



Návrh potřebných úprav výrobní linky a realizace úprav v prototypu

Vytvořeno v rámci projektu

**Rozvoj datových politik v oblasti zlepšování kvality
a interoperability dat veřejné správy**
CZ.03.4.74/0.0/0.0/15_025/0013983

KA 6: Návrhy a realizace prostředí pro ontologické konceptuální datové modelování

Výstup: C5V6



1. Definice výstupu

"Typ výstupu: Dokument s návrhem úprav výrobní linky, dokument s dokumentací upraveného prototypu, zdrojové kódy vytvořeného prototypu vydané též jako open source (otevřený zdrojový kód).

Na základě zkušeností z C5V3, C5V4 a C5V5 budou analyzovány, navrženy a realizovány potřebné úpravy výrobní linky. Tyto úpravy budou probíhat již v průběhu řešení C5V3, C5V4 a C5V5. Potřebné úpravy budou zdokumentovány."

2. Úvod

Prototyp výrobní linky navržený v rámci C5V3 je v rámci tohoto výstupu dále rozvíjen, a to na základě uživatelské zpětné vazby. V průběhu testování prototypu výrobní linky členy projektového týmu, zástupci OHA MVČR a pracovníky MMR bylo identifikováno několik klíčových změn a rozšíření. Tyto požadavky jsou od uživatelů sbírány a průběžně zpracovávány v repozitáři <https://github.com/opendata-mvcr/sgov-assembly-line>.

Kromě těchto změn, došlo rovněž k vývoji nového prototypového nástroje Showlt pro zobrazování publikovaných slovníků široké veřejnosti, neboť veřejné uživatelské rozhraní systému Termlt se ukazuje jako příliš komplikované pro neškolené uživatele.

Další výraznou změnou v prototypu Výrobní linky byla změna architektury. Motivací pro tuto změnu byl neefektivní samostatný vývoj jednotlivých komponent (zejména pro aplikaci Termlt). Cílem změny je opuštění konceptu pracovních prostorů a jeho nahrazení vidličkami (forky), tedy draft verzemi slovníku. First class citizenem výrobní linky se tak stane samotný slovník.

Konečně dokument obsahuje i seznam navrhovaných features vhodných pro další vývoj výrobní linky pro její produkční nasazení.

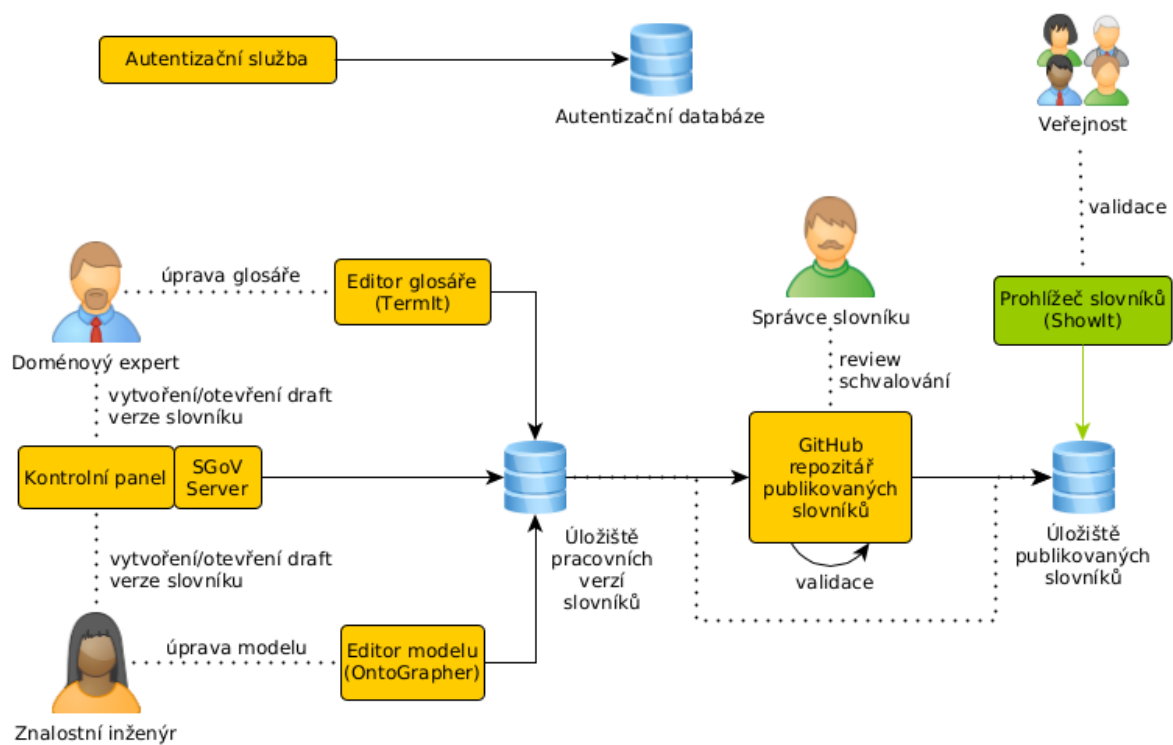
3. Showlt

Systém Showlt slouží k jednoduchému zobrazení pojmů a jejich vyhledávání v sémantickém slovníku pojmů. Je nasazen na adrese

<https://slovník.gov.cz/prohlížíme> .

Zdrojové kódy systému Showlt jsou k dispozici v GIT repozitáři <https://github.com/opendata-mvcr/sgov-browser> (commit 4181c110995062c953e10d8f7c5e66e0de49397e).

Začlenění systému Showlt do architektury výrobní linky je znázorněno na následujícím diagramu zeleně:



Obrázek 1: Architektura výrobní linky se zeleně vyznačenými změnami pro umožnění přístupu veřejnosti k obsahu Sémantického slovníku pojmů pomocí aplikace Showlt.

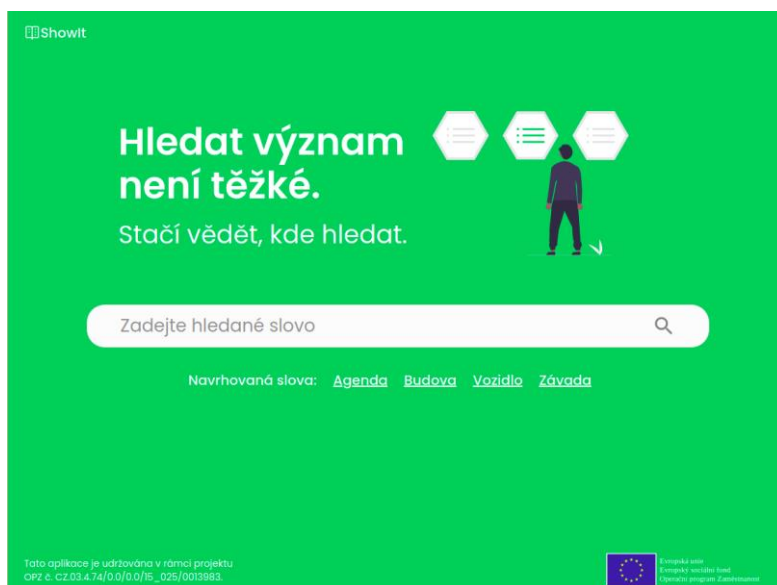
Základní práce se systémem Showlt

Nástroj Showlt nahrazuje read only verzi nástroje Termt, umožňuje vyhledávat "slova", ke kterým přiřazuje pojmy s konkrétním významem.

Pojmy v nástroji Showlt jsou sdruženy ve slovnících. Slovník se zpravidla zabývá jedním legislativním nebo jiným dokumentem, datovou sadou, agendou, atd. Názvy pojmů v rámci slovníku musí být jedinečné.

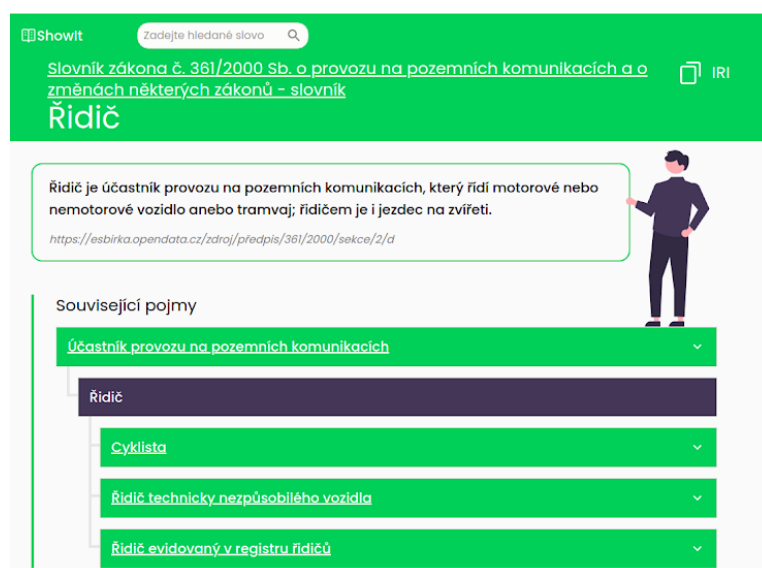
Pojmy obsahují definici, její zdroj, hierarchické a modelové (sémantické) vztahy vzhledem k ostatním pojmům, ať už ze stejného, nebo z jiných slovníků.

Úvodní stránka slouží k vyhledávání slov.



Obrázek 2: Titulní stránka aplikace ShowIt.

Např. po vyhledání slova “řidič” se zobrazí jeho detail pojmu s názvem řidič (pokud je jediný):



Obrázek 3: Příklad detailu pojmu Řidič v aplikaci ShowIt po vyhledání slova Řidič.

V případě nejednoznačnosti (kdy se jedno slovo vyskytuje v různých významech ve více zdrojích) má uživatel možnost kliknutím na příslušný zdroj vybrat požadovaný význam.



Obrázek 4: Příklad slova s definicemi v různých kontextech - tedy slova, které je vyjádřeno několika různými pojmy.

4. Změna architektury výrobní linky

Cílem změny architektury je odstranit z procesu pracovní prostory a nahradit je **draft verzemi jednotlivých slovníků**. Změna architektury se dotýká celé výrobní linky, ačkoliv největší změny jsou na straně SGoV serveru a v aplikaci TermIt. Motivací pro změnu je neefektivní vývoj jednotlivých komponent ve výrobní lince a vylepšení user experience (UX) celého systému. Cílem je, aby uživatel, který chce upravit nějaký slovník, musel udělat pouze minimální počet nutných kroků ve výrobní lince, než se dostane k samostatné úpravě. Ve stávajícím prototypu to znamená otevřít kontrolní panel, vytvořit pracovní prostor, přidat slovník do pracovního prostoru, otevřít pracovní prostor v nástrojích určených pro editaci slovníků – TermIt nebo OntoGrapher, provést změny a ty publikovat.

