
Využití celexového čísla pro rešerši právních informací

Use of celex numbers for legal information retrieval

Jakub Harašta

Ústav práva a technologií, Právnická fakulta Masarykovy univerzity

Martin Loučka

Legal Technologies, s.r.o.

Abstrakt

Od 1. července 2013 je Úřední věstník EU vydávaný v elektronické podobě v rámci platformy EUR-Lex. Tato platforma je tak vyčerpávajícím zdrojem informativních i autentických informací z práva EU. Navzdory tomuto pokroku, na který v ČR stále ještě čekáme, čelí EUR-Lex občasným výtkám ve vztahu k nedostatečné uživatelské přívětivosti a přehlednosti. Izolovaná směrnice či jiný dokument totiž nepředstavuje právní informaci, ale právní datum. Vzhledem ke specifickému postavení rozhodovací praxe Soudního dvora EU jsme schopni dosáhnout komunikace právní informace až ve chvíli provázání rozhodovací praxe a textu směrnice. Přístup k souvisejícím dokumentům by tedy měl být maximálně jednoduchý a měl by umožňovat jednoduchou orientaci. EUR-Lex toto v současné době neumožňuje, kdy judikaturu navázanou na nařízení či směrnici zobrazuje pouze pomocí tzv. celexového čísla. V tomto příspěvku se věnujeme struktuře celexového čísla a možnému využití existující XML struktury EUR-Lexu pro automatizaci některých vyhledávacích funkcí. Náš příspěvek vychází z následujících premis:

- relevantní právní informace obsahují i stanoviska generálních advokátů, která není možné v současné podobě jednoduše vyhledat (v seznamu souvisejících dokumentů nejsou uváděny)*
- relevantní právní informací je existence nezodpovězené předběžné otázky napadené k SDEU, tuto informaci ale v současné podobě zobrazování souvisejících dokumentů není možné pohodlně zjistit*
- na základě existující metadatové struktury EUR-Lexu (zejména celexových čísel jako unikátních identifikátorů jednotlivých dokumentů) je možné automatizovat některé vyhledávací úkony a napravit shora uvedené nedostatky současného vyhledávání.*

Klíčová slova: *právní informace, celexové číslo, EUR-Lex*

Abstract

From 1 July 2013 the EU Official Journal is published in electronic form through the EUR-Lex platform. Therefore, this platform is comprehensive

source of authentic and non-authentic information regarding EU law. Despite this progress, that is however still not present in the Czech Republic, the EUR-Lex is facing occasional criticism because of the lack of user-friendliness and clarity. Isolated directive or other document can not be considered a legal information, but merely the legal datum.

Given the specificity of the decisions of the Court of Justice of the EU, we are able to achieve communication of the resulting legal information only after linking the decision with the text of the respective directive. Therefore, access to these related documents should be extremely simple and should allow for easy orientation of the user. Currently, EUR-Lex does not provide this comfort as the case law related to its directive or regulation is displayed only by so-called celex number. In this paper we discuss the structure of celex number and the possibility to use existing XML structure maintained by EUR-Lex for automatization of some search functions. Our paper is based on the following premises: Relevant legal information includes the opinion of the advocate general, which is currently impossible to locate easily (these opinions are not present in the list of related documents).

Relevant legal information is the information about existence of unresolved preliminary questions brought before the CJEU; however, this information can not be easily ascertained given the current form of displaying of the related documents.

Based on existing metadata structure in which EUR-Lex operates (especially celex numbers as unique identifiers of individual documents), it is possible to automate some research operations and to fix the above mentioned shortcomings of the current search.

Keywords: *legal information, legal information retrieval, EUR-Lex, celex number*

Úvod a motivace

Od 1. července 2013 je Úřední věstník EU vydáván v elektronické podobě v rámci platformy EUR-Lex.¹ Ten jako platforma pro zpřístupňování práva EU existoval i dříve, ale právě od zmíněného data se stal i zdrojem autentických informací o právu EU. Tento vývoj, na který v České republice ještě čekáme,² představuje krok správným směrem, kterým je právě snadný a bezplatný přístup k autentickým právním informacím. Navzdory tomuto posunu se ale EUR-Lex stále nestal dominantním způsobem přístupu k právu EU. Domníváme se, že je to zčásti kvůli na první pohled nepřehlednému systému označování dokumentů celexovým číslem. Tento unikátní identifikátor však může sám o sobě mimo jiné umožňovat snadný přechod mezi jednotlivými fázemi soudního řízení před Soudním dvorem Evropské unie (dále jen „SDEU“) a v kombinaci s robustní metadatovou strukturou EUR-Lexu může usnadňovat vyhledávání relevantních informací.

K plnému využití metadatové struktury je potřeba značná míra předvedění o úloze identifikátorů a celkových možnostech systému. Namísto výkladu o jeho možnostech však v tomto textu představujeme demonstraci možností EUR-Lexu vytvořením jednoduchého systému pro navracení právních informací. Konkrétně se jedná o extrakci všech dostupných dokumentů z každého jednotlivého řízení před SDEU, které je vztažené k uživatelem definované směrnicí.

V jednotlivých částech textu postupně představíme (1) specifické postavení judikatury SDEU v právu, (2a) strukturu celexových čísel a (2b) metadatovou strukturu EUR-Lexu. Dále popíšeme (3) základní vztahy a postupy, které jsme využili v rámci demonstrace možností. Text je pak uzavřen (4) závěrem, ve kterém nastíníme problémy, se kterými jsme se při využití metadatových záznamů v rámci demonstračního cvičení setkali.

