism, and the Great Depression, 1877-1939</dc:coverage>
<dc:type>Photograph</dc:type>
<dc:format>1 photograph : negative, b&w ; 4 x 5 in.</dc:format>
<dc:format>Image</dc:format>
<dc:identifier>local-cont-no: 81-00423-7</dc:identifier>
```

Mezi základní rysy DC patří jednoduchost, mezinárodní použití možnost individuální modifikace.

4.9.3 CMIS

V roce 2008 daly společnosti EMC, IBM a Microsoft první návrh specifikace standardu, který měl definovat rozhraní pro komunikaci v rámci aplikací pro správu podnikového obsahu. Tento návrh byl podstoupen společnostem Alfresco, SAP, Oracle a Open Text a stal se základem pro standard **CMIS** (Content Management Interoperability Services) společnosti OASIS[8], která ho v roce 2009 prohlásila za standard.

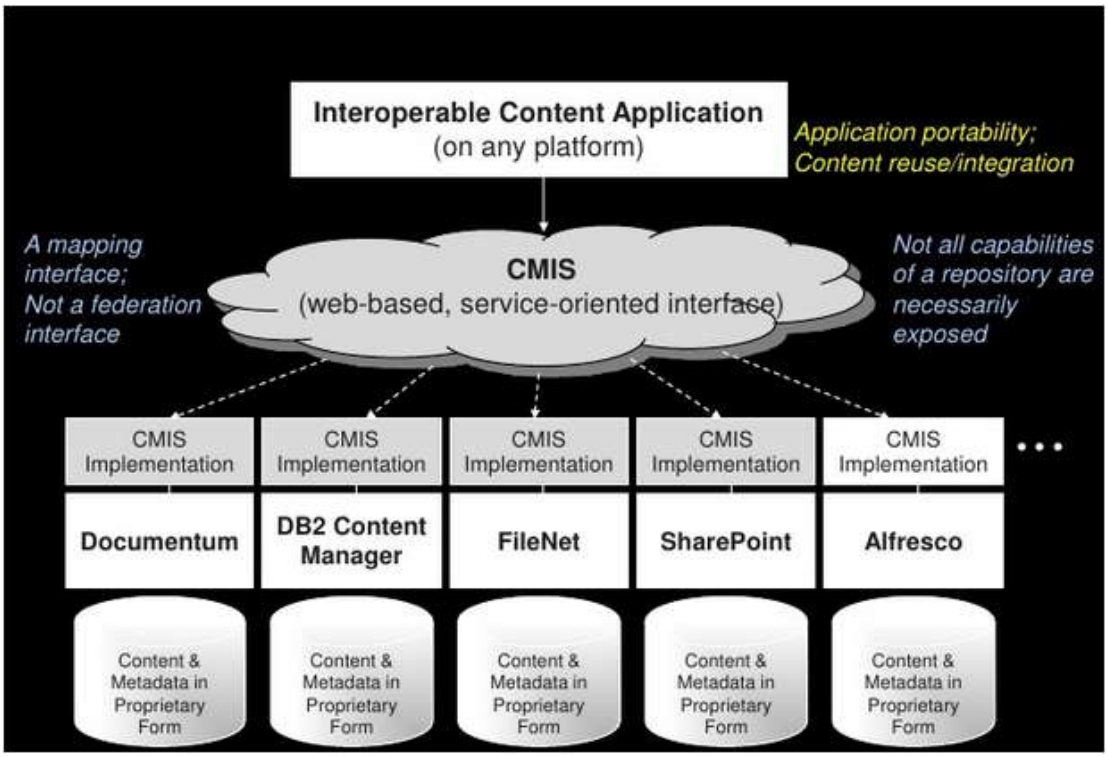
Počátky ECM řešení nalezneme v stand-alone systémech, které řešily správu dokumentů s omezenou funkcionalitou, bez návaznosti na jiné systémy. Postupem času docházelo k propojení na jiné systémy (účetní, výrobní systémy apod.), a integrace v rámci podniků se stávaly komplikovanější. Tento standard umožní snazší a rychlejší včlenění do IT infrastruktury podniku, což s sebou mimo jiné nese i vyšší procento úspěchu implementace a nižší náklady. Mimo jiné nahrazuje standard **ODMA** (Open Document Management API), který se snažil o standardizaci komunikace mezi desktopovými aplikacemi a DMS systémy. Také má podnikům zajistit to, že nebude již na svém dodavateli softwaru závislý a může přejít na jiný produkt.

CMIS je technickou specifikací pro komunikaci s ECM úložištěm pomocí webových služeb (komunikační protokol je Simple Access Protokol (SOAP) nebo Represtational State Transfer (Rest/ATOM)). Standard CMIS je navržen tak, aby pracoval s existujícími uložišti, kde je obsah uchován a zpřístupňován uživatelům. (Švík 2009)

Mezi jeho cíle patří oddělení služeb a obsahu, což umožňuje spravovat samotný obsah a vyvíjet základní aplikace zcela odděleně, podporovat vývoj kompozitních aplikací. Mezi jeho výhody patří nezávislost na použité platformě a programovacím jazyku (Švík 2009).

Jak je patrné z obrázku 3 (Newton 2009), standard umožňuje integraci různých uložišť, takže je možné jedním rozhraním přistupovat k úložišti od firmy IBM, ECM apod. se základními funkcionalitami, jako jsou přidání, checked in/checked out (vyzvednuní/vrácení), změny metadat a přístupových práv. Zároveň je snazší napojení na podnikové procesy (BPM) bez vývoje speciálních konektorů.

[8] Společnost OASIS, název opět pochází z prvních písmen anglických slov The Organization for the Advancement of Structured Information Standards, je neziskové sdružení, jehož cílem je vývoj standardů týkajících se bezpečnosti, internetu věcí (Internet of Things), energetiky, správy obsahu, krizového řízení a dalších věcí.



Obrázek 3: Architektura CMIS (Newton 2009)

V odborné literatuře se můžeme setkat s názorem, že jeho výkon je horší než používání nativních rozhraní pro programování. Je jistě oprávněný, problémy však způsobuje špatná aplikace v rámci produktu, ne špatná architektura ve standardu.

4.9.4 AGRkMS

Pod názvem se skrývá australský standard týkající se správy záznamů (The Australian Standard on Records Management). Obsahuje informace, které se týkají záznamů, a kontextové informace, které popisují zachycení a používání záznamů.

Standard stanovuje druhy metadat, které by měly organizace u záznamu evidovat ve strukturové podobě. Tyto metaúdaje by měly identifikovat objekty, zajistit autenticitu obsahu, strukturu a základní požadavky pro správu záznamů. To zajistí spolehlivé, smysluplné a přístupné záznamy, které splňují obchodní potřeby a důkazní požadavky.

4.9.5 Schema.org

Schema.org patří mezi otevřený slovník metadat. Byl spuštěn v roce 2011 a je součástí tzn. otevřených dat (Linked Open Data), která jsou součástí sémantického webu. Tento slovník je ve formátu RDF[9], což tvůrcům webových stránek umožňuje sémanticky

[9] Je zkratka z anglických slov Resource Description Framework, neboli systém popisu zdrojů. Využívá se pro modelování informací v různých syntaxích.

označit text a zvýšit schopnost strojů porozumět obsahu, což ve svém důsledku může vést lepšímu vyhledávání v rámci webových vyhledávačů.

Je používán například v rámci katalogu WordCat, díky čemuž se jeho katalogizační záznamy dostaly více ke koncovým uživatelům – katalog byl zaindexován vyhledávačem Google, který záznamy převedl mimo jiné i do své části Google Books (Schema.org 2015).

4.9.6 DBpedia

DBpedia je komunitní projekt, jehož snahou je získat strukturované informace z Wikipedie a zpřístupnit je na webu. Projekt umožňuje klást sofistikovanější dotazy než na klasické Wikipedii a odkazovat se na jiné informace, než umožňuje Wikipedie. Díky tomu komunita doufá, že bude snazší využívat velké množství informací, které se na Wikipedii nachází, a budou použity novým, zajímavým způsobem. Mohlo by to mít dopad na zlepšení navigace v encyklopedii samotné. (Riley)

Znalostní báze a ontologie začínají hrát důležitou roli pro podnikové vyhledávání a integraci informací. Je složité znalostní báze, které obsahují pouze určitou doménu, udržovat aktuální. Naproti tomu stojí projekt Wikipedia, který dnes patří k centrálním zdrojům lidského poznání, udržovaný tisíci přispěvateli. Využití DBpedia ontologie v rámci ECM domény může podnikům přinést zlepšení ve vyhledávání podnikového obsahu.

Strojově čitelná Wikipedie má již českou verzi.

4.9.7 OWL

Ontologie je možné modelovat více jazyky, například **OWL** (Web Ontology Language), vyvíjen jako standard organizací W3C. OWL se dělí na třídy (množina obsahující instance), vlastnosti (binární relace mezi instancemi) a instance (konkrétní objekt v doméně). Mimo jiné má prostředky pro vyjádření kardinality.

OWL má podkategorie (Klempa 2006):

- OWL Full – obsahuje všechny konstrukce jazyka.
- OWL DL – také obsahuje všechny konstrukce jazyka, ale s různými omezeními, například instance třídy nemůže být zároveň třídou. Umožňuje maximální schopnost vyjádření bez ztráty výpočetní úplnosti.
- OWL Lite – je podobné OWL DL, ale má ještě více omezení. Hodí se například pro použití v aplikacích, kde postačuje klasifikační hierarchie a jednoduché omezení.

Na obrázku 4 uvádím příklad vyjádřené ontologie Pizza v jazyce OWL.

