

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra informatiky a výpočetní techniky

Bakalářská práce

Zpracování dat o vědeckých projektech EU

ZÁPADOČESKÁ UNIVERZITA V PLZNI

Fakulta aplikovaných věd
Akademický rok: 2022/2023

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Lucie LEXOVÁ**
Osobní číslo: **A22B0096P**
Studijní program: **B0613A140015 Informatika a výpočetní technika**
Specializace: **Informatika**
Téma práce: **Zpracování dat o vědeckých projektech EU**
Zadávající katedra: **Katedra informatiky a výpočetní techniky**

Zásady pro vypracování

1. Seznamte se s repozitáři s vědeckými výsledky evropských projektů, zejména s ohledem na strukturu dostupných dat.
2. Seznamte se s metodami vizualizace dat v bibliografických a podobných nástrojích.
3. Navrhněte nástroj, který umožní zpracovávat data z databáze Cordis nebo podobných zdrojů a zobrazit vztahy mezi záznamy uživatelů.
4. Navržené řešení implementujte.
5. Vytvořené řešení otestujte především s ohledem na uživatelskou přívětivost a použitelnost a na přesnost zobrazených dat.

Rozsah bakalářské práce: **doporuč. 30 s. původního textu**
Rozsah grafických prací: **dle potřeby**
Forma zpracování bakalářské práce: **tištěná/elektronická**

Seznam doporučené literatury:

Dodá vedoucí bakalářské práce

Vedoucí bakalářské práce: **Ing. Richard Lipka, Ph.D.**
Katedra informatiky a výpočetní techniky

Datum zadání bakalářské práce: **3. října 2022**
Termín odevzdání bakalářské práce: **4. května 2023**

L.S.

Doc. Ing. Miloš Železný, Ph.D.
děkan

Doc. Ing. Přemysl Brada, MSc., Ph.D.
vedoucí katedry

V Plzni dne 25. října 2022

Prohlášení

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně a výhradně s použitím citovaných pramenů.

V Plzni dne 21. června 2023

Lucie Lexová

Abstract

The European Union supports scientific projects with funding. The CORDIS platform allows you to browse the data of projects, including collaboration of the organisations and the reports which are related to the projects. The most important framework programmes for research and innovation in recent years are FP7 (Framework Programme 7) and H2020 (Horizon 2020).

This bachelor thesis includes an analysis of datasets from CORDIS, an overview of some existing searching tools, and the proposal of an application that is able to import datasets, search between projects, organisations and reports, visualise the data appropriately and provide the user an overview of the links between records.

Abstrakt

Evropská unie podporuje finančními prostředky vědecké projekty. Platforma CORDIS umožňuje procházení dat dotovaných projektů, včetně organizací, které na projektech spolupracují a reportech, které z projektů vychází. Nejvýznamnějšími rámcovými programy posledních let pro výzkum a inovace jsou programy FP7 (Framework Programme 7) a H2020 (Horizon 2020).

Tato práce obsahuje analýzu vybraných datasetů z aplikace CORDIS, seznámení se s podobnými, již existujícími nástroji pro vyhledávání a návrh aplikace, která je schopná datasety projektů, organizací a reportů importovat a následně v nich vyhledávat, vhodně vizualizovat data a poskytnout uživateli přehled o spojitosti mezi záznamy.

Obsah

1	Úvod	7
2	Fondy EU	8
2.1	Rámcové programy EU	8
2.2	Horizon 2020	9
2.3	Horizon Europe	10
3	Analýza dat evropských projektů	12
3.1	CORDIS	12
3.2	Data programu FP7	12
3.3	Data programu Horizon 2020	16
4	Metody vizualizace dat vyhledávacích nástrojů	18
4.1	Analýza vyhledávání výsledků vybraných nástrojů	19
4.1.1	ACM Digital Library	19
4.1.2	Cambridge University Press	19
4.1.3	DOAJ	20
4.1.4	Europeana	21
4.1.5	ScienceDirect	22
4.1.6	SCOAP ³	24
4.2	Zhodnocení vyhledávání u zkoumaných nástrojů	24
4.3	Vizualizace výsledků u zkoumaných nástrojů	25
5	Vizualizace dat evropských projektů	26
6	Návrh aplikace	29
6.1	Sekce Přehled importovaných datasetů	29
6.2	Sekce Vyhledávání	29
7	Implementace	30
7.1	Další použité technologie	30
7.1.1	Spring Framework	30
7.1.2	JPA	30
7.1.3	JSF	31
7.2	Spuštění aplikace	31
7.3	Hlavní stránka	31
7.4	Přehled importovaných datasetů	32

7.5	Vyhledávání	33
8	Testování	36
8.1	Test uživatelské přívětivosti	36
8.1.1	Opravy po získání zpětné vazby	37
8.2	Test rychlosti importu datasetů	37
8.3	Test přesnosti zobrazených dat	38
9	Závěr	42
10	Příloha A	43
11	Příloha B - Uživatelská příručka	50
11.1	Spustění aplikace	50
11.2	Přehled importovaných datasetů	51
11.3	Vyhledávání	53
	Literatura	56

1 Úvod

Evropská unie (dále jen EU) ročně vkládá do dotací miliardy eur. Finanční prostředky mohou čerpat všechny státy, které jsou součástí EU a výjimečně státy se statutem přidružené země. Hlavním cílem fondů EU je zajistit členským státům finanční prostředky, které povedou k minimalizaci rozdílů ve společenské a ekonomické sféře a to jak mezi státy, tak mezi jednotlivými regiony. Kromě strukturálních fondů, jež slouží především k obecnému zvýšení vyspělosti státu, existují i unijní fondy, mezi kterými najdeme i programy pro dotace vědeckých projektů.

Unijních programů je celá řada, patří sem například Horizont Evropa pro výzkum a inovace nebo Program v oblasti zdraví zaměřen na zlepšení zdraví občanů. Vzhledem k transparentnosti EU jsou veškerá data o dotacích dostupná na oficiálních stránkách, kde si je může kdokoliv zobrazit.

Jedním z cílů práce bude s těmito daty pracovat, nejvíce nás budou zajímat data související s vědeckými projekty. Dále budeme řešit relevantnost dat pro případné vyhledávání informací o projektech, seznámíme se s metodami vizualizace dat v podobných a již existujících nástrojích. Po analýze navrhne aplikaci, která dokáže importovat dostupná data a vrátit uživateli zpracované informace podle jeho požadavku. Po implementaci aplikace bude nutné testování, získání zpětné vazby a kontrola přesnosti zobrazených dat.

Výsledná aplikace by se v praxi dala použít k nalezení organizace, se kterou může univerzita spolupracovat. Možností bude vyhledání organizace podle lokace nebo podle spolupráce na projektech ve vědeckém oboru, který bude v souladu se zaměřením projektu univerzity.

2 Fondy EU

Dotace od Evropské unie pokrývají širokou škálu programů a projektů. Mezi nejvýznamější strukturální fondy, určené k čerpání prostředků pro zvýšení vyspělosti státu, patří fond pro regionální rozvoj (European Regional Development Fund), fond pro sociální rozvoj (European Social Fund) a fond pro soudržnost (Cohesion fund) [4].

EU dále finančně podporuje projekty, které se týkají vzdělávání (například Erasmus+), životního prostředí (Environment and Climate Action), kultury (Creative Europe) nebo výzkumu a inovací (Horizon 2020) [5].

Žádost o dotace podléhá přísnému řízení, za které je zodpovědná Evropská komise. Specialisté vyhodnotí přijaté žádosti. Každý program, pod který žádost spadá, může mít jiné postupy vyhodnocování. Poté komise vybírá nejlepší projekty dokud nevyčerpá rozpočet daného programu [6]. Finanční příspěvek může být poskytnut v různých formách. Nejčastější formou je grant, dále EU nabízí půjčky, akcie, odměny pro výherce soutěží a další [7].

2.1 Rámcové programy EU

První rámcový program (Framework Programme 1, FP1) byl navržen již v roce 1982 [9]. Program se měl týkat předmětů jako zemědělství, průmysl, suroviny, energie, výzkum a životní podmínky. Následující rok byl program EU schválen pro období 1984-1987 s rozpočtem 3,75 miliard eur [9]. Po FP1 následoval druhý rámcový program FP2, jehož záměr byl velmi podobný předchozímu programu, ale s větším důrazem na výzkum a technologie. Pro FP2 bylo na období 1987-1991 vyhrazeno 5,4 miliard eur [9]. V letech 1991-1994 byl v EU platný třetí rámcový program FP3 s rozpočtem 6,6 miliard eur [9]. Došlo k výraznému zaměření na nové technologie jako informační a komunikační, biotechnologie a technologie nových materiálů. V podobném duchu platil v období 1994-1998 rámcový program FP4 s rozpočtem 13,2 miliard eur a v období 1998-2002 FP5 s rozpočtem 13,7 miliard eur [1, 9].

Během platnosti FP5 vznikl v roce 2000 nový program, Evropský výzkumný prostor (European research area, ERA), zaměřený na výzkum a inovace, lidské zdroje, infrastrukturu a sociální vědy [1, 10]. Rámcový program FP5 pomáhal výzkumy spadající pod ERA realizovat. Při navrhování FP6 se komise rozhodla více zaměřit na výzkum a spolupráci členských státu

na vědeckých projektech. FP6 s rozpočtem 17,9 miliard eur pro období 2002-2006 obsahoval podprogramy přímo zaštitující ERA [1, 9].

V roce 2005, při navrhování FP7, bylo komisí rozhodnuto, že je nutné program zaměřit pouze na aktivity výzkumu a inovací. Trvání programu bylo prodlouženo na sedm let, tedy od roku 2007 do 2013, stejně jako Víceletý finanční rámec (Multiannual Financial Framework) [1]. Rozpočet byl stanoven ve výši 50 miliard eur [1].

Při navrhování osmého rámcového programu navrhla Evropská komise spojení všech existujících programů zaměřujících se na výzkum do jednoho [1]. Pro efektivnost a snažší přístup vědecké obce k dotacím vznikl v roce 2013 komplexní program pro výzkum a inovace Horizon 2020 [11].

2.2 Horizon 2020

Horizon 2020 (popřípadě zkráceně H2020, česky Horizont 2020) byl v letech 2014-2020 největší evropský program pro výzkum a inovace. Jednalo se v pořadí o osmý rámcový program tohoto typu. Na celé sedmileté období byly vyhrazeny finanční prostředky ve výši téměř 80 miliard eur [8]. Program má tři pilíře - Vynikající věda, Vedoucí postavení v průmyslu a Společenské výzvy [1].

Vynikající věda má za úkol přimět ke spolupráci nejlepší evropské vědce a rozvíjet pole vědy. Kromě základního výzkumu podporuje pobyty na výzkumných pracovištích a dotuje nákladná zařízení jako jsou výkoné lasery, medicínský materiál, letadla či monitorovací stanice [1]. Dalším úkolem tohoto pilíře je motivovat vědce, podporovat výzkum nových technologií a udržet tím Evropu konkurenceschopnou, což povede ke vzniku nových vědeckých pracovišť a dalším spolupracím na technologiích budoucnosti.

Pilíř Vedoucí postavení v průmyslu se orientuje na investice do užitečných technologií jako je mikroelektronika, informační a komunikační technologie, biotechnologie či nanotechnologie [1]. Je potřeba podporovat podniky a motivovat je ke spolupráci s veřejným sektorem. Tím se přispívá nejen k inovativnosti, ale i k novým pracovním pozicím a snižování nezaměstnanosti.

Společenské výzvy představují soubor výzev, u kterých se předpokládá, že investice do nich mohou přinést užitek občanům. Pilíř zahrnuje především výzvy týkající se zdraví. Jedním z cílů je podporovat zdravotnický výzkum a nové způsoby léčby. Další výzvy se týkají udržitelnosti, recyklace a životního prostředí obecně, proto EU podporuje bioenergie, odvětví založená na biotechnologiích a "zelenou ekonomiku"¹ [1]. Soubor zahrnuje i výzvy pro

¹zelená ekonomika - ekonomika s minimálními negativními dopady na životní prostředí

ochranu a bezpečnost občanů. V zájmu EU je podpora výzkumu a inovací takových technologií, které přispějí k bezpečnosti a zároveň ochrání soukromí občanů [1].

2.3 Horizon Europe

Na program Horizon 2020 přímo navazuje devátý rámcový program s názvem Horizon Europe (Horizont Evropa). Jeho přípravy začaly v roce 2018, kdy komise EU a členské státy navrhly hlavní cíle programu. Na období 2021-2027 je připraven rozpočet ve výši 95,5 miliard eur [12]. Strategie nového rámcového programu, Horizon Europe, má následující body - zvýšení inovačního potenciálu v EU, zajištění větší transparentnosti, širší synergie s ostatními programy EU a usnadnění přístupu k programu prostřednictvím digitální transformace [2, 3].

Díky získaným zkušenostem s organizací předchozího programu bylo možné přinést do Horizon Europe vylepšení. Jedním z nich je vytvoření rady pro inovace (European Innovation Council), která bude mít za úkol podporovat průlomové, byť rizikové, projekty a inovativní start-upy [12]. Další novinkou je definování misí (EU Missions), jejichž řešení výrazně ovlivní evropskou společnost. Mise zahrnují boj proti rakovině, adaptace na klimatické změny, podporu "chytrých měst"², zajištění zdravé, netoxické půdy a jídla a ochranu oceánů a vodních ploch [12]. Program by měl efektivněji přistupovat k novým partnerstvím jak z veřejného, tak soukromého sektoru a dát organizacím možnost spolupracovat i se státy mimo EU [12]. Součástí programu je nařízená otevřenost (Open Science Policy), tedy povinné sdílení publikací a výsledků výzkumů [12].

Struktura programu je rozdělena do tří pilířů - Vynikající věda, Globální výzvy a konkurenceschopnost evropského průmyslu a Inovativní Evropa. Napříč pilíři je blok zvaný Rozšiřování účasti a posílení Evropského výzkumného prostoru [13].

Pilíř Vynikající věda se příliš neliší od pilíře předchozího programu Horizon 2020. Pod pilíř spadá především spolupráce vědců na inovativních vědeckých projektech a pobyty na výzkumných pracovištích. Akce Marie Skłodowska-Curie se mimo jiné snaží o přilákání mladých talentovaných studentů do světa vědy [12].