1 Postavení a specifika judikatury SDEU

SDEU má vzhledem ke svému postavení specifickou funkci. Nejčastějším druhem řízení, ze kterého národní praxe čerpá právní závěry, je řízení o předběžné otázce.³ V tomto zvláštním druhu řízení soud nerozhoduje ve věci samé, neřeší samotný spor mezi stranami. SDEU v tomto druhu řízení reaguje na otázku položenou národním soudem k tomu, jakým způsobem je nutné vykládat v konkrétním řízení unijní právo. SDEU zde funguje jako těleso zajišťující jednotný výklad práva ve všech členských státech. Obecně se tedy tato rozhodnutí pohybují na vyšší míře abstrakce, než je běžné, a výlučně odpovídají na otázku o výkladu položenou národním soudem. Soud může odpovědět dvěma způsoby – může dospět k rozhodnutí, kterým na položenou otázku explicitně odpoví, nebo může vydat usnesení, kterým otázku prohlásí za již vyřešenou a připojí k tomu úvahu na téma aplikovatelnosti již existující rozhodovací praxe na danou věc.

V současné době chybí literatura, která by se extenzivně a jasně vypořádávala se závazností judikatury SDEU.⁴ Obecně ale platí, že závěry rozhodovací praxe SDEU je nutné zohledňovat, protože v odpovědi na předběžnou otázku soud interpretuje ustanovení konkrétního právního předpisu. Pro pochopení bude zřejmě nejjednodušší, pokud problematiku

¹ Na základě Nařízení Rady (EU) č. 216/2013 ze dne 7. března 2013, o elektronickém vydávání *Úředního věstníku Evropské unie*

² Elektronické zpřístupňování práva je plánováno v podobě projektu elektronické Sbirky zákonů a mezinárodních smluv. K tématu STUPKA, Václav. E-Sbirka a e-Legislativa. *Revue pro právo a technologie*, 2014, roč. 5, č. 9, s. 185-196. Předpisy směřující k vytvoření elektronické platformy autentických právních informací i v České republice prošly dne 15. 6. 2016 Senátem, s účinností se pak počítá od 1. 1. 2020.

³ Uvádí BOBEK, Michal, KÜHN, Zdeněk a kol. *Judikatura a právní argumentace. 2. přepracované a aktualizované vydání*. Praha: Auditorium, 2013. S. 287.

⁴ Konstatuje Bobek 2013 op. cit., s. 304.

přiblížíme na případu naší demonstrace. V rámci dalšího textu bude popsán systém, který navrácí rozhodovací praxi SDEU, která je navázána na konkrétní směrnici. Směrnice představuje právní nástroj směřující k harmonizaci národních úprav členských států. Členské státy musí na základě evropské směrnice upravit národní legislativu tak, aby odpovídala cílům ve směrnici stanoveným. Národní legislativu je pak nutné vykládat a aplikovat orgány členského státu v souladu se směrnicí. Směrnice jako taková ale nestanovuje konkrétní prostředky a nástroje, kterými má být cílů dosaženo. Konkrétní prostředky tak jsou na jednotlivých členských státech – směrnice poskytuje pouze poměrně abstraktní rámec pro jejich implementaci. V rámci tohoto abstraktního rámce pak může docházet k pochybnostem a diskuzím o tom, jaká konkrétní národní úprava je ještě v souladu, případně jakým způsobem ji aplikovat tak, aby v souladu opravdu byla. V případě pochybností může národní soud položit předběžnou otázku a SDEU na tuto otázku odpoví. Výklad, který SDEU v rámci odpovědi provede, pak představuje de facto obohacení textu příslušné směrnice. V situaci vymezené předběžnou otázkou SDEU odpovídá a tato odpověď se stává výkladovým vodítkem pro další aplikaci příslušného ustanovení.

Specifickým prvkem judikatury SDEU jsou i stanoviska generálního advokáta. Pozice generálního advokáta pochází z frankofonního prostředí.⁵ V našem prostředí k nim tedy chybí bezprostřední ekvivalent a snad v důsledku toho jsou stanoviska často opomíjena. Stanoviska generálního advokáta jsou přitom zdrojem nestranných a nezávislých odůvodněných stanovisek,⁶ která soud často následuje.⁷ Stanovisko generálního advokáta tak představuje expertní rozbor dané problematiky a může tak suplovat odlišné stanovisko,⁸ které známé např. z českého Ústavního soudu. Může představovat alternativní či širší úvahy, kterým se SDEU ve svých rozhodnutích spíše vyhýbá.

Z výše uvedeného jsme pro účely naší demonstrace dovedli, že existence samotného rozhodnutí představuje extrémně důležitý právní názor, ale nestačí pro kompletní pochopení problematiky. Odpověď soudu je v řízení přímo a zcela zásadně ovlivněna položenou otázkou. Stejně tak je, pro kompletní pochopení kontextu daného případu, nezbytné věnovat pozornost i stanovisku generálního advokáta, pokud bylo ve věci vydáno.⁹ Z toho důvodu se tedy v rámci demonstrace nezaměřujeme pouze na vztah směrnice a rozhodnutí SDEU, ale na vztah směrnice a celého případu před SDEU – tedy od znění položené předběžné otázky, přes stanovisko generálního advokáta (pokud je přítomné), až po ukončení řízení v podobě rozhodnutí či usnesení SDEU.

2a Celexová čísla

Jedním z problematičtějších aspektů přístupu k právu EU je identifikace konkrétních dokumentů. Tento problém je možné popsat ve dvou rovinách, kdy obě jsou organizací EU vlastní. První rovinou je mnohost typů dokumentů, které EU vydává a které je nutné publikovat. Druhou je pak mnohojazyčnost EU. Pro demonstraci problémů s označováním typů dokumentů je možné popsat dva základní dokumenty určující právní rámec ochrany osobních údajů – *směrnice Evropského parlamentu a Rady 95/46/ES ze dne 24. října o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů* a *nařízení Evropského parlamentu a Rady (EU) 2016/679 ze dne 27. dubna 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů*. Oba dokumenty jsou jiného druhu (směrnice, resp. nařízení), obsahují

⁵ K tomu BOBEK, Michal. A Fourth in the Court: Why Are There Advocates General in the Court of Justice. *Cambridge Yearbook of European Legal Studies*, 2012, roč. 14, s. 529-561. S. 529-531.