```
<owl:Class rdf:about="#Pizza">
   <rdfs:label xml:lang="en">Pizza</rdfs:label>
   <rdfs:subClassOf>
      <owl:Restriction>
         <owl:onProperty rdf:resource="#hasBase"/>
         <owl:someValuesFrom rdf:resource="#PizzaBase"/>
      </owl:Restriction>
   </rdfs:subClassOf>
   <rdfs:subClassOf rdf:resource="#Food"/>
   <owl:disjointWith rdf:resource="#PizzaTopping"/>
</owl:Class>
<owl:ObjectProperty rdf:about="#hasTopping">
   <rdf:type
   rdf:resource="http://www.w3.org/2002/07/owl#InverseFunctionalPropert
   "/>
   <rdfs:comment xml:lang="en">
      Note that hasTopping is inverse functional because isToppingOf is
      functional
   </rdfs:comment>
   <rdfs:domain rdf:resource="#Pizza"/>
   <rdfs:subPropertyOf rdf:resource="#hasIngredient"/>
   <rdfs:range rdf:resource="#PizzaTopping"/>
   <owl:inverseOf rdf:resource="#isToppingOf"/>
</owl:ObjectProperty>
<owl:Class rdf:about="#PizzaBase">
   <rdfs:label xml:lang="pt">BaseDaPizza</rdfs:label>
   <rdfs:subClassOf rdf:resource="#Food"/>
   <owl:disjointWith rdf:resource="#PizzaTopping"/>
   <owl:disjointWith rdf:resource="#Pizza"/>
</owl:Class>
```

Obrázek 4: Příklad OWL ve formátu RDF/XML (Riley)

4.9.8 SKOS

SKOS je zkratka z anglických slov Simple Knowledge Organization System, což je standard pro sdílení, publikování a reprezentaci taxonomie, heslářů, tezaurů, řízených slovníků nebo klasifikačních schémat (Riley). Jako jiné formáty i tento byl vyvíjen organizací W3C.

Jádrem standardu je pojem (concept), který může být uspořádán do skupin či schémat. Ze standardu RDF používá tzv. RDF trojici (RDF triples), subjekt – vlastnost – objekt. Umožňuje vytvářet hierarchickou (nadřazený/podřazený termín) nebo asociační strukturu (preferovaný/nepreferovaný termín). Často bývá používán ve spojení s jazykem OWL.

Jeho základní prvky jsou (Semantic web c2017):

- prefLabel: preferovaný termín;

- altLabel: alternativní název, používán například pro překlad termínu do cizího jazyka;
- hiddenLabel: skrytý název, může být používán pro vyhledávání či automatickou indexaci, ale není koncovému uživateli zobrazován, může obsahovat například gramaticky špatně napsaný termín;
- broader: vyjádření vztahu nadřazenosti;
- narrower: vyjádření vztahu podřízenosti;
- closeMatch: vyjádření vztahu mezi termíny, které jsou si natolik podobné, že mohou být při vyhledávání zaměněny;
- scopeNote: poznámka pomáhající vyjádřit obsah pojmu;
- topConceptOf: vyjadřuje, že daný pojem je na nejvyšší úrovni.

SKOS je například používán v Národní technické knihovně, v Polytématickém strukturovaném hesláři nebo Mnohojazyčném tezauru Evropské unie (Eurovoc), nebo v tezauru AGROVOC.

Tato kapitola obsahuje zákony, normy a standardy, které se technologií ECM a prací s dokumenty z pohledu třídění zabývají. V kapitole není proveden výčet všech existujících objektů. Hlavním kritériem výběru bylo, zda jsou tyto objekty zmiňovány v praktické části práce. K dalším zajímavým projektům patří například PREMIS, MODS, TEI, BIBFRAME, DoD 5015.2 nebo národní standard pro elektronické systémy spisové služby.

Na dokumenty mají vliv kromě předpisů popsaných výše i zákony dotýkající se specifické oblasti podnikání. Jako příklad uveďme zákon č.127/2005 Sb. o elektronických komunikacích, který například nařizuje, že dokument typu podrobný výpis hovoru je nutné skartovat po době, kdy mu zaniklo právo na reklamaci služby, v tomto případě je to půl roku.

PRAKTICKÁ ČÁST

5 Analýza současného stavu ECM z pohledu klasifikace

Stejně jako jiná odvětví i ECM doména prošla bouřlivým vývojem i v roce 2016, kdy došlo k prodeji produktu Documentum firmě Opentext[10]. Kromě toho společnost Gartner[11] zveřejnila v prosinci, že termín ECM je zastaralý a nahradila ho termínem content services.[12]

Vývoj, kterým doména ECM prošla, je zachycen v tabulce 2. Autorem je John Mancini (Mancini, 2017) vizionář a CEO[13] organizace AIIM.

Tabulka 2: Vývoj domény ECM (předloženo na základě (Mancini 2017))

	Systémy záznamů	Systémy zapojení	Systémy porozumění
	1987 - 2007	2007 - 2017	2017 -
Technologie	ERM, ECM, BPM	SaaS, SharePoint, Cloud	noSQL, business intelligence, Hadoop
Problém	Papírová exploze	Exploze obsahů	Exploze dat
Správa informací	Dokumenty a obrázky	Dokumenty všech typů	Data s obsahem jako podmnožina
Lidé	Limitovaný specialista na dokumenty	Znalostní pracovníci	Informační profesionálové
Proces	Zaměření na výkon a na cíl	Tým a ad-hoc spolupráce	Zákaznická zkušenost
IT role	Vývoj a nasazení	Konfigurace a iterace	Zapojení a interpretace
Architektura	Management záznamů	Informační architektura	Chaotické a dynamické architektury
Kupující	Oddělení ve velkých společnostech	Rozšíření na střední společnosti	Mnoho typů řešení

Zaměříme-li se na klasifikaci, pak jedním z klíčových kritérií úspěchu implementace technologie ECM je správný výběr věcného třídění dokumentů. To potvrzuje i studie

[10] Prodejce, který dlouhodobě nabízí řešení pro správu podnikového obsahu. Společnost Gartner tuto firmu dlouhodobě řadí mezi takzvané vůdce v ECM doméně.
[11] Gartner je poradenská a výzkumná společnost dlouhodobě se zaměřující na oblast informačních technologií.
[12] Během psaní této diplomové práce probíhá živá diskuze mezi odborníky, zda je nutné zavedenou terminologii měnit, co je obsahem nového termínu a nepřišla-li firma Gartner s něčím, co tu již dávno je a to jako jedna ze součástí ECM domény.
[13] Zkratka pochází z anglických slov Chief Executive Officer a je u nás rovna pracovní pozici generálního ředitele.

společnosti Gartner z roku 2006, která zjistila, že 70% nových projektů zavádějících technologie ECM, bylo neúspěšných, neboť nedostatečně investovaly do části týkající se třídění dokumentů. I přesto, že je studie relativně stará, jak je patrné ze studií níže, moc se nezměnilo.

V organizacích se můžeme setkat se třemi obecnými styly návrhu třídění (Stakhov, 2012):

- obsahově zaměřené – navržené třídy klasifikace vycházejí z obsahu či formy samotného dokumentu (korespondence, smlouvy, faktury, …);
- organizačně zaměřené – vycházejí z organizační struktury oddělení ve firmě (IT, personální oddělení, obchodní oddělení, …);
- funkčně zaměřené – organizace dokumentů z funkčního nebo procesního pohledu (z pohledu skartace dokumentů do podoby návrhu plánu skartace se stejným jménem).

Do ECM technologií začínají pronikat nástroje Webu 2.0, jelikož jsou uživateli vyžadovány. Jsou na to přivyklí z běžného brouzdání webovými stránkami. Z pohledu třídění se jedná o našeptávání při vyhledávání, používání tagů a jejich zobrazení v mraku tagů, třídění dle relevance.

Kromě klasických číselníků, heslářů a tezaurů se setkáme s fasetovým tříděním, které se promítá do způsobu vyhledávání dokumentů.

Ze studie organizace AIIM (Miles, 2014), která byla provedena v polovině roku 2014 a zúčastnilo se jí 203 respondentů, vyplývá několik zajímavých výsledků:

- 73 % respondentů souhlasí, že pravděpodobně polovina jejich uloženého obsahu nemá pro obchod přidanou hodnotu, ale nevědí, která to je;
- 62 % respondentů cítí, že obsah uložených záznamů se ročně navýší o 30 %. Průměrný odhad je okolo 35 %;
- 27 % implementací ECM mělo těžký začátek z hlediska klasifikace, často z důvodu neochoty uživatelů třídit obsah pod požadovanými metaúdaji či z důvodu jejich absence. 41% respondentů chce svou stávající klasifikaci optimalizovat.

Ze závěru studie je patrné, že mezi hlavní problémy dotazovatelů patří klasifikace obsahů a jejich odmazávání. Sdílené diskové úložiště a poštovní klienti jsou ve velkém nepořádku (firmy netuší, jaký obsah je zde uložen). ECM technologie na základě slabých pravidel používání, provedených migracích, nedostatečné podpoře a vyžadování firemní politiky v raných verzích softwaru, na tom není nejlépe. Manuální vyčištění nepořádku není reálné minimálně proto, že sami uživatelé při vkládání nového obsahu nejsou bezchybní. Řešením je využít automatické nástroje a roboty, které umí analyzovat obsah a na základě něho opravit či doplnit chybějící megadata. Některé z těchto nástrojů umožňují i přesunutí (např. na levnější storage) nebo výmaz obsahů.

Podobná studie (Miles, 2015), také od společnosti AIIM a opět se dotýká informačního managementu z roku 2015, které se účastnilo 398 respondentů, vyplývá, že při vytváření klasifikace, která je součástí informační politiky, přes 40 % firem využívá služeb externích specialistů. Odhadují, že mezi 50 – 60 % uložených dokumentů je typu ROI, čož potvrdil výstup ze studie o rok dříve. Automatická klasifikace vykazuje nekonzistentnost, nicméně už 21 % dotázaných má s automatickou klasifikací úspěch. Přes 50 % přiznává, že nastavení pravidel pro automatickou klasifikaci ve specializovaných softwarech je obtížné. Když byly respondenti dotazováni, která z použitých metod klasifikace je více přesná, zda-li manuální nebo automatická, pak se ve 34 % uvádí, že automatická klasifikace je přesnější.

Studie trhu od AIIM (Miles, 2016) z roku 2016 se zúčastnilo 266 respondentů. Studie mimo jiné zjišťovala, zda mají organizace známé problémy se špatnou správou obsahu, zjistila:

- přes 50 % obsahu je využito pouze jednou, i když by mohlo být využité vícekrát;
- přes 20 % má problém být v souladu s legislativními či jiným standardy;
- přes 50 % je nutné vynaložit duplicitní úsilí s tvorbou obsahu;
- přes 60 % musí vynaložit mnoho času, věnovaného na hledání relevantního obsahu.

Z této kapitoly vyplývá, že i přes to, že informační technologie poskytuje nástroje pro správu podnikového obsahu, stále zde existují problémy s řešením základních problémů týkající se práce s dokumenty (duplicitní obsah, problém s vyhledáváním obsahu, zajištění

zákonných požadavků, likvidace obsahu, …). Pomoci s nimi může dobře vytvořená a implementovaná klasifikace. V rámci technologie ECM se začínají používat nástroje, které umožňují automatické třídění podnikového obsahu.

6 Metodika klasifikace zákaznických dokumentů ve firmě

Jak je z výše uvedené analýzy stavu patrné, dobrá klasifikace dokumentů je pro firmu zásadní. Nacházíme se v době, kdy nároky na datová úložiště rostou. Aby z nich firma mohla mít užitek, musí uložené dokumenty dokázat řídit, vyhledat/použít ve správný čas, ale také ve správný čas zlikvidovat. Co je pro dokumenty specifické (je jejich výhodou i nevýhodou), je to, že se v organizaci stěží hledá jeho jediný vlastník, neboť dokumenty vytváří či s nimi pracuje více zaměstnanců a většinou si na ně (ze svého úhlu pohledu) dělá i nároky. Tento problém nenajdeme např. u účetního softwaru, kde je jeho hlavním uživatelem finanční oddělení, nebo u personálního softwaru, se kterým pracuje jen oddělení lidských zdrojů. Říká se, že CRM[14] software je srdcem firmy. Domnívám se, že pak ECM je jeho mozkem a z důvodu roztříštěnosti vlastníků bývá často opomíjen.

Oblast třídění dokumentů je často upozaďována či opomíjena i samotnými konzultanty dodávaného softwaru, kteří předpokládají, že má podnik dostatečné znalosti k řešení problému, a z toho důvodu mu při návrhu pomáhají minimálně.

Tato kapitola obsahuje **metodiku klasifikace zákaznických dokumentů** ve firmě. Zásady, které jsou v této metodice popisovány, platí pro firmy, které jsou procesně orientované na zákazníka, nicméně čerpat z této metodiky mohou i jinak orientované firmy, jen je nutné některé části upravit, aby vyhovovaly koncovému uživateli.

Hlavní cíle metodiky klasifikace zákaznických dokumentů jsou:

- popsat postup, jakým způsobem vytvořit a spravovat klasifikaci zákaznických dokumentů ve firmě;[15]
- upozornit na problematiku třídění dokumentů ve firmě a proč není vhodné tuto oblast podceňovat.

Způsoby, jak vytvořit v organizaci klasifikaci, jsou tři: zakoupení komerčního nebo volně dostupného odborného třídění[16], návrh třídění na základě analytických nástrojů a vytvoření vlastního systému. Může se na první pohled zdát, že první dvě varianty vycházejí levněji, ale není tomu tak. Zakoupené třídění či výstupy z analytických nástrojů není možné

[14] Pod termímen CRM (Customer Relationship Management) je chápán systém, který podporuje řízení vztahu se zákazníky.
[15] Zákaznické dokumenty v podniku jsou ty dokumenty, které se obsahově váží k externím zákazníkům firmy, přímo k jednotlivým subjektům, nebo mají obecnější zaměření na obchodní či marketingovou strategii pro jednotlivé zákaznické segmenty.
[16] Webová stránka Taxonomy Warehouse (http://www.taxonomywarehouse.com/) obsahuje různé typy existujících klasifikací.

převzít, jak jsou. Je nutné je projít, zhodnotit stávající stav a potřeby organizace do budoucna a dle toho provést úpravy, což patří mezi časově náročné činnosti.

Vytvoření interního řešení nahrává tomu, že organizace za pomoci kvalitních specialistů je schopna sama řešení vytvořit, neboť má k dispozici veškeré potřebné informace (zná procesy, své zaměstnance, potřeby, strategii). Zaměstnanci – uživatelé – jsou již přítomni procesu tvorby, což pomáhá vytvářet pozitivní postoj ke změně (s uživateli se vede dialog a změna není nařízena direktivně shora, je čas na odstranění rezervovaného přístupu ke změně). Pro malé a střední podniky je tato cesta nejdostupnější.

Metodika je primárně určena pro IT architekty, solution architekty, IT analytiky, ale informace z ní mohou čerpat i projektoví manažeři, specialisté na archivaci či informační specialisti.

Obsahem metodiky je postup, jak vytvořit zákaznickou klasifikaci dokumentů ve firmě, její implementace, ověření a následné vylepšování. Metodika se zabývá obsahy, které jsou v digitální podobě. Uložení a klasifikace papírových dokumentů je mimo její rámec.

Metodika je rozdělena do dvou oddílů, v první části je popsán samotný proces a druhá část specifikuje, co je součástí jednotlivých úseků procesů. Metodika může být využita jako celek, anebo lze použít pouze relevantní části, dle situace, ve které se podnik nachází.

Metodika může být nápomocná v těchto případech:

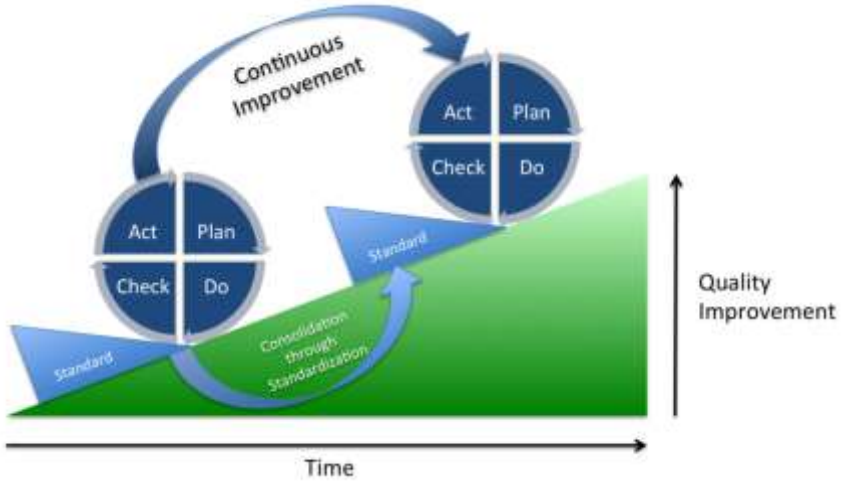
- při implementaci či upgradu ECM technologie nebo její části,
- při migraci dokumentů,
- při včleňování nového softwaru (např. na základě akvizice či fúze firmy) nebo terminaci zastaralého softwaru,
- při vytváření informační politiky firmy,
- při vývoji změn v rámci ECM technologie,
- při skartaci dokumentů,
- optimalizaci úložní kapacity.

6.1 Proces klasifikace zákaznických dokumentů

Navržený proces je rozdělen do čtyř hlavních částí a vychází z PDCA metody, IBM Smarter Process Method a z norem ISO 23081, ISO 15489 (normy jsou blíže popsány v teoretické části).

PDCA metoda, někdy též nazývaná Demingův cyklus, neboť autorem je Wiliam Edwars Deming. Je to metoda, která se může použít pro neustálé vylepšování procesů, služeb nebo výrobků. Skládá se ze 4 částí: naplánuj (plan) – proveď (do) – ověř (check) – jednej (act)., Počáteční písmena jednotlivých částí jsou obsaženy v názvu metody. (Sedláček 2011)

V prvním kroku cyklu se definuje cíl, zároveň i nástroje a způsob řešení. Samotná implementace navrženého plánu je realizována v druhém kroku, proveď. V kroku „ověř“ se provádí kontrola, zda byl cíl dosažen a v čem se implementace liší od původního plánu. Poslední krok, jednej, má za cíl zjistit, zda je zavedená změna úspěšná a nic nebrání v jejím používání v běžném provozu (je prohlášena za standard). Pokud není úspěšná, je nutné projít celým cyklem znovu. Ve chvíli, kdy je proces dokončen a implementace je uvedena do běžného provozu, je možné provádět další změny, viz obrázek PDCA cyklus. Cykličnost je základním principem této metody. Díky ustavičnému opakování se proces, výrobek, služba apod. neustále vylepšuje a díky tomu může firma dosahovat svých cílů. (Deming Cycle 2009)

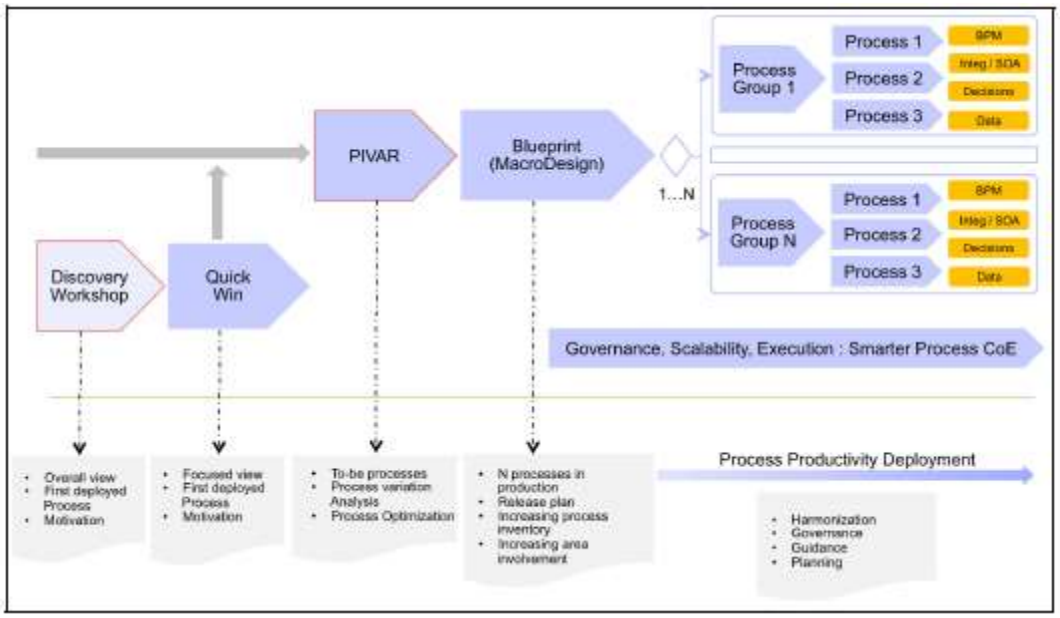


Obrázek 5: PDCA cyklus (Deming Cycle 2009)

Metodika si z toho konceptu hlavně přejímá myšlenku ověřování a vylepšování.

IBM Smarter Process Method je metodika, kterou používá firma IBM pro implementaci BPM[17] projektů. Znázorněna na obrázku 6.

Na začátku projektu se uskuteční workshop, na kterém se do detailu probere jeden reprezentativní proces a stanoví se počáteční odhady. Následně se realizuje pilot, na kterém se ukáže rychlost projektu, možnosti technologie a základní přidaná hodnota pro klienta. Následně se pokračuje analýzou současného stavu a vytvoří se hrubý návrh celého řešení (blueprint). Projekt se následně dělí do pracovních skupin, které mají za úkol vývoj jednotlivých firemních procesů. Metodika počítá s agilním způsobem vývoje a testování.

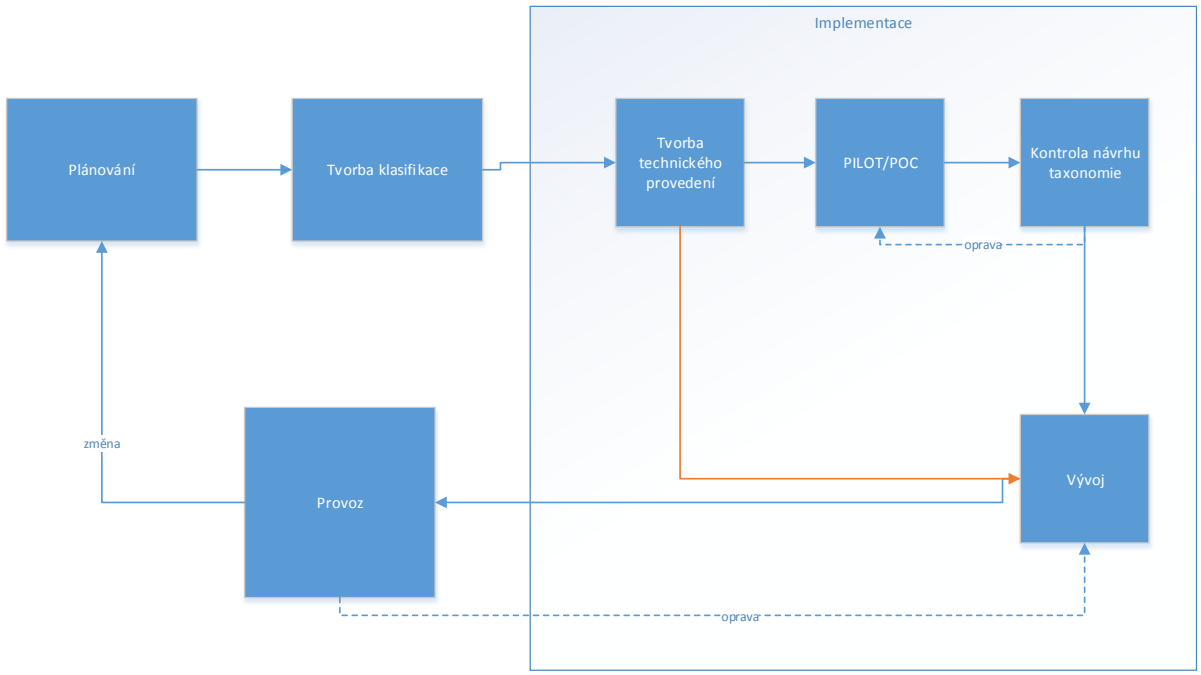


Obrázek 6: IBM Smarter Process Method (Arsanjani 2015)

Z metodiky se čerpá z toho důvodu, že na dokumenty je nazíráno procesním pohledem, zároveň je zde prvek pilotního vzorku, na kterém je možné otestovat navržené řešení. Myšlenka postupného dodávání vývoje a interakce uživatelů během celého procesu je v IBM metodice vnímána pozitivně – na rozdíl od standardního vodopádového projektového řízení, kde je vše dodáno najednou (Arsanjani 2015).

Mnou navržená metodika obsahuje tyto základní etapy: plánování s analýzou stávajícího stavu, tvorbu obecného návrhu klasifikace, implementace a provoz. Graficky je znázorněna na obrázku 7.

[17] Zkratka z anglických slov Business Process Management, nástroj, který umožňuje návrh, vývoj a správu procesů v podniku. (Arsanjani 2015)



Obrázek 7: Proces klasifikace dokumentů (zdroj: autorka)

Ve fázi **plánování** se provádí analýza stávajícího stavu, stanovuje se obsah a cíle dodávky.

Při **tvorbě klasifikace** se hlouběji analyzuje stávající stav, vytváří se obecný model třídění a definuje se vhodná oblast pro fázi Pilot/POC[18].

Vzhledem k faktu, že ve fázi plánování provádíme pouze hrubou analýzu aktuálního stavu, která je následována návrhem detailního provedení, zdá se, že nám fáze studie proveditelnosti chybí. Způsobů, jak tento nedostatek řešit, je několik. Buď bude firmě stačit analýza z fáze plánování a studie proveditelnosti se provede mezi fázemi provedení a tvorbou detailního provedení, nebo bude studie proveditelnosti vedena zvlášť jako samostatný projekt, na jehož základě se rozhodne o realizaci, nebo se zcela vynechá a bude vytvořeno pouze detailní provedení. Osobně se přikláním k nevytváření studie, neboť tato fáze zbytečně navýší časové i finanční náklady procesu. Pokud bude provedena do detailu, bude se téměř shodovat s detailním designem. Bude-li vytvořena povrchně, neodhalí skryté problémy a z větší části nebude v rámci detailního provedení použita.

Ve fázi **implementace** nalezneme tři části. První z nich je **tvorba technického provedení.** V této části se ze získaných dat z prvních částí vytváří návrh technického řešení.

Tvorba klasifikace a technického provedení nejsou realizovány v jedné etapě, aby nedocházelo k vynechání návrhu obecného modelu třídění, ze kterého se následně odvodí

[18] Zkratka POC je složena z anglických slov proof of koncept, tento termín lze přeložit jako ověření konceptu.

technické provedení (model bude upraven dle možností systémů). Vytvoření návrhu obecného třídění dokumentů je důležité ze dvou důvodů: může být použit jako výchozí bod pro architektonické potřeby týkající se změn v doménách a zároveň je výzvou, aby byl model implementován co nejvíce v plném rozsahu. Při vývoji není dobré jít nejsnazší možnou cestou. Vyplatí se hledat různé vhodné alternativy pro řešení zadání apod. V případě včleňování tohoto pracovního postupu do projektové metodiky organizace je možné, aby oba kroky byly součástí jedné fáze, ale je nutné zde zachovat souslednost a v rámci etapy nejdříve provést návrh klasifikace, následně technický návrh.

Fáze **PILOT/POC** má ověřit, že navržený koncept klasifikace je správný a pro organizaci použitelný. Pro ověření můžeme použít buď pilot nebo POC. Rozdíl mezi těmito pojmy je v tom, že v rámci pilotní části se vybere část z celého řešení, která následně projde fázemi vývoje, testováním a nasazením do produkce. Kdežto POC nasadí koncept pouze na test. V obou případech je vhodná součinnost koncových uživatelů, která může mít různé podoby:

- uživatelé se v rámci testování chovají, jako by šlo o reálný provoz;
- uživatelé po nějakou dobu v běžném provozu pracují jak v novém prostředí, tak ve starém prostředí (časově náročné);
- uživatelé se se systémem setkají až po jeho spuštění do provozu (během fáze testů je jejich zapojení minimální).

Mezi výhody pilotního projektu patří to, že vybere-li se vhodná část a kvalitně se naimplementuje, pak jde o tzv. rychlou výhru. Uživatelé mají v relativně krátké době systém k dispozici, mohou ho vyzkoušet. Na základě jejich zkušeností jsou pak provedeny změny do dalších etap, uživatelé získají k systému kladný vztah a cítí větší odpovědnost za úspěšné dokončení celého projektu. K nevýhodám patří to, že jde o software, který je již nasazen na produkční prostředí. Musí se tedy vyřešit způsob, jak např. provést změny v navržené kategorizaci. Projeví-li se něco jako nevyhovující, je nutné vyřešit, v jakém produkčním režimu systém poběží (za jakých podmínek se o něj bude starat tým, který odpovídá za provoz a do jaké míry má odpovídat očekávaným produkčním metrikám apod.), jak zajistit integraci na jiné aplikace (bude např. nutné, aby skenovací software prováděl uložení do obou systémů?) či v jakém modu poběží stávající systém. Mezi výhody POC patří, že je možné zjistit, nakolik je navržený koncept proveditelný, je možné uživatelům demonstrovat funkčnost systému a navrženou klasifikaci. Nevýhodou je, že pokud se při testech či

demonstraci používají produkční data, pak je nutné po ukončení zajistit jejich odstranění. Výběr způsobu se odvíjí od organizace a situace, osobně se více přikláním k pilotu, neboť je velmi těžké získat plnou pozornost a podporu uživatelů na chod systému, který není ještě reálně v provozu.

Část **kontrola návrhu třídění** je zde vyčleněna zvlášť z toho důvodu, že jde o důležitý milník v procesu. Je nutné na základě výstupů z Pilotu/POCu rozhodnout, zda jsou původní navržená třídění pro firmu akceptovatelná. Pokud ne, musí se zjistit, co není funkční, a podniknout kroky k opravě. Až poté je možné přistoupit k poslednímu kroku, a to je **vývoj** zbývajících částí. Způsob, jakým bude tento krok proveden, záleží na velikosti projektu a projektových zvyklostech dané organizace (je možné vývoj dodávat agilně či vodopádovým projektovým řízením)[19].

Po vývoji, testování a instalaci zbylého vývoje se dostáváme k poslední fázi, tou je **provoz.** Na začátku této fáze by se mělo ověřit, že je vše dodáno dle specifikace detailního provedení, opravit chyby vzniklé vývojem. I zde by mělo dojít k ověření, že navržená taxonomie odpovídá a splňuje očekávání. Pokud tomu tak není, měla by se provést její oprava (způsob instalace opravy záleží na podniku a procesech, které má nastaveny pro drobné vývojové změny). V rámci běžného provozu se pravidelně sleduje, je-li třídění stále aktuální, nedochází-li k nějakým systémovým nebo i věcným chybám.

Jsou-li evidovány nějaké požadavky na změnu/drobný vývoj, nebo dochází v systému k chybám, je nutné je opravit. Začíná celý proces od začátku, tedy od fáze plánování, s tím rozdílem, že při drobné změně je ve fázi implementace přeskočena část POC (na obr. vyobrazeno červenou čarou).

Celý proces metodiky je navržen tak, aby ho bylo možné s co nejmenší problémy začlenit do procesu, který má organizace nastaven pro implementaci IT vývoje a jeho změn. Nicméně zároveň reflektuje specifické potřeby, které si klasifikace zákaznických dokumentů vyžaduje.

Metodika u klasifikace zákaznických obsahů počítá s procesním pohledem, neboť toto je jedna z klíčových vlastností těchto dokumentů, i když je někdy opomíjena. Na rozdíl například od dokumentů z oddělení lidských zdrojů – ty využívá zpravidla pouze toto oddělení a je zároveň jejich vlastníkem. Proto mohou být jejich dokumenty tříděny na základě

[19] Základní rozdíl mezi vodopádovým řízením a agilní metodikou je v tom, že agilní metodika je ve stálé interaci, nebo-li vrací se ke stejným krokům (vývoj, test, oprava), kdežto vodopádový model postupuje v sekvencích po jednotlivých krocích (návrh, vývoj, test,…). (Malkusová 2016)

organizační struktury[20]. U zákaznických dokumentů je někdy těžké hledat vlastníka, neboť jsou v určité fázi životního cyklu využívány různými odděleními, jak bude níže rozvedeno.

V metodice je několik kroků, které mají za cíl ověřit, že navržená klasifikace je plně funkční. Je to z toho důvodu, že se musí počítat s delší reakční dobou uživatelů, než je například zvykem u softwaru typu front-end[21]. Veškeré změny ve třídění se neprojeví okamžitě, může to trvat i roky. Reakční čas se odvíjí od délky firemních procesů, které s dokumenty pracují (např. oddělení, které vymáhá pohledávky, si změny ukládání většinou všimne až v době, kdy nemůže potřebné dokumenty vyhledat, což je někdy i několik let od ukončení projektu, kterým se původní proces změnil).

6.2 Fáze plánování

Z pohledu projektového řízení se nacházíme ve stavu specifikace zadání projektu. V praxi se s velkou pravděpodobností nesetkáme s projektem, který by měl jako hlavní cíl dodat zákaznickou či celopodnikovou klasifikaci dokumentů. Většinou je tato potřeba spuštěna jinými faktory – viz příklady uvedené níže. A to je také důvod, proč bývá velmi často podceňována, neboť málokdo si uvědomuje návazné souvislosti (např. jednoduchý požadavek zobrazení určitého typu dokumentu zákazníkovi na portálu může mít za následek změnu skenovacího procesu, změnu metadat, integraci několika softwarů. Něco, co z uživatelského pohledu vypadalo jako drobný požadavek na vývoj, se najednou ukázalo jako složitý a nákladný projekt).

Mezi hlavní cíle této části patří zjištění základního stavu (kde stojíme), stanovení si cílů (kam chceme dojít), ideálně i způsob měření dosaženého cíle a rozsah celé aktivity, vydefinování týmu (s kým).

6.2.1 Stanovení rozsahu aktivity a cílů

Cíl je buď shodný s cílem celého projektu, který změnu vyvolal, nebo je z něho odvozen jako podcíl. Cíl by měl být co nejvíce konkrétní, pro business[22] pochopitelný a měl by mít přidanou hodnotu. Cílem by nemělo být vytvoření firemní klasifikace k dokumentům, to je pouze prostředek k jeho dosažení. Hlavním cílem projektu může být například zavedení

[20] Pro dokumenty personálního oddělení je více hodné použít při třídění organizační pohled, který může být kombinován s procesním pohledem, neboť toto oddělení spravuje nejen dokumenty týkající se jejich klientů (zaměstnanců), ale i dokumenty k některým interním procesům.
[21] Softwary ve firmě můžeme dělit na back-end a front-end. Back-endový software provádí činnosti, které nejsou zřetelné, viditelné pro uživatele. Front-endový software je jeho opakem, je to ta část, se kterou se uživatel setkává – grafické rozhraní, čelní panel u tiskárny apod. (Hlavenka, 1997)
[22] V této práci je termín používán ve smyslu interního zákazníka, důvodem je odlišení mezi interním a externím zákazníkem. Termín je v tomto významu běžně v praxi používán.

specializovaného softwaru pro potřeby procesu vymáhání, z něhož může být odvozen podcíl zrychlení jeho části, která se zabývá kompletací důkazních dokumentů daného dlužníka (vydefinovat taková metadata, aby byly dokumenty bez problémů vyhledatelné).

Ze stanovených cílů je pak patrné, jaký rozsah bude celá aktivita mít. Zde je nutné dávat pozor na fakt, že z pohledu dokumentů se nebudeme například zabývat pouze jejich částí (pouze dokumenty týkající se vymáhání), ale vždy je nutný doménový pohled. Ve firmě můžeme dokumenty rozdělit do několika oblastí, nejzásadnější je doména zákaznických dokumentů, ale nalezneme zde i dodavatelské dokumenty, personální či provozní dokumenty.

To znamená, že pokud řešíme část zákaznických dokumentů, pohlížíme i na drobnou změnu z globálního pohledu a ověříme, že změna v klasifikaci s sebou neponese neočekávané negativní dopady (např. změna hodnoty v metaúdaji „Časové_Razítko = Ne“ způsobí, že dokumenty nebudou označovány elektronickou pečetí[23], což by vedlo k tomu, že není prokazatelné, že byl dokument v čase editován a nemusel by být v soudním procesu uznán jako důkaz).

Z obecného pohledu může kvalitní klasifikace dokumentů přinést:

- vylepšení firemního procesu (např. může dojít ke zkrácení vymáhacího procesu, zlepšení kvality zákaznického servisu, …);
- rozšíření horizontálních kompetencí zaměstnanců (např. zaměstnanec na Front Office již nemusí čekat na odpověď z Back Office, neboť je dokument s požadovanými informacemi dostupný online);
- snížení ceny hardwaru a softwaru (např. nebude docházet k duplikaci dokumentů, dojde ke zkvalitnění integrace mezi programy, ukládání na pomalé/rychlé disky, dostupnost v online/offline režimu);
- snížení rizika auditních nálezů či zákonných postihů (např. nebude problém s anonymizací osobních údajů, neoprávněné či předčasné likvidaci dokumentů);
- snazší určení odpovědnosti a vlastnictví obsahů;
- zprůhlednění práce jednotlivých oddělení, sjednocení společného komunikačního jazyka (např. nemělo by docházek k tomu, že jsou dokumenty tříděny pod synonymními názvy služeb či produktů).

[23] Dříve, dle zákona o 227/2000 o elektronickém podpisu Sb., se používal termín elektronická značka.

6.2.2 Výpočet návratnosti investic a stanovení metrik

Návratnost investic se vypočítává na celý projekt, ale vzhledem k tomu, že správné nastavení klasifikace má dopad na kvalitu dodaného projektu, je možné i zde stanovit, jaké úspory by měla změna přinést. Příklady jsou uvedeny níže.

Vrátíme-li se k uvedenému příkladu týkajícího se vymáhání, pak je možné vypočítat, o kolik se zkrátí proces vyhledání důkazních dokumentů, což s sebou nese nejen zkrácení celého procesu, ale i snížení objemu práce na daného zaměstnance. V případě, že bude nutné celý proces vyhledávání více z automatizovat (vyhledávání bude provádět software namísto uživatele), pak kvalitní klasifikace umožní paralelní vyhledávání v různých uložištích dokumentů pod synonymy.

Jako zkvalitnění služby zákazníkům může být hodnocena dostupnost jejich dokumentů přes zákaznický portál (faktury, smlouvy, pohledávky, ….), což sníží komunikace na front end office. To generuje snížení pracovní zátěže, může vést k personálním úsporám, nebo k horizontálnímu rozříšení kompetencí zaměstnanců, jejich rekvalifikaci na jinou pozici ve firmě.

Pokud se firma rozhodne vydat cestou bez papírové dokumentace, pak můžeme například přistoupit k podpisu smluv v elektronické podobě. Kromě výhody téměř okamžité dostupnosti dokumentu jak uvnitř firmy, tak na zákaznickém portále, to s sebou přináší i snížení nákladů na snímání, archivaci originálů a jejich následnou skartaci, snížení rizika ztráty originálu (např. některé originály z poboček na centrálu nikdy nedorazí a tyto smlouvy jsou pak potenciálně nevymahatelné). Na druhou stranu to s sebou nese požadavky na zajištění dlouhodobé prokazatelnosti autenticity, integrity, pravosti a čitelnosti elektronického dokumentu, což se projeví v nákladech na změny v hardwaru či softwaru. Mimo jiné sem vstupuje i otázka, zda je vhodné zákazníkům na portále poskytovat přímo originál smlouvy, nebo z důvodu bezpečnostního rizika pouze jeho kopii.

Pokud budeme vycházet z průzkumu uvedeného v knize *Enterprise Content Management Solutions* (Forquer 2005), ze kterého vyplynulo, že 30 % zaměstnanců tráví 10 – 20 % svého času hledáním informací. Máme-li firmu o 1 200 zaměstnancích, pak se toto týká 360 z nich. Předpokládáme, že jejich průměrná mzda je 26 000 Kč měsíčně, pokud

zaměstnanci tráví hledáním pouze 10 %, stojí to firmu 936 000 Kč měsíčně, tedy 11 232 000 Kč ročně.[24]

Jak je z čísel patrné, jakékoliv zrychlení procesu vyhledávání je pro firmu přínosem, ušetřením času koncových uživatelů. Zrychlení může být dosaženo mimo jiné tím, že uživatel může použít přednastavenou vyhledávací šablonu, použije automatické předvyplnění dat ve vyhledávacím formuláři, nebo se mu sníží se počet kliků, které musí v systému udělat, než se dostane k požadovanému dokumentu.

Abychom mohli ověřit, že aktivita splnila cíl, musíme být schopni to změřit. Nastavení metrik s tématikou návratnosti úzce souvisí. Pokud jsme si stanovili, že se daný proces zrychlí, pak je nutné proces před změnou a po změně změřit a výsledky mezi sebou porovnat. Další možností je stanovit metriku, která říká, že se sníží chybovost nedohledaných relevantních elektronických dokumentů. Po nasazení změny porovnáme počty nevyhledaných relevantních dokumentů před a po změně. Jedna z metrik může měřit zrychlení odezvy vyhledávání nebo ukládání dokumentů. Aby se uživateli usnadnilo ukládání dokumentů, mohou být některá metadata automaticky předvyplněná, nebo dotahována z jiného softwaru. Také zde můžeme změřit, zda se uložení před změnou a po změně rychlilo (mimo jiné se sníží i chybovost způsobená lidským faktorem).

V budoucnu mohou být tyto metriky použity jako podklady pro uzavření dohody o úrovni poskytovaných služeb (SLA[25]).

6.2.3 Vydefinování pracovního týmu

V rámci fáze vydefinování pracovního týmu je nutné si stanovit, které lidské zdroje a v jakém rozsahu bude aktivita potřebovat. Na první pohled se může zdát, že se hlavně dotýká pouze odborníků z IT, zadavatele aktivity a případně bude nutná součinnost vlastníka dokumentů. Vzhledem ke komplexnosti činnosti je vhodné zjistit a předem si zajistit i čas na konzultace zástupců jednotlivých stakeholderů[26] (nebo aspoň těch, kterých se změna bezprostředně dotýká).

Z pohledu práce se zákaznickými dokumenty obvykle dělíme stakeholdery na:

[24] Statistické údaje týkající se neefektivnosti při rozhodování se různí, například firma Delphi ve své studii (Delphi Group 2002) z roku 2002 uvádí, že 49 % zaměstnanců tráví každý den méně než 2 hodiny vyhledáváním informací, které jsou rozhodující pro výkon jejich činností a 37 % zaměstnanců uvádí délku mezi 2 – 4 hodinami. Je na podniku, zda pro výpočet využije některé ze statisticky dostupných dat nebo zda provede vlastní výzkum týkající se délky času spojené s vyhledávání informací.
[25] SLA je zkratka z anglických slov Service Level Agreement. Jedná se o dohodu mezi zákazníkem a dodavatelem o úrovni poskytovaných služeb.
[26] Stakeholder je každá zainteresovaná strana, která má jakoukoliv vazbu na prováděnou aktivitu v podniku.

- obchodní oddělení,
- finanční oddělení,
- právní oddělení,
- odborné specialisty dle typu firmy (konzultanti, technici, zaměstnanci call centra, zaměstnanci back office, projektový manažeři,…),
- externí subjekty (zákazníci, partneři, státní organizace,…),
- office management (recepční, sekretářky, podatelna, správa spisové služby a archivu,…),
- IT specialisty zajišťující doménu ECM (vývojáři, alanytici, administrátoři),
- management firmy.

Zástupce jednotlivých stakeholderů je vhodné využít nejen pro získání potřebných informací, ale zároveň jako prostředníky, kteří budou informovat své kolegy o probíhající aktivitě. Na základě zapojení zástupců do procesu tvorby se sníží i negativní postoj, který s sebou změna přináší. Zároveň je nutné počítat s tím, že každý jednotlivý stakeholder má na práci s dokumenty jiný pohled, z toho důvodu vzniknou třecí plochy, se kterými je nutné se vyrovnat.

6.2.4 Analýza současného stavu

Pro objektivní posouzení výchozího stavu je nezbytné mít představu o stávající ECM architektuře v podniku (kde a jak jsou zákaznické dokumenty využívány, kdo tyto softwary používá), jaká je IT strategie (Jsou tyto softwary určené k terminaci, k upgradu? Je zde vyvíjen tlak na sjednocení ECM a větší integraci s jinými nástroji?).

Jaká je závazná legislativa a normy, které musí firma z pohledu dokumentů splňovat a nakolik se jí to daří.

Mezi další zajímavé zdroje informací o současném stavu patří návody, postupy práce s dokumenty nebo nálezy auditu.

6.2.4.1 ECM model zralosti

ECM model zralosti (Pelz-Sharpe 2009) je nástroj, který nám pomůže dokreslit aktuální stav ve firmě.

Pro potřeby této metodiky bude využita pouze část, která ovlivňuje oblast klasifikace dokumentů. Model zralosti je vhodný i z toho důvodu, že umožňuje stanovení budoucího

stavu (cíle). Mimo jiné je možné ho využít i při vytváření informační strategie, neboť zobrazuje, kam se firma v dané oblasti může v budoucnu posunout.

Model je rozdělen na 3 základní kategorie: člověk, informace, systém. Dále se pak člení do 13 dimenzí s 5 úrovněmi zralosti: neřízený, vznikající, formativní, provozní, proaktivní.

Z pohledu klasifikace dokumentů jsou důležité tyto oblasti:

- klasifikace – do jaké míry podnik klasifikuje obsah;
- životní cyklus – úplnost řízení životního cyklu dokumentu;
- governance – rozsah pravidel týkajících se informačního managementu;
- opětovné využití – rozsah, s jakým jsou dokumenty opětovně využívány;
- vyhledatelnost – schopnost najít správný dokument ve správný čas.

Podrobná specifikace dle úrovní zralosti je v tabulce 3. Podnik by měl ke každé z pěti dimenzí přiřadit jednu úroveň zralosti z pohledu stávajícího stavu a jednu úroveň zralosti z pohledu budoucího stavu (po dodání projektu).

Každý podnik je jedinečný a z toho důvodu není možné a reálné, aby měl ve všech oblastech nejvyšší úroveň zralosti, potřebná úroveň je dána velikostí firmy (malé podniky nepotřebuji sofistikovaná řešení), strategickými cíli a aktuálními potřebami práce s obsahem.

Tabulka 3: Úroveň zralosti ECM (zdroj: upraveno dle (Pelz-Sharpe 2009))

	Neřízený	Vznikající	Formativní	Provozní	Proaktivní
Klasifikace	Neexistuje formální klasifikace	Třídění na úrovni oddělení; začátek tagování obsahu	Řízené slovníky; počátek firemní klasifikace	Existuje seznam všech úložišť a firemního obsahu; vytvořeno globální třídění	Odstranění nepotřebných dokumentů; vývoj folksonomie; průběžné hodnocení stavu metadat
Životní cyklus	Management životního cyklu neexistuje	Většina obsahu je archivována nahodile; některá oddělení mají v zárodku skartační politiku	Vývoj skartační politiky, zavedení spisové služby	Spisová služba (papír i elektro)	Všechny typy dokumentů procházejí formálním životním cyklem
Governance	Neexistují ani pravidla, ani postupy	Žádné nebo malé množství formálních politik; politiky nejsou centralizované	Vývoj informační politiky a kodifikace postupů	Pravidla a postupy jsou široce rozšířené	Pravidelné přezkoumávání stavu a případné potřebné opravy; zpětná vazba od klíčového zákazníka
Opětovné využití	Rutinní duplikace obsahu	Obsahy jsou stále rutinně duplikovány	Počáteční analýza obsahu	Změna v používání dokumentů napříč systémy a kanály	Obsahy jsou používány napříč systémy a kanály

Vyhledatelnost	Zaměstnanci tráví dlouhý čas při vyhledávání obsahu v různých portálech	Aplikace základních metadat; hledání, jakým způsobem vyladit vyhledávání přes portály	Realizace vyhledávací technologie; analýza vyhledávání a další ladění vyhledávání	Vývoj federativního a/nebo celopodnikového vyhledávání	Vyhledávání a klasifikace jsou na celofiremní úrovni

6.2.4.2 SWOT analýza

SWOT analýza je analytický nástroj, který je možné využít pro zhodnocení vnějších a vnitřních faktorů mající dopad na úspěch konkrétního záměru. Velmi často se používá na poli marketingu a strategického řízení. Vznik je datován do šedesátých letech 20. století a jejím autorem je Albert Humphrey.

Podobně jako v případě PDCA metodiky je název složen z počátečních písmen jednotlivých faktorů: silné stránky (strengths), slabé stránky (weaknesses), příležitosti (opportunities) a hrozby (treats).

Tento nástroj je výhodný pro sjednocení a vyhodnocení poznatků, které byly získány postupy uvedenými výše.[27]

Níže je zobrazen praktický příklad využití analýzy z pohledu klasifikace zákaznických dokumentů.

Tabulka 4: příklad SWOT analýzy (zdroj: autorka)

Silné stránky: • Existence základních postupů pro práci se zákaznickými dokumenty (dokumenty jsou ukládány na jedno místo, s definovanou syntaxí) • Existence číselníků v rámci jednotlivých oddělení • Nastavení procesu uložení zpráv přijatých přes datovou schránku	**Slabé stránky:** • De-duplikace obsahu • Špatná vyhledatelnost dokumentů • Nejednotnost klasifikace – uživatelsky nepřívětivé vyhledávání
Příležitosti: • Zefektivnění stávajících procesů • Zvýšení kvality zákaznických služeb (přístup k dokumentům přes portál 24/7/365) • Vytvoření komunikační platformy pro výměnu zákaznických dokumentů s partnery	**Hrozby:** • Zákonné nedodržení norem (problém s anonymizací a skartací dat) • Únik/zneužití zákaznických dat (špatná klasifikace citlivého obsahu)

Pro případy, že by tento nástroj nebyl plně vyhovující (bylo by nutné zmapovat rozdíly mezi stávajícím a cílovým stavem, navrhnout a zhodnotit varianty řešení), je možné využít **GAP analýzu**[28].

[27] K nevýhodám tohoto nástroje obecně patří subjektivnost, statičnost, riziko zkreslení a vytvoření chybných závěrů.

6.3 Tvorba obecného návrhu klasifikace

Obecný návrh klasifikace je vize – ideální stav, jak by dokumenty měly být tříděny, pokud by neexistovala technická, uživatelská a jiná omezení.

Před tím, než se pustíme do vytváření samotné koncepce podnikového třídění, musíme sebrat data, ze kterých budeme následně při tvorbě vycházet. Jako vstupní vhled je možné využít výstup z fáze plánování.

Jako další zdroje informací lze využít:

- spisový a skartační plán organizace,
- dokumentace k firemním procesům,
- návody uživatelů u stávajících systémů,
- standardy, normy a legislativa platné pro dané odvětví
- dokument obsahující informační politiku podniku,
- stávající metadata v používaných systémech,
- využité analytické nástroje v organizaci,
- nálezy auditů,
- strategie podniku – přináší informace o budoucích cílech firmy, je nutné třídění vytvořit tak, aby bylo možné do ní nové změny začlenit bez velkých problémů.

Pro získání dat od koncových uživatelů je vhodné využít tyto metody:

