**MASARYKOVA UNIVERZITA**

**FAKULTA INFORMATIKY**



**DRYAD Repository – mezinárodní**

 **repozitář výzkumných dat**

URL: <http://datadryad.org/>

Projekt PV070

**Alena Rinaglová**

UČO: 438339

Brno, 3. ročník B-FI PLIN,

3. 12. 2017

# Charakteristika repozitáře

Dryad je digitální interdisciplinární repozitář neboli úložiště dat, které umožňuje přijímat práce akademiků z různých institucí. Zaměřuje se na oblast medicíny a biologie a umožňuje mimo jiné i archivaci, šíření a sdílení, jako ostatní repozitáře. Obsahuje výzkumné a akademické práce spojené s daty z oblasti přírodních a medicínských věd. Nachází se zde převážně recenzovaná data k přidruženým článkům, ale i data nerecenzovaných zdrojů, jako jsou např. dizertační práce nebo monografie. [1]

Dryad je současně nezisková organizace otevřená pro širokou veřejnost a také členská organizace řízená časopisy, vydavateli, vědeckými společnostmi a dalšími zúčastněnými stranami.

# Cíle repozitáře

Hlavním cílem projektu je poskytování infrastruktury a podpora opakovaného použití dat odborné literatury. Cílem společnosti Dryad je podpora světa, ve kterém jsou výzkumná data otevřeně k dispozici, integrována do odborné literatury a běžně znovu využívána k vytváření dalších znalostí. [2]

# Popis projektu

Idea Dryadu vznikla v květnu roku 2007 jako výsledek práce na workshopu National Evolutionary Synthesis Centre (NESCent) s názvem *„Digital data preservation, sharing, and discovery: Challenges for Small Science Communities in the Digital Era.“* [3]

Zveřejnění projektu proběhlo v lednu v roce 2008, tedy necelý rok po jeho vzniku.

## Samostatný projekt Dryad UK

Dryad UK je jednoletý projekt, jenž vznikl díky finanční podpoře Britské knihovny a Oxfordské Univerzity v Anglii ve spolupráci s centrem digitálních služeb a NESCentem fungujícím do roku 2011, kdy pomohl zřídit Dryad jako trvalý mezinárodní repozitář pro výzkumné datové sady biologických věd spojených s vydáváním časopisecky publikovaných článků. [4]

## Financování projektu

Počáteční financování pokryla Národní vědecká nadace National Evolutionary Synthesis Centre spolu s dalšími americkými partnery.[[1]](#footnote-1)

Ačkoliv je Dryad nezisková organizace, od roku 2013 společnost zavedla takzvané *submission fees*, tj. poplatky za podání od členů a předkladatelů dat, díky nimž je schopna poskytnout volný přístup k údajům. Dryad vybírá tzv. *Data Publishing Charg*e (DPC) ve výši 120 USD na jeden datový balíček (ke splatnosti po jeho zveřejnění). Další možné poplatky se vztahují k datům s kapacitou vyšší, než je 20 GB. Poplatky byly navrženy tak, aby pokryly základní funkce pro uchování dat.

Lidé ze zemí s nízkou a středně vysokou ekonomickou úrovní podle World Bank si mohou zažádat o výjimku a placení poplatku se vyhnout. Česká a Slovenská republika se v současné době řadí do skupiny s vysokou ekonomickou úrovní.[[2]](#footnote-2) [5]

Další možnost pro jednotlivce, jak neplatit DPC je využití sponzora (např. časopisu), ovšem pouze v případě nepřekročení vymezené kapacity.

# Fungování instituce

Vnitřní politiku výzkumného repozitáře řídí dvanáctičlenná správní rada, kterou volí její členové. Členem může být nezávislý časopis, společnost nebo vydavatel, vzdělávací instituce, knihovna, financující organizace nebo jiná organizace podporující misi Dryad.

Společnost koordinuje politiky sdílení dat a snaží se podporovat dlouhodobou udržitelnost svého úložiště. [6]

Dryad najdeme také v oficiálním seznamu registru výzkumných repozitářů re3data.org (Registry of Research Data Repository).

