D4 – Vzorové údaje

Projekt financovaný Európskou úniou Next GenerationEU prostredníctvom Plánu obnovy a odolnosti SR v rámci projektu č. 09I05-03-V02-00049.

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
ÚRAD PODPREDSEDU VLÁDY SLOVENSKEJ REPUBLIKY PRE PLÁN OBNOVY A ZNALOSTNÚ EKONOMIKU
VAIA VÝSKUMNÁ A INOVAČNÁ AUTORITA
VÝSKUMNÁ AGENTÚRA

Úvod

Projekt „**Automatizácia analýzy právnych textov na základe strojového učenia**“ (ďalej len „**ALTAML**“) predstavuje kľúčovú iniciatívu zameranú na integráciu inovatívnych prístupov v oblasti spracovania údajov a následnej analýzy, so zameraním na oblasť práva, ktorá zahŕňa aj právo informačných a komunikačných technológií. Cieľom tohto projektu je vyvinúť a overiť účinné metódy automatizovanej analýzy právnych textov pomocou techník strojového učenia. To zahŕňa najmä vývoj nástrojov, ktoré uľahčujú spracovanie a analýzu veľkých objemov údajov vo forme rôznych právnych dokumentov, ako aj extrakciu relevantných informácií (atribútov) z takýchto dokumentov vrátane identifikácie kľúčových pojmov, odkazov na právne predpisy a ďalších atribútov.

Cieľom projektu ALTAML je teda prispieť k efektívnejšiemu prístupu k právnym informáciám a k zrýchleniu právnych procesov, čím sa zabezpečí vyššia miera právnej istoty a zlepší sa dostupnosť právnych textov, výsledkov ich analýzy a relevantných právnych informácií pre odbornú aj širokú verejnosť.

V rámci pracovného balíka KPB1 bol vytvorený výstup **D4 - Vzorové údaje** - vytvorenie súboru vzorových údajov. Vzhľadom na rozsiahly súbor zdrojov bude na demonštračné účely zverejnená len časť spracovaných dokumentov.

Obsah

Úvod **1**

1 Zdroje vzorových údajov **3**

2 Opis vzorových údajov **4**

1 Zdroje vzorových údajov

Datasety obsahujú rozhodnutia podľa zdroja získania, a to v 3 formátoch - pdf, xml a html.

Zdroje získavania rozhodnutí boli popísané v správe a sú to tieto:

1) justice.gov.sk - webová stránka Ministerstva spravodlivosti - rozhodovacia činnosťou všeobecných súdov, Najvyššieho súdu SR a Najvyššieho správneho súdu SR,
2) nsud.sk - webová stránka Najvyššieho súdu SR - rozhodovacia činnosťou súdu,
3) nssud.sk - webová stránka Najvyššieho správneho súdu SR - rozhodovacia činnosťou súdu,
4) ustavnysud.sk - webová stránka Ústavného súdu SR - rozhodovacia činnosťou súdu,
5) nsud.sk - webová stránka Najvyššieho súdu SR - Zbierka rozhodnutí a stanovísk NS SR,
6) nssud.sk - webová stránka Najvyššieho správneho súdu SR - Zbierka stanovísk a rozhodnutí NSS SR,
7) ustavnysud.sk - webová stránka Ústavného súdu SR - Zbierka nálezov a rozhodnutí ÚS SR.

2 Opis vzorových údajov

Legenda:

source	zdroj získania rozhodnutia
url	adresa konkrétneho rozhodnutia
id	identifikácia v našej databáze
sp_zn	spisová značka rozhodnutia
date	dátum rozhodnutia
typ	typ rozhodnutia
titul	názov rozhodnutia (buď získaný zo zdroja alebo vytvorený wk)
summary	skrátený obsah rozhodnutia vytvorený wk
autor	súd, ktorý rozhodnutie vydal
pdf	názov získaného pôvodného rozhodnutia vo formáte pdf
xml	názov xml súboru
html	názov html súboru
zip	názov zipu, v ktorom sú rozhodnutia

1.	**Ministerstvo spravodlivosti SR**
source	justice.gov.sk
URL	https://www.justice.gov.sk/sudy-a-rozhodnutia/sudy/rozhodnutia
ID	JUD4287149SK
sp_zn	37Ca/1/2016
date	9.12.2022
typ	Rozsudok
titul	Občianske právo - Zmluva o dielo; Občianske
summary	Ochrana autorského práva a zakázanie
autor	Okresný súd Košice I
pdf	Rozsudok_37Ca-1-2016.pdf
xml	ju_4287149_sk-JUD4287149SK.xml
html	ju_4287149_sk-JUD4287149SK.html
zip	01.zip