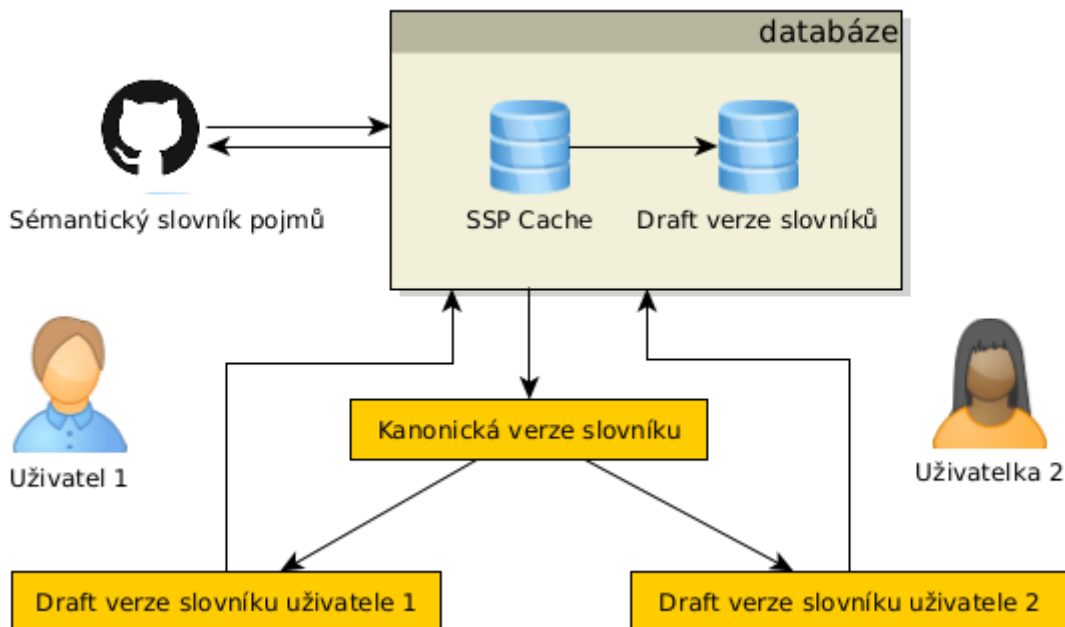
V navrhované úpravě to znamená otevřít kontrolní panel, vytvořit nový nebo vyhledat existující slovník, který chce uživatel upravit a otevřít ho v TermItu nebo OntoGrapheru. Důležitým krokem je sjednocení backendu aplikace TermIt používané v rámci výrobní linky s verzí aplikace TermIt, která je vyvíjena jako samostatná aplikace. Toto sjednocení vede k lepšímu a efektivnějšímu vývoji nástroje TermIt i mimo výrobní linku (kde už nyní probíhá nemalá část vývoje).

Změnu architektury bylo kvůli zachování funkčního testovacího nasazení velice složité implementovat jednorázově, proto byla rozdělena do několika dílčích kroků:

1. sjednocení datové struktury pracovního prostoru a slovníku
 - a. veškeré pomocné kontexty z pracovního prostoru je potřeba přesunout pod slovníkový kontext,
 - b. veškerá metadata (autor atd...) také,
2. Implementace podpory sjednocené struktury pro SGoV server,
3. Implementace podpory sjednocené struktury pro nástroje TermIt a OntoGrapher,
4. Implementace sledování změn (kdo jsou přispěvatelé pro jednotlivé draft verze slovníků a kdy v ní provedli poslední změnu),
5. Implementace podpory předchozích bodů v nové verzi kontrolního panelu.

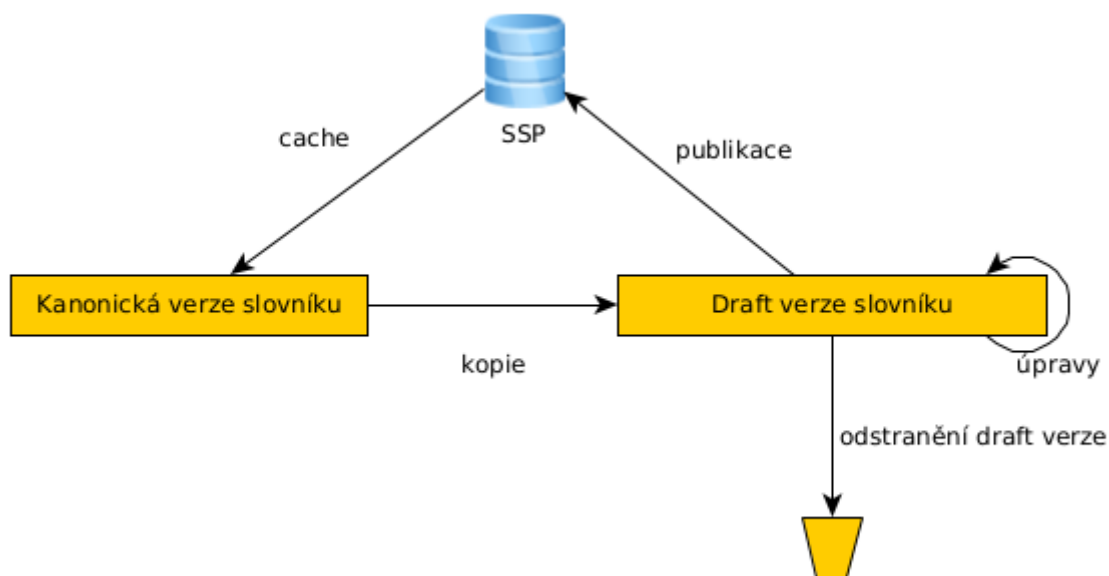


Praktické dopady těchto změn ovlivňují procesy a životní cyklus změn ve slovníku následujícím způsobem.



Obrázek 5: *Single source of true* pro Sémantický slovník pojmů je v GitHub repozitáři. Na základě jeho obsahu se aktualizuje databáze, která obsahuje tzv. SSP Cache (kopii SSP) a zároveň udržuje draft verze jednotlivých slovníků. Různí uživatelé (Uživatel 1 a Uživatelka 2) vytváří kopie z Kanonické verze a jejich úpravou vznikají draft verze těchto slovníků, které jsou ukládány v samostatných kontextech do databáze.

Sémantický slovník pojmů (SSP) obsahuje publikované verze slovníků veřejné správy. Na jeho základě vzniká tzv. SSP cache, která obsahuje tzv. kanonické slovníky, tedy slovníky v takové podobě, v jaké se nachází v SSP. V okamžiku, kdy uživatel edituje slovník, výrobní linka vytvoří jeho draft verzi. Po dokončení úprav je draft verze odeslána k publikaci a po úspěšném procesu publikace je nová verze zveřejněna v Sémantickém slovníku pojmů a kanonická verze v SSP cachi je nahrazena touto novou verzí, jak je vidět na diagramu na Obrázku 6. Po publikaci je dle životního cyklu vhodné draft verzi smazat.



Obrázek 6: Životní cyklus slovníku a jeho verzí. Při editaci slovníku Výrobní linka vytváří kopii jeho Kanonické verze jako Draft verzi slovníku. Po úpravách vzniká návrh k publikaci, po jehož schválení nebo zamítnutí může být Draft verze dále upravována a/nebo smazána.

Okamžikem schválení Draft verze nahrazuje Kanonickou verzi slovníku v SSP a je aktualizována i v SSP Cachi.

Různí uživatelé mohou vytvářet různé verze stejného kanonického slovníku a mohou si do nich navzájem i přispívat. Uživatelka 2 z Obrázku 5 může za určitých podmínek editovat i Draft verzi slovníku uživatele 1 a naopak. Z toho vyplývá, že při procesu publikace je potřeba řešit možné konflikty mezi různými verzemi draftů (viz [Nástroj pro kontrolu a publikaci slovníků](#)) a bylo je potřeba řešit i při řešení dílčích kroků, především při implementaci změn v nástroji Termlt.

Praktickým dopadem změny architektury je i předávání slovníků aplikacím (Termlt a OntoGrapher). V původní verzi prototypu bylo předáno IRI pracovního prostoru (metadatový kontext) aplikacím pro úpravu slovníků a modelů. V nové verzi je předáváno IRI samotného draftu slovníku (slovníkový kontext). V případě manipulace s více slovníky najednou je tedy nutné předávat všechny slovníkové kontexty.

Ve stávající podobě kontrolního panelu jsou stále vytvářeny pracovní prostory, ale aplikacím je předáván již pouze seznam slovníkových kontextů. Nová verze kontrolního panelu, která již bere v úvahu změnu architektury je popsána v kapitole [Kontrolní panel - nová verze](#)



5. Dokumentace nové verze Výrobní linky

Tato kapitola dokumentuje změny ve výrobní lince od verze prototypu popsané ve výstupu C5V3, odevzdaném v rámci ZOR 052022. Popisuje změny ve funkcionalitě jednotlivých nástrojů v rámci výrobní linky.

High-level architektura

Původní high-level architektura byla postavená na správě slovníků v rámci pracovních prostorů. Nově je first class citizenem v rámci výrobní linky slovník a jeho úpravami vznikají daft verze těchto slovníků. To výrazně přepracovává původní koncept high-level architektury tak, jak je popsán ve výstupu C5V3.

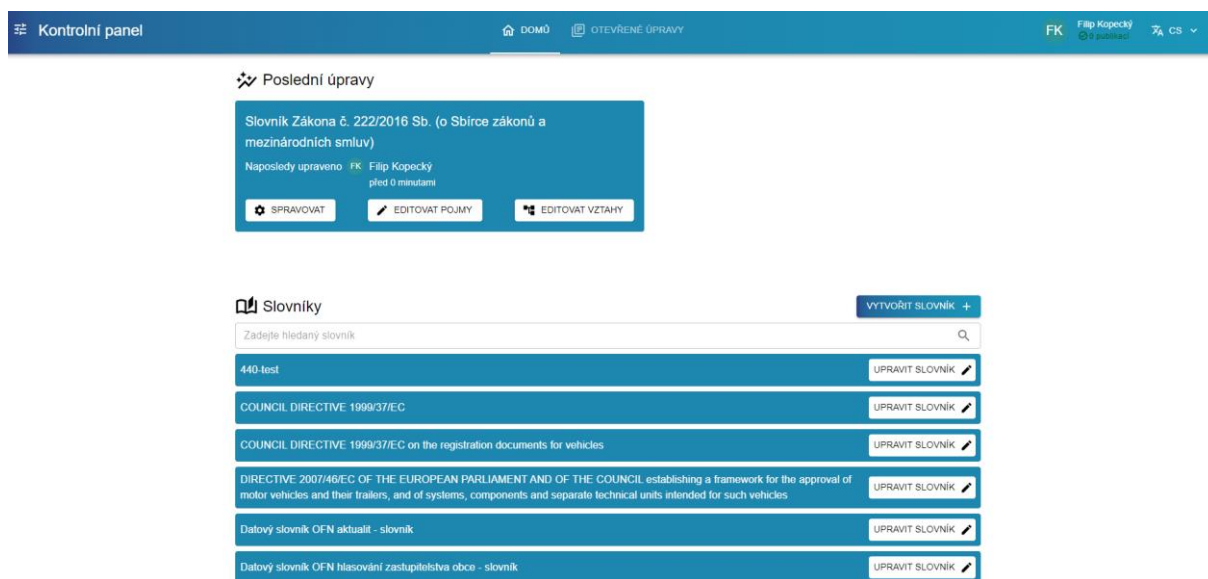
Kontrolní panel - nová verze

Jak již bylo zmíněno výše, nová architektura VL odstraňuje pracovní prostory a slovník se tedy stal first class citizenem celé výrobní linky. Bylo tedy nutné upravit vzhled Kontrolního panelu, tak aby tyto změny ovlivnily uživatele co možná nejméně a jeho práci zjednodušily.