⁶ Tuto povinnost má přímo stanovenou ve čl. 252 Smlouvy o fungování EU.

⁷ Bobek uvádí cca 70-80%. Viz Bobek 2013 op. cit., s. 300 a literaturu tam citovanou.

⁸ Bobek 2013 op. cit., s. 301.

⁹ Od roku 2001 je možné věc rozhodnout bez stanoviska generálního advokáta, které bylo do té doby pravidlem a nutností pro každé řízení.

číslování v očividně jiných číselných řadách¹⁰ a svoji legitimitu odvozují z jiného institucionálního rámce (Evropská společenství, resp. Evropská unie). Pro ilustraci problémů s mnohojazyčností je nutné si uvědomit, že shora uvedené názvy směrnice 95/46/ES a nařízení 2016/679 mají svoji alternativu ve všech úředních jazycích EU. To celkovou orientaci dále znesnadňuje.

Podobná situace je i v rámci identifikace judikatury, kdy tradice některých států velí identifikovat pořadovým číslem věci, jako je běžnou praxí např. v České republice, zatímco jiné státy preferují označení stranami, např. Anglie. Tím nemusí být zcela jasné, jaký dokument je zamýšlen a pro jeho přesnou identifikaci je tak nutné jej opatřit řadou různých identifikátorů, což znesnadňuje zpracování textu – spory stejných stran se totiž mohou opakovat nebo se tytéž strany mohou přit v jiných věcech, takže je zapotřebí uvádět např. i datum rozhodnutí. Rozhodnutí je také možné identifikovat i odkazem na tištěnou sbírku rozhodnutí SDEU – ta navíc před 1. 5. 2004 nevycházela, z celkem pochopitelných důvodů, česky.¹¹

Aby tyto dvě roviny problému nekomplikovaly orientaci v systému, probíhala od počátku evidence dokumentů pomocí celexových čísel. Celexové číslo představuje identifikaci každého dokumentu v jednotném¹² formátu. Tento identifikátor navíc nezohledňuje jazykovou verzi dokumentu – lze tak říci, že identifikace probíhá na ontologické úrovni díla standardního referenčního rámce IFLA.¹³

Celexové číslo každého dokumentu má tvar SRRRRD(D)ČČČ(Č), kde

- „S“ představuje znak identifikující sektor, do kterého dokument náleží. Může se jednat o 1 (zakládací smlouvy), 2 (mezinárodní smlouvy), 3 (sekundární právo), 4 (doplňující právní předpisy), 5 (návrhy), 6 (judikatura), 7 (vnitrostátní prováděcí opatření), 8 (vnitrostátní judikatura), 9 (parlamentní interpelace), o (konsolidované dokumenty), C (další dokumenty zveřejňované v řadě C Úředního věstníku) a E (dokumenty Evropského sdružení volného obchodu).
- „RRRR“ představuje čtyři znaky identifikující rok vydání dokumentu
- „D(D)“ představuje jeden či dva znaky identifikující druh předpisu v rámci sektoru¹⁴
- „ČČČ(Č)“ představuje tři nebo čtyři znaky identifikující číslo dokumentu.

Výše deklarovaná jednotná forma identifikátoru má výjimku v rámci sektoru 1 (viz příklad níže), kterým jsou zakládací smlouvy. Zatímco v ostatních sektorech je entitou identifikovanou celexovým číslem dokument, v rámci sektoru zakládacích smluv je identifikovanou entitou článek příslušné smlouvy. Druh předpisu totiž v rámci tohoto sektoru nepředstavuje obecnou třídu (např. směrnice L či nařízení R v rámci sektoru sekundárního práva 3), ale konkrétní mezinárodní smlouvu. Pro účely naší demonstrace je tato pochopitelná¹⁵ nekonzistentnost naštěstí zcela irelevantní.

¹⁰ Od 1. 1. 2015 došlo k harmonizaci číslování právních aktů EU, což zpřehledňuje systém *pro futuro*, ale neřeší velké množství různých záznamů z minulosti. Viz Harmonizace číslování právních aktů EU, dostupné z:

<http://eur-lex.europa.eu/content/tools/elaw/OA0614022CSD.pdf>.

¹¹ Obširněji Bobek 2013 op. cit., s. 296-299.

¹² *Sic!* K tomu dále v textu.

¹³ Srov. standardní referenční rámec IFLA - IFLA Study Group on the Functional Requirements for Bibliographic Records. Functional Requirements for Bibliographic Records [online]. IFLA Section on Cataloguing: Mnichov [cit. 26. 3. 2015]. Dostupné z: <http://www.ifla.org/files/assets/cataloguing/frbr/frbr.pdf>. S. 12.

¹⁴ Veškeré druhy dokumentů v rámci jednotlivých sektorů jsou dostupné na http://eur-lex.europa.eu/content/tools/TableOfSectors/types_of_documents_in_eurlex.html?locale=cs

¹⁵ Smluv je výrazně menší množství, než jakýchkoli dalších dokumentů, z důvodu odlišného procesu, který vede k jejich vzniku.

Příklad:

Celexové číslo 31996L0009 identifikuje směrnici 96/9/ES o právních ochraně databází – 3 je sektorem sekundárního práva, L značí směrnici, 0009 značí 9 a 1996 značí rok vydání.