2.	**Najvyšší súd SR**
source	https://www.nsud.sk/rozhodnutia/
URL	https://www.nsud.sk/rozhodnutia/4cdo2402020/
ID	JUD4347552SK
sp_zn	4Cdo/240/2020
date	28.3.2023
typ	Uznesenie

titul	určenie autorského práva s nárokom na náhradu nemajetk. ujmy
summary	Odmietnutie dovolania v spore o určenie autorského práva a náhradu nemajetkovej ujmy
autor	Najvyšší súd SR - senát
pdf	751519.9c7c1a.pdf
xml	ju_4347552_sk-JUD4347552SK.xml
html	ju_4347552_sk-JUD4347552SK.html
zip	02.zip

3.	**Najvyšší správny súd SR**
source	https://www.nssud.sk/
URL	https://www.nssud.sk/wp-content/uploads/7Svk492022.pdf
ID	JUD4566407SK
sp_zn	7Svk/49/2022
date	27.3.2024
typ	Rozsudok
titul	Udelenie oprávnenia na výkon kolektívnej správy práv zvukárov k spracovaným zvukovým záznamom
summary	Zrušenie rozhodnutia o kolektívnej správe práv zvukových majstrov
autor	Najvyšší správny súd SR - senát
pdf	Rozsudok_37Ca-1-2016.pdf
xml	ju_4287149_sk-JUD4287149SK.xml
html	ju_4287149_sk-JUD4287149SK.html
zip	03.zip

4.	**Ústavný súd SR**
source	https://www.ustavnysud.sk/rozhodovacia_cinnost/rozhodnutia
URL	https://www.ustavnysud.sk/docDownload/4444fd9a-3648-4660-92dd-73b1150a1bbc
ID	JUD2616598SK
sp_zn	III. ÚS 651/2016 - 90
date	28.11.2017
typ	Nález
titul	Porušenie základných práv pri poskytovaní monitoringu médií bez súhlasu, čo zahŕňalo neoprávnené vyhotovovanie rozmnožení článkov
summary	Porušenie základných práv MAFRA Slovakia v súvislosti s autorskými právami
autor	Ústavný súd SR - senát
pdf	III. ÚS 651_2016.pdf
xml	ju_2616598_sk-JUD2616598SK.xml
html	ju_2616598_sk-JUD2616598SK.html

zip	04.zip

5.	**Zbierka rozhodnutí a stanovísk Najvyššieho súdu SR**
source	https://www.nsud.sk/zbierka-stanovisk-a-rozhodnuti/
URL	https://www.nsud.sk/data/att/676/637297.f1bf1b.pdf
ID	JUD4242507SK
sp_zn	2 CdoGp 1/2020
date	24.2.2022
typ	Uznesenie
titul	Absolútne majetkové autorské práva. Odúmrť. Voľné dielo.
summary	K otázke prechodu dedičstva v prípade absolútnych majetkových autorských práv, či pripadne štátu alebo sa stane voľným
autor	Najvyšší súd SR - senát
pdf	637297.f1bf1b.pdf
xml	ju_4242507_sk-JUD4242507SK.xml
html	ju_4242507_sk-JUD4242507SK.html
zip	05.zip

6.	**Zbierka stanovísk a rozhodnutí Najvyššieho správneho súdu SR**
source	https://www.nssud.sk/zbierka-stanovisk-a-rozhodnuti-najvyssieho-spravneho-sudu-slovenskej-republiky/
URL	https://www.nssud.sk/wp-content/uploads/Zbierka_2_2023.pdf
ID	JUD4390154SK
sp_zn	33 D 17/2021
date	17.11.2022
typ	Rozhodnutie
titul	Disciplinárne previnenie; viazanosť právnym názorom Ústavného súdu Slovenskej republiky
summary	porušenie povinnosti sudcu pri nerešpektovaní rozhodnutia ústavného súdu
autor	Najvyšší správny súd - senát (ostatné)
pdf	Zbierka_3_2023.pdf
xml	ju_4390154_sk-JUD4390154SK.xml
html	ju_4390154_sk-JUD4390154SK.html
zip	06.zip

