

Portál OpenAIRE

Martin Lochman

Ústav informačních studií a knihovnictví FF UK v Praze

E-mail: martin.lochman@gmail.com

Záznam původního příspěvku:

LOCHMAN, Martin. Portál OpenAIRE. *Ikaros* [online]. Březen 2015, roč. 19, č. 3 [cit. 2015-04-01]. ISSN 1212-5075. Dostupný z: <http://ikaros.cz/portal-openaire>.

English title:

OpenAIRE portal

English abstract:

The article focuses on OpenAIRE portal, a service provider over digital archives and repositories, which was launched in 2009 with a singular goal – to provide unlimited access to research outputs and scholarly literature in Europe. The first part of the text briefly describes the history of the portal, its structure from the technical standpoint and the set of available services. The second segment analysis the user interface as well as the search options. Lastly, statistical data and a final assessment are provided.

Klíčová slova:

Informační služby, metadata, digitální archivy, přebírání záznamů, OAI-PMH

Keywords:

Information services, metadata, digital archives, metadata harvesting, OAI-PMH

1. Úvodní informace

Portál OpenAIRE (angl. OpenAIRE portal) představuje výsledek jednoho z nejvýznamnějších projektů služeb nad digitálními archivy a institucionálními repozitáři s otevřeným přístupem v oblasti vědy, výzkumu a inovací v 21. století. Portál je financován z rozpočtu Evropské komise a v současné době je řízen Konsorciem OpenAIRE (angl. OpenAIRE consortium)¹.

Hlavními cíli uvedeného projektu byly²:

- vybudování podpůrných struktur pro výzkumné pracovníky, kteří jsou povinni uložit publikované výsledky projektů ze 7. Rámcového programu Evropské unie
- zřízení a provoz virtuální infrastruktury pro recenzované články a další formy výsledků vědecké činnosti
- využití vědeckých a výzkumných dat a experimentální ověření manipulace s těmito daty ve spolupráci s vědeckými komunitami

¹ *OpenAIRE* [online]. OpenAIRE Consortium, 2009- [cit. 2015-04-01]. Dostupný z: <https://www.openaire.eu/>.

² TKACÍKOVÁ, Daniela. Projekt OpenAIRE – Výzva a příležitost i pro Českou republiku. In: *INFORUM 2010: 16. konference o profesionálních informačních zdrojích, Praha, 25–27. 5. 2010* [online]. Praha: Albertina Icome Praha, 2010 [cit. 2015-04-01]. Dostupný z: <http://www.inforum.cz/pdf/2010/tkacikova-daniela-1.pdf>.

- udržitelnost e-infrastruktury vybudované v rámci projektu OpenAIRE.

Portál, který v rámci projektu vznikl, umožňuje skrze jednotné uživatelské rozhraní zprostředkovat přístup k výsledkům vědecké činnosti ve formě elektronických dokumentů rozmanitých typů a formátů v celoevropském měřítku. Dále nabízí dílčí služby typu statistik, signálního informování, personalizovaného uživatelského rozhraní, propojení na sociální sítě a dokonce i vlastního virtuálního úložiště Zenodo³, které směřují k podpoře principů otevřeného přístupu v celoevropském měřítku.

K počátku dubna 2015 portál OpenAIRE nabízel přístup již k více než 9 900 000 metadatových záznamů dokumentů a více než 6 800 souborů dat (angl. dataset) pocházejících z 579 zdrojových archivů a časopisů s otevřeným přístupem. K více než 9,5 miliónům těchto primárních dokumentů (96 % celkového počtu) je garantován přístup v režimu otevřeného přístupu. Nejvíce příspěvků pochází z Nizozemska (615 tisíc záznamů)⁴.

2. Vývoj

Historie portálu je spojena s tříletým projektem Evropské komise OpenAIRE (zkratka pro Infrastrukturu s otevřeným přístupem pro výzkum v Evropě, angl. Open Access Infrastructure for Research in Europe), který byl realizován mezi prosincem 2009 a listopadem 2012 a jehož cílem bylo vytvořit infrastrukturu pro komunikaci vědeckých informací v souladu s pilotním programem pro OA Evropské komise (angl. European Commission Open Access Pilot) a pokyny pro OA Evropské výzkumné rady (angl. European Research Council Open Access guidelines). Úspěšné završení projektu představilo nejen výše uvedený portál, ale také tzv. Evropský systém helpdesků (angl. European Helpdesk System), založený na distribuované síti národních a regionálních kontaktních osob za účelem spolupráce s knihovnickými organizacemi zainteresovanými v záležitostech OA (např. LIBER, SPARC Europe), a vlastní digitální úložiště určený odborníkům, kteří nemají přístup do institucionálních nebo předmětových repozitářů.

V prosinci 2011 byl zahájen návazný projekt OpenAIREplus (zkratka pro 2. Generaci Infrastruktury s otevřeným přístupem pro výzkum v Evropě, angl. 2nd Generation of Open Access Infrastructure for Research in Europe) financovaný 7. Rámcovým programem Evropské komise (angl. The EC 7th Framework Programme), jehož záměrem bylo rozšířit záběr původní infrastruktury nejen směrem vyššího počtu uživatelů, ale i typologie odborných informací. Zahrnuty měly být materiály – včetně nerecenzovaných publikací a výzkumných souborů dat – ze všech vědeckých disciplín v členských státech EU. Mezi dílčí cíle projektu patřilo také propojování souborů dat a publikací napříč vědeckými disciplínami pomocí sémantických vztahů, tzv. „přidružených publikací“ (angl. Enhanced publications) vytvořených koncovými uživateli a zajištění interoperability s již existujícími virtuálními infrastrukturami jako Mendeley, DataCite a ORCID⁵. Projekt měl původně trvat 30 měsíců (do 31. května 2014), ale byl prodloužen až do 31. prosince 2014.