Druhý pilíř, Globální výzvy a konkurenceschopnost evropského průmyslu, je rozdělen do šesti jednotlivých klastrů podle oboru:

²chytré město - město využívající moderní technologie pro zvýšení blahobytu občanů, avšak bez zátěže na životní prostředí

1. Zdraví,
2. Kultura, kreativita a inkluzivní společnost,
3. Občanská bezpečnost ve společnosti,
4. Digitální oblast, průmysl a vesmír,
5. Klima, energetika a mobilita,
6. Potraviny, biohospodářství, přírodní zdroje, zemědělství a životní prostředí [14].

Pod pilíř dále spadá Společné výzkumné středisko (Joint Research Centre), jež poskytuje nezávislé vědecké důkazy a technické poradenství [2, 15].

Pilíř Inovativní Evropa se zaměřuje na podporu inovativních produktů a ekosystémů. Patří sem již zmíněná Evropská rada pro inovace, která pomáhá průlomovým projektům. Evropské inovační ekosystémy (European Innovation Ecosystems) a Evropský inovační a technologický institut (European Institute of Innovation and Technology) se snaží o spojení vzdělávání, výzkumu a inovací dohromady a dosáhnout tak největšího přínosu pro všechny zúčastněné [2, 14].

Blok Rozšiřování účasti a posílení Evropského výzkumného prostoru podporuje členské státy EU v jejich výzkumech a inovacích [14]. Cílem je odstranění bariér, jež státu brání úplnému využití potenciálu ve vědě [16]. Pod tuto platformu spadají Rozšiřování účasti a šíření excelence (Widening participation and spreading excellence) a Reforma a zlepšování evropského systému výzkumu a inovací (Reforming and Enhancing the European R&I system) [14].

3 Analýza dat evropských projektů

Na oficiálních stránkách EU můžeme najít databáze dotovaných projektů a jejich výsledků. Každý program má pro přehlednost svou vlastní webovou stránku s odkazy na stažení datasetů. Nás budou nejvíce zajímat datasety projektů pro výzkum a inovace dostupné na CORDIS (The Community Research and Development Information Service). Cíl této kapitoly je seznámení se s repozitáři s vědeckými výsledky evropských projektů.

3.1 CORDIS

CORDIS je přední zdroj dat projektů pro výzkum a inovace. Zahrnuje všechny programy od FP1 po Horizon 2020 [17]. Aplikace umožňuje snadné vyhledávání napříč všemi programy pomocí slov a filtrů. Vyhledávat můžeme nejen projekty, ale i publikace a výsledky. Pokud bychom chtěli získat data celého programu, na stránkách CORDIS jsou aktuálně ke stažení pouze data mladších programů (FP7 a H2020), starší programy (FP1-FP6) jsou dohledatelné ve webových archivech EU [18]. Další možností je využití portálu EU Open Data, kde můžeme dohledat data všech programů po zadání klíčového slova "cordis"[19]. Datasety projektů jsou nejčastěji ke stažení ve formátech CSV a XML.

Aplikace kromě vyhledávání umožňuje uživatelům zveřejňovat články související s projekty. Každý článek projde před zveřejněním přísnou kontrolou, kterou mají na starosti editoři CORDISu [19]. Vybrané příspěvky mohou být publikovány v CORDIS magazínu Research*eu nebo jako součást brožury CORDIS Results Packs [20].

3.2 Data programu FP7

Datasety FP7 jsou ke stažení ve formátech CSV, XLSX pro EXCEL a jako ZIP archiv s jednotlivými XML soubory. Dostupná jsou data o projektech, organizacích a reportech [21]. Po stažení všech tří datasetů ve formátu CSV a následném vložení do databázového nástroje MySQL Workbench můžeme všechny informace prohlížet.

První soubor (FP7 Projects) obsahuje následující atributy projektů:

1. *rcn* - zkratka pro Record Control Number, jedná se o jedinečné číslo záznamu;
2. *id* - jedinečné číslo projektu vytvořené při vzniku dohody o grantu;
3. *acronym* - zkratka vytvořená z názvu projektu;
4. *status* - má tři možnosti, CAN (canceled - zrušený), CLO (closed - uzavřený), ONG (ongoing - probíhající);
5. *programme* - specifikace podprogramu, například SPACE, ENERGY, TRANSPORT;
6. *topics* - kód zaměření a zároveň pod jaký podprogram projekt spadá, například kód AAT-2007-1.1-03 říká, že projekt spadá pod program spolupráce a transportu a týká se pohonných systémů. Tyto údaje je nutné hledat ve speciálních tabulkách dostupných na EU Open Data;
7. *frameworkProgramme* - rámcový program projektu, v tomto případě pouze FP7;
8. *title* - název projektu;
9. *startDate* - datum začátku projektu;
10. *endDate* - datum ukončení projektu;
11. *projectUrl* - adresa webových stránek projektu;
12. *objective* - krátký popis projektu a jeho cíle;
13. *totalCost* - celkové finanční náklady [€];
14. *ecMaxContribution* - maximální finanční příspěvek od EU [€];
15. *call* - kód řízení, například FP7-AAT-2007-RTD-1 znamená, že projekt FP7 s podprogramy spolupráce a transport spadá pod řízení RTD (Research, Technology and Development) projektů;
16. *fundingScheme* - určuje typ akce, například CP-FP (Collaborative Project - Focused research Project) je označení pro malý až střední výzkumný projekt zahrnující spolupráci;
17. *coordinator* - přední organizace, která projekt vede;
18. *coordinatorCountry* - kód země, ve které koordinátor sídlí;

19. *participants* - spoluúčastníci projektu, pokud je jich více, jsou odděleny středníkem;
20. *participantCountries* - kódy zemí, ve kterých spoluúčastníci sídlí, popřípadě odděleny středníkem;
21. *subjects* - kód oboru projektu [21].

Dataset (FP7 Organizations) s informacemi o organizacích obsahuje atributy:

1. *projectRcn* - číslo záznamu projektu;
2. *projectID* - id projektu - vazba na projekt z výše uvedeného souboru (FP7 Projects);
3. *projectAcronym* - zkratka projektu;
4. *role* - pouze jedna z možností coordinator/participant;
5. *id* - identifikační číslo organizace;
6. *name* - název organizace;
7. *shortName* - zkrácený název organizace, popřípadě u organizací s krátkým celým názvem bývá *shortName* ekvivalentní s *name*;
8. *activityType* - zde je několik možností, HES (Higher or Secondary Education Establishments), REC (Research Organisations), PRC (Private for-profit entities), PUB (Public bodies) a OTH (Other);
9. *endOfParticipation* - konec účasti organizace, může nabývat pouze pravdivostních hodnot true/false;
10. *ecContribution* - výše finančního příspěvku od EU [€];
11. *country* - kód země, kde organizace sídlí;
12. *organizationUrl* - adresa webových stránek organizace;
13. *contactForm* - odkaz na kontaktní formulář na stránkách EU;
14. - 24. *street*, *city*, *postCode*, *vatNumber*, *contactType*, *contactTitle*, *contactFirstNames*, *contactLastNames*, *contactFunction*, *contactTelephoneNumber*, *contactFaxNumber* - údaje o organizaci a kontaktních osobách [21].

Soubor (FP7 Reports) s informacemi o výsledcích projektů má atributy:

1. *rcn* - zkratka pro Record Control Number, jedná se o jedinečné číslo záznamu;
2. *language* - zkratka jazyku, kterým je report napsán;
3. *title* - nadpis reportu;
4. *teaser* - krátká upoutávka;
5. *summary* - souhrn reportu;
6. *workPerformed* - popis odvedené práce;
7. *finalResults* - finální výsledky;
8. *lastUpdateDate* - datum poslední změny;
9. *country* - kód země, ze které report pochází;
10. *projectID* - id projektu reportu - vazba na projekt z výše uvedeného souboru (FP7 Projects);
11. *projectAcronym* - zkratka projektu;
12. *programme* - specifikace podprogramu, například SPACE, ENERGY, TRANSPORT;
13. *topics* - kód zaměření a zároveň pod jaký podprogram projekt (k němuž se report vztahuje) spadá, například kód AAT-2007-1.1-03 říká, že projekt spadá pod program spolupráce a transportu a týká se pohonných systémů. Tyto údaje je nutné hledat ve speciálních tabulkách dostupných na EU Open Data;
14. *relatedFile* - cesta k souboru v dokumentaci CORDISu;
15. *url* - odkaz na stránky reportu;

Po rozboru datasetů můžeme vytvořit ER (Entity Relationship, entitně vztahový) model, který je součástí přílohy (10.1). Z analýzy je zřejmé, že jsou některé údaje nadbytečné. Projekt má dvě jedinečná čísla *rcn* a *id*, přičemž by stačilo pouze jedno z nich. Jelikož známe celý název projektu, je pro nás *acronym* nepotřebný. Data týkající se specifikace projektu, *programme*, *call* a *fundingScheme*, často obsahují složité, těžko rozklíčovatelné kódy, které jsou pro nás nepotřebné. Dataset projektů dále obsahuje informace o příslušných

organizací. Tyto informace bychom mohli vynechat, protože se vyskytují i v datasetu organizací, jsou tedy v relaci duplicitní. Poslední atribut, *subject*, který má popisovat obor projektu, bohužel v tabulkách není vyplněný, proto i ten je pro nás nepotřebný. Můžeme tedy navrhnout následující soubor atributů tak, abychom zachovali nejdůležitější informace: *id, status, topics, frameworkProgramme, title, objective, startDate, endDate, projectUrl, totalCost, ecMaxContribution*.

Pokud se zaměříme na dataset organizací, i zde bychom mohli některé informace označit za nadbytečné. Záznamy obsahují data o projektu, *projectRen, projectID* a *projectAcronym*. Vzhledem k propojení tabulek je pro nás dostačující jedna jedinečná identifikace projektu a ostatní informace můžeme dohledat v datasetu projektů. Atribut *shortName* je nepotřebný, protože známe celý název organizace. Za nadbytečné by se daly považovat vedlejší údaje o organizaci a kontaktních osobách (viz výše 14.-24. bod datasetu organizací). Pokud bychom chtěli organizaci kontaktovat, bude praktičtější si kontakt na organizaci vyhledat na internetu. Problém těchto údajů je následující: pokud si organizace určí více kontaktních osob, záznamy se až na kontaktní osobu duplikují, čili nelze u jedné organizace určit více kontaktních osob v jednom záznamu. Vede to jednak k zbytečně většímu objemu dat a také obtížnému určení primárního klíče tabulky organizací. Můžeme navrhnout zjednodušený, ale dostačující soubor atributů: *id, name, activityType, endOfParticipation, ecContribution, country, city, organizationUrl*.

Jelikož je mezi projekty a organizacemi vazba typu M:N, můžeme vytvořit vazební tabulku (včetně role organizace na projektu) a samotný záznam organizace si již nemusí držet identifikátor projektu. Nebudeme tedy muset uchovávat duplicitní data organizací, jako v původním datasetu. Vazební tabulka by obsahovala atributy: *projectId, organizationId, role*.

Jeden report je vázán na jeden projekt. Dataset reportů obsahuje informace o projektu, které už z tabulky projektů známe. Stačí nám tedy pouze vazba na identifikátor projektu a ostatní data dohledáme z relace. Údaje *workPerformed* a *finalResults* jsou většinou nevyplněné. Datum poslední změny *lastUpdateDate* je pro nás nepotřebný. Dostačující soubor atributů by mohl vypadat následovně: *id, language, title, teaser, summary, country, projectID, relatedFile, reportUrl*.

3.3 Data programu Horizon 2020

Stejně jako u datasetů FP7, jsou i datasey programu Horizon 2020 ke stažení ve formátech CSV, XLSX pro EXCEL a ZIP. Můžeme si stáhnout pět

souborů s daty o projektech, organizacích, publikacích, reportech a výstupech [22]. Následně se zaměříme na data o projektech, organizacích a reportech.

Soubor (H2020 Projects) s informacemi o projektech má totožné atributy jako dataset projektů pod programem FP7. Hodnoty se liší u následujících atributů:

1. *status* - nabývá jednu ze tří hodnot, SIGNED (podepsaný), CLOSED (uzavřený) a TERMINATED (ukončený);
2. *programme* - specifikace programu, musí se dohledat ve speciálních tabulkách, například H2020-EU.3.1.7. znamená, že se projekt věnuje inovativní medicíně;
3. *topics* - kód zaměření projektu, například IMI2-2014-01-01 říká, že předmětem projektu jsou přístupy k terapii při onemocnění diabetem 1. typu;
4. *frameworkProgramme* - rámcový program projektu, zde ve všech případech H2020;
5. *call* - kód řízení, takřka identický s kódem zaměření;
6. *fundingScheme* - určuje typ akce, například IMI2-RIA je označení pro projekt týkající se inovativní medicíny a je RIA (Research and Innovation Action) [22].

Zdroj informací o organizacích (H2020 Organisations) má opět totožné atributy jako dataset organizací FP7. Jediný rozdíl v hodnotách atributů je následující:

1. *role* - jedna z možností coordinator/participant/partner.

Dataset o reportech (H2020 Reports) má všechny atributy stejné jako dataset reportů FP7.

Součástí přílohy (10.2) je ER model programu H2020, ze kterého je zřejmý méně komplikovaný primární klíč tabulky organizace. Je nutné zdůraznit, že to je vzhledem k tomu, že data týkající se kontaktních osob jsou v tomto datasetu prázdná, tudíž nedochází k duplicitám atributu *id*. Všechny výše analyzované datové sady jsou analogické k předchozím datovým sadám FP7 a jejich závěry jsou stejné. I zde jsou některá data nadbytečná a mohli bychom u nich zjednodušit soubory atributů. Detailní výpis a jak by zjednodušené soubory atributů mohly vypadat viz kapitola 3.2.

4 Metody vizualizace dat vyhledávacích nástrojů

V dnešní době je dostupné velké množství elektronických informačních zdrojů. Pokud jsou zdroje seriózní a ověřitelné, slouží k získání informací například pro vědu a výzkum. Elektronické informační zdroje můžeme rozdělit podle typů na:

- online katalogy = elektronické katalogy knihoven;
- databáze (profesionálních informací);
- oborové portály;
- digitální knihovny, e-knihy, e-časopisy [23].

Další dělení může být podle typu databáze:

- fulltextové - zahrnují kompletní text knihy, článku atd., též je můžeme najít pod názvem *plnotextové*;
- bibliografické – obsahují jen abstrakty, citace, odkazové informace, klíčová slova ap.;
- faktografické - obsahují konkrétní data, např. statistiky nebo numerické hodnoty;
- obrazové – zahrnují data fotografií či obrázků;
- multimediální – pro zpřístupnění audio/video souborů;
- smíšené - kombinace výše uvedených typů databází do jedné [23][24].