Titulní obrazovka nového Kontrolního panelu zobrazuje seznam všech slovníků k dispozici v SSP. Chce-li uživatel daný slovník upravit, stačí aby pouze kliknul na tlačítko upravit a slovník se připraví k úpravám. Toto zjednodušení urychluje práci s Kontrolním panelem, jelikož dříve bylo nezbytné založit nový pracovní prostor, pojmenovat ho a nakonec slovník přidat do pracovního prostoru. Nyní je možné všechny tyto kroky učinit pomocí jediného tlačítka.

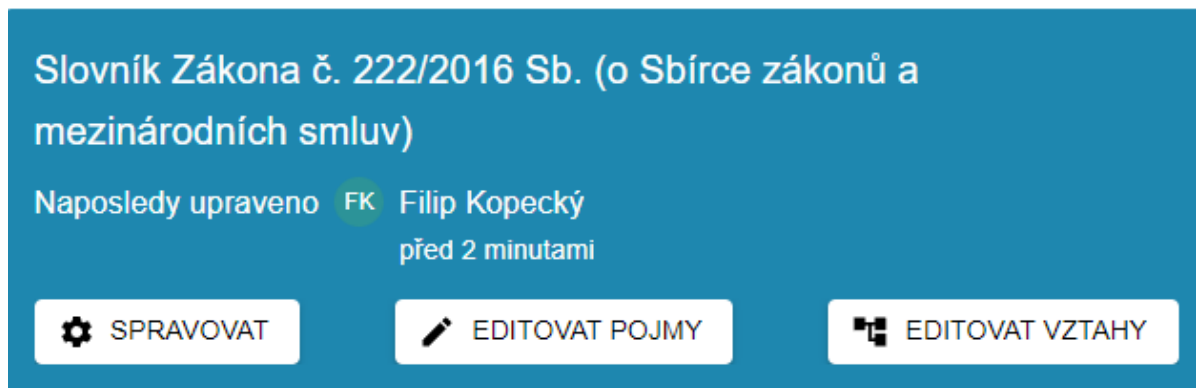
Dalším zrychlením práce uživatelů s Kontrolním panelem je personalizovaná nabídka naposledy upravovaných slovníků. Z této nabídky je možné slovníky ihned otevřít v TermItu či OntoGrapheru. Oproti předchozí verzi, která zobrazovala všechny otevřené pracovní prostory, uživatelé nejsou nuceni hledat jejich úpravy v dlouhém seznamu všech otevřených pracovních prostorů.

Hlavním cílem implementace těchto změn bylo doručit uživatelům co možná nejkratší cestu k zahájení nových úprav či pokračování v těch předchozích. Tyto use-casy byly identifikovány jako nejčastější případy užití Kontrolního panelu a proto jsou k dispozici hned na titulní obrazovce a jako takové jsou vyvolatelné jedním kliknutím tlačítka.



Obrázek 7: Úvodní stránka nového Kontrolního panelu Výrobní linky.

Pochopitelně je možné upravovat více slovníků najednou, stejně jako tomu bylo v předchozí verzi Kontrolního panelu. Všechny nově upravované slovníky vytváří tzv. projekty. Do existujícího projektu je možné přidávat další slovníky – přesněji jejich kopie. Projekty jsou logická seskupení změn dovolující upravovat více slovníků najednou. Přidávání slovníků je možné pod správou změn daného slovníku. (tlačítko “spravovat”).



Obrázek 8: Detail rozpracovaného projektu na úvodní stránce.

To nás přeměruje na správu změn projektu, ve které můžeme jednoduše vidět jaké slovníky upravujeme a případně k nim další slovníky přidat.



Slovník Zákona č. 222/2016 Sb. (o Sbírce zákonů a mezinárodních smluv)

Upraveno před 3 minutami

EDITOVAT POJMY

EDITOVAT VZTAHY

PUBLIKOVAT

SMAZAT PROJEKT

Upravuje

Slovník Zákona č. 222/2016 Sb. (o Sbírce zákonů a mezinárodních smluv) ODEBRAT

PŘIDAT SLOVNÍK

Obrázek 9: Detail rozpracovaného projektu po jeho rozkliknutí.

Seznam všech projektů lze nalézt pod kartou *otevřené úpravy*. Informace o jejich poslední úpravě je zobrazovaná v uživatelsky přívětivé variantě, která je případně upozorní na změny které jsou dlouho otevřené. (jeden z hlavních problémů ve VL, kde uživatelé nechávali pracovní prostory dlouho otevřené bez publikace). Důvodem pro včasné publikace je mimo jiné i to, že draft verze slovníku se může časem odlišovat od kanonické verze.


Název	Upraveno	Spolupracovali	Hledat
TEST_MS	před 13 minutami	MŠ	SPRAVOVAT
test-stenclak	před 8 dny	ŠS	SPRAVOVAT
test-trello-470	před 16 dny	AK	SPRAVOVAT
test-trello-2	⚠ před 21 dny	AB	SPRAVOVAT

Obrázek 10: Seznam otevřených úprav upozorňuje uživatele na projekty, které nebyly dlouho publikovány.

Název nového projektu odpovídá názvu slovníku, který je v něm upravován. V případě přidání více slovníků do jednoho projektu to ale nedává smysl a projekt by měl být přejmenován. To je možné udělat v kartě projektu kliknutím na ikonu tužky u názvu projektu.



Informační koncepce ČR ✎


 Tomáš Šedivec
Modified 8 hours ago

 EDIT TERMS

 EDIT RELATIONS

 PUBLISH

 DELETE PROJECT

 Modifies

Informační koncepce ČR

 REMOVE

ADD VOCABULARY

Obrázek 11: *Přejmenování projektu je možné kliknutím na ikonu tužky vedle jejího zobrazeného názvu.*

Aplikace Termlt - nová verze

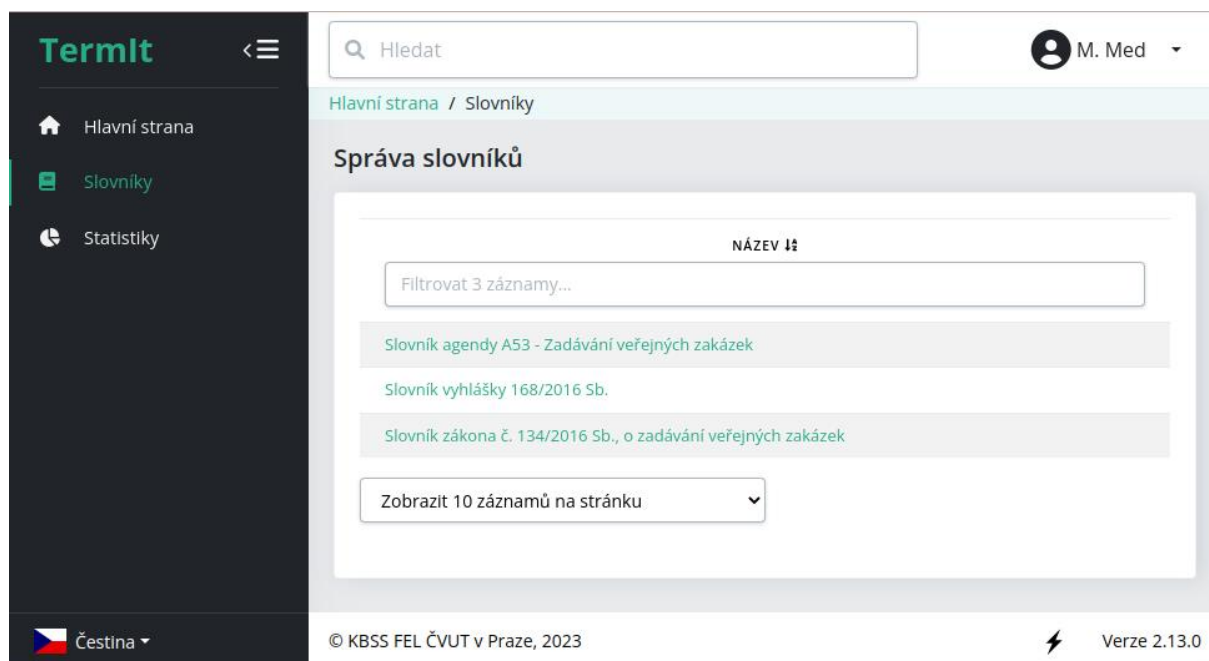
Jedním z důležitých kroků změny architektury byla implementace podpory sjednocené struktury architektury v aplikaci Termlt. Součástí tohoto úkolu je sjednocení backendové části aplikace se standalone verzí, která je provozována na Institutu plánování a rozvoje hlavního města Prahy (IPR). Důsledkem sjednocení bude jednodušší implementace změn a výhoda udržování pouze jednoho repozitáře pro backend aplikace Termlt. Nevýhodou je to, že stávající verze architektury nepodporuje všechny funkcionality, které nabízí frontend aplikace Termlt ve standalone verzi (pro pořádek, všechny funkce nepodporoval ani starý frontend v rámci KODI). Nepodařilo se tedy sjednotit celou frontendovou část a stále je potřeba udržovat dva repozitáře. Ale nová verze frontendu vychází ze standalone aplikace a funkce, které architektura nepodporuje, jsou v nové verzi skryté. Jejich následná implementace souvisí toliko s podporou v architektuře.

Základní a nejdůležitější změnou v nové verzi je to, že aplikace nepřebírá identifikátor pracovního prostoru, ale seznam slovníků (resp. jejich draft verzí). Předávání slovníků je na první pohled vidět v URL adrese, ale ve skutečnosti je uložené i v rámci session. Znamená to, že si aplikace slovníky pamatuje i v okamžiku, kdy z HTTP GET requestu zmizí.