K 1. lednu 2015 byla zahájen navazující projekt OpenAIRE2020 (angl. The OpenAIRE2020 Project). Projekt je financován z rozpočtu Evropské komise a jeho hlavním cílem je podpora programu Horizont 2020 (angl. Horizon 2020) zprostředkováním infrastruktury, v jejímž

³ Zenodo [online]. Meyrin (Švýcarsko): CERN, 2009- [cit. 2015-04-01]. Dostupný z: <http://www.zenodo.org/>.

⁴ OpenAIRE Monitoring. In: *OpenAIRE* [online]. OpenAIRE Consortium, 2009- [cit. 2015-04-01]. Dostupný z: <https://www.openaire.eu/overall-stats/statistics/infra-monitoring>.

⁵ Project factsheets. In: *OpenAIRE* [online]. OpenAIRE Consortium, 2009- [cit. 2015-04-01]. Dostupný z: <https://www.openaire.eu/about/general-information/openairefactsheet-40>.

rámci budou vědecké publikace a výsledky výzkumů ukládány, propojovány a zpřístupňovány.

V současné době je do projektu zapojených již 31 států v Evropě.

3. Struktura

Datová infrastruktura portálu OpenAIRE je založena na balíčku nástrojů D-NET Software Toolkit⁶. Ten nabízí otevřené prostředí orientované na služby, ve kterém mohou být komplexní procesy zpracování dat snadno navrženy, modifikovány a obohaceny o nové funkce⁷. V současné době tak struktura portálu sestává ze tří následujících vrstev⁸:

- **Vrstva pro správu dynamické sady heterogenních datových zdrojů:** Hlavní funkcí je sběr a integrace metadat z institucionálních repozitářů kompatibilních s pokyny OpenAIRE (angl. OpenAIRE guidelines) pomocí protokolu OAI-PMH. Za účelem pozdějšího propojování záznamů a jejich sémantického obohacování však portál také čerpá data z databáze Evropské komise E-CORDA (oficiální seznam projektů tzv. Sedmého rámcového programu EC FP7) a adresáře OpenDOAR (oficiální seznam institucionálních repozitářů).

V současné době portál OpenAIRE sklízí metadata všech typů otevřeně přístupných dokumentů z mnoha institucionálních repozitářů, z portálů a agregátorů na národní úrovni i z archivů OA časopisů. Kromě evropských zdrojů pokrývá také řadu zdrojů mimoevropských včetně archivu arXiv.org.

- **Vrstva pro správu informačního prostoru:** Zde probíhají procesy „normalizace“ (resp. transformace a čištění) a propojování získaných dat a metadat. Vrstva zároveň zprostředkuje funkce automatického odvozování vztahů mezi záznamy publikací a projekty. S cílem poskytnout co nejpřesnější statistické informace jsou její součástí také deduplikační procesy umožňující identické záznamy a autory identifikovat a sloučit.
- **Vrstva pro podporu přístupu do informačního prostoru:** Uvedená vrstva je zodpovědná za zprostředkování funkcí zajišťujících přístup koncových uživatelů (vyhledávání, prohlížení a nadstavbové služby) a systémů (všechny entity v informačním prostoru jsou dosažitelné pomocí aplikačních programových rozhraní REST, OAI-PMH, OAI-ORE a SRW/CQL)⁹.

⁶ Na tomto technologickém řešení byla založena i architektura bývalého portálu DRIVER (zkratka pro „Vize infrastruktury digitálních repozitářů pro Evropský výzkum, angl. Digital Repository Infrastructure Vision for European Research). Ten z ekonomických důvodů zanikl a roku 2014 byla jeho data sloučena s centrální metadatovou databází OpenAIRE.

⁷ *D-NET LAB* [online]. Athens (Řecko): University of Athens, 2006- [cit. 2015-04-01]. Dostupný z: <http://www.d-net.research-infrastructures.eu/>.

⁸ MANGHI, Paolo et al. OpenAIREplus: the European Scholarly Communication Data Infrastructure. *D-Lib Magazine* [online]. September/October 2012, vol. 18, no. 9/10 [cit. 2015-04-01]. ISSN 1082-9873. Dostupný z: <http://www.dlib.org/dlib/september12/manghi/09manghi.html>.

⁹ CASTELLI, Donatella a Paolo MANGHI. The OpenAIRE Project – Open Access Infrastructure for Research in Europe. *Ercim News* [online]. January 2010, no. 80 [cit. 2015-04-01]. ISSN 0926-4981. Dostupný z: <http://ercim-news.ercim.eu/en80/es/openaire>.

4. Dílčí služby

Vedle základních možností vyhledávače nad digitálními archivy, tedy dotazového vyhledávání a prohlížení rejstříků, portál OpenAIRE nabízí několik dílčích služeb orientovaných jak na autory, resp. poskytovatele obsahu (včetně provozovatelů repozitářů), tak i na koncové uživatele¹⁰.

Lokace akademických institucí/repozitářů: Vědcům a odborníkům, kteří mají zájem zveřejnit svůj výzkum prostřednictvím portálu, OpenAIRE nabízí jednoduchou funkci lokalizace instituce nebo repozitáře, s nimiž jsou afilováni. Tato funkce zahrnuje jak dotazové vyhledávání opatřené našeptávačem, tak i interaktivní mapu všech repozitářů zapojených do projektu.