- Dotazníkové šetření – umožňuje rychle získat základní informace od velkého počtu uživatelů. Může být koncipováno tak, že se analyzuje celá tematická šíře nebo se zjišťuje pouze její část. Je lepší použít větší množství krátkých dotazníků (s určitým časovým odstupem) než jeden velký. Ideální je dotazníky tématicky přizpůsobit podle skupin (uživatelé z oddělení personálního manegementu nebudou schopni zodpovědět otázky na téma ukládání zákaznické projektové dokumentace apod.).
- Řízené rozhovory – aby byl rozhovor co nejvíce přínosný, je nutné, aby se na něj dotazovaný předem připravil. Rozhovor by měl být veden podle předem připravených tematických oblastí a vše by mělo být zaznamenáno – ideálně

[28] GAP analýza, někdy též diferenční analýza, je metodika umožňující rozhodování a řešení problému.

zaznamenávat do tabulkové formy, aby pak bylo snazší s výstupy pracovat. Rozhovor by měl být veden tam, kde nebudou dotazovatel ani tázaný rušeni.

- Pozorování – velmi efektivní metoda. Analytik, který bude nějakou dobu u uživatele přítomen, pak bude schopen kvalifikovaně zodpovědět dotazy, provádět prvotní uživatelské testy, aniž by byla nutná součinnost koncového uživatele.

Není nutné vybrat pouze jednu z výše uvedených metod, lze použít jejich kombinaci.

V rámci sběru dat bychom měli získat veškeré dostupné informace o způsobu ukládání, využívání či archivování zákaznických obsahů. K tomuto účelu nám slouží tabulka níže.

Tabulka 5: Způsob práce se zákaznickým obsahem (zdroj: autorka)

Oblast zákaznických dokumentů	Faktury	Smlouvy	…
Oddělení (vlastník dokumentu)			
Popis (k jakým účelům se dokument používá, jaké informace obsahuje,…)			
Je dokument v el. podobě originálem?			
V jakých formátech se může vyskytovat? (doc, docx, pdf,…)?			
Používá dokument i jiné oddělení (je jeho autorem, jeho adresátem,…)?			
S jakými metaúdají se ukládá? (seznam využívaných číselných hodnot metadat)			
S jakými metaúdaji je nejčastěji vyhledáván?			
Kde (v jakém systému) je dokument uložen?			
Je ukládána poslední/finální verze dokumentu, nebo je dokument verzován?			
Podléhá dokument schvalovacímu řízení?			
Je vkládán dokument do systému uživatelem, nebo je ukládán systémem? Pokud systémem, jde o synchronní či asynchronní komunikaci?			
Pokud je dokument vytvářen z formuláře/ze šablony, název/ID šablony			
Je dokument používán pouze v daném systému, nebo propagován do jiných systémů?			
Z jakých typů rozhraní se k němu přistupuje? (přes webové rozhraní, přes mobilní aplikace, přes desktop aplikaci,…) Popř. jaké rozhraní by			

uživatelé do budoucna pro svoji práci uvítali?			
V jakých procesech je dokument použit a v jakém vztahu?			
Jaká jsou přístupová opravnění k dokumentu? (kdo smí dokument vidět, editovat, měnit metadata, mazat,…)			
Stupeň utajení			
Je dokument online po celou dobu, nebo se archivuje a je po určité době dostupný pouze v offline?			
Na základě čeho a po jaké době je možné dokument skartovat?			
Je dokument ukládán pod zákaznickým číslem? Existuje-li ve firmě více úrovní zákaznických čísel, pod kterým z nich?			
Je požadována AIPČ (autenticita, integrita, pravost, čitelnost) dokumentu, pokud ano, jakým způsobem?			

Při sběru dat, který bude nejspíše prováděn metodou rozhovoru, je vhodné si ke každému typu dokumentu vyžádat jeho reálný příklad, ve kterém je vhodné si poznačit údaje, pod kterými je do systému ukládán a jak se v systému jmenují (ne vždy se název údaje na papíru bude rovnat názvu metaúdaje v systému).

Tabulka může být verifikována či doplněna na základě studie zdrojů informací, jejíž seznam je uveden výše.

Z pohledu metadat je dobré vytvořit ještě jednu tabulku, která bude párovat jednotlivé názvy používaných jednotlivými stakeholdery, popř. bude nést informaci, zda se přes tento metaúdaj i vyhledává nebo se pouze zobrazuje ve výsledcích zobrazení. Příklad tabulky je níže.

Tabulka 6: Příklad synonymních názvů u metadat (zdroj: autorka)

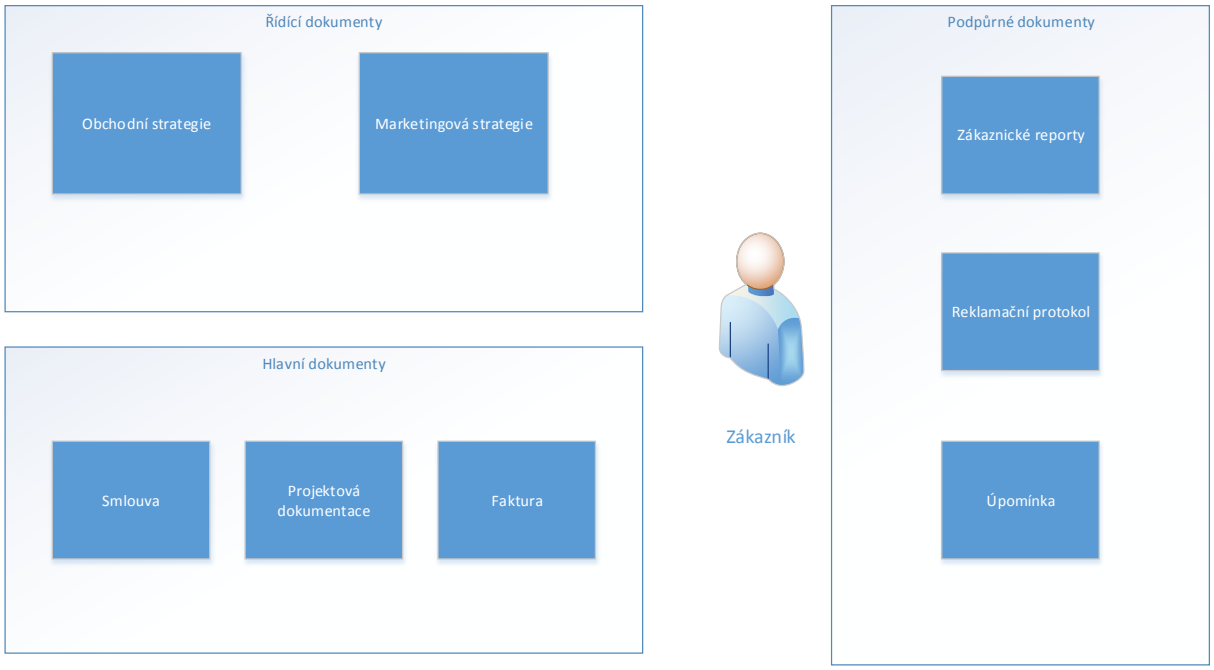
Metaúdaje			**Vyhledává se přes něj v ECM?**	**Zobrazuje se ve výsledcích v ECM?**
Číslo zákaznické smlouvy (papír)	Číslo zákazníka (CRM)	Číslo kontraktu (ECM)	Ano	Ne
Číslo faktury (papír)	Interní číslo faktury (ERP)	Číslo daňového dokladu (ECM)	Ano	Ano
Číslo obchodního případu (CRM)	Externí ID (ERP)		Ne	Ano

Pozn. Tabulka mapuje metadata z pohledu uživatele. V případě, že existují metadata, která si mezi sebou vyměňují systémy při interakci a mají jiný název, je vhodné tuto tabulku rozšířit i o tyto údaje.

6.3.1 Procesní mapa dokumentů

Jak je výše popsáno, na zákaznické dokumenty je vhodné nahlížet procesním pohledem. Je to z toho důvodu, že dokumenty jsou často jejich součástí a v rámci procesů se podnikem pohybují. Protože většinou nejde o jeden proces, v rámci kterého je s dokumentem v organizaci pracováno, je nutné vytvořit procesní mapu dokumentů.

Použijeme-li analogické dělení, jako je tomu u procesů, pak můžeme dokumenty rozdělit na hlavní, řídící a podpůrné. Jak toto dělení může ve firmě vypadat, uvádím na obrázku 8.



Obrázek 8: Rozdělení dokumentů (zdroj: autorka)

Obrázek 9 demonstruje, jakým způsobem se s dokumentem ve firmě pracuje (jeden dokument jako součást různých procesů) a jak může procesní mapa dokumentů vypadat. Obchodník vytvoří objednávku, jejíž podepsanou formu následně obdrží podatelna. Nejpozději zde je pro obchodníka objednávkový proces ukončen. S dokumentem ovšem dále pracují jiná oddělení, například call centrum, které ho potřebuje při řešení problému (např. služba není dle domluvených podmínek ve smlouvě aktivní), proces končí po dořešení problému. V poslední řadě může pak s dokumentem pracovat zaměstnanec, který má na

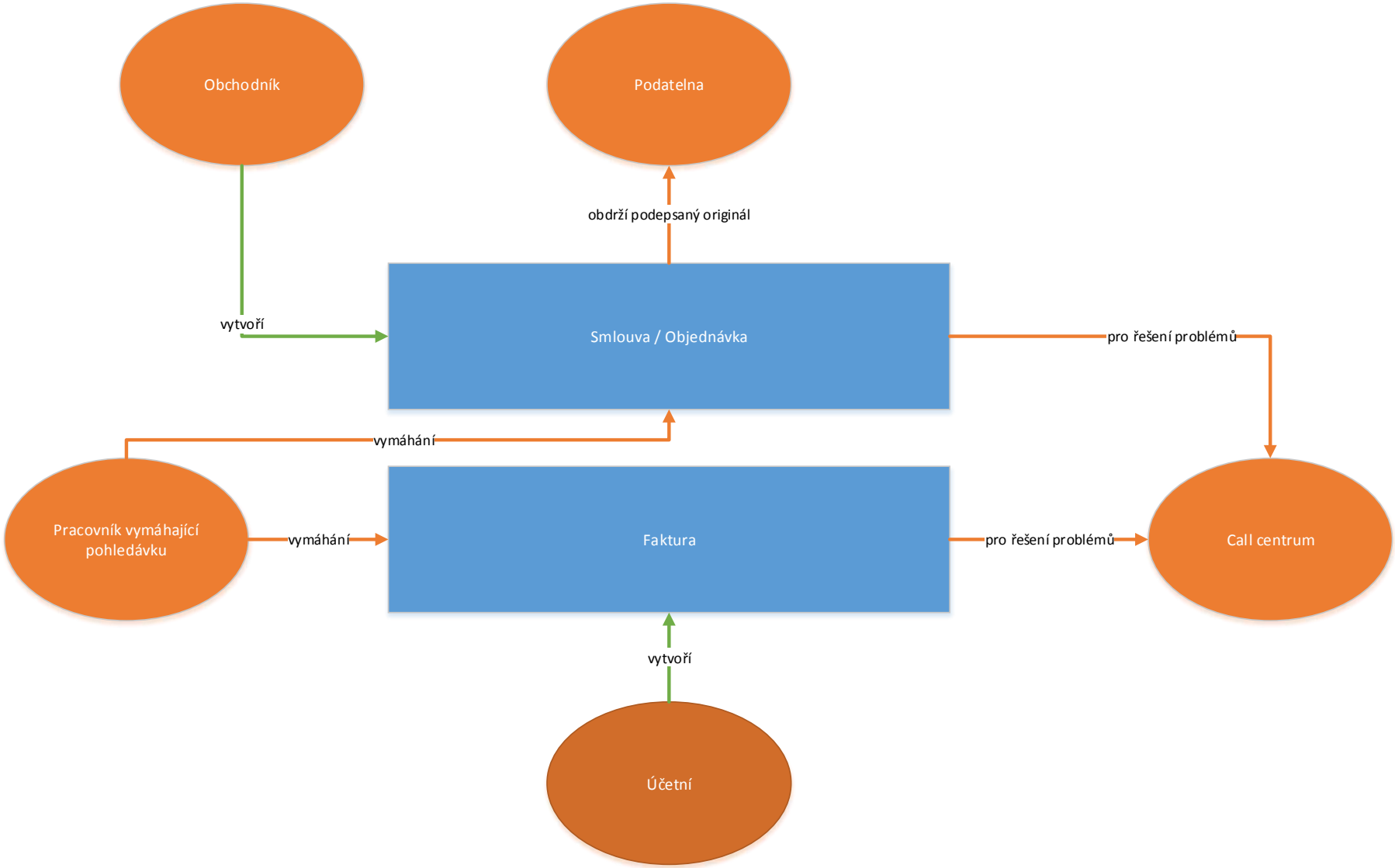
starosti vymáhací proces, neboť zákazník za poskytnuté služby nezaplatil, a tak bude firma platbu vyhánat soudní cestou, tzn. objednávku mají k dispozici právníci.

Krom přehledové procesní mapy je možné pro přehlednost vytvořit procesní mapu pro každý dokumentový typ.

Mapa dokumentů by měla být doplněna o dokumentaci jednotlivých typů dokumentů v tabulkové podobě, pro její tvorbu je možné vycházet z vyplněné tabulky 7.

Tabulka 7: Příklad tabulkového zápisu typů dokumentů (zdroj:autorka)

ID dokumentu	Název dokumentu
DAN	Daňový doklad
DAN 01	Faktura
DAN 01.01	Podrobný výpis za jednotlivé využívané služby
DAN 01.02	Příloha k faktuře
DAN 02	Dobropis
SML	Smlouva
SML 01	Rámcová smlouva
SML 02	Dodatek týkající se poskytovaných služeb
OBJ	Objednávka
OBJ	Objednávka služeb
OBJ	Objednávka produktů



Obrázek 9: Procesní mapa dokumentů (zdroj: autorka)

Atributy popisující dokumentové typy jsou představeny níže, opět je vhodné pro přehlednost použít tabulkovou formu.

Tabulka 8: Popis dokumentu (zdroj:autorka)

Název dokumentu			
Číslo dokumentu		**Číslo/a procesu/ů**	
Popis dokumentu			
Stručný popis dokumentu + co je jeho obsahem.			
Metadata			
BS metadata, pod kterými se dokument ukládá BS metadata, pod kterými se dokument zobrazuje Administrativní metadata			
Oprávnění			
Stupeň utajení		**Osobní údaje**	**Ano/Ne**
Právo tvorby dokumentu		**Právo zobrazit dokument**	
Právo editace dokumentu		**Právo mazat dokument**	
Technické údaje			
Typy použitých MimeType: **Vstupní systém:** **Výstupní systém:** (ideálně u každého vyjmenovat seznam používaných metadat)			
Archivace / Skartace			
Typ skartace: (podmíněná, nepodmíněná)			

Skartační znak: (A – archiv , S – skart, V - výběr) **Délka skartační doby:** **Začátek plynutí skartační doby:**
Informace o originále
Informace o tom, zda je či není dokument originál. Není-li originál, napsat, jaká metadata jsou externími klíči na uloženém dokumentu, na základě něhož se dokument v archivu dohledává
Metriky
Poznámky
Vlastník dokumentu
Legislativa, metodiky, standardy
Popis specifik a povinností vyplávajících ze zákona, z interních metodik (typ vzorového formuláře), či norem

6.3.2 Popis postupu tvorby firemního třídění

Při tvorbě klasifikace je nutné nezapomínat na její čtyři základní cíle: **identifikovat obsah, jeho správu, pomoc při vyhledávání obsahu a sledovat jeho využívání**, více viz teoretická část této práce (kapitola 3.3.1).

V teoretické části je uvedeno 9 pravidel, která by měla navržená kategorizace splňovat, aby se tím realizovaly hlavní cíle. Ke každému pravidlu byla přidána sada

kontrolních otázek, které by se při tvorbě neměly nespouštět ze zřetele. Seznam je vhodné použít i při finální kontrole navržené klasifikace.

- **Intuitivní** – Používají uživatelé třídění bez problémů? Dokážou předvídat, v jaké kategorii naleznou potřebný dokument?
- **Jednoznačnost** – Neexistuje více možností, kam daný dokument uložit? Jsou uživatelé nuceni se při ukládání dokumentu rozhodovat, kam ho zatřídit?[29]
- **Pohostinnost** – Neexistuje dokument, který není možné do kategorizace zatřídit, nebo-li jsou všechny dokumenty ve firmě klasifikovatelné? Nemají uživatelé tendenci při ukládání volit nevhodné či obecné kategorie (např. výběr z položky jiné, ostatní apod.)?
- **Konzistetnost a předvídatelnost** - orientují se uživatelé rychle a bez chyb v rámci klasifikace?
- **Relevantnost** – Používá klasifikace stejný jazyk jako její uživatelé? Odráží způsoby uspořádání informací ve firmě?
- **Šetrnost** – Není v třídění redundance?
- **Smysluplnost** – Splňuje vyhledávání dokumentů princip přesnosti?
- **Trvanlivost** – Jak často bude nutné provádět úpravy, revize navrženého třídění?
- **Vyváženost** – Neobsahují některé kategorie větší množství dokumentů? Nedochází zde k přetížení? Pokud ano, je to způsobeno špatně navrženou klasifikací?

Jak je tomu i v jiných procesech, klíčovým (a zároveň i rizikovým) prvkem procesu je zákazník. V našem případě jakýkoliv uživatel, který potřebuje manipulovat s firemním obsahem. Proto by měl mít vždy možnost se k navržené firemní klasifikaci vyjádřit, neboť jiné třídění vytvoří specialista pro daný obor, zcela jiný pohled má laik a specifický pohled má také konzultant specializující se na tvorbu klasifikace. Každý z nich používá rozdílný jazyk (slang), který je ovlivněn firemním prostředím. Navržená kategorizace by měla tento

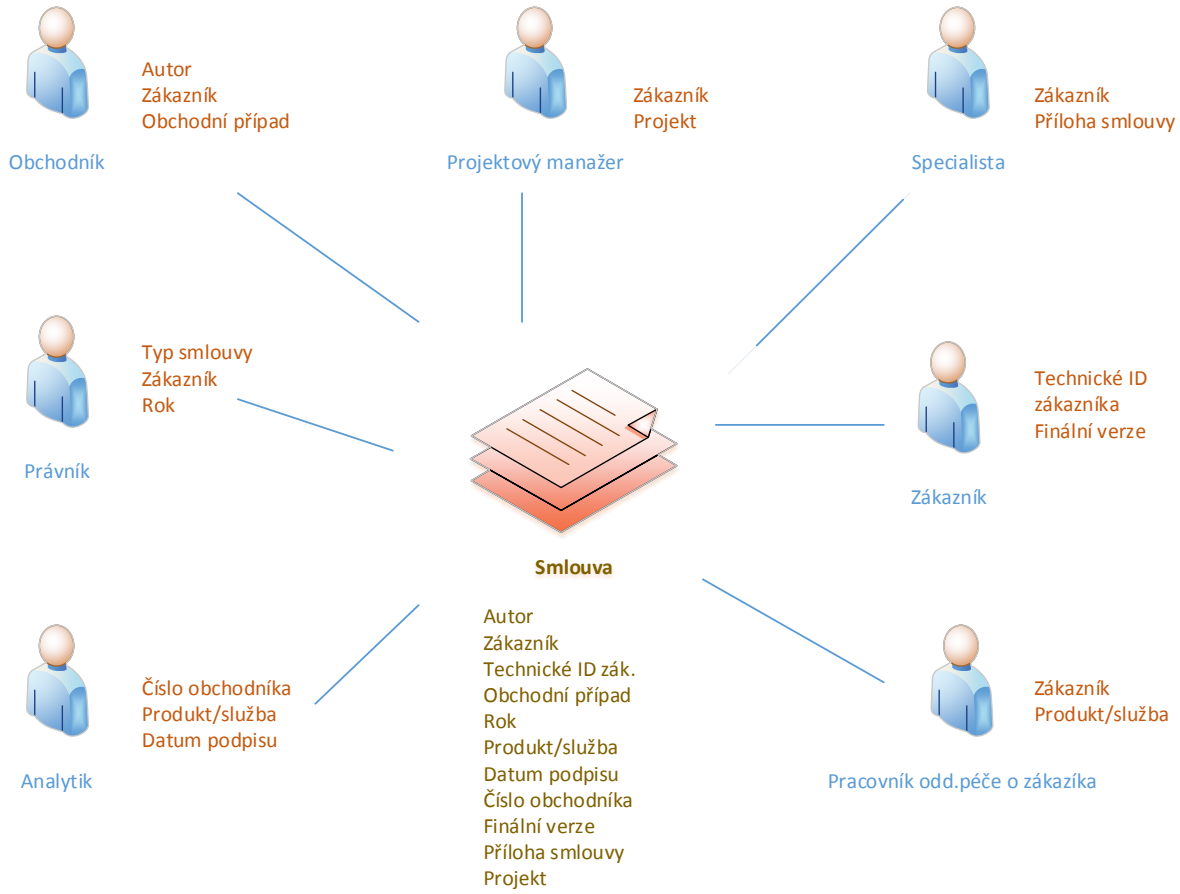
[29] Zde je tato vlastnost myšlena ve smyslu, že navržené třídění je jednoznačné, neexistují zde synonymní, hononymní výrazy či únikové výrazy typu ostatní, další apod. Není tím myšleno, že je nutné dokument jednoznačně uložit pouze do jedné kategorie.

jazyk odrážet a neměla by (ani to není jejím cílem) uživatele učit používání správné terminologie daného oboru.

Nejspíše je to překvapivé, ale při tvorbě klasifikace se musí začít od konce, tzn. od vyhledávání dokumentů. Neboť chceme-li dokument nalézt, musí obsahovat metaúdaje, se kterými uživatelé v běžném provozu pracují. Je tedy nutné zajistit, aby dokument metadata obsahoval a aby byla plněna.

V tomto případě nás zajímá, jakým způsobem uživatel dokumenty specifikuje. Každý uživatel bude mít na třídění jiný pohled, který se odvíjí od jeho zkušeností a vnímání reality, dostupných informací o dokumentu a pracovním procesu.

Jako modelový příklad různého přístupu uživatelů k metaúdajům poslouží dokument typu smlouva. Obchodník bude potřebovat smlouvu zákazníka, například proto, že by ji rád použil jako vzorovou při tvorbě smlouvy pro jiného klienta. Bude nejspíše dokument hledat pod metadaty autor dokumentu, konkrétní zákazník nebo obchodním případem. Právní oddělení bude mít odlišný pohled v tom, že jej zajímá typ smlouvy, zákazník a rok. Oddělení péče o zákazníka, které s ním řeší problém na poskytované službě či produktu, bude naopak hledat pod zákazníkem a produktem/službou. Projektového manažera nezajímá ani produkt či obchodní případ, ale chce najít dokument pod projektem. Zákazník může k dokumentu přistupovat pomocí portálu, v tomto případě mu chce firma zpřístupnit finální verzi dokumentu, pro zvýšení bezpečnosti je zobrazení realizováno pod technickým číslem zákazníka (tedy pod jiným, než se kterým pracují uživatelé). Součástí smlouvy je i příloha, která technicky specifikuje poskytovanou službu/produkt (např. lokace, způsob zapojení, reakční doby, metriky,…). Pouze tuto část chceme zpřístupnit specialistovi, který má zákazníka na starosti, zvl. v případě, že se může jednat o subdovatele. V poslední řadě nám vstupuje do hry i uživatel, který zajišťuje tvorbu reportů, toho například z pohledu smlouvy zajímá číslo obchodníka, datum podepsání a kód produktu/služby. Grafické vyobrazení business pohledů práce s dokumentem je zobrazen níže.



Obrázek 10: Business pohled na smluvní dokument (zdroj: autorka)

Potřebné informace získáme metodami a ze zdrojů popsaných výše. Aby bylo možné s informacemi dále pracovat – budou použity pro tvorbu datového modelu ve fázi implementace, zaznamenáme business pohledy v tabulkové podobě. Tabulka je k dispozici níže.

Toto je zjednodušená verze, která bude v rámci implementace rozšířena o jiné hodnoty. V tabulce nejsou ještě uvedeny konečné názvy metaúdajů (jde o tvar, který je nyní používán v systémech nebo ho takto nazvali sami uživatelé). Pojem „použito“ je myšleno jak z pohledu vyhledávání, tak z pohledu výsledku vyhledávání. Do pole „hodnota“ zaznamenáme specifikaci, co se do pole vyplňuje[30], můžeme zde vypsat i hodnoty z již existujícího číselníku, jak tomu je u příkladu metaúdaje „typ dokumentu“. Do pole vyhledávání zaznamenáme, jakým způsobem se daný dokument vyhledává.

[30] Někdy je metaúdaj nazván genericky a nabývá v rámci jednotlivých typů dokumentů různých hodnot. Například číslo zákazníka může obsahovat různé formáty zákaznických ID, což může být způsobeno změnou CRM systémů, různými migracemi apod.

Tabulka 9: Příklad business pohledů (zdroj: autorka)

Třída			**Smluvní dokumenty**			**Daňové doklady**	
Název dokumentu			*Smlouva na produkt*	*Smlouva za poskytnutí služeb*	*Technické přílohy*	*Faktura*	*…*
Popis			smluvní dokument	smluvní dokument	dokument obsahuje technickou specifikaci	faktura	
Uživatel			obchodník, zam. péče o zákazníka, oddělení vymáhání, projektový manažer, právník, zákazník	obchodník, zam. péče o zákazníka, oddělení vymáhání, projektový manažer, právník, zákazník	zam. péče o zákazníka, projektový manažer, právník, **specialista**, zákazník	zam. péče o zákazníka, oddělení vymáhání, zákazník	
Vyhledávání			Vyhledat smlouvu pod obchodníkem, zákazníkem, daným obchodním případem. Vyhledat dokument pod zákazníkem a danou službou.	Autor = Obchodník AND Zákazník = (ID zákazníka) AND Obchodní případ = (ID obchodního případu); Zákazník = (ID zákazníka) AND Projekt = (ID projektu) AND Typ dokumentu = SMLSL	Vyhledat dokument technickou přílohu pod zákazníkem.	Zákazník = (ID zákazníka) AND Typ dokumentu = DANFAK	
Metadata	**Autor**	Použito	Ano	Ano	Ano	Ano	
		Hodnota	obchodník	obchodník	specialista	účetní	
	Zákazník	Použito	Ano	Ano	Ano	Ano	
		Hodnota	ID zákazníka	ID zákazníka	ID zákazníka	ID zákazníka	
	Číslo obchodníka	Použito	Ano	Ano	Ne	Ne	
		Hodnota	ID ochodníka	ID obchodníka	ID obchodníka	ID obchodníka	
	Typ dokumentu	Použito	Ano	Ano	Ano	Ano	
		Hodnota	SMLPR[31]	SMLSL	SMSLPRTECH	DANFAK	

[31] Použitá zkratka není reálná, jde o modelový příklad.

Zápis může být proveden buď v přirozeném jazyce, nebo formálním zápisem – oba případy jsou uvedeny v tabulce. Tabulku je možné si upravit dle potřeby, například se může metaúdaj „použito“ již více specifikován, zda je metaúdaj vyhledáván, nebo se zobrazuje ve výsledcích vyhledání, nebo se používá pouze na úrovni detailního pohledu.

Vezmeme-li vytvořené tabulky z předchozí části a tabulku analýzy business pohledů, jsme nyní schopni navrhnout firemní klasifikaci zákaznických dokumentů. Ve zkratce lze říci, že se sjednotí názvy termínů, odstraní se synonyma, vydefinují se číselníky.

V teoretické části je popsáno, jakým způsobem můžeme dělit metadata v rámci ECM domény. Pro potřeby metodiky je zásadní třídění metadat (tabulka 10). Metadata můžeme rozdělit na **systémová metadata** (jsou již v systému definována, hodnoty do nich mohou být vkládány uživatelem nebo systémem) a **klientská metadata** (jsou to veškerá metadata, která byla vytvořena na základě potřeb businessu v daném systému). Následně jsou tyto oblasti děleny dle účelu použití, které používá norma ISO 23081-1. Ke každé části jsou přidána základní metadata zákaznických dokumentů, která jsou představena níže.

Tabulka 10: Třídění metadat pro potřeby metodiky (zdroj: autorka)

Systémová metadata	Metadata	Klientská metadata	Metadata
-		**Metadata pro obchod**	Číslo zákazníka
Metadata pro ochranu	Zadržení (hold)	-	Skartační znak Fixní doba Datum výmazu
Popisná metadata	Interní identifikátor Název Jazyk Autor Datum vložení Datum úpravy Verze	**Popisná metadata**	Popis/Abstrakt Typ dokumentu Jazyk Autor Název Externí identifikátor
Metadata pro vyhledávání	Formát souboru	**Metadata pro vyhledávání**	Druh dokumentu Třídění
Metadata pro správu oprávnění	Oprávnění	**Metadata pro správu oprávnění**	Úroveň utajení

Z tabulek nám vychází, jaké oblasti[32] zákaznických dokumentů ve firmě nalezneme a s jakými vlastnostmi. Na první pohled je patrné, že dokumenty v rámci jedné třídy mají téměř shodné metaúdaje než dokumenty z jiné třídy. Nicméně vzhledem k tomu, že v této práci řešíme pouze část firemních dokumentů (zákaznické dokumenty), nenalezneme zde takové rozdíly, kdybychom například srovnávali dokumentové třídy z jiných oblastí.

V rámci zákaznických dokumentů se setkáme s těmito třídami:

- strategické a marketingové dokumenty,
- obchodní dokumenty,
- smluvní dokumenty,
- daňové dokumenty,
- provozní dokumenty (jejich členění se odvíjí od typu obchodu dané organizace – projektová dokumentace, technická dokumentace, …),
- komunikace s klientem (papírová korespondence, e-mailová korespondence, zápis z telefonického rozhovoru, krátké textové zprávy, obsahy ze sociálních sítí, poukazy, předávací protokoly,…),
- dokumenty související s vymáhacím procesem.

Při návrhu metadat je nutné dodržet, kromě jiného, několik **obecných principů**:

- Použitý jazyk u metadat by měl být jeden (některé softwary dovolují použít funkcionalitu, která umožňuje zobrazovat název metaúdaje v různých jazycích), používat pouze jeden tvar jmen a sloves (např. u podstatných jmen použít 1. pád, u sloves 3. osobu) a sjednotit u jmen číslo (buď singular nebo plural).
- Pokud to již interní směrnice podniku nestanovuje, pak definovat syntaxi zápisu (způsob použití malých a velkých písmen, oddělovačů, prefixů a sufixů).
- Nepoužívat synonymní ani homonymní tvary, neboť tyto tvary mohou uživatele při práci zmást, dávají prostor k diskuzím a k dohadům (tzn. používat jednoznačně vysvětlitelné termíny).

[32] Termín oblast je dále nahrazen termínem třída, který odpovídá terminologii používané pro klasifikaci dokumentů. Třída je množinou jedinců určitého typu.

Sloučit ty metaúdaje, které obsahují stejné údaje (např. číslo zákazníka), jen jsou pod různými třídami pojmenovány jinak – viz mapovací tabulka synonymních názvů výše.

V případě, že metaúdaj bude mít předdefinovaný seznam možností, je opět nutné provést jejich sjednocení. Tento seznam může obsahovat kód[33], a nebo samotný popis. Výhoda kódu je ta, že věcný obsah kódu se může časem měnit (rozšiřovat/zužovat jeho obsah), aniž by bylo nutné provést změny v systému. Nevýhodou je, že uživatel musí znát obsah kódu, aby jej mohl použít[34]. Popis je pro uživatele pochopitelný, ale je u něj nutná častější revize. Kódy mohou mít podobu číselnou, abecední, alfanumerickou. Při abecední podobě se vychází z popisu hodnoty (např. smlouva = SML, faktura = FAK,…), pomáhá to lehčímu zapamatování. I zde by se měl dodržovat jednotný postup tvorby (např. první písmeno každého slova z popisu). V seznamu by neměly být termíny, které by pak mohly sloužit k tomu, že je uživatel zvolí, když neví, pod jakou hodnotu dokument uložit (např.: další, jiný, neodpovídá, nespecifikováno,…).

Tvorbu metadat bych rozdělila na dva extrémy: generický – minimální počet metadat, a detailní – velké množství specifických metadat použitelných z větší části pouze pro jednu třídu dokumentů. Oba extrémy mají své výhody a způsoby použití, při tvorbě zákaznické dokumentace je ale dobré najít ideální poměr mezi extrémy.

Generický přístup volí obecné názvy metadat, díky tomu jsou opakovatelně použitelné přes všechny třídy. Jako příklad můžeme uvést: číslo dokumentu, datum, specifikace objektu. Chceme-li uložit fakturu, do čísla dokumentu napíšeme číslo faktury, do data datum splatnosti a specifikaci objektu nastavíme na hodnotu faktura. U smlouvy budeme postupovat obdobně, specifikace objektu bude smlouva, v čísle dokumentu bude číslo zákaznické smlouvy a do data vyplníme, od kdy je smlouva platná. Mezi výhody tohoto systému patří to, že jsou metadata opakovatelně využitelná a je jich málo. Většinou se tento přístup používá v případě, že uživatel nepřistupuje přímo k úložišti, ale přes jiný software, tzn. je v pozici back-endu.

Detailní přístup naopak volí co nejvíce specifické názvy metadat, které jsou použitelné v rámci třídy a minimálně mimo ni. U tohoto typu se setkáme s velkým množstvím popisných metadat. To většinou sledujeme, pokud uživatelé pracují přímo se softwarem, který umožňuje práci s dokumenty, tzn. technologie ECM je zde v pozici front-endu. K výhodám front-endu

[33] Nepodařilo se mi najít studii, která by se zabývala tím, zda si uživatelé lépe pamatují kódy složené z čícel, textovu podobu, nebo kombinaci obojího. Z toho důvodu jsem provedla malou anketu mezi odborníky, kteří se účastnili připomínkování práce, a zjistila jsem, že nejvíce oblíbené jsou číselné kódy, dále pak smíšené a až na konci čistě textové.

[34] Nebo musí být naprogramována funkcionalita, která uživateli zobrazí popis, ale do systému uloží kód.

patří možnost dohledatelnosti pod různými metadaty, měly by zde být pokryty všechny vyhledávací strategie businessu. V praxi platí, že váže-li se k druhu dokumentu velké množství metadat, přestávají uživatelé nezajímavá data plnit. Příklad detialního třídění je uveden v příloze č. 3, kde jsou metadata tříd *Article* a *CreativeWork*, společná metadata mají pouze ta, která dědí z nadřazené třídy Thing. Třídy pocházejí z ontologie Schema.org.

V dnešní době existují seznamy metadat v různých normách, standardech (viz uvedené příklady v teoretické části – kapitola 4.8 a 4.9), které jsou buď na obecné úrovni, nebo definují metaúdaje se specifickým záměrem. Při tvorbě zákaznických metadat můžeme z těchto dokumentů vycházet. V rámci této práce bylo provedeno hrubé porovnání některých z nich, které se nachází v příloze č. 2[35]. Ze srovnání jsou patrná základní metadata, která nalezneme u většiny. Jedná se o metaúdaje: *identifikátor, název, autor, popis/abstrakt, pokrytí/místo, externí identifikátor, typ dokumentu, klasifikace, datum, oprávnění, formát, jazyk.*

Při podrobném prozkoumání zjistíme, že u porovnávaných dokumentů je možné použít členění metadat dle ISO 23081-1, a to, které je součástí tabulky č. 10. Zajímavým zdrojem metaúdajů, který je v tabulce uveden, je DBpedia – termín dokument. Nejedná se o normu či standard v klasickém pojetí, ale je to ontologie[36]. Vzhledem k tomu, že se ontologie v ničem zásadním neodchyluje od ostatních standardů, je to důkazem toho, že jak metadata v normách, tak ontologie sledují stejný cíl a to určit vlastnosti/metadata, které co nejblíže vyjádří obsah uloženého dokumentu.

Z přílohy č. 2 se vychází při sestavování minimální sady metadat (metadatové údaje z tabulky jsou dále rozšířeny o specifické parametry typické pro zákaznické dokumenty).

Při ukládání obsahu do systému systém k dokumentu vygeneruje jedinečný **identifikátor**, který má v celém systému pouze tato položka (shodné ID nedostane ani verze daného objektu). Jeho struktura a způsob generování jsou různé. V praxi se pro tento účel používá GUID, globálně jedinečný identifikátor (z aj. globally unique identifier) je to 128bitové číslo sloužící k identifikaci objektů v IT systémech[37].

[35] Srovnání metaúdajů není v některých případech přesné, neboť to jejich povaha použití neumožňuje, vztahy mezi nimi mají synonymní povahu. Z toho důvodu můžeme v porovnání v rámci jedné normy najít i více termínů. Pro účely této práce je to plně dostačující, smyslem srovnání bylo najít metadata, která jsou pro všechny v průniku.

[36] Původně jsem zamýšlela pro tento účel využít termíny z ontologie Schema.org, ale vzhledem k jejich složitosti (velké množství metaúdajů k objektu) to nebylo pro tyto účely vhodné.

[37] Tento termín pochází od firmy Microsoft, synonymní varianta termínu je UUID (universally unique identifier).

Podobně je na tom i atribut **externí identifikátor**, který má za cíl odkazovat, pod jakým identifikátorem je objekt dohledatelný a dostupný jinde. Typickým příkladem použití je *ISBN*[38] a nebo *DOI*[39] v diginálním světě. V podnikové praxi se potkáme s jeho využitím v několika podobách. Za prvné může být plněn hodnotami, jež jsou interními identifikátory shodného objektu v jiném systému, zvláštní využití má při migraci dokumentů mezi systémy. Dále to může být externí identifikátor pro všechny zúčastněné strany, k jeho generování dochází v případě, že je v rámci nějakého procesu potřeba identifikace daného objektu dříve, než objekt interní identifikátor dostane. Takovým metadatem může být například hodnota pole z *čárového kódu*[40], který je připevněn na analogový nosič dříve, než je vytovřen záznam v systému.

Formát souboru v digitální podobě patří mezi systémová metadata, pro tento účel je nejčastěji využívána tzn. Mime Type, což je internetová norma[41]. Zajímavostí je, že vznikla původně pro potřeby rozšíření elektronické pošty o nové možnosti přes SMTP. Odtud původ anglického názvu, který je složen z počátečních písmen anglických termínů Multipurpose Internet Mail Extension (výceúčelové rozšíření internetové pošty). Tento metaúdaj má několik způsobů využití. Systém může na základě této hodnoty volit, v jakém typu prohlížeče zvolený obsah otevře. Vzhledem k faktu, že se objevují nové formáty a nynější postupně zastarávají (např. staré texty v T 602), je nutné pro dlouhodobou archivaci provádět konverzi obsahu do archivních formátů[42], což zajistí integritu a přečtení dokumentů i v budoucnu. Pro firemní procesy může být tato vlastnost nápomocná při identifikaci finálního dokumentu, který může mít jinou podobu než jeho rozpracované varianty (rozpracované varianty mohou mít podobu WORD dokumentu, kdežto finální dokument může být uložen již ve fomátu PDF či ve formě TIFF). Na základě toho je možné nastavit pravidelné automatické odmazávání variant dokumentu, nebo větší granualitu na úrovni přístupu k dokumentu (k Word dokumentu

[38] ISBN je číselný kód, který jednoznačně identifikuje knihy. Zkratka pochází z anglických slov International Standard Book Number (mezinárodní standardní č íslo knihy).
[39] DOI jsou počáteční písmena z anglických termínů Digital Object Identifier (identifikátor digitálního objektu), který má opět za cíl jednoznačne idenfitikovat objekt.
[40] Čárový kód je v dnešní době postupně nahrazován QR kódem, který umožňuje nést větší množství dat než je tomu u klasického 2D kódu, což umožňuje snazší vyčítání dat při snímání dokumentů. Zkratka pochází z angl. slov Quick Response (rychlá odpověď). V dnešní době se mezi odborníky vede diskuze, zda tento prvek není již zastaralý a není-li lepší ho při implementaci technologie ECM zajišťující digitalizaci dokumentu zcela vynechat (ekvivalenty v aj. Imaging, Image Management apod.).
[41] Nejde o normy v klasickém smyslu, jsou spíše na úrovni doporučení.
[42] Archivními formáty jsou PDF/A-1a pro textově statické, obrazové a kombinované dokumenty; formát TIFF pro statické obrázkové dokumenty, formát PNG pro bezeztrátovou kompresi rastrové grafiky. Specifikováno v normách ISO 19005-1: 2005, ISO 19005-2: 2011 a ISO/IEC 15948:2004.

mohou mít přístup pouze obchodníci, kdežto jeho finální podobu již mohou vidět například i jiní zaměstnanci).

Metaúdaj **název** většinou obsahuje krátký a výstižný popis, co dokument obsahuje (např. Smlouva o díle s firmou ABC, Předávací protokol k projektu Ahoj,…). Někdy se můžeme setkat, zvláště při ukládání dokumentů na sdílené disky, s jeho formalizovanou podobou (např. Smlouva_o_dilo_ABC_ver1,…).

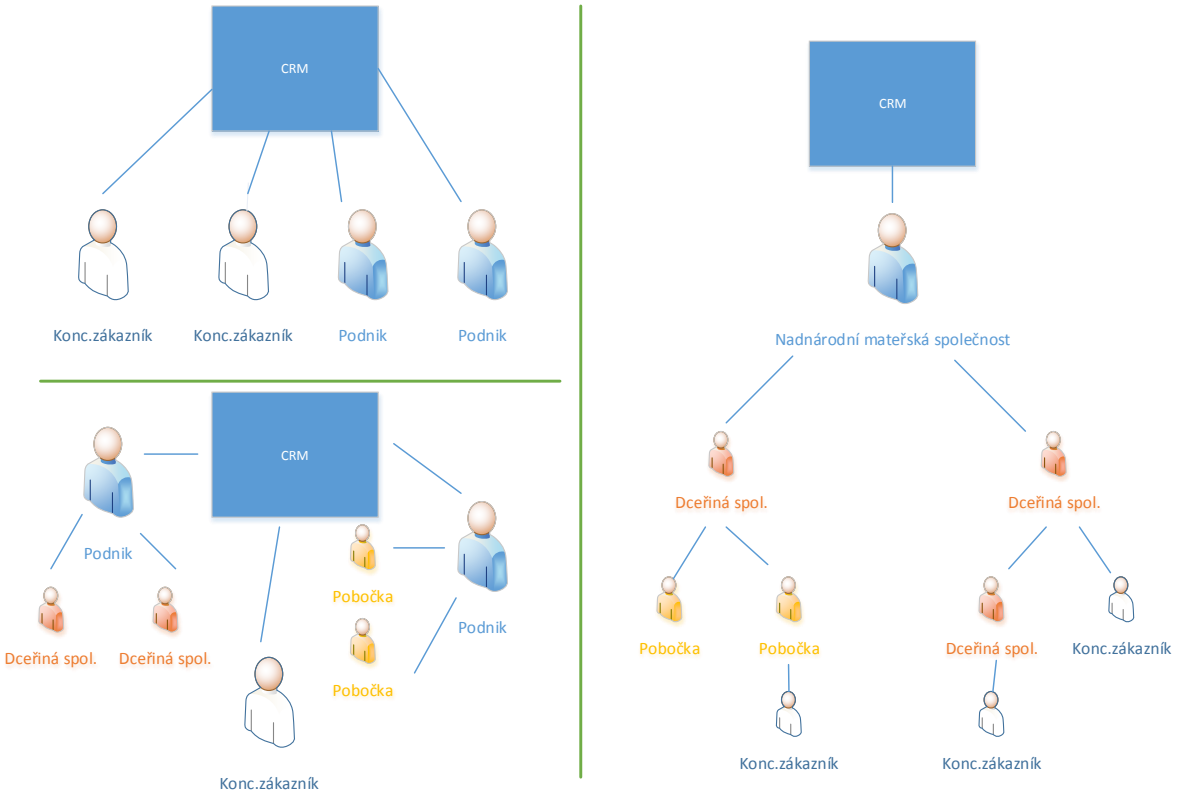
Metaúdaj **autor** patří jak do systémových, tak klientských metadat. Je to z toho důvodu, že na něj existují dva pohledy. Pokud se na atribut autor díváme z pohledu systému, pak je pro něj autorem uživatel, který dokument do systému vložil[43]. Což nemusí být stejný subjekt z pohledu intelektuálního vlastnictví dokumentu (není to subjekt, který má na dokument intelektuální právo). V případě potřeby, pro zajištění obou pohledů, je nutné vytvořit tento metaúdaj ještě jednou.

Poslední obecný metaúdaj můžeme označit **jazyk**, který specifikuje, v jakém jazyce je obsah/dokument napsán. Může být vyplňován automaticky systémem na základě analýzy textu, nebo je při jeho uládání vyplněn uživatelem. Při jeho identifikaci většinou nedochází k chybám, nicméně u tohoto pole můžeme narazit na problém, pokud je dokument psán ve více jazycích (některé ECM systémy neumožňují vložit do pole více než jen jednu hodnotu). Zde bychom mohli použít pravidlo, že pro dokument bude vybrán ten jazyk, který převažuje, nicméně toto pravidlo nám u některých dokumentů nepomůže (např. u vícejazyčných smluv či obchodních nabídek, kde je zastoupení jazyků stejnoměrné).

Mezi klientské metaúdaje, které jsou specifické pro zákaznikcé dokumenty, patří **ID zákazníka**. Při identifikaci zákazníka je vhodnější využívat jeho ID z CRM[44] systému, neboť obchodní název se může v čase měnit, kdežto změny ID v CRM nejsou časté (i když vyjímky existují i zde). Na první pohled není zřejmé, že by s hodnotou tohoto metaúdaje mohl být problém. Opak je pravdou, zvláště u velkých korporací či podniků se specifickou službou/produktem se setkáváme se složitou zákaznickou strukturou, kdy se v rámci CRM systémů setkáváme na různé úrovni, a životního cyklu zákazníka s různými ID, které daného zákazníka specifikují.

[43] Je to subjekt, neboť autorem nemusí být pouze fyzická osoba, ale i jiný systém, který při přihlašování do systému prošel jeho mechanizmy autentifikace a autorizace.
[44] CRM je zkratka z Customer Relationship Management a označují se jí systémy, které řídí vztahy se zákazníky.

Z pohledu životního cyklu klienta se můžeme setkat s případem, že potenciální klient má jiné ID, než ho bude mít jako aktivní klient, a zcela jiné ho bude mít neaktivní klient[45]. V případě zákaznické hierarchie se v rámci organizací potkáváme s různým větvením, které odráží potřeby organizace a je ovlivněno typy zákazníků, které organizace má (jiné členění bude, pokud se firma pohybuje v maloobchodní síti a většina jejich zákazníků jsou koncoví spotřebitelé, jiné, pokud většina jejich zákazníků jsou organizace, jiné, pokud jde o malé podniky, a ještě jiné, jedná-li se o mezinárodní korporace). Příklady zákaznických hierarchií je na obrázku 11.

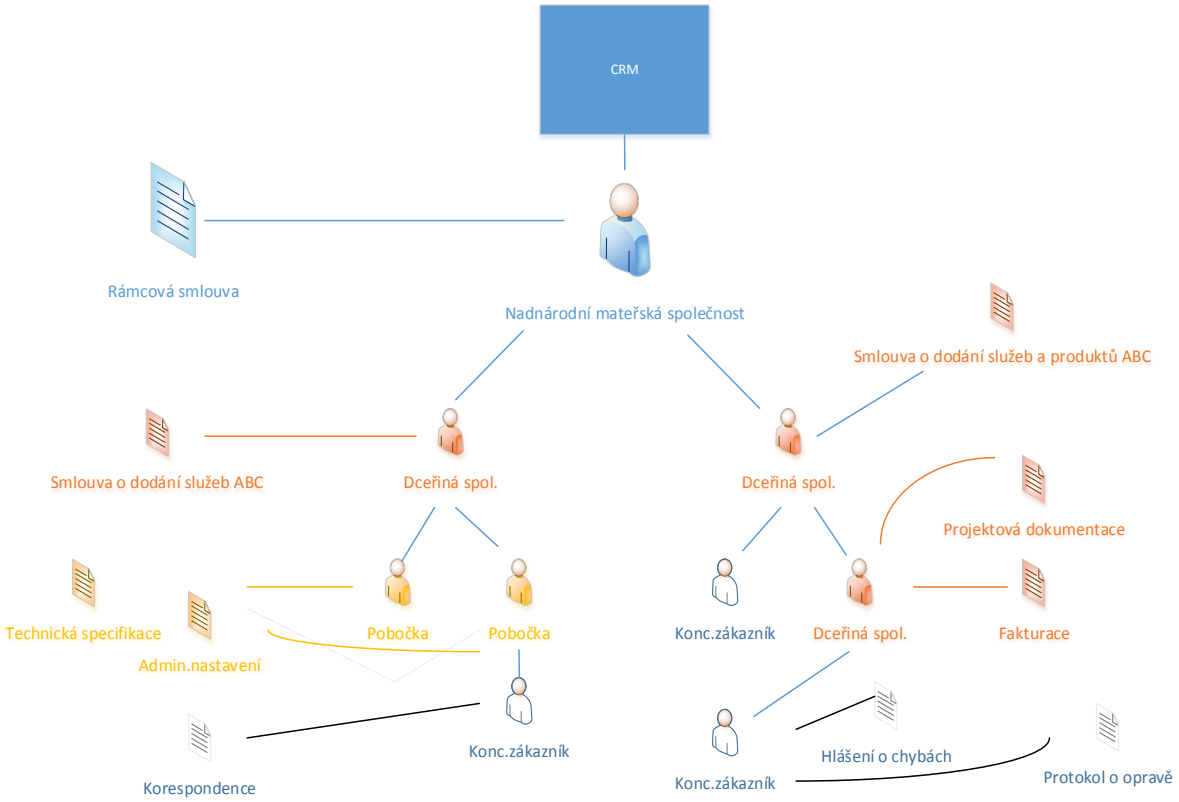


Obrázek 11: Příklady zákaznických hierarchií (zdroj: autorka)

U jednoduché zákaznické struktury s dokumenty není problém, protože jsou všechny přirazeny k danému zákazníkovi. U složitějších, jako je tomu na obrázku vlevo dole, se s menšími problémy, který typ dokumentu je pod jakou hierarchií kategorizován, se již setkat můžeme. Záleží na společnosti, zda bude chtít mít všechny dokumenty navázané k nejvyšší organizační struktuře, či je začne ukládat na nižší hierarchickou úroveň. Pokud zvolí druhou možnost, bude nutné vytvořit více metaúdajů typu zákazník i v ECM.

Příklad, jak mohou být dokumenty navázány na hierarchickou strukturu, je uveden na obrázku 12.

[45] I když změna může být pouze provedena formou, že se přidá prefix nebo sufix k neměnnému středu.



Obrázek 12: Provázanost dokumentů na zákaznickou hierarchii (zdroj: autorka)

S touto provázaností musí počítat i procesy, které s dokumenty pracují. Je tím myšleno, že oddělení, které má na starosti vymáhací proces, nemůže chtít vyhledat fakturu pod identifikací koncového zákazníka nebo, že uživatelé, kteří mají na starosti péči o zákazníka, nenaleznou hlášené chyby pod identifikací společnosti apod. V rámci této struktury je nutné procesně ošetřit změny zákaznických entit (koncový uživatel může být navázán pod společnost ABC, později pod CDE, dceřiná firma zanikne a její následovnicí bude zcela jiná firma, která nespadá pod stávající mateřskou společnost). V rámci těchto procesů je nutné ošetřit i následnou práci s dokumenty, budou-li dostupné pod původním objektem, bude-li nutné je připojit pod nové zákaznické entity. Tyto přesuny mají následně dopad i na způsob skartace. Při rozhodování, jaké zákaznické identifikátory použít, je nutné mít nazřeteli, že ECM není CRM systém a není správné do něj přenášet celou složitou zákaznickou strukturu.

Zákaznické dokumenty jsou jedním ze zdrojů tzv. 360stupňového pohledu na zákazníka. Cílem pohledu je získat o zákazníkovi kompletní přehled aktivit spojených se společností. Ten je získáván z různých zdrojů, jejichž prostřednictvím může zákazník společnost kontaktovat.

Výše je zmíněn proces skatace, pro tyto účely by měl být v systému nastaven minimálně metaúdaj **skartační znak**, který může nabývat hodnot A (archiv), V (výběr)[46], S (skart), a **skartační lhůta** (doba, při jejímž průběhu slouží dokument provozním potřebám podniku, po jejím uplynutí s ním bude naloženo dle skartačního znaku). Nejspíše se setkáme ještě s metaúdaji zajišťujícími počátek plynutí skartační doby a metadaty, které lokalizují v archivu papírové originály (patří sem ID fyzického dokumentu, ID archivní krabice).

Dalším důležitým metaúdajem je **datum výmazu**, i to patří mezi obojaká metadata. Existuje jako systémový metaúdaj, nebo ho při ukládání vyplňuje uživatel či systém (před samotným uložením může být v procesu uvedeno, že daný dokument je možné smazat za XSZ měsíců), je tedy nutné ho v systému vytvořit. Specifikuje, od kdy je možné dokument bez problémů ze systému odstranit. Na něj bývají navázány procesy, které by měly výmaz dokumentu zajistit.

K posledním základním metaúdajům patří **úroveň utajení**. Používá se pro označení, do jaké míry dokument obsahuje citlivé informace. V příloze č. 4 je uvedena tabulka, která porovnává, jakých hodnot v rámci norem nabývá, je zde patrné, že panuje shoda. Pokud norma pro své účely potřebuje oblast více vymezit, používá k tomu jiné doprovodné metaúdaje. Mezi uvedenými normami je specifická klasifikace DIN 66399, která označuje dokumenty třemi číselnými úrovněmi. Organizace může využít jednu z variant uvedených v příloze, anebo si vytvoří své vlastní třídění. Domnívám se, že většině firem by mohlo vyhovovat třídění uvené v tabulce níže.

[46] Tato hodnota definuje, že je nutné v rámci skartačního řízení posoudit, nabývá-li dokument hodnoty A (je tedy archiválií a bude trvale uložen v archivu) nebo S (nemá trvalou hodnotu a je možné ho zničit).

Tabulka 11: Návrh zákaznické klasifikace utajovaných informací (zdroj: autorka)

Název	Kód	Popis
Veřejné	00	obsah/dokumenty běžně dostupné na webové prezentaci firmy, nabídky,…
Poloveřejné	05	Obsah/dokumenty dostupné přes portály (zákaznický, dodavatelský, partnerský)
Interní	10	Obsah/dokumenty, které jsou dostupné většině zaměstnanců (obchodní nabídky, obchodní ceníky, …)
Důvěrné	40	Obsah/dokumenty, které jsou již vyhrazeny pouze určité skupině (mohou sem patřit projektové dokumenty, reporty, speciální obchodní nabídky,…)
Tajné	50	Obsah/dokumenty, které jsou určené managementu (obchodní, marketingové plány, ...)
Přísně tajné	100	Obsah/dokumenty, které mají informace, jež jsou pro společnost kritické (strategické dokumenty, plány nových produktů/služeb, …)
Utajované podle zákona	101	Vymezené zákonem 412/2005 Sb., nutné před skartačním řízením odtajnit

V tabulce č. 9 je uveden i příklad kódu, pokud by se v implementační fázi rozhodlo, že některé seznamy budou i pro uživatele obsahovat pouze kódy, ne popis.

Existuje sada metadat, která se v určitých situacích též váží k zákaznickým dokumentům a jejich využití záleží na situaci a podniku. Těmito metaúdaji jsou:

- popis/abstrakt – u většiny zákaznických dokumentů je plně postačující název dokumentu/obsahu. Toto pole bývá uživateli používáno při ukládání verzí, specifikují v něm, k jakým změnám v dokumentu došlo;
- označení produktu/služby;
- číslo obchodního případu;
- označení projektu;
- oblast pokrytí – pole může obsahovat informace týkající se zeměpisného určení místa, licencí/autorských práv k dílu, legislativní územní působnost;
- jiný identifikátor – tento druh údaje neodkazuje na stejný dokument dostupný jinde, ale věcně k sobě váže určité dokumenty. Mezi takové identifikátory patří číslo jednací nebo referenční číslo nebo jde o případy, kdy v systému neexistuje dokument jako celek (nicméně klíčoví uživatelé na něj jako na celek nahlížejí), ale je fyzicky rozdělen na více částí, pak se tyto části mohou na sebe odkazovat

shodným identifikátorem (například smlouva není uložena jako celek, ale pod jedním systémovým ID nalezneme tělo dokumentu, pod jiným přílohy dokumentu, …).

Pro potřeby běžného provozu je dobré pamatovat i na metaúdaje, které zaručí snažší identifikaci chyby a její opravu, nebo mohou sloužit k reportovacím účelům. Mezi tyto metaúdaje například patří číslo importní dávky, název aplikace, která import provedla (nutné zvláště v případech, kdy import neprovádí sama, ale zprostředkovatel, který provádí import i za jiné systémy nebo se pro systémové importy přistupuje do systému pod jedním systémovým účtem). Tento typ metadat obsahuje standard T40 (AIIM T40 1991), seznam metadata je k dispozici v příloze č. 5.

Poslední sadou metaúdajů jsou ty, které se používají v systému pro přidělování oprávnění týkajících se práce s dokumenty (možnost vytvoření, editace, verzování, výmaz,…). Způsoby a možnosti se odvíjí od typu použitého softwaru, jejich správné nastavení je pro práci se zákaznickými dokumenty klíčové. Neboť nebude-li provedeno dobře, pak citlivé informace o zákaznících uvidí nepovolané osoby (což nemusí být pouze zaměstnanci, ale i dodavatel nebo jiný zákazník), může dojít k jejich nechtěné editaci či dokonce odstranění.

V teoretické části (kapitola 4.9.2) je představen standard Dublin Core, který je v různých projektech používán jako zdroj metadat. Vzhledem k jeho rozšířenosti jsem se rozhodla srovnat základní metadata z DC s navrženými základními metadaty pro třídění zákaznických dokumentů (tabulka 12):

Tabulka 12: Párování metadat mezi Dublin Core a navrženými zákaznickými metadaty (zdroj: autorka)

Dublin Core	Zákaznická metadata
Název	Název
Tvůrce	Autor
Předmět	Druh dokumentu, třídění
Popis	Popis / Abstrakt
Vydavatel	-
Přispěvatel	-
Datum	Datum vložení, datum úpravy
Typ	
Formát	Formát
Identifikátor	Interní identifikátor, Externí identifikátor
Zdroj	-
Jazyk	Jazyk
Vztah	-
Pokrytí	-
Práva	-
-	Číslo zákazníka
-	Skartační znak
-	Datum výmazu
-	Fixní doba
-	Zadržení
-	Oprávnění
-	Úroveň utajení

Z tabulky vyplývá, že navržená metadata pro zákaznické dokumenty jsou bez problému párovatelná na DC.

6.3.3 Tvorba předmětového selekčního jazyka pro zákaznické dokumenty

Pro účely této kapitoly byl použít open-source program Protégé ve verzi 5.1.0 a 4.3.0 s instalovanými plug-iny SKOS, SOVA, OWL Viz, OntoGraf a VoWL. Mezi výhody používaného softwaru patří možnost uložení třídění do RFT/XML datového formátu, který může být následně bez velké námahy upraven a používán v ECM technologii. Dále podpora OWL jazyka (blíže teoretická část této práce – kapitola 4.9.7).

Navržené přemětové selekční jazyky pro zákaznické dokumenty jsou konstruovány na obecné úrovni, aby našly uplatnění v různých typech organizací. Z toho důvodu není navržen

pouze jeden typ selekčního jazyka, ale více. Používání stejné sady termínů zároveň umožňuje porovnání mezi nimi, což dává organizaci možnost vybrat si a následně doupravit typ vhodný pro ni.

Před tím, než budou představeny jednotlivé předmětové selekční jazyky, pozastavíme se nad terminologií formát, typ, druh dokumentu. Jak je patrné z tabulky 13, kde je provedeno mapování metadat jednotlivých standardů, pro tyto termíny neexistuje jednotná shoda, jak se má metaúdaj jmenovat a co je jeho obsahem.[47]

Problematiku můžeme rozdělit na formát, druh dokumentu a obsah (věcné třídění). V dokumentovém formátu je uveden odkaz na fyzické ztvárnění, kdežto v druhu dokumentu spíše na typ vyjádření (formy). Pod obsahem je zmíněna hlavní myšlenka, význam, který je prostřednictvím dokumentu komunikován. Níže v tabulce uvádím, jak jsou jednotlivé pojmy nazývány v různých normách, názvy jsou ponechány v originálním znění, neboť se domnívám, že při překladu by mohlo dojít k posunutí významu.

Tabulka 13: Srovnání standardů z pohledu formy a druhu dokumentu (zdroj: autorka)

Název standardu	Formát dokumentu	Druh dokumentu
AIIM – TR40	Input media format Output media format	Document type
RDA	Media terms	Content terms
MARC	Physical Description Fixed Field - 007 Physical Description - 300	Genre Term List
OCLC WorldCat.org	Format/Facet	Format/Facet
ISBD – Area 0	Media Type Term	Content Type Term
AGRkMS	-	Document Forms Scheme
Library of Congres	-	General Terms Hierarchies
Dublin Core	Format	Type

[47] Roztříštěnost názvu termínu ve standardech je dáno jejich zaměřením a složitostí (některý standard si vystačí pouze s jedním metaúdajem, jiný jich má implementováno několik a jde až do takového detailu, že ho zajímá počet stránek, příloh,…).