7.	**Zbierka nálezov a rozhodnutí Ústavného súdu SR**

source	https://www.ustavnysud.sk/rozhodovacia_cinnost/zbierka-n%C3%A1lezov-a-uznesen%C3%AD
ID	JUD3178033SK
sp_zn	PL. ÚS 9/2016
date	7.2.2018
typ	Nález
titul	Hudobné kvóty
summary	Spor o ústavnosť imitácie slovenských kultúrnych prvkov v rozhlasovom vysielaní, definícia slovenského hudobného diela prostredníctvom národných prvkov
autor	Ústavný súd SR - plénum
pdf	Dokument-28.pdf
xml	ju_3178033_sk-JUD3178033SK.xml
html	ju_3178033_sk-JUD3178033SK.html
zip	07.zip

Súčasťou tohto dokumentu sú aj 2 prílohy:

1. Data sets – description I (Časť I: domény „Autorské právo" a „Ochranné známky / Nekalá súťaž")

- Dokument popisuje 17 dátových balíkov vytvorených pre domény autorské právo a ochranné známky/nekalá súťaž, ktoré obsahujú dokumenty vo formáte html.
- Dokumenty boli do balíkov vybrané na základe fulltextových dotazov alebo dotazov do metadát v databáze, pričom samotné dotazy boli formulované podľa zadania riešiteľov projektu.
- Materiál je členený do štyroch častí: zadania partnerov s odkazmi na čísla balíkov (napr. ustanovenia Autorského zákona č. 185/2015 Z. z., § 442a a § 458a Občianskeho zákonníka, smernice 2001/29/ES a 2006/115/ES, judikatúra SD EÚ vo veciach C-162/10 a C-306/05, zákon č. 506/2009 Z. z. o ochranných známkach a § 44 až § 48 Obchodného zákonníka), podrobný prehľad jednotlivých balíkov so znením meta-dotazov a full-textových dotazov, popis výsledkových sád a zoznam príloh. Ku každému balíku prináleží prehľad vo formáte .xlsx (datovy-balik-1.xlsx až datovy-balik-17.xlsx) s hárkom „altaml" obsahujúcim všetky dokumenty a vo väčšine prípadov aj hárkom „po_redukcii" so zúženým výberom podľa analýzy spisovej značky (pre autorské právo registre „Ca" a „Co", pre ochranné známky „CbPV" a „CoPV"). Dokument zároveň vysvetľuje význam jednotlivých stĺpcov výsledkových sád (dotaz, spisová značka, ID v databáze, dátum rozhodnutia, prameň, právna oblasť, register, ECLI a ďalšie).

2. Data sets – description II (Časť II: doména „Digitálny dôkaz v trestnom konaní")

- Dokument popisuje 14 dátových balíkov pripravených pre doménu digitálny dôkaz v trestnom konaní, opäť s dokumentmi v html formáte vybranými prostredníctvom fulltextových a metadátových dotazov podľa zadania riešiteľov.
- Zadania vychádzajú najmä z ustanovení Trestného poriadku (§ 3, § 89a, § 90, § 91, § 92, § 96d, § 100 – § 118), zákona č. 236/2017 Z. z. o európskom vyšetrovacom príkaze, Dohovoru o počítačovej kriminalite, smernice 2013/40/EÚ o útokoch na informačné systémy a nariadenia (EÚ) 2023/1543 o európskych príkazoch na predloženie a uchovanie elektronických dôkazov.
- Popri samostatných vyhľadávacích kritériách boli použité aj kumulatívne kritériá, kde sa jednotlivé ustanovenia kombinovali s kľúčovými slovami z oblasti digitálnych dát (obsah komunikácie, prevádzkový a lokalizačný údaj, dátový nosič, mobilný telefón, server, cloud, e-mail, GPS údaje, šifrovaná komunikácia, kryptomena a ďalšie). Rovnako ako v prvej časti sú aj tu ku každému balíku priradené výsledkové sady v prehľadoch (datovy-balik-18.xlsx až datovy-balik-31.xlsx) s plnou aj zredukovanou verziou; redukcia sa vykonala analýzou spisovej značky so zameraním na trestnú agendu (registre „T", „To", „Tdo", „Tk", „Nt" a „ÚS"). Preškrtnuté položky v prehľade signalizujú dotazy s nulovým alebo nepodstatným výsledkom vyhľadávania.