Zenodo: V případě, že daný repozitář s portálem nespolečně pracuje, popřípadě se jedná o autora, který není afilován s žádnou akademickou institucí, OpenAIRE umožňuje dokumenty uložit do vlastního úložiště Zenodo provozovaného při Evropské organizaci pro jaderný výzkum CERN.

Propojení publikací: Registrovaní uživatelé mají možnost propojit výsledky výzkumu ve formě publikací a souborů dat s projekty, v jejichž rámci výzkum proběhl, a entitami, které ho financovaly.

Připojení k OpenAIRE: Provozovatelé mohou prostřednictvím jednoduchého rozhraní registrovat své archivy a repozitáře v síti OpenAIRE. Nejprve jsou však nuceni otestovat jejich kompatibilitu se směrnicemi OpenAIRE, teprve poté je možné přistoupit k samotné registraci.

Statistiky: Služba OpenAIRE Monitoring uživatelům nabízí každodenně aktualizované statistické informace zahrnující nejen celkové počty publikací, souborů dat a repozitářů, ale i dílčí data zohledňující jejich jednotlivé charakteristiky (geografický původ, jazyk, zdrojový archiv, typ, datum) a žebříčky (top 20 převládajících jazyků a top 20 zdrojových archivů). Statistiky lze zobrazit ve formě graficky přívětivých tabulek, sloupcových grafů a kruhových diagramů).

OpenAIRE newsletter: Registrovaní uživatelé se mohou zdarma přihlásit k odběru online newsletteru, jehož obsah se vztahuje k aktuálnímu dění v oblasti politiky otevřeného přístupu, aktivit OpenAIRE, vědecké komunikace a správy výzkumných dat.

5. Vyhledávání

Vyhledávání metadatových záznamů dokumentů a souborů dat lze v rámci portálu OpenAIRE realizovat dvěma hlavními způsoby: klasickým **dotazovým vyhledáváním** pomocí vyhledávacího okna a alternativním **prohlížením** rejstříků.

První z uvedených způsobů nabízí dvě varianty: jednoduché a pokročilé vyhledávání. Základem první z uvedených je jediné vyhledávací okno, do kterého je zadán informační dotaz jednoslovné či víceslovné povahy. Jednoduchý režim (viz obr. 1) umožňuje pracovat jak se základními booleovskými operátory AND, OR a NOT, tak i se sadou syntaktických prvků zahrnujících zástupné znaky, pravostranné/levostranné omezení a specifikaci dotazu použitím uvozovek.

¹⁰ Introducing the functionalities. In: *OpenAIRE* [online]. OpenAIRE Consortium, 2009- [cit. 2015-04-01]. Dostupný z: <https://www.openaire.eu/functionalities/functionalities/intro-to-functionality/>.

The screenshot shows the OpenAIRE search interface. At the top, there are social media icons and links for 'NEWSLETTER', 'LOG IN', and 'REGISTER'. The main navigation bar includes 'PARTICIPATE DEPOSIT, JOIN', 'SEARCH PUBLICATIONS, DATA, PROJECTS', 'STATISTICS OA, PROJECTS, TOPICS', 'SUPPORT FAQ, HELPDESK, GUIDES', and 'OPEN ACCESS IN EUROPE'. Below this, there are links for 'Home', 'Search', and 'Find'.

The search interface features a search bar with a 'Search keywords' input field and a search button. Below the search bar, there are filter tabs: 'All', 'Publications', 'Data', 'projects', 'People', 'Organizations', and 'Datasources'. The 'All' tab is currently selected.

On the right side, there is a 'Search' section with a 'Find' link and a list of filter options: 'all', 'publications', 'data', 'projects', 'people', 'organizations', and 'data sources'. Below this list, there are links for 'OpenAIRE Data Providers' and 'Information'.