Projdeme-li všechny vydefinované seznamy, zjistíme, že ve formátu fyzický popis objektu a u druhu dokumentu často dochází k míchání obsahu s formou. Zajímavě to má řešen katalog OCLC WorlCat.org, který používá hierarchické třídění, kdy na první úrovni je druh dokumentu a v druhé úrovni je formát. Jeho třídění je v příloze č. 6.

Při návrhu předmětového selekčního jazyka, jsem se potýkala se stejným problémem – nedařilo se mi formu od obsahu oddělit. Např. pod pojmem smlouva se definuje jak samotný obsah dokumentu, tak jeho forma (která bude mít uvnitř podniku ustálenou podobu-strukturu).

Metaúdaj formát dokumentu může být v rámci celé klasifikace vyčleněn na roveň obecným metadatům, která jsou specifikována výše. Může také být součástí věcného třídění, nebo není specifikován vůbec, neboť jeho používání by přinášelo organizaci minimální informační hodnotu (podnik se setkává s omezenými typy formátů). V příloze č. 7 uvádím seznamy formátů uvedené v různých standardech.

Seznamy, které se týkají druhu dokumentu, představené v tabulce č. 13, nejsou v příloze zastoupeny, neboť v některých případech (např. Kongresová knihovna) se jedná o několikastránkový dokument. Jedinou vyjímku tvoří číselník organizace *State Archives and Records a jeho specifikace druhu dokumentu* (State Archives and Records NSW). Důvod je ten, že v seznamu jsou obsažené typy dokumentů, se kterými se v rámci třídění podnikového obsahu (nejen zákaznického) často setkáváme a v některých případech z něho navržené věcné třídění vychází. Je k dispozici v příloze č. 8.

Při návrhu věcného selekčního jazyka je nutné dodržet principy a zásady popsané výše. Jen ve zkratce: je nutné neustále si uvědomovat, pro koho je klasifikace tvořena, kdo bude její konzument, využít či se inspirovat již stávajícími tříděním, vyhnout se hononymním a synonymním termínům. Třdění by nemělo mít více jak 6 – 12 hlavních tříd a 2 – 3 úrovně podtříd.

Jak již bylo zmíněno, zákaznické dokumenty můžeme rozdělit do oblastí, ze kterých budeme vycházet při tvorbě předmětového selekčního jazyka:[48]

- strategické a marketingové dokumenty – neobsahují přímo odkaz na reálného zákazníka. Jde o dokumenty, které obsahují marketingové a strategické cíle, jaké typy zákazníků by firma chtěla mít, v jakém segmentu, jakým způsobem toho docílí a jaký typ produktu/služby jim chce nabízet;

[48] Seznam je možné rozdělit do dvou kategorií – na dokumenty o zákaznících a dokumenty zákazníků.

- obchodní dokumenty – dokumenty dotýkající se obchodních procesů (nabídky, prezentace produktů/služeb, ceníky,…);
- smluvní dokumenty – dokumenty, na jejichž základě vznikl smluvní vztah (rámcová smlouva, smlouva na produkt/službu, závazná objednávka,…);
- daňové dokumenty – typy zákaznického vyúčtování (faktura, daňový doklad, vrubopis, dobropis,…);
- provozní dokumenty[49] – toto členění se odvíjí od podnikání dané společnosti (projektová dokumentace, technická dokumentace, protokoly o předání, administrátorská nastavení, reporty, návody, zápisy ze schůzek,…);
- komunikace s klientem – může mít různé podoby – od klasické papírové korespondence, přes e-mailovou korespondenci, datové zprávy, zprávy ze sociální sítě. Patří sem například různé typy poukazů, memoranda, newslettery, logy, přepisy/zápisy z telefonického rozhovoru;
- dokumenty související s vymáhacím procesem – ty obsahy, které slouží jako důkaz v rámci vymáhacího pohledávkového procesu.

6.3.3.1 Řízený seznam termínů

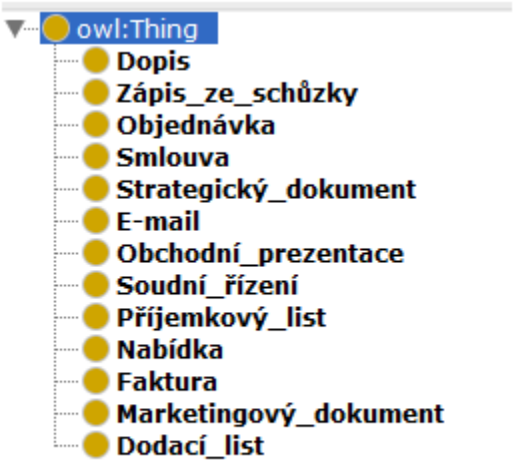
Patří mezi nejjednodušší formu třídění, jde o množinu termínů, které mají základní vztahy mezi sebou (podobné vlastnosti, účel pro které byly vytvořeny – nákupní seznam). Důvod, proč byly tyto termíny seskupeny, by měl být bez problému rozeznatelný.

K výhodám patří, že uživatelé nemají problém s porozuměním termínům a dokáží je použít při vyhledávání či ukládání.

Jeho nevýhody jsou: seznam by neměl obsahovat velké množství termínů (ideálně 12 - 15)[50]. Jejich používání v systému práci zpomaluje (uživatel musí hledaný termín najít a označit z nabídkového seznamu metaúdaje). Pokud se včas nepodchytí potřeba rozšíření či přechod na jiný typ třídění, pak případně následná reindexace je náročná.

[49] Termín provozní dokumenty byl vybrán, aby vymezil ty dokumenty, které vznikají v rámci běžného provozu a dotýkají se zákazníka. Je to oblast, která je pro každý podnik specifická a liší se i co do rozsahu (v některém podniku se s ní skoro nesetkáme, jinde existuje různá typová škála).
[50] I když jsou samozřejmě výjimky, např. seznamy kódující zemí či použitý jazyk.

Tento typ je nečastěji používán v malých firmách, většinou je vyvinut interně (intuitivně) bez hluboké analýzy a s nízkým požadavkem na zatížení lidského zdroje. Příklad, jak takový seznam může vypadat, je vidět na obrázku 13.[51]



Obrázek 13: Zákaznické třídění - typ řízený seznam termínů (zdroj: autorka)

Některé položky mohou být použity i vícekrát, bude-li podnik chtít rozlišovat i formu dokumentu (např. papírová faktura, elektronická faktura, papírová smlouva, elektronická smlouva,…).

6.3.3.2 Taxonomie

Termín taxonomie, jak již bylo v teoretické části vysvětleno (kapitola 3.3.1), je používán ve dvou významech – v širším pojetí, kdy jsou pod tímto termínem chápány věcné selekční jazyky v podnikovém prostředí, a užším pojetí, kdy jde o typ věcného selekčního jazyka. Taxonomie je na cestě mezi řízeným seznamem a tezaurem. Dá se říci, že je to seznam, který je již dělen hierarchicky. Mezi termíny existují nadřazené, podřazené a souřadné vztahy.

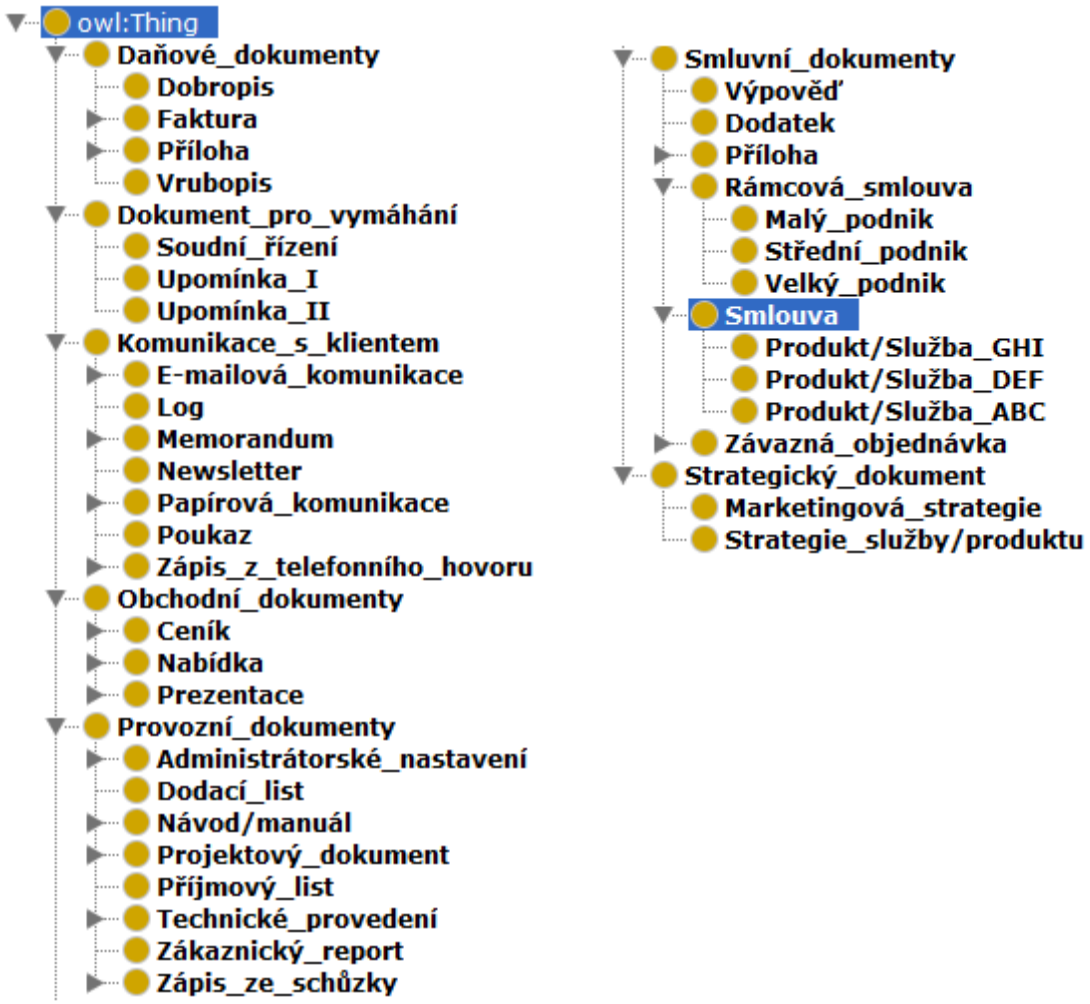
Struktura taxonomie podněcuje spíš k prohlížení sbírky dokumentů než k vyhledávání. K výhodám taxonomie patří uživatelsky přívětivé a snadno pochopitelné třídění[52]. Také to, že obsahuje ty termíny, které nyní firma využívá (to ale může být za určité situace bráno i jako nevýhoda).

[51] Podtržítka v názvu třídění jsou použita dle jmenné konvence Protégé softwaru, v tomto formátu by se neměla objevit na produkčním prostředí.
[52] Nicméně i zde narážíme na to, že každý člověk klasifikuje universum jinak, a tak pro někoho může být navržené třídění nepřehledné nebo těžko pochopitelné (pro dělení do jednotlivých podskupin by použil jiný princip).

K jejím nevýhodám patří omezení, co se týká velikosti (hierarchický strom by neměl být extrémně košatý – jak do šířky použitých termínů, tak do hloubky používaných úrovní). V systémech je většinou koncipován tak, že je možné dokument zatřídit pouze pod jedním termínem. To vede ke kompromisům, kam uložit dokument, který je zároveň nabídkou a objednávkou v jednom. Uživatel je nucen procházet více složkami a zpomaluje se tak proces vyhledávání.

Tento typ třídění opět nalezneme v menších či středních organizacích, ale může se vyskytnout i v korporacích, kde bude součástí systému, který eviduje omezenou část zákaznických dokumentů (např. právní oddělení ve svém systému bude mít evidovány veškeré zákaznické smlouvy). Jako je tomu u řízeném seznamu, tak i tady bývá vyvinut interně – většinou na základě toho, že seznam začal být natolik široký, že se v něm špatně orientovalo, a proto bylo nutné rozdělit ho do logických skupin.

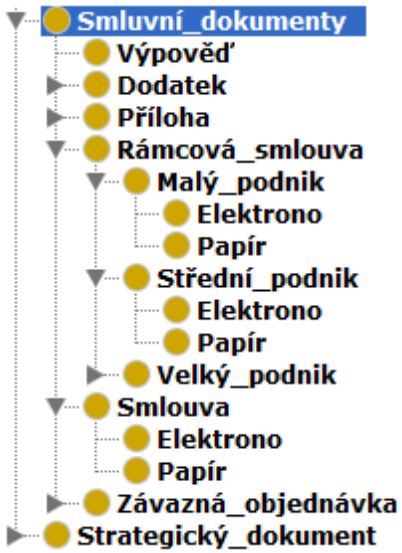
Níže je zobrazen návrh zákaznického třídění. Jak je z obrázku patrné, některé termíny z předchozího seznamu byly použity ve stejném tvaru, jiné byly upraveny. Na třetí úroveň byl odkaz na produkt/službu, ke které se dokumenty váží. Otevřený seznam je zobrazen pod termínem služba. Stejný seznam je připojen i k termínům, které před sebou mají vodorovnou šedivou šipku, pro lepší orientaci, ale zůstaly nerozbaleny. Produkt/službu takto vážeme v případě, že nám systém neumožní tuto informaci vložit do jiného metaúdaje (např. je toto třídění realizováno prostřednictví sdílených disků, kde se využívají složky, které jsou v sobě vnořeny). Otázka k diskuzi je, zda popis služby/produktu při třídění využívat či ne. Odvíjí se od situace v podniku, pokud se produkty/služby často mění a je jich hodně, pak je vhodné tento typ informace neuvádět, nebo používat oblast, ne přímo typ (např. pojištění, pečivo, sportovní nářadí, zájezdy, ...).



Obrázek 14: Taxonomie zákaznických dokumentů (zdroj: autorka)

Další variantou by mohlo být, že na třetí úrovni bude místo produktu/služby uveden rok (velmi často se s časovým údajem setkáme u smluvní dokumentace).

Navržené třídění kombinuje věcný pohled s jinými charakteristikami dokumentu. Tento přístup je v taxonomii častý. Obrázek 15 byl rozšířen o formu dokumentu. I zde se budeme potýkat s problémem – někomu by u dimenze formy vyhovalo, kdyby byla na vyšší úrovni.



Obrázek 15: Taxonomie zákazických dokumentů s formou dokumentu (zdroj: autorka)

6.3.3.3 Fasetová klasifikace

Fasetová klasifikace, jak je v teoretické části uvedeno (kapitola 3.2.1), patří mezi postkoordinovaný selekční jazyk. Vztah mezi termíny vzniká až při formulaci rešeršního dotazu (při vyhledávání). Faseta je metaúdaj, který popisuje jednu vlastnost.

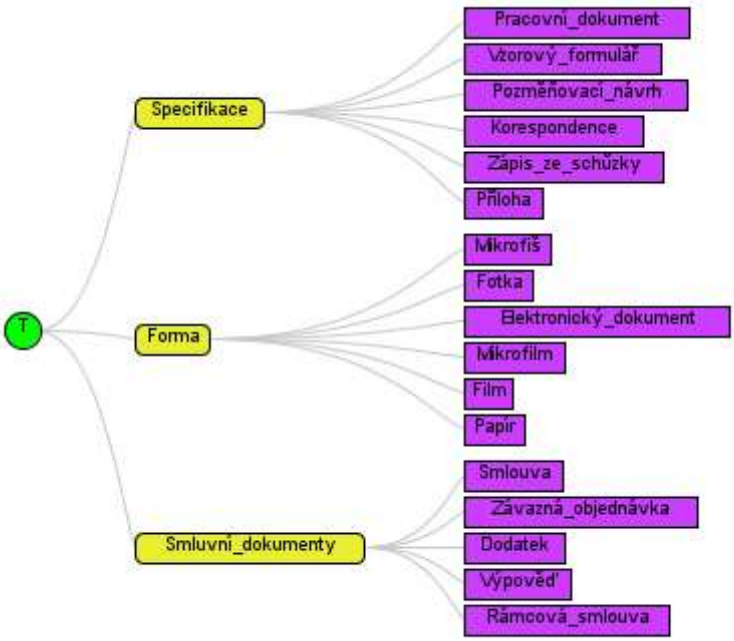
Jde o typ třídění, které podporuje vyhledávání. To, že vztah mezi termíny vzniká v rámci formulace dotazu, umožňuje jasně formulovat dotaz za pomoci proximitních či jiných operátorů (uživatel není odkázán na vztahy, které jsou dány předem samotným tříděním – např. taxonomií). Dále je jeho kladem překonání omezení, které plyne z toho, že každý uživatel při vytváření hierarchického třídění použije jiný princip dělení – dokumenty můžeme hledat jak pod produktem/službou, tak pod dokumentovým typem či jeho formou. Použití funkce vyhledávání je rychlejší než prohlížení s hierarchickou strukturou, která může mít omezení počtu objektů vnořených v rámci jedné úrovně, jak je dále v implementační části rozebíráno. Fasetové třídění je nápomocné i při filtrování či třídění výsledného seznamu položek.[53]

K nevýhodám patří, že toto třídění neumožňuje prohlížení, takže uživatel musí znát obsah a způsob klasifikace daného úložiště. Je tedy méně intuitivní, ne každý uživatel má s formulací rešeršního dotazu zkušenosti.

Z pohledu jeho použitelnosti záleží na zvolené ECM technologii a uživatelích (preferují-li či ne vyhledávání). Můžeme se s tímto typem třídění setkat jak u malých,

[53] Úspěšností fasetového třídění pro vyhledávání se zabývá studie (Uddin, 2007).

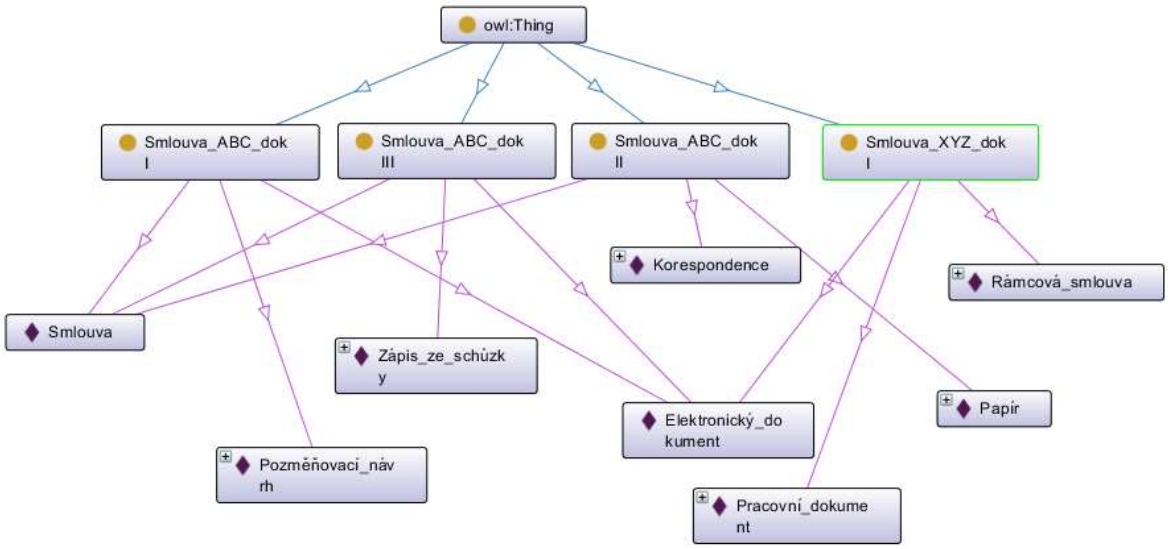
středních, tak i velkých podniků. Fasetová klasifikace nepatří mezi složité třídění, takže bývá opět vyvíjena uvnitř samotné firmy.



Obrázek 16: Příklad fasetového zákaznického tříděně – smluvní dokumenty (zdroj: autorka)

Fasetové zákaznické třídění je na obrázku 16. Bylo aplikováno pouze na jednu část, smluvní dokumenty. Většinou se fasetové třídění implementuje tak, že jeden z metaúdajů je hlavní a ostatní ho dále rozvíjejí (v ideálním případě uživatel u druhého pole dostane v nabídce už jen ty hodnoty, které jsou relevantní k vybrané hodnotě z prvního pole – např. k hodnotě závazná objednávka nebude možné vybrat pozměňovací návrh). Pokud bychom celé zákaznické třídění dali pouze do dvou faset, nabídka hodnot v těchto seznamech by byla nepřehledná a uživatel by se těžko orientoval, jaké hodnoty k sobě přiřadit.

Na rozdíl od hierarchie, pokud to systém umožní, je možné hodnotu z jednoho pole kombinovat s hodnotou z jiného pole. Příklady předmětové katalogizace zákaznických dokumentů používajících fasetové třídění je na obrázku 17.



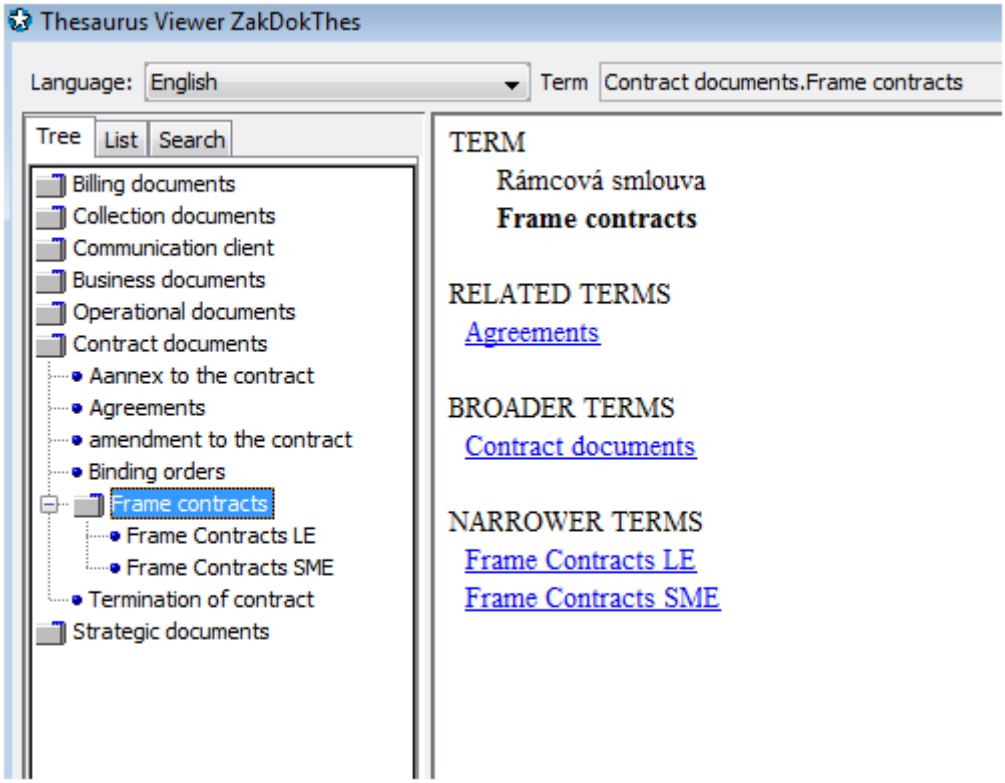
Obrázek 17: Příklady předmětové kategorizace pro zákaznické smlouvy (zdroj: autorka)

6.3.3.4 Tezaurus

Tezaurus patří mezi deskriptorové selekční jazyky, bližší specifikace je v teoretické části této práce (kapitola 3.2.3). Může se zdát, že je stejný jako taxonomie, neboť někdy je možné jím procházet stejně jako hierarchickým tříděním. Tento typ zobrazení je umožněn díky tomu, že deskriptory[54] také mezi sebou mají nadřazený (BT), podřazený (NT) vztah. Ale kromě jiného zde nalezneme odkaz na nepreferovaný termín a s tím související ekvivalentní vztahy. Používaný termín USE odkazuje z nedeskriptoru na deskriptor, kdežto US opačně, tzn. z deskriptoru na nedeskriptor. Vztah asociace bývá vyjádřen zkratkou RT. Zkratka SN se používá pro vysvětlující poznámky.

Ke kladům tezauru patří přesné vyjádření obsahu dokumentu a přesná konstrukce selekčního dotazu, což je umožňěno na základě vztahů mezi temrmíny, které zachycuje. Další výhodou je možnost procházet jím v hierarchické podobě i prohlížet v rejstříkové podobě. V neposlední řadě i to, že umožňuje tvorbu s více jazykovými variantami. To je patrné z obrázku 18, kde je vytvořený zákaznický tezaurus zobrazen v anglickém jazyce.

[54] Tezaurus obsahuje preferované termíny, tzv. Deskriptory, a nepreferované termíny, tzv. nedeskriptory.



Obrázek 18: Zákaznický tezaurus v anglickém jazyce (zdroj: autorka)

K jeho nevýhodám patří hlavně to, že je složitý na konstrukci. Není snadné ho ve stávajících ECM technologiích používat (out-of-box řešení s tímto typem klasiface nepočítají – viz níže)[55]. Pro některé uživatele může být koncept tezauru nepochopitelný, a z toho důvodu se jim s ním může špatně pracovat.

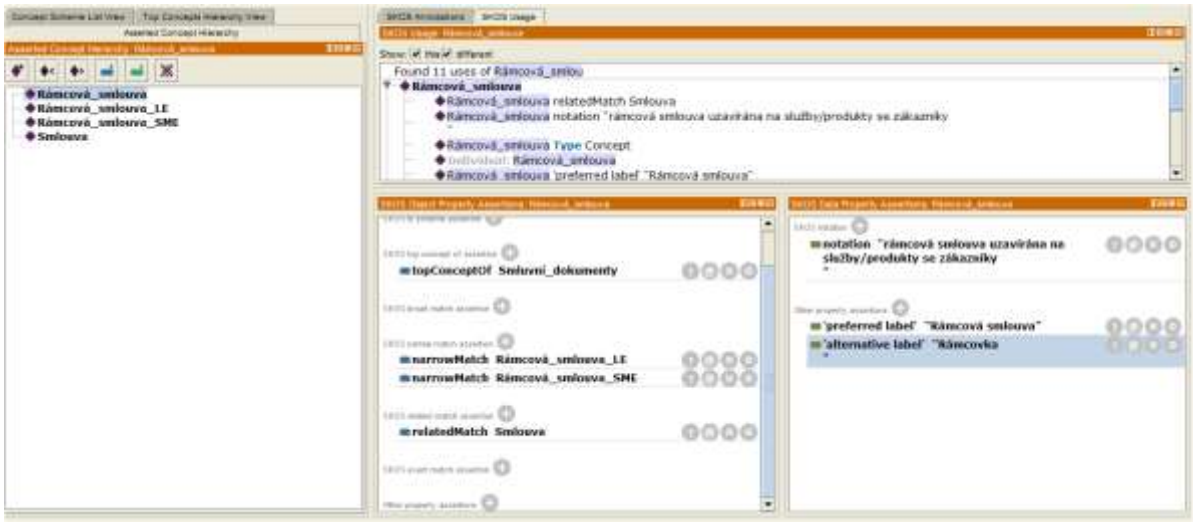
S tezaurem se setkáváme ve velkých korporacích, většinou jsou vyvinuty s externí pomocí, neboť, jak již bylo řečeno, jejich konstrukce není snadná, což je podloženo i existencí několika norem (kapitola 3.2.3). Není vždy nutné tezaurus vyvíjet, ale mohou být využity stávající oborové tezaury, které se dle potřeby korporací upraví. Zajímavostí je, že se v praxi nesetkáme s tím, že by firma poptávala tezaurus, spíše poptává taxonomii, ale při bližší analýze se zjistí, že je potřeba tezaurus. Nejspíše proto se lidé, kteří se zabívají tvorbou tezauru, v anglo-americkém prostředí nazývají „taxonomist“.

Při tvorbě tazauru je vhodné mít na paměti několik zásad: nepoužívat široké a obecné termíny, vyhnout se nesoběstačným termínům (např. ostatní, další,…), nepoužívat jeden termín pro dva koncepty (vektor – jednou pod matematikou, jindy pod epidemiologií), používat pouze ty termíny, které budou využity při vyhledávání či klasifikaci. Tezaurus by měl být složen z podstatných jmen v množném čísle, u víceslovného termínu použít běžně

[55] Termín out-of-box je u programů používán pro funkce, které pracují téměř ihned po instalaci a nepotřebují velké úpravy, vývoj či složitou konfiguraci.

používané pořadí slov, doporučuje se používat malá písmena, a pokud je to možné, vyhnout se akronymům.

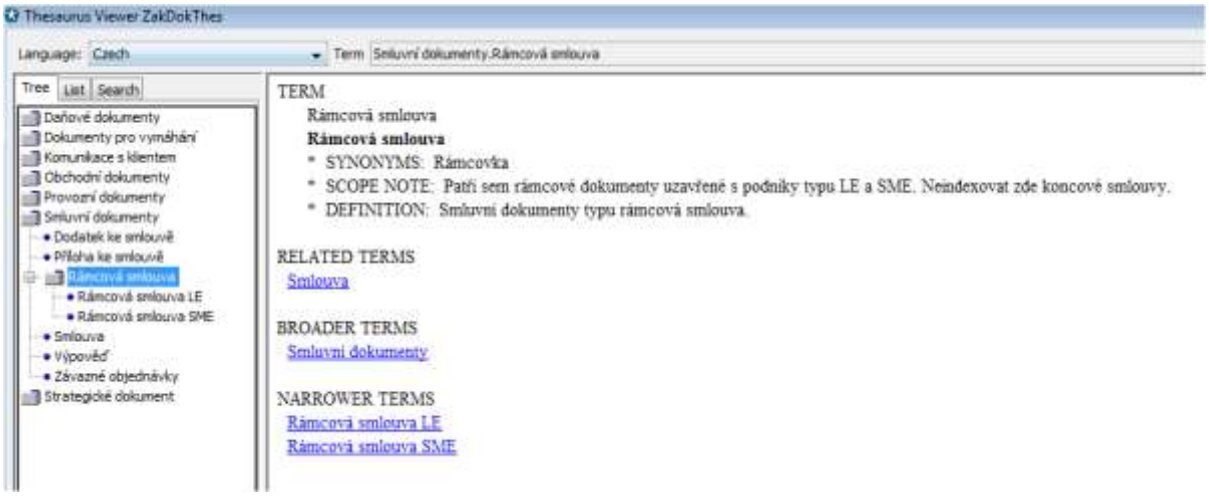
Jsou dvě metody tvorby tezauru: tvorba od shora dolů, anebo zdola nahoru. Pokud tvoříme shora dolů, neboli od obecného ke specifickému, pak na začátku se stanoví hlavní oblasti, v ideálním případě by jejich počet neměl přesáhnout dvacet. Do těchto hlavních oblastí se řadí termíny, které obsahují vztahy popsané výše. Tvoříme-li zdola nahoru, neboli od specifického k obecnému, vytvoříme seznam termínů[56], tyto termíny následně začleňujeme do skupin. Při vytváření je nutné vyhnout se nesprávnému určení nadřazeného/podřazeného vztahu mezi termíny, používání slangu/hantýrky, kterému uživatelé nerozumí, špatnému zařazení termínu v hierarchii (bývá to způsobeno nesprávným pochopením termínu), nedostatečné úrovni (plochá konstrukce) anebo naopak přetížení některých větví termíny. Stručný návod, jak postupovat při tvorbě tezauru, je v literatuře (Craven, 2008).



Obrázek 19: Tezaurus - deskriptor rámcová smlouva v Protegé (zdroj:autorka)

Příklad zákaznického tazauru je zachycen na obrázku 19. Záznam je vytvořen v aplikaci Protégé v pluginu SKOS. Protože se při práci začaly projevovat funkční chyby pluginu a plug-in neumožňuje jinou vizualizaci, byl nakonec tezaurus vytvořen v nástroji ThManager ve verzi 2.0. Na obrázku 20 je zobrazen deskriptor rámcová smlouva v hierarchické podobě.

[56] Tento seznam může být vytvořen na základě termínů, které jsou použity u již zatříděných dokumentů v podniku.



Obrázek 20: Hierarchické zobrazení tezauru (zdroj: autorka)

Mezi výhody tezauru patří možnost jeho zobrazení nejen v hierarchické, ale i v rejstříkové podobě. Způsob zobrazení v rejstříkové podobě se stejným termínem je na obrázku 21.



Obrázek 21: Rejstříková podoba zákaznického tezauru (zdroj: autorka)

6.3.3.5 Ontologie

Posledním typem klasifikace je ontologie. Jak již bylo v teoretické části představeno (kapitola 3.3.2), zcela nepatří mezi předmětové selekční jazyky, ale vzhledem k její vzrůstající využitelnosti jsem se ji rozhodla do seznamu zařadit. Autoři odborných článků ontologii často připodobňují k tezauru.[57]

[57] Osobně tento názor nesdílím, domnívám se, že ontologie je podobná myšlenkovým mapám či znalostnímu grafu firmy Google.

Tezaurus má s ontologií podobné to, že je možné v ní specifikovat neomezené množství vazeb a na rozdíl od tezauru není omezena jen na několik. Množství a typy vazeb záleží na účelu, pro který je ontologie vyvíjena.

Ontologie není určena pro běžné uživatele, ale pro pročítačové zpracování (automatická indexace, sémantický web, klasifikace e-mailů, apod.). K jejím výhodám patří otevřenost, která umožňuje její vytvoření pro široké spektrum potřeb a vícepoužití v rámci organizace. Dále zlepšuje navigaci a vyhledávání. Přispívá ke konzistenci používaných pravidel.

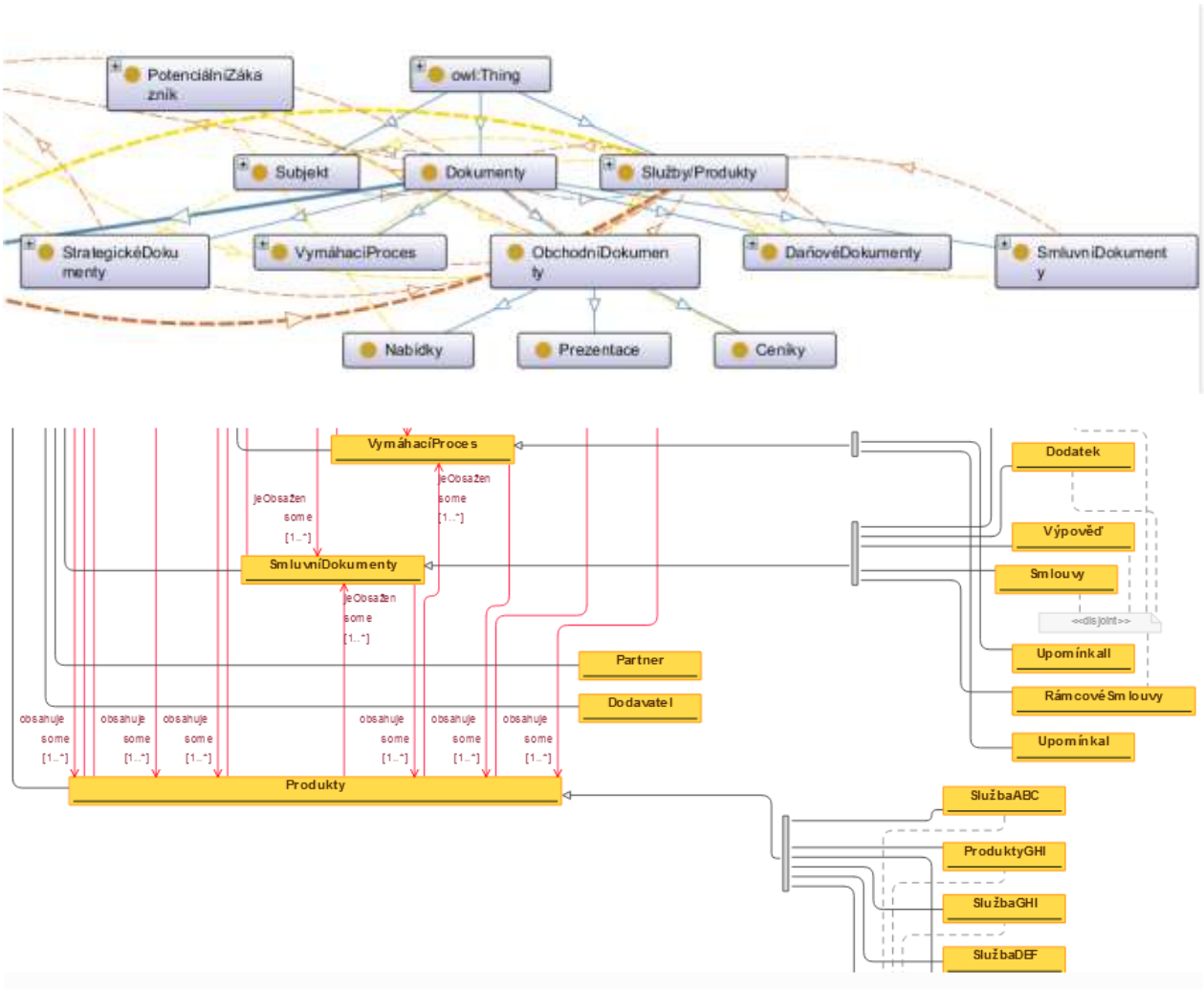
Její hlavní nevýhodou je obtížnost při její tvorbě. Je těžké najít vytvořenou ontologii, která by odpovídala potřebám organizace. Velikost ontologie je nepřímo úměrná její specifičnosti. Dále zde jsou problémy při vizualizaci.

Stejně jako u tezauru se s využitím ontologie setkáme spíše ve velkých podnicích, i když ve specifických případech může sloužit i v jiných typech podniků.

Při tvorbě je nutné si uvědomit dvě skutečnosti: neexistuje správná cesta nebo metodologie, jak ontology vytvořit. A opět se vše odvíjí od potřeb podniku. Před tvorbou je nutné si definovat, co má být obsahem ontologie, a zjistit, zda není možné pozměnit stávající ontolgii.[58] Na začátku si napíšeme seznam termínů, které jsou důležité a chceme je do ongologie zanést. Následně vytvoříme hierachii tříd, můžeme využít metodu postupující shora dolů (od obecného ke specifickému) nebo zdola nahoru (od specifického k obecnému). Je možné oba přístupy kombinovat. Následně jsou k třídám přidány vlastnosti. K jednotlivým třídám lze přidávat restrikce, které vymezují, které instance do dané třídy patří. V OWL najdeme tři kategorie: ohraničení kvantifikátoru (Quantifier restriction), kardinální ohraničení (Cardinality restriction) a ohraničení máVlastnost (hasValue restriction). (Klempa, 2006)

Při tvorbě je vhodné se vyhnout mít stále na zřeteli: ontologie nebude obsahovat všechny informace či vlastnosti z použité domény, není nutné nastavovat u všech tříd, že jsou vzájemně disjunktivní, používání stejného čísla u jmen (buď množné, nebo jednotné), pro rozeznání tříd a vlastností je vhodné používat prefix (v aj has-) či sufix (v aj -of), tzn. v češtině má-vlastnost (např.: má výrobce, má polevu) je-vlastností (např.: je výrobcem, je poleva), nepoužívat zkratky. Příklad navržené zákaznické ontologie na obrázku č. 22 demonstruje možnosti vizualizačních pluginů v nástroji Protégé.

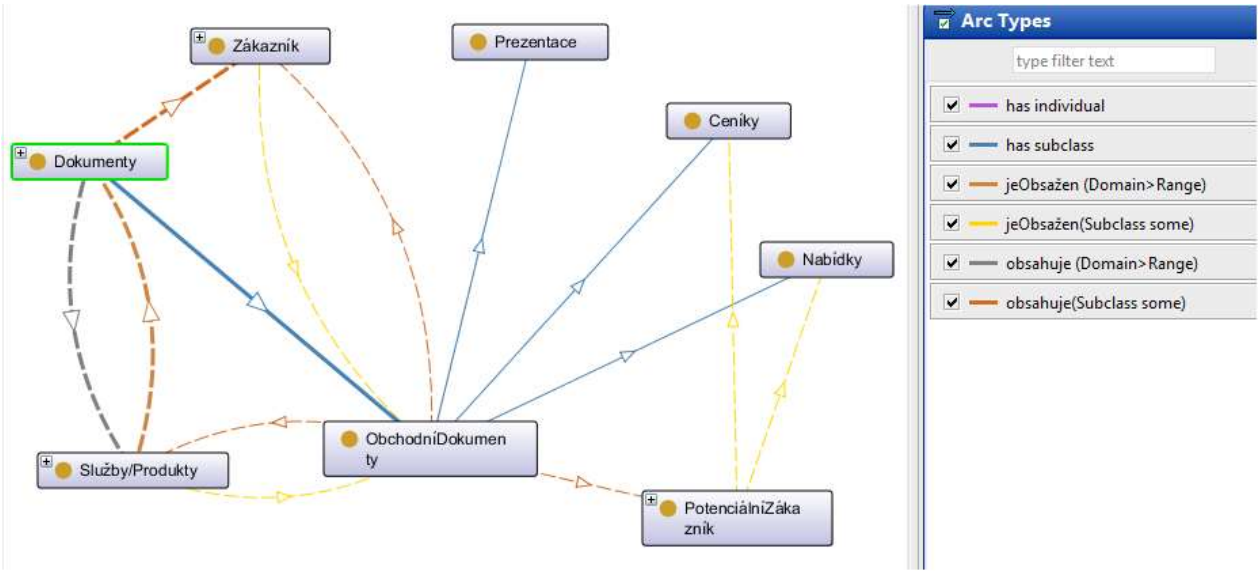
[58] Knihovny ontologií jsou dostupné např. na http://www.daml.org/ontologies/ nebo na http://www.ksl.stanford.edu/software/ontolingua/ .



Obrázek 22: Část navržené zákaznické ontologie (zdroj: autorka)

Z obrázku je patrné, jak může mít ongologie složité vazby mezi jednotlivými objekty. Z toho důvodu je níže zobrazena pouze její část, obchodní dokumenty. Je zřetelné, že obchodní dokumenty patří do třídy dokumenty, nadřazeny objektům jsou prezentace, ceníky, nabídky. Dále je zde čitelný vztah mezi obchodními dokumenty, které obsahují zákazníka, stejně jako potenciálního zákazníka a služby/produkty.

V příloze č. 9 je k dispozici celá navržená ontologie zákaznických dokumentů ve formátu OWL.



Obrázek 23: Výsek z ontologie - obchodní dokumenty (zdroj:autorka)

Jak může nasazení ontologie zlepšit automatickou indexaci, je rozebráno v odborném článku Chaminda Manjula Wijewickrema (Wijewickrema 2014).

Tato podkapitola měla za cíl ozřejmit, jaké jsou možnosti při tvorbě předmětového selekčního jazyka. Seznamuje s jednotlivými typy, jaké jsou jeho klady, zápory a využití.

Jedním z vedlejších přínosů této podkapitoli je možnost sledovat dané termíny, jak se v rámci jednotlivých typů věcných selekčních jazyků mění (některé zůstávají, jiné se zcela vytrácí či jsou použity v jiném tvaru). Proměna a způsob použití termínu může být jedním z vodítek, jaký typ třídění by měl podnik použít. Obecně můžeme říci, že by podnik měl zvážit:

- Jak detailní chce mít třídění – stačí několik termínů s obecnými názvy, nebo specifické termíny?
- Jakým způsobem chce s tříděním pracovat – bude použito pro prohlížení či vyhledávání? Od toho se pak odvíjí, zda bude volit prekoordinovaný či postkoordinovaný selekční jazyk.
- Kdo má s tříděním při kategorizaci pracovat – bude katalogizaci provádět člověk, nebo stroj?

6.4 Implementace

V rámci podkapitoly implementace se budeme zabývat jejími dvěma částmi (tvorbou technického provedení a kontrolou navrženého třídění). Ostatní podčásti této fáze nejsou z pohledu metadat zajímavé, odvíjí se od typu nasazované technologie a od projektové metodiky dané organizace.

Ve fázi **implementace** je nutné definovat, která podmnožina zákaznického obsahu bude použita pro fázi PILOTu/POC, na které se ověří, že navržené třídění splňuje očekávání organizace. Je vhodné vybrat část, která je relativně samostatná, není zde složitá integrace na procesy či aplikace (není součástí velkého počtu procesů, není náročná na vstupní data, není nutné provádět velké technické změny do návazných aplikací). Měly by to být dokumenty, které jsou využívané koncovým uživatelem téměř okamžitě po jejich vložení do systémů, anebo se může jednat o zákaznický obsah, s nímž jsou nyní ve firmě problémy. Uživatelé tak zažijí přínosy, které jim nasazení klasifikace potažmo technologie ECM umožňuje. Pokud nebude systém ve firmě používán ve větší míře jako back-end, ale držitelé budou aktivně používat jeho uživatelské prostředí (byť může být i vnořeno ve formě iframu či widgetu do jiné aplikace), je vhodné, aby byla vybrána oblast, která umožní uživateli seznámit se s tímto prostředím. Díky tomu může navíc do další implementační fáze poskytnout požadavky týkající se úpravy GUI[59] při práci s daty.

V této fázi by se měli definovat vlastníci jednotlivých zákaznických obsahů, pokud tak nebylo učiněno již dříve. S velkou pravděpodobností bude těžké najít jediného vlastníka dat pro všechny dokumenty, zvláště jedná-li se o velkou společnost. Proto je vhodné oblast rozdělit do menších celků, nejspíše se využije dělení podle jednotlivých druhů dokumentů (faktury, smlouvy,…). Ale i zde může docházet ještě k dalšímu dělení na menší podčásti (např. jiný vlastník bude mít pod sebou smlouvy pro velkoobchod, jiný smlouvy pro maloobchod, …). Pokud tomu tak je, je nutné na to při definici používaných metadat pamatovat zejména proto, aby bylo možné v systému ta která data jednotlivých vlastníků vyhledat, což může být mimo jiné nutné i z pohledu nastavení oprávnění nad obsahy. Tento požadavek s sebou přináší nutnost generovat změny do stávajících podnikových procesů (bude nutné stanovit práva a omezení vlastníků). V rámci této problematiky by se měla organizace zamyslet, zda zároveň neustanoví oddělení či specialisty v jednotlivých

[59] Zkratka GUI pochází z anglických slov graphical user interface, překládané jako grafické uživatelské rozhraní. Toto rozhraní umožňuje koncovému uživateli práci s daty, součástí tohoto rozhraní bývají grafické prvky jako například tlačítka, rolety, ikony,…

oblastech/odděleních, kteří by zodpovídali za datovou kvalitu metadat, a správce, který bude mít na starosti architekturu celého vytvořeného třídění. V podnicích v USA se setkáme s oddělením, které tuto oblast zastřešuje, většinou má toto oddělení na starosti i spisovou službu, archivační proces, tvorbu a správu informační politiky (objevují se zde názvy pracovních pozic jako document manager, records and information manager, enterprise information and records manager,…).

6.4.1 Tvorba technického provedení

Nacházíme se v situaci, kdy máme vytvořenu „ideální" klasifikaci zákaznických dokumentů dané organizace. Ideální ve smyslu, že byla navržena bez omezení, která s sebou již samotné aplikace přináší. Je vytvořena tak, jak by měla být postavena, pokud by neexistovala technická či jiná omezení. Při tvorbě technického provedení, ale následně i při testování a v provozu je nutné mít stále na zřeteli 3 faktory, které ovlivní, zda bude navržená klasifikace kvalitní. Jsou to: **lidé** (uživatelé systému), **procesy** (podpora všech podnikových procesů), **technologie** (aktuální a budoucí stav aplikací v podniku).

Při tvorbě detailního technického provedení potřebujeme znát, jaký je stávající stav a vize IT architektury ve firmě. V organizaci může existovat (zvl. ve velkých podnicích) velké množství různých typů úložišť (od sdíleného úložiště na discích, přes DMS systémy až k proprietálním uložištím v jednotlivých podnikových systémech typu ERP, CRM, HRM,…). Z pohledu uživatele nás bude zajímat, zda k těmto systémům přistupuje jednotlivě nebo prostřednictvím portálového řešení, jehož výhodou je přístup z jednoho prostředí k více zdrojům.

Měli bychom si při analýze stavu klást otázky: Existuje ve firmě více ECM technologií či jiných úložišť, kam jsou archivovány dokumenty? Pokud ano, plánuje se jejich postupné slučování s cílem jednotného uložiště nebo jednotného úložiště pro určitý typ dokumentů (např. všechny zákaznické dokumenty budou uloženy v softwaru Documentum, personální dokumenty budou uloženy v SAPu apod.)? Je možné stanovit, jaké typy dokumentů mají být do kterého úložiště ukládány? Pokud existuje více uložišť pro zákaznické dokumenty, je možné sjednotit názvy, číselníky používaných metaúdajů – je možné do nich navržené obecné třídění implementovat? Nebo bude implementována pouze část, která je součástí probíhající aktivity?[60]

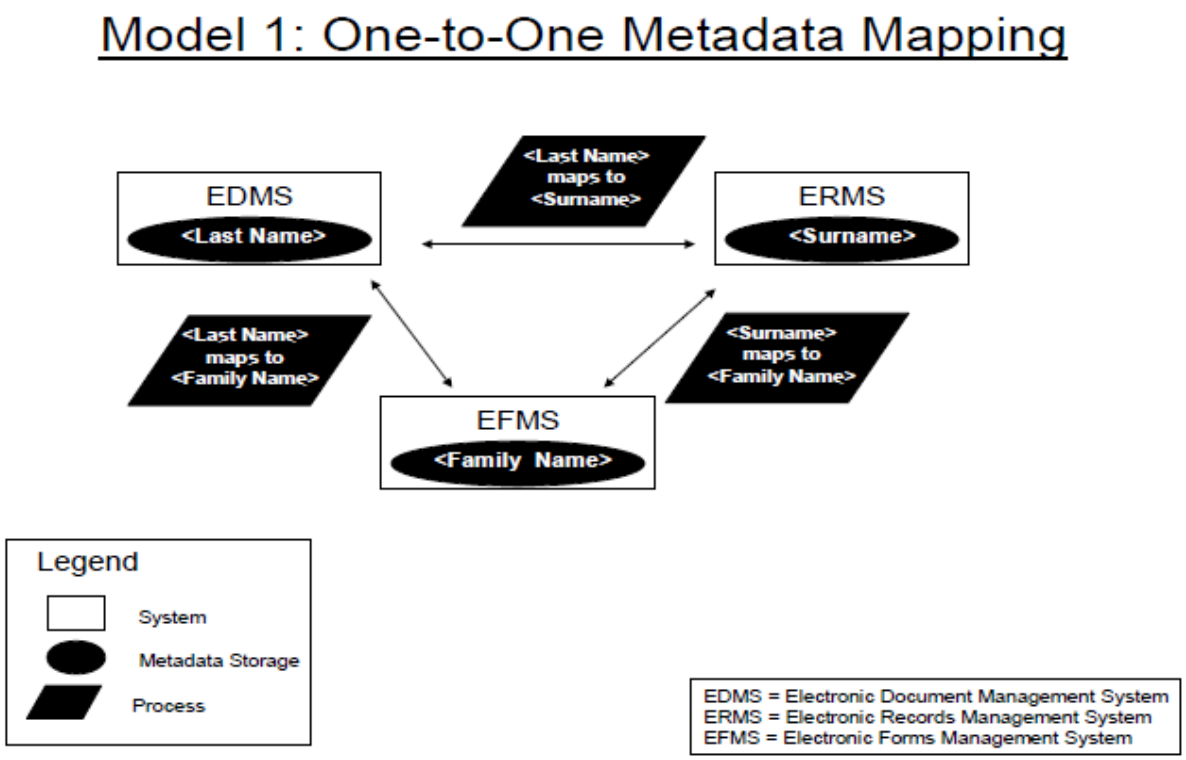
[60] Existence více uložišť v podniku zvyšuje nároky na správu, rozvoj, podporu a vzájemnou integraci, což navyšuje náklady na IT v podniku.

Z pohledu integrace jednotlivých úložišť nás zajímá, zda budou dokumenty a metadata dostupné pouze z dané aplikace, nebo budou přístupné i odjinud. Způsob a úroveň integrace (např. datová, integrace aplikací, procesů,…) jsou podmíněny opět organizací a jejími pravidly a zvyky. Specifickou oblastí je federativní vyhledávání, které má své využití v portálových řešeních. Existují zde dva přístupy: a) front-endová aplikce předá dotaz jednotlivým úložištím a počká na odpovědi a výsledek následně zobrazí uživateli, b) front-endová aplikace si u sebe udržuje centrální index a provádí vyhledávání v něm (tzn. je nutné použít mechanismus, který zajistí přenesení změn v úložiších do front-endu). V praxi se můžeme potkat i s hybridním řešením, což je kombinace obou přístupů (některé aplikace poskytují data do centrálního indexu, v jiných aplikacích probíhá on-line vyhledávání)[61]. Jednou z cest, jak zjednodušit integrační proces, je využít standard CMIS, standard umožňující přístup a sdílení dat mezi různými víceúložišti ECM (kapitola 4.9.3).

Zajímavý pohled na integraci metadat nalezneme ve standardu ANSI/AIIM/ARMA TR48-2006, který představuje tři způsoby, jak může být realizována výměna metadat mezi systémy.

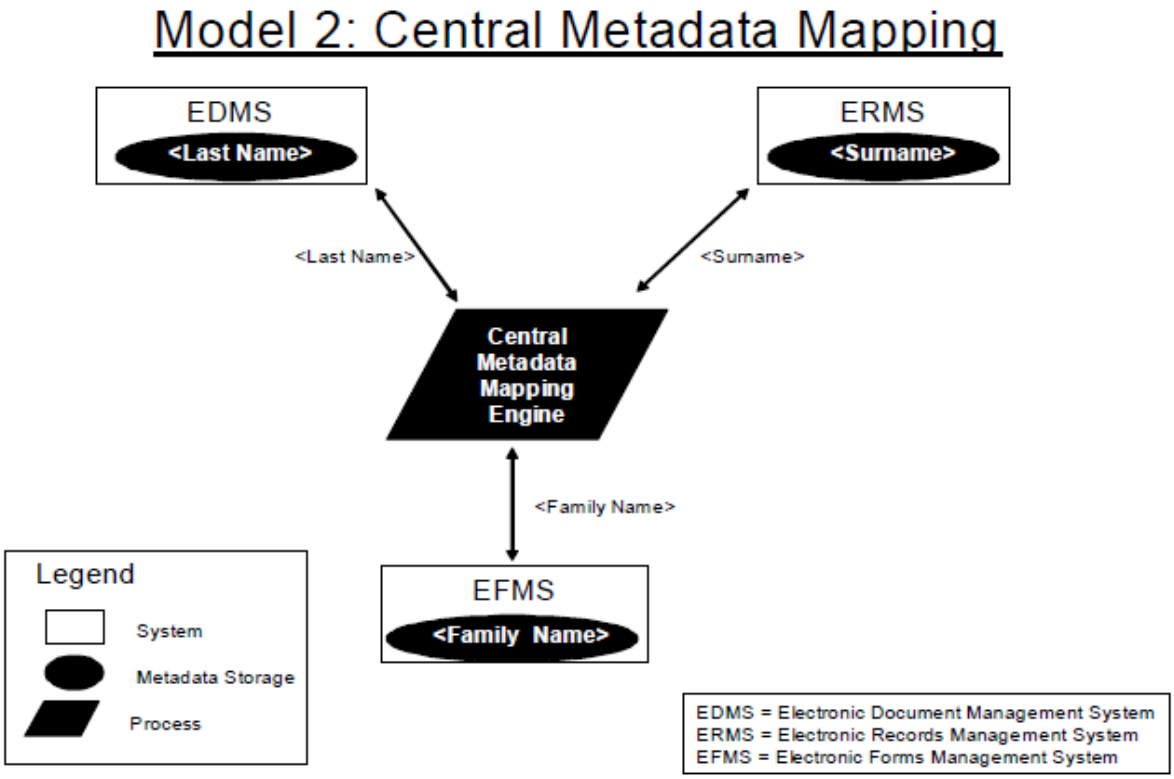
První z nich je mapování jeden na jednoho. V rámci aplikací, které spolu komunikují, se vyberou ta metadata, která jsou pro oba systémy shodná, a provede se věcné mapování metaúdaje na metaúdaj, viz obr. níže. Při změně metadatové struktury se musí zajistit, aby se změnou počítal i cílový systém.

[61] V knihovnicko-informačním světě se setkáme s protokolem OAI-PHM, jehož účelem je umožnit sklizeň metadatových záznamů z různých depozitářů.



Obrázek 24: Mapování metadata jeden na jednoho (zdroj: AIIM TR48)

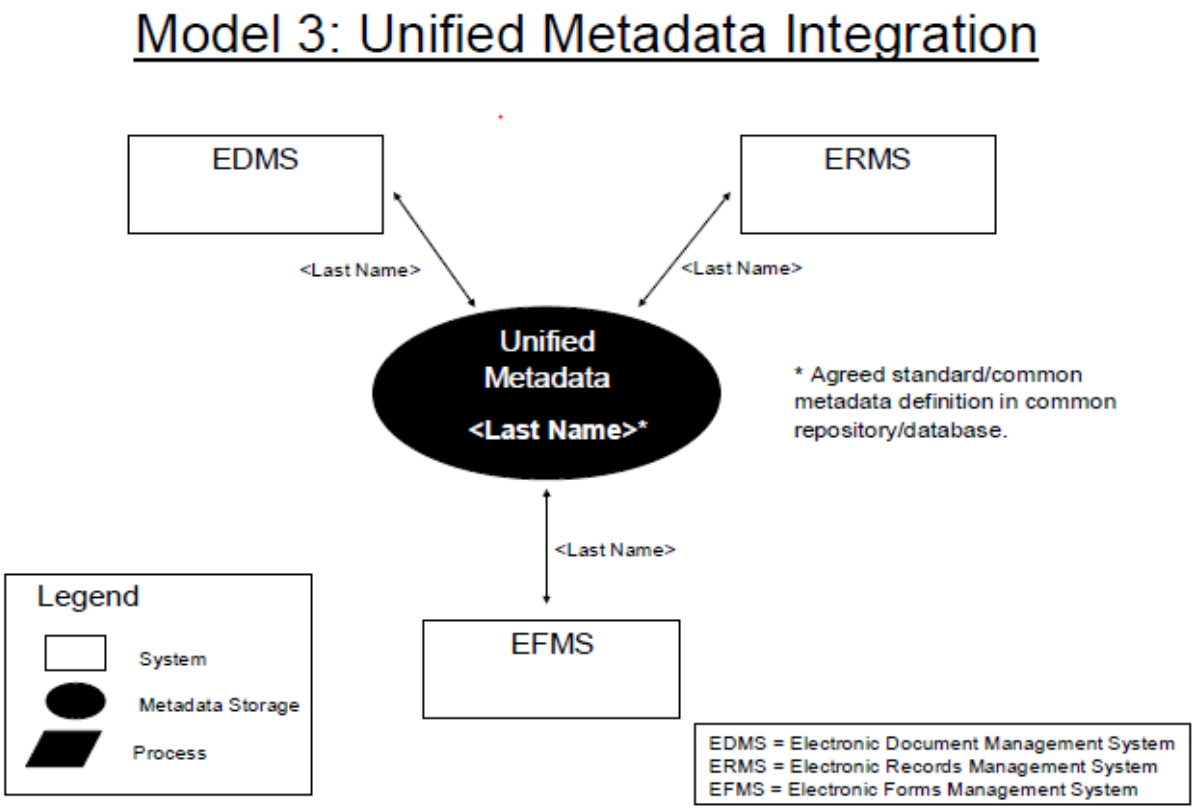
Druhým způsobem je centrální mapování metadat. V rámci tohoto řešení je vytvořena aplikační platforma, se kterou dotčené programy přes své aplikační programové rozhraní komunikují. Jde o typ integrace jeden na mnoho (one-to-many), neboli jedno úložiště se takto integruje s více úložišti. Výhodou je, že aplikace si nemusí pro každý typ úložiště vytvářet své vlastní programové rozhraní, ale stačí jí pouze jediné – napojení na centrální mapování metadat. Vše je znázorněno na obrázku 25.



Obrázek 25: Centrální mapování metadat (zdroj: AIIM TR48)

Posledním typem integrace je sjednocená metadatová integrace. V tomto případě dochází k tomu, že ta metadata, která jsou využívána jinými aplikacemi, jsou převedena do jednotného skladu – metadata informační architektury a dochází ke skutečnému sdílení metadat. Jde o integraci typu mnoho na mnoho (many-to-many) – viz obr. níže.

Na tomto druhu integrace je nejvíce patrné, proč je nutné vytvřit obecný rámec zákaznického třídění, neboť se velmi zjednodušeně dá říct, že sjednocená metadatová integrace se rovná obecnému rámci třídění.



Obrázek 26: Jednocená metadatová integrace (zdroj: AIIM TR48)

Nyní jsme ve fázi, kdy se musí vize klasifikace zákaznických dokumentů přetavit do datového modelu metadat, který bude následně implementován v úložišti či úložištích. Opět využijeme data, která byla získána v předcházejících fázích. Zůstaneme u tabulkového zobrazení, viz tabulka 14.

Pole, které říká, v jakém systému je metaúdaj držen, je tam z toho důvodu, že vycházíme z obecného návrhu třídění, které definuje veškeré potřebné vlastnosti u jednotlivých druhů evidovaných dokumentů. Nicméně v této fázi může dojít k domluvě s klíčovými uživateli, že potřebná metadata mají v jiných systémech a není nutné je tedy u dokumentu uchovávat (např. stačí dokument zaindexovat pouze pod jedním číslem zákazníka, není nutné u něho mít identifikaci služby/produktu, neboť dokument je přímo svázán se službou u daného zákazníka v CRM systému). Nejspíše by se mohlo zdát, že by bylo vhodné tyto metaúdaje v datovém modelu určeném přímo pro ECM neuvádět. To ale není správné řešení, po odmázání se tím přijde o celkový pohled na klasifikaci, který by při budoucích požadavcích na změnu mohl chybět. Zároveň slouží jako kontrola, že nebyl žádný metaúdaj vynechán.

Důvody, proč jsou některá metadata mimo ECM technologie, jsou různé, mohou to například být:

- Technické omezení ECM technologie (omezený počet klientsky definovaných metadat, omezení šířky databázového řádu tabulky, do které se hodnoty zapisují,…).
- Procházejí častou změnou, což by znamenalo aktualizovat hodnoty i v ECM.
- Většina uživatelů tento metaúdaj nepotřebuje, nevyhledává se přes něj a minoritní uživatelé jsou schopni si tento údaj bez problému získat z jiné aplikace.

Zároveň je toto pole rozšířeno o informaci, jaká aplikace řídí kmenová data (tzn. jaký systém data do pole generuje, údržba hodnot číselníku,…).

Tabulka 14: Návrh datového modelu (zdroj: autorka)

		Název metaúdaje		
Zobrazovaný název		Cislo faktury	Cislo smlouvy	…
Systémový název		cislo_fak	cislo_sml	
Popis/poznámka		vygener. číslo daňového dokladu	přidělené číslo obchodní smlouvy	
Datový typ		string	Integer	
Délka pole		20	10	
Možnost vložit multi hodnoty?		Ne	Ano	
Výběr hodnot z číselníku?		Ne	Ne	
Vytvořit v dtb index? (tzn. využíván pro vyhledávání)		Ano	Ano	
Vynutit povinnost vyplnění		Ne	Ne	
Metaúdaj je shodný s …		Faktura ID (ERP)	Contract ID (CRM)	
Proces - vstup	Fakturace - archivace dokladu	Ano	Ne	
Proces - vyhledávání	Fakturace - zobrazení faktury na portále	Ano	Ne	
Metaúdaj je dostupný v systému /master dat je systém		ECM/ERP	ECM/CRM	
Architektura vnitřního členění	složka - Faktury	Ano	Ne	
	složka - SME smlouvy	Ne	Ne	

Pole „architektura vnitřního členění systému“ by mělo být upraveno dle implementované či používané ECM technologie. Některé technologie používají zařazování

obsahu do tříd, jiné do složek, knihoven, na položky apod. Tyto objekty samy o sobě mají název a mohou být doprovázeny vlastními metadaty.

V případě, že je v poli „výběr pole z číselníku“ = Ano, pak je vhodné tento číselník[62] vydefinovat v separátím listě a na tento slovník se zde pouze odkázat. Používat číselník je výhodné v případě, že se jedná o malé množství hodnot, které se často nemění. Jeho používáním zamezíme uživatelské kreativitě, uživatel je nucen vybrat jednu z nabízených položek (nemůže do pole vepisovat volně tvořené termíny). To vede k jednotnému používání správné formy termínu. Opět je nutné při tvorbě dbát zásad, které byly již výše popsány, rozhodnout, zda bude číselník ve formě zkratek či popisu, vyhnout se obecným termínům, které by sváděly uživatele k použití, místo, aby dokument indexovali pod správným termínem.

Mezi specifická metadata patří ta, která obsahují předmětové selekční jazyky (např. řízený seznam termínů, taxonomii,…). V rámci podniku může být použit pouze jeden typ, nebo kombinace více typů. Taxonomii je možné realizovat v rámci jednoho pole v aplikaci, ale i použít více polí (jednotlivé hierarchické úrovně rozdělit do jednotlivých polí).

V teoretické části této práce, v kapitole vymezující základní pojmy, jsou uvedeny termíny spojené s vyhledávání dokumentu. Způsob, jaký koncovému uživateli vyhovuje, je nutné brát na zřetel (kapitola 2.7). Pokud uživatelé preferují prohlížení, volíme taxonomii nebo řízený slovník termínů, preferují-li analytické vyhledávání, pak je vhodnější využít fasetové vyhledávání či tezaurus. Specifické je mobilní vyhledávání, u něhož musíme počítat s několika faktory: uživatel chce rychlý a relevantní výsledek, provádí jednoduchá krátká hledání, je omezen časem či přístupem na internet (např. připojení je pomalé, stahování velkého objemu dat je drahé, omezení na zobrazované formáty dat,…). V tomto případě nutné filtrování dat může být provedeno automaticky či manuálně samotným uživatelem. Filtrovat je možné na základě aktuálního času, dle geografického místa, kde se zrovna uživatel nachází nebo dle přednastavených referencí uživatele. Pro mobilní přístup se používá jiné GUI, které se liší nejen graficky, ale i i funkčně (některé funkce systému nejsou přístupné). Protože je nutné, aby se uživatel dostal k informacím rychle (bez zbytečných kliků navíc) a při vyhledávání nebyl zahlcen nerelevantními dokumenty, je vhodné pro mobilní přístup vydefinovat speciální vkládací a vyhledávací šablony. Mobilní přístup bude například potřebovat technik, který je u zákazníka a šetří s ním způsobenou škodu. Pro práci v terénu

[62] Termín číselník je v běžné praxi hojně využíván, jde o aplikaci pojmu řízený slovník, který je možné používat i na jiná metadata, než jsou pouze metadata spojená s předmětovým tříděním.

mu při vyhledávání stačí, aby se zobrazily pouze dokumenty obsahující technickou dokumentaci. Potřebuje-li reklamovaný stav zdokumentovat, dostačuje, aby při vytváření dokumentu vyplnil pouze základní metaúdaje a připojil fotodokumentaci, podrobné vyplnění provede, až bude pracovat přes počítač.

Při návrhu jmen metadat musíme brát na zřetel nejen principy popsané výše (stejný jazyk, shodný tvar podstatných jmen apod…), ale i specifika daného softwaru. U IBM FileNet P8 platí, že při vytváření objektů se nesmí začínat mezerou či číslem, názvy tříd smí obsahovat abecedu, číslice od 0 do 9 a ze speciálních znaků pouze spodní podtržítko. Dále je nutné se při vytváření třídy vyhnout rezervovaným prefixům (Cm, DITA, RM, EDM, EDISC, ICC), které jsou používané samotným systémem a v budoucnu by mohlo dojít k nechtěným kolizím při upgradu či drobných změnách (je možné, že by se vytvořil zákaznický objekt, který by měl v budoucnosti shodný název se systémovým objektem, což by zapříčinilo nefunkčnost jedné či více komponent). (IBM Knowledge Center)

Pole „povinnost vyplnění“ specifikuje, že tento metaúdaj je nutné při zakládání dokumentu vždy vyplnit a nezáleží na tom, jakým způsobem je dokument do systému ukládán (zda ho přes GUI ukládá uživatel, nebo systém přes webové služby,…).

Někteří dodavatelé ECM technologie mohou poskytnout vlastní vzorové tabulky, které jsou psány v terminologii systému a reflektují jeho vnitřní architekturu. V tomto případě je možné využít tyto tabulky a dle potřeby je rozšířit o pole z tabulky 14. V příloze č. 10 je příklad z dokumentace MS SharePoint 2013 pro plánování metadat.

Při definici datového pole bychom kromě výše popsaného měli u některých typů (např. data) vydefinovat vstupní masku (např. jak se bude zapisovat datum: 1.1.2017, nebo 01.01.2017, nebo 2017/01/01,…). Při definici metaúdaje datové pole je nutné vzít na zřetel nejen reálné hodnoty (jsou-li číselné, textové,…), ale i způsob, jakým se vyhledává a následně pracuje. Neboť pokud uživatel potřebuje při vyhledávání použít operátor „začíná s“, pak není možné pole definovat jako číselné, ale jako textový typ apod. Pro referenci o způsobu vyhledávání pomůže tabulka č. 9.

U některých polí může systém automaticky generovat data prostřednictvím externího identifikátoru, který je potřebný v návazných procesech při práci s dokumentem, automaticky předgenerované prefixy či sufixy.

Při specifikaci datových polí a číselníku by se mělo mít na paměti, že některé typy metadat jsou definovány různými normami a pro následnou integraci s jinými systémy, je vhodné používat dle stanoveného standardu. Níže jsou vyjmenovány některé z nich:

- ČSN EN ISO 3166-1: Kódy pro názvy zemí a jejich části - Část 1: Kódy zemí
- ČSN ISO 8601: Datové prvky a formáty výměny - Výměna informací - Zobrazení data a času
- ČSN ISO 19115: Geografická informace – Metadata

Nejvíce variabilní typ pole, do kterého je možné vložit jakýkoliv volně tvořený textový řetězec, je pole typu textový řetězec (v programovacích, značkovacích jazycích datový typ string, u databázových systémů nazývaný varchar). K jeho výhodám patří to, že nepotřebuje pravidelnou úpravu z pohledu zastarávání hodnot, jako je tomu u číselníku, není omezen na jeden typ informace (může obsahovat jak text, tak číslice, časový údaj apod.). Nicméně díky této vlastnosti je těžší na něm nastavit kontroly a uživatelé v něm bývají kreativní. K jeho nevýhodám patří, vzhledem k tomu, že nebývá formalizováno, že se pak hůř využívá pro automatizované systémy či analýzy. Uživatelé místo toho, aby vyplňovali jiné metaúdaje, budou raději využívat toto pole s nestrukturovanými informacemi. Existuje využití, kdy je volný text žádoucí a není možné ho zcela (nebo pouze v omezené míře) nahradit formalizovanými metaúdaji. Příkladem mohou být anotace, název obsahu apod.

Pro práci v uživatelském rozhraní bude nutné nastavit vkládací šablony. Pro její specifikaci je opět vhodné využít tabulkovou formu, vzor s příkladem využití je k dispozici níže.

Tabulka 15: Definice vkládacích šablon v GUI (zdroj: autorka)

Metaúdaj		**Číslo smlouvy**	**Číslo zákazníka**	**…**
ID šablony	1			
Interní ID šablony	{GUID}			
Název šablony	Smluvní dokument			
Popis	použít při vkládání smluvních dokumentů			
Architektura vnitřního členění (do složky, knihovny,…)	Smlouvy (složka)			
Povinnost metaúdaje		Ano	Ano	

Přednastavené hodnoty		Ne	Ne	
Kontrola vkládané hodnoty		Ne	začíná 55, délka max. 10 znaků	
Popis/nápověda k poli			Hodnota musí začínat 55, s max. délkou 10 znaků	
Oprávnění umožňující práci se šablonou	Pravni_odd_all, Recepcni_all			
Oprávnění umožňující editaci	Admin_pravni			
Poznámka				

Pro práci v uživatelském rozhraní bude nutné nastavit také vyhledávací šablony. Pro její specifikaci je opět vhodné využít tabulkovou formu, vzory s příkladem využití je k dispozici níže.

Tabulka 16: Definice vyhledávacích šablon v GUI (zdroj: autorka)

Architektura vnitřního členění (do složky, knihovny,…) vstup		**Smlouvy (složka)**				
		Metadata - vstup	**Operátor**	**Přednastavená hodnota**	**Metadata - výstup**	**Třídit dle**
ID šablony	1					
Interní ID šablony	{GUID}					
Název šablony	Smluvní dokument					
Popis	použít při vyhledávání					
Oprávnění umožňující práci se šablonou	Pravni_odd_all, Recepcni_all					
Oprávnění umožňující editaci	Admin_pravni_odd					
Poznámka						
		Číslo smlouvy	Začíná s		Číslo smlouvy	3
		Číslo zákazníka	Začíná s		Číslo zákazníka	1
		Typ smlouvy	Rovná se	SML_HL	Typ smlouvy	2
					Autor	4
					Datum vložení	5

Navržené tabulky je potřeba upravit podle možností a architektury, které cílový produkt podporuje. Na první pohled se může zdát, že vytváření navržených tabulek výše je zbytečné a jen se tím vytváří administrativní zátěž – proč zadané hodnoty nezafixovat na schůzce s klíčovými uživateli, přes e-mailovou komunikaci, či po telefonu. Vytvoření tabulkových vzorů s sebou nese několik výhod:

- Tabulková prezentace je srozumitelná i pro klíčové uživatele, je možné na ní pracovat v offline podobě (konzultant ECM nemusí být při definování, které bude vedeno spíše na uživatelské úrovni, přítomen).
- Na základě tabulek je možné v určitý čas změny zafixovat a předat k nastavení v systému (předejde se tím případným dohadům, kdo a proč změnil vyhledávací šablonu či datový typ metadata. Nastavení v systému nemusí provádět seniorní ECM konzultat, ale může být provedeno juniorem – což sekundárně snižuje náklady na aktivitu).
- Tabulky mohou být dále udržovány a aktualizovány, jsou ideálně dostupné jak oddělení, která mají na starosti provoz, tak klíčovým uživatelům, businessovým analytikům (snadný přístup k informacím, jak je co v systému nastaveno, bez nutnosti kontaktovat vývoj či provoz).