Celexové číslo 12012E101 identifikuje článek 101 smlouvy o fungování EU – 1 je sektorem primárního práva, tedy základacích smluv, E značí smlouvu o fungování EU, protože zde nemáme druh předpisu jako obecnou třídu. Smlouva byla vydána v roce 2012 a konkrétně se jedná o její článek 101.

V demonstraci, která je předmětem dalšího výkladu nás zajímá pouze provázanost směrnic (celexová čísla ve tvaru **3RRRRLČČČČ**) a jednotlivých fází řízení před SDEU, konkrétně předběžné otázky (**6RRRRCNČČČČ**), stanoviska generálního advokáta (**6RRRRCČČČČČ**) a rozhodnutí SDEU (**6RRRRCJČČČČ**) nebo usnesení SDEU (**6RRRRCOČČČČ**).

2b Metadata

Bez právních informačních systémů si vyhledávání právních informací v podstatě nelze představit. Vyhledávání pomocí booleovského modelu pro vyhledávání je již nějakou dobu standardem. Jeho využívání má však jisté nedostatky, které výborně ilustruje starší experiment Blaira a Marona.¹⁶ Praktikující právníci v něm byli vyzváni k rešerši informací v rámci systému – pokus měl být ukončen samotným právníkem ve chvíli, kdy nabyl dojmu, že se mu podařilo ze systému navrátit alespoň 75% dokumentů.¹⁷ Napříč pokusy 26 právníků se podařilo vyhledat průměrně pouze 20%.¹⁸ Každý jednotlivý právník tedy měl ukončit svoje snažení ve chvíli, kdy měl pocit, že se mu podařilo získat převážnou většinu relevantních dokumentů z dané databáze. Jakkoli je tento experiment staršího data, dle našeho názoru ilustruje do značné míry i dnešní situaci – vyhledávání pomocí booleovského modelu se totiž soustředí na slova, zatímco právníky zajímají spíše koncepty.¹⁹ Jedním z postupů, který byl v minulosti identifikován jako možná cesta vpřed, tedy ke zvýšení úspěšnosti vyhledávání právních informací, je využití citační analýzy.²⁰

Na narůstající počet dokumentů v databázích, který pozorujeme od opuštění papíru jako média pro komunikaci soudních rozhodnutí,²¹ samozřejmě tvůrci právních informačních systémů reagují, ale v České republice pouze nedostatečně.²² Citační analýza za jako taková má původ v pracích *Citation Indexes for Science*²³ a *Networks of Scientific Papers*²⁴. V poslední dekádě pak evidujeme zvýšený výskyt citační analýzy pro odhalování vlastností sítí právních dokumentů – převážně soudních rozhodnutí. Setkáváme se s pracemi zaměřenými na citační analýzu rozhodnutí Federálního odvolacího soudu USA²⁵ a Nejvyššího soudu

¹⁶ David C. Blair a M. E. Maron. An evaluation of retrieval effectiveness for a full-text document-retrieval system. *Communications of the ACM*, 1985, roč. 28, č. 3, s. 289-299.

¹⁷ Tamtéž s. 291

¹⁸ Výsledky podrobně na s. 293-295.

¹⁹ Viz MOENS, Maria-Francine. Innovative techniques for legal text retrieval. *Artificial Intelligence and Law*, 2001, roč. 9, č. 1, s. 29-57, s. 29. Tento problém se dá překonat právě využitím alternativních metod vyhledávání – viz výklad SCHWEIGHOFER, Erich. The Revolution in Legal Information Retrieval or The Empire Strikes Back. *Journal of Information, Law and Technology*, 1999, roč. 4, č. 1, nestránkováno. Část 6.5 Connectionist information Retrieval.

²⁰ Obecně SMITH, Linda C.. Citation Analysis. *Library Trends*, 1981, roč. 30, č. 1, s. 83-106. S. 96-97. V právní doméně potom GEIST, Anton. The Open Revolution: Using Citation Analysis to Improve Legal Text Retrieval. *European Journal of Legal Studies*, 2010, roč. 2, č. 3, s. 137-145.

²¹ Srov. WIDDISON, Robin. New Perspectives in Legal Information Retrieval. *International Journal of Law and Information Technology*, 2002, roč. 10, č. 1, s. 41-70.

²² K problematické extrakci známých setů rozhodnutí z existujících právních informačních systémů viz MYŠKA, Matěj, HARAŠTA, Jakub, LOUTOCKÝ, Pavel a Jakub MÍŠEK. *Revue pro právo a technologie*, 2016, roč. 7, č. 13, s. 147-189. S. 159-169.

²³ GARFIELD, Eugene. Citation Indexes for Science. *Science*, 1955, roč. 122, č. 3159, s. 108-111.

²⁴ PRICE, Derek J. de Solla. Network of Scientific Papers. *Science*, 1965, roč. 149, č. 3683, s. 510-515.

²⁵ LANDES, William M. LESSIG, Lawrence a Michael E. SOLIMINE. Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges. *Journal of Legal Studies*, 1998, roč. 27, č. 2, s. 271-332.

USA.²⁶ Zájem se pak postupně přesouvá i do citační analýzy rozhodovací praxe Evropského soudu pro lidská práva²⁷ a Soudního dvora Evropské unie.²⁸ Citační analýza, jakkoli je často označovaná za relativně jednoduchou techniku, se ukázala pro právní prostředí poměrně náročnou – to z důvodu chybějícího citačního standardu, jehož nepřítomnost je zejména v ČR obzvláště markantní,²⁹ nebo celkové roztržičnosti citace, kdy jsou relevantní informace přítomny implicitně.³⁰ Samotná extrakce citací se tak může proměnit v náročný úkol a proto je obzvláště vhodné přidávat k judikatuře metadata takovým způsobem, jakým činí EUR-Lex – tím jsou vlastně seznamy citovaných dokumentů připraveny k použití.