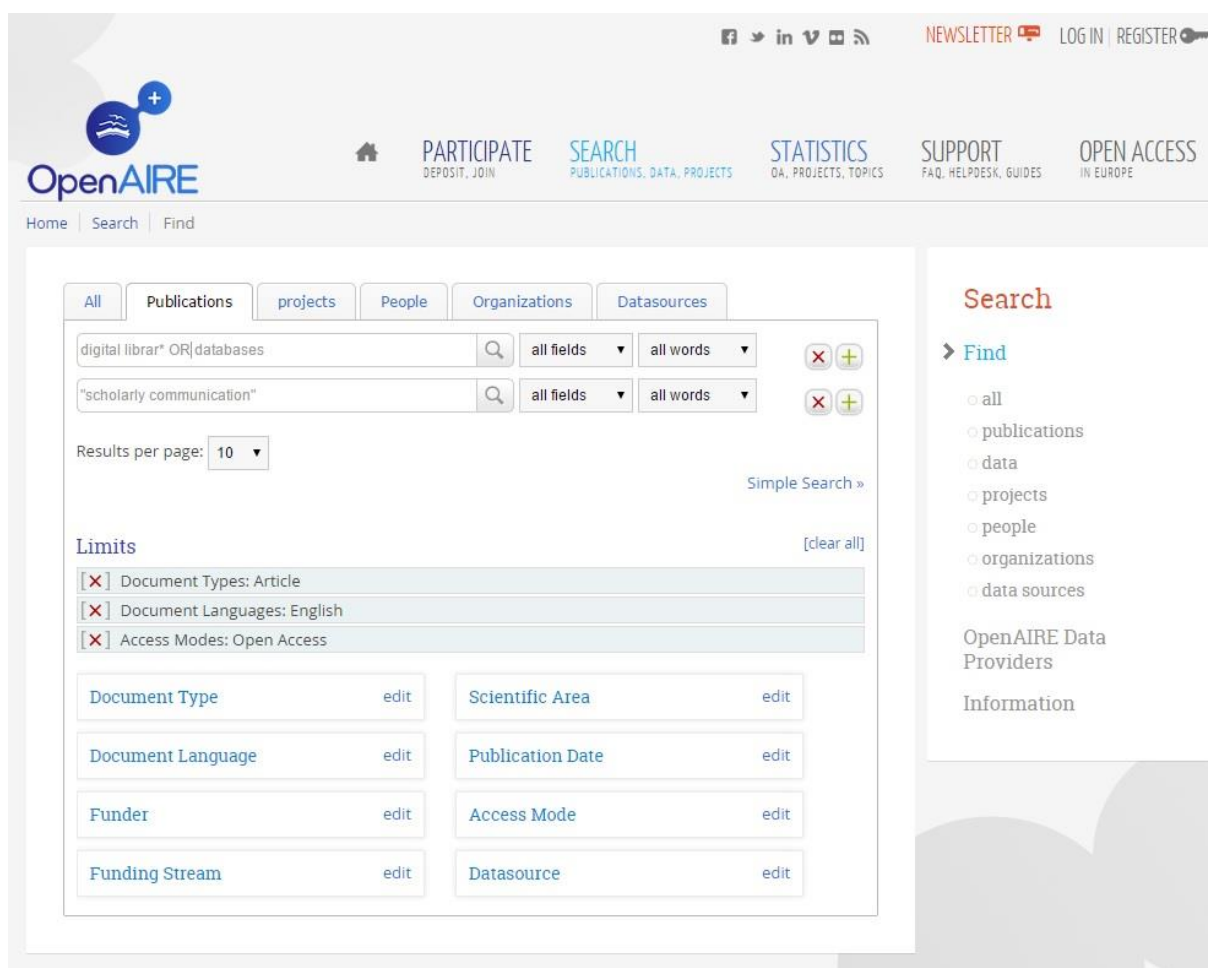
The main content area displays a table of search results. The table has four columns: 'Document Type', 'Document Language', 'Publication Year', and 'Datasource'. Below the table, there are two more sections: 'Funder' and 'Access Mode'.

| Document Type | Document Language | Publication Year | Datasource |
|--------------------------------------------|---------------------------------------|-------------------------------|----------------------------------------------------|
| Article (7048079) | English (6611482) | 2012 (909467) | Europe PubMed Central (3017041) |
| Unknown (1225846) | Undetermined (826668) | 2013 (893494) | DOAJ-Articles (1737620) |
| Conference object (310838) | Spanish (495230) | 2011 (836142) | arXiv.org e-Print Arc... (1006629) |
| Doctoral thesis (306184) | German (316172) | 2010 (727515) | INRIA a CCSD electron... (332759) |
| Other (296623) | Portuguese (297300) | 2014 (666434) | DSpace @ Cambridge (186835) |
| View more ⇨ | View more ⇨ | View more ⇨ | View more ⇨ |

| Funder | Access Mode | Context |
|----------------------------------------------|---------------------------------------|-----------------------------|
| European Commission (100069) | Open Access (9532871) | ec (100069) |
| Wellcome Trust (10702) | Restricted (315143) | uk (10702) |
| | Closed Access (66905) | egi (3595) |
| | not available (20703) | |
| | Embargo (3241) | |

Obrázek 1 - Náhled rozhraní jednoduchého vyhledávání

Pokročilé vyhledávání (viz obr. 2) lze zpřístupnit prostřednictvím tlačítka „více vyhledávacích možností“ (angl. More search options). Ve své standardní podobě nabízí jedno vyhledávací okno (přidat lze však mnoho dalších) a dvě rozbalovací nabídky, které specifikují typ pole (všechna pole, název, autor, vydavatel a předmět) a vztahy mezi zadanými klíčovými slovy (všechna slova a jakákoliv slova – možnosti odpovídající booleovským operátorům AND a OR). Uživatelé mohou dále vyhledávání zúžit výběrem **typu dokumentu**, **jazyka**, **entity**, která výzkum financovala, **vědecké oblasti**, **data publikování**, **způsobu přístupu** a **zdrojového archivu**. V neposlední řadě je možné nastavit počet záznamů na zobrazenou stránku výsledků.



Obrázek 2 - Náhled rozhraní pokročilého vyhledávání

Režim prohlížení je možné realizovat výběrem prvku ze sedmi stejných skupin, které jsou k dispozici k specifikaci vyhledávacího procesu v pokročilém vyhledávání, tedy: **typ dokumentu, jazyk, entita**, která výzkum financovala, **vědecká oblast, rok publikování, způsob přístupu a zdrojový archiv**.

Nutné je však uvést, že obě výše uvedené varianty, tj. jednoduché dotazové vyhledávání i prohlížení rejstříků, lze provést v následujících obsahových kategoriích¹¹: **publikace, data, projekty, lidé, organizace a datové zdroje**.

Poté, co systém provede vyhledávání nad indexem metadatové databáze, je uživatelům zobrazena příslušná množina metadatových záznamů, uspořádaná podle klesající obsahové relevance (viz obr. 3). Záznamy jsou standardně zobrazeny ve zkrácené podobě a obsahují pouze následující údaje: název příspěvku (vybavený hypertextovým odkazem vedoucím na úplný záznam), jméno autora (popř. autorů), rok publikování a abstrakt. Graficky – formou zámku v pravé části záznamu – je také indikován způsob přístupu¹².