Způsob, jakým bude navržené třídění implementováno do dané technologie ECM, záleží na jeho možnostech a případných možnostech investic do vývoje. Asi se nesetkáme s tím, že bychom měli problém s realizací seznamu či taxonomie, ale již můžeme mít problém s integrací tezauru či ontologie. Podobné je tomu i u možnosti klientsky definovat metaúdaje, což je funkce, které robustní ECM technologie umožňují, ale i zde mohou být omezení. S nemožností velkých změn a klientských nastavení se setkáváme u technologií, které jsou určené pro malé až střední podniky, zvl. kdy je DMS komponenta jedním z modulů podnikového softwaru (další moduly mohou být lidské zdroje, práce s projekty, …).

Níže předkládám návazné funkcionality, které mohou být v cílové ECM technologii dostupné, některé z nich vychází z principu Webu 2.0[63].

Zajímavou funkcí je **podržet**[64], jejímž cílem je zabránit tomu, aby byl dokument smazán. Je jedním z nástrojů spadajících do procesu archivace a skartace či pravidlům, které

[63] Termín odkazuje na novou etapu vývoje webových stránek, která se vyznačuje tím, že se ze statických stránek stal prostor pro společnou tvorbu a sdílení různých informací.
[64] V anglickém jazyce se používá termín hold.

mají zajistit shodu s právními předpisy[65] [66]. Jeho učebnicovým příkladem je zabránit výmazu dokumentu, pokud patří do důkazní dokumentace v rámci právního sporu. K odstranění příznaku podržet dojde až ve chvíli, kdy je spor ukončen (pozbyl důvod dokument dále archivovat) a je možné ho bez následků ze systému vymazat.

Kromě oficiálního firemního třídění umožují uživatelům některé systémy, aby dokumenty kategorizovali pod svojí vlastní klasifikací. Pro tyto účely se používají **klíčová slova**[67] nebo **značky** (tagy). Otázka, zda povolit uživateli vlastní třídění či ne, patří opět mezi ty, na které neexistuje jednotná odpověd a odvíjí se od situace a typu podniku. Mezi důvody proč ano, patří například:

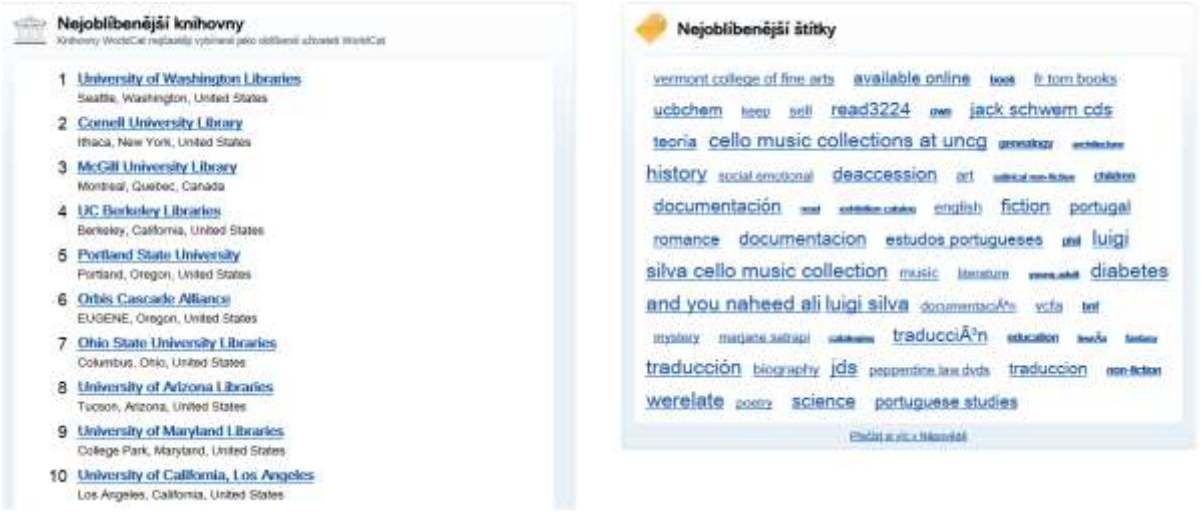
- každý člověk má vlastní vnímání reality, tzn. i jiný způsob práce, tedy i jiný způsob třídění dokumentů[68], to že si bude moci obsahy třídit dle svých zvyků, mu práci usnadní a urychlí;
- centrálně zpravované třídění nikdy nebude moci rychle reagovat na změny a požadavky[69];
- vyřeší se tím případné kolize u malých skupin, které mají specifické potřeby.

Nevýhody mohou být spíše technického rázu, i když se může vyskytnout případ, že skupina uživatelů si vytvoří vlastní klasifikaci, kterou bude používat a přestane dodržovat celofiremní standard. Druhým extrémem je používání nesmyslných termínů, které se k obsahu nehodí, což následně vede ke snižování relevance při vyhledávání. K technickým problémům může například patřit:

- problém s výkonem systému;
- co dělat s tagy uživatele, který odejde;
- pokud v rámci ECM existuje vlastní pracovní prostor (kam se mohou i vkládat obsahy), co s tímto prostorem po jeho odchodu.

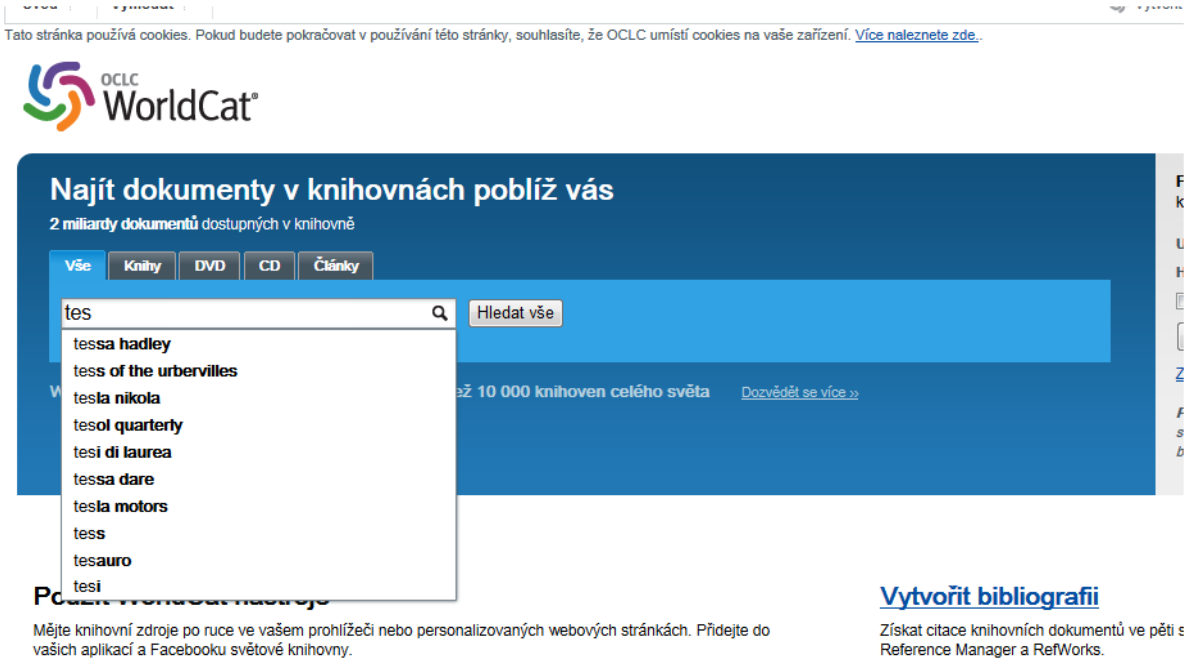
[65] V anglickém jazyce se používá termín compliance.
[66] Jeho typické využití je v rámci spisových služeb či správě záznamů (record management).
[67] Klíčová slova svou podstatou připomínají předmětová hesla, ale na rozdíl od nich jde o použití volného textu bez řízeného slovníku. V praxi se velmi často setkáváme s tzv. autorskými klíčovými slovy, což jsou termíny, které vydefinoval sám autor (není tomu jinak ani u vysokoškolských prací).
[68] Je to pozorovatelné už na tom, jakým způsobem lidé pracují se svým e-mailovým klientem (někdo třídí do složek, někdo si e-maily označuje, někdo netřídí vůbec).
[69] Toto je známý problém v informačně knihovnické oblasti, že třídění nejsou dostatečně pružná, aby reagovala na změny ve vývoji a v poznání světa (podléhají složitému schvalovacímu řízení). Na druhou stranu je to zároveň i jejich klad, zvyšuje se tím jejich kvalita, že v nich nenajdeme nesmyslné termíny, které měly jepičí život.

K funkci komunitního značkování se váže možnost zobrazení nejvíce využívaných **značek v oblaku** – čím je slovo více používané, tím je vizuelně větší. Na obrázku dole je příklad z OCLC OPACu[70].



Obrázek 27: Využití tzn. tag clouds v OCLC OPACu (WorldCat.org 2001-2017)

Další vlastnost, která urychlí nejen vyhledávání, ale i ukládání dokumentu do systémů, je **našeptávání**. Uživateli jsou při psaní nabízeny termíny aktuálně dle toho, co právě píše. Příklad na obrázku níže je opět z OCLC OPACu.

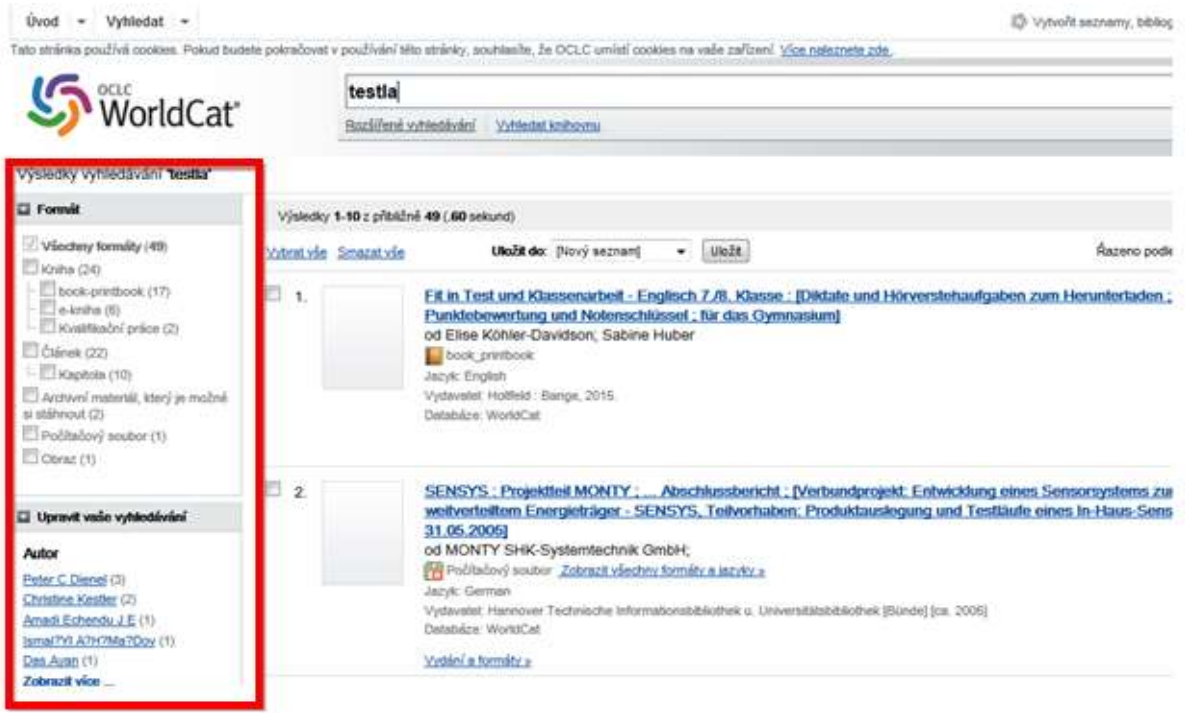


Obrázek 28: Funkce automatického našeptávání v OCLC OPACu (WorldCat.org 2001-2017)

[70] OPAC je zkratka používaná v oblasti knihovnictví a označuje veřejně přístupný online katalog (zkratka z angličtiny – online public acces catalog).

Technologie dále může umožnit třídit záznamy dle zvoleného metaúdaje (někdy může pole volit a měnit uživatel, jindy jsou předem nastaveny na vyhledávací šabloně a neměnné).

Vyhledávání také může ulehčit nasazení faset, které mohou být v systému používány ve formě filtru ve výsledcích vyhledávání – viz obrázek 29. Uživatel má možnost snadno a rychle upravit svůj dotaz a dostat se k lepším výsledkům. Protože jsou většinou nabízeny u dané fasety pouze ty hodnoty, které obsahují vyhledané výsledky, přináší to uživateli rychlý přehled o tom, s jakou množinou pracuje (např. z obrázku níže je patrné, že nejčastěji je slovo obsažené v preprintech a pokud chtěl uživatel najít nějaké obrázky, musí svůj dotaz položit úplně jinak).



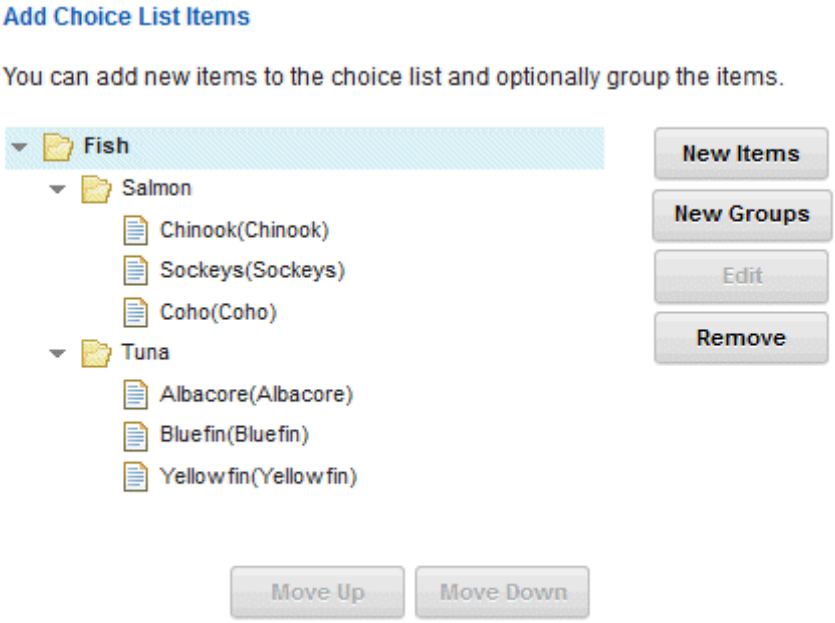
Obrázek 29: Možnosti úpravy vyhledávacího dotazu v OCLC OPACu (WorldCat.org 2001-2017)

S ulehčením vyhledávání se pojí i možnost označit obsah jako oblíbený, což následně uživateli urychlí přístup k dokumentu bez nutnosti jeho opakovaného vyhledávání.

Při ukládání obsahů do systémů se krom našeptávání můžeme setkat i s funkcí automatické klasifikace, která může být prováděna buď na základě předem definovaného pravidla (např. dokument oddělení péče o zákazníky ponese vždy tuto hodnotu v metaúdaji oddělení,…) nebo jsou termíny extrahovány přímo z plného textu (rozeznává se zde například váha jednotlivých slov – je-li slovo z názvu, v abstraktu, jak často se slovo v textu vyskytuje, blízko určitých slov,…).

Nyní se blíže podíváme na tři produkty od různých vendorů, jakým způsobem umožňují začlenění navrženého třídění.

Jedním z vendorů ECM technologie je společnost IBM, jejímž produktem je mimo jiné FileNet P8. V případě, že bychom v rámci technologie chtěli implementovat zákaznické třídění ve formě taxonomie, můžeme využít při nastavování metaúdaje výběr ze seznamu hodnot. Ty je možné nastavit i v hierarchické formě, což je doloženo obrázkem 30. Uživatel při ukládání a vyhledávání pracuje až s koncovými hodnotami – systém neumožní uložit či vyhledat dokument přes nadřazené uzly. Využijeme-li příkladu níže a uživatel bude chtít vidět všechny dokumenty typu Salmon, bude muset vybrat a nastavit vyhledávač takto: Fish = Chinook OR Fish = Sockeys OR Fish=Coho.



Obrázek 30: IBM FileNet P8 hierarchický seznam hodnot (IBM 2014)

Problém nastane, pokud bychom zde chtěli implementovat třídění ve formě tezauru. Chceme-li se vyhnout programování, pak mezi základní funkce patří nastavení synonymních výrazů, ale pouze pro fultextové vyhledávání[71]. Drobnému vývoji se nevyhneme ani u funkce automatické klasifikace či u možnosti značkování. Drobný je tento vývoj v tom smyslu, že produkt s těmito funkcionalitami má v základu připraveno vývojové rozhraní. (IBM 2014)

Podobně jako FileNet P8, tak i SharePoint 2013, který produkuje firma Microsoft, umožňuje vytvářet taxonomii. I zde je možné vytvářet hierarchický seznam hodnot, který je

[71] Při fultextovném vyhledávání se hledaný výraz neporovnává s metaúdaji vážícími se k dokumentu, ale s výskytem hledaného termínu přímo v samotném dokumentu.

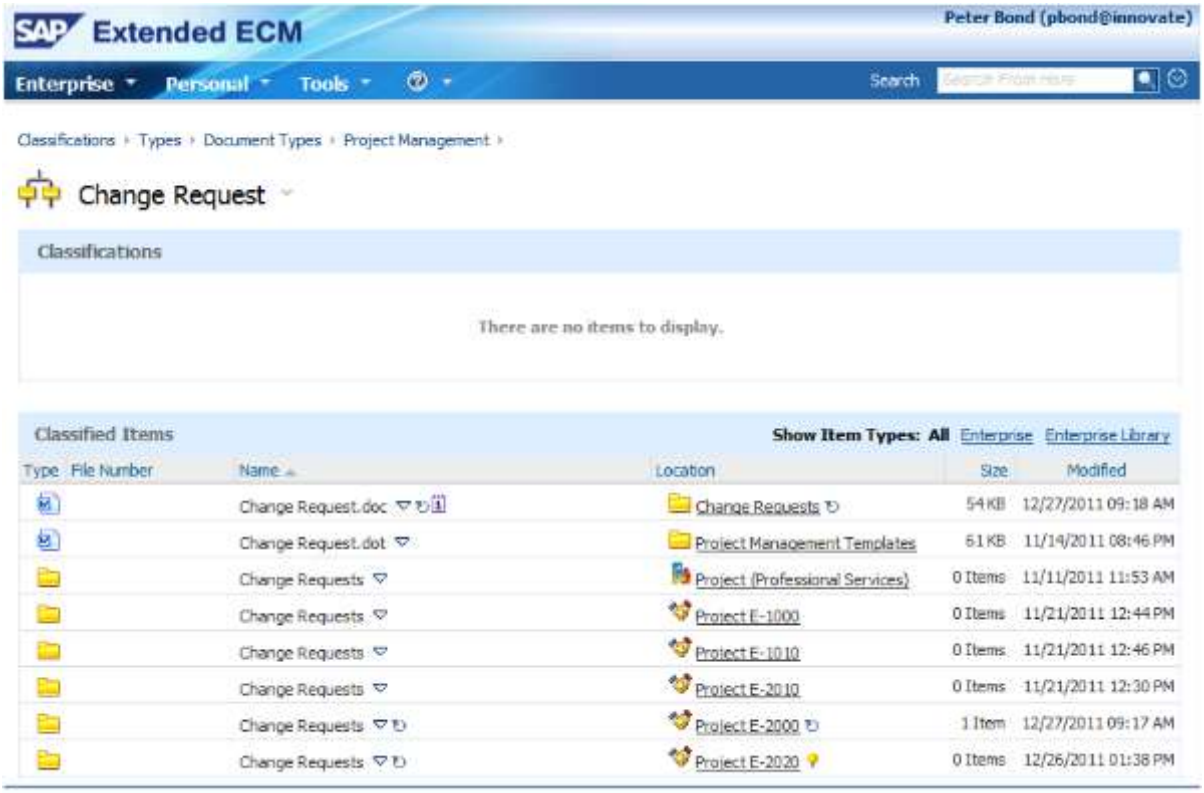
spravován centrálně komponentou Term Store Management Tool. I zde je uživatel omezen tím, že může do vlastnosti dokumentu vybrat poslední hodnotu v uzlu.

Technologie SharePointu je produktu FileNet podobná i v tom, že umožňuje rozšířit vyhledávání přes synonymní výrazy. Import se provádí přes CSV, který má vydefinovaný formát. Zajímavé je, že používaný tezaurus (takto je funkcionalita nazývána v aplikaci) není možné ze systému vyexportovat.

I pro SharePoint umožňuje automatickou indexaci, většinou se jedná o vývoje třetí strany.(Microsoft 2016)

Pokud nás zajímá funkce podržet, nalezneme ji u obou produktů popisovaných výše.

Posledním produktem je rozšíření SAPu o technologii ECM od firmy OpenText. Toto řešení umožňuje vytvářet hierarchické taxonomie, příklad viz obrázek 31. Na první pohled se uživatel může domnívat, že prochází klasickou složkovou strukturou, ale není tomu tak. Dokumenty zůstávají uloženy na svém původním místě (složce), sem jsou pouze odkazovány.



Obrázek 31: Open text taxonomie (OpenText 2012)

Mimo podporu vytváření taxonomie umožňuje využívat i funkci faset (komponenta Facet Browsing). V rámci komponenty dochází ke kombinování funkce prohlížení mezi

složkami s vyhledáváním. Metadata jsou zde využívána k tomu, aby se generovala taxonomie za běhu (podobné, jako na obrázku č. 29). (OpenText 2012)

Tento krátký exkurz mezi vendory ECM systémů[72] odhalil dvě skutečnosti. V rámci svých aplikací nepoužívají terminologicky správné označení jednotlivých funkcí, a některé názvy mohou být až zavádějící (např. název tezaurus u funkce, která rozšiřuje vyhledávání o synonyma).

Dále je zajímavé pozorovat, jakým způsobem je stavěna vnitřní architektura systémů, která je podmíněna jejich historickým vývojem. U produktu firmy IBM je patrný sklon k vytváření tříd a vyhledávání, u Microsoftu naopak k vytváření hierarchií ve formě složek a brouzdání[73]. Pozici někde mezi zaujímá produkt od firmy OpenText, který není čistě vyhrazen ani na složky, ani na vyhledávání přes metadata. Jejich způsob přidělování metadat k dokumentům je zajímavý (u FileNetu je toto řízené přes objekt třídy, u SharePointu naopak přes objekt knihovna). OpenText umožňuje na vkládaný dokument odvodit či zdědit metadata na základě složky, vkládací šablony, nebo od typu dokumentu.

Na to, že každý z produktů má jiné vnímání reality, by se mělo pamatovat již ve fázi výběru produktu, neboť z pohledu dodávaných funkcionalit jsou na tom podobně. Pokud uživatelé ve firmě jsou zvyklí a inklinují k prohlížení/ke složkové hierarchii, spíše pro ně bude jednodušší přechod na aplikaci typu SharePoint než FileNet a naopak.

Na srovnání výrobců softwarů jsme mimo jiné demonstrovali, že je vhodné nejdříve vytvořit obecné třídění, neboť tak vznikne objektivní obraz vnímání práce s obsahem v organizaci. Třídění bude následně upraveno dle vybraného produktu. Pokud se tato fáze přeskočí a začne se třídění vytvářet na základě možností dodávaného produktu, hrozí, že v systému sice bude vše nastaveno a připraveno, ale uživatelé nedokáží práci v systému pochopit a sžít se s ní, což může vést k tomu, že systém nebude nikdy ve firmě přijat a plně využíván (bude obcházen, uživatelé k němu budou zastávat negativní postoj a budou s ním pracovat neradi).

Mimo jiné by mohlo v rámci implementace docházet k různým nedorozuměním, díky nestejnému používání terminologie, kdy klíčový zákazník očekává tezaurus a místo toho

[72] Mimochodem firmy IBM a OpentText, se dle firmy Gartner, která pravidelně zveřejňuje analýzy týkající se ECM domény, i za rok 2016 patří mezi vůdce, kdežto Microsoft mezi vyzyvatele.
[73] Což je na jednu stranu i jedna z nevýhod, díky limitaci 5 000 položek na knihovnu, kdy jsou pak podniky nuceny hledat řešení, jak toto omezení obejít (často je to spojeno i s performance problémy). S podobným omezením je možné se potkat i u aplikace FileNet, kde je doporučeno použít maximálně 200 položek na stejné úrovni. Pokud je potřeba poskytnout více položek, pak je vhodné zvážit vývoj speciální vyhledávací komponenty.

dostane rozšířené vyhledávání o synonymní výrazy[74]. Tím, že se nejdříve vytvoří obecný rámec, dojde k sjednocení terminologie na straně klienta, který bude na dodavatele softwaru mít už detailně zpracované požadavky, co potřebuje zapracovat a v jaké podobě.

6.4.2 Kontrola navrženého třídění

Jedním z kroků metody PDCA je provedení kontroly. Tento krok je součástí projektových metodik či změnových požadavků v podnicích. Z pohledu navrženého třídění zákaznických dokumentů se zaměříme opět na to, čím je tato oblast specifická. Obsah rozdělíme na dva druhy kontroly, jedna probíhá před instalací probíhající aktivity do produkčního prostředí (a je součástí běžných testů), druhá probíhá již v rámci produkčního běhu (a je součástí pilotního provozu aktivity).

Na začátku je nutné vrátit se k cílům, které byly stanoveny v kroku plánování, a odvodit z těchto cílů otázky, jejichž zodpovězení povede k plnění cíle a na které se bude v rámci testování hledat odpověď. Využijeme zde příklad podcíle uvedeného v kapitole Plánování (kapitola 6.2.1), kdy uživatel potřebuje zajistit kompletaci dokumentů pro vymáhání dlužníka. Pak kontrolní otázky budou:

- Jaký typ dokumentů uživatel potřebuje a pod jakými metadaty jsou ukládány?
- Jakým způsobem je schopen zjistit hodnotu daného metaúdaje? (např. ID zákazníka ze CRM systému,…).
- Kolik vyhledávacích dotazů musí uživatel specifikovat, aby dostal všechny potřebné dokumenty (stačí mu pouze jeden pro všechny potřebné druhy dokumentů, nebo musí svůj dotaz několikrát upravovat, aby dostal všechny)?

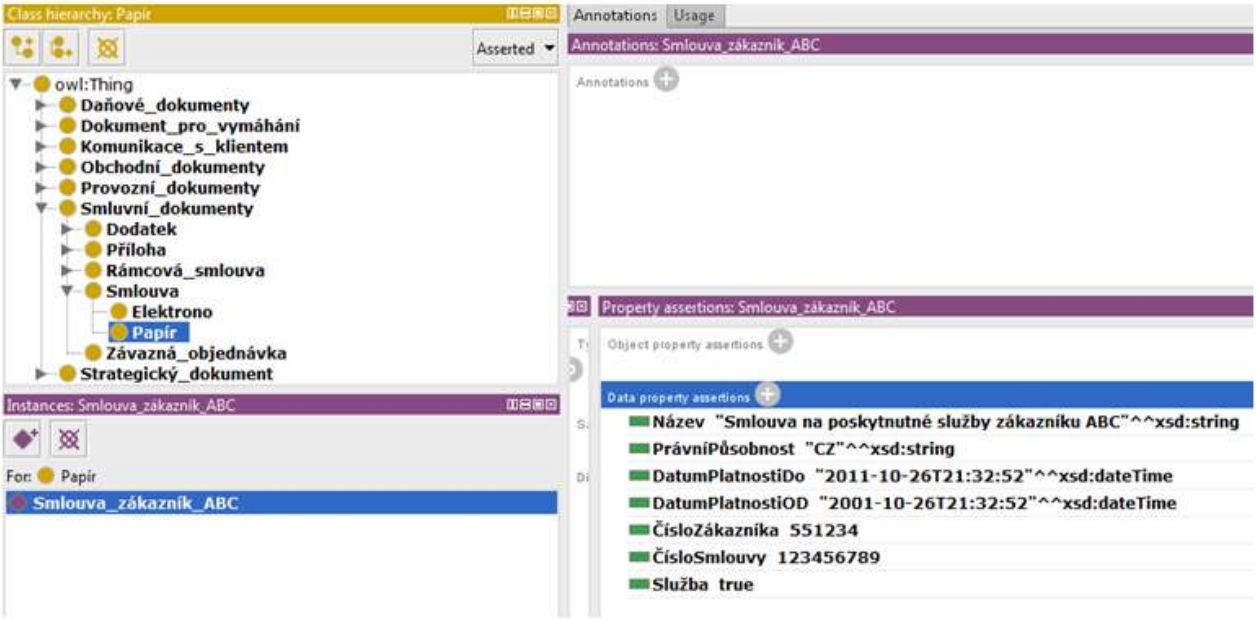
Odpovědi na kontrolní otázky nám pouze říkají to, zda byl splněn cíl projektu. Což není pro potřeby zákaznických dokumentů dostatečné, v rámci této metodiky na ně pohlížíme procesně, tzn. je nutné otestovat i to, že navržené změny nepovedeou k omezení/ k problémům při práci s dotčenými oblastmi dokumentů i pro jiné klíčové uživatele. Vracíme se zde ke kapitole, kde byly definovány dokumenty i s návazností na procesy a jednotlivé business pohledy na práci s dokumentem. Je tedy nutné ke každému dotčenému typu dokumentu otestovat všechny jeho způsoby použití (např. to, že se změní způsob snímání

[74] Osobně bych se u této funkcionality obávala, že místo předpokládaného přínosu, že uživatel dostane i dokumenty, které by pod hledaným termínem nebyly, neboť používají jinou terminologii, obávala opaku, bude zavalen informacemi – ještě větším počtem dokumentů, než by měl původně, kdy ho jen představa prohlížení odradí. Věřím, že tato komponenta může být pro specifický případy vhodná a přínosná, ale masovému použití bych se raději vyhnula.

dokumentu A pro potřeby vymáhání, nebude mít dopad na to, že ho jiná oddělení nenajdou). Tato kontrola by se měla dělat v obou fázích, tzn. jak ve fázi testování, tak ve fázi pilotního provozu, kde nedochází ke generování testovacích dat na produkci, ale je nutná komunikace se všemi zúčastněnými stranami (stakeholdery), zda nemají nějaké problémy, které jsou zapříčiněny nasazením této aktivity (tzn. zjistit, zda se s novým typem dokumentu / třídění při své běžné práci již setkali a nějak je to v práci neomezilo).

Ve fázi testování je nutné zaměřit se na přípravy testovacího vzorku dat, ve kterém by měly být zastoupeny všechny dokumenty podléhající změnám. Měly by být uloženy a vyhledány pod všemi variantami, které mohou nastat (tzn. uložit smlouvu o dílo pod metaúdaji A a B, ale i pod C, D, i když tato varianta nastává pouze v případě, že jde o VIP zákazníka apod.). Toto je obzvlášť nutné dodržet, dochází-li v rámci aktivity k migraci dat, jejichž metadata budou s velkou pravděpodobností upravována dle používaného třídění v cílovém úložišti.[75]

Pro první testování, zda je navržená klasifikace správná, můžeme použít prostředí, ve kterém se třídění navrhuje. Na obrázku 32 je v nástroji Protége založen objekt smlouva zákazníka ABC a k němu jsou přiřazeny metaúdaje s hodnotami a jejich datovým typem.



Obrázek 32: Příklad metadat u objektu zákaznické smlouvy (zdroj: autorka)

[75] Migrace dat z jiného úložiště s sebou nejspíše ponese i požadavek na změny ve stávajícím třídění, neboť se s velkou pravděpodobností nepovede namapovat stávající metadata a metadata z migrovaného úložiště v pomru1:1, zvl. pokud migrace dat byla spuštěna akvizicí/fúzí firmy (přece jen budou mít rozdílné druhy dokumentů, hlavně v provozní oblasti).

V případě, že se při modelování třídění používal jazyk OWL, je možné vyvinout jednoduchou aplikaci, ve které si uživatel bude moci simulovat vyhledávání či vkládání dokumentu. Něco podobného bylo vytvořeno pro projekt Pizza jako nadstavba nad aplikaci Prétegé:



Obrázek 33: Aplikace umožňující práci s navrženou ontologií pro běžného uživatele (The University of Manchester, 2013)

Další ověření, že navržené třídění je správné, je výběr zástupce všech stakeholderů (potřebujeme dvě skupiny), kterých se změny týkají, a provedeme s nimi tyto testy. První skupina zástupců bude testovat, nakolik je navržený systém intuitivní. Dostanou za úkol uložit dokumenty, které pro výkon své práce potřebují, následně budou požádáni i o vyhledávání dokumentů. Druhá skupina zástupců bude také ukládat a vyhledávat dokumenty, s tím rozdílem, že před akcí dostanou pokyny, jak s dokumentem pracovat. Na základě toho se blíže seznámí se nástrojem. Druhé testování ukáže, jak rychle si uživatel osvojí práci v novém nástroji, zda nejsou chyby ve třídění. V obou případech měříme chybovost a rychlost. Uživatelé by měli být schopni ve druhém případě správně uložit dokument v 90 % v daném časovém limitu (časový limit může být stanoven na základě času z reálného běhu nebo z časů první skupiny).[76]

[76] V ideálním případě se stejný test provede i na stávajícím prostředí a následně se mezi sebou porovnají. I když se na první pohled může zdát, že na stávajícím prostředí budou ukazatele lepší, neboť je pro uživatele prostředí známé, nemusí tomu tak být. Na druhé straně je nutné přihlédnout k rychlosti odezvy starého prostředí, kde bude mnohonásobně více uložených dat, což vyhledávání zpomalí.

Obsahy by měly být ukládány stejným způsobem, jako tomu bude na produkci (tzn. bude-li je do systému generovat a ukládat jiný systém, pak by to tyto systémy měly dělat již nyní). Pokud nedojde k shodnému uložení a provedou se pouze integrační funkční testy, pak se v rámci produkčního běhu setkáme s případem, že z pohledu funkčního vše zapadá, ale protože se generují špatné hodnoty (např. v rámci číselníku, nebo budou dodávána data, která neprojdou validací týkající se délky a formátu – typickým příkladem je telefoní číslo, které může být ukládáno ve zkrácené 9místné formě, nebo obsahovat kód země buď ve formátu +420 nebo 00420), dokumenty nemohou být uloženy nebo vyhledány pod správnými hodnotami.

V případě, že jsou data do systému generována jinou aplikací, je vhodné provést ověření, že hodnoty v metaúdajích se rovnají hodnotám i v těle uloženého dokumentu. Zjistí se tak, že mapování metaúdajových polí mezi systémy je provedeno korektně. Při ukládání uživatelem není tento krok zásadní, neboť předpokládáme, že uživatel ví, jakým způsobem dokument zaindexovat. Nicméně je toto vhodné provést i zde, neboť se tím odhalí případné problémy s nepochopením obsahů jednotlivých metaúdajů a způsobem jejich používání – na základě toho se pak může navržené třídění ještě revidovat, či to může být jeden z bodů, který je nutný zavést do dokumentace a upozornit na to uživatele při školení.

Při testování automatické klasifikace, fultextového vyhledávání či u migrací je vhodné u testovacího vzorku dat ověřit, že došlo ke správnému přiřazení termů, že se vrací dokumenty, které hledaný termín v těle dokumentu obsahují[77]. Nicméně bychom se měli v testech zaměřit i na dva identifikátory. Jsou to koeficient úplnosti a koeficient přesnosti.

Koeficient **úplnosti** je poměr mezi počtem nalezených relevantních záznamů k celkovému počtu relevantních dokumentů v databázi, kdežto koeficient **přesnosti** je poměr počtu nalezených relevantních záznamů k celkovému počtu vrácených záznamů. Koeficienty jsou vůči sobě ve vztahu nepřímé úměrnosti.

Koeficient úplnosti je vhodné sledovat například u migrovaných dokumentů, kdy pro daný informační dotaz nalezneme stejný odpovídající počet dokumentů před migrací (ve zdrojovém systému) i po migraci v novém systému (v cílovém systému). Během migrace dochází k různým transformacím, ke generování metaúdajů či jejich výmazu. Na základě této kontroly jsme schopni odhalit chyby, které manipulace s metaúdaji může přinést (například:

[77] V začátcích fultextového vyhledávání bylo možné setkat se s případy, že aplikace nebyly připraveny pro mezinárodní prostředí (neočekávali, že by se v psaném textu objevily i jiné znaky, než které jsou uvedeny v ASCII tabulce).

špatné transformace/vygenerované metaúdaje, problém s neviditelnými znaky u hodnot v metaúdaji, který běžný uživatel v GUI nevidí, ale které zapříčiní to, že hledaný dokument není zařazen do vráceného výsledku).

Koeficient přesnosti může odhalit chyby v navrženém třídění. V případě, že detailní rešeršní dotaz vrátí velké množství nerelevantních dokumentů, pak navržené třídění nemá správnou úroveň granuality (např. potřebujeme-li vrátit aktivní smlouvu u zákazníka ABC a ve výsledcích se krom aktivních smluv vrací i návrhy smluv, staré neaktivní smlouvy, pak je chyba v klasifikaci. Zde je nutné vyloučit to, že by zadaný rešeršní dotaz byl na obecné úrovni).

V rámci pilotního provozu bychom se kromě kontroly koeficientu přesnosti či splněných cílů měli zaměřit na to, jak je klasifikace využívána. Zda je navržená klasifikace pro uživatele předvídatelná, intuitivní, nedochází-li k nepochopení a špatnému používání metaúdajů, což by snížilo jejich vyhledávání – viz 9 pravidel, která by mělo třídění splňovat.

Opět bychom se měli ubezpečit, že navržená změna negativně neovlivnila některý ze stávajících procesů dokumentů podléhajících změně a že jsou pro všechny zainteresované strany dokumenty dostupné stejně, jako tomu bylo před změnou.

Mezi další způsoby, jak je možné ověřit, že navržená změna není krokem zpět, patří porovnat práce uživatele před a po změně. Doslova je tím míněno, kolik kliků musí pro vyhledání / uložení dokumentu v novém systému udělat. Pokud změna přináší větší uživatelský dyskomfort (např. nutné zadávat při ukládání hodnoty, které byly dříve automaticky generovány nebo pro získání dokumentu se místo jednoho rešeršního dotazu musí provést dva i více), měla by se odhalit příčina a následně provést minimálně zmírnění tohoto nekomfortu (např. dogenerování hodnot, přednastavení hodnot na určitou hodnotu – zvážit, zda je daný metaúdaj potřebný a není možné ho přestat používat, zjemnit kategorizaci, aby uživatel nemusel zadávat více rešeršních požadavků).

Při ověřování by se nemělo zapomínat ani na kontrolu, zda navržená klasifikace splňuje informační politiku a business strategie. Dále, že jsou využívány uznávané mezinárodní číselníky (např. kódy zemí, jazyka, …).

V neposlední řadě je nutné – jak při kontrole v rámci testů, tak v rámci pilotního projektu – provést kontrolu oprávnění nad dokumenty. Z pohledu klasifikace je důležité, aby dokument mohl vkládat pouze určený uživatel, změnu hodnot v metaúdaji opět pouze oprávněný uživatel (toto oprávnění je zbytečné ponechávat všem uživatelům, obyčejný

uživatel typu čtenář toto oprávnění nepotřebuje a je možné mu ho odebrat. Vyhneme se tak problémům s nechtěnými změnami, které jsou většinou zapříčiněny tím, že uživatel buď neví, na co ikona je, a nedopatřením ji použije, nebo je zvídavý a testuje, co mu systém povolí).

6.5 Provoz

Úspěšným ukončením pilotního provozu práce na zákaznické klasifikaci nekončí, právě naopak. Jak již bylo výše napsáno, mezi hlavní cíle klasifikace patří identifikovat obsah, zajistit jeho správu, pomoc při vyhledávání obsahu a sledovat jeho využívání. Z toho důvodu se musí nastavit kontrolní mechanismy, které zajistí, že vše výše popsané bude splněno. Tyto mechanizmy představíme níže.

V praxi dnešní doby zaznamenáváme vliv řídit a dodávat IT ve formě služby (což podporují rozřířené rámce používané v organizací ITIL a COBIT). Z toho důvodu jsou kontrolní mechanismy představeny ve čtyřech navržených službách (viz kapitoly níže začínající názvy ECM_01 až ECM_04), které mají zajistit základní cíle klasifikace (vymezené v kapitole Popis postupu tvorby firemního třídění). Podnik může tyto služby zavést jako součásti celofiremní služby technologie ECM, použít jen některé, či využít pouze některé vybrané metriky a začlenit je do stávajících služeb.

Než se začneme podrobněji věnovat popisu navržených služeb, je vhodné si představit jejich společné vlastnosti: při integraci služeb do stávajích procesů je nutné stanovit, kdo ponese zodpovědnost za to, aby služby byly dodávány v požadovaném čase a kvalitě. Tento typ pracovní činnosti mívá v kompetencích servisní manažer. Pokud tomu tak není, mělo by tuto jeho roli převzít oddělení, do jehož náplně spadá i informační management. Dále by každá ze služeb měla mít stanoven eskalační proces a zdokumentováno, jakými systémy a komponenty v IT je služba dodávána. S tím se pojí otázka dokumentace systému. Dokumenty (tabulky), které vznikly na základě této metodiky, se doporučuje uržovat v aktuální podobě, protože ulehčí analýzu při řešení chyb v rámci běžného provozu a při budoucím vývoji.

Pro každou službu by se měla stanovit matice odpovědností (RACI matice), jejíž příklad je uveden v tabulce 17. Je použita RACI matice shodně s COBITem 5.

Tabulka 17: Raci matice pro služby klasifikace (zdroj: autorka)

ECM	Role				
Aktivity	Vlastník dat	Aplikační správce	Servisní manažer	Klíčový uživatel	Informační manažer
Identifikovat obsah	I	R	A	C	I/C
Vyhledávání obsahu	I	R	A	C	I/C
Správa obsahu	I	R	A	C	I/C
Využívání obsahu	I	R	A	C	I/C

V poli aktivita se uvedou všechny úkoly, které jsou se službou spojeny, a určí se, jaký vztah má ta která role k dané aktivě. Používají se počáteční písmena z aglických termínů: **r** (responsible) – role, které na úkolech pracují, **a** (accountable) – role, které nesou odpovědnost za aktivitu jako celek, **c** (consulted) – role, se kterou je možné aktivitu konzultovat a **i** (informed) - role, která je informována o postupu a konečném výsledku.

Každá ze služeb by měla mít definovaného vlastníka, kterému se pravidelně reportuje, zda je služba dodávana v očekávané kvalitě (splňuje metriky, dostupnost,…). Nejspíše jím bude stejný zaměstnanec, který je zároveň vlastníkem dokumentů. Stavem služby se mohou zabývat také klíčoví businessmani, jež potřebují bezproblémový provoz pro podporu svých klíčových procesů.

Dále je službám nutné určit jejich prioritu z celofiremního pohledu (podstandard, standard, nadstandard, kritická), doba provozu, dostupnost a spolehlivost.

S klíčovými uživateli a vlastníky by měly probíhat pravidelné schůzky, jejichž cílem je získat zpětnou vazbu na poskytované služby (spokojenost, návrhy na změnu). V případě, že je technologie používání celofiremní a má hodně uživatelů, pak je vhodné pravidelnou spokojenost sledovat pomocí krátkého dotazníku.

6.5.1 Služby zajišťující běžný provoz

V podkapitolách níže jsou vymezeny čtyři základní služby, jejichž posláním je zajistit kvalitní podporu podnikového procesu nebo činnosti, které pro své fungování potřebují práci s dokumenty a metadaty.

6.5.1.1 ECM_01 Služba identifikující obsah

Cílem služby je zaručit, že bude možné dokument správně určit, tzn. bude mít správná metadata a správné lokační údaje k elektronické verzi obsahu. Službu budou poskytovat všechny komponenty, které zajišťuji ukládání do systému (ať už sychronně či asychronně, dávkově či po kusech). Při ukládání se musí zajistit, že nedojde k vygenerování a přiřazení nesprávných hodnot. Některé potenciální chyby je možné minimalizovat kontrolou správného formátu dat a používáním číselníku.

Nevýhodou číselníku je, že se musí udržovat v aktualizované podobě. Je tedy nutné najít způsob, jak toto zabezpečit, zvláště, pokud k aktualizaci dochází pravidelně v dávkové kontrole (kontrolovat, že pravidelný update probíhá, kontrolovat logy importu, že se nehromadí na vstupu, obsah, který není možné uložit z důvodu špatné hodnoty v číselníku).

Kromě validace dat na vstupu je nezbytné nastavit kontrolu validace dat nad databází. Například je možné do nějakého metaúdaje vkládat volný text, ale ukládají se do něj pouze číselné hodnoty v rámci určitého typu/procesu dokumentu, je žádoucí vytvořit report, který bude obsahovat informace o dokumentech, jež mají jiné hodnoty, než jsou číselné.[78] Pokud má nějaký metaúdaj začínat prefixem či končit sufixem, provádět pravidelně kontrolu, že tomu tak je.

Mít nastaven proces, jakým způsobem bude provedena reindexace dokumentů, které byly nedopatřením uloženy s nesprávnou hodnotou.

Pokud existuje automatické obohacování metadata (nad dokumenty můžou probíhat různé typy následného obohacování na základě nějaké spouštěcí aktivity) či automatická klasifikace, musí se sledovat, že dané komponenty nevykazují funkční chyby. K zásadní reindexaci dochází při migracích dokumentů nebo změně datového a klasifikačního schématu. Vždy je nutné provést kontrolu a mít připraven postup, jak vrátit zpět původní hodnoty, pokud změny byly provedeny špatně.

Vzhledem k tomu, že služba obsahuje i identifikaci obsahu v úložišti, je nutné sledovat, že nedochází místo na discích a je možné data ukládat. Některé aplikace umožňují kontrolu konzistence dat, jejímž cílem je zjistit, že uložená informace o lokaci dokumentu v databázi odpovídá uložení dokumentu na úložišti. Není-li tomu tak, provedou opravu.

[78] Tento případ je možné ošetřit i tak, že pro daný druh dokumentu nebude možné vložit do pole jiný znak než číslo.

Také by se mělo monitorovat, zda se v logu aplikace nevyskytuje velké množství dokumentů, které není možné zobrazit/stáhnout – u jednotlivých případů bude nejspíše při vyhledávání uživatelem použito jedinečné ID dokumentu, které bylo časem smazáno. Nicméně pokud se v logu začne objevovat velké množství případů, že není možné dokument najít/stáhnout, pak to signalizuje problém, který může být způsoben nepřístupností diskového pole, neoprávněným odmazáním dat apod.

Metriky služby:

- celková kapacita diskového pole, kolik z toho je volného místa;
- počet chybných dokumentů, které nebylo možné uložit z důvodu neexistence hodnoty v číselníku;
- počet chybně uložených dokumentů, které nemají číselnou podobu v metaúdaji, který umožňuje volně vkládat jakýkoliv text;
- počet dokumentů, které nebylo možné nalézt přes ID, externí ID;
- počet dokumentů, které nebylo možné zobrazit či stáhnout;
- pravidelná kontrola konzistence dat a počtu vad.

6.5.1.2 ECM_02 Služba zajišťující správu obsahu

Cílem služby je podpora dokumentu ve všech fázích jeho života, tzn. pro podporu všech potřebných klientských procesů zajišťuje adekvátní metadata.

Jednou z těchto funkcí je zadržení (hold) dokumentu proti nechtěnému výmazu. Pokud se v organizaci využívá, je potřeba zajistit kontrolní mechanizmy, které budou monitorovat, že komponent pracuje jak má (včas zajistí zadržení dokumentu, nebo se zadržení zruší).

Některé systémy umožní dokumenty tzv. **rezervovat**, což znamená, že dokument je zamknutý pro uživatele, který v něm aktuálně provádí změny. Stává se, že někdy tento příznak u dokumentu zůstane, i když uživatel dokument needituje. Je vhodné dlouhodobé rezervace obsahu sledovat a provádět uvolnění obsahu.

V rámci ECM_02 jsou monitorovány všechny funkce, které podporují proces skartace a archivace v organizaci. Je nutné kontrolovat, že obraz papírového dokumentu, který byl pořízen na vstupu, nese externí identifikátory pro lokaci papírového originálu. Dále, že

dokumenty, které mají nastaveno datum výmazu, jsou po uplynutí data ze systému skutečně odstraněny.[79]

Metriky služby:

- počet dokumentů ve stavu zadržení;
- počet dokumentů podléhajících skartaci;
- velikost uvolněného diskového pole po odmazání dokumentů po skartaci;
- počty dokumentů v jednotlivých dokumentových stavech (v procesu, vydáno, rezervováno, nahrazeno);
- pravidelné zálohování klasifikace.

6.5.1.3 ECM_03 Služba zajišťující vyhledávání obsahu

Služba zajišťující vyhledávání obsahu má za cíl zajistit kvalitu a funkčnost dodávané komponenty, která zajišťuje vyhledávání nad zákaznickým obsahem.

V rámci této služby by měla pravidelně probíhat školení pro uživatele, zvláště nové zaměstnance. Měly by existovat návody a příručky, jak s různými typy dokumentů pracovat, zvl. z pohledu vyhledávání dokumentů. V ideálním případě existují i návody ve formě tzv. kapesních návodů[80]. Měly by existovat personalizované návody relevantní pro jednotlivé typy klíčových uživatelů, neboť obecné návody jsou většinou pro běžné uživatele nepochopitelné. Těžko si představují, jakým způsobem mají pracovat se svým typem dokumentů[81] (proto by obchodník měl mít manuál s praktickými radami, jakým způsobem vyhledat smlouvy, návrhy nabídek…, kdežto účetní by ve své příručce nalezla, jakým způsobem vyhledat daňový doklad apod.). Tyto návody mohou být dostupné jak z intranetové stránky daného oddělení, tak například z vlastní intranetové stránky k ECM technologii, anebo mohou být přímo dostupné v systému. V ideálním případě ECM systém umožňuje mít vlastní návody k jednotlivým polím, která se při vyhledávání používají (např. po odkliknutí obrázku otazníku se uživatel doví, že toto pole má význam XYZ, je ukládáno ve formátu ABC). Tato nápověda je nápomocná zvláště u migrovaných dat, kdy je možné do ní dát rady jako: Do tohoto pole byly migrovány hodnoty z pole DFG ze systému RFT.

[79] Vzhledem k tomu, že externí identifikátor lokace papírového originálu může být pouze v metadatech dokumentu, je nutné před výmazem dokumentu z elektronického úložiště vytvořit seznam, na jehož základě bude následně vyhledán a skartován papírový originál.
[80] Většinou jde o malou kartičku ve formátu A5, max. A4, na které jsou popsány tipy, jak se systémem pracovat.
[81] Toto je pochopitelné, protože většinou neznají celistvý pohled na celou architekturu a unikají jim tedy souvislosti.

Uživatelé uvítají přítomnost kontaktní osoby či e-mailové adresy, na kterou se mohou v případě problémů s vyhledáváním obrátit o pomoc. U většiny případů uživatelé neví, jak je dokument klasifikován a nedokáží ho tedy najít, případně neznají obecné principy, jak dokumenty vyhledávat.[82]

Na základě toho, jakou odezvu systém poskytuje (jak dlouho trvá vyhledání, zda nedochází k time-outům), se podniknou kroky k nápravě (vytvoření indexu nad databází, zvýšit výkon databáze či serveru, nebo změnit klasifikaci, zajistit uživatelům, aby mohli zadat dotaz, který je více specifický, apod.). Svou roli tu hraje, zda jsou dokumenty ukládány na rychlé či pomalé disky, jsou-li dostupné v online režimu či offline.

Z pohledu vyhledávání je nutné sledovat v logu ty selekční dotazy, které končí sdělením, že pro zadaný výraz nebyl nalezen žádný objekt. V případě většího výskytu podobných informačních dotazů to může značit, že stávající klasifikace je zastaralá a je nutné ji rozšířit o nové termíny.

Log je zajímavý i z toho pohledu, jakým způsobem se v systému vyhledává – jaké jsou nejpoužívanější metaúdaje a v jakých skupinách, na základě toho je pak možné provést změny na úrovni databázových indexů, přednastavených vyhledávacích šablon.

Další způsob, jak se dovědět, že klasifikace je zastaralá, je sledování přírůstků dokumentů pro jednotlivé termíny nebo třídy. Pokud se nějaká třída/termín přestal využívat či naopak je pod něj ukládán velký počet obsahů, mělo by se zjistit, co je příčinou.

Metriky služby:

- počet školení,
- rychlost odezvy systému po zadání rešeršního dotazu,
- počet nedokončených řešeršních systémů, za které může aplikace sama (tzn. nebyly zrušeny uživatelem),
- průměrný počet vyhledaných záznamů na informační dotaz, rozdělený dle předmětové klasifikace zákaznického třídění,
- počet dokumentů pro dokumentové třídy.

6.5.1.4 ECM_04 Služba zjišťující používání obsahu

Tato služba má za cíl monitorovat, jakým způsobem je obsah používán.

[82] Neznají práci s boolean či proximitními operátory, práci se zástupnými znaky při vyhledávání, řazení /třídění výsledků na výstupu apod.

V rámci této služby může docházet k odmazávání provozních dokumentů, které mají krátkou dobu životnosti (např. informační dokument, který je vygenerován pro technického pracovníka a obsahuje informace o klientovi, zakázce a domluveném čase výjezdu). Podobně na tom jsou i neaktivní verze dokumentu (např. pracovní verze objednávek, nabídek, smluv, projektové dokumentace) či různé koncepty, které nebyly realizovány (nikdy se z nich nestala smlouva, objednávka,…). Tyto dokumenty mohou být po určité době ze systému bez problémů odstraněny. Pokud jsou nastaveny procesy odmazávání nad těmito obsahy, musí se zabepečit monitoring této komponenty, že pracuje korektně.

Podobně může fungovat i komponenta, která bude mít za úkol přenášet dokumenty, ke kterým se již aktivně nepřistupuje (nikdo je nevyhledává, nečte a neprovádí se změny nad metaúdaji) na levnější varianty úložišť. Dá se říct, že se provede jejich elektronická archivace (analogie k papírové formě, dokumenty jsou ukládány do spisovny) a čekají na dobu, kdy budou dle skartačního plánu ze systému trvale odstraněny. V rámci této služby by měly být stanoveny povinné údaje, které musí obsah mít, aby ho bylo možné korektně skartovat.[83]

V rámci této služby mohou být nastaveny různé typy reportů nad metaúdaji. Například pro potřeby obchodního oddělení se může pravidelně generovat seznam všech uložených naskenovaných zákaznických smluv, který je následně porovnán s reportem z CRM, se kterými zákazníky smlouva byla uzavřena, díky čemuž se odhalí chybějící, které smlouvy nepřišly k archivaci. Následně může být spuštěn proces, který zajistí získání smlouvy. Nebo může existovat report pro potřeby kontroly dodavatelské faktury u externě dodávané služby zajišťující digitalizaci dokumentů.

Do této služby spadá i oblast týkající se monitoringu s nulovým výsledkem ve vyhledávání, který byl popsán ve službě ECM_03.

Do služby spadá správa a monitoring komponent, které zajišťují zobrazení obsahů (pro různý formát obsahu může být použit jiný prohlížeč, např. obrázky mohou být otevírány v aplikaci, která je instalována na PC, PDF dokumenty v plug-inu v internetovém prohlížeči atd.).

Metriky služby:

- počet dokumentů určených k výmazu dle data splatnosti,

[83] Způsoby, jak se s řešením vyrovnat, mohou být různé, dokumenty mohou být přesunuty z rychlých na pomalé disky nebo z online na offlinový přístup. Dále může docházet k přesunu mezi úložišti, kdy jedno úložiště slouží pro běžnou práci s dokumenty (v rámci technologiec ECM se jedná o DMS komponentu), pokud se s dokumentem již aktivně nepracuje, může být přesunut jinam a po vypršení skartační doby se provede likvidace nebo trvalá archivace, což s sebou může nést přesun do úložiště typu archiv.

- počet reportů a jejich četnost,
- dostupnost aplikací, které umožňují zobrazení / stáhnutí obsahu,
- počet dokumentů migrovaných na levnější úložiště,
- počet odmazaných neaktivních verzí a velikost získaného úložného prostoru.

V organizaci mohou existovat i jiné specializované služby, kde jedna z komponent budou dokumenty. Většinou jsou takto sledovány kritické procesy (např. vygenerování a zpřístupnění zákaznických faktur na portále, v rámci toho bude mít technologie ECM stanovenou dobu, za jak dlouho je povinna vygenerované dokumenty uložit a následně je zobrazit na portálu).

6.5.2 Implementace změn

Tvorba zákaznické klasifikace je nekonečný proces, je nutné ověřovat její aktuálnost a zajistit včlenění nových metaúdajů (ať už je to vyvoláno změnami prodávané služby/produktu, fúzí či akvizicí, nebo vývojem terminálu a standardu v oboru, ve kterém se firma pohybuje). Požadavky na změnu mohou být vyvolány stakelhodermy, procesy a technologií.

Není vhodné provádět změny v již implementovaném třídění rychle a bez rozmyslu. Každá změna by měla být testována a mělo by se ověřit, že nedojde k narušejí nějakého firemního procesu, který s dokumenty pracuje.

Z toho důvodu je níže uvedena tabulka, která pomůže při evaluaci požadované změny a napomůže rozhodnutí, zda změnu implementovat a jakým způsobem.

Tabulka 18: Tabulka evalující dopady změny v třídění (zdroj: autorka)

Název dokumentu	***[ABC]***
Popis, co dokument obsahuje	…
Je upravován již stávající druh dokumentu, nebo jde o zcela nový druh (poddruh)?	
Je možné nový typ dokumentu klasifikovat jako jemu podobné obsahy? Pokud ano, jakému typu je podobný	
Jakým způsobem je dokument ukládán?	
Pod jakými metaúdaji bude ukládán?	
Je nutné změnit vkládací šablonu?	
Jakým způsobem bude dokument vyhledáván (návaznost na jinou aplikaci)?	
Pod jakými metaúdaji bude dokument vyhledáván?	
Je nutné změnit vyhledávací šablonu?	
Kdo bude vlastníkem tohoto typu dokumentu?	
Jaké bude mít metaúdaje z pohledu archivace a skartace?	
Bude pravidelně odmazáván?	
Návaznost na firemní procesy	
Způsob monitoringu, metriky kvality,…	
Klíčoví uživatelé, definovat oprávnění nad dokumentem	

Prostřednictvím této tabulky je možné nastartovat dialog nad požadovanými změnami, na jehož konci bude rozhodnuto, jakým způsobem bude obsah klasifikován a zda bude či nebude nutné provádět v stávajícím třídění nějaké změny.

Po změně se nesmí zapomenout aktualizovat uživatelská, administrativní a analytická dokumentace (např. datový model, návody pro uživatele,…).

V kapitole jsme představili metodiku klasifikace zákaznických dokumentů ve firmě. Vyjmenovali jsme a specifikovali kroky nezbytné při návrhu, implementaci, testování klasifikace. V oblasti provozu byly blíže popsány služby, které mají zajistit bezproblémový chod klasifikace. V rámci fáze tvorby obecného návrhu třídění byla blíže popsána specifika týkající se tvorby předmětového selekčního jazyka.

7 Závěr

Diplomová práce se zabývá tříděním podnikového obsahu a jeho dopady na ukládání, vyhledávání, správu a podporu celého životního cyklu.