Dryad je založen na otevřeném úložišti - tzv. open source, softwaru DSpace. Software byl vyvinut Massachusettským technologickým institutem a společností Hewlett-Packard v roce 2002, což je přibližně stejná doba, kdy vznikl samotný repozitář.

Repozitářový software DSpace slouží jako digitální archivní systém pro akademické, neziskové organizace, které vytvářejí přístupné digitální repozitáře [7]. Je zaměřený na dlouhodobé ukládání a uchovávání digitálního obsahu. Některé jeho funkce byly speciálně upraveny pro Dryad.

Webové stránky jsou vytvořeny v HTML v souladu s dodatkem v americkém zákoně z roku 1973 o přístupnosti datového obsahu s názvem *Section 508.[[3]](#footnote-3)* Tento zákon upravuje technologický způsob tvorby webových stránek tak, aby byly přístupné pro osoby s tě-lesným postižením a pomohly zmírnit bariéry v informačních technologiích pro tyto osoby. Zákon se vztahuje na všechny federální společnosti, které pracují nebo používají elektronické informační technologie. [8]

Dryad poskytuje k nahlédnutí tzv. *Voluntary Product Accessibility Template* (VPAT) z roku 2013, který slouží jako důkaz, že systémový software i hardware splňuje požadavky ze zákona. Dokument obsahuje seznam kritérií stanovených šablonou VPAT, která jsou Dryadem doplněna o informací o jejich podpoře. [[4]](#footnote-4)

Při tvorbě přístupného softwaru může docházet k mnoha problémům. Např. lidé trpící barvoslepostí mohou mít problém rozlišovat různé informace z webových stránek pouze podle odlišení jejich barev, naopak pro zdravé jedince je tato metoda spolehlivá a přehledná. Další problém nastává, pokud je prohlížení softwaru možné pouze s použitím počítačové myši a není podporováno jeho ovládání pomocí klávesnice, kterou používá mnohem více lidí.

Dalším možným problémem pro osoby se zrakovou vadou nebo slepotou je chybějící podpora sdělování podstatných informací poslechovou cestou. Jestliže produkt splňuje preference uživatele ve výše zmíněných problematických bodech, považuje se za dostačující a splňující podmínky Section 508. [9]

## Podmínky šíření dat a citace zdrojů z repozitáře

Podmínky pro šíření a citaci zdrojů jsou stanoveny za pomoci licencí společnosti Creative Commons a Science Commons. Autor svým potvrzením dává souhlas k legálnímu užití jeho díla na základě podmínek, které si sám stanovil, popř. vybral již sestavenou licenci v souladu se zákony dané země.

Dryad podporuje licenci Creative Commons Zero, což znamená, že uživatel, který používá data z repozitáře, se vzdává práv na autorství a všech souvisejících práv pro celou zemi ve prospěch veřejné domény v rozsahu, jež povoluje zákon. Uživatel smí dílo šířit, upravovat

a kopírovat, a to bez nutnosti žádat o souhlas původního autora. [10]

# Poskytnutá data a metadata

Dryad zapisuje metadata pomocí standardu pro popis digitálních objektů Dublin Core (DC). Záměrem takto zvoleného standardu je usnadnit vyhledávání elektronických zdrojů. (Včetně www stránek.) [11]

Jak popisuji v oddíle o podmínkách šíření a citaci zdrojů pro uživatele, na veškerá data, která jsou v repozitáři uložena se vztahuje licence Creative Commons Zero, tedy úplné vzdání práv na autorství.

Ze statistiky z roku 2014 se v repozitáři vyskytuje na 20 000 článků, které byly staženy přibližně 600 000krát. Počet autorů se pohybuje okolo 23 000 a počet vědeckých časopisů okolo 360. [12]

**Ke dni 15. 11. 2017 vypadá statistika následovně:** [13]

|  |
| --- |
| **Statistika (2017)** |
| **–** | **Celkově** | **Měsíčně (30 dní)** |
| **Datové balíčky** | 19 600 | 422 |
| **Datové soubory** | 62 900 | 1 300 |
| **Časopisy** | 650 | 145 |
| **Autoři** | 69 800 | 6 700 |
| **Počet stažení** | 2 179 629 | 45 600 |

Při porovnání obou statistik, mezi nimiž je časový rozestup přibližně 4 roky, vidíme, jak rychle se Dryad rozšiřuje.