Dále se hodí zmínit rozdělení elektronických informačních zdrojů na primární a sekundární. Primární zdroje jsou nosiče přímých informací a faktů. Mohli bychom sem zařadit například knihu (včetně elektronické verze), celý článek, report výzkumu a podobně. Sekundární zdroje obsahují informace o těchto jich zmíněných primárních zdrojích, tzn. obsahují pouze odkazy, citace a další bibliografické informace. Příkladem může být primární zdroj, který by pro nás představoval celý text článku a sekundární zdroj, kde už bychom z článku našli pouze abstrakt. Pokud chceme v datech vyhledávat, musíme zvážit parametry jako téma, klíčová slova, jazyk nebo geografické

omezení, časové omezení atd. Podle zvážení těchto parametrů, pak vyplňujeme vyhledávací pole, popřípadě tvoříme dotazy, ve vybraném nástroji (např. v databázi) [23].

Západočeská univerzita v Plzni nabízí registrovaným uživatelům přístup k licencovaným elektronickým informačním zdrojům. V následujících řádcích se seznámíme s některými z nich. Výběr především zahrnuje zdroje z oblasti vědy a technologií, vědecké články a časopisy. Cílem kapitoly je se seznámit s metodami vizualizace dat v bibliografických a podobných nástrojích, abychom byli schopni snadno navrhnout vlastní nástroj pro vizualizaci.

4.1 Analýza vyhledávání výsledků vybraných nástrojů

4.1.1 ACM Digital Library

ACM (Association for Computing Machinery) Digital Library je platforma umožňující vyhledávání mezi více než třemi a půl miliony publikacemi jako jsou časopisy a knihy z oblasti počítačových věd. U vybraných vydavatelů umožňuje full-textové vyhledávání [25].

Při otevření rozšířeného vyhledávání se nám otevře stránka s poli pro vyplnění. Jsou zde i tooltipy s nápovědou u jednotlivých polí. Pravá část stránky zahrnuje detailní nápovědu, jak filtr použít a jak tvořit fráze pro full-textové vyhledávání.

Všechna pole rozšířeného vyhledávání:

- *Search items from* - základní informace, odkud chceme data vyhledat;
- *Search Within* - umožňuje vybrat libovolné množství atributů a u nich zadat vyhledávací text, můžeme tedy například zadat klíčová slova titulku nebo konkrétní jméno autora;
- *Filters* - umožňuje vyfiltrovat konkrétního vydavatele, agenturu a další;
- *Publication Date* - ke specifikaci data vydání publikace, můžeme využít přednastavený rozsah nebo si zadat vlastní.

4.1.2 Cambridge University Press

Cambridge University Press spadající pod Cambridge Core je online služba, která byla navržena ke snadnému a rychlému procházení obsahu pro akademický výzkum. Zahrnuje více než 1,8 milionů článků a přes 46 tisíc knih [27].

Hlavní stránka nabízí vyhledávání podle klíčových slov a nebo možnost vybrat jeden z předmětů vyhledávání jako je například právo, matematika nebo psychologie. Po výběru předmětu můžeme vybrat konkrétní magazín nebo knihu v oblasti, která nás zajímá. Například při výběru předmětu "Počítačové vědy" dostaneme nabídku rozdělenou na magazíny a knihy. V magazínech je nabídka konkrétních magazínů jako "Theory and Practice of Logic Programming", "Probability in the Engineering and Informational Sciences" nebo "Mathematical Structures in Computer Science" a po rozkliknutí dostaneme detail magazínu včetně posledních článků, jež v něm byly publikovány. Co se týče knih, tam můžeme vybrat jednu z oblastí (např. "Software Engineering and Development"). Po kliknutí na oblast nás aplikace přeměruje na stránku s rozšířeným vyhledáváním. Ve zvolené oblasti nyní můžeme knihy filtrovat podle parametrů:

- *Search* - základní vyhledávání klíčových slov přes všechny atributy;
- *Access* - umožňuje vybrat, jestli chceme zobrazit obsah, který je přístupný jen uživatelům nebo všem;
- *Content type* - rozhodnutí, zda chceme hledat knihu nebo jen kapitoly z knih;
- *Author* - autor knihy;
- *Publication Date* - ke specifikaci data vydání publikace, můžeme využít přednastavený rozsah nebo si zadat vlastní;
- *Tags* - vyhledávání přes tagy;
- *Publishers* - volba vydavatele.

4.1.3 DOAJ

DOAJ, celým názvem Directory of Open Access Journal, je platforma k prohlížení vědeckých časopisů s otevřeným přístupem. Momentálně DOAJ nabízí vyhledávání mezi téměř 9 miliony odbornými články [28].

Na vyhledávací stránce nejprve vybereme, jestli chceme vyhledávat časopisy nebo články. Poté pokračujeme zadáním klíčových slov a výběrem, jestli chceme hledat klíčová slova přes všechny atributy nebo zvolíme nějaký konkrétní. Výběr atributů se pro články a časopisy liší. Po kliknutí na potvrzující tlačítko se načte stránka s rozšířeným vyhledáváním. Pokud vyhledáváme časopisy, nalezené položky nyní můžeme dále filtrovat přes pole:

- *SUBJECTS* - možnost výběru z 527 oblastí;

- *LANGUAGES* - volba jazyka;
- *LICENSES* - volba licence;
- *PUBLISHERS, PUBLISHERS' COUNTRIES* - možnost výběru vydavatele a jeho země;
- *PEER REVIEW TYPES* - výběr typu recenze;
- *DATE ADDED* - volba data přidání publikace, lze vybírat pouze celé roky.

Rozšířené vyhledávání článků má pole:

- *SUBJECTS* - možnost výběru z 516 oblastí;
- *JOURNALS* - výběr časopisu, ve kterém se článek nachází;
- *JOURNAL HAS THE SEAL* - časopis splňuje všechna kritéria platformy DOAJ a je oceněn jako jeden z nejlepších zdrojů - volba z možností ano/ne;
- *YEAR OF PUBLICATION* - volba data přidání publikace, lze vybírat pouze celé roky, je ekvivalentní k "DATE ADDED"u vyhledávání časopisů viz výše.

4.1.4 Europeana

Web Europeana je full-textová databáze se záznamy digitálních evropských kulturních dědictví. Záměrem je poskytovat data pro inspiraci a informovat uživatele o kultuře a historii. Web sbírá média od více než 4 tisíc různých evropských institucí. Dohromady je zde téměř 57 milionů textů, obrázků, hudby, videí a 3D vizualizací [30].

Kromě vyhledávání klíčových slov přes všechny atributy můžeme použít rozšířené filtry. Na stránce pro hledání se v pravé části nachází níže vypsaná pole:

- *THEME* - výběr oblasti jako je umění, móda, hudba ap.;
- *TYPE OF MEDIA* - volba média - obrázek/text/hudba/video/3D;
- *CAN I USE THIS?* - možnost filtrovat média, které lze dále použít bez dalších podmínek;
- *PROVIDING COUNTRY* - výběr příslušné země;

- *LANGUAGE* - jazyk média;
- *PROVIDING INSTITUTION* - výběr instituce, která médium poskytuje;
- *COLOUR* - výběr barvy;
- *IMAGE ORIENTATION* - volba orientace obrázku - "Portrait"/"Landscape";
- *IMAGE SIZE* - výběr rozlišení obrázku;
- *FILE FORMAT* - formát média, například MP4 nebo JPEG;
- *RIGHTS STATEMENT* - možnost výběru médií, na které se nevztahují omezení autorských práv;

Dále je zde možnost přidat k vyhledávání klíčová slova pro daný atribut. Můžeme tedy například zadat klíčová slova pro název média a zároveň jiná klíčová slova pro popis.

4.1.5 ScienceDirect

ScienceDirect spadá pod vydavatelství Elsevier a umožňuje full-textové vyhledávání v článcích, časopisech a knihách. Momentálně databáze zahrnuje přes 4,7 tisíce časopisů a přes 34,3 tisíce knih [29].

Na platformě ScienceDirect můžeme najít čtyři různé vyhledávací stránky. První je hlavní stránka aplikace, kde jsou pouze tři vyhledávací pole:

- *Find articles with these terms* - možnost zadání klíčových slov, které se nachází v článku;
- *In this journal or book title* - zadání názvu časopisu nebo knihy;
- *Author(s)* - zadání autora/ů.

Po vyplnění těchto tří polí se dostaneme na stránku s výsledky, kde po levé straně najdeme sloupec s poli pro upřesnění vyhledávání:

- *Years* - výběr roku vydání;
- *Article type* - určení typu článku, například "Product review" nebo "Book chapters";
- *Publication title* - název časopisu nebo knihy;
- *Access type* - přístupnost, můžeme zvolit "Open access & Open archive".

Další možností je vyhledávání po kliknutí na "Journals & Books" v hlavní stránce. Zde se nachází následující prvky pro filtrování výsledků:

- *Filter by journal or book title* - zadání názvu časopisu nebo knihy;
- *Domain* - skupina zaměření (např. "Computer Science");
- *Subdomain* - zaměření konkrétně (např. "Computer Vision and Pattern Recognition"), výběr závisí na zvoleném "Domain";
- *Publication type* - typ publikace;
- *Journal status* - možnost volby, jestli časopis přijímá další příspěvky;
- *Access type* - přístupnost, můžeme vybrat, jestli chceme pouze otevřený nebo i předplacený obsah.

Pro ScienceDirect je samozřejmostí i rozšířené vyhledávání. Ikona otazníku umožňuje proklik na nápovědu včetně vysvětlení všech pravidel zadávání klíčových slov. Tato čtvrtá možnost zahrnuje následující:

- *Find articles with these terms* - možnost zadání klíčových slov, které se nachází v článku;
- *In this journal or book title* - zadání názvu časopisu nebo knihy;
- *Year(s)* - zadání konkrétního roku nebo rozsahu let;
- *Author(s)* - zadání autora/ů;
- *Author affiliation* - příslušnost autora (např. Harvard);
- *Volume(s)* - počet svazků, možnost zadání jednoho čísla nebo rozsahu;
- *Issue(s)* - počet vydání, možnost zadání jednoho čísla nebo rozsahu;
- *Page(s)* - počet stran, možnost zadání jednoho čísla nebo rozsahu;
- *Title, abstract or author-specified keywords* - zadání klíčových slov, které budou obsaženy v názvu, abstraktu nebo v detailu autora;
- *Title* - název publikace;
- *References* - reference publikace;
- *ISSN or ISBN* - kód ISSN nebo ISBN publikace.

4.1.6 SCOAP³

SCOAP³ se prezentuje jako úložiště článků o jaderné fyzice. Články lze stahovat a dále šířit díky licenci CC-BY [26].

Nejprve je součástí vyhledávání pole pro zadání klíčových slov a pole pro výběr magazínu. Pokud některý magazín vybereme, objeví se stránka s dalšími vyhledávacími prvky a odpovídajícími články, které vyhovují vyhledávání. Na této stránce jsou následující vyhledávací pole:

- *Search* - základní vyhledávání klíčových slov přes všechny atributy;
- *Year* - umožňuje vybrat jeden konkrétní rok nebo rozsah let, kdy byl článek publikován;
- *Country / Region / Territory* - pole pro výběr geografického území, odkud pochází autor;
- *Journal* - specifikace, ve kterých magazínech chceme články vyhledávat.

4.2 Zhodnocení vyhledávání u zkoumaných nástrojů

V každém z nástrojů je součástí vyhledávání pole pro klíčová slova. V některých nejen přes všechny atributy, ale je možné zvolit atribut konkrétně, popřípadě jich lze zadat více najednou a klíčová slova mezi atributy rozdělit. Na platformách ACM Digital Library a DOAJ je možné ve frázích používat logické operátory jako AND a OR, v ACM Digital Library navíc i operátor NOT. V nástroji Europeana je umožněna alespoň negace příznakem "does not contain" (možný ekvivalent k operátoru NOT). ACM Digital Library a ScienceDirect obsahují obsáhlou nápovědu k filtrování a tvorbě frází klíčových slov.

V nástrojích ACM Digital Library, DOAJ, Europeana a ScienceDirect se nachází tlačítko k resetu všech navolených filtrů. Některá pole pro filtrování obsahují mnoho možností k výběru a u nástrojů se liší jejich seřazení. Některé aplikace je ukazují v pořadí podle zastoupení výsledků (včetně čísla počtu záznamů v závorce za možností) - výhodou je, že jako uživatel dopředu vím, kolik záznamů mohu vyhledat, nevýhodou je, že absence abecedního seřazení zneprůjemňuje vyhledání konkrétní možnosti. U nástrojů DOAJ a Europeana oceňuji filtrování v možnostech vybraných parametrů přes textové vyhledávání. Pokud tedy mám možnost filtrovat podle státu (například

země, ze které pochází autor), mohu do textového vyhledávání napsat část názvu státu a možnosti se automaticky zredukuje na ty, které část názvu obsahují. Uživatelsky přívětivá by byla možnost státu vyhledávat pouze přes dvoumístný alfabetský kód, jelikož odlišné volby jako "Czechia" a "Czech Republic" mají kód státu shodný.

4.3 Vizualizace výsledků u zkoumaných nástrojů

Ve všech zkoumaných aplikacích byly výsledky vyhledávání vizualizovány podle stejného vzoru. Vždy je primární název, typicky jde o titulek publikace, poté jsou zobrazeny sekundární údaje. Pro přehlednost mezi všemi výsledky je sekundárních údajů méně, popřípadě jsou zkrácené. Součástí přílohy (10.3) jsou obrázky, jak náhodné výsledky vyhledávání z jednotlivých nástrojů vypadají.

Konkrétní výsledek jde rozkliknout pro zobrazení všech známých detailů. Ve zkoumaných nástrojích se v detailu objevoval celý text abstraktu, bližší údaje o autorech nebo počet stažení dat záznamu. V ACM Digital Library lze zobrazit reference, citace a popřípadě komentáře uživatelů. V nástroji Cambridge University Press jsou například dostupné grafy počtu návštěvníků záznamu a odkazy na související obsah. Detail magazínu v aplikaci DOAJ umožňuje zobrazit (kromě základních údajů) všechny články, které do magazínu patří. Zajímavá vizualizace v detailu záznamu je u stránky Europeana. Tam najdeme souřadnice a mapu s vyznačenou lokací, odkud daný záznam pochází. ScienceDirect v detailu záznamu zobrazuje citace a také celý obsah článku v PDF.