Jak je z kapitoly, zabývající se analýzou současného stavu patrné, podniky se v dnešní době potýkají s problémy dotýkající se správy dokumentů (duplikace obsahu, problém s vyhledáváním relevantního dokumentu, zajištění zákonných požadavků, …). Jedním ze základních prvků, které mohou pomoci s odstraněním těchto problémů, je používání kvalitního třídění.

Z toho důvodu bylo hlavním cílem této diplomové práce navržení metodiky klasifikace zákaznických dokumentů v podniku.

Tento hlavní cíl jsem v rámci kapitoly 6 naplnila tak, že jsem definovala a popsala:

- samotný proces
- fázi plánování, jejíž součástí je stanovení rozsahu aktivity a cílů, výpočet návratnosti investic, stanovení metrik, definice pracovního týmu
- tvorbu obecného návrhu klasifikace, včetně definice základních metadat
- fázi implementace, jejíž součástí je tvorba technického provedení a způsob kontroly navrženého třídění
- fázi provozu, která obsahuje návrh služeb a postup při implementaci změn

V rámci podkapitoly 6.3.3 jsem splnila dílčí cíl práce, vytvořila předmětové selekční jazyky pro zákaznické dokumenty v podniku, které jsou součástí metodiky. Oproti původnímu záměru jsem se odchýlila tím, že jsem nevytvořila pouze jeden, ale několik druhů (řízený seznam termínů, taxonomii, fasetové třídění, tezaurus, ontologii). Důvodem je umožnění metodiky variability použití, zároveň jsou na jednotlivých druzích demonstrovány silné a slabé stránky, což umožňuje podniku výběr nejvhodnějšího druhu třídění. Při tvorbě selekčních jazyků jsem zjistila, že je nemožné oddělit formu dokumentu od obsahu.

Pro potřeby tvorby metodiky jsem mimo jiné:

- porovnala standardy specifikující metadata pro dokumenty (příloha č. 2)
- porovnala standardy vymezující stupně utajovaných informací (příloha č. 4)
- navrhla jsem a zdůvodnila procesní mapu dokumentů (kapitola 6.3.1)

- porovnala jsem produkty od tří dodavatelů z pohledu použitelnosti pro implementaci navrženého třídění

V rámci některých části metodiky je i naznačen myšlenkový poustup, způsob, jak se k návrhu dospělo. Důvodem, proč v metodice není pouze výsledek, je, že vzhledem k rozsahu použití metodiky nemusí být výsledek pro podnik použitelný, ale může vyjít z informací, na jejichž základě výsledek vznikl a sám je rozvíjel dle své potřeby (např. kapitola 6.3.3, kapitola 6.3.2).

Pojetí práce je svým rozsahem a obsahem unikátní jak v českém jazyce, tak v anglické literatuře. Nepodařilo se mi v rámci rešerše dohledat podobně zaměřenou práci, většina odborných prací, obsahující postup tvorby třídění, zůstávala na obecné úrovni celého podniku nebo detailně rozebírala jednotlivé techniky (např. využití pravděpodobností metody při automatické klasifikaci dokumentů). Práce vychází z informačně-knihovnické vědy, z ní používá teorii katalogizace a obohacuje ji o možnosti využití technologie ECM z oboru informační technologie.

Navržená metodika má napomoci firmám při tvorbě a správě jejich třídění dokumentů. Při vytváření metodiky byl kladen důraz na její využití v praxi, což se projevuje v její struktuře nebo používané terminologii.

Pro další vědecké zkoumání mám následující náměty: upravit či navrhnout metodiku pro jiný druh podnikového obsahu (např. dokumenty personálního oddělení, dokumenty v rámci znalostní báze), využití metodiky v praxi, možnosti využití předmětového selekčního jazyka BSC mimo komponentu správu záznamů v rámci ECM, možnosti použití ontologie s jiným typem selekčního jazyka a jejích vliv na zvýšení kvality klasifikace podnikového obsahu, rozšíření metodiky o nástroje automatické klasifikace a o metody začleňování jiné klasifikace do stávající.

Při psaní této práce jsem se v odborné literatuře často setkávala s názorem, že tvorba třídění je z části věda a z části umění. Zpočátku jsem k tomuto názoru byla spíše skeptická, a to díky předchozímu studiu, kdy jsem se učila o klasifikaci vědy s následnou praxí. Čím déle jsem se tématem zaobírala v širším významu (nad rámec reálného podniku), jsem zjistila, že způsobů, jak k této problematice přistupovat, uchopit ji a dosáhnout cíle, je mnoho a jejich správné propojení v daném místě, čase a v podniku se spíše než vědě začíná podobat umění. K této činnosti je zapotřebí mít pro tuto práci zvýšený cit, zvláštní vlohy, ne každý, i když se naučí teorii, tak bude dobrým v této činnosti.

8 Terminologický slovník

Termín	Zkratka	Význam
Business Classification Scheme	BSC	klasifikační schéma v hierarchické struktuře, které má na hlavní úrovni obchodní funkce a na nižších se rozpadá na aktivity podniku (Plymouth University2017)
Business Process Management	BPM	nástroj, který umožňuje návrh, vývoj a správu procesů v podniku. (Arsanjani 2015)
Content Management Interoperability Services	CMIS	je technickou specifikací pro komunikaci s ECM úložištěm pomocí základní sady obecných webových služeb (Švík 2009)
data		symboly nebo znaky, které představují nezpracovaná fakta nebo čísla, tvoří základ informací (ARMA International 2007)
dokument		zaznamenaná informace nebo objekt, se kterým je možné zacházet jako s celkem (ISO 15489-1, 2001)
Electronic Records Management System	ERMS	elektronická správa elektronických a neelektronických záznamů využívající počítačové vybavení a software v souladu s uznávanými zásadami a postupy správy dokumentů (AIIM a ARMA, 2004)
Enterprise content management	ECM	strategie, metody a nástroje používané k zachycení, správě, uložení, uchovávání a dodávání obsahu a dokumentů vztahujících se k organizačním procesům. Nástroje a strategie ECM umožňují správu nestrukturovaných informací dané organizace bez ohledu na jejich umístění (AIIM 2011)
folksonomie		schéma kategorizace obsahu internetu, kdy uživatelé označují obsah WWW stránek popisky (tagy) podle svého výběru za účelem utřídění těchto stránek a jejich rychlejšího vyhledávání(Havlová 2003)
indexace		proces vyjádření obsahu dokumentu pomocí prvků selekčního jazyka, obvykle s cílem umožnit zpětné vyhledávání (Balíková 2003b)
informace		zpráva podávaná ústně, písemně nebo jiným způsobem, s pomocí signálů nebo technických prostředků (Malá československá encyklopedie 1989, s. 35)
katalogizace		jmenný a věcný popis dokumentů, někdy chápaný úžeji pouze jako jmenný popis dokumentů, jindy obecněji jako tvorba katalogizačních záznamů pro různé druhy katalogů a jejich organizace. V době budování lístkových katalogů nauka o knihovních katalozích a o metodách jejich budování včetně jejich organizace. (Vodičková 2003)
klasifikace		uspořádání pojmů nebo jiných entit do tříd a dalších skupin (podtříd apod.) za účelem vyjádření jejich sémantických vztahů. Třídy a další skupiny jsou obvykle vyjádřeny notací. Nebo přidělování notací (znaků tříd) klasifikačního systému za účelem vyjádření obsahu dokumentu (Balíková 2003c)

klíčová slova		významové slovo nebo slovní spojení, obvykle ve formalizované podobě, vybrané z názvu nebo textu dokumentu a sloužící jako věcný selekční údaj (Balíková 2003d)
metadata / metaúdaj		strukturované nebo částečně strukturované informace, které umožňují tvorbu, správu a používání záznamů v čase v rámci domény nebo i mezi doménami (ISO 23081-2, 2009)
Modelové požadavky na správu elektronických dokumentů	MoReq2	popisuje modelové požadavky pro správu elektronických záznamů se zaměřením na funkční požadavky systémů elektronické spisové služby (MoReq2 specification 2008)
obsah		souhrn všech vlastností, které vyjadřují charakter předmětu nebo pojmu (Komenského slovník naučný 1938, s. 350)
ontologie		základní část filozofie, věnována otázce bytí jako takového a nejobecnějším kategoriím jsoucna (Anon. 1999, s. 464)
prohlížení		základní způsob pohybu v informačním prostoru služby WWW prohlížením obsahu webových stránek s využitím hypertextových odkazů. Je protikladem a doplňkem k přímému vyhledávání informací (Sklenák 2003)
ROT obsah		je to obsah, který je redundantní, zastaralý a triviální (Mancini 2011)
Simple Knowledge Organization System	SKOS	standard pro sdílení, publikování a reprezentaci taxonomie, heslářů, tezaurů, řízených slovníků nebo klasifikačních schémat (Riley)
taxonomie		druh řízeného slovníku, který má hierarchii (nadřazený termín/podřazený termín), ale ne nutně související vazby, jako je tomu u standardního tezauru (Bhat 2004)
třída		je množinou jedinců určitého typu (Horridge 2011)
věcný popis		proces tvorby věcných selekčních údajů a jejich přidělování dokumentům. Nauka o věcných katalozích a o metodách jejich tvorby. (Balíková 2003f)
Web Ontology Language	OWL	jazyk umožňující modelovat ontologii (Klempa 2006)
zákazník		oddělení v rámci firmy, která stanovují požadavek na proces (účetnictví) a přijímají výstup procesu (finanční výkaz) od jednoho nebo více interních nebo externích dodavatelů (Customer 2007)
záznam		informace, které společnost nebo jednotlivec vytváří, přijímá a uchovává jako důkazy a informace, a to na základě právních povinností nebo během obchodních transakcí (ISO 15489-1, 2001)
znalost		schopnost člověka nebo jakéhokoli jiného inteligentního systému uchovávat, komunikovat a zpracovávat informace do systematicky a hierarchicky uspořádaných znalostních struktur. Znalost je charakterizována schopností abstrakce a generalizace dat a informací (Jonák 2003)

9 Seznam literatury a použitých zdrojů

9 Ingredients to a Successful Taxonomy [online], 2012. Silver Spring: AIIM, 6 s. [cit. 2017-04-17]. How-to Guide. Dostupné z: http://info.aiim.org/successful-taxonomy-nine-keys-to-success?_ga=1.256573640.2103637144.1492462039

ABDUS, Sattar Chaudhry, 2010. Assessment of taxonomy building tools. *The Electronic Library* [online]. **28**(6), 769-788. ISSN 02640473.

AIIM a ARMA, 2004. ANSI/AIIM/ARMA TR48-2004, *Framework for Integration of Electronic Document Management Systems and Electronic Records Management Systems*. Silver Spring: AIIM.

AIIM MS60, *Standard - Electronic folder interchange datastream*, 1996. Silver Spring: AIIM.

AIIM TR40, *Technical Report – Suggested index fields for documents in electronic image management (EIM) environments*, 1995. Silver Spring: AIIM, 15 s.

AIIM Training Programs: How to Use Metadata and Taxonomies to Improve Access to Information, Silver Spring: AIIM, 23 s.

AIIM, 2011. Your Information Management and Collaboration Ressource [online]. Silver Spring: AIIM [cit. 2011-03-23]. Dostupné z: http://www.aiim.org/What-is-ECM-Enterprise-ContentManagement

AIIM, 2016. Toolkit: How to Determine the Right Type of Taxonomy for Your Business, 2016. In: *AIIM: The Global Community of Information Professionals* [online]. Silver Spring: AIIM [cit. 2017-04-17]. Dostupné z: http://www.aiim.org/Resources/Toolkits/How-to-Determine-Right-Taxonomy

AIIM, 2017a. Toolkit: How to Develop a Search Strategy and Support Team. In: *AIIM: The Global Community of Information Professionals* [online]. Silver Spring: AIIM [cit. 2017-04-17]. Dostupné z: http://www.aiim.org/Resources/Toolkits/How-to-Develop-Search-Strategy

AIIM, 2017b. What is Enterprise Search. In: *AIIM Glosary* [online]. Silver Spring: AIIM [cit. 2016-03-17]. Dostupné z: http://www.aiim.org/What-is-Enterprise-Search

AIIM/ARMA TR 48-2006, *Revised Framework for Integration of Electronic Document Management Systems and Electronic Records Management Systems*. Washington: ANSI.

Anon., 1999. Ontologie. In: *Všeobecná encyklopedie v osmi svazcích*. Praha: Diderot. Encyklopedie Diderot. ISBN 80-902-5557-4.

ARMA International, 2007. data. In Glossary of Records and Information Management Terms [online]. [s.l.] : ARMA International [cit. 2011-03-22]. Dostupné z: http://www.arma.org/standards/glossary/

ARSANJANI, Ali, Nakul BHARADE, Magnus BORGENSTRAND, Philipp SCHUME, J. Keith WOOD a Vyacheslav ZHELTONOGOV, 2015. *Business Process Management Design Guide Using IBM Business Process Manager* [online]. New York: IBM [cit. 2017-04-17]. RedBooks. Dostupné z: http://www.redbooks.ibm.com/redbooks/pdfs/sg248282.pdf

Atribut, 2005-2017. In: *ABZ.cz: slovník cizích slov* [online]. Ostrava : ABZ knihy [cit. 2016-06-27]. Dostupné z: http://slovnik-cizich-slov.abz.cz/web.php/slovo/atribut

Australian Government Recordkeeping Metadata Standard: (AGRkMS), 2015 [online]. Version 2.2. Canberra: National Archives of Australia [cit. 2017-04-17]. ISBN 978-1-922209-06-1. Dostupné z: http://www.naa.gov.au/Images/AGRkMS-Version-2.2-June-2015_tcm16-93990.pdf

Australian Government security classification system, 2015. In: *The Protective Security Policy Framework: PSPF* [online]. Barton: Attorney-general's Department [cit. 2017-04-17]. Dostupné z: https://www.protectivesecurity.gov.au/informationsecurity/Pages/AustralianGovernmentSecurityClassificationSystem.aspx

BALÍKOVÁ, Marie, 2003a. Deskriptorový selekční jazyk. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna ČR, 2003- [cit. 2016-06-27]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000001516&local_base=KTD

BALÍKOVÁ, Marie, 2003b. Indexace. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna ČR, 2003- [cit. 2016-06-27]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000001549&local_base=KTD

Balíková, Marie, 2003c. Klasifikace. In: KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV) [online]. Praha : Národní knihovna ČR, 2003- [cit. 2017-04-18]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000001558&local_base=KTD.

BALÍKOVÁ, Marie, 2003d. Klíčové slovo. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna ČR, 2003- [cit. 2016-06-27]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000001563&local_base=KTD

BALÍKOVÁ, Marie, 2003e. Postkoordinovaný selekční jazyk. In: KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV) [online]. Praha : Národní knihovna ČR, 2003- [cit. 2017-04-18]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000001606&local_base=KTD.

BALÍKOVÁ, Marie, 2003f. Věcná katalogizace. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna ČR, 2003- [cit. 2016-06-27]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000001657&local_base=KTD

BALÍKOVÁ, Marie, 2003g. Věcné pořádání informací. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna ČR, 2003- [cit. 2016-06-27]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000001658&local_base=KTD

BAYONA-ORÉ, Sussy, Jose A. CALVO-MANZANO, Gonzalo CUEVAS a Tomas SAN-FELIU, 2014. Critical success factors taxonomy for software process deployment. *Software Quality Journal* [online]. 22(1), 21-48 [cit. 2017-04-17]. DOI: 10.1007/s11219-012-9190-y. ISSN 09639314. Dostupné z: http://link.springer.com/10.1007/s11219-012-9190-y

BHAT, M. H. 2014. Development and deplyoment of taxonomies. In: *Information Studies*. vol. 20, no. 2, pp. 97-111. Library, Information Science & Technology Abstracts, EBSCO*host*, viewed 10 April 2017.

BHAT, M. H. and S. M. Shafi, 2014. Taxonomies in knowledge organisation – Need, description and benefits. In: *Annals Of Library & Information Studies*. **61**(2), s. 102-111. Library, Information Science & Technology Abstracts, EBSCO*host*, viewed 10 April 2017.

BRATKOVÁ, Eva a Helena KUČEROVÁ, 2015. K otázkám metadatového popisu systémů organizace znalostí. Knihovna: knihovnická revue [online]., **26**(1), s. 5-36 [cit. 2017-04-09]. Dostupný

z WWW: <http://knihovnarevue.nkp.cz/archiv/2015-01/recenzovane-prispevky/k-otazkam-metadatoveho-popisu-systemu-organizace-znalosti>. ISSN 1801-3252.

BRATKOVÁ, Eva a Helena KUČEROVÁ. Systémy organizace znalostí a jejich typologie. *Knihovna* [online]. **25**(2), s. 5-29 [cit. 2017-04-17]. Dostupný z WWW: <http://knihovna.nkp.cz/knihovna142/142005.htm>. ISSN 1801-3252.

BRIDGES, Jeffrey D, 2007. Taking ECM From Concept to Reality. Information Management Journal [online]. **41**(6), s. 30-32,34,36,39. ISSN 15352897.

BRIOLA, Daniela, Alessandro AMICONE a Dante LAUDISA, 2013. *Ontologies in industrial Enterprise Content Management Systems: the EC2M Project*. IARIA. ISBN 978-1-61208-273-8. Dostupné také z: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.678.8704

CADENCE GROUP, 2008. Taxonomies: The Backbone of Enterprise Content Management. *New Images*. Gaithersburg,: ARMA, **20**(3), 1-6.

CISCO, S. & W. JACKSON, 2005. Creating Order out of Chaos with Taxonomies. *Information Management Journal*. **39**(3), s. 44-50. Academic Search Complete, EBSCO*host*, viewed 10 April 2017.

CISCO, Susan a Wanda K. JACKSON, 2005. Creating Order out of Chaos with Taxonomies. *Information Management Journal*. 2010, **39**(3), s. 44-50. ISSN 1535-2897.

CONNELLY, J. 2011. Shedding Light on Building and Implementing Successful Taxonomies. In: *Information Management Journal*. **45**(3), s. 42-47. Academic Search Complete, EBSCO*host*, viewed 10 April 2017.

CONNELLY, James, 2011. Shedding Light on Building and Implementing Successful Taxonomies. *Information Management*. **45**(3), s. 42-47,51.

CRAVEN, Tim, 2008. Thesaurus Construction. In: *Tim Craven* [online]. London: The University of Western Ontario [cit. 2017-04-17]. Dostupné z: http://publish.uwo.ca/~craven/677/thesaur/main00.htm

Customer, 2007. In: *BusinessDictionary.com* [online]. Washington: WebFinance [cit. 2016-06-27]. Dostupné z: http://www.businessdictionary.com/definition/customer.html#ixzz4CioWmzDK

ČESKÁ REPUBLIKA, Obchodní zákoník, 1991. In: *Sbírka zákonů*. Praha: Ministerstvo vnitra, ročník 1991, částka 98, číslo 513. Dostupné také z: http://aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=2510

ČESKÁ REPUBLIKA, Zákon o archivnictví a spisové službě a o změně některých zákonů, 2004. In: *Sbírka zákonů*. Praha: Ministerstvo vnitra, ročník 2004, částka 173, číslo 499. Dostupné také z: http://aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=4477

ČESKÁ REPUBLIKA, Zákon o dani z přidané hodnoty, 2004. In: *Sbírka zákonů*. Praha: Ministerstvo vnitra, ročník 2004, částka 78, číslo 235. Dostupné také z: http://aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=4382

ČESKÁ REPUBLIKA, Zákon o ochraně osobních údajů a o změně některých zákonů, 2000. In: *Sbírka zákonů*. Praha: Ministerstvo vnitra, ročník 2000, částka 32, číslo 101. Dostupné také z: http://aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=3420

ČESKÁ REPUBLIKA, Zákon o ochraně utajovaných informací a o bezpečnostní způsobilosti, 2005. In: *Sbírka zákonů*. Praha: Ministerstvo vnitra, ročník 2005, částka 143, číslo 412. Dostupné také z: http://aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=4741

ČESKÁ REPUBLIKA, Zákon o účetnictví, 1991. In: *Sbírka zákonů*. Praha: Ministerstvo vnitra, ročník 1991, částka 107, číslo 563. Dostupné také z: http://aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=2519

ČESKÁ REPUBLIKA, *Zákon občanský zákoník*, 2012. In: . Praha: Ministerstvo vnitra, ročník 2012, částka 33, číslo 89, s. 1368. Dostupné také z: http://aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=6144

DCMI Home: Dublin Core Metadata Initiative [online], c1995-2017. DCMI [cit. 2017-04-17]. Dostupné z: http://dublincore.org/

DELPHI GROUP, 2002. *Taxonomy and Content Classification: Market Milestone Report* [online]. Andover: Delphi Group [cit. 2017-04-17]. Dostupné z: http://lsdis.cs.uga.edu/SemanticEnterprise/Delphi_LingoMotorfinal.pdf

Deming Cycle: The Wheel of Continuous Improvement, 2009. In: *Total Quality Management: Just another WordPress.com weblog* [online]. WordPress [cit. 2017-04-17]. Dostupné z: https://totalqualitymanagement.wordpress.com/2009/02/25/deming-cycle-the-wheel-of-continuous-improvement/

DLM Forum [online], c2017. [cit. 2017-04-17]. Dostupné z: http://www.dlmforum.eu/

Document, 2016. In: *DBpedia Mappings* [online]. [cit. 2017-04-17]. Dostupné z: http://mappings.dbpedia.org/server/ontology/classes/Document

DOČKALOVÁ, Pavlína, 2017. Stupně utajení podle normy DIN 66399. In: *Office2000.cz: Modrerní kancelář bez starostí* [online]. Moderní kancelář bez starostí [cit. 2017-04-17]. Dostupné z: https://www.office2000.cz/clanek/55-stupne-utajeni-podle-normy-din-66399

DoD 5015.02: Electronic Records Management Software Applications Design Criteria Standard, 2007. Arlington: Department of Defense. Dostupné také z: http://www.dtic.mil/whs/directives/corres/pdf/501502std.pdf

DOFKOVÁ, Jitka, 2004. Ranganathanovo dvojtečkové třídění a jeho uplatnění v současné knihovnické praxi. KIVI: Knihovnictví a Informační Věda Informuje [online]. 2004[cit. 2016-06-20]. ISSN 1214-7265. Dostupné z: http://www.phil.muni.cz/kivi/clanky.php?cl=41

DOLEČEK, Marek, 2014. Pohledávky. In: *BusinessInfo.cz: Oficiální portál pro podnikání a export* [online]. Praha: CzechTrade [cit. 2017-04-16]. Dostupné z: http://www.businessinfo.cz/cs/clanky/pohledavky-ppbi-51112.html#!&chapter=5

Dublin Core: Czech homepage [online], 2006. Brno: Masarykova univerzita [cit. 2017-04-17]. Dostupné z: http://webserver.ics.muni.cz/dublin_core/index.html

DUCHARME, Bob, 2011. Improve your taxonomy management using the W3C SKOS standard. In: *IBM developerWorks: IBM´s resource for developers and IT professionals* [online]. IBM [cit. 2017-04-16]. Dostupné z: https://www.ibm.com/developerworks/library/x-skostaxonomy/

Email Protective Marking Standard for the Australian Government, 2012. version 2012.3. AGIMO, 32 s. ISBN 978-1-922096-05-0. Dostupné také z: http://www.finance.gov.au/files/2012/04/EPMS2012.3.pdf

ERSKINE, Richard W., c2009. Demystifying Taxonomies: Understanding taxonomies and their use in content management. In: *DELL EMC* [online]. EMC: Massachusetts [cit. 2017-04-16]. Dostupné z: https://www.emc.com/collateral/emc-perspective/h4736-demystifying-taxonomies-ep.pdf

FAGAN, Jody Condit, 2010. Usability Studies of Faceted Browsing: A Literature Review. *Information Technology and Libraries*. 2010, **29**(2), s. 58-66. ISSN 07309295.

FILOUŠOVÁ Blanka, 2002. MDT jako univerzální pořádací soustava. *KIVI* [online]. 2002, [cit. 2016-06-19]. Dostupné z: http://www.phil.muni.cz/kivi/clanky.php?cl=28&akce=t

FORQUER, Bill, Peter JELINSKI a Tom JENKINS, 2005. *Enterprise Content Management Solutions: What You Need to Know*. Waterloo: pen Text Corporation. ISBN 0973066261.

Göker, A., & Davies, J., 2009. *Information retrieval : searching in the 21st century.* Chichester U.K.: John Wiley & Sons, Inc. 2009.

Havlová, Jaroslava, 2003. folksonomie. In: KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV) [online]. Praha : Národní knihovna ČR, 2003- [cit. 2017-04-18]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000014605&local_base=KTD.

HAYSTAC, 2014. *The Unified Information Governance Imperative: A Pro-Active Solution to Managing Unstructured information*. 1. AIIM: Silver Spring. Dostupné z: http://www.aiim.org/Research-and-Publications/Research/White-Papers/Haystac_Clustering

HEDDEN, Heather, 2010. Taxonomies and controlled vocabularies best practices for metadata. *Journal of Digital Asset Management* [online]. **6**(5), s. 279-284. ISSN 17436540.

HEDDEN, Heather, 2016a. *The accidental taxonomist*. Second edition. New Jersey: Information Today, 469 s. ISBN 9781573875288.

HEDDEN, Heather, 2016b. The Accidental Taxonomist: Free Taxonomy Management Software. In: *The Accidental Taxonomist* [online]. [cit. 2017-04-16]. Dostupné z: http://accidental-taxonomist.blogspot.cz/2016/02/free-taxonomy-management-software.html

HLAVA, Marjorie M.K., 2014. The Taxobook: Applications, Implementation, and Integration in Search. *Synthesis Lectures on Information Concepts, Retrieval, and Services* [online]. **6**(4), 1-156 [cit. 2017-04-17]. DOI: 10.2200/S00604ED1V03Y201410ICR037. ISSN 1947945x. Dostupné z: http://www.morganclaypool.com/doi/abs/10.2200/S00604ED1V03Y201410ICR037

HLAVENKA, Jiří, 1997. *Výkladový slovník výpočetní techniky a komunikací: 5500 pojmů z oblasti výpočetní techniky : přes 7000 křížových vazeb : výklad anglických a českých odborných pojmů*. 3. vyd. Praha: Computer Press. ISBN 8072260235.

HORRIDGE, Matthew, 2011. *A Practical Guide To Building OWL Ontologies: Using Protege 4 and CO-ODE Tools* [online]. Edition 1.3. Manchester: University Of Manchester [cit. 2017-04-17]. Dostupné z: http://mowl-power.cs.man.ac.uk/protegeowltutorial/resources/ProtegeOWLTutorialP4_v1_3.pdf

How to Develop Taxonomies: Support Navigation, Information Discovery, and Findability [online], 2012. Silver Spring: AIIM [cit. 2017-04-17]. Dostupné z: http://info.aiim.org/how-to-create-taxonomies-to-support-navigation-and-discovery?_ga=1.188919016.2103637144.1492462039

IBM Knowledge Center: IBM FileNet P8, Version 5.2, c2014[online]. IBM [cit. 2017-04-17]. Dostupné z: https://www.ibm.com/support/knowledgecenter/SSNW2F_5.2.0

IBM, c2014. Choice lists: IBM FileNet P8, Version 5.2. In: *IBM Knowledge Center* [online]. IBM [cit. 2017-04-17]. Dostupné z: https://www.ibm.com/support/knowledgecenter/SSNW2F_5.2.0/com.ibm.p8.ce.admin.tasks.doc/choicelists/chl_concepts_choice_lists.htm

Intelligence Community Authorized Classification and Control Markings: CAPCO, 2011. Washington: Office of The Director of National Intelligence. Dostupné také z: https://fas.org/sgp/othergov/intel/capco_reg_v5-1.pdf

ISBD Area 0: Content Form and Media Type Area, 2009. Haag: IFLA. Dostupné také z: https://www.ifla.org/files/assets/cataloguing/isbd/area-0_2009.pdf

ISO 15489-1, *Information and documentation - Records management: Part 1: General*, 2001. Geneva: ISO, 19 s.

ISO 15489-2, *Information and documentation - Records management: Part 2: Guidelines*, 2001. Geneva: ISO, 39 s.

ISO 15836-1, *Information and documentation - The Dublin Core metadata element set*, 2009. Second edition. Geneva: ISO.

ISO 23081-1, *Information and documentation - Records management processes -Metadata for records: Part 1: Principles*, 2006. Geneva: ISO, 20 s.

ISO 23081-2, *Information and documentation - Records management processes -Metadata for records: Part 2: Conceptual and implementation issues*, 2009. Geneva: ISO, 33 s.

ISO/TC46/SC11N800R: Building a metadata schema – where to start, 2008. Revision. Baltimore: NISO. Dostupné také z: http://www.niso.org/apps/group_public/download.php/5271/N800R1_Where_to_start_advice_on_creating_a_metadata_schema.pdf

JONÁK, Zdeněk, 2003. znalost. In: KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV) [online]. Praha : Národní knihovna ČR, 2003- [cit. 2017-04-17]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000498&local_base=KTD.

KLEMPA, Tomáš, 2006. Opis jazyka OWL pre reprezentáciu ontológií. In: *Ivan Kapustík* [online]. Bratislava: Ústav informatiky a softvérového inžinierstva [cit. 2017-04-17]. Dostupné z: http://www2.fiit.stuba.sk/~kapustik/ZS/Clanky0607/klempa/index.html

Komenského slovník naučný : 7. díl, Kvášňovice-Mildschuh, 1938. Praha : Nakl. a vyd. Komenského slov. nauč., 1938. 607 s.

KUČEROVÁ ZRÁLÍKOVÁ, Václava, 2011. *Současné možnosti řešení správy a oběhu dokumentů ve firmě*. Praha. Diplomová práce. Univerzita Karlova, Filozofická fakulta, Ústav informačních studií a knihovnictví.

KUČEROVÁ, Hana, 2011. *Vyšší odborná škola informačních služeb* [online]. c2011 [cit. 2017-04-17]. Definice informace. Data - informace - znalosti. Dostupné z: http://info.sks.cz/users/ku/ZIZ/inform1.htm

LAMBE, Patrick, 2007. *Organising knowledge: taxonomies, knowledge and organisational effectiveness*. Oxford: Chandos. ISBN 9781843342274.

LARRIVEE, Bob. *Using Analytics: automating processes and extracting knowledge* [online]. Silver Spring: AIIM, 2016 [cit. 2016-03-17]. Dostupné z: http://www.aiim.org/Resources/Research/Industry-Watches/2016/2016_Dec_Content%20Analytics

LIBRARY OF CONGRESS, 2011. Term and Code List for RDA Media Types. In: *Library of Congress* [online]. Washington: Library of Congress [cit. 2017-04-17]. Dostupné z: https://www.loc.gov/standards/valuelist/rdamedia.html

Malá československá encyklopedie : 2. díl, D-CH, 1985. 1. vyd. Praha : Academia, 1985. 969 s.

MALKUSOVÁ, Tereza, 2016. Je lepší vyvíjet vodopádem nebo scrumem? In: *AITOM* [online]. Praha: AITOM [cit. 2017-04-17]. Dostupné z: https://www.aitom.cz/co-je-noveho/je-lepsi-vyvijet-vodopadem-nebo-scrumem

MANCINI, John F., 2011. 5 myths about ECM ROI. In: *Digital Landfill: Overcoming Information Chaos* [online]. Silver Spring: AIIM [cit. 2017-04-16]. Dostupné z: http://info.aiim.org/digital-landfill/newaiimo/2011/08/17/5-myths-about-ecm-roi

MANCINI, John F., 2011. 5 myths about ROT - Redundant, Obsolete, and Trivial Files. In: *Digital Landfill: Overcoming Information Chaos* [online]. Silver Spring: AIIM [cit. 2017-04-16]. Dostupné z: http://info.aiim.org/digital-landfill/newaiimo/2011/09/20/5-myths-about-rot-redundant-obsolete-and-trivial-files

MANCINI, John F., 2017. Content management a quest for new ground. In: *Digital Landfill: Overcoming Information Chaos* [online]. Silver Spring: AIIM [cit. 2017-04-16]. Dostupné z: http://info.aiim.org/digital-landfill/content-management-a-quest-for-new-ground

MARTENS, D, PROVOST, F, 2014. Explaining data-driven document classifications. In: *MIS Quarterly*. **38**(1), s. 73-A6. ISSN: 02767783.

MAYHEW, Beth, 2014. Do you know the 9 Principles of Classification? In: *AIIM Community* [online]. Silver Spring: AIIM [cit. 2017-04-17]. Dostupné z: http://community.aiim.org/blogs/beth-mayhew/2014/08/06/do-you-know-the-9-principles-of-classification

MICROSOFT, 2016. Plan terms and term sets in SharePoint Server 2013. In: *TN Resources and Tools for IT Professionals: Technet* [online]. Washington: Microsoft [cit. 2017-04-17]. Dostupné z: https://technet.microsoft.com/en-us/library/ee519604.aspx

MILES, Doug 2014. *IG Policy versus IG Reality: bringing your wild content uder control* [online]. Silver Spring: AIIM [cit. 2016-03-17]. Dostupné z: http://www.aiim.org/Research-and-Publications/Research/AIIM-White-Papers/IG-Policy-IG-Reality

MILES, Doug, 2015. *Information Governance: too important to be left to humans* [online]. Silver Spring: AIIM [cit. 2016-03-17]. Dostupné z: http://www.aiim.org/Resources/Research/Industry-Watches/2015/2015_Dec_Information-Governance

MILES, Doug, 2016. *Information Management: state of the Industry 2016* [online]. Silver Spring: AIIM [cit. 2016-03-17]. Dostupné z: http://www.aiim.org/SearchPage?keyword=Information+Management+-+state

MOLNÁR, Zdeněk, 2012. *Pokročilé metody vědecké práce*. Zeleneč: Profess Consulting. Věda pro praxi (Profess Consulting). ISBN 9788072590643.

MoReq2 specification: model requirements for the management of electronic records, update and extension, 2008. Luxembourg: Office for Official Publications of the European Communities. INSAR, 8. ISBN 9789279097720.

Musen, M.A, 2015. *The Protégé project: A look back and a look forward*. AI Matters. Association of Computing Machinery Specific Interest Group in Artificial Intelligence, **1**(4), June 2015. DOI: 10.1145/2557001.25757003.

NEWTON, John, 2009. CMIS and Interoperability: AIIM 2009. In: *Share and Discover Knowledge on LinkedIn SlideShare* [online]. Mountain View: LinkedIn [cit. 2017-04-17]. Dostupné z: https://www.slideshare.net/johnnewton/cmis-and-interoperability-aiim-2009

NOY, Natalya F. a Deborah L. MCGUINNESS, *Ontology Development 101: A Guide to Creating Your First Ontology* [online]. Stanford: Stanford University, 25 s. [cit. 2017-04-17]. Dostupné z: http://protege.stanford.edu/publications/ontology_development/ontology101.pdf

OPENTEXT, 2012. *OpenText Extended ECM for SAP Solutions eBook* [online]. Waterloo: OpenText [cit. 2017-04-17]. Dostupné z: http://www.opentext.es/file_source/OpenText/en_US/PDF/OpenText_Extended_ECM_for_SAP_Solutions_10_-_eBook.pdf

PAPÍK, Richard, 2014. Rešeršní strategie a rešeršní služby II.: Vyhledávání informací v elektronických informačních zdrojích z hlediska používaných termínů. *Duha* [online]. **28**(1) [cit. 2017-04-16]. Dostupný z WWW: <http://duha.mzk.cz/clanky/resersni-strategie-resersni-sluzby-ii-vyhledavani-informaci-v-elektronickych-informacnich-zdr>. ISSN 1804-4255.

PELZ - SHARPE, Alan, Apoorv DURGA, David SMIGIEL, Erik HARTMAN a Tony BYRNE, 2009. *ECM 3: ecm maturity model*. 1.

PETERKA, Jiří, 2016. Unijní eIDAS přichází: O co přijdeme u elektronických podpisů?. *Lupa.cz (www.lupa.cz): server o českém Internetu* [online]. [cit. 2017-04-16]. ISSN 1213-0702. Dostupné z: http://www.lupa.cz/clanky/unijni-eidas-prichazi-o-co-prijdeme-u-elektronickych-podpisu/?utm_expid=.1rnVC9uKTLGPIiC_juvx9A.0&utm_referrer=http%3A%2F%2Fwww.lupa.cz%2Fserialy%2Feidas-elektronicke-podpisy-a-sluzby-vytvarejici-duveru%2F

PETERKA, Jiří, 2016a. EIDAS a elektronické podpisy: jak poznáme, o jaký podpis jde? *Lupa.cz (lupa.cz): server o českém Internetu* [online]. Praha: Internet Info [cit. 2017-04-18]. ISSN 1213-0702. Dostupné z: http://www.lupa.cz/clanky/eidas-a-elektronicke-podpisy-jak-poznam-o-jaky-podpis-jde/

PETERKA, Jiří, 2016b. EIDAS: Elektronické značky a pečetě a rekviem za datovou zprávu. *Lupa.cz (www.lupa.cz): server o českém Internetu* [online]. [cit. 2017-04-16]. ISSN 1213-0702. Dostupné z: http://www.lupa.cz/clanky/eidas-elektronicke-znacky-a-pecete-a-rekviem-za-datovou-zpravu/?utm_expid=.1rnVC9uKTLGPIiC_juvx9A.0&utm_referrer=http%3A%2F%2Fwww.lupa.cz%2Fserialy%2Feidas-elektronicke-podpisy-a-sluzby-vytvarejici-duveru%2F

PINCHER, Michael, 2010. A guide to developing taxonomies for effective data management. *ComputerWeookly.com: Information Technology News* [online]. Newton: TechTarget [cit. 2017-04-

17]. Dostupné z: http://www.computerweekly.com/feature/A-guide-to-developing-taxonomies-for-effective-data-management

Planková, Jindra, 2003. indexátor. In: KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV) [online]. Praha : Národní knihovna ČR, 2003- [cit. 2017-04-18]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000727&local_base=KTD.

PLYMOUTH UNIVERSITY, 2017. Business Classification Scheme. In: *Plymouth University* [online]. Plymouth: Plymouth University [cit. 2017-04-16]. Dostupné z: https://www.plymouth.ac.uk/uploads/production/document/path/5/5952/Business_Classification_Scheme_v1.0.pdf

10 Seznam obrázků, tabulek a grafů

10.1 Seznam obrázků

Obrázek 1: Závislost mezi daty, informacemi a znalostmi (Kučerová 2011) 8
Obrázek 2: Příklad BSC Univerzity v Plymouthnu (Plymouth University 2017) 15
Obrázek 3: Architektura CMIS (Newton 2009) 41
Obrázek 4: Příklad OWL ve formátu RDF/XML (Riley) 43
Obrázek 5: PDCA cyklus (Deming Cycle 2009) 51
Obrázek 6: IBM Smarter Process Method (Arsanjani 2015) 52
Obrázek 7: Proces klasifikace dokumentů (zdroj: autorka) 53
Obrázek 8: Rozdělení dokumentů (zdroj: autorka) 68
Obrázek 9: Procesní mapa dokumentů (zdroj: autorka) 70
Obrázek 10: Business pohled na smluvní dokument (zdroj: autorka) 75
Obrázek 11: Příklady zákaznických hierarchií (zdroj: autorka) 83
Obrázek 12: Provázanost dokumentů na zákaznickou hierarchii (zdroj: autorka) 84
Obrázek 13: Zákaznické třídění - typ řízený seznam termínů (zdroj: autorka) 92
Obrázek 14: Taxonomie zákaznických dokumentů (zdroj: autorka) 94
Obrázek 15: Taxonomie zákazických dokumentů s formou dokumentu (zdroj: autorka) 95
Obrázek 16: Příklad fasetového zákaznického tříděně – smluvní dokumenty (zdroj: autorka) 96
Obrázek 17: Příklady předmětové kategorizace pro zákaznické smlouvy (zdroj: autorka) 97
Obrázek 18: Zákaznický tezaurus v anglickém jazyce (zdroj: autorka) 98
Obrázek 19: Tezaurus - deskriptor rámcová smlouva v Protegé (zdroj:autorka) 99
Obrázek 20: Hierarchické zobrazení tezauru (zdroj: autorka) 100
Obrázek 21: Rejstříková podoba zákaznického tezauru (zdroj: autorka) 100
Obrázek 22: Část navržené zákaznické ontologie (zdroj: autorka) 102
Obrázek 23: Výsek z ontologie - obchodní dokumenty (zdroj:autorka) 103
Obrázek 24: Mapování metadata jeden na jednoho (zdroj: AIIM TR48) 107
Obrázek 25: Centrální mapování metadat (zdroj: AIIM TR48) 108
Obrázek 26: Jednocená metadatová integrace (zdroj: AIIM TR48) 109
Obrázek 27: Využití tzn. tag clouds v OCLC OPACu (WorldCat.org 2001-2017) 117
Obrázek 28: Funkce automatického našeptávání v OCLC OPACu (WorldCat.org 2001-2017) 117
Obrázek 29: Možnosti úpravy vyhledávacího dotazu v OCLC OPACu (WorldCat.org 2001-2017) 118
Obrázek 30: IBM FileNet P8 hierarchický seznam hodnot (IBM 2014) 119
Obrázek 31: Open text taxonomie (OpenText 2012) 120
Obrázek 32: Příklad metadat u objektu zákaznické smlouvy (zdroj: autorka) 123
Obrázek 33: Aplikace umožňující práci s navrženou ontologií pro běžného uživatele (The University of Manchester, 2013) 124

10.2 Seznam tabulek

Tabulka 1: Dublin Core Metadata (Dublin Core 2006) 39
Tabulka 2: Vývoj domény ECM (předloženo na základě (Mancini 2017)) 45
Tabulka 3: Úroveň zralosti ECM (zdroj: upraveno dle (Pelz-Sharpe 2009)) 62
Tabulka 4: příklad SWOT analýzy (zdroj: autorka) 64
Tabulka 5: Způsob práce se zákaznickým obsahem (zdroj: autorka) 66
Tabulka 6: Příklad synonymních názvů u metadat (zdroj: autorka) 67
Tabulka 7: Příklad tabulkového zápisu typů dokumentů (zdroj:autorka) 69
Tabulka 8: Popis dokumentu (zdroj:autorka) 71
Tabulka 9: Příklad business pohledů (zdroj: autorka) 76
Tabulka 10: Třídění metadat pro potřeby metodiky (zdroj: autorka) 77
Tabulka 11: Návrh zákaznické klasifikace utajovaných informací (zdroj: autorka) 86
Tabulka 12: Párování metadat mezi Dublin Core a navrženými zákaznickými metadaty (zdroj: autorka) 88
Tabulka 13: Srovnání standardů z pohledu formy a druhu dokumentu (zdroj: autorka) 89
Tabulka 14: Návrh datového modelu (zdroj: autorka) 110
Tabulka 15: Definice vkládacích šablon v GUI (zdroj: autorka) 113
Tabulka 16: Definice vyhledávacích šablon v GUI (zdroj: autorka) 114
Tabulka 17: Raci matice pro služby klasifikace (zdroj: autorka) 128
Tabulka 18: Tabulka evalující dopady změny v třídění (zdroj: autorka) 135

11 Přílohy

Příloha č. 1: Rozšířený seznam prvků Dublin Core (Dublin Core 2006) _____ 153

Příloha č. 2: Porovnání metadata (zdroj: autorka) _____ 156

Příloha č. 3: Schema.org – příklady tříd CreativeWork a Article (Schema.org) _____ 161

Příloha č. 4: Stupně utajovaných informací _____ 167

Příloha č. 5: Metadata ze standardu T40 (AIIM T40 1991)_____ 168

Příloha č. 6: Formát/Fasety OCLC WorldCat.org _____ 169

Příloha č. 7: Dokumentové formáty _____ 170

Příloha č. 8: State Archives and Records: Document form metadata scheme (State Archives and Records NSW) _____ 172

Příloha č. 9: Navržená zákaznická ontgologie (zdroj: autorka)_____ 176

Příloha č. 10: Příklad vzorové tabulky pro plánování metadat v MS SharePoint 2013 (Microsoft 2016) _____ 203

Příloha č. 1: Rozšířený seznam prvků Dublin Core (Dublin Core 2006)

Název prvku	Anglický překlad	popis
Abstrakt	Abstract	Shrnutí obsahu zdroje.
Přístupová práva	accessRights	Informace o tom, kdo má přístup ke zdroji nebo označení úrovně zabezpečení zdroje.
Metoda přírůstku	accrualMethod	Metoda, pomocí níž jsou jednotky přidávány do sbírky.
Periodicita přírůstku	accrualPeriodicity	Frekvence, ve které jsou jednotky přidávány do sbírky.
Strategie výběru přírůstků	accrualPolicy	Taktika/strategie, podle které se řídí přidávání jednotek do sbírky.
Uživatelské určení	audience	Třída entity, pro kterou je zdroj určen nebo je pro ni užitečný.
Dostupný	available	Datum (často rozmezí), kdy zdroj bude nebo byl dostupný.
Bibliografická citace	bibliographicCitation	Bibliografický odkaz zdroje.
Odpovídá (čemu)	conformsTo	Odkaz na zavedený standard, kterému zdroj odpovídá.
Vytvořen	created	Datum vytvoření zdroje.
Datum schválení	dateAccepted	Datum schválení zdroje (např. these na katedře univerzity, článku v časopise apod.)
Datum copyrightu	dateCopyrighted	Datum zapisu copyrightu.
Datum odevzdání	dateSubmitted	Datum odevzdání zdroje (např. these, článku atd.).
Úroveň vzdělání	educationLevel	Obecný údaj popisující potřebnou úroveň vzdělání nebo školení. Nebo přesnější údaj o odborném zaměření a stupni pokročilosti vzdělání cílové skupiny/uživatelů ve smyslu jejího/jejich pokroku v rámci vzdělání nebo výcviku.
Rozsah	extent	Velikost nebo trvání zdroje.
Má formát	hasFormat	Popisovanému zdroji předcházel odkazovaný zdroj, který má vpodstatě stejný intelektualní obsah prezentovaný v jiném formátu.
Obsahuje část	hasPart	Popisovaný zdroj obsahuje odkazovaný zdroj a to fyzicky, nebo logicky.
Má verzi	hasVersion	Popisovaný zdroj má verzi, edici nebo adaptaci, a to jmenovitě odkazovaný zdroj.

Vzdělávací metoda	instructionalMethod	Proces používaný k vštípení znalosti, názoru a dovedností, pro který je zdroj navržen.
Je formátem	isFormatOf	Popisovaný zdroj má stejný intelektuální obsah jako odkazovaný zdroj, ovšem prezentovaný v jiném formátu.
Je částí	isPartOf	Popisovaný zdroj je fyzickou nebo logickou části odkazovaného zdroje.
Je odkazován	isReferencedBy	Popisovaný zdroj je odkazován, citován nebo jinak zmiňován odkazovaným zdrojem.
Je nahrazen	isReplacedBy	Popisovaný zdroj je nahrazen, vyměněn nebo vyřazen odkazovaným zdrojem.
Je vyžadován	isRequiredBy	Popisovaný zdroj je vyžadován odkazovaným zdrojem, a to buď fyzicky nebo logicky.
Vydán	issued	Datum formalního vydání (tj. publikování) zdroje.
Je verzí	isVersionOf	Popisovaný zdroj je verzí, edicí nebo adaptací odkazovaného zdroje. Změna verze naznačuje podstatné změny v obsahu, nikoliv změnu formy.
Licence	license	Právní dokument udělující oficiální povolení k nakládání se zdrojem.
Zprostředkovatel	mediator	Třída entity, která zprostředkovává přístup ke zdroji a pro kterou je zdroj určený nebo užitečný.
Médium	medium	Materiál nebo fyzický nosič zdroje.
Modifikovaný	modified	Datum, kdy byl zdroj změněn.
Původ	provenance	Údaj o jakékoliv změně ve vlastnictví a správě zdroje od jeho vytvoření, která byla podstatná pro jeho autenticitu, integritu a interpretaci.
Odkazuje	references	Popisovaný zdroj odkazuje, cituje nebo jinak zmiňuje odkazovaný zdroj.
Nahrazuje	replaces	Popisovaný zdroj nahrazuje, vyměňuje nebo vyřazuje odkazovaný zdroj.