¹¹ Ve všech kategoriích najednou lze provést pouze jednoduché vyhledávání a prohlížení, nikoliv vyhledávání pokročilé.

¹² Otevřený přístup je znázorněn ikonou oranžového odemčeného zámku, zatímco omezený (angl. Restricted) a embargovaný (angl. Embargoed) obsah indikuje šedý uzamčený zámek.

The screenshot shows the OpenAIRE search interface. At the top, there are navigation links for Home, Search, Find, and projects. The main header includes the OpenAIRE logo and navigation tabs for PARTICIPATE, SEARCH, STATISTICS, SUPPORT, and OPEN ACCESS. A search bar contains the text 'digital library'. Below the search bar, there are tabs for 'All', 'Publications', 'Data', 'projects', 'People', 'Organizations', and 'Datasources'. The search results are displayed in a list format, with each entry showing the title, author(s), and a brief description. The left sidebar contains several filter categories: Document Type (Article, Unknown, Book, Master thesis, Part of book or chapt...), Document Language (English, Undetermined, Dutch; Flemish, Spanish, Lithuanian), Funder (European Commission, Wellcome Trust), Funding Stream (Cooperation, Ideas, People, Immunology and Infect..., Molecules, Genes and ...), and Scientific Area (HEALTH, ERC, PEOPLE, ICT, KBBE).

Obrázek 3 - Náhled množiny výsledků vyhledávání a faset údajů

S výslednými záznamy lze dále pracovat prostřednictvím četných faset údajů (angl. Refine by), které jsou zobrazeny v sloupci v levé části obrazovky. Valná většina údajů je ryze formální povahy: typ dokumentu, jazyk dokumentu, entita financující výzkum, tzv. Funding Stream, rok publikace, způsob přístupu a zdrojový archiv. Jediným reprezentantem věcných údajů je vědecká oblast¹³.

¹³ Kategorizace tohoto údaje vychází ze Sedmého rámcového programu EC FP7, který vyčleňuje čtyři základní bloky, resp. programy – Spolupráce (angl. Cooperation), Nápady (angl. Ideas), Lidé (angl. People) a Kapacity (angl. Capacities) – a pátý specifický program zaměřený na nukleární výzkum a výcvik (angl. Nuclear research and training). V rámci uvedených pěti programů je definováno celkem 22 vědeckých oblastí, podle kterých lze výsledky vyhledávání dále zúžit. V samotném systému jsou uvedeny pod zkratkami: ICT pro informační a komunikační technologie, ENV pro životní prostředí, atd.

Přímou součástí rozhraní zobrazujícího výsledky vyhledávání je také okno jednoduchého vyhledávání a prostřednictvím tlačítka „více vyhledávacích možností“ mohou uživatelé také ihned zpřístupnit vyhledávání pokročilé.

Úplný záznam dokumentu (viz obr. 4) obsahuje kromě základní sady metadat (viz níže v podkapitole 6.) následující položky: použitá literatura (angl. References), související data (angl. Related Data) – v případě, že je publikace propojena s určitým souborem dat – a podobné publikace (angl. Similar Publications) – které na základě sémantické podobnosti textu zobrazí seznam tematicky podobně laděných publikací. K dispozici je dále možnost záznam sdílet na sociální síti – Facebooku, Twitteru, Google+ nebo LinkedIn, vygenerovat bibliografickou citaci v jednom z nejvyžívanějších formátů včetně BibTeX, Harvard, APA a IEEE a zpřístupnit záznam samotného zdrojového archivu včetně seznamů dokumentů a souborů dat a dílčích statistik. Klíčová je položka „Stáhnout z“ (angl. Download from), pod níž je uveden seznam odkazů, z nichž lze stáhnout kopii primárního dokumentu. Pomocí tlačítek „Propojit s projektem“ (angl. Link to project) a „Propojit se souborem dat“ (angl. Link to dataset) mohou registrovaní uživatelé využít výše popsané funkce propojení publikace se soubory dat nebo projekty.

The screenshot shows the OpenAIRE interface for a specific document. At the top, there is a navigation bar with the OpenAIRE logo and links for PARTICIPATE, SEARCH, STATISTICS, SUPPORT, and OPEN ACCESS. Below this is a breadcrumb trail: Home | Search | Find. The main content area displays the document title "DIGITAL ERA: UTILIZE OF CLOUD COMPUTING TECHNOLOGY IN DIGITAL LIBRARY" by Raghunadha Reddy, T. (2012). It includes metadata such as the publisher (International Journal of Digital Library Services), language (English), and type (Article). A detailed subject list is provided. A text abstract describes the paper's focus on cloud computing in digital libraries. On the right side, there are options to share, download from DOAJ-Articles, and cite the article in BibTeX format. At the bottom, there are tabs for References, Related Data, and Similar Publications, with a message indicating "No references."