5 Vizualizace dat evropských projektů

Po analýze dat evropských projektů, respektive organizací a reportů, už víme, jaké hodnoty datasety obsahují. Nyní potřebujeme najít scénáře, jak v praxi může vyhledávání v datech probíhat, abychom později navrhli aplikaci, která bude uživatelsky přívětivá a umožní uživateli vhodně výsledná data vizualizovat.

Základem by mělo být vyhledávání v samotných entitách - projektech, organizacích a reportech. Z analýzy dat připomeneme atributy jednotlivých entit.

Projekt má atributy: *id, status, topics, frameworkProgramme, title, objective, startDate, endDate, projectUrl, totalCost, ecMaxContribution*.

Organizace má atributy: *id, name, activityType, endOfParticipation, ecContribution, country, city, organizationUrl*.

Report má atributy: *id, language, title, teaser, summary, country, projectID, relatedFile, reportUrl*.

Primárním použitím, je zkoumání projektů s vybraným zaměřením a přehled organizací, které na něm spolupracují. Nejdůležitějším atributem projektu jsou tedy jeho zaměření - *topics*. Dále by nás u projektu mohl zajímat *status* a popřípadě i rámcový program, do kterého spadá - *frameworkProgramme*. U projektů jsou důležité atributy *title* a *objective*, ve kterých už bude vyhledávání komplikovanější, proto bychom přes tyto dva atributy mohli hledat pomocí klíčových slov.

U samotných organizací je podstatný název. Vzhledem k počtu možných záznamů, rozmanitosti a délce názvů bude nejlepší i zde vyhledávat přes klíčová slova. V praxi by bylo užitečné filtrovat organizace geograficky, tzn. podle atributu *country* a *city*, popřípadě i podle typu organizace - *activityType*.

Co se týče reportů, po bližším zkoumání dat jsem nenarazila na report, který by nebyl v anglickém jazyce. Stále je ale možnost vyhledávat přes kód země (*country*), ze které report pochází. Pro atributy *title, teaser* a *summary* bude vzhledem k jejich rozsahu opět nejpraktičtější textové vyhledávání přes klíčová slova.

Kromě vyhledávání v samotných entitách budeme potřebovat i kombinované vyhledávání. Musíme tedy umožnit:

- vyhledávání projektů podle kritérií organizací;

- vyhledávání projektů podle kritérií reportů;
- vyhledávání organizací podle kritérií projektů;
- vyhledávání reportů podle kritérií projektů.

Uživatel musí mít možnost vyhledávat podle klíčových slov přes všechny entity, ale i jednotlivě.

Výše uvedené poznatky by měly vést k pokrytí následujících scénářů vyhledávání:

- *Základní*
 - všechny projekty vybraných kritérií (zaměření/status/rámcový program) + odpovídající klíčová slova v názvu/popisku projektu;
 - všechny organizace vybraných kritérií (typ/stát) + odpovídající klíčová slova v názvu/měště organizace;
 - všechny reporty vybraných kritérií (stát) + odpovídající klíčová slova v názvu/upoutávce/souhrnu reportu.
- *Kombinované*
 - všechny projekty na nichž pracovala alespoň jedna organizace splňující kritéria organizací;
 - všechny projekty u kterých je znám alespoň jeden report splňující kritéria reportů;
 - všechny organizace u nichž platí, že každá z nich pracovala na alespoň jednom projektu splňující kritéria projektů;
 - všechny reporty, které se týkají projektu splňující kritéria projektů.

U kombinovaného vyhledávání lze zadat klíčová slova pro všechny entity nebo jen pro jednu vybranou.

Ukázka konkrétních scénářů vyhledávání:

- všechny projekty se zaměřením "Science in Society", rámcového programu FP7, statusu "SIGNED", na kterých spolupracovala alespoň jedna organizace z České republiky, která patří do typu "Research Organisations" a zároveň je u projektu znám alespoň jeden report z České republiky;

- všechny organizace z České republiky typu "Research Organisations", které spolupracovaly alespoň na jednom projektu se zaměřením "Science in Society", rámcového programu FP7 a statusu "SIGNED";
- všechny reporty, které pochází z České republiky a týkají se projektu se zaměřením "Science in Society", rámcového programu FP7 a statusu "SIGNED";

Výsledek vyhledávání může obsahovat hodně záznamů, proto by bylo vhodné zobrazit jen stručné údaje, u kterých budeme moci rozkliknout detail s kompletními informacemi. Stejný princip byl u všech zkoumaných podobných nástrojů viz kapitola 4.3.

Detail projektu by měl kromě všech atributů z databáze zobrazovat i list organizací, které se na projektu podílely a seznam všech reportů, které s projektem souvisí. U detailu organizace by mělo být vedle známých atributů viditelné, na jakých projektech spolupracovaly včetně jejich role, jestli šlo o koordinátora nebo pouze účastníka. Dále by bylo přínosné, kdyby detail organizace vizualizoval seznam dalších organizací, se kterými byla vybraná instituce v součinnosti a počet projektů, na kterých společně pracovaly. Budeme schopni vidět, se kterou organizací dochází k nejčastější spolupráci. Detail reportu by měl kromě známých atributů zobrazovat záznam projektu, kterého se report týká. Nad seznamy souvisejících objektů zobrazujeme, kolik položek obsahují. Například v detailu organizací vidíme nad seznamem projektů jejich počet.

6 Návrh aplikace

Aplikace by měla mít dvě hlavní části:

- přehled importovaných datasetů;
- vyhledávání.

Obecně se budeme snažit o jednoduše a čistě vypadající aplikaci, která bude intuitivní na použití a uživatelsky přívětivá. Během implementace a návrhů designu se budeme snažit využít poznatky z analýzy vybraných bibliografických nástrojů z kapitol 4.2 a 4.3.

6.1 Sekce Přehled importovaných datasetů

Tato sekce by měla umožňovat snadný import nového datasetu. Umožníme tedy import dat ve formátu CSV stažených ze zdroje CORDIS. Nově vložený import by měl být viditelný v přehledu importů. Součástí přehledu, by mělo být datum, kdy k importu došlo, typ entit, které import obsahuje a kolik záznamů pod něj spadá. Z datasetů budeme ukládat pouze potřebná data (viz kapitoly 3.2 a 5). Uživatel by měl mít možnost importované datasety z aplikace smazat, do přehledu tedy můžeme přidat ikonu pro smazání odpovídajících záznamů.

6.2 Sekce Vyhledávání

Po importu dat do aplikace v nich budeme chtít vyhledávat. Vytvoříme tedy stránku s filtrováním a textovým polem pro klíčová slova. Při vytváření jednotlivých parametrů vyhledávání budeme vycházet z kapitoly 5. Přidáme možnost všechny vyhledávací pole resetovat. U zobrazených výsledků vyhledávání budeme moci rozkliknout stránku s detailními informacemi, opět návrh vychází z kapitoly 5.

7 Implementace

Aplikace byla vyvíjena na operačním systému Microsoft Windows 10 za použití vývojového prostředí IntelliJ IDEA. Jedná se o webovou server-side aplikaci. Serverová část, která řídí logiku softwaru, komunikuje s databází a zpracovává požadavky, je naprogramovaná v jazyce Java s použitím nástroje Maven, pro správu projektu a snadné definování závislostí. Klient běžící ve webovém prohlížeči posílá požadavky serverové části a zobrazuje uživatelské rozhraní pro interakci s uživatelem. Využité technologie v klientské části jsou především HTML, CSS a JavaScript. Během vývoje jsem klienta zobrazovala ve webovém prohlížeči Google Chrome. Ke správě databázových kontejnerů byla využita open-source platforma Docker s použitím relačního databázového systému PostgreSQL. Pro uchování historie změn v kódu a zajištění bezpečnosti proti ztrátě dat jsem soubory aplikace průběžně nahrávala do verzovacího systému Git (konkrétně GitHub).

Struktura souborů aplikace je rozsáhlá, proto jsem se rozhodla, dát obrázek struktury jako součást přílohy (10.9).

7.1 Další použité technologie

7.1.1 Spring Framework

Spring Framework je open-source technologie pro vývoj aplikací v jazyce Java. Poskytuje sadu nástrojů, knihoven a funkcí pro snadnější vytváření programu. Spring framework využívá *kontejnery* zodpovědné za vytváření a správu objektů [31].

7.1.2 JPA

Pro usnadnění komunikace s databází je použita technologie JPA (Java Persistence API). JPA poskytuje vrstvu nad relační databází, kontroluje persistenci a umožňuje číst data bez nutnosti psaní přímých dotazů do databáze. Psaní přímých dotazů je možné, ale stačí v dotazech používat atributy z třídy pro danou entitu. Navíc usnadňuje vytváření a změny databázových tabulek, protože umí převádět Java objektový model na relační databázový - tento přístup se nazývá objektově relační mapování (ORM) [32].

7.1.3 JSF

JSF (JavaServer Faces) je technologie pro vývoj webových uživatelských rozhraní. Díky JSF můžeme přistupovat k datům a logice na straně serveru a na klientské části je zobrazit. V HTML kódu můžeme používat JSF komponenty a díky nim snadno zajistit interakci mezi serverou a klientskou částí [33].

7.2 Spuštění aplikace

Spuštění probíhá za pomoci technologií Maven a Docker. Přesný postup spuštění samotné aplikace je popsán v uživatelské příručce, která je součástí přílohy (11).

Během spuštění se aplikace konfiguruje. Dochází k nastavování pravidel, jako například mapování stránek na přehlednější názvy bez přípon *.html*. Dále probíhá konfigurace databáze. V tuto chvíli se vytvoří všechny potřebné tabulky (pokud neexistují). Proběhne řada kontrol existence záznamů a pokud dané záznamy neexistují, vloží se do databáze data číselníků států a zaměření projektů. K tomuto importu využíváme soubory vygenerované nástrojem MySQL s explicitně vypsánými dotazy. Všechny potřebné soubory s příponou *.sql* se nachází ve složce *queries*. Ke kontrole existence využíváme SQL dotazy typu *COUNT* a při importu dotazy typu *INSERT*.

Zmíněné číselníky jsou dostupné ke stažení v repozitáři CORDIS. Problém nastal při pokusu importu zaměření projektů. Oba zkoumané rámcové programy (FP7 a H2020) mají vlastní číselníky a kódy zaměření se liší. Bylo proto nutné importovat oba datasey. Dohromady je záznamů zaměření několik tisíc a stažené datasey měly vysokou chybovost, kvůli nevhodně vloženým uvozovkám. Bohužel zde nebyla jiná možnost, než je všechny projít a ručně opravit.

7.3 Hlavní stránka

Hlavní stránka je vizuálně hodně minimalistická. Nalezneme zde hlavičku s názvem aplikace a možnost přepínání překladů mezi angličtinou a češtinou. Obě varianty dávají smysl, jelikož aplikace bude používána především česky mluvícími uživateli, ale veškerá data evropských projektů jsou v angličtině. Zvolení anglického jazyka proto může působit méně rušivě při používání aplikace. Další výhodou je použití překladových souborů. Všechny textové konstanty, které jsou v aplikaci viditelné, jsou přehledně v jednom souboru

a je triviální je upravit. Klientská část poté přebírá konstanty přes klíč v závislosti na aktuálně zvoleném jazyce.

Ve centrální části hlavní stránky se nachází dvě velká tlačítka, která nás přesměrují na stránku s přehledem importovaných datasetů, respektive stránku pro vyhledávání v datech.

7.4 Přehled importovaných datasetů

Vzhled sekce Přehled importovaných datasetů včetně níže popisovaných elementů jako jsou tlačítka a pop-upy je součástí uživatelské příručky (11.2).

Sekce *Přehled importovaných datasetů* obsahuje tlačítka s ikonou "+" pro import nového datasetu. Po kliknutí se otevře pop-up s poli pro výběr souboru ve formátu CSV a výběr typu obsahu. Vybrat můžeme jedu z možností Projekty/Organizace/Reporty. Po vyplnění polí můžeme kliknout na tlačítka "Importovat". V tu chvíli se na klientské části zobrazí loader pro indikaci načítání a na serverovou část se pošle soubor k importu a vybraný typ.

Metoda pro import nového datasetu se rozsáhlá. Na začátku vytvoříme model importu, který zahrnuje atributy jako název souboru, typ obsahu a datum, kdy k importu dochází. Dále se kontroluje hlavička CSV souboru. Ta musí obsahovat všechny povinné atributy podle vybraného typu. Pokud je hlavička v pořádku, dojde k procházení dalších řádků importovaného souboru. Jeden záznam může být na více řádků, proto jsme museli algoritmus upravit tak, aby načítal jen do chvíle, dokud se na další řádce nenachází následující záznam. Po načtení záznamu se kontroluje jeho obsah a snažíme se z dat vytvořit odpovídající model. Vytvořenému modelu nastavíme identifikátor importu. Jestliže záznam obsahuje všechna potřebná data, je přidán na list úspěšně načtených záznamů a algoritmus pokračuje, dokud neprojdeme celý soubor. Následně list záznamů uložíme do databáze. Pokud se jedná o soubor s organizacemi, ukládáme i list vazeb mezi organizacemi a projekty.

Znamená to tedy, že všechna data z načteného souboru obsahují identifikátor importu. Ten držíme ze dvou důvodů. Prvním důvodem je možnost odstranění importovaného datasetu - odstraníme všechny záznamy, u kterých se atribut id importu shoduje s identifikátorem odstraňovaného záznamu. Druhý důvod je poskytnutí informace o importu v detailu záznamu v sekci *Vyhledávání*. Pokud si tedy rozklikneme detail záznamu, najdeme zde informaci, z jakého importu pochází.

Během implementace jsem zjistila, že mají stažené datasety z platformy

CORDIS vysokou chybovost. Jednalo se například o nevhodně přidané uvozovky nebo nadbytečné znaky, které znemožňovaly parsování textového řetězce na číslo. Po analýze chyb jsem algoritmus rozšířila, aby s těmito chybami dokázal pracovat a záznam bylo možné uložit. Stále se však objevovaly záznamy, které uložit nešlo. Například v datasetu organizací rámcového programu FP7 bylo množství záznamů, ve kterých nebyl vyplněn atribut id, který je pro aplikaci zásadní. S vedoucím práce jsme se proto rozhodli aplikaci rozšířit o možnost chybné záznamy stáhnout, aby případně mohly být uživatelem opraveny.

Metoda pro import tedy obsahuje i vytváření souboru s chybnými záznamy. Nejprve se přidá hlavička z importovaného souboru a pokud se některý ze záznamů nepodaří převést na odpovídající model, zapíše se do souboru chybných záznamů.

