

Využití celexového čísla pro automatizaci rešerší

Jakub Harašta / Martin Loučka
Ústav práva a technologií PrF MU

EUR-Lex

- Publikační platforma pro legislativní a nelegislativní dokumenty EU
- Starý vs. nový EUR-Lex
- Související dokumenty, jazykové verze apod.

Celexové číslo

- Relikt z původního systému
- 31996L0009, 62012CJ0293
- Identifikuje dokument v systému bez ohledu na jeho jazykovou verzi
- 12012E101
 - Atypická struktura, vede až na úroveň jednotlivých článků
 - Čl. 101 SFEU

Náš cíl

- Vytvořit nástroj, který využije existujících vazeb mezi dokumenty (skrze celexové číslo) a bude vizualizovat průběh jednotlivých řízení navázaných na konkrétní směrnici, tedy:
 - 1) Je k dispozici dokument, který řízení zahájil?
62012CN0293
 - 2) Je k dispozici stanovisko GA?
62012CC0293
 - 3) Jak řízení skončilo?
62012CO0293 nebo 62012CJ0293

Postup rešerše I

- Načtení směrnice od uživatele (rok a číslo směrnice)
- Stažení generovaného XML a získání
 - Názvu směrnice
 - Používaných typů identifikátorů v rámci daného XML
 - Seznam připojených dokumentů souvisejících s judikaturou
- Odeslání základního výstupu na stranu uživatele

Postup řešení II

- Iterace všemi položkami a postupné dotazování na některé další informace
- Vše dynamicky vykreslováno do stránky
- Veškerý provoz přes náš server
 - Možnost cachování a průběžného zrychlení běhu

Dotazování XHR

- Na co se ptáme:
 - Existence dokumentů podle dalších souvislostí
 - Názvy rozhodnutí SDEU
 - Existence připojených případů
- Proč:
 - XML neobsahuje všechny dokumenty a občas „nemluví“ pravdu
 - XML neobsahuje názvy judikátů
 - XML neobsahuje připojené případy v podobě, kterou bychom uměli jednoduše zpracovat

Problémy

- XML na straně EUR-Lexu má nízké desítky tisíc řádků
 - Pro nás je použitelných cca 20
- XML neobsahuje základní metadata o připojených dokumentech
- Repozitář CELLAR je obsažen v XML jako zdroj kompletního obsahu, ale nelze ho využít pro zjišťování existence dokumentů
 - Několik neúspěšných HTTP požadavků vyústí v zablokování přístupu
- Informace je někdy nutné hledat z jiných zdrojů
 - Chybí metadata u některých dokumentů z důvodu neschopnosti EUR-Lexu zpracovávat vlastní standardy
- Téměř neexistuje dokumentace, existující části mohou být zavádějící

Další postup

- Určitě:
 - Snaha o optimalizaci dynamického vykreslování
 - Jazykové verze
 - Snaha o odstranění falešně negativních výsledků
- Optimálně:
 - Zmapování XML struktury EUR-Lexu a její provázání s procesními pravidly
 - Rozdělování judikatury podle druhu řízení
 - Implementace ECLI

Závěr

- Cíl: jednoduchý skript pro vykreslování průběhu řízení a zjišťování nerozhodnutých předběžných otázek
- Výsledek: poměrně robustní JavaScript/AJAX se sérií desítek XHR
- Zkušební provoz
- Odstraňování nedostatků

Poděkování

- Kolegům z PrF MU za testování první verze a pomoc při odstraňování nedostatků
- Kolegům z AION CS za objasnění některých „tajů“ EUR-Lexu
- Vám za pozornost

Celex.lawportal.cz

jakub.harasta@law.muni.cz

martin.loucka@mail.muni.cz