Příkladem předání seznamu jednotlivých draft verzí slovníků Agendy A53, vyhlášky č. 168/2016 Sb. a zákona č. 134/2016 Sb., o zadávání veřejných zakázek je následující URL:

<https://slovník.gov.cz/modelujeme/v-nastroji/termit/#/vocabularies?vocabulary=https://slovník.gov.cz/legislativní/sbírka/168/2016/verze/0474d597-eca8-4db0-ac0c-05eb0a2302bb&vocabulary=https://slovník.gov.cz/agendový/slovn-k-agendy-a53-zad-v-n-ve-ijn-ch-zak-zek/verze/73558a4f-6a39-41de-9e7f-57591297882b&vocabulary=https://slovník.gov.cz/legislativní/sbírka/134/2016/verze/6a3074df-d8e6-4d44-ac25-303f416cd80d>

Úvodní stránka aplikace Termlt s předanými uvedenými draft verzemi slovníků je na následujícím obrázku:



Obrázek 12: Seznam slovníků projektu otevřených v aplikaci Termlt.

Většina problémů, které nastaly při změně architektury spočívaly v tom, že standalone verze Termlta pracuje nad živou databází, zatímco nová architektura umožňuje pracovat s kanonickým slovníkem a libovolným počtem jeho draft verzí. Tím pádem může dojít k situaci, že jeden pojem má několik reprezentací – verzí. Tato komplikace může nastat na několika místech, ale vždy se jedná o propojení mezi pojmy. Při každém výskytu tohoto problému bylo potřeba pojem vybírat nejen podle identifikátoru, ale i podle slovníkového kontextu, ve kterém se pojem nachází. To částečně zkomplikovalo implementaci, ale problém je vyřešen.

Nové funkce

Mezi nové funkce patří všechny funkce, které jsou obsaženy ve sjednoceném backendu a které nebyly omezeny novou architekturou nebo frontendem. Jedná se především o:

- podporu komentářů včetně přehledu komentovaných pojmů a odpovědí na komentáře,
- historii změn,
- validaci úplnosti pojmů,
- stav pojmu (rozpracovaný nebo schválený).

V současné verzi chybí podpora uložených dokumentů a textové analýzy. Důvodem je nutnost řešení správy dokumentů v rámci celé výrobní linky, tedy nejen aplikace Termlt. Dokumenty, které se objevují ve výrobní lince by měly mít vlastní správu, ať už se jedná o textové nebo jiné dokumenty přiřazené ke slovníkům nebo například diagramy z aplikace OntoGrapher.



● Část veřejné zakázky je zadávána v zadávacím postupu [IRI](#)

Definice

Byl zahájen zadávací postup na zadání části veřejné zakázky

Typ pojmu [Typ vztahu](#)

Pojmy se stejným významem

Nadřazené pojmy [Veřejná zakázka](#) [L-SGOV-134/2016](#)

Podřazené pojmy

Související pojmy [Část veřejné zakázky](#) [L-SGOV-134/2016](#)
[Zadávací postup](#)

Doplňující poznámka

Slovník [Slovník agendy A53 - Zadávání veřejných zakázek](#)

Stav pojmu [ROZPRACOVANÝ](#)

[Komentáře](#) 0 [Další atributy](#) 0 [Kontrola](#) [Historie změn](#)

Obrázek 13: *Detail pojmu v aplikaci TermIt.*

Uživatelské prostředí

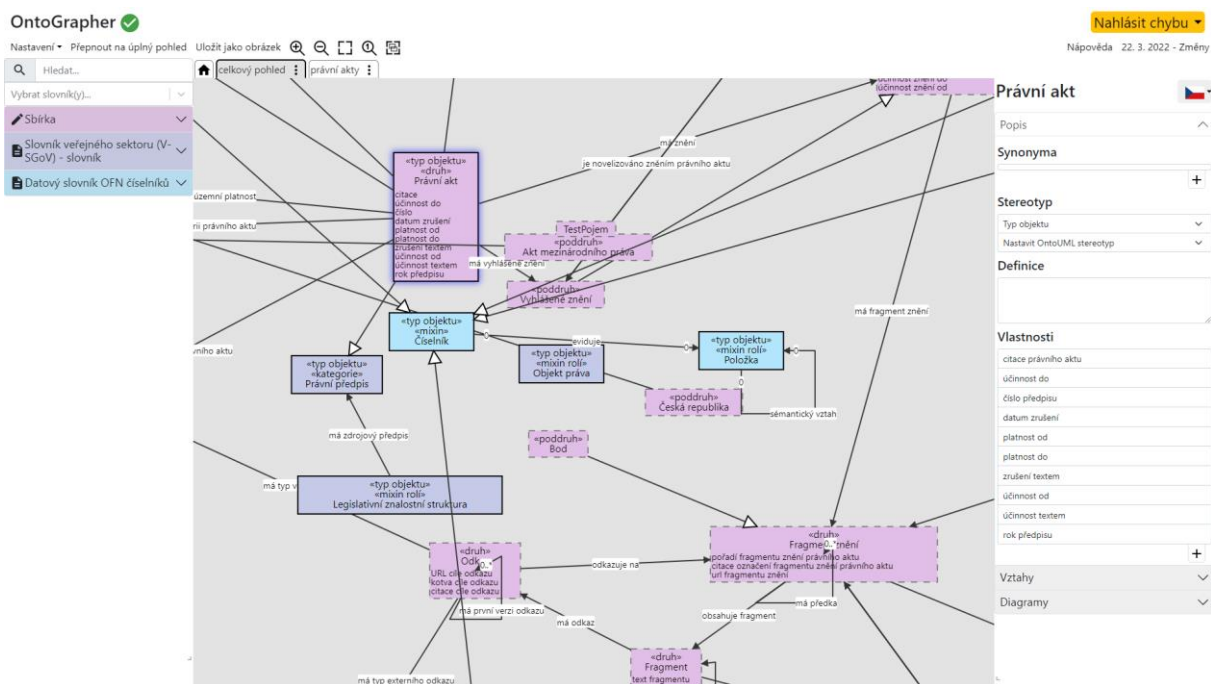
Nová verze uživatelského prostředí odpovídá frontendové verzi standalone aplikace TermIt s několika výjimkami.

- Standalone TermIt pracuje živě nad všemi slovníky v databázi. Nový TermIt pracuje pouze se slovníky, které jsou mu předány a pouze ty vypisuje v seznamu slováků. Nicméně statistiky jsou zobrazovány nad celou SSP cache.
- Byla odstraněna karta Dokument. Je to způsobeno absencí vazby mezi slovníkem a dokumentem v architektuře výrobní linky.
- Bylo odstraněno tlačítko Analyzovat u jednotlivých slovníků. Analýza definic probíhá automaticky a s absencí vazby mezi slovníky a dokumenty nedává zatím smysl analyzovat dokumenty.
- Změnila se barevná identita aplikace TermIt ze světle modré na zeleno černou.

Ukázka nového uživatelského prostředí je na obrázcích 12 a 13.

Aplikace OntoGrapher - Nová verze

Od verze OntoGrapheru vydanou pro výstup C5V3 se s dalším vývojem nástroj vylepšil v několika aspektech - kromě podpory již zmíněné upravené architektury nebo oprav chyb hlášených uživateli v průběhu používání nástroje se také podařilo obohatit nástroj o další uživateli žádané funkce, zvýšit stabilitu, a upravit uživatelské prostředí.



Obrázek 14: Ukázka nového uživatelského prostředí aplikace OntoGrapher.

Nové funkce

OntoGrapher s pomocí úpravy architektury SSP nyní při publikaci navíc ukládá a načítá diagramy a další informace (úplné kardinality, zobrazované jméno). Uživatel tedy může pokračovat v práci na slovníku/slovnících i po publikaci ve stejném stavu, jako předtím.

Samotné diagramy se mohou exportovat jako obrázek, tj. buď rastrovou verzí jako soubor PNG nebo vektorovou verzí jako soubor SVG. Snímek má podle verze bílý/průhledný podklad, aby se dal jednoduše zakomponovat do dalších prací, například článků, zpráv, či prezentací. Uživatel si může zvolit, zda-li do SVG exportu přidat i další informace o diagramu - jmenovitě IRI pojmů a forma zobrazení diagramu (více viz Uživatelské prostředí).

Uživatelské prostředí

Vzhled a průběh používání nástroje prošly významnou evolucí. Cílem je zpříjemnění zkušenosti uživatelů i navržení rozhraní tak, aby některé elementární chyby uživatele, které by vedly k nevalidnímu modelu, nešly provést. Přidaly se animace a popisky, a zjednodušily se některé interakce, aby bylo jednoznačnější, pro co která část uživatelského prostředí slouží.

Pro práci se samotnými diagramy a jejich metadaty vznikla nová "domovská karta". Je reakcí na zpětnou vazbu uživatelů v důsledku funkce ukládání do a načítání diagramů z SSP, kde velmi rychle vzniklo velké množství diagramů, které nešlo jednoduše spravovat. V této kartě může uživatel

- Otevírat a zavírat diagramy, tj. zobrazovat/schovávat je na liště s diagramy,
- vytvářet a mazat diagramy,
- upravovat metadata diagramů (název, popis, asociované slovníky).



Při prvním otevření draft verzí slovníků si tedy uživatel otevře/vytvoří jen ty diagramy, se kterými chce pracovat bez toho, aby mu zbylé diagramy plnily lištu.

Práce v samotných diagramech (tj. na “plátně”) se upravila – interakce s myší jsou lehce pozměněné; byla zavedena podpora multi-touch touchpadů (např. u MacBooků) a přidala se funkce multi-selekce – možnost vybrat více pojmů/vztahů a vykonávání některých akcí najednou.

OntoGrapher poskytuje uživatelům dva způsoby zobrazení diagramu: Úplný pohled, ve kterém jsou vidět všechny pojmy (včetně vlastností a vztahů) a spojení na základní úrovni, a Kompaktní pohled, ve kterém se vlastnosti zobrazují pod pojmy a vztahy se místo přesné reprezentace zobrazují jako jednoduché “čáry”.

V nynější verzi uživatelského prostředí se Kompaktní mód vyzdvihuje jako primární mód zobrazení modelu. V tomto módu je již možné dělat všechny úpravy, které předtím šly provádět pouze v Úplném pohledu - tj. plná manipulace s vlastnostmi, spojení s matčíným pojmem včetně kardinalit, a mazání vztahů.

Další vývoj

V současné verzi OntoGrapheru se při každé změně ukládá do metadat slovníku informace o poslední změně - datum a čas změny s jejich autorem, ale ještě ne povaha samotné změny. V budoucnosti by se tato funkce mohla rozšířit o ukládání i této změny a dále i všech předchozích změn v době mezi vytvořením draft verze a publikací slovníku. Tato podrobnější data by se dala využít pro další vylepšení, např. “cestování v čase” mezi změnami pojmů/vztahů/diagramů apod. pro ovrácení nechtěných změn nebo lepší pochopení, jací autoři dělali jaké změny.