Obrázek 4 - Ukázka metadatového záznamu ve výstupním prezentačním formátu

6. Metadatový záznam

Standardním formátem metadatových záznamů sklizených portálem OpenAIRE je **formát OpenAire** (angl. OpenAire Format) verze 0.2. V prefixu metadatového záznamu ve schématu

XML je uvedeno „oaf“. Vzhledem k různorodosti popisu v repozitářích, ze kterých jsou sklízena metadata, se jednotlivé záznamy do jisté míry odlišují délkou. Přesto lze identifikovat prvky, které jsou společné všem:

- **Název:** název, pod kterým je záznam v systému vyhledatelný. Je vybaven hypertextovým odkazem, který vede na záznam v primárním archivu.
- **Autor:** fyzická osoba, která nese primární odpovědnost za dílo.
- **Typ dokumentu:** přesně definovaný typ primárního dokumentu, který záznam popisuje

Vedle výše uvedených prvků se v metadatových záznamech vyskytují také následující:

- **Předmět** (v rámci záznamu v OpenAIRE uveden v angl. Jako „Subject“): obsahuje jednoslovné, popř. víceslovné termíny vyjadřující obsah primárního dokumentu. Vzhledem k rozmanitosti zdrojových archivů v nich ale panuje značná nejednotnost – na základě průzkumu bylo možné zjistit, že v některých případech je užito neřízených klíčových slov přímo zprostředkovaných autory, naopak v dalších je aplikováno Třídění JITA pro knihovní a informační vědu (angl. JITA Classification System of Library and Information Science) a Třídění Kongresové knihovny (v záznamech se vyskytují nejen alfanumerické notace, ale i přesné slovní ekvivalenty). V závislosti na provenienci jsou také některá předmětová hesla uvedena v národních jazycích (viz obr. 5, kde je vedle klíčových slov ve francouzštině aplikováno také Třídění Kongresové knihovny).
- **Abstrakt:** redukováný text vystihující obsahovou stránku dokumentu v národním jazyce.
- **Rok:** časový údaj vztahující se k publikování primárního dokumentu. V záznamu je uveden v závorce za autorovým jménem.
- **Nakladatel:** entita, resp. organizace nebo instituce, která dokument publikovala.
- **Typ dokumentu:** přesně definovaný typ primárního dokumentu.
- **Jazyk:** jazyk primárního dokumentu
- **Identifikátor:** trvalý identifikátor informačního zdroje. Podobně jako v případě předmětových hesel i zde panuje značná nejednotnost – značné množství záznamů je opatřeno identifikátorem DOI; užito je však také například i identifikátoru databáze PubMed Central (angl. PMC ID).

The screenshot shows the OpenAIRE interface for a specific article. At the top, there are navigation links for Home, Search, and Find. The article title is 'Deux approches du risque d'inondation en France' by Ledoux, Bruno; Bonnefoy, Jean-luc (1994). The publisher is the French National Center for Scientific Research (CNRS), and the journal is M@ppemonde. The language is French, and the type is Article. The subjects listed include CATASTROPHE, FRANCE, INONDATION, RISQUE NATUREL, PRÉVENTION, LCC:Mathematical geography, Cartography, LCC:GA1-1776, LCC:Geography. Anthropology. Recreation, LCC:G, DOAJ:Geography, DOAJ:Earth and Environmental Sciences, LCC:Mathematical geography. Cartography, LCC:GA1-1776, LCC:Geography. Anthropology. Recreation, LCC:G, DOAJ:Geography, DOAJ:Earth and Environmental Sciences, LCC:Mathematical geography. Cartography, LCC:GA1-1776, LCC:Geography. Anthropology. Recreation, LCC:G, LCC:Mathematical geography. Cartography, LCC:GA1-1776, LCC:Geography. Anthropology. Recreation, LCC:G, LCC:Mathematical geography. Cartography, LCC:GA1-1776, LCC:Geography. Anthropology. Recreation, LCC:G.

Below the subjects, there is a short abstract in French: 'En France, les sources d'informations centralisées sur les risques naturels sont rares. Pourtant, une politique de prévention engagée par l'État nécessite de disposer d'une information à l'échelle nationale, d'évaluer les enjeux et de mobiliser les acteurs. Deux sources d'informations sont décrites et exploitées, qui permettent, par une représentation cartographique nationale, d'amorcer analyses et débats.'

There are two buttons: 'Link to project' and 'Link to dataset'. Below these are tabs for 'References', 'Related Data', and 'Similar Publications'. The 'References' tab is active, showing 'No references.' There are also social media sharing options (Facebook, Twitter, LinkedIn) and a 'Share - Bookmark' section. A 'Download from' section lists the source as 'French National Center for Scientific Research (CNRS) / M@ppemonde'. A 'Cite this article' section shows a BibTeX citation: '@misc{Ledoux, Bruno_Bonnefoy, Jean-luc_1994, title={Deux approches du risque d'inondation en France}, publisher={French National Center for Scientific Research (CNRS)}, author={Ledoux, Bruno and Bonnefoy, Jean-luc}, year={1994}, month={Nov}}'. A 'Collected from' section lists 'DOAJ-Articles'.

Obrázek 5 - Ukázka metadatového záznamu s autorskými klíčovými slovy a termíny z Třídění Kongresové knihovny

7. Uživatelský přístup

Valná většina dílčích služeb portálu OpenAIRE (vyhledávání a prohlížení záznamů, statistické údaje) je uživatelům k dispozici bez nutnosti vytvoření vlastního účtu. Registrace vyžaduje naprosté minimum vstupních údajů (jméno a příjmení, heslo, e-mailovou adresu a verifikaci, že se jedná o živého uživatele prostřednictvím testu CAPTCHA). Registrovaní uživatelé mají přístup k výše uvedeným funkcím propojování publikací, souborů dat a projektů, identifikace vlastních prací a přihlášení se k nim (včetně možnosti nastavit přístupová práva), vytváření akcí (angl. Create an Event)¹⁴ a k sadě informačních a propagačních materiálů o portálu (angl. OpenAIRE NOAD Information and Dissemination KIT). Samozřejmostí je služba signálního informování formou RSS kanálu a propojení na sociální síť.

