

Bc. Zdeněk
Častorál

Diplomová práce

Inženýrská informatika
Softwarové inženýrství
2021/2022

Vedoucí práce:
Ing. Petr Příbýl
CCA Group a.s.

Generování dokumentů z uživatelských šablon

Abstrakt

Tato diplomová práce se zabývá problematikou automatického generování dokumentů dosazováním dat do uživatelských šablon. Cílem této práce je vytvořit nový softwarový produkt či použít a upravit již existující softwarový produkt třetí strany, který umožňuje vytvářet a editovat šablony dokumentů a následně generovat dokumenty nad připravenou sadou dat do formátu Microsoft Word. Zadavatelem této práce je společnost CCA Group a.s.

Práce nejprve shrnuje obecnou problematiku automatizace dokumentů, především se zaměřuje na automatické generování dokumentů. Dále je analyzováno stávající řešení zadavatele včetně jeho nedostatků a poté jsou detailně specifikovány požadavky na nové řešení. V další části práce je zdůvodněn výběr varianty vytvoření nového softwarového produktu, poté je popsán jeho návrh a implementace. Na závěr je uveden průběh testování implementovaného softwarového produktu.

Úvod

Zadavatelská společnost realizuje řadu projektů v oblasti informačních technologií. Jeden z nich se zabývá správou a řízením dokumentů. Právě z tohoto důvodu vznikl požadavek na produkt, který bude umožňovat automatickou tvorbu dokumentů prostřednictvím dosazování dat do uživatelských šablon. Společnost CCA Group a.s. takový produkt řadu let vyvíjí a provozuje. V současné době uvažuje o kompletně novém softwarovém řešení realizovaném v moderních technologiích.

Cílem této práce je vytvořit nový softwarový produkt či použít a upravit již existující softwarový produkt třetí strany, který nahradí stávající řešení zadavatele, bude snadno rozšiřitelný a udržitelný a bude implementován v moderních technologiích.

Východiska, analytická část

Nejprve byla provedena analýza stávajícího řešení zadavatele, při které byly identifikovány nedostatky tohoto řešení. Na základě těchto poznatků byla vypracována specifikace požadavků na řešení nové.

Dále byly analyzovány existující softwarové produkty řešící požadovanou problematiku. Z této analýzy vyplynulo, že žádný z nich není vhodný pro potřeby zadavatele a došlo tedy k rozhodnutí implementovat vlastní řešení. Pro potřeby realizace bylo nutné vyřešit problematiku následujících oblastí:

- struktura vstupních dat a jejich popis,
- formální jazyk pro definici šablon,
- lexikální a syntaktická analýza šablon a jejich parsování,
- dosazení vstupních dat do šablon,
- uživatelské rozhraní editoru šablon.

Hlavní aspekty realizace

V realizační části práce byla definována struktura vstupních dat tak, aby byla dostatečně gene-

rická pro různorodá data. Jako formát zápisu byl zvolen JSON. Dále byla navržena gramatika formálního jazyka pro definici šablon, která zpracovává jazykové konstrukce pro: doplňování dat na určená místa v šabloně, definici direktiv, které ovlivňují způsob generování dokumentu a definici metod, které modifikují doplňovaná data.

Ukázka zdrojového textu

Příklad jazykové konstrukce pro definici direktivy foreach v šabloně dokumentu (slouží k realizaci opakovacího bloku).

```
foreach ($p in $file.participants)
  ${p.firstName} ${p.lastName}
#end
```