V rámci naší demonstrace se nezabýváme vlastnostmi citačních sítí soudních rozhodnutí,³¹ ale přistupujeme k nim podobně jako Geist.³² Naším cílem tedy je zefektivnit vyhledávání právních informací a zajistit vysokou míru spolehlivosti takto vrácených dokumentů. Metadatový záznam v podobě identifikace citovaného dokumentu je pro efektivnější vyhledávání zcela zásadní – citace mezi dokumenty totiž implikují jistou souvislost. Může se jednat o totéž téma, které může být definováno jinak co do rozsahu, může se jednat o dopad na interpretaci téhož předpisu nebo využití stejných argumentů.

Je také nutné poznamenat, že se v rámci naší demonstrace v této chvíli zabýváme pouze specifickým druhem reference uvedené v EUR-Lexu, kterou je reference, kdy byla směrnice nebo její výklad ovlivněna případem řešeným před SDEU.³³ V případě, že se judikatura vztahuje ke konkrétní směrnici, objevuje se v metadatovém záznamu příslušné směrnice údaj o tom, že byla některá její část interpretována SDEU (viz příklad níže). S pomocí extrakce těchto údajů z metadatového záznamu příslušné směrnice je možné identifikovat judikaturu SDEU, která se k dané směrnici vztahuje.

Příklad:

Část metadatového záznamu shora uvedené směrnice 96/9/ES identifikované celexovým číslem 31996L0009 obsahuje údaje

A07P1 Interpreted by 62002CJ0046

A07P5 Interpreted by 62002CJ0203

A07P1 Interpreted by 62002CJ0203

A07 Interpreted by 62002CJ0203

A07P1 Interpreted by 62002CJ0338

A01P2 Interpreted by 62002CJ0444

A07P1 Interpreted by 62002CJ0444

²⁶ FOWLER, James H. a Sangick JEON. The authority of Supreme Court precedent. *Social Networks*, 2008, roč. 30, č. 1, s. 16-30. Také FOWLER, James H. et al. Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court. *Political Analysis*, 2007, roč. 15, č. 3, s. 324-346. Také LUPU, Yonatan a James H. FOWLER. Strategic Citations to Precedent on the U.S. Supreme Court. *Journal of Legal Studies*, 2013, roč. 42, č. 1, s. 151-186.

²⁷ LUPU, Yonatan a Erik VOETEN. Precedent in International Courts: A Network Analysis of Case Citations by the European Court of Human Rights. *British Journal of Political Science*, 2012, roč. 42, č. 2, s. 413-439.

²⁸ DERLÉN, Mattias a Johan LINDHOLN. Goodbye van Gend en Loos, Hello Bosman? Using Network Analysis to Measure the Importance of Individual CJEU Judgments. *European Law Journal*, 2014, roč. 20, č. 5, s. 667-687. Také SADL, Urška a Yannis PANAGIS. What is Leading Case in EU Law? An Empirical Analysis. *European Law Review*, 2015, roč. 40, č. 1, s. 15-34.

²⁹ Viz BAŇOUCH, Hynek a Zdeněk KÜHN. O publikaci a citaci judikatury aneb proč je někdy judikatura jako císařovy nové šaty. *Právní rozhledy*. 2005, roč. 13, č. 13, s. 484-491. Také Myška 2016 op. cit.

³⁰ Viz KRÍŽ, Vincent, HLADKÁ, Barbora, DĚDEK, Jan a Martin NEČASKÝ. Statistical Recognition of References in Czech Court Decisions. In: GELBUKH, Lexander, ESPINOZA, Félix Castro a Sofia N. GALICIA-HARO (eds.). *Human-Inspired Computing and its Applications. MICAI 2014*. Springer: Berlin, Heidelberg, 2014, s. 51-61. Problémy jsou shrnuté na s. 53-54.

³¹ Výzkum tímto směrem vedli např. Fowler 2008 op. cit. nebo KONIARIS, Marios, ANAGNOSTOPOULOS, Ioannis a Yannis VASSILIOU. Legislation as a complex network: Modelling and analysis of European Union legal sources. In: HOEKSTRA, Rinke (ed.). *JURIX 2014: The Twenty-Seventh Annual Conference*. Amsterdam: IOS Press, 2014, s. 143-152.

³² Geist 2010 op. cit., s. 142-143.

³³ Srov. Koniaris 2014 op. cit., s. 147.

Tento zápis značí, že rozhodnutí SDEU ve věci C-203/02 interpretovalo čl. 7, čl. 7 odst. 1 a čl. 7 odst. 5 v rámci dané směrnice, rozhodnutí SDEU ve věci C-338/02 a C-46/02 interpretovaly čl. 7 odst. 1 a rozhodnutí C-444/02 interpretovalo čl. 1 odst. 2 a čl. 7 odst. 1 směrnice 96/9/ES.

3 Demontrace

Následující část textu bude věnována rozboru fungování aktuální implementace našich poznatků o fungování EUR-Lexu do praxe. Vycházíme z toho, že metadatová struktura EUR-Lexu, stejně jako používání celexových čísel pro identifikaci dokumentu, je pro běžného uživatele problematicky uchopitelná. Cílem demonstrace je tedy umožnit vrácení jednotlivých rozhodnutí SDEU v podobě, která umožní snadnější orientaci. Představení výsledků tak probíhá v kontextu daném směrnicí a i jinými dokumenty v rámci téhož řízení.

Z hlediska implementace lze aplikaci rozdělit na dvě základní části, které ve vzájemném propojení získávají a prezentují relevantní výsledky z EUR-Lexu uživateli. Před nimi samozřejmě stojí jednoduchý iniciační formulář, který má za úkol přijmout data od uživatele. Uživatel identifikuje směrnici, ke které chce provést rešerši, pomocí roku jejího vydání a jejího pořadového čísla.