Po zpracování celého importovaného souboru a uložení dat do databáze vrátíme na klientskou část model s výsledky. Na klientské části se pop-up s loaderem změní a uživatel je informovaný o počtu úspěšně uložených záznamů. Pokud se během importu některé záznamy nepodařilo uložit, uživatel vidí jejich počet a může si tlačítkem stáhnout soubor s chybnými záznamy.

Dále v sekci nalezneme tabulku importovaných datasetů, kde se v podstatě zobrazují záznamy z tabulky všech importů. Najdeme zde údaje jako název importovaného souboru, typ obsahu, datum importu a počet záznamů, které se úspěšně uložily. Na konci řádky záznamu najdeme ikonu pro smazání, jejíž princip je popsán výše.

7.5 Vyhledávání

Vzhled sekce Vyhledávání včetně níže popisovaných elementů jako jsou tlačítka a pole vyhledávání je součástí uživatelské příručky (11.3). V příloze (10.10) najdeme vizualizaci výsledků a detailů záznamů.

Obsah sekce *Vyhledávání* by se dal rozdělit do tří částí. Na začátku stránky se nachází pole pro zadání klíčových slov s možností výběru, jestli chceme klíčová slova vyhledávat ve všech entitách nebo pouze v konkrétní vybrané. Uživatelská příručka vysvětluje použití logických operátorů v textu klíčových slov. V levé části se nachází sloupec pro filtrování. Zbytek stránky je prostorem pro přehled výsledků vyhledávání.

Vyhledávání klíčových slov a filtrování je propojené. Po kliknutí na tlačítko "Filtrovat", se na serverou část pošle dotaz zahrnující jak klíčová slova, tak parametry filtrů. Funkcionalita vyhledávání vychází z kapitol 5 a 6.

Ve filtrovacím sloupci můžeme vybrat typy entit (Projekty, Organizace, Reporty). Pokud není žádný typ vybrán, automaticky vyhledáváme mezi všemi zmíněnými. Následují parametry filtrování, jež se vztahují pouze k projektům. Pro výběr zaměření projektů stačí zaškrtnout libovolné množství checkboxů s názvy. Vzhledem k vysokému počtu možností, je nad skupinou checkboxů textové pole, do kterého můžeme zadat část názvu zaměření a možnosti se zredukuje na ty, kde název zaměření obsahuje řetězec zadaného textu. Dále můžeme vybírat z rámcových programů a statusů.

Parametry filtrování pouze pro organizace zahrnují výběr typu, kde můžeme vybrat libovolné množství zaškrtnutím checkboxů. Pod výběrem typu najdeme skupinu checkboxů pro filtrování podle státu, ze kterého organizace pochází. Opět můžeme zvolit libovolné množství. Nad skupinou checkboxů se nachází textové pole pro redukci možností, funkcionality je stejná jako u redukce zaměření projektů viz výše. Zadání můžeme buď část názvu státu nebo pouze dvoumístný alfabetský kód.

K reportům se vztahuje pouze jeden filtrovací parametr a to pro výběr státu, který funguje stejně jako pro organizace.

Možnosti ve filtrovacích parametrech se načítají ze serverové části. Při popsání spuštění aplikace jsem zmiňovala, že se do databáze načítají dva číselníky - pro státy a zaměření projektů. Serverová část data z číselníků zpracuje a pošle na klientské rozhraní. Problém je, že uživateli zobrazujeme pouze název zaměření, ale identifikace je podle jejich kódu. Nastává situace, kdy se kód zaměření liší, ale název je stejný. Tím docházelo k duplicitním názvům v možnostech, proto se během zpracování záznamů na serveru záznamy sloučí a vrátíme list pouze jedinečných názvů. Při zpracování dat států pro organizace a reporty nevracíme pouze list se jménem státu, ale tvoříme řetězec složený ze jména státu a kódu, například "Czechia (CZ)".

Při zpracování možností rámcových programů vybíráme z databázové tabulky projektů pouze jedinečné názvy programů. Stejným způsobem vytváříme list statusů. Zde bylo nutné položky listu sjednotit, protože oba rámcové programy používají jiná označení. Bez úpravy by měl uživatel na výběr dvě různé možnosti se stejným významem. Co se týče možností typů organizace, v databázi se pod daným atributem nachází jen pro uživatele nepochopitelné zkratky. Změnila jsem tedy položky listu a zkratky přepsala na celé názvy dohledané v repozitáři CORDIS.

Na konci sloupce filtrování najdeme čtyři checkboxy pro využití kombinovaného vyhledávání:

- Projekty podle kritérií organizací;
- Projekty podle kritérií reportů;

- Organizace podle kritérií projektů;
- Reporty podle kritérií projektů.

Kliknutím na tlačítko "Filtrovat" vytvoříme v JavaScriptu ze všech filtrovacích parametrů JSON objekt a posíláme ho jako textový řetězec serverové části. Metoda na serveru řetězec převede zpět na JSON a získá hodnoty, které uživatel pro filtrování vybral. Poté probíhá vyhledávání pro všechny nebo jen vybrané typy entit. Začínáme vyhledáváním všech id záznamů, které odpovídají klíčovým slovům. Vytváříme dotazy nad databází, zda se klíčová slova vyskytují v datech vybraných atributů. Získané identifikátory následně dále redukuje podle vybraných filtrů. Když získáme list id záznamů, které vyhovují klíčovým slovům a všem zadaným filtrům, tak vytvoříme list modelů. Tyto modely obsahují pouze základní údaje pro zobrazení ve výsledcích vyhledávání. Hotový list modelů vrátíme klientské části, která výsledná data vizualizuje.

Základní údaje projektu: název, zaměření, rámcový program, status, první tři řádky popisu.

Základní údaje organizace: název, typ, stát.

Základní údaje reportu: nadpis, stát, první tři řádky upoutávky.

Na výsledek vyhledávání můžeme kliknout, čímž zavoláme JavaScript metodu, která vezme identifikátor rozkliknutého výsledku a otevře jeho detail v novém okně prohlížeče. Serverová část nyní v databázi najde záznam podle id a vytvoří kompletní model.

Projekty kromě základních údajů v detailu obsahují: celkové náklady [€], maximální příspěvek [€], datum začátku, datum konce, URL, celý popis, id záznamu v databázi, informace o importu, list organizací včetně jejich rolí na projektu, list reportů.

Organizace kromě základních údajů v detailu obsahují: město, URL, id záznamu v databázi, informace o importu, list projektů včetně jejich role, list organizací včetně počtu projektů vzájemné spolupráce.

Reporty kromě základních údajů v detailu obsahují: jazyk, URL, celou upoutávku, souhrn, id záznamu v databázi, informace o importu, projekt.

8 Testování

Během implementace byla back-endová část aplikace testovaná za pomoci ladící konzole samotného vývojového prostředí IntelliJ IDEA a front-endová část za pomoci vývojářského nástroje uvnitř prohlížeče Google Chrome.

8.1 Test uživatelské přívětivosti

Aplikace byla předvedena třem respondentům, se kterými jsem funkčnost osobně testovala na svém vlastním přenosném zařízení. Po krátkém vysvětlení, čeho se aplikace týká, jsem je nechala s aplikací samostatně pracovat. Zkoumala jsem intuitivnost a uživatelskou přívětivost vytvořeného nástroje. Respondenti si zkusili importovat testovací soubor (pro rychlejší import a nerozhození databáze nevhodnými daty), který po importu mohli zkusit smazat. Poté zkoušeli filtrovat data v sekci Vyhledávání. Zajímaly mě především negativní aspekty, abych na základě zpětné vazby mohla aplikaci opravit/vylepšit.

Jejich ohlasy byly následující:

- pochvala přehlednosti úvodní stránky a možnosti přepínat mezi češtinou a angličtinou;
- příjemný celkový vzhled webu a interaktivních prvků, jako je změna vzhledu elementů při najetí myši;
- příliš nezajímavá hlavička webu s názvem aplikace;
- špatně fungující tooltip u tlačítka pro návrat na hlavní stránku;
- čistý vzhled tabulky přehledu importovaných datasetů;
- málo přehledný sloupec pro filtrování výsledků;
- složité vybírání zaměření projektu;
- neintuitivní volby pro kombinované vyhledávání;
- absence loaderu pro kliknutí na tlačítka filtrovat, není jasné, jestli se aplikace zasekla, nebo se načítají výsledky;
- dobrá přehlednost v detailech entit.

8.1.1 Opravy po získání zpětné vazby

Po získání zpětné vazby a pozitivních ohlasů na vzhled aplikace, jsem se rozhodla nedělat žádné zásadní designové změny. Souhlasím, že hlavička stránky by mohla pro uživatele vypadat atraktivněji, pokud bych zde použila místo ikony nějaký obrázek. Díky testerům jsem zjistila, že je u jednoho tlačítka špatně fungující tooltip. V průběhu implementace jsem se rozhodla tooltipy odstranit a konkrétně u tohoto tlačítka jsem zapomněla smazat příslušnou část kódu. Tento bod tedy považuji za opravený.

Další výhrady byly ke sloupci pro filtrování. Rozhodla jsem se obalit filtrování pro jednotlivé entity do rámečků s nadpisem, aby se přehlednost zvýšila. Rovněž jsem změnila popisky pro možnosti kombinovaného vyhledávání. Princip kombinovaného vyhledávání je celkově složitější a změna popisků nepomůže natolik, aby bylo uživateli jasné, jak tyto možnosti fungují, proto bude nejlepší odkázat na uživatelskou příručku. Posledním bodem byla absence loaderu. Loader jsem měla již implementovaný v pop-upu pro import nového datasetu. Přenesla jsem tedy část kódu a přidala stejně vypadající element i do výsledků vyhledávání. Po kliku na tlačítko "Filtrovat" se na stránce objeví indikátor, že se výsledky teprve načítají.

Negativní ohlasy byly na složité filtrování mezi zaměřeními projektu. S tím souhlasím, ale bohužel není jiná možnost, jak zaměření více sloučit, aby bylo možností méně a volba tak byla přehlednější. Zde jsem provedla pouze malou změnu, a to abecední seřazení, které jeden z respondentů navrhl jako možné vylepšení.

8.2 Test rychlosti importu datasetů

V následující tabulce je přehled, jak dlouho trvalo importovat všechny záznamy datasetů. Čas zahrnuje čtení ze souboru formátu CSV, kontroly, vytvoření odpovídajících modelů, samotné ukládání do databáze a případné zapisování chybných zápisů do souboru pro pozdější stažení uživatelem. Test byl proveden na zařízení HP - notebooku s 16 GB RAM, procesorem AMD Ryzen 7 4700U 2,0 GHz, s operačním systémem Windows 10 Home 64-bit.

Tabulka 8.1: Přehled doby trvání importů.

Název souboru	Velikost	Doba trvání	Počet záznamů	Počet uložených záznamů
cordis-fp7projects.csv	58,4 MB	1,04 min	25 778	25 767
cordis-h2020projects.csv	65,2 MB	1,24 min	30 562	30 562
cordis-fp7organizations.csv	51,0 MB	7,19 min	144 111	19 196
cordis-h2020organizations.csv	41,4 MB	30,31 min	138 610	34 999
cordis-fp7reports.csv	415,1 MB	13,31 min	21 875	21 875
cordis-h2020reports.csv	110,7 MB	1,09 min	17 251	17 250

Z tabulky je zřejmá vyšší doba zpracování dat organizací. Důvodem je duplicita záznamů s rozdílem kontaktní osoby - již rozebráno v kapitole 3.2. Proto je zde i velký rozdíl mezi počtem záznamů z datasetu a počtem uložených záznamů do databáze. Při ukládání organizací také musíme vzít v potaz, že se současně ukládají data do vazební tabulky pro vztah organizací a projektů. Déle trvalo i zpracování reportů z rámcového programu FP7, kde je příčinou velikost dat. Reporty programu FP7 ve svém souhrnu mnohdy obsahují celý text obsahu reportu, zatímco reporty programu H2020 obsahují jen zkrácenou část.

8.3 Test přesnosti zobrazených dat

Nástroj CORDIS umožňuje vyhledávání v projektech a reportech, není možné vyhledávat přímo organizace. Nyní zkusíme vyhledat náhodný projekt a report na stránce CORDIS a poté zkusíme odpovídající záznamy najít v naší aplikaci. Výsledky z obou aplikací porovnáme.

Tabulka 8.2: Porovnání projektů

	Projekt v CORDIS	Projekt v naší aplikaci
Název	FLEXIBLE PRODUCTION TECHNOLOGIES AND EQUIPMENT BASED ON ATMOSPHERIC PRESSURE PLASMA PROCESSING FOR 3D NANO STRUCTURED SURFACES	FLEXIBLE PRODUCTION TECHNOLOGIES AND EQUIPMENT BASED ON ATMOSPHERIC PRESSURE PLASMA PROCESSING FOR 3D NANO STRUCTURED SURFACES
Zaměření	NMP-2007-3.5-1 - Processes and Equipment for High Quality Industrial Production of 3-Dimensional Nanosurfaces	Processes and Equipment for High Quality Industrial Production of 3-Dimensional Nanosurfaces (NMP-2007-3.5-1)
Program	FP7	FP7
Status	Closed project	SIGNED
Celkové náklady [€]	10 447 885,73	10447885.73
Maximální příspěvek [€]	7 400 000,00	7400000.0
Datum začátku	1 June 2008	01. 06. 2008
Datum konce	30 November 2012	30. 11. 2012
URL	http://www.n2p-project.eu/	http://www.n2p-project.eu/
Popis	Outstanding progress has...	Outstanding progress has...
Počet účastněných organizací	22	22
Název odpovídajícího reportu	Final Report Summary - N2P (FLEXIBLE PRODUCTION TECHNOLOGIES AND EQUIPMENT BASED ON ATMOSPHERIC PRESSURE PLASMA PROCESSING FOR 3D NANO STRUCTURED SURFACES)	Final Report Summary - N2P (FLEXIBLE PRODUCTION TECHNOLOGIES AND EQUIPMENT BASED ON ATMOSPHERIC PRESSURE PLASMA PROCESSING FOR 3D NANO STRUCTURED SURFACES)

Z tabulky je patrné, že je většina atributů shodná, některé mají pouze jiné formátování. Odlišné hodnoty byly pro status projektu. Moje aplikace neoznačila projekt jako uzavřený, což může být způsobeno neaktuálními daty v aplikaci. Stránka v repozitáři CORDIS má uvedené datum poslední změny 10. března 2023. Je tedy možné, že od doby, kdy jsem datasety stahovala, mohlo dojít v záznamu k aktualizaci.