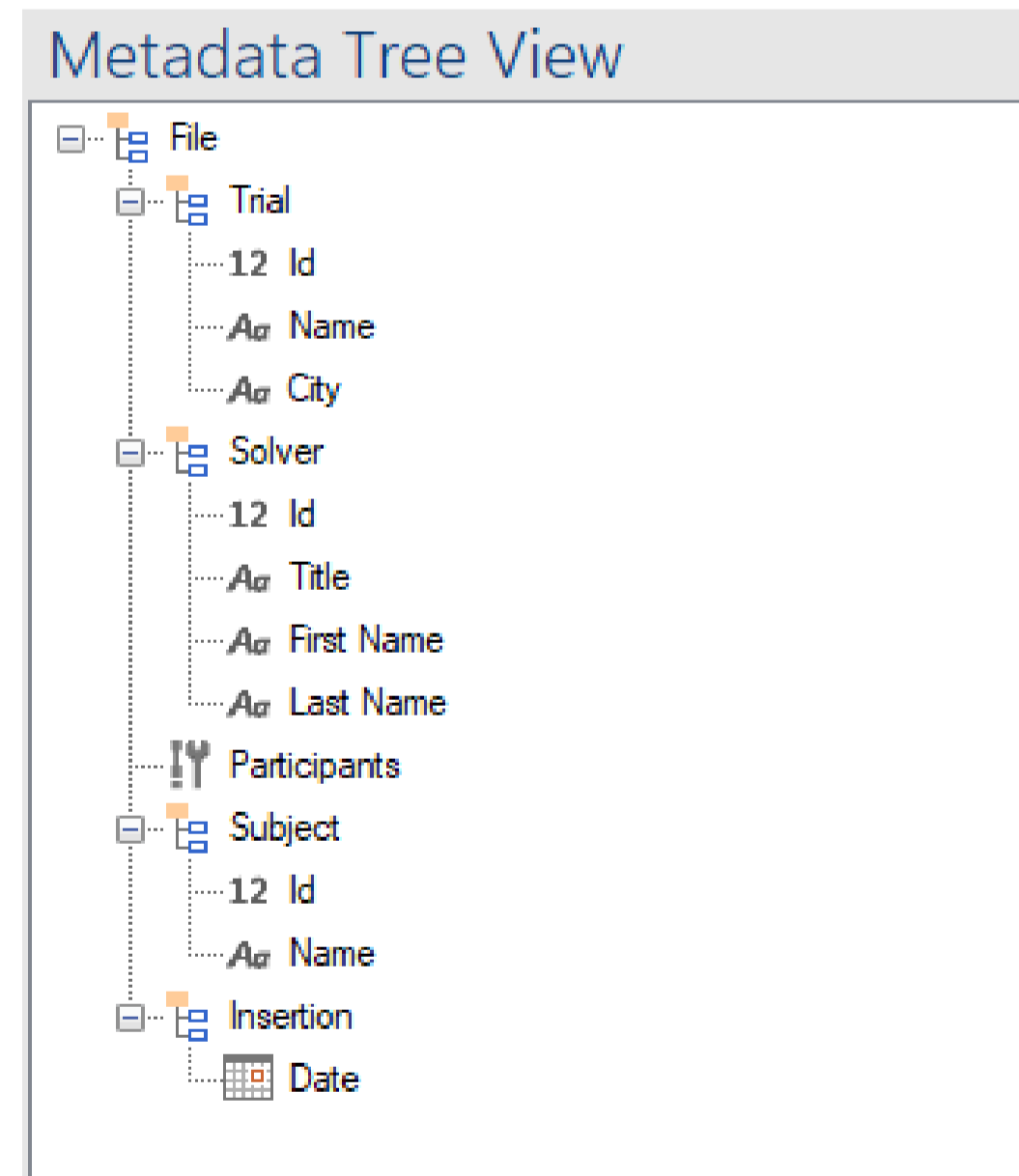
Pro syntaktickou a lexikální analýzu šablon a následné parsování byl použit nástroj ANTLR.

Vzhledem ke skutečnosti, že šablona dokumentu i samotný vygenerovaný dokument jsou ve formátu Microsoft Word, byla pro dosazování dat použita knihovna Microsoft Office Interop Word.

Jako uživatelské rozhraní pro vytváření a editaci šablon byl použit přímo produkt Microsoft Word, do kterého bylo implementováno rozšíření (Add-In) pomocí nástroje VSTO. Toto rozšíření umožňuje pohodlné vytváření šablon dokumentů včetně sledování syntaktických chyb.



Náhled nástrojové lišty rozšíření



Náhled zobrazení struktury vstupních dat

Dosažené výsledky

Implementovaný softwarový produkt byl otestován a je plně funkční. V současné době je připraven k nasazení do testovacího prostředí zadavatele.

Závěr

Cíl práce byl splněn. Výstupem je nový softwarový produkt, který umožňuje vytvářet a editovat šablony dokumentů a následně generovat dokumenty nad vstupními daty do formátu Microsoft Word. Do budoucna je plánováno nasazení do provozního prostředí a nahrazení stávajícího řešení zadavatele.