Zmíněné dvě základní části nejsou z časového hlediska v průběhu zpracování požadavku oddělené – jejich oddělení spočívá spíše v technické a logické stránce. Zatímco jedna část aplikace provádí komunikaci se serverem EUR-Lexu a ve své podstatě zajišťuje mj. připojení do centrálního úložiště EUR-Lexu, kterým je repozitář CELLAR, druhá slouží pro vyhodnocení přijatých dat a jejich prezentaci uživateli. První část tedy získaná data žádným způsobem nehodnotí a nečiní z nich závěry – výjimkou jsou výjimečné situace spočívající v neexistenci uživatelem zadané směrnice nebo zadání směrnice, ke které neexistuje žádná judikatura SDEU. Až druhá část systému následně překládá hrubá data do uživatelsky přívětivé podoby a rozhoduje také, co které údaje znamenají. Na tomto místě znovu předesíláme, že hranice těchto částí je na první pohled poměrně neostrá a za účelem jasnosti naší demonstrace dochází ke zvýraznění oddělení.

3a Komunikace a získávání dat

Komplikovanost získávání dat, která pro potřebu naší demonstrace získáváme ze tří různých zdrojů, je dle našeho názoru výrazně určena neoptimálním zpracováním příslušné dokumentace k API³⁴ EUR-Lexu, technickým omezením při dotazech na samotnou existenci různých dokumentů a vysokou mírou komplikovanosti metadatové XML struktury, která je vrácena v odpovědi na rešeršní dotazy. Právě z toho důvodu tak pracujeme hned se třemi druhy dat, abychom obešli některé technické a organizační limity EUR-Lexu.

Prvním zdrojem našich informací je samotné API EUR-Lexu. Tento zdroj využíváme k získání základních informací o směrnici, jako je ověření její existence, a některých primárních charakteristik, jako je její název zobrazovaný na stránce výstupu rešerše. Mimo těchto dat získáváme z tohoto zdroje také seznam všech dokumentů, které interpretují uživatelem zadanou směrnicí a zároveň spadají do sektoru judikatury. Zajímavé je, že v jinak obsáhlém XML výstupu, což je dáno i nutností obsáhnout různé jazykové verze, je k prováděcí judikatuře pouze málo údajů – identifikace probíhá pouze celexovým číslem. Tento stav je na jednu stranu velmi výhodný, protože nám umožňuje jednoznačně identifikovat dokument a vyhnout se tak některým problémům, které jsme popsali výše. Na druhou stranu jsme ale nuceni další informace, jako je spisová značka, názvy stran sporu nebo datum rozhodnutí,

³⁴ Application Programming Interface. Jedná se o standardní rozhraní pro komunikaci se serverem na úrovni aplikací.

získávat alternativně – spisovou značku extrahujeme z celexového čísla a další údaje opatřujeme z jiných zdrojů.

Druhým zdrojem, ze kterého čerpáme data, jsou dokumenty, které obsahují pravidelně aktualizované seznamy všech soudních řízení historicky vydaných SDEU. V nich jsou uvedeny základní informace o daných rozhodnutích, jako je populární název (který se někdy kryje s názvy stran), spisová značka nebo datum rozhodnutí. Nejužitečnější informací obsaženou v tomto zdroji je identifikace věcí, které byly soudem spojeny k projednání ve stejném řízení (např. z důvody podobné předběžné otázce). Nevýhodou tohoto dokumentu je formát, který je standardní HTML pro zobrazení jako webové stránky a není primárně určen pro strojové zpracování. Za použití vhodných regulárních výrazů jsme ovšem schopni i tato data spolehlivě číst.

Třetím zdrojem, se kterým pracujeme je centrální repozitář CELLAR. Ten využíváme ke zprostředkování přístupu k plným textům dokumentů, které v aplikaci prezentujeme.

3b Prezentace dat uživateli

Tato část aplikace má dva základní úkoly. Prvním z nich je vyhodnocování dat, která jsou získána prostřednictvím výše popsaného mechanismu a odpovídající vykreslování výsledků do prohlížeče uživatele aplikace. Druhým pak je zasílání dodatečných požadavků na základě aktuálního stavu rešerše a opět jejich vyhodnocování a vykreslování výsledků.

Jako podobu vykreslení výsledků jsme zvolili tabulku, v rámci které jednotlivé řádky znamenají jednotlivá řízení před SDEU. Výhodou tohoto řešení je unifikace, kdy uživatel obdrží informace o jednotlivých řízeních v podobě, která mu umožní rychlejší orientaci.

Samotný řádek tabulky je rozdělen na tři hlavní bloky. V prvním jsou zobrazeny informace o předběžné otázce zahajující řízení, ve druhém informace o stanovisku generálního advokáta, a ve třetím bloku jsou informace ohledně vydaného konečného rozhodnutí či usnesení ve věci. Existující dokumenty jsou indikovány zelenou barvou. Jedním kliknutím je možné zobrazit jejich plné znění z repozitáře CELLAR. Neexistující dokumenty jsou indikovány červenou barvou a uživateli je poskytnut očekávaný dotaz plynoucí z běžného řazení dokumentů v repozitáři CELLAR pro případné manuální ověření reálné neexistence dokumentu.

3c Proces rešerše

Rešerši uživatel iniciuje zadáním požadavku, v jehož rámci specifikuje směrnici jejím rokem a číslem. V první fázi je na základě uživatelem zadaných hodnot vytvořeno celexové číslo příslušné směrnice a je ověřena její přítomnost v systému. Pak jsou načtena základní data z API EUR-Lexu – těmi je název směrnice a také seznam řízení, kterými byla tato směrnice interpretována.