Tabulka 8.3: Porovnání reportů

	Report v CORDIS	Report v naší aplikaci
Název	Final Report Summary - N2P (FLEXIBLE PRODUCTION TECHNOLOGIES AND EQUIPMENT BASED ON ATMOSPHERIC PRESSURE PLASMA PROCESSING FOR 3D NANO STRUCTURED SURFACES)	Final Report Summary - N2P (FLEXIBLE PRODUCTION TECHNOLOGIES AND EQUIPMENT BASED ON ATMOSPHERIC PRESSURE PLASMA PROCESSING FOR 3D NANO STRUCTURED SURFACES)
Stát	Germany	Germany (DE)
Jazyk	-	en
URL	-	-
Teaser	-	Outstanding progress has...
Obsah	Outstanding progress has...	Outstanding progress has...
Název odpovídajícího projektu	FLEXIBLE PRODUCTION TECHNOLOGIES AND EQUIPMENT BASED ON ATMOSPHERIC PRESSURE PLASMA PROCESSING FOR 3D NANO STRUCTURED SURFACES	FLEXIBLE PRODUCTION TECHNOLOGIES AND EQUIPMENT BASED ON ATMOSPHERIC PRESSURE PLASMA PROCESSING FOR 3D NANO STRUCTURED SURFACES

Atributy zkoumaných reportů jsou shodné, jen moje aplikace zobrazuje *Jazyk* a *Teaser* navíc.

V kapitole Implementace (7) jsem popisovala problematiku importu chybných záznamů. Při pokusu o importování datasetu reportů rámcového programu H2020 jeden záznam nebylo možné uložit z důvodu nevhodně vložených uvozovek. Dohledala jsem tento záznam a nyní zkusíme porovnat projekt, kterého se report týká.

Tabulka 8.4: Porovnání projektů s neimportovaným reportem

	Projekt v CORDIS	Projekt v naší aplikaci
Název	A REVOLUTIONARY SAFE, QUIET, LONG RANGE, PERSONAL VERTICAL TAKEOFF AND LANDING (VTOL) DEVICE	A REVOLUTIONARY SAFE, QUIET, LONG RANGE, PERSONAL VERTICAL TAKEOFF AND LANDING (VTOL) DEVICE
Zaměření	EIC-SMEInst-2018-2020 - SME instrument	SME instrument (EIC-SMEInst-2018-2020)
Program	H2020	H2020
Status	Closed project	CLOSED
Celkové náklady [€]	71 429,00	71429.0
Maximální příspěvek [€]	50 000,00	50000.0
Datum začátku	1 August 2019	01. 08. 2019
Datum konce	30 November 2019	30. 11. 2019
URL	https://www.audaxaero.com/	https://www.audaxaero.com/
Popis	World urban population is...	World urban population is...
Počet účastněných organizací	1	1
Název odpovídajícího reportu	Periodic Reporting for period 1 - Audax Aero (A REVOLUTIONARY SAFE, QUIET, LONG RANGE, PERSONAL VERTICAL TAKEOFF AND LANDING (VTOL) DEVICE)	-

V tabulce si můžeme všimnout, že data projektů odpovídají, ale naše aplikace nemá v databázi uložený záznam reportu, proto se v detailu projektu nezobrazí.

9 Závěr

V této práci jsme nejprve zjistili, jak fungují dotace Evropské unie pro výzkum a inovace. Seznámili jsme se s nejdůležitějšími rámcovými programy a jejich pilíři. Repozitáře dotovaných projektů, organizací a reportů jsou dostupné na platformě CORDIS, kde jsme si tyto data mohli stáhnout ve formátu CSV.

Získaná data rámcových programů FP7 a H2020 jsme poté analyzovali. Zjistili jsme jaká je jejich struktura, jaké obsahují atributy a jaký je jejich význam. Také jsme mezi záznamy projektů, organizací a reportů našli vzájemné vazby. U všech atributů datasetů jsme zhodnotili důležitost a navrhli jsme dostačující soubory atributů pro další použití v naší aplikaci.

V další kapitole jsme se seznámili s metodami vyhledávání a vizualizací dat v bibliografických a podobných nástrojích. U několika vybraných nástrojů jsme zkoumali, jak funguje rozšířené vyhledávání včetně vyhledávání klíčových slov a jak vypadá vizualizace výsledků. Zhodnotili jsme jejich uživatelskou přívětivost, abychom později mohli sami navrhnout aplikaci, která bude vhodně vyhledávat a vizualizovat data.

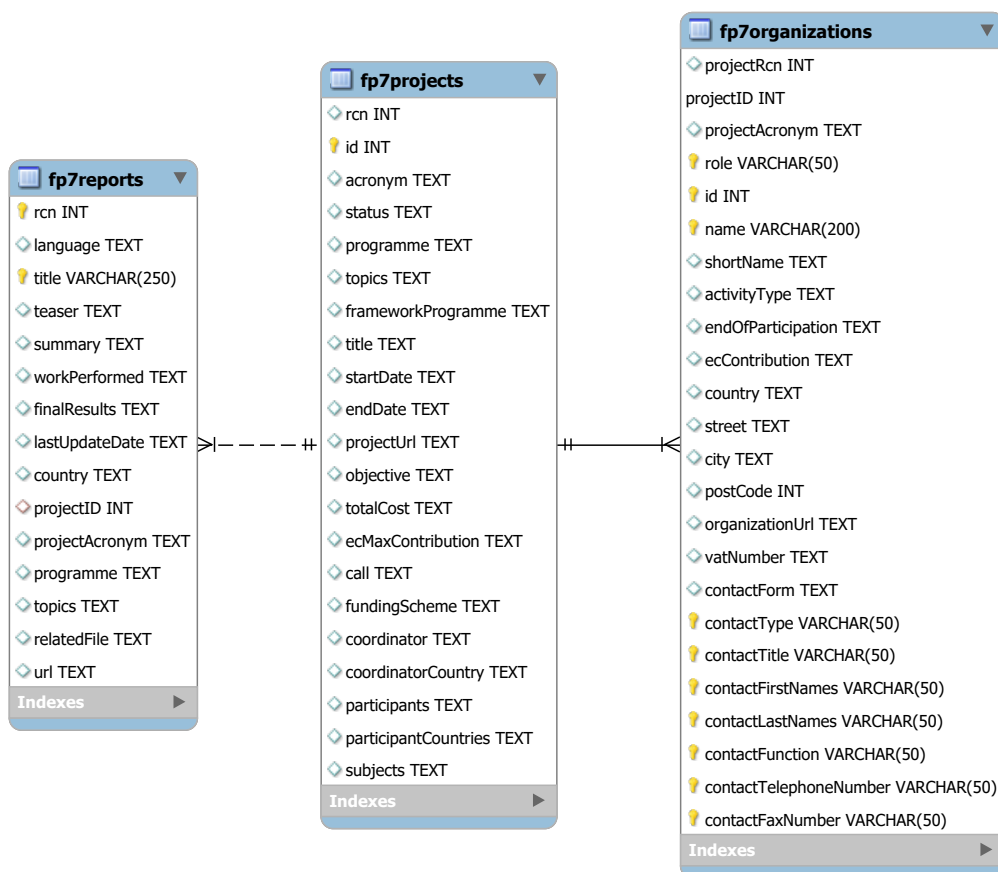
Dále jsme diskutovali nad možnými scénáři vyhledávání v datech evropských projektů, organizací a reportů. Navrhli jsme základní vyhledávání v jednotlivých entitách i kombinované vyhledávání, které umožní složitější dotazy. Kvůli předpokládanému velkému množství záznamů jsme se rozhodli ve výsledcích vyhledávání zobrazit pouze základní údaje. Detail záznamu bude viditelný až po rozkliknutí, čímž zajistíme přehlednost a zároveň možný přístup ke kompletním informacím.

Po vyhodnocení scénářů vyhledávání dat jsme přešli k návrhu aplikace. Představili jsme její dvě hlavní části a jaká by měla být funkcionality. Základem je možnost importovat datasety (typicky z repozitáře CORDIS) a v uložených datech vyhledávat, vizualizovat výsledky a zobrazit vztahy mezi záznamy. Poté jsme popsali samotnou implementaci, zvolené technologie a strukturu kódu.

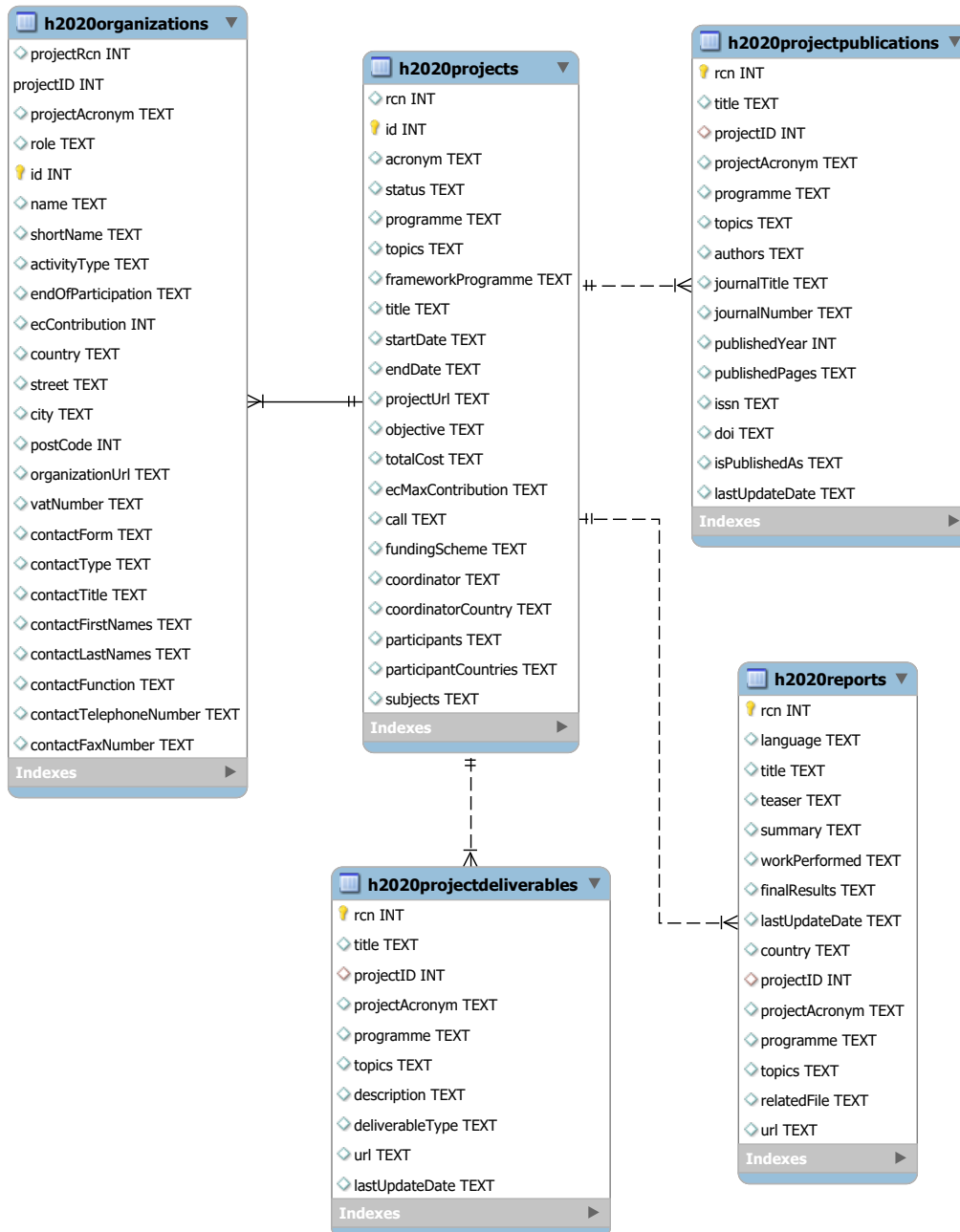
Poslední kapitola se věnovala testování. Aplikaci jsme nechali otestovat respondenty a po jejich zpětné vazbě jsme provedli opravy. Otestovali jsme rychlost importu datasetů na mém zařízení a v závěru jsme zkontrolovali přesnost zobrazených dat.

10 Příloha A

Obrázek 10.1: ER model programu FP7



Obrázek 10.2: ER model programu H2020



Obrázek 10.3: Výsledek vyhledávání nástroje ACM Digital Library

POSTER
October 2020

[Teaching primary school children about computer viruses: preliminary results of an intervention study](#)

Katerína Tsarava, Manuel Ninaus, Tereza Hannemann, Kristína Volná, + 2

WIPSCe '20: Proceedings of the 15th Workshop on Primary and Secondary Computing Education • October 2020, Article No.: 21, pp 1–2 • <https://doi.org/10.1145/3421590.3421660>

Children increasingly use computing devices. However, it is unclear whether they have basic knowledge of security-related issues such as computer viruses and, in case they do not, what they can learn about them. In this study, we first examined ...

1 81 Highlights

” + •

Obrázek 10.4: Výsledek vyhledávání nástroje Cambridge University Press

[Observed half-lives of \$^3\text{H}\$, \$^{90}\text{Sr}\$ and \$^{137}\text{Cs}\$ in hydrosphere in the Vltava river basin in vicinity of NPP Temelín \(Czechia\)](#)

J.C. Barescut, J.C. Gariel, J.M. Péres, E. J. Hanslík, V. Jedináková-Křížová, M. Brtvová, D. Ivanovová, E. Kalinová, B. Sedlářová, J. Svobodová, P. Šimonek, H. Tomášková

Journal: *Radioprotection* / Volume 40 / Issue S1 / May 2005

Published online by Cambridge University Press: 17 June 2005, pp. S759-S764

Print publication: May 2005

Article [Get access](#) [Export citation](#)

[View abstract](#)

Obrázek 10.5: Výsledek vyhledávání nástroje DOAJ

Psychology and its Contexts

Psychologie a její Kontexty

Published by *University of Ostrava* in **Czechia** Accepts manuscripts in **Czech**

Journal subjects
Philosophy. Psychology. Religion: Psychology

Last updated on 23 Apr 2020
[Website](#)

No charges
[Author retains all rights](#)
[CC BY](#)

Obrázek 10.6: Výsledek vyhledávání nástroje Europeana



Obrázek 10.7: Výsledek vyhledávání nástroje ScienceDirect

Research article ● Open access

71 **Key technologies of artificial intelligence in electric power customer service**
Global Energy Interconnection, December 2021, ...
Xingping Wu, Xusheng Liu, Yeteng An

 [View PDF](#) [Abstract](#) ▼ [Export](#) ▼

Obrázek 10.8: Výsledek vyhledávání nástroje SCOAP³

- [Observation of four-top-quark production in the multilepton final state with the ATLAS detector](#)

Aad, G. ; Abbott, B. ; Abeling, K. ; Abicht, N. ; Abidi, S. ; et al - 2023-06-12

This paper presents the observation of four-top-quark ($\bar{t}t\bar{t}t$) production in proton-proton collisions at the LHC. The analysis is performed using an integrated luminosity of 140 fb^{-1} at a centre-of-mass energy of 13 TeV collected using the ATLAS detector. Events containing two leptons wi ...