Vyžaduje	requires	Popisovaný zdroj vyžaduje odkazovaný zdroj na podporu své funkčnosti, výkonu nebo pro zachování souvislosti obsahu.
Držitel práv	rightsHolder	Osoba nebo organizace vlastnící nebo spravující práva na zdroj.
Prostorový	spatial	Prostorová charakteristika intelektuálního obsahu zdroje.
Obsah	tableOfContents	Seznam dílčích jednotek obsahu zdroje.

Časový	temporal	Časová charakteristika intelektuálního obsahu zdroje.
Platný	valid	Datum (často rozmezí) platnosti zdroje.

Příloha č. 2: Porovnání metadata (zdroj: autorka)

ISO 23081		MoReq2 - 2.úroveň	AIIM TR40	Australian e-mail standard (AGIMO)	ANSI/AIIM/ARMA TR48-2004	Dbpedia - document
Identity	entity type	-	-	entity type	-	-
	agreggation	-	-	-	-	-
	registration identifier	system_identifier	primary key	identifier	Unique Identifier	identifier
	-	system_identifier_ rendition	-	-	-	-
Description	abstract	abstract	abstract / subject	description	Description/Abstract/ Subject	description/hasAnotation
	-	author	name(s) of author(s) or editor(s)	-	Author/Creator/Origi nator/	author/writer
	classification	classification	-	keyword	Key Words	subjectTerms/DCC/skos:no tation
	-	copy_recipient	-	-	-	
	jurisdiction	-	copyright	juridistion / coverage	Coverage/Scope	licence
	-	date	document date and time	date range	Date Created/Date Published	completationDate/
	external identifiers	external_identifier	document reference / Keywords or secondary keys	related entity	Rendition Number	document number/references
	place	place	-	location	Coverage/Scope/Loc ation	
	-	recipient	-	-	-	
	-	sender	preparers	-	contributor	chiefEditor/composer/musi cComposer/narrator/mainC

						haracter/staring/translator
	title	title	title	name	title	OriginalTitle/alternativeTitle
	-	-	-	-	Source	Source
	-	-	organization name	contact	From or Sender or Originator/Publisher	publisher
	-	record_type	document type	categories	Document Type	-
	-	-	comments	-	-	-
	-	counter_signature	-	-	-	-
	-	-	number of pages	-	-	-
	-	-	page identifier	-	-	biboPages
	-	-	revision series	-	-	-
	-	-	revision ID	-	-	-
	-	-	user defined field	-	Use-Defined Fields	-
	-	-	owners	-	Office of Origin/Origination Organization	Producer/ProductionCompany
	-	-	-	medium	-	-
Event plan	-	abstract	-	contact	-	-
	event type	-	-	-	-	-
	event relation	agent	-	-	-	-
	event date/time	date	-	-	-	dateLastUpdated
	event description	event_description	-	-	-	-
	event trigger	event_trigger	-	-	-	runtime
	-	status	-	-	-	-
	-	volume	-	-	status	-

Event history	event description	abstract	-	change history	Change History / Preservation and Migration History	-
	event relation	-	-	-	-	-
	event type	-	-	-	-	-
	event identier	-	-	-	-	-
	event date/time	date	-	-	Date Available/Date Closed/Date Cutoff/Date Declared or Filed/Date Modified/Date Received or Acquired/	-
	-	transfer_or_destroy	-	-	-	-
	-	transfer_to	-	-	-	-
	-	-	-	disposal	disposal	-
	-	-	-	-	Disposal Action	-
	-	-	-	-	Disposal Action Date	-
	-	-	-	-	File Code/Number	-
	-	disposal_hold		-	-	-
Use	access	-	-	permission	Accessibility	-
	-	-	-	position	-	-
	audience	-	-	-	audience	-
	integrity	-	-	integrity check	-	-
	documentary form	-	-	document form	-	-
	rights	-	-	rights	rights	-
	-	administrator	-	-	-	-

	-	inactive	-	-	-	-
	language	language	language	language	language	OriginalLanguage
	-	-	-	security classification	Security Classification	-
	-	-	-	security caveat	Supplemental Markings	-
	-	status	-	-	-	-
	technical enviroment	technical_environment	-	format / extent	Format or Application/media type	fileName/fileSize/galleryImage
	-	-	-	precedence	-	-
	-	-	-	Dissemination Limiting Markers		-
					Version Number	-
					Relationships/Links	-
					Signed By/Signator	-
					To/Addressee/CC/BCC	-
					Vital Record Indicator	
Relation	-	agent	-	-	-	-
	-	applies_to_agent	-	-	-	-
	-	applies_to_class	-	-	-	-
	-	cross_referenced_to	-	-	-	-
	-	disposal_hold	-	-	-	-
	-	entity_agent	-	-	-	-
	-	has_redaction	-	-	-	-
	-	has_role	-	-	-	-

	-	has_user	-	-	-	-
	-	is_child_of	-	-	-	-
	-	is_member_of	-	-	-	-
	-	is_redaction_of	-	-	-	-
	-	is_parent_of	-	-	-	-
	-	previous_fully_qualified_classification_code	-	-	-	-
	-	r&d_schedule	-	-	-	-
	-	record_type	-	-	-	-
-	-	-	-	-	-	basedON
-	-	-	-	-	-	cites
-	-	-	-	-	-	previousWork
-	-	-	-	-	-	subsequentWork

Příloha č. 3: Schema.org – příklady tříd CreativeWork a Article (Schema.org)

4.1 Article

Property	Expected Type	Description
	Properties from Action	
actionStatus	ActionStatusType	Indicates the current disposition of the Action.
agent	Organization or Person	The direct performer or driver of the action (animate or inanimate). e.g. *John* wrote a book.
endTime	DateTime	The endTime of something. For a reserved event or service (e.g. FoodEstablishmentReservation), the time that it is expected to end. For actions that span a period of time, when the action was performed. e.g. John wrote a book from January to *December*. Note that Event uses startDate/endDate instead of startTime/endTime, even when describing dates with times. This situation may be clarified in future revisions.
error	Thing	For failed actions, more information on the cause of the failure.
instrument	Thing	The object that helped the agent perform the action. e.g. John wrote a book with *a pen*.
location	Place or PostalAddress or Text	The location of for example where the event is happening, an organization is located, or where an action takes place.
object	Thing	The object upon the action is carried out, whose state is kept intact or changed. Also known as the semantic roles patient, affected or undergoer (which change their state) or theme (which doesn't). e.g. John read *a book*.
participant	Organization or Person	Other co-agents that participated in the action indirectly. e.g. John wrote a book with *Steve*.
result	Thing	The result produced in the action. e.g. John wrote *a book*.
startTime	DateTime	The startTime of something. For a reserved event or service (e.g. FoodEstablishmentReservation), the time that it is expected to start. For actions that span a period of time, when the action was performed. e.g. John wrote a book from *January* to December. Note that Event uses startDate/endDate instead of startTime/endTime, even when describing dates with times. This situation may be clarified in future revisions.
target	EntryPoint	Indicates a target EntryPoint for an Action.
	Properties from Thing	
additionalType	URL	An additional type for the item, typically used for adding more specific types from external vocabularies in microdata syntax. This is a relationship between something and a class that the thing is in. In RDFa syntax, it is better to use the native RDFa syntax - the 'typeof' attribute - for multiple types. Schema.org tools may have only weaker understanding of extra types, in particular those defined externally.
alternateName	Text	An alias for the item.
description	Text	A description of the item.
disambiguatingDescription	Text	A sub property of description. A short description of the item used to disambiguate from other, similar items. Information from other properties (in particular, name) may be necessary for the description to be useful for disambiguation.

Property	Expected Type	Description
image	ImageObject or URL	An image of the item. This can be a URL or a fully described ImageObject.
mainEntityOfPage	CreativeWork or URL	Indicates a page (or other CreativeWork) for which this thing is the main entity being described. See background notes for details. Inverse property: mainEntity.
name	Text	The name of the item.
potentialAction	Action	Indicates a potential Action, which describes an idealized action in which this thing would play an 'object' role.
sameAs	URL	URL of a reference Web page that unambiguously indicates the item's identity. E.g. the URL of the item's Wikipedia page, Freebase page, or official website.
url	URL	URL of the item.

4.2 CreativeWork

Property	Expected Type	Description
	Properties from CreativeWork	
about	Thing	The subject matter of the content.
accessibilityAPI	Text	Indicates that the resource is compatible with the referenced accessibility API (WebSchemas wiki lists possible values).
accessibilityControl	Text	Identifies input methods that are sufficient to fully control the described resource (WebSchemas wiki lists possible values).
accessibilityFeature	Text	Content features of the resource, such as accessible media, alternatives and supported enhancements for accessibility (WebSchemas wiki lists possible values).
accessibilityHazard	Text	A characteristic of the described resource that is physiologically dangerous to some users. Related to WCAG 2.0 guideline 2.3 (WebSchemas wiki lists possible values).
accountablePerson	Person	Specifies the Person that is legally accountable for the CreativeWork.
aggregateRating	AggregateRating	The overall rating, based on a collection of reviews or ratings, of the item.
alternativeHeadline	Text	A secondary title of the CreativeWork.
associatedMedia	MediaObject	A media object that encodes this CreativeWork. This property is a synonym for encoding.
audience	Audience	An intended audience, i.e. a group for whom something was created. Supersedes serviceAudience.
audio	AudioObject	An embedded audio object.
author	Organization or Person	The author of this content or rating. Please note that author is special in that HTML 5 provides a special mechanism for indicating authorship via the rel tag. That is equivalent to this and may be used interchangeably.
award	Text	An award won by or for this item. Supersedes awards.
character	Person	Fictional person connected with a creative work.
citation	CreativeWork or Text	A citation or reference to another creative work, such as another publication, web page, scholarly article, etc.
comment	Comment	Comments, typically from users.
commentCount	Integer	The number of comments this CreativeWork (e.g. Article, Question or Answer) has received. This is most applicable to works published in Web sites with commenting system;

Property	Expected Type	Description
		additional comments may exist elsewhere.
contentLocation	Place	The location depicted or described in the content. For example, the location in a photograph or painting.
contentRating	Text	Official rating of a piece of content—for example,'MPAA PG-13'.
contributor	Organization or Person	A secondary contributor to the CreativeWork or Event.
copyrightHolder	Organization or Person	The party holding the legal copyright to the CreativeWork.
copyrightYear	Number	The year during which the claimed copyright for the CreativeWork was first asserted.
creator	Organization or Person	The creator/author of this CreativeWork. This is the same as the Author property for CreativeWork.
dateCreated	Date or DateTime	The date on which the CreativeWork was created or the item was added to a DataFeed.
dateModified	Date or DateTime	The date on which the CreativeWork was most recently modified or when the item's entry was modified within a DataFeed.
datePublished	Date	Date of first broadcast/publication.
discussionUrl	URL	A link to the page containing the comments of the CreativeWork.
editor	Person	Specifies the Person who edited the CreativeWork.
educationalAlignment	AlignmentObject	An alignment to an established educational framework.
educationalUse	Text	The purpose of a work in the context of education; for example, 'assignment', 'group work'.
encoding	MediaObject	A media object that encodes this CreativeWork. This property is a synonym for associatedMedia. Supersedes encodings.
exampleOfWork	CreativeWork	A creative work that this work is an example/instance/realization/derivation of. Inverse property: workExample.
fileFormat	Text or URL	Media type, typically MIME format (see IANA site) of the content e.g. application/zip of a SoftwareApplication binary. In cases where a CreativeWork has several media type representations, 'encoding' can be used to indicate each MediaObject alongside particular fileFormat information. Unregistered or niche file formats can be indicated instead via the most appropriate URL, e.g. defining Web page or a Wikipedia entry.
funder	Organization or Person	A person or organization that supports (sponsors) something through some kind of financial contribution.
genre	Text or URL	Genre of the creative work or group.
hasPart	CreativeWork	Indicates a CreativeWork that is (in some sense) a part of this CreativeWork. Inverse property: isPartOf.
headline	Text	Headline of the article.
inLanguage	Language or Text	The language of the content or performance or used in an action. Please use one of the language codes from the IETF BCP 47 standard. See also availableLanguage. Supersedes language.
interactionStatistic	InteractionCounter	The number of interactions for the CreativeWork using the WebSite or SoftwareApplication. The most specific child type of InteractionCounter should be used. Supersedes

Property	Expected Type	Description
		interactionCount.
interactivityType	Text	The predominant mode of learning supported by the learning resource. Acceptable values are 'active', 'expositive', or 'mixed'.
isAccessibleForFree	Boolean	A flag to signal that the publication is accessible for free. Supersedes free.
isBasedOn	CreativeWork or Product or URL	A resource that was used in the creation of this resource. This term can be repeated for multiple sources. For example, http://example.com/great-multiplication-intro.html. Supersedes isBasedOnUrl.
isFamilyFriendly	Boolean	Indicates whether this content is family friendly.
isPartOf	CreativeWork	Indicates a CreativeWork that this CreativeWork is (in some sense) part of. Inverse property: hasPart.
keywords	Text	Keywords or tags used to describe this content. Multiple entries in a keywords list are typically delimited by commas.
learningResourceType	Text	The predominant type or kind characterizing the learning resource. For example, 'presentation', 'handout'.
license	CreativeWork or URL	A license document that applies to this content, typically indicated by URL.
locationCreated	Place	The location where the CreativeWork was created, which may not be the same as the location depicted in the CreativeWork.
mainEntity	Thing	Indicates the primary entity described in some page or other CreativeWork. Inverse property: mainEntityOfPage.
mentions	Thing	Indicates that the CreativeWork contains a reference to, but is not necessarily about a concept.
offers	Offer	An offer to provide this item—for example, an offer to sell a product, rent the DVD of a movie, perform a service, or give away tickets to an event.
position	Integer or Text	The position of an item in a series or sequence of items.
producer	Organization or Person	The person or organization who produced the work (e.g. music album, movie, tv/radio series etc.).
provider	Organization or Person	The service provider, service operator, or service performer; the goods producer. Another party (a seller) may offer those services or goods on behalf of the provider. A provider may also serve as the seller. Supersedes carrier.
publication	PublicationEvent	A publication event associated with the item.
publisher	Organization or Person	The publisher of the creative work.
publishingPrinciples	URL	Link to page describing the editorial principles of the organization primarily responsible for the creation of the CreativeWork.
recordedAt	Event	The Event where the CreativeWork was recorded. The CreativeWork may capture all or part of the event. Inverse property: recordedIn.
releasedEvent	PublicationEvent	The place and time the release was issued, expressed as a PublicationEvent.
review	Review	A review of the item. Supersedes reviews.
schemaVersion	Text or URL	Indicates (by URL or string) a particular version of a schema used in some CreativeWork. For example, a document could declare a schemaVersion using an URL such as http://schema.org/version/2.0/ if precise indication of schema

Property	Expected Type	Description
		version was required by some application.
sourceOrganization	Organization	The Organization on whose behalf the creator was working.
spatialCoverage	Place	The spatialCoverage of a CreativeWork indicates the place(s) which are the focus of the content. It is a subproperty of contentLocation intended primarily for more technical and detailed materials. For example with a Dataset, it indicates areas that the dataset describes: a dataset of New York weather would have spatialCoverage which was the place: the state of New York. Supersedes spatial.
sponsor	Organization or Person	A person or organization that supports a thing through a pledge, promise, or financial contribution. e.g. a sponsor of a Medical Study or a corporate sponsor of an event.
temporalCoverage	DateTime or Text or URL	The temporalCoverage of a CreativeWork indicates the period that the content applies to, i.e. that it describes, either as a DateTime or as a textual string indicating a time period in ISO 8601 time interval format. In the case of a Dataset it will typically indicate the relevant time period in a precise notation (e.g. for a 2011 census dataset, the year 2011 would be written "2011/2012"). Other forms of content e.g. ScholarlyArticle, Book, TVSeries or TVEpisode may indicate their temporalCoverage in broader terms - textually or via well-known URL. Written works such as books may sometimes have precise temporal coverage too, e.g. a work set in 1939 - 1945 can be indicated in ISO 8601 interval format format via "1939/1945". Supersedes datasetTimeInterval, temporal.
text	Text	The textual content of this CreativeWork.
thumbnailUrl	URL	A thumbnail image relevant to the Thing.
timeRequired	Duration	Approximate or typical time it takes to work with or through this learning resource for the typical intended target audience, e.g. 'P30M', 'P1H25M'.
translator	Organization or Person	Organization or person who adapts a creative work to different languages, regional differences and technical requirements of a target market, or that translates during some event.
typicalAgeRange	Text	The typical expected age range, e.g. '7-9', '11-'.
version	Number or Text	The version of the CreativeWork embodied by a specified resource.
video	VideoObject	An embedded video object.
workExample	CreativeWork	Example/instance/realization/derivation of the concept of this creative work. eg. The paperback edition, first edition, or eBook. Inverse property: exampleOfWork.
	Properties from Thing	
additionalType	URL	An additional type for the item, typically used for adding more specific types from external vocabularies in microdata syntax. This is a relationship between something and a class that the thing is in. In RDFa syntax, it is better to use the native RDFa syntax - the 'typeof' attribute - for multiple types. Schema.org tools may have only weaker understanding of extra types, in particular those defined externally.
alternateName	Text	An alias for the item.
description	Text	A description of the item.
disambiguatingDescription	Text	A sub property of description. A short description of the item used to disambiguate from other, similar items. Information from other properties (in particular, name) may be necessary

Property	Expected Type	Description
		for the description to be useful for disambiguation.
image	ImageObject or URL	An image of the item. This can be a URL or a fully described ImageObject.
mainEntityOfPage	CreativeWork or URL	Indicates a page (or other CreativeWork) for which this thing is the main entity being described. See background notes for details. Inverse property: mainEntity.
name	Text	The name of the item.
potentialAction	Action	Indicates a potential Action, which describes an idealized action in which this thing would play an 'object' role.
sameAs	URL	URL of a reference Web page that unambiguously indicates the item's identity. E.g. the URL of the item's Wikipedia page, Freebase page, or official website.
url	URL	URL of the item.

Příloha č. 4: Stupně utajovaných informací

4.1. Porovnávací tabulka různého použití ve standardech

Zákon 412/2005	MoReq 2	Australian Government information security management	USA CAPCO Register
Přísně tajné	Top Secret	Top Secret	Top Secret
Tajné	Secret	Secret	Secret
Důvěrné	Confidential	Confidential	Confidential
Vyhrazené	Restricted		
	Unclassified		Unclassified
		Protected	
	Caveat: NATO Eyes Only; WEU Eyes Only	Sensitive: Cabinet; Personal; Legal and Sensitive	
	Descriptor: Commercial; Personnel; Management; Audit and Accounts		

4.2. Norma DIN 66399 (Dočkalová 2017)

Úroveň 1 – Jedná se o běžně dostupné nebo interní informace, které jsou určené pro široké spektrum uživatelů, a u nichž chceme, aby byly nečitelné. Ochrana osobních dat je garantována. Pod první úroveň utajení řadíme dokumenty v 1. (katalogy, brožury), 2. (směrnice, obchodní podmínky) a 3. (obchodní nabídky, obecné ceníky) stupni bezpečnosti.

Úroveň 2 – Do této kategorie řadíme vysoce citlivé, osobní a důvěrné dokumenty, jež jsou vymezeny malé skupině osob. Jejich zneužití má na organizaci výrazný dopad a může vést k porušení zákona či smluvních závazků. Ochrana osobních dat je vymáhána striktními nařízeními. Do této úrovně ochrany patří dokumenty se 3. (obchodní nabídky, ceníky), 4. (cenové nabídky, obchodní výsledky, zdravotní dokumentace) a 5. (patenty, stavební dokumentace, ID) stupněm bezpečnosti.

Úroveň 3 – Sem patří nanejvýš citlivé, důvěrné a přísně tajné informace. Přístup k nim má jen velmi omezený okruh osob, které se musí bezpečným způsobem identifikovat. Zneužití těchto dat je vážné, ohrožující existenci organizace a může vést k porušení utajených obchodů a transakcí. Ochrana osobních údajů je garantována v maximální možné míře. Řadíme sem dokumenty s 5. (patenty, stavební dokumentace, ID), 6. (tajné dokumenty vyžadující mimořádné opatření) a 7. (přísně tajné dokumenty vyžadující mimořádné bezpečnostní opatření – armáda, tajné služby) stupněm bezpečnosti.

Příloha č. 5: Metadata ze standardu T40 (AIIM T40 1991)

Metadata, která zaručí integritu při importu obsahu do systému	Metadata vhodná pro vyhledávání
Batch sequence number	Abstract
Creation date and time	Name(s) of author(s) or editor(s)
Local filing date and time	Comments
Expire date and time	Document date and time
Start date and time	Document type
Document size (in bytes)	Document reference
Input media format	Keywords or secondary keys
Input media location	Number of pages
Indexing operator ID	Languages
Image record migration date	Organization name
Image record destruction date	Preparers
Release date and time	Owners
Revision history	Page identifier
Output media format	Primary key
Output media location	Record series
Orientation of page	Revision ID
Output volume label	Title
References to other documents	Subject
Superseded documents	Copyright
Local file references	User defined field
Source document scaling ratio	
Source document size (print dimensions)	
Security code	
Authorization	
Security classification	
Scan divize ID	
Scanner operator IT	
Status	
Distribution list	

Příloha č. 6: Formát/Fasety OCLC WorldCat.org

Archival Material		Kit
Downloadable Archival Material		Map
Article		eMap
Chapter		Music
Downloadable article		CD
Audiobook		eMusic
CD		LP
eAudiobook		Cassette
LP		Musical Score
Cassette		Downloadable Musical Score
Book		Newspaper
Braille		
Continually Updated Resource		eNewspaper
eBook		Object
Large Print		Sound Recording (Other)
Microform		Toy
Thesis/Dissertation		Video
Computer File		Bluray
Encylopedia		eVideo
Game		DVD
		Film
Video Game		VHS
Image		Visual material
Interactive Multimedia		Downloadable Visual Material
Internet Resource		
Website		
Downloadable Image		
Journal/Magazine		
eJournal/eMagazine		

Příloha č. 7: Dokumentové formáty

7.1 AIIM – TR40

input media format	output media format
unknown	uspecified
paper	fixed hard disk
microfilm	removable hard disk
microfiche	floppy disk
aperture card	floptical
film jacket	WORM
fax (compressed record, as received electronically)	CD-ROM (includes all CD technologies)
digital recording media	rewritable optical disk
slides	optical tape
film	DAT (digital audio tape)
	QIC (quarte-inch catridge) tape
	3780 cartridge tape
	nine-track magnetic tape
	video tape

7.2 RDA (Library of Congress 2011)

audio
computer
microform
microscopic
projected
stereographic
unmediated
video
other
unspecified

7.3 ISBN Area – 0 (ISBN Area 0, 2009)

Media Type Term	Definition and Scope of Term
Audio	**For audio player-enabled resources** Media used to store recorded sound, designed for use with a playback device such as a turntable, audiocassette player, CD player, MP3 player, or iPod. Includes media used to store digitally encoded as well as analogue sound.
Electronic	**For computer-enabled resources** Media used to store electronic files, designed for use with a computer. Includes media that are accessed remotely through file servers, as well as direct-access media such as computer tapes and disks.
Microform	**For microform reader-enabled resources** Media used to store reduced-size images, not readable to the human eye, and designed for use with a device such as a microfilm or microfiche reader. Includes both transparent and opaque micrographic media.
Microscopic	**For microscope-enabled resources** Media used to store minute objects, designed for use with a device such as a microscope to reveal details invisible to the naked eye.
Projected	**For projector-enabled resources** Media used to store moving or still images, designed for use with a projection device such as a motion picture film projector, slide projector, or overhead projector. Includes media designed to project both two-dimensional and three-dimensional images.
Stereographic	**For stereograph viewer-enabled resources** Media used to store pairs of still images, designed for use with a device such as a stereoscope or stereograph viewer to give the effect of three dimensions.
Video	**For video player-enabled resources** Media used to store moving or still images, designed for use with a playback device such as a videocassette player or DVD player. Includes media used to store digitally encoded as well as analogue images.
Multiple media	For resources comprising mixed media where three or more media types are applicable.
Other media	If none of the terms listed above applies to the type of carrier and intermediation device required to convey, use, or perceive the content of the resource(s) being described, the term "other media" is given.

Příloha č. 8: State Archives and Records: Document form metadata scheme (State Archives and Records NSW)

Document form	Scope
Agenda	A list of issues or activities used as a schedule or program for an event, conference, forum or meeting.
Agreement	A summary or record of an arrangement between two or more parties. Use **Contract** for a binding legal agreement in the form of a contract.
Blog	A website with regular entries of commentary, descriptions of events or other material such as graphics or video. Entries are commonly displayed in reverse-chronological order.
Blog post	An entry made on a blog. See **Blog** for further information.
Brief	A request for a consultant to undertake a job. Use **Briefing note** for a report that gives the recipient information to make informed decisions or to be up-to-date about an issue. Use **Project brief** for a statement that describes the purpose, cost, time performance requirements and constraints of a project.
Briefing note	A report that either gives the recipient the information they need to know to make an informed decision or provides information so the recipient is up-to-date on an issue. Use **Brief** for a request for a consultant to undertake a job.
Business plan	A blueprint for how a business will run and what future direction the business will take.
Cabinet Minute	A minute controlled and circulated under The Department of Premier and Cabinet's direction and supervision to authorised recipients only.
Checklist	A written enumeration of actions to be taken, meant to aid memory rather than provide detailed instruction. Use **Dataset** for bibliographic data or catalogues. Use **Index** for a listing of items or entries provided for reference or navigation purposes.
Contract	A binding, legal agreement between two or more parties for the delivery of a product, provision of a service or management of a resource. Use **Agreement** for other for a summary or record of an arrangement between two or more parties which is not legally binding.
Dataset	Structured information encoded in lists, tables, databases etc (for example spreadsheets, databases, GIS data). Data may be numeric, spatial, spectral, statistical or structured text (including bibliographic data and database reports).
Diary	Information arranged in calendar order documenting appointments and engagements. Use **Journal** for information arranged in calendar order documenting events, business or proceedings.
Digital certificate	Any form of electronic code that describes or provides permission to access a resource.
Digital signature	Any form of electronic code used to simulate the security properties of a handwritten signature or to establish authenticity.
Drawing	A sketch or design made with pen, pencil or crayon. Use **Plan** for a drawing made to scale to represent a structure, machine or building.
Electronic message	Any electronically mediated communication. This includes electronic mail, text messages, instant messages, electronic voice messages, electronic video messages and computer conferencing, but excludes scanned versions of written or printed messages.
Enquiry	A request for information.
Facsimile	A copy or reproduction of an item that is as true-to-the-original source as possible. A facsimile transferred by means of a telecommunications technology operating over the telephone network is also known colloquially

	as a ‘fax’.
Fact sheet	A summary of information about a product, service, organisation, event or topic.
Form	A structured solicitation of input from a user (for example, comments, a survey or an order).
Government Gazette	A regular formal publication produced by government that may include vacancies, appointments, bulletins, notices and legislative directives.
Guidelines	The primary purpose of the resource is to present factual information advice or guidance about an organisation, event or service. Most general advisory pages on government websites will be of this document type. Use **Instruction** for resources that provide direction rather than information.
Homepage	The introductory page or major entry point for a site on the internet. In most cases an organisation will have only one resource of this document type, except where there is likely to be a public perception that a distinct business unit standard alone as an organisational entity. Use **Web page** for lower-level pages and **Blog** for a website blog.
House notes	Documents compiled for Ministers attending Parliament which provide brief outline of departmental functions or particular issues.
Index	Any listing of items or entries provided for reference or navigation purposes, including an inventory, register or directory. Use **Checklist** for a written enumeration of actions to be taken.
Instruction	Resources in which the primary purpose is to provide instructions or directions (for example, how to write a report or how to register for a service). Includes manuals, handbooks, tutorials and quizzes. Use **Guidelines** for resources that have primarily informational content. [Note: Should be distinguished from **Procedure** or removed if **Procedure** is a more suitable term for the organisation.]
Invoice	A written list of merchandise, with prices, delivered or sent to a buyer.
Journal	A record or register of events, business or proceedings. Use **Diary** for information arranged in calendar order documenting appointments and engagements.
Lease	A written agreement where the owner of an asset grants a second party the right to its possession and use for a specified period and under specified conditions, in return for periodic payments.
Legal advice	A formal opinion regarding the substance or procedure of the law by an officer of the court, ordinarily in exchange for financial or other tangible compensation.
Letter	A written or printed communication addressed to a person or a number of persons, including scanned versions of written or printed communications.
Log	A chronological listing of actions, observations, data or transactions.
Map	A visual representation of an area highlighting relationships between elements of that space. Use **Plan** for a drawing made to scale to represent a structure, machine or building.
Media release	A resource specifically designed to provide a brief public statement on an issue or event, via the mass media.
Meeting minutes	A summary record of proceedings of a formal meeting.
Memorandum	A note describing something to be remembered or acted upon in the future.
Metadata	Data describing the context, content, structure and organisation of records and other information.
Ministerial briefing note	Formal note providing advice or information to the portfolio Minister or other government organisations on aspects of an organisation’s operations or matters concerning them.
Minute	A form of correspondence acting as an official note or memorandum, usually recording an action or decision, or seeking approval for a course of action. [Note: Should be distinguished from M**emorandum** or removed if **Memorandum** is a more suitable term for the organisation.]

Moving image	A form of visual representation other than text, involving moving pictures, animation, video or film, with or without audio. For some resources it may be appropriate to use a combination of document types. (For example: video recording of an exhibition opening – specify document types as 'promotion; moving image').
Note for file	A summary or record of a less formal meeting or ad hoc discussion (including by telephone).
Plan	A drawing made to scale to represent a structure, machine or building. Includes architectural, engineering and mechanical plans. Use **Business plan** for a blueprint for how a business will run and what future direction the business will take. Use **Plan** for a visual representation of an area highlighting relationships between elements of that space.
Podcast	A series of digital media files (either audio or video) made available on a website for download.
Policy	A major formal publication detailing a course or line of action adopted and pursued by the organisation. Includes public accountability documents such as corporate directions and other strategic plans. Use **Report** for resources that convey the results of an inquiry, account for activities or document speeches and presentations.
Position description	A document describing a position in detail including its purpose, functions and responsibilities and the knowledge, skills and abilities required to perform the job competently. Position descriptions are used in evaluating positions and copies are given to applicants.
Presentation	Any form of visual and/or verbal communication used to show or explain a topic to an audience.
Procedure	A sequence of action or instructions to be followed. [Note: Should be distinguished from **Instruction** or removed if **Instruction** is a more suitable term for the organisation.]
Project brief	A statement that describes the purpose, cost, time performance requirements and constraints of a project.
Project plan	A formal accepted document outlining the aims, costs and time sequencing of a project. A project plan is signed off by key stakeholders and becomes the basis for monitoring and controlling the project.
Promotion	Descriptive or marketing information about an organisation or material that promotes its products, services, activities or collections (for example, 'What's New' pages, brochures). Does not include media releases.
Proposal	A written offer from a seller to a prospective buyer. Includes an offer regarding work to be completed.
Purchase order	A commercial document issued by a buyer to a seller, indicating types, quantities, and agreed prices for products or services the seller will provide to the buyer.
Questions on notice	Notice of questions to members of Parliament, signed by a member, and handed to one of the Clerks-at-the-Table during a sitting of the House.
Questions without notice	Questions asked by members rising in their place and obtaining the call of the chair (the President). Questions must seek information or press for action in relation to matters with which Ministers are officially connected, to proceedings pending in the House or to any other issues within Ministers' responsibilities.
Quote	A statement of the price of goods or services.
Register	A formal, written record of acts, events, names etc, often kept in a prescribed manner.
Report	An account describing organisational activity, a speech or presentation. Includes statements of the organisation's opinion, a decision or the results of an enquiry. Use **Dataset** for database reports.
Roster	A list allocating names to particular tasks, duties or work times.
RSS feed	A means of providing an automatic summary of frequently updated web content in a standardised format. Abbreviation of Really Simple Syndication

	feed.
Schematic	A diagram that represents the properties of a system using abstract, graphic symbols rather than realistic pictures. A schematic usually omits all details that are not relevant to the information the schematic is intended to convey, and may add unrealistic properties that aid comprehension. For example, a subway map intended for riders may represent a subway station with a dot; the dot doesn't resemble the actual station at all but gives the viewer the information he needs without unnecessary visual clutter. Use **Still image** for still visual representation including images, photographs, diagrams, maps and graphics.
Software	Computer programs in source or compiled form that may be available for installation on another machine.
Sound recording	Primarily audio representation that may be ambient, effects, music, narration or speech. For some audio resources it may be appropriate to use a combination of document types. (For example: sound recording of a presentation – specify document types as 'presentation; sound recording'.)
Still image	A primarily still visual representation other than text. Includes electronic and physical representations such as images, photographs, diagrams, maps and graphics. For digital representations of physical resources, use a more specific document type where possible. (For example: a scanned media release – use 'media release; still image'.) Use **Schematic** for a diagram that represents the properties of a system using abstract, graphic symbols rather than realistic pictures.
Tender - request for quotation	An invitation to suppliers to take part in a bidding process to bid on specific products or services. Often known as a RFQ.
Tender - request for tender	A structured invitation to suppliers for the supply of products and/or services. Often known as an RFT.
Tweet	A post or status update on Twitter, a micro-blogging service.
Web page	A resource of information suitable for the internet that can be accessed through a web browser. A web page may provide navigation to other web pages via hypertext links. Use the more specific document type where possible (For example: Use 'Home page' for the introductory page or major entry point for a site, use 'Wiki' for a web page wiki).
Wiki	A page or collection of web pages designed to enable anyone accessing it to contribute or modify content using a simplified markup language. Wikis are often used to create collaborative websites, intranets and knowledge management systems.

Příloha č. 9: Navržená zákaznická ontgologie (zdroj: autorka)

<?xml version="1.0"?>

<rdf:RDF xmlns="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#"

xml:base="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91"

xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"

xmlns:owl="http://www.w3.org/2002/07/owl#"

xmlns:xml="http://www.w3.org/XML/1998/namespace"

xmlns:xsd="http://www.w3.org/2001/XMLSchema#"

xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">

<owl:Ontology rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91"/>

<!--

///

//

// Object Properties

//

///

-->

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen -->

<owl:ObjectProperty rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen">

<owl:inverseOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<rdfs:domain rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Subjekt"/>

<rdfs:domain rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty"/>

<rdfs:range rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dokumenty"/>

</owl:ObjectProperty>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje -->

<owl:ObjectProperty rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje">

<rdfs:domain rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dokumenty"/>

<rdfs:range rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Subjekt"/>

<rdfs:range rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty"/>

</owl:ObjectProperty>

<!--

//

//

// Classes

//

//

-->

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Ceníky -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Ceníky">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ObchodníDokumenty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#DaňovéDokumenty -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#DaňovéDokumenty">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dokumenty"/>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Zákazník"/>

</owl:Restriction>

</rdfs:subClassOf>

<rdfs:subClassOf>