Očekávají se i další úpravy uživatelského prostředí na základě následných zkušeností uživatelů. Kromě toho se plánuje i rozšíření “domovské karty”, např. o sekci pro diskuzi práce na draft verzi slovníku či prohlížení výše zmíněných změn.

Dále je ve vývoji funkce pro export diagramu i do výpisu pojmů v něm - nyní se uvažuje o formátech TXT a CSV, kde by každý formát měl jiný způsob a jiné množství informací o daných pojmech. U TXT by šlo o jednoduchý výpis napodobující ten z výstupu C2V11, sekce Odvození evidence údajů agendy, příklad z bodu 4. S formátem CSV (nebo XLSX) se export kromě názvu pojmu obohatí o definici, zdroj (tj. reference na původ pojmu) a typ - stejným způsobem budou popsány i vlastnosti a vztahy.

Současně v OntoGrapheru chybí podpora pro Registr práv a povinností, jednak ve formě výše zmíněného exportu pojmů, ale především v možnosti používat jinou sadu stereotypů - základní slovník pro tyto účely nestačí. Pojmosloví pro modelování RPP (vycházející z metodiky C2V11) a základního slovníku (vycházejícího z UFO) se také někdy překrývá, což je matoucí pro uživatele. Ideálně by OntoGrapher měl podporovat vyměnitelné top-level ontologie, ve kterém mají stereotypy i vlastní pravidla, které upřesňují jejich použití a podle kterých se mohl tvořený model validovat.



6. Nástroj pro kontrolu a publikaci slovníků

Jak bylo řečeno výše, při zaměření se na slovník jako first class citizen dochází k nutnosti rozlišovat mezi různými draft verzemi stejného kanonického slovníku. Po změně architektury je dalším krokem usnadnění publikace slovníků uživatelsky přívětivým způsobem a to včetně řešení případných konfliktů. Návrhem a tvorbou prototypu řešení se v rámci své diplomové práce zabývají Bc. Filip Kopecký a Bc. Michal Švagr.

Popis současného stavu

Současný proces publikace draftů slovníků začíná stisknutím tlačítka "Publikovat" v Kontrolním panelu. Po stisknutí tlačítka se spustí sada aktivit, které převedou draft slovníku do podoby struktury kanonického slovníku a nahraje se do GitHub repozitáře SSP ve formě takzvaného Pull Requestu. Podobná akce se stane také s drafty příloh slovníku. GitHub notifikuje uživatele zodpovědné za revize změn o novém Pull Requestu, kde mohou vidět změny provedené v draftu od aktuálního kanonického slovníku v podobě textových rozdílů RDF. Mohou se tedy pak rozhodnout, zda změny zamítnou a zašlou autorovi připomínky nebo změny schválí a tím se propíší do SSP repozitáře nedlouho poté do SSP cache.

Tento proces má v sobě několik úskalí. Jak jste si mohli všimnout zahrnuje využití nástroje GitHub třetí strany. Tento nástroj využívá jiné přihlášení než Výrobní linka, je tedy zapotřebí, aby byli uživatelům zodpovědným za revize změn vytvořeny účty a přidělena patřičná práva. Propojení účtů mezi těmito dvěma autorizačními službami v současné době také neexistuje, což zhoršuje přehlednost a dohledatelnost komunikace vztažené k provedeným změnám. Přesto že GitHub podporuje v Pull Requestech komentování i na úrovni jednotlivých řádků RDF ne všichni uživatelé Výrobní linky mají zapnuté oznámení z GitHub je tedy aktuální praxí zasílat dotazy na změny nebo vysvětlení zamítnutí změn pomocí emailu. Tím se komunikace úplně odpoutává od Výrobní linky a je tedy dohledatelná pouze v emailových klientech uživatelů. Zasílané emaily jsou navíc psány ručně s tím, že je zapotřebí specifikovat o jaké části modelu a změně se komunikuje.

To nás dostává k dalšímu problému revize změn, které uživatelé vidí v podobě RDF. Kvůli faktu, že RDF serializace do Trig (aktuálně používané pro reprezentaci kanonických slovníků) není normalizovaná i malá změna může způsobit velké množství přesunů, které jsou pak v textové reprezentaci GitHubem zvýrazňovány jako změny. Je tedy zapotřebí aby se uživatelé zodpovědní za revize změn orientovali nejen v modelování slovníků a nástroji GitHub, ale aby také dobře znali syntax RDF, rozpoznali změny od přesunů a pochopili význam dopadu změn na celý slovník i bez vizuální reprezentace propojení jednotlivých pojmů.

Návrh nástroje

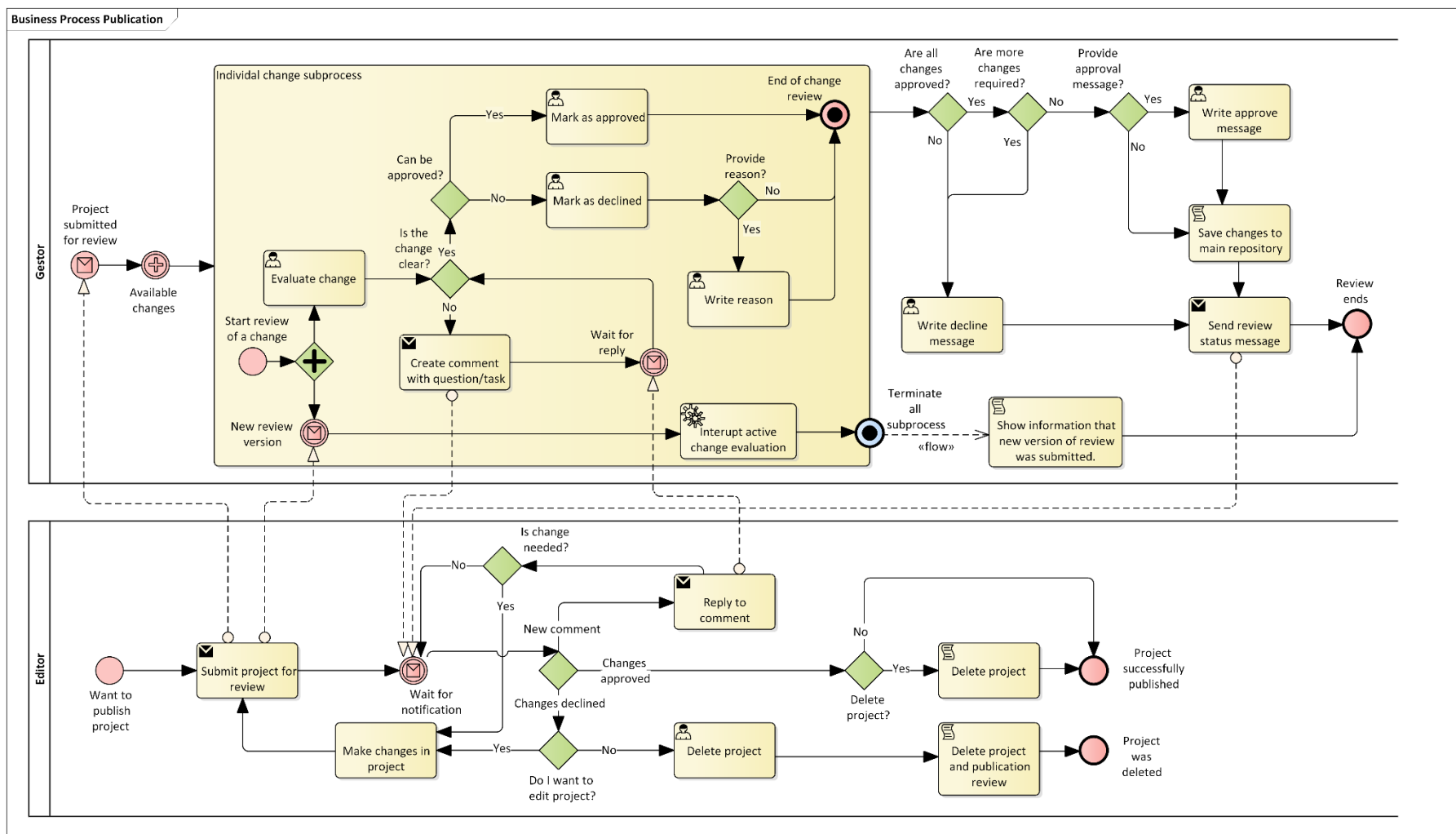
Je tedy navrhován nový nástroj jménem CheckIt, který má za cíl tyto problémy vyřešit. Bude se tedy jednat o nástroj umožňující správu rolí uživatelů zodpovědných za revize slovníků (z pohledu této revize), komunikaci ohledně jednotlivých změn v podobě komentářů v samotné



aplikace a zobrazování změn ve formě diagramů (pro změnu vztahů) a barevného odlišení nových/změněných/smazaných parametrů pojmů.

Byl navrhnout nový proces opět začínající stisknutím tlačítka v Kontrolním panelu, který však notifikuje uživatele Výrobní linky zodpovědné za revizi (Gestory) slovníku/ů určených k publikaci. Ty v nástroji CheckIt zrevidují a následně schválí nebo zamítnou každou změnu (např. nový název a popis pojmu). Když bude mít k jakékoliv změně dotaz napíše k ní komentář a uživatelé Výrobní linky zodpovědní za vytvoření změny budou notifikováni. Na komentář mohou odpovědět nebo v případě nutnosti upravit svůj model a aktualizovat verzi k publikaci. Ve chvíli, kdy budou všechny změny schváleny, může je Gestor publikovat. Publikace znamená promítnutí změn do kanonických slovníků v SSP. V případě, že Gestor se změnou nesouhlasí a zamítne ji, může revizi ukončit s udáním důvodu zamítnutí. Tyto důvody jsou následně zaslány autorům změn, kteří se mohou rozhodnout, zda chtějí draft slovníku smazat a začít se změnami od začátku nebo pokračovat v úpravách. Procesní diagram nástroje CheckIt je zobrazen na obrázku 15.

Pro popis změn byla vytvořena nová ontologie pro popis změn dostupná na <https://github.com/kbss-cvut/popis-zmen-ontology> a její schéma je na obrázku 16. Hlavní entitou je Publikační kontext, ve kterém jsou sdruženy všechny změny. Jednotlivé změny mají přiřazen jeden z následujících typů: vytvořeno, odstraněno, upraveno nebo vráceno. Změny jsou reprezentovány jako RDF trojice (triply). V případě změny (tedy ne vzniku nového triplu) mohou obsahovat volitelný odkaz na nový objekt. Změny odkazují na kontext (např. slovník) ze kterého jsou. Schvalovací proces provádí uživatel s rolí Gestora pro daný slovník, tedy s přiřazeným slovníkovým kontextem.