Následně je tento strohý dílčí výstup odeslán jako mezistupeň do prohlížeče uživatele. Seznam je postupně doplňován načítanými podrobnostmi, mezi které patří identifikace jednotlivých případů stranami nebo indikace stavu řízení prostřednictvím barevného semaforu. Existence nezodpovězené předběžné otázky – ta se v systému projeví existencí dokumentu typu CN, ale neexistencí dokumentu CJ nebo CO – je zcela zásadní informací, které jsme schopni v rámci naší aplikace získat,³⁵ protože implikuje vývoj interpretace předpisu v dohledné době. Silnou stránkou aplikace je také schopnost detekovat spojení

³⁵ Dle našeho povědomí v českých komerčních právních informačních systémech podobná funkce neexistuje.

řízení s jiným – v takovém případě je odpovídající položka přesunuta do výsledku řízení, které je vedené jako hlavní, s poznámkou, že došlo ke spojení.

Výsledkem je tedy zobrazení kompletního přehledu judikatury SDEU k rešeršované směrnici, přičemž uživateli jsou přehledně zobrazeny následující údaje (viz obrázek níže):

- Číslo věci
- Populární název
- Stav řízení a existence jednotlivých dílčích dokumentů
- Odkaz na plné znění všech existujících dokumentů
- Informace o spojených řízeních

Obrázek:

Na obrázku je výstup ze systému v případě rešerše směrnice 95/46/ES. Na obrázku je tak postupně vidět (1) řízení ve věci C-70/10, kde je v systému k nalezení předběžná otázka, stanovisko generální advokáta a závěrečné rozhodnutí; (2) řízení ve věci C-582/14, kde je v systému možné nelézt znění předběžné otázky a stanovisko generální advokáta, ale nikoli závěrečné rozhodnutí či usnesení a je tak možné očekávat brzkou změnu názoru, týkajícího se některé části rešeršované směrnice; a (3) rozhodnutí ve spojených věcech C-141/12 a C-372/12, kdy v systému nalezneme předběžné otázky pro obě původně samostatná řízení, dále stanovisko generální advokáta pro řízení spojená a také finální rozhodnutí SDEU.

Case Scarlet Extended (C-70/10)		
Reference	AG's opinion	Judgment
62010CN0070	62010CC0070	62010CJ0070
Full text	Full text	Full text

Case Breyer (C-582/14)		
Reference	AG's opinion	Order
62014CN0582	62014CC0582	62014CO0582
Full text	Full text	Full text

Case YS and others (C 141/12 and C 372/12)		
Reference	AG's opinion	Judgment
62012CN0141	62012CC0141	62012CJ0141
Full text	Full text	Full text

Joined case: [62012CN0372*](#)

4 Závěr

V rámci naší demonstrace jsme identifikovali některé problémy, kterým bychom se nyní rádi věnovali. Jedním z nich je právě dvojí možnost ukončení řízení ve věci před SDEU. Zatímco primárním způsobem ukončení řízení je rozhodnutí, z procesních pravidel plyne i možnost ukončit řízení usnesením. Rozhodnutí se překládají do všech úředních jazyků EU, zatímco usnesení nikoli. Tato specifická forma ukončení řízení se pak projeví jako existující dokument, ale při snaze vyvolat dokument z centrálního repozitáře CELLAR v jazyce, ve kterém dokument neexistuje, nedojde k vrácení dokumentu jako takového, ale jeho XML dokumentu bez příslušného jazykového znění. Užitečnost tohoto XML záznamu je poměrně

vysoká, protože nám umožňuje určit základní vlastnosti usnesení bez znalosti konkrétního jazyka. Jsme tak schopni říci, na jaké jiné dokumenty dané usnesení odkazuje. Tento problém nám ale neumožňuje tvrdit, že je předložená demonstrace uživatelsky přívětivá, protože po uživateli opět vyžaduje jistou úroveň expertních znalostí.

Dalším problémem je nekonzistentní přítomnost dokumentů obsahující předběžnou otázku. Některé dokumenty se vrací jako neexistující, protože se v minulosti do systému nezaváděly samostatně. Text otázek je možné dohledat v rámci rozhodnutí (typ CJ). Z existence další fáze řízení, tedy stanoviska generálního advokáta (typ CC) nebo rozhodnutí (typ CJ) je možné usuzovat, že předběžná otázka byla položena, ale v systému se daný dokument nevyskytuje samostatně. Pro budoucí implementaci našich postupů tak zřejmě bude nutné vyřešit mj. i otázku, zda vlastně dokumenty typu CN zobrazovat vždy, nebo je zobrazovat pouze v případě, že po nich nenásleduje dokument typu CJ/CO, který by řízení končil. V případě druhého zmiňovaného postupu bychom tak poskytovali informaci o existenci nevyřešené předběžné otázky, ale už neposkytovali informaci o znění předběžné otázky ve chvíli, kdy je v systému přítomno rozhodnutí, přičemž v tomto případě je znalosti znění předběžné otázky pro právní praxi nepochybně přínosnější. Při zpětném pohledu se vykreslování samostatných dokumentů pro předběžné otázky jeví jako nadbytečné a pro uživatele matoucí.

Třetím problémem je problematické získávání komplexních informací o řízeních. Při snaze získat údaje o spojených řízeních jsme se stali obětí Hlavy 22 – pokud máme řízení a potřebujeme zjistit, ke kterému bylo připojeno (a pokud vůbec), potřebujeme metadatový záznam onoho neznámého rozhodnutí. I z toho důvodu jsme nakonec přikročili k výše popsanému druhému datovému zdroji, přes jisté těžkosti s jeho automatizovaným zpracováním. Na druhé straně, považujeme ho, po obsahové stránce a datové bohatosti, za téměř inspirativní pro budoucí zdokonalení EUR-Lexu.