Published in: **European Physical Journal C**, Volume 83 (2023) Issue 6 (Pages 1-35) by Springer

DOI: 10.1140/epjc/s10052-023-11573-0

[XML](#) [PDF/A](#)

Obrázek 10.10: Vizualizace projektu ve výsledku vyhledávání

Projekt FLEXIBLE PRODUCTION TECHNOLOGIES AND EQUIPMENT BASED ON ATMOSPHERIC PRESSURE PLASMA PROCESSING FOR 3D NANO STRUCTURED SURFACES

Zaměření: Processes and Equipment for High Quality Industrial Production of 3-Dimensional Nanosurfaces (NMP-2007-3.5-1)

Program: FP7
Status: SIGNED

Outstanding progress has been made in recent years in developing novel structures and applications for direct fabrication of 3D nanosurfaces. However, exploitation is limited by lack of suitable manufacturing technologies. In this project we will develop innovative in-line high throughput technologies based on atmospheric pressure surface and plasma technologies. The two identified approaches to direct 3D nanostructuring are etching for manufacturing of nanostructures tailored for specific applications, a...

Obrázek 10.11: Vizualizace organizace ve výsledku vyhledávání

Organizace ZAPADOCESKA UNIVERZITA V PLZNI

Typ: Higher or Secondary Education Establishments
Stát: Czechia (CZ)

Obrázek 10.12: Vizualizace reportu ve výsledku vyhledávání

Report Final Report Summary - N2P (FLEXIBLE PRODUCTION TECHNOLOGIES AND EQUIPMENT BASED ON ATMOSPHERIC PRESSURE PLASMA PROCESSING FOR 3D NANO STRUCTURED SURFACES)

Stát: Germany (DE)

Outstanding progress has been made in recent years in developing novel structures and applications for direct fabrication of 3D nano-surfaces. However, exploitation is limited by lack of suitable manufacturing technologies. In this project we develop...

Obrázek 10.13: Detail projektu

EU PROJECTS Databáze vědeckých projektů dotovaných Evropskou unií

English | Čeština

Detail projektu

Projekt FLEXIBLE PRODUCTION TECHNOLOGIES AND EQUIPMENT BASED ON ATMOSPHERIC PRESSURE PLASMA PROCESSING FOR 3D NANO STRUCTURED SURFACES

Zaměření: Processes and Equipment for High Quality Industrial Production of 3-Dimensional Nanosurfaces (NMP-2007-3.5-1)

Program: FP7
Status: SIGNED

Celkové náklady [€]: 10447885.73
Maximální příspěvek [€]: 7400000.0
Datum začátku: 01. 06. 2008
Datum konce: 30. 11. 2012

URL: <http://www.n2p-project.eu/>

Outstanding progress has been made in recent years in developing novel structures and applications for direct fabrication of 3D nanosurfaces. However, exploitation is limited by lack of suitable manufacturing technologies. In this project we will develop innovative in-line high throughput technologies based on atmospheric pressure surface and plasma technologies. The two identified approaches to direct 3D nanostructuring are etching for manufacturing of nanostructures tailored for specific applications, and coating. Major impact areas were selected, demonstrating different application fields. Impact Area 1 focuses on structures for solar cell surfaces. Nanostructured surfaces have the potential to improve efficiencies of cells by up to 25% (rel), having dramatic impact on commercial viability. Impact Area 2 focuses on biocidal surface structures. Increasing concerns about infections leading to the conclusion, that only multi-action approaches for control of infection transfer can be effective. We plan to combine such surfaces with 3D nanostructures, which will both immobilise and deactivate pathogenic organisms on surfaces. Impact Area 3 is the direct growth of aligned carbon nanotubes on electrode surfaces. The material is under investigation for use in high load capacitors which are seen as key components for energy storage systems, e.g. for Hybrid Electric Vehicle. Impact Area 4 focuses on tailored interfaces to achieve durable adhesion on polymer surfaces by 3D nanostructuring and coating. Target is to reduce energy consumption by introducing lightweight materials. The N2P partners have been chosen to ensure a strong capability to exploit and disseminate the outcomes. Involved end-user industries represent high market value segments: photonics, aeronautics, automotive

Počet záznamů: 22

Organizace	Reporty
Organizace AFS ENTWICKLUNGS + VERTRIEBS GMBH Role organizace: participant Typ: Private for-profit entities Stát: Germany (DE)	
Organizace AIRBUS DEFENCE AND SPACE GMBH Role organizace: participant Typ: Private for-profit entities Stát: Germany (DE)	
Organizace BAUR PRUF- UND MESSTECHNIK GMBH Role organizace: participant Typ: Private for-profit entities Stát: Austria (AT)	
Organizace CVD TECHNOLOGIES LIMITED Role organizace: participant Typ: Private for-profit entities	

Obrázek 10.14: Detail organizace

EU PROJECTS
English | Čeština

Databáze vědeckých projektů dotovaných Evropskou Unií

←
Detail organizace

Organizace ZAPADOCESKA UNIVERZITA V PLZNI

Typ: Higher or Secondary Education Establishments
 Stát: Czechia (CZ)
 Město: PILSEN
 URL: <http://www.zcu.cz>

Id záznamu v databázi: 999843894
 Informace o importu: cordis-h2020organizations.csv (11. 06. 2023 19:10:50)

Projekty Spolupracující organizace

Počet záznamů: 24

Projekt A service platform for aggregation", processing and analysis of urban and regional planning data

Role organizace: participant
 Zaměření: SME initiative on Digital Content and Languages (ICT-2011.4.1)
 Program: FP7
 Status: CLOSED

Urban and Regional Planning data sets are not aggregated so far", and thus it is very difficult to use them for any other purpose than for printing of simple publishing by the authorities that they were created by. Creating time series or comparative analyses on these data sets is not yet possible; researchers, spati...

Projekt Affordable High-Performance Green Redox Flow Batteries

Role organizace: participant
 Zaměření: Advanced Redox Flow Batteries for stationary energy storage (LC-BAT-4-2019)
 Program: H2020
 Status: SIGNED

The objective of HIGREEW (Affordable High-performance Green Redox flow batteries) is to design, develop and validate an advanced redox flow battery, based on new water-soluble low-cost organic electrolyte compatible with optimized low resistance membrane and fast electrodes kinetics for a high energy densi...

Obrázek 10.15: Detail reportu

EU PROJECTS
English | Čeština

Databáze vědeckých projektů dotovaných Evropskou Unií

←
Detail reportu

Report Final Report Summary - N2P (FLEXIBLE PRODUCTION TECHNOLOGIES AND EQUIPMENT BASED ON ATMOSPHERIC PRESSURE PLASMA PROCESSING FOR 3D NANO STRUCTURED SURFACES)

Stát: Germany (DE)
 Jazyk: en
 URL:

Teaser: Outstanding progress has been made in recent years in developing novel structures and applications for direct fabrication of 3D nano-surfaces. However, exploitation is limited by lack of suitable manufacturing technologies. In this project we develop...

Outstanding progress has been made in recent years in developing novel structures and applications for direct fabrication of 3D nano-surfaces. However, exploitation is limited by lack of suitable manufacturing technologies. In this project we develop innovative in-line high throughput technologies based on atmospheric pressure surface and plasma technologies. The two identified approaches to direct 3D nano-structuring are etching for manufacturing of nano-structures tailored for specific applications and coating. The overall project aims are:- Bridge the gap between advanced R&D results in nano-technology and industrial production- Demonstrate introduction of nano-technology into high impact application areas- Establish a sustainable network of leading European players, comprising industrial oriented R&D supplier and equipment supplier End-user industries The targeted applications are grouped in two main areas with wide interest both commercial and societal:- Energy production, e.g. solar energy storage, e.g. for e-car Energy savings; e.g. lightweight vehicles- Health Medicine, e.g. biocidal walls for hospitals or Hygiene, e.g. in high visitor traffic areas Four major impact areas were selected, demonstrating different application fields: Impact Area 1 - Photovoltaics is aimed on efficient power production based on solar cells, one of the major and fastest growing renewable energies. The overall efficiency of such devices is significantly affected by an optimised "light management" to improve the energy harvesting properties of the solar cell. We are aiming for two areas with high commercial impact: (1a) Crystalline Silicon Solar Cells currently comprising a major market share for power production and (1b) Thin Film Silicon Solar Cells having high potential for

Projekt

Projekt FLEXIBLE PRODUCTION TECHNOLOGIES AND EQUIPMENT BASED ON ATMOSPHERIC PRESSURE PLASMA PROCESSING FOR 3D NANO STRUCTURED SURFACES

Zaměření: Processes and Equipment for High Quality Industrial Production of 3-Dimensional Nanosurfaces (NMP-2007-3.5-1)
 Program: FP7
 Status: SIGNED

Outstanding progress has been made in recent years in developing novel structures and applications for direct fabrication of 3D nanosurfaces. However", exploitation is limited by lack of suitable manufacturing technologies. In this project we will develop innovative in-line high throughput technologies based ...

11 Příloha B - Uživatelská příručka

11.1 Spustění aplikace

Pro spuštění aplikace je nutné mít nainstalovaný Maven. Nejprve zkompi-lujeme klientskou část. V příkazovém řádku se nasměrujeme do adresáře aplikace a zadáme následující příkaz:

```
mvn clean package
```

Dále je nutné mít nainstalovaný Docker. Spuštění kontejneru provedeme přes příkazový řádek. V adresáři aplikace zadáme příkaz:

```
docker-compose up --build
```

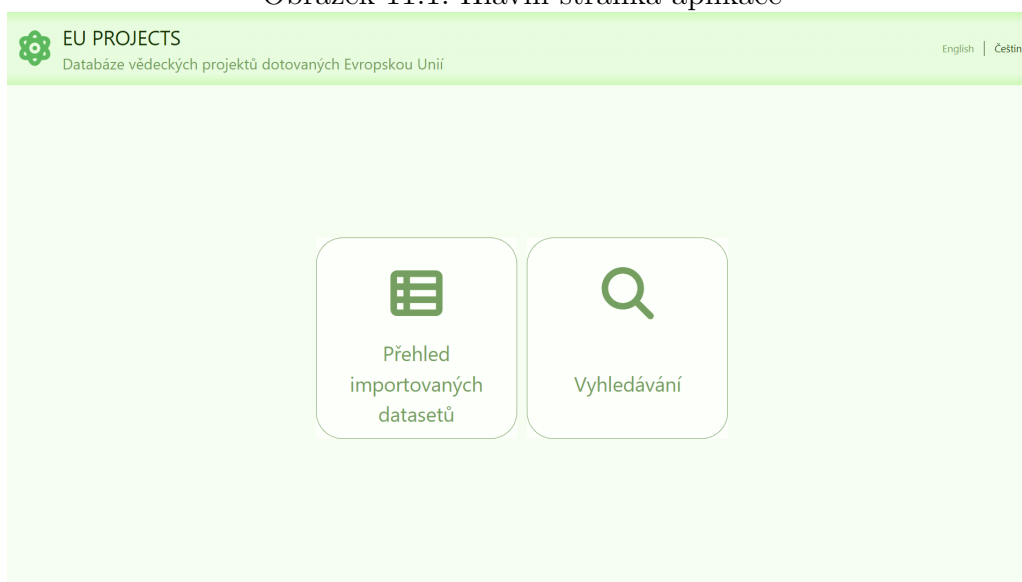
Konfigurační soubory Dockeru uvnitř naší aplikace po zadání příkazu spustí kontejner pro použití.

Nyní stačí otevřít webový prohlížeč a zadat adresu:

```
localhost:8085
```

Po načtení aplikace se dostaneme na hlavní stránku.

Obrázek 11.1: Hlavní stránka aplikace



11.2 Přehled importovaných datasetů

Na hlavní stránce vybereme sekci "Přehled importovaných datasetů". Zde vidíme všechny doposud importované datasety včetně základních informací.