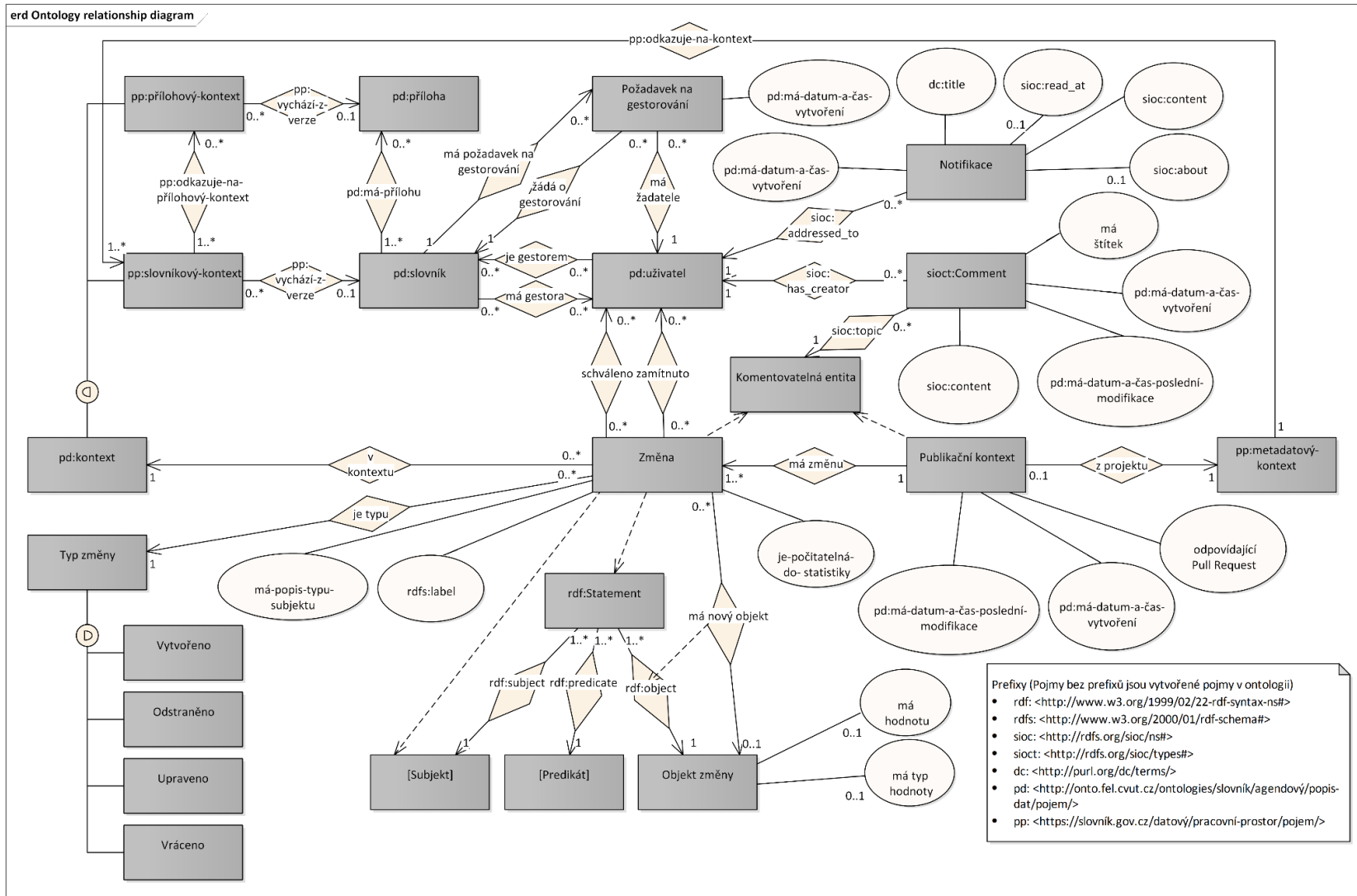


Obrázek 15: Procesní diagram zpracování změn v nástroji CheckIt. Znárodnující zaslání projektu k revizi editorem, revizi změn gestory a následné schválení/zamítnutí projektu s příslušnými notifikacemi o akcích. Diagram také znázorňuje možnost diskuze nad jednotlivými změnami



Evropská unie
Evropský sociální fond
Operační program Zaměstnanost

a možnost zaslání upraveného projektu v průběhu revize.





Evropská unie
Evropský sociální fond
Operační program Zaměstnanost

Obrázek 16: *Schéma ontologie pro popis změn.*



Proběhla také rešerše nástrojů pro porovnávání RDF. Z nástrojů Python RDFLib, Protégé Desktop s pluginem na porovnávání RDF, OWLDiff, porovnávání Normalizovaného RDF a Apache JenaAPI byly vybrán Apache JenaAPI, kvůli jeho výkonnosti na větších datových sadách, objektové reprezentaci RDF dat umožňující flexibilní vyhledávání parametrů pojmů k reprezentaci změn v kontextu okolí, schopnosti rozpoznávat nezměněné blank nody a jeho využití již v současných nástrojích Výrobní linky.

Prototyp nástroje CheckIt určený pro obhajobu diplomových prací běží na adrese <https://onto.fel.cvut.cz/modelujeme/v-nastroji/checkit/> se stejnými přihlašovacími údaji, jaké platí pro demo verzi výrobní linky.

7. Funkční požadavky pro další vývoj výrobní linky

Další vývoj Výrobní linky je podrobně rozmyšlen a naplánován. V následující kapitole je seznam funkčních požadavků pro další vývoj. Některé z nich stihly být naimplementovány ještě v závěrečné fázi projektu (duben/květen 2023). Z důvodu zachování integrity textu jsou přesto uvedeny v tomto výčtu.

Autentifikace pomocí Centrálního agendového a informačního systému (CAIS)

Během roku 2023 je nově zaváděn Centrální agendový a informační systém (CAIS). Jeho součástí by měla být identifikace zaměstnanců veřejné správy. Jedním z funkčních požadavků pro implementaci do systému Výrobní linky je integrace přihlašování uživatelů z veřejné správy pomocí autentizační služby systému CAIS. V současné době veřejná správa používá autentizační systém Jednotného identitního prostoru informačních systémů veřejné správy a Katalog autentizačních a autorizačních služeb (JIP/KAAS). Jeho použití pro informační systémy, kde interní uživatelé jsou zaváděni úřady, které nejsou správci těchto informačních systémů je použití autentizačního systému JIP/KAAS povinné. Podle všeho se to ale Výrobní linky netýká, přestože to jistě přináší řadu výhod. Jedná se především o využití stávajícího systému, který identifikuje úředníky s jejich oblastí působnosti. Díky tomu by bylo možné přiřazovat uživatelské role ke konkrétním slovníkům ve správě daného úřadu.

Správa uživatelských rolí pro jednotlivé slovníky

V současné verzi prototypu Výrobní linky nejsou uživatelské role řešeny pro jednotlivé slovníky, nebo jejich skupiny. V praxi to znamená, že kdokoliv se přihlásí do výrobní linky má právo měnit všechny kopie slovníků ve všech projektech. V rámci prototypu to není problém, pro plné nasazení jako služby (nebo systému) veřejné správy je nutné tento problém vyřešit. Předpokládáme vznik několika uživatelských rolí. V praxi by se mělo jednat o právo prohlížet



slovníky (to by asi nemělo být omezeno), právo slovníky editovat a dále vlastnické právo, umožňující publikaci a správu práv ostatních uživatelů. Nabízí se několik řešení:

Uživatelské skupiny

Uživatelé budou organizováni ve skupinách podle pracoviště. Uživatelská práva ke kopiím slovníků jsou přiřazena konkrétní skupině a všichni její členové mohou ke slovníku přistupovat, editovat ho, komentovat nebo publikovat.

Zakladatel je vlastníkem

Uživatel, který vytvořil kopii slovníku pro ní získává vlastnické právo. Ostatním uživatelům může udělovat právo k editaci, nebo vlastnické právo předat.

Vlastnictví v rámci skupiny

Podobně jako v prvním případě mají členové skupiny přístup k projektu a všem slovníkům v něm, ale zároveň mají mezi sebou rozdělená práva. Jeden nebo více uživatelů jsou vlastníky skupiny a mohou ostatním uživatelům přidávat práva k editaci všech slovníků v organizaci, nebo přímo ke konkrétním kopiím konkrétních slovníků.

Všechny řešení vyžadují zavedení uživatelských rolí, jejich přiřazování uživatelům a tedy i úpravu ontologie. Kromě toho je potřeba zavést systém pro správu uživatelských rolí a to mimo autentizační službu, která by měla být na systému nezávislá.

V delším horizontu by správa slovníků mohla být přiřazována na základě pracovního zařazení uživatele a jeho vztahu k agendě/zákonu. To souvisí s propojením s autentizačním Centrálním agendovým a informačním systémem, u kterého se předpokládá, že bude organizovat uživatele ve vztahu k agendám a zákonům, které se dotýkají jeho pracovní náplně.

Správa dokumentů

V jedné z předchozích verzí systému Termlt byla funkce anotace dokumentu a jeho obsahu pojmy ze slovníků a textová analýza dokumentů. Diagramy vytvářené v nástroji OntoGrapher jsou v tomto kontextu také považovány za dokumenty. Z pohledu architektury celého systému Výrobní linky dává smysl aby byly dokumenty spravovány v samostatném nástroji, ke kterému budou přistupovat ostatní nástroje výrobní linky. K tomuto účelu je možné použít některé existující nástroje pro správu dokumentů. Cílem je, aby bylo možné vytvořit entitu typu dokument (existuje v ontologii a je k tomuto účelu využívána ve standalone verzi systému Termlt), opatřit ji metadaty a dalšími anotačními vlastnostmi a přiřadit k ní konkrétní soubory (ve formátech HTML, PDF, SVG, ale i dalších). Nástroj musí zároveň umožnit nahrát a uchovat tyto soubory a také je náležitě oannotovat. Součástí řešení v dlouhodobějším horizontu musí být integrace s eSbírkou.

Textová analýza dokumentů v nástroji Termlt

Po vyřešení předchozího požadavku je možné v upravené podobě vrátit do nástroje Termlt anotační komponentu. Analytická komponenta systému Termlt umožňuje analyzovat



dokumenty a k nim náležící soubory a dokumenty k nim přiřazené. Tato komponenta analyzuje text (v současné době ve formátu HTML), vyhledává v něm výskyty pojmů z vybraných slovníků a na základě lematických pravidel navrhuje možné výskyty pojmů, které se v textu nachází a ve slovníku ne. Pojmy, jejichž výskyt je v textu potvrzen jsou dále použity k anotaci samotného dokumentu.