Uživatelé se také mohou dostat do přímého kontaktu s provozovateli portálu prostřednictvím propracovaného kontaktního formuláře nabízejícího následující kategorie pro specifikaci dotazu: poskytovatelé dat, otevřený přístup, vědci, správci přístupu, služby portálu a tvůrci politik (angl. Policy makers).

¹⁴ Tato funkce však není určena běžnému uživateli, nýbrž primárně účastníkům projektu OpenAIRE. Slouží k vkládání informací o aktuálním dění (konference, semináře, workshopy, etc.) spojeném s projektem a souvisejícími projekty.

8. Webové rozhraní

Webové rozhraní portálu OpenAIRE (viz obr. 6) je koncipováno velmi přehledně a jednoduše. Ústřední částí hlavní stránky je dynamické okno, které atraktivní formou prezentuje základní informace o portálu včetně stručných aktuálních statistik. V horní části obrazovky jsou k dispozici základní odkazy na **dílčí služby, vyhledávání a prohlížení, statistiky, podporu a často kladené otázky** a informace o **otevřeném přístupu v Evropské unii**. Nechybí také odkazy na sociální sítě, newsletter a přihlášení k uživatelskému účtu.

Bezprostředně pod výše popsáním dynamickým oknem se nachází rozhraní jednoduchého vyhledávání, které umožňuje prohledávání jak metadatové databáze, tak i webových stránek portálu. V dolní části obrazovky lze využít informačních odkazů rozdělených do 4 kategorií: výzkumníci (angl. Researchers), poskytovatelé dat (angl. Data Providers), správci výzkumu (angl. Research Managers) a financující agentury (angl. Funding Agencies). Uživatelsky atraktivní jsou také sekce nejnovějších zpráv týkajících se otevřeného přístupu, kalendář zachycující konference a semináře prosazující principy OA a nejnovější tweet z twitterového účtu OpenAIRE.

Webové rozhraní v plném rozsahu, včetně funkčních součástí a dynamických prvků, je lokalizováno výhradně do anglického jazyka.

The screenshot shows the OpenAIRE website homepage. At the top, there are social media icons (Facebook, Twitter, LinkedIn, YouTube, RSS) and links for NEWSLETTER, LOG IN, and REGISTER. The main navigation bar includes:

- OpenAIRE logo
- Home icon
- PARTICIPATE (DEPOSIT, JOIN)
- SEARCH (PUBLICATIONS, DATA, PROJECTS)
- STATISTICS (OA, PROJECTS, TOPICS)
- SUPPORT (FAQ, HELPDESK, GUIDES)
- OPEN ACCESS (IN EUROPE)

 Below the navigation bar, there are four main content boxes for different OpenAIRE versions:

- OpenAIRE basic**: ONLY OPEN ACCESS CONTENT, via driver oai set.
- OpenAIRE 2.0**: ONLY EC FUNDED CONTENT, via ec_fundedresources oai set.
- OpenAIRE 2.0 +**: OPEN ACCESS AND EC FUNDED CONTENT, via driver and ec_fundedresources oai set.
- OpenAIRE 3.0**: OPEN ACCESS AND/OR EC FUNDED AND/OR NATIONAL/OTHER FUNDED CONTENT, via openaire oai set.

 A search bar with a dropdown menu for 'Repositories' and a search icon is located below these boxes. The page also features four main user groups with their respective icons and lists of links:

- Researchers**: Deposit publications & data, Link research to funder/project, Download author's guide for OA.
- Data Providers**: Register & validate content, Get guidelines for interoperability, Enrich repository content.
- Research Managers**: Learn about EC & national policies, Create project reports, Get statistical information.
- Funding Agencies**: Learn how to embrace OA, Participate in OpenAIRE, Find aggregated statistics.

 At the bottom, there are sections for 'Latest News' (including a success story about disseminating research via OpenAIRE) and 'Mark your calendar' (with a list of events for March 10th and 19th). A blue banner on the right side of the bottom section promotes a 'Pilot Launched: Funding Open Access for Post-Grant FP7 Publications' and a 'Gold Open Access Pilot' for post-grant FP7 publications.

Obrázek 6 - Hlavní stránka webového rozhraní portálu OpenAIRE

9. Statistiky

Tabulka 1 nabízí dílčí statistické údaje o portálu OpenAIRE. K jejich vypracování bylo využito údajů zprostředkovaných přímo provozovateli portálu prostřednictvím propracovaného statistického nástroje i rozhraní pokročilého vyhledávání a prohlížení. Data jsou aktuální k 1. dubnu 2015.

Portál OpenAIRE záznamy sklízí nejen přímo z repozitářů při akademických institucích, ale i z databází agregátorů. Právě ti patří mezi největší poskytovatele dat, jak je uvedeno níže.

Nejvíce záznamů dokumentů pochází z Nizozemska¹⁵ (615 959); hojně jsou také zastoupeny Velká Británie (583 792), Německo (570 418) a Španělsko (549 825). Původem z České republiky je v portálu vyhledatelných 90 795 záznamů. Nejméně zastoupenou zemí je naopak Slovensko s pouhými 839 záznamy. Je nutné uvést, že do těchto statistiky nejsou započítány databáze agregátorů, pouze institucionální repozitáře a digitální archivy.

Z jazykového hlediska však převládá angličtina (6 634 011), dále španělština (497 892), němčina (325 970) a překvapivě také portugalština (299 465). Záznamů dokumentů v češtině je v systému zastoupeno prozatím 73 190.