```
        <owl:Restriction>
            <owl:onProperty
rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>
            <owl:someValuesFrom
rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-
91#Služby/Produkty"/>
        </owl:Restriction>
      </rdfs:subClassOf>
   </owl:Class>



   <!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#DodacíListy -->

   <owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-
91#DodacíListy">
      <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-
ontology-91#KomunikaceSKlientem"/>
   </owl:Class>



   <!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dodatek -->

   <owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-
91#Dodatek">
      <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-
ontology-91#SmluvníDokumenty"/>
   </owl:Class>
```

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dodavatel -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dodavatel">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Subjekt"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dokumenty -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dokumenty">

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Zákazník"/>

</owl:Restriction>

</rdfs:subClassOf>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dopisy -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dopisy">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#KomunikaceSKlientem"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#E-maily -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#E-maily">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#KomunikaceSKlientem"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Faktury -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Faktury">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#DaňovéDokumenty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#KomunikaceSKlientem -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#KomunikaceSKlientem">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dokumenty"/>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Zákazník"/>

</owl:Restriction>

</rdfs:subClassOf>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty"/>

</owl:Restriction>

</rdfs:subClassOf>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#MarketingováStrategie -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#MarketingováStrategie">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#StrategickéDokumenty"/>

<owl:disjointWith rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProdejníStrategie"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Nabídky -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Nabídky">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ObchodníDokumenty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Návody -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Návody">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#KomunikaceSKlientem"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ObchodníDokumenty -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ObchodníDokumenty">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dokumenty"/>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#PotenciálníZákazník"/>

</owl:Restriction>

</rdfs:subClassOf>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Zákazník"/>

</owl:Restriction>

</rdfs:subClassOf>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty"/>

</owl:Restriction>

</rdfs:subClassOf>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Opisy -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Opisy">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#DaňovéDokumenty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Partner -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Partner">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Subjekt"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#PotenciálníZákazník -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#PotenciálníZákazník">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Subjekt"/>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Ceníky"/>

</owl:Restriction>

```
        </rdfs:subClassOf>

        <rdfs:subClassOf>

            <owl:Restriction>

                <owl:onProperty
rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>

                <owl:someValuesFrom
rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Nabídky"/>

            </owl:Restriction>

        </rdfs:subClassOf>

    </owl:Class>




    <!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Prezentace -->


    <owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-
91#Prezentace">

        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-
ontology-91#ObchodníDokumenty"/>

    </owl:Class>




    <!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProdejníStrategie -->


    <owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-
91#ProdejníStrategie">

        <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-
ontology-91#StrategickéDokumenty"/>

    </owl:Class>
```

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProduktyABC -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProduktyABC">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProduktyDEF -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProduktyDEF">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProduktyGHI -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProduktyGHI">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProjektovéDokumenty -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProjektovéDokumenty">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#KomunikaceSKlientem"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#PříjmovéListy -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#PříjmovéListy">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#KomunikaceSKlientem"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Příloha -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Příloha">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SmluvníDokumenty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#RámcovéSmlouvy -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#RámcovéSmlouvy">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SmluvníDokumenty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SlužbaABC -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SlužbaABC">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SlužbaDEF -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SlužbaDEF">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SlužbaGHI -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SlužbaGHI">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Smlouvy -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Smlouvy">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SmluvníDokumenty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SmluvníDokumenty -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SmluvníDokumenty">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dokumenty"/>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Zákazník"/>

</owl:Restriction>

</rdfs:subClassOf>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty"/>

</owl:Restriction>

</rdfs:subClassOf>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Spis -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Spis">

<owl:equivalentClass>

<owl:Class>

<owl:unionOf rdf:parseType="Collection">

<rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dodatek"/>

<rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Faktury"/>

<rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#RámcovéSmlouvy"/>

<rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Smlouvy"/>

<rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#UpomínkaI"/>

<rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#UpomínkaII"/>

<rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#VymáhacíProces"/>

<rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Výpověď"/>

</owl:unionOf>

</owl:Class>

</owl:equivalentClass>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#StrategickéDokumenty -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#StrategickéDokumenty">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dokumenty"/>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Zákazník"/>

</owl:Restriction>

</rdfs:subClassOf>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty
rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<owl:someValuesFrom
rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-
91#Služby/Produkty"/>

</owl:Restriction>

</rdfs:subClassOf>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Subjekt -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-
91#Subjekt"/>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#TechnickéProvedení
-->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-
91#TechnickéProvedení">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-
ontology-91#KomunikaceSKlientem"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#UpomínkaI -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-
91#UpomínkaI">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#VymáhacíProces"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#UpomínkaII -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#UpomínkaII">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#VymáhacíProces"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Vrubopisy -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Vrubopisy">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#DaňovéDokumenty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#VymáhacíProces -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#VymáhacíProces">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dokumenty"/>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Zákazník"/>

</owl:Restriction>

</rdfs:subClassOf>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#obsahuje"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty"/>

</owl:Restriction>

</rdfs:subClassOf>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Výpověď -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Výpověď">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SmluvníDokumenty"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ZákaznickýReport -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ZákaznickýReport">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#KomunikaceSKlientem"/>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Zákazník -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Zákazník">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Subjekt"/>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#DaňovéDokumenty"/>

</owl:Restriction>

</rdfs:subClassOf>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#KomunikaceSKlientem"/>

</owl:Restriction>

</rdfs:subClassOf>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ObchodníDokumenty"/>

</owl:Restriction>

</rdfs:subClassOf>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SmluvníDokumenty"/>

</owl:Restriction>

</rdfs:subClassOf>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#StrategickéDokumenty"/>

</owl:Restriction>

</rdfs:subClassOf>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#VymáhacíProces"/>

</owl:Restriction>

</rdfs:subClassOf>

</owl:Class>

<!-- http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty -->

<owl:Class rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Služby/Produkty">

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#DaňovéDokumenty"/>

</owl:Restriction>

</rdfs:subClassOf>

<rdfs:subClassOf>

<owl:Restriction>

<owl:onProperty rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#KomunikaceSKlientem"/>

</owl:Restriction>

```
    </rdfs:subClassOf>

    <rdfs:subClassOf>

      <owl:Restriction>

        <owl:onProperty
rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>

        <owl:someValuesFrom
rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-
91#ObchodníDokumenty"/>

      </owl:Restriction>

    </rdfs:subClassOf>

    <rdfs:subClassOf>

      <owl:Restriction>

        <owl:onProperty
rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>

        <owl:someValuesFrom
rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-
91#SmluvníDokumenty"/>

      </owl:Restriction>

    </rdfs:subClassOf>

    <rdfs:subClassOf>

      <owl:Restriction>

        <owl:onProperty
rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>

        <owl:someValuesFrom
rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-
91#StrategickéDokumenty"/>

      </owl:Restriction>

    </rdfs:subClassOf>

    <rdfs:subClassOf>

      <owl:Restriction>

        <owl:onProperty
rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#jeObsažen"/>
```

<owl:someValuesFrom rdf:resource="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#VymáhacíProces"/>

</owl:Restriction>

</rdfs:subClassOf>

</owl:Class>

<!--

///

//

// General axioms

//

///

-->

<rdf:Description>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AllDisjointClasses"/>

<owl:members rdf:parseType="Collection">

<rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Ceníky"/>

<rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Nabídky"/>

<rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Prezentace"/>

</owl:members>

</rdf:Description>

<rdf:Description>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AllDisjointClasses"/>

<owl:members rdf:parseType="Collection">

```
        <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#DodacíListy"/>

        <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dopisy"/>

        <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#E-maily"/>

        <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Návody"/>

        <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProjektovéDokumenty"/>

        <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#PříjmovéListy"/>

        <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#TechnickéProvedení"/>

        <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ZákaznickýReport"/>

    </owl:members>

</rdf:Description>

<rdf:Description>

    <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AllDisjointClasses"/>

    <owl:members rdf:parseType="Collection">

        <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Dodatek"/>

        <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Příloha"/>

        <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#RámcovéSmlouvy"/>

        <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Smlouvy"/>

        <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Výpověď"/>

    </owl:members>

</rdf:Description>

<rdf:Description>
```

```
        <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AllDisjointClasses"/>

        <owl:members rdf:parseType="Collection">

            <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Faktury"/>

            <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Opisy"/>

            <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#Vrubopisy"/>

        </owl:members>

    </rdf:Description>

    <rdf:Description>

        <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AllDisjointClasses"/>

        <owl:members rdf:parseType="Collection">

            <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProduktyABC"/>

            <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProduktyDEF"/>

            <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#ProduktyGHI"/>

            <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SłużbaABC"/>

            <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SłużbaDEF"/>

            <rdf:Description rdf:about="http://www.semanticweb.org/vkz/ontologies/2017/2/untitled-ontology-91#SłużbaGHI"/>

        </owl:members>

    </rdf:Description>

</rdf:RDF>



<!-- Generated by the OWL API (version 4.2.6.20160910-2108) https://github.com/owlcs/owlapi -->
```

Příloha č. 10: Příklad vzorové tabulky pro plánování metadat v MS SharePoint 2013 (Microsoft 2016)

Microsoft® Office SharePoint® Server 2013 Term sets worksheet	
Fill this worksheet by using the following topics: Plan terms and term sets	
Prepared by:	
Date:	

Term set	**Description**	**Owner**	**Group**
cost center	All valid cost centers in the company	Manuel Oliveira	HR term sets
product name	The names of all products the company sells. For released products, this is the official name from Marketing. For pre-release products, this can be a code name.	Anna Misiec	marketing term sets
job title	Official job titles	Ben Miller	HR term sets
company offices	The name of every facility at which the company maintains a presence.	Facilities	facilities term sets