Pojmy

Článek 3

Pojmy Metropolitního plánu

(1) Pojmy užívané v Metropolitním plánu a samostatně nevymezené vycházejí z platných právních předpisů, zejména ze znění zákona č. 183/2006 Sb., o územním plánování a stavebním řádu (stavební zákon), ve znění pozdějších předpisů (dále jen „stavební zákon“), ze speciálních právních předpisů dotýkajících se územního plánování a staveb a z Pražských stavebních předpisů^[2] (dále také „PSP“), nejsou-li takové pojmy samostatně vymezeny ani stanoveny právními předpisy, jsou použity v jejich obvyklém obecném významu.

(2) Jinou stavbou se pro účely Metropolitního plánu rozumí mimo budov všechny stavby, zařízení a opatření ve smyslu stavebního zákona^[3].

(3) Budovou^[4] se pro účely Metropolitního plánu rozumí nadzemní stavba spojená se zemí zpravidla pevným základem, prostorově kompaktní a vůči vnějšímu prostředí převážně uzavřená obvodovými stěnami a střešní konstrukcí. Budovy vytvářejí zástavbu a jejich umístění je regulováno Metropolitním plánem.

(4) Zástavbou se pro účely Metropolitního plánu rozumí budova nebo soubor budov rámci jednoho nebo více stavebních bloků.

Článek 4

Základní pojmy

Základními pojmy se pro účely Metropolitního plánu, dalších územně plánovacích dokumentací a územně plánovacích podkladů pro území Prahy rozumí:

a) Metropolitním plánem Územní plán hlavního města Prahy pořizovaný pro celé území Prahy,^[5]

Obrázek 17: Nástroj textové analýzy dokumentů na příkladu textu Metropolitního slovníku hlavního města Prahy ve verzi 3.4 ze standalone verze TermIta, která je provozována pro Institut plánování a rozvoje hlavního města Prahy Elektrotechnickou fakultou ČVUT.


Seznam všech slovníků do ShowIta

Nástroj ShowIt slouží k prohlížení obsahu Sémantického slovníku pojmů a je plně popsán v kapitole [ShowIt](#). Na základě zpětné vazby uživatelů se ukázalo, že samotné vyhledávání v pojmech nemusí všem vyhovovat a že část uživatelů vyžaduje možnost vyhledávání v seznamu všech slovníků, které jsou součástí SSP. Součástí úkolu je citlivě zapracovat ovládací prvek do uživatelského prostředí nástroje ShowIt. Tento požadavek stihl být zapracován ještě v rámci projektu.



ShowIt


Hledat význam není těžké. Stačí vědět, kde hledat.



Zadejte hledané slovo

Navrhovaná slova: [Agenda](#) [Budova](#) [Vozidlo](#) [Závada](#)
Nebo si prohlédněte [dostupné slovníky](#)

Tato aplikace je udržována v rámci projektu
OPZ č. CZ.03.4.74/0.0/0.0/15_025/0013983.



Evropská unie
Evropský sociální fond
Operační program Zaměstnanost

ShowIt

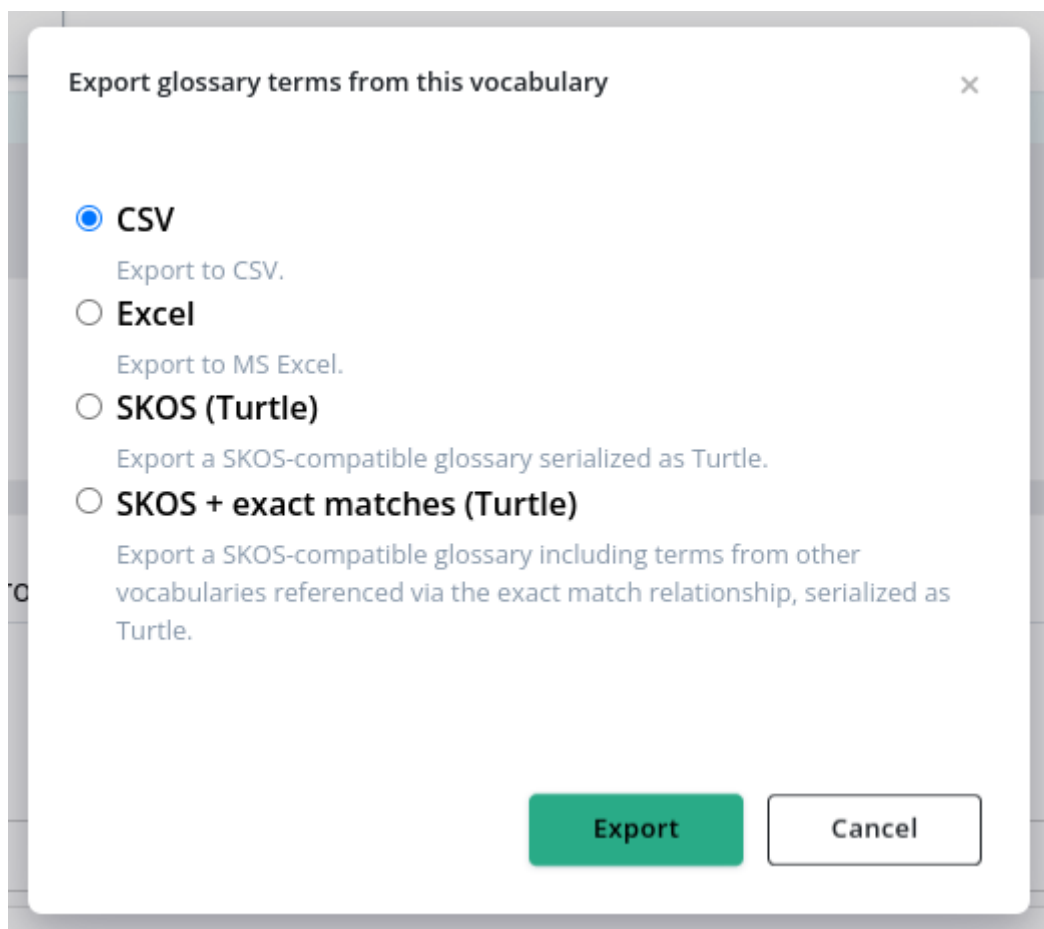
Slovníky

- 440-test
- COUNCIL DIRECTIVE 1999/37/EC
- COUNCIL DIRECTIVE 1999/37/EC on the registration documents for vehicles
- ČSN 73 0831
- ČSN 73 0875
- ČSN 73 2604 Ocelové konstrukce – Kontrola a údržba ocelových konstrukcí pozemních a inženýrských staveb
- ČSN 75 6081
- ČSN ISO/IEC 20000-2 - Informační technologie - Management služeb
- Datový slovník číselníku jazyků
- Datový slovník číselníku států

Obrázek 18: Zapracování zobrazení dostupných slovníků do úvodní stránky Výrobní linky a seznam slovníků s možností vyhledávání.

Export/import slovníku z/do VL, příp. jednotlivých nástrojů.

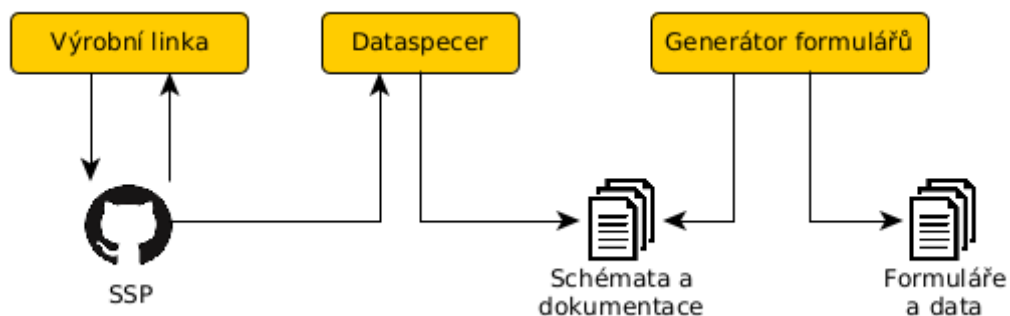
Úpravy nástroje TermIt při aplikaci nové architektury způsobily vyřazení i další komponenty - exportu slovníků. Tato komponenta byla ještě v rámci do projektu do nástroje vrácena. Export je umožněn ve formátu CSV, Excel, SKOS v serializaci Turtle nebo SKOS v serializaci Turtle včetně vazeb exact match.



Obrázek 19: Dialogové okno v aplikaci Termt umožňující export slovníku ve vybraných formátech.

Integrace nástroje Dataspecer do Výrobní linky

Nástroj Dataspecer je jedním z výstupů projektu KODI, konkrétně jeho cíle 3. Nástroj je dostupný na adrese <https://slovník.gov.cz/generujeme/> a umožňuje na základě slovníků ze Sémantického slovníku pojmů vytvářet datová schémata a dokonce i jejich dokumentaci. Nástroj je popsán ve výstupu C3V8 - [Návrh zajištění podpory dokumentace datových sad generované z ontologických konceptuálních datových modelů agend v NKOD.](#)



Obrázek 20: Z výstupů Výrobní linky, tedy Sémantického slovníku, je pomocí nástroje



Dataspecer možné generovat datová schémata včetně dokumentace. Dalším logickým krokem je na základě vygenerovaných schémat generovat formuláře těmto schématům odpovídající. Vyplňováním formulářům pak poskytovatelé získají data, v ideálním případě propojená a otevřená a zároveň odpovídající datovým schématům vycházejícím ze slovníku.

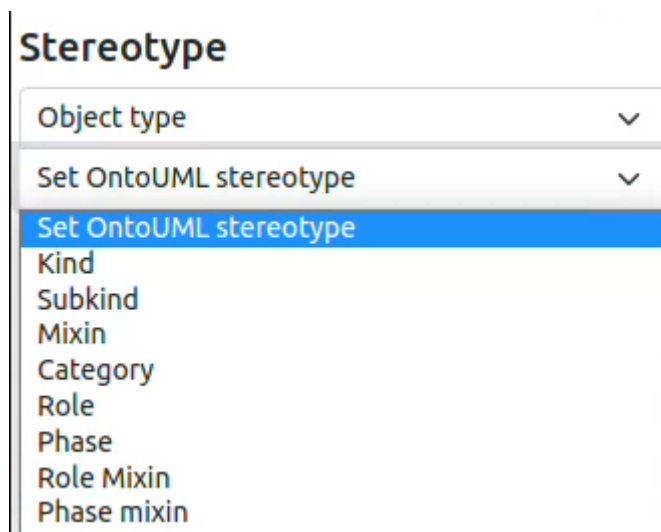
Cílem tohoto požadavku je zařazení nástroje Dataspecer do workflow za výrobní linku tak, jak je vidět na obrázku 20. Souvisí s tím i následující požadavek na Generátor formulářů. Zařazení dataspeceru vyžaduje změny na straně tohoto nástroje, jedná se především o přihlašování uživatelů za použití autentizační služby (ať už stávající, nebo pomocí Centrálního agendového a informačního systému) a použití uživatelských rolí, které by zabránilo tomu, aby byly schémata editovatelná všemi uživateli. Na straně Výrobní linky očekáváme zařazení tlačítka do Kontrolního panelu, které uživatele přesměruje do nástroje Dataspecer. Další možností by bylo zařadit tlačítka k detailu slovníku v nástroji ShowIt. Dataspecer by načítal rovnou slovník, ze kterého do něj byl uživatel přesměrován.