Přes shora uvedené problémy, se kterými jsme se setkali, považujeme metadatovou bohatost EUR-Lexu za bezprecedentní v rámci volně dostupných systémů právních informací. Máme za to, že je zde značný potenciál pro aktivismus v oblasti zpřístupňování dat. Kromě námi demonstrovaného spojování a jednoduché vizualizace řízení máme za to, že je možné vytěžovat komplexní citační vztahy mezi jednotlivými dokumenty, potažmo přikročit k sofistikovanějším vizualizačním technikám nebo tvořit automaticky anotované texty jednotlivých právních předpisů.

Dostupnost

V textu popsaný nástroj je k dispozici na celex.lawportal.cz, kód potom na GitHubu (<https://github.com/martinloucka/celex>).

Poděkování

Tento text byl podpořen Masarykovou univerzitou v rámci projektu specifického výzkumu Metodologie citační analýzy judikatury (MUNI/A/0940/2015). Předchozí fáze jsme měli možnost představit národnímu publiku v rámci konference ČPIT 2015 (*Využití celexového čísla pro automatizaci rešerší*) a mezinárodnímu publiku v rámci konference Cyberspace 2015 (*Access to Legal Information in EU: Case of Celex Numbers*). Rádi bychom poděkovali účastníkům obou konferencí za jejich připomínky k naší práci.

Reference

- BAŇOUCH, Hynek a Zdeněk KÜHN. O publikaci a citaci judikatury aneb proč je někdy judikatura jako císařovy nové šaty. *Právní rozhledy*, 2005, roč. 13, č. 13, s. 484-491
- BOBEK, Michal, KÜHN, Zdeněk a kol. *Judikatura a právní argumentace. 2. přepracované a aktualizované vydání*. Praha: Auditorium, 2013.
- BOBEK, Michal. A Fourth in the Court: Why Are There Advocates General in the Court of Justice. *Cambridge Yearbook of European Legal Studies*, 2012, roč. 14, s. 529-561.
- David C. Blair a M. E. Maron. An evaluation of retrieval effectiveness for a full-text document-retrieval system. *Communications of the ACM*, 1985, roč. 28, č. 3, s. 289-299.
- DERLÉN, Mattias a Johan LINDHOLN. Goodbye van Gend en Loos, Hello Bosman? Using Network Analysis to Measure the Importance of Individual CJEU Judgments. *European Law Journal*, 2014, roč. 20, č. 5, s. 667-687.
- FOWLER, James H. a Sangick JEON. The authority of Supreme Court precedent. *Social Networks*, 2008, roč. 30, č. 1, s. 16-30.
- FOWLER, James H. et al. Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court. *Political Analysis*, 2007, roč. 15, č. 3, s. 324-346.
- GARFIELD, Eugene. Citation Indexes for Science. *Science*, 1955, roč. 122, č. 3159, s. 108-111.
- GEIST, Anton. The Open Revolution: Using Citation Analysis to Improve Legal Text Retrieval. *European Journal of Legal Studies*, 2010, roč. 2, č. 3, s. 137-145.
- KONIARIS, Marios, ANAGNOSTOPOULOS, Ioannis a Yannis VASSILIOU. Legislation as a complex network: Modelling and analysis of European Union legal sources. In: HOEKSTRA, Rinke (ed.). *JURIX 2014: The Twenty-Seventh Annual Conference*. Amsterdam: IOS Press, 2014, s. 143-152.
- KŘÍŽ, Vincent, HLADKÁ, Barbora, DĚDĚK, Jan a Martin NEČASKÝ. Statistical Recognition of References in Czech Court Decisions. In: GELBUKH, Alexander, ESPINOZA, Félix Castro a Sofía N. GALICIA-HARO (eds.). *Human-Inspired Computing and its Applications. MICAI 2014*. Springer: Berlin, Heidelberg, 2014, s. 51-61.
- LANDES, William M. LESSIG, Lawrence a Michael E. SOLIMINE. Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges. *Journal of Legal Studies*, 1998, roč. 27, č. 2, s. 271-332.
- LUPU, Yonatan a Erik VOETEN. Precedent in International Courts: A Network Analysis of Case Citations by the European Court of Human Rights. *British Journal of Political Science*, 2012, roč. 42, č. 2, s. 413-439.
- LUPU, Yonatan a James H. FOWLER. Strategic Citations to Precedent on the U.S. Supreme Court. *Journal of Legal Studies*, 2013, roč. 42, č. 1, s. 151-186.
- MOENS, Maria-Francine. Innovative techniques for legal text retrieval. *Artificial Intelligence and Law*, 2001, roč. 9, č. 1, s. 29-57.
- MYŠKA, Matěj, HARAŠTA, Jakub, LOUTOCKÝ, Pavel a Jakub MÍŠEK. *Revue pro právo a technologie*, 2016, roč. 7, č. 13, s. 147-189.
- PRICE, Derek J. de Solla. Network of Scientific Papers. *Science*, 1965, roč. 149, č. 3683, s. 510-515.

SADL, Urška a Yannis PANAGIS. What is Leading Case in EU Law? An Empirical Analysis. *European Law Review*, 2015, roč. 40, č. 1, s. 15-34.

SCHWEIGHOFER, Erich. The Revolution in Legal Information Retrieval or The Empire Strikes Back. *Journal of Information, Law and Technology*, 1999, roč. 4, č. 1, nestránkováno.

SMITH, Linda C.. Citation Analysis. *Library Trends*, 1981, roč. 30, č. 1, s. 83-106.

STUPKA, Václav. E-Sbírka a e-Legislativa. *Revue pro právo a technologie*, 2014, roč. 5, č. 9, s. 185-196.

WIDDISON, Robin. New Perspectives in Legal Information Retrieval. *International Journal of Law and Information Technology*, 2002, roč. 10, č. 1, s. 41-70.