Zdrojovým archivem s jednoznačně největším počtem sklizených záznamů je Europe PubMed Central (3 017 086). Značné množství záznamů článků (vedle záznamů periodik) je sklízeno také z DOAJ: Directory of Open Access Journals (1 755 162), arXiv.org e-Print Archive (1 009 948) a HAL-Inria (337 179). České repozitáře jsou do portálu zapojeny celkem 4, přičemž nejvíce záznamů poskytuje Digitální repozitář Národního úložiště šedé literatury.

| | |
|---------------------------------------------------------------------------|--------------------------|
| Celkový počet záznamů v indexu | 9 980 369 |
| Celkový počet záznamů dokumentů s otevřeným přístupem | 9 588 911 |
| Celkový počet záznamů souborů dat v indexu | 6 883 |
| Celkový počet zdrojových archivů | 579 |
| Počet záznamů pocházejících z České republiky | 90 795 |
| Převládající druh dokumentu | Článek |
| Převládající rok publikování dokumentu | 2012 (911 337) |
| Převládající jazyk dokumentu | Angličtina |
| Zdrojový archiv s nejvyšším počtem poskytnutých záznamů | Europe PubMed Central |
| Zdrojový archiv s nejvyšším počtem poskytnutých záznamů české provenience | Digitální repozitář NUŠL |
| Země s nejvyšším počtem záznamů | Nizozemsko |

Tabulka 1 - Statistické údaje o portálu OpenAIRE k 1. dubnu 2015

10. Vyhodnocení

Portál OpenAIRE je bezpochyby významným počinem v oblasti vývoje vyhledávačů a služeb nad digitálními archivy s otevřeným přístupem v celoevropském měřítku. Sklizení metadatových záznamů primárních dokumentů vysoké informační hodnoty, jejich sémantické obohacování o vazby na jiné záznamy a zprostředkování pestré nabídky návazných služeb z něj činí významný nástroj vědecké komunikace.

Jako nesporné přednosti portálu lze vyzdvihnout jak kvalitativní – graficky kvalitně zpracované, přívětivé a přehledné uživatelské rozhraní nabízející vyčerpávající informace o všech aspektech OpenAIRE včetně často kladených otázek (FAQ), příruček a newsletteru, velmi podrobné statistiky zahrnující možnosti znázornění formou grafů a tabulek

¹⁵ Prvenství Nizozemsku patří zejména díky repozitářům univerzit ve Wageningen, Eindhovenu a Utrechtu, které každý poskytují více než 100 000 záznamů.

a v neposlední řadě nadstavbové služby orientované na individuální uživatele i provozovatele repozitářů – tak kvantitativní charakteristiky – celkovou velikost indexu čítající více než 10 miliónů záznamů, celkový počet zdrojových archivů (téměř 600) a celosvětový záběr.

Na druhou stranu je nutné uvést několik nedostatků. Ty spočívají zejména ve vyhledávání informací. Rozhraní pokročilého vyhledávání je poměrně obtížně zpřístupnitelné (uživatel je nucen nejprve zvolit jednu z šesti výše uvedených kategorií; vyhledávání tedy není možné vztáhnout na všechny najednou) a třídění výsledků je užitím faset je (s jedinou výjimkou) omezeno na kategorie ryze formálního charakteru. Kvalitu vyhledávání také značně komplikuje fakt, že metadatové záznamy se liší svou podrobností a v mnoha případech postrádají klíčové prvky.

Literatura

CASTELLI, Donatella a Paolo MANGHI. The OpenAIRE Project – Open Access Infrastructure for Research in Europe. *Ercim News* [online]. January 2010, no. 80 [cit. 2015-04-01]. ISSN 0926-4981. Dostupný z: <http://ercim-news.ercim.eu/en80/es/openaire>.

D-NET LAB [online]. Athens (Řecko): University of Athens, 2006- [cit. 2015-04-01]. Dostupný z: <http://www.d-net.research-infrastructures.eu/>.

Introducing the functionalities. In: *OpenAIRE* [online]. OpenAIRE Consortium, 2009- [cit. 2015-04-01]. Dostupný z: <https://www.openaire.eu/functionalities/functionalities/intro-to-functionality/>.

MANGHI, Paolo et al. OpenAIREplus: the European Scholarly Communication Data Infrastructure. *D-Lib Magazine* [online]. September/October 2012, vol. 18, no. 9/10 [cit. 2015-04-01]. ISSN 1082-9873. Dostupný z: <http://www.dlib.org/dlib/september12/manghi/09manghi.html>.

OpenAIRE [online]. OpenAIRE Consortium, 2009- [cit. 2015-04-01]. Dostupný z: <https://www.openaire.eu/>.

OpenAIRE Monitoring. In: *OpenAIRE* [online]. OpenAIRE Consortium, 2009- [cit. 2015-04-01]. Dostupný z: <https://www.openaire.eu/overall-stats/statistics/infra-monitoring>.

Project factsheets. In: *OpenAIRE* [online]. OpenAIRE Consortium, 2009- [cit. 2015-04-01]. Dostupný z: <https://www.openaire.eu/about/general-information/openairefactsheet-40>.

TKAČÍKOVÁ, Daniela. Projekt OpenAIRE – Výzva a příležitost i pro Českou republiku. In: *INFORUM 2010: 16. konference o profesionálních informačních zdrojích, Praha, 25–27. 5. 2010* [online]. Praha: Albertina Icome Praha, 2010 [cit. 2015-04-01]. Dostupný z: <http://www.inforum.cz/pdf/2010/tkacikova-daniela-1.pdf>.

Zenodo [online]. Meyrin (Švýcarsko): CERN, 2009- [cit. 2015-04-01]. Dostupný z: <http://www.zenodo.org/>.