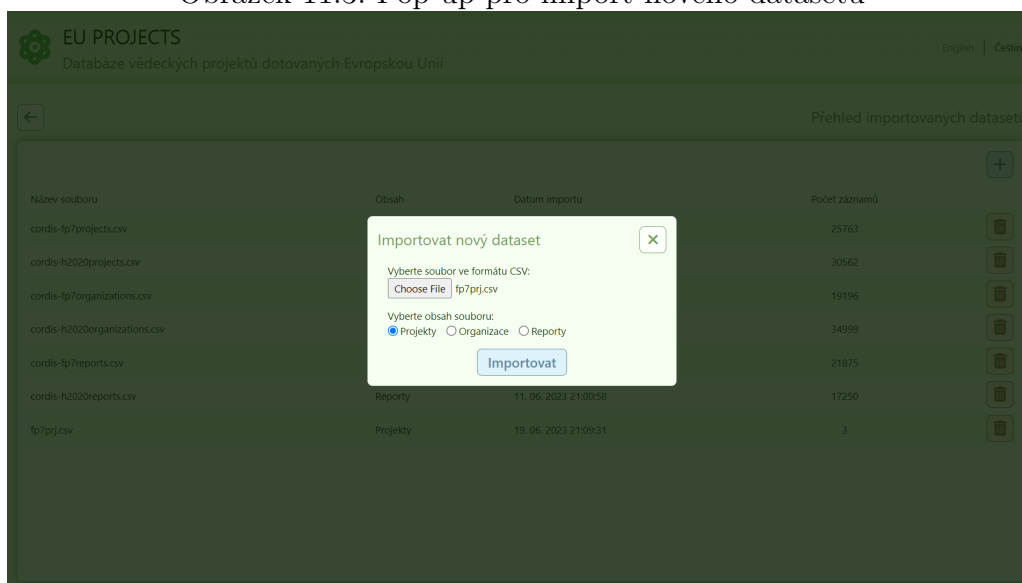
Obrázek 11.2: Stránka Přehled importovaných datasetů



Název souboru	Obsah	Datum importu	Počet záznamů	
cordis-fp7projects.csv	Projekty	11. 06. 2023 18:55:31	25763	
cordis-h2020projects.csv	Projekty	11. 06. 2023 18:57:22	30562	
cordis-fp7organizations.csv	Organizace	11. 06. 2023 19:01:46	19196	
cordis-h2020organizations.csv	Organizace	11. 06. 2023 19:10:50	34999	
cordis-fp7reports.csv	Reporty	11. 06. 2023 20:47:05	21875	
cordis-h2020reports.csv	Reporty	11. 06. 2023 21:00:58	17250	
fp7pj.csv	Projekty	19. 06. 2023 21:09:31	3	

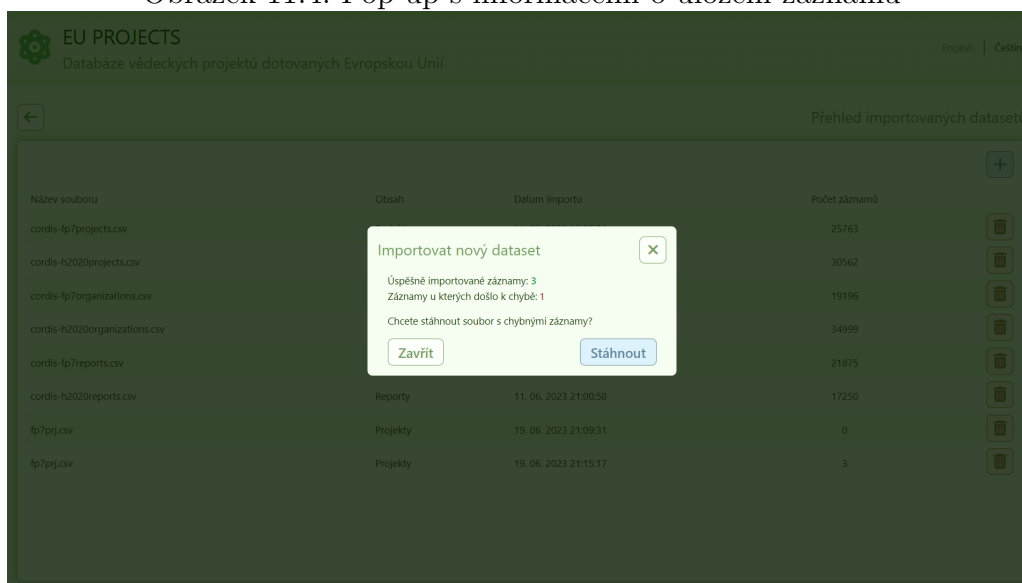
Pro přidání nového datasetu klikneme na tlačítko s ikonou "+" v pravém horním rohu. Objeví se pop-up pro výběr souboru a volbou typu dat. Po vyplnění polí klikneme na tlačítko "Importovat".

Obrázek 11.3: Pop-up pro import nového datasetu



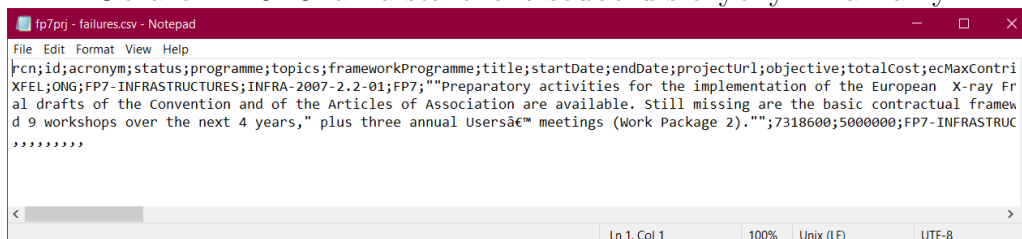
Po zpracování a uložení dat se pop-up změní. Vidíme počet úspěšně importovaných záznamů, popřípadě můžeme stáhnout soubor s chybnými záznamy k možné opravě.

Obrázek 11.4: Pop-up s informacemi o uložení záznamů



Níže je obrázek, jak může vypadat stažený soubor s chybným záznamem. Obsahem je hlavička z původního souboru následovaná chybnými záznamy (zde jen jeden chybný záznam).

Obrázek 11.5: Ukázka staženého souboru s chybnými záznamy



11.3 Vyhledávání

Vyhledávat můžeme pomocí filtrů nebo klíčových slov. Vedle textového pole pro klíčová slova můžeme vybrat, jestli chceme pomocí klíčových slov hledat ve všech entitách nebo vybereme nějakou konkrétní. Součástí zadání klíčových slov mohou být použity logické operátory AND, OR a NOT. Je důležité je zapisovat velkými písmeny. Vyšší prioritu má operátor AND, poté dojde k zpracování OR. Operátor NOT slouží k negaci příslušného textu. Pokud chceme hledat přesnou frázi, můžeme text zabalit do uvozovek.

Příklady vyhledávání s logickými operátory:

- *computer AND NOT coding* - najde všechny záznamy, které obsahují slovo *computer* a zároveň neobsahují slovo *coding*;
- *computer AND machine* - najde všechny záznamy, které obsahují slovo *computer* a zároveň slovo *machine*;
- *computer OR machine OR coding* - najde všechny záznamy, které obsahují slovo *computer* nebo slovo *machine* nebo slovo *coding*;
- *computer OR machine AND NOT coding* - najde všechny záznamy, které obsahují slovo *computer* nebo slovo *machine* a zároveň neobsahují slovo *coding*.

Na konci sloupce filtrování najdeme čtyři checkboxy pro využití kombinovaného vyhledávání. Níže je stručné vysvětlení.

Filtrovat:

- Projekty podle kritérií organizací - všechny projekty na nichž pracovala alespoň jedna organizace splňující kritéria organizací;
- Projekty podle kritérií reportů - všechny projekty u kterých je znám alespoň jeden report splňující kritéria reportů;
- Organizace podle kritérií projektů - všechny organizace u nichž platí, že každá z nich pracovala na alespoň jednom projektu splňující kritéria projektů;
- Reporty podle kritérií projektů - všechny reporty, které se týkají projektu splňující kritéria projektů.

Pro zobrazení výsledků klikneme na tlačítko lupy vedle textového pole pro klíčová slova nebo na konci sloupce s filtrovacími parametry.

Obrázek 11.6: Stránka Vyhledávání

The screenshot displays the 'EU PROJECTS' search interface. At the top, the header includes the logo, the text 'EU PROJECTS Databáze vědeckých projektů dotovaných Evropskou Unií', and language options 'English | Čeština'. Below the header, a search bar contains the text 'zapadoceska'. To the left of the search results is a 'Filtrování' (Filtering) sidebar with the following sections:

- Typ:** Projekty, Organizace, Reporty
- Projekty Zaměřené:** 2014: Framework Partnership Agreement, 2014: Policy environment for FET Flagships, 2D CONSULTATION EEN BBC H2020, 2nd Generation of design tools for ocean energy devices and arrays
- Program:** H2020, FP7

The main search results area, titled 'Výsledky vyhledávání', shows 'Počet záznamů: 1' (Number of records: 1). The single result is an organization: **Organizace: ZAPADOCESKA UNIVERZITA V PLZNI**. Below the organization name, it specifies 'Typ: Higher or Secondary Education Establishments' and 'Stát: Czechia (CZ)'. A 'Filtrovat' (Filter) button is located at the bottom of the sidebar.

Literatura

- [1] EU. *HORIZON 2020 ve stručnosti - Rámcový program EU pro výzkum a inovace*. Publications Office of the European Union, 2014. ISBN 978-92-79-38910-8.
- [2] EU. *Horizon EUROPE*. Publications Office of the European Union, 2020. ISBN 978-92-76-24754-8.
- [3] EU. *PARTICIPATION IN HORIZON EUROPE: KEEPING IT SIMPLE - competitive, fair and transparent, easy to access*. Publications Office of the European Union, 2020. ISBN 978-92-76-22119-7.
- [4] *Informace o fondech* [online]. DotaceEU.cz, 2018. [cit. 2020/12/21].
Dostupné z: <https://dotaceeu.cz/cs/evropske-fondy-v-cr/informace-o-fondech>.
- [5] *EU funding* [online]. European Union, 2019. [cit. 2020/12/21]. Dostupné z: https://europa.eu/european-union/about-eu/funding-grants_en.
- [6] *Selection process* [online]. European Commission, 2019. [cit. 2020/12/21].
Dostupné z: https://ec.europa.eu/info/funding-tenders/how-eu-funding-works/how-get-funding/selection-process_en.
- [7] *Types of funding* [online]. European Commission, 2019. [cit. 2020/12/21].
Dostupné z: https://ec.europa.eu/info/funding-tenders/how-eu-funding-works/types-funding_en.
- [8] *Funding programmes and open calls* [online]. European Commission, 2020. [cit. 2020/12/29]. Dostupné z: https://ec.europa.eu/info/research-and-innovation/funding/funding-opportunities/funding-programmes-and-open-calls_en.
- [9] *The Research Framework Programmes* [online]. State Secretariat for Education, Research and Innovation SERI, 2020. [cit. 2021/01/04].
Dostupné z: <https://www.sbfi.admin.ch/sbfi/en/home/research-and-innovation/international-cooperation-r-and-i/eu-framework-programmes-for-research/previous-eu-framework-programmes.html>.
- [10] *European research area (ERA)* [online]. European Commission, 2020. [cit. 2021/01/04]. Dostupné z: https://ec.europa.eu/info/research-and-innovation/strategy/era_cs.

- [11] *EU Framework Programmes for Research & Innovations* [online]. Reinforcing CAP, 2017. [cit. 2021/01/04]. Dostupné z: <https://www.recap-h2020.eu/eu-framework-programmes-for-research-and-innovation/>.
- [12] *IMPLEMENTATION STRATEGY FOR HORIZON EUROPE, Version 1.0* [online]. European Commission, 2020. [cit. 2021/01/20]. PDF. Dostupné z: https://ec.europa.eu/info/sites/info/files/research_and_innovation/strategy_on_research_and_innovation/documents/ec_rtd_implementation-strategy_he.pdf.
- [13] *Horizon Europe* [online]. European Commission, 2021. [cit. 2021/01/19]. Dostupné z: https://ec.europa.eu/info/horizon-europe_en.
- [14] *Horizont Evropa* [online]. European Commission, 2019. [cit. 2021/01/19]. PDF. Dostupné z: https://ec.europa.eu/info/sites/info/files/research_and_innovation/strategy_on_research_and_innovation/presentations/horizon_europe_cs_investice_utvarejici_nasi_budoucnost.pdf.
- [15] *Joint Research Centre* [online]. European Commission, 2021. [cit. 2021/01/20]. Dostupné z: <https://ec.europa.eu/info/departments/joint-research-centre>.
- [16] *Nový rámcový program Horizon Europe* [online]. Národní portál pro evropský výzkum, 2019. [cit. 2021/01/20]. Dostupné z: <https://www.evropskyvyzkum.cz/cs/nastroje-spoluprace/ramcove-programy/9-rp>.
- [17] *About CORDIS* [online]. European Commission, 2021. [cit. 2021/01/25]. Dostupné z: <https://cordis.europa.eu/about>.
- [18] *Archives* [online]. European Commission, 2021. [cit. 2021/01/25]. Dostupné z: <https://cordis.europa.eu/about/archives>.
- [19] *CORDIS services* [online]. European Commission, 2021. [cit. 2021/01/25]. Dostupné z: <https://cordis.europa.eu/about/services>.
- [20] *CORDIS content* [online]. European Commission, 2021. [cit. 2021/01/25]. Dostupné z: <https://cordis.europa.eu/about/content>.
- [21] *CORDIS - EU research projects under FP7 (2007-2013)* [online]. EU Open Data Portal, 2018. [cit. 2021/01/26]. Dostupné z: <https://data.europa.eu/euodp/en/data/dataset/cordisfp7projects>.
- [22] *CORDIS - EU research projects under Horizon 2020 (2014-2020)* [online]. EU Open Data Portal, 2021. [cit. 2021/01/26]. Dostupné z: <https://data.europa.eu/euodp/en/data/dataset/cordish2020projects>.

- [23] *J. Fiala, Elektronické informační zdroje: využití pro život I.* [online]. Masarykova Univerzita, 2023. [cit. 2023/04/08]. Dostupné z: https://is.muni.cz/elportal/estud/ff/js07/informace/materialy/pages/eiz_opora.pdf.
- [24] *Elektronický zdroj* [online]. WikiKnihovna, 2023. [cit. 2023/04/08]. Dostupné z: https://wiki.knihovna.cz/index.php/Elektronick%C3%BD_zdroj.
- [25] *About the ACM Digital Library* [online]. ACM Digital Library, 2023. [cit. 2023/04/08]. Dostupné z: <https://dl.acm.org/about>.
- [26] *About the repository* [online]. SCOAP3, 2023. [cit. 2023/04/08]. Dostupné z: <https://scoap3.org/scoap3-repository/>.
- [27] *About Cambridge University Press* [online]. Cambridge University Press, 2023. [cit. 2023/04/08]. Dostupné z: <https://www.cambridge.org/core/about>.
- [28] *About DOAJ* [online]. DOAJ, 2023. [cit. 2023/04/08]. Dostupné z: <https://www.doaj.org/about/>.
- [29] *Journals & Books* [online]. ScienceDirect, 2023. [cit. 2023/04/08]. Dostupné z: <https://www.sciencedirect.com/browse/journals-and-books>.
- [30] *Europeana About* [online]. Europeana, 2023. [cit. 2023/04/08]. Dostupné z: <https://www.europeana.eu/en/about-us>.
- [31] *Spring Framework - Overview* [online]. tutorialspoint, 2023. [cit. 2023/06/12]. Dostupné z: https://www.tutorialspoint.com/spring/spring_overview.htm.
- [32] *What is JPA? Introduction to Java persistence* [online]. InfoWorld, 2023. [cit. 2023/06/12]. Dostupné z: <https://www.infoworld.com/article/3379043/what-is-jpa-introduction-to-the-java-persistence-api.html>.
- [33] *What is JSF? Introducing JavaServer Faces* [online]. InfoWorld, 2023. [cit. 2023/06/12]. Dostupné z: <https://www.infoworld.com/article/3322533/what-is-jsf-introducing-javascript-faces.html>.