Dokončení a tvorba nástroje pro generování formulářů pro sběr dat odpovídajících SSP

Cílem je vytvořit nástroj Správce otevřených dat, který umožní tvorbu dat pomocí formulářů generovaných z OFN specifikací, nebo přímo z SSP. Hlavní scénář je tvorba dat dle OFN – aplikace by měla být umožnit vytvoření formuláře na základě specifikace OFN, umožnit uživateli tento formulář vyplnit a exportovat data kompatibilní s OFN a připravená pro publikaci do Národního katalogu otevřených dat (NKOD). Nástroj byl v rámci projektu vytvářen, ale nebyl dokončen. Pro dokončení nástroje je potřeba doplnit integraci s nástrojem Dataspecer – formulář půjde generovat z konfigurace projektu definovaném v Dataspeceru (jak je vidět na obrázku 20) a umožnit ukládání dat ve formátu JSON-LD, který odpovídá požadavkům na otevřená data ze strany OFN. Požadavky na funkcionalitu byly analyzovány v dokumentu [Generování formulářů z SGOV](#). Aktuální stav nástroje umožňuje zobrazení formulářů pro OFN Turistické cíle, Aktuality, Události a Sportoviště.

Rozšíření a validace stereotypů pro OntoGrapher a TermIt odpovídající RPP

V dokumentu [Metodika definice údajů vedených v agendě](#) jsou definovány stereotypy pojmů pro modelování pojmů agend. V současné podobě nástrojů ve Výrobní lince nejsou tyto stereotypy podporovány. V případě nástroje TermIt je podporována pouze základní sada stereotypů podle Unified Foundational Ontology (UFO). Pro nástroj OntoGrapher je podporována ještě základní sada stereotypů podle OntoUML.



Obrázek 21: Seznam stereotypů podle OntoUML v nástroji OntoGrapher.

Metodika obsahuje doporučené postupy pro definici údajů vedených v agendě v rámci procesu ohlašování a registrace agend podle zákona č. 111/2009. Pro popis pojmů v agendách je v dokumentu doporučen konceptuální model. Reprezentace pojmů do konceptuálního modelu probíhá následujícím způsobem:

- Pojem reprezentující **osoby** určitého druhu klasifikuje jako **typ subjektu**.
- Pojem reprezentující **věci** určitého druhu klasifikuje jako **typ objektu**.
- Pojem reprezentující určitou **rolí** osob klasifikuje jako **typ role**.
- Pojem reprezentující určitou **fázi** osob klasifikuje jako **typ fáze**.
- Pojem reprezentující **vztahy** určitého druhu mezi osobami nebo mezi osobami a věcmi klasifikuje jako **typ vztahu**.
- Pojem reprezentující **události** určitého druhu klasifikuje jako **typ události**.
- Pojem reprezentující **činnosti** určitého druhu klasifikuje jako **typ činnosti**.
- Pojem reprezentující **vlastnosti** určitého druhu charakterizující osoby, věci, role, vztahy, události nebo činnosti klasifikuje jako **typ vlastnosti**.

Stereotypy na pravé straně není v OntoGrapheru možné pojmům přiřadit. Pro zjednodušení modelování agend by bylo vhodné tyto stereotypy přidat do nástrojů Termlt i OntoGrapher. Všechny stereotypy by mělo být možné vztáhnout ke stereotypům dle UFO i OntoUML.

V nástroji OntoGrapher je seznam stereotypů řešen v JSON souborech.

Dalším problémem je validace. Dokument definuje podmínky, za jakých vznikají vztahy mezi pojmy s různými stereotypy. Validace může být řešena dvěma způsoby, buď na straně OntoGrapheru, nebo při publikaci na straně SGoV serveru. Podmínkou může být například to, že vztah směřuje vždy směrem od osoby k jiné osobě nebo k věci. Podle prvního způsobu by OntoGrapher upozornil na nesoulad v případě, že by tato podmínka nebyla dodržena. Validace při publikaci by nechala modeláře vymodelovat v OntoGrapheru téměř cokoli (pokud by to neporušovalo podmínky vnitřně nastavené v OntoGrapheru), ale publikace by neproběhla. V logu validace budou uvedeny důvody, proč validace neprošla. K validaci je vhodné použít např. SHACL.



Modulární rozšiřování stereotypů a jejich validace

Jedná se o rozšíření předchozího funkčního požadavku na možnost přidání libovolné sady pojmů jako stereotypů pro projmy ve slovníku. Pro libovolný slovník by mělo být možné vybrat sadu pojmů a používat ji jako sadu stereotypů v nástrojích výrobní linky. Modulárním způsobem bude možné použít jako stereotypy libovolné pojmy ze Sémantického slovníku pojmů a stejným způsobem vytvořit i větší počet sad stereotypů, a to i pro jediný slovník.

Konfigurace implicitního jazyka slovníku

Aktuálně celý deployment výrobní linky má implicitní jazyk pro tvorbu pojmů, který je aplikován na každý slovník. Do budoucna by bylo užitečné mít možnost vytvářet slovníky v různých jazycích (typicky angličtina nebo čeština). Implementace by byla možná se zachováním implicitního jazyka pro celé nasazení výrobní linky s možností měnit jazyk pro jednotlivé slovníky.

Integrace s eSbírkou, nástroj pro analýzu HTML stránek, nebo nástroj pro konverzi dokumentů do HTML podoby, která je analyzovatelná.

V rámci práce na projektu probíhala spolupráce na návrhu schémat datových sad v RDF formátech pro publikaci eSbírkou. Samotná eSbírka není v současné době ještě hotová, ale počítáme v dalším rozvoji výrobní linky s tím, že jednou bude hotová. Cílem našich návrhů bylo mimo jiné i to, aby bylo legislativní dokumenty eSbírkou možné anotovat pomocí pojmů ze Sémantického slovníku pojmů a na druhou stranu aby bylo přímo v dokumentech eSbírkou možné vybírat pojmy a vytvářet glosář. Pro integraci výrobní linky s eSbírkou bude potřeba použít anotační nástroj a nástroj textové analýzy (po vzoru podobných nástrojů ze standalone verze nástroje TermIt, viz kapitola [Textová analýza dokumentů](#)) nad HTML dokumenty samotné eSbírkou. Toho by bylo možné dosáhnout pluginem buď do internetového prohlížeče, nebo do samotné eSbírkou. Prototyp takového pluginu vznikl v rámci [studentské diplomové práce na Fakultě elektrotechnické ČVUT](#).

Batch editing

Při tvorbě slovníků pomocí nástrojů výrobní linky je řada atributů pojmů, které by bylo možné editovat dávkově. To by mohlo výrazně zefektivnit a zrychlit tvorbu slovníků v nástrojích. Jedná se například o hromadnou úpravu stereotypu pojmu nebo kardinality vztahu nebo typu vlastnosti.



8. Posouzení kvality výstupu na základě uživatelského testování

Na základě specifikace tohoto výstupu má být jeho kvalita posouzena třemi vybranými pracovníky Odboru hlavního architekta MVČR nebo jiných subjektů. Vybraní pracovníci už během projektu používali výrobní linku k modelování vybraných agend. Pro zhodnocení výstupu jim byly položeny tři otázky:

1. Popište scénář, k jehož řešení využíváte Výrobní linku. Které její nástroje využíváte k jakým činnostem?
2. Které části výrobní linky a jejich funkce oceníte nejvíce?
3. Které funkce Vám v současné podobě výrobní linky chybí? Které z nich jsou zásadní pro to, abyste mohli použít výrobní linku k řešení vámi uvedeného scénáře/scénářů?

Na základě odpovědí na tyto otázky byly vyhodnoceny funkce, které by do výrobní linky měly být přidány. Zároveň získáváme přehled o tom, k řešení jakých scénářů je výrobní linka používána a které její funkcionality jsou nejdůležitější.

Testování nakonec provedly čtyři osoby. Z výsledků testování vychází, že všichni uživatelé výrobní linky, kteří posuzovali kvalitu výstupu, používají výrobní linku jako celek, tedy TermIt pro tvorbu glosáře a OntoGrapher pro tvorbu významových vztahů mezi objekty, ačkoliv to popisují různě. Jejich cílem je vytvoření uceleného slovníku, nebo také tvorba grafického vyjádření modelu. Uživatelé pak nejvíce pracují v modelovacím nástroji OntoGrapher a používají ho i k tvorbě glosáře. Návrhů pro vylepšení výrobní linky se sešlo více.

Jako největší problém se jeví propisování změn z OntoGrapheru do nástroje TermIt. Uživatelům také chybí pro nástroje dostatečná nápověda nebo návod pro práci. Uživatelé by ocenili propojení nástrojů výrobní linky na zákonné normy nebo legislativní dokumenty. Dalším požadavkem je podpořit import/export v jednoduchém výměnném formátu (např. CSV). To bohužel z důvodu komplexity RDF není možné v takovém rozsahu, aby se těm neztrácely informace. Poslední připomínkou je nedokonalý export dat z výrobní linky směrem k registraci údajů do Registru práv a povinností.

Další připomínky, které testéři uváděli (například dočasné vypuštění exportu do CSV, excelu nebo do formátu Turtle, nebo propojení pojmů k jejich definici nebo výskytu v dokumentu) jsou buď již nově naimplementované, nebo se s nimi počítá. Návrhy na změny i hlášení chyb (bugů) probíhá nadále ve formě GitHub issues v repozitáři <https://github.com/datagov-cz/sgov-assembly-line/issues>.

9. Zdrojové kódy nástrojů výrobní linky

Výrobní linka je vytvářena jako Open source a zdrojové kódy jejích součástí jsou udržovány v repozitářích na GitHubu. Odkazy na zdrojové kódy všech nástrojů a návod na nasazení celé výrobní linky je README souboru v centrálním repozitáři <https://github.com/datagov-cz/sgov-assembly-line>.



Reference

- [1] - [Návrh zajištění podpory dokumentace datových sad generované z ontologických konceptuálních datových modelů agend v NKOD](#), výstup C3V8 projektu,
- [2] - [Generování formulářů z SGOV](#), detailní analýza požadavků na nástroj pro generování formulářů,
- [3] - [Metodika definice údajů vedených v agendě](#), výstup C2V11 projektu,