##

## Identifikace datových objektů uložených v repozitáři

Veškerá data v repozitáři mají přidělen identifikátor. Dryad používá tzv. Data Object Identifier (DOI) skrze DataCite. Jedná se o trvalý, globálně jedinečný identifikátor objektu, který je nezbytnou součástí citovaných dat. [14]

##  Postup zpracování dat od uživatele

Data může uživatel vkládat dvěma způsoby. Za podpory vybraného časopisu anebo nahráním dat do depositáře přímo. Postup zpracování dat se v obou případech mírně liší.

Pokud se uživatel rozhodne nahrát svá data s pomocí vědeckého časopisu, musí nejprve dostat schválení od editora, poté data putují ke správci dat, automaticky se přidělí Data Object Identifier (DOI) a nakonec se data publikují. Postup bez podpory časopisu je takřka stejný, místo editora data kontroluje pouze správce. [15]

Dryad nabádá předkladatele dat, aby své materiály vkládali ve formátech ASCII nebo HTML pro jejich snadnější uložení a opětovné užití. [16]

# Externí kontakty a dokumentace

* Dotazy veřejnosti: help@datadryad.org [[5]](#footnote-5)
* Náměty na zlepšení, popř. kritika: <http://datadryad.org/feedback>
* Pro zájemce o členství: director@datadryad.org
* Vlastní blog na: <https://blog.datadryad.org/>
* Veškerá dokumentace: <http://wiki.datadryad.org/Main_Page>

# Zhodnocení projektu a jeho přínosu

Dlouhodobá ochrana vědeckých dat a jejich sdílení je podstatnou součástí vědeckého poznání, které díky tomu můžeme opětovně využívat k dalším výzkumným účelům. Pro komunitu vědců se jedná o nenahraditelný zdroj kvalitních dat. Díky uloženým informacím mohou vědci dále zkoumat, potvrzovat nebo vyvracet již zavedené hypotézy ve vztahu k vlastní práci v dané oblasti.

V archivovaných datech mohou být implicitně obsaženy i odpovědi na otázky dalších vědců.

Uložení takových dat také umožňuje zabránit jejich ztrátě a zaručit udržení vysoké kvality informací v nich uložených v trvanlivější formě, než na jiných (analogových) nosičích. (Vědci často o svá data přijdou pouhým zálohováním nebo jeho nevhodným umístěním na nosič.)

Repozitář Dryad umožňuje uživatelům využívat uložených dat k výše zmíněným účelům, a tím přispívá k šíření vědomostí mezi lidmi. Přínos projektu vidím právě ve schopnosti nabídnout volně dostupné informace z medicínského prostředí a biologie za účelem rozšiřovat poznání mezi lidmi.

# Seznam literatury/zdrojů, URL

[1] Hrabal, Jan. In: *Repozitáře vědeckých dat*. [online]. [cit. 2017-11-17]. Dostupné z: <http://ltp.knihovna.cz/?p=385>

[2] Data Dryad: In: *The organization: Overview* [online]. [cit. 2017-11-17]. Dostupné z: <http://datadryad.org/pages/organization>

[3] Wikipedia: In: *Dryad Repository* [online]. [cit. 2017-11-17]. Dostupné z: [https://en.wikipedia.org/wiki/Dryad\_(repository)](https://en.wikipedia.org/wiki/Dryad_%28repository%29)

[4] The Digital Curation Centre: In: *Dryad UK* [online]. [cit. 2017-11-28]. Dostupné z: <http://www.dcc.ac.uk/projects/dryaduk>

[5] Word Bank: In: *Word Bank Country and Lending Groups* [online]. [cit. 2017-11-28]. Dostupné z: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519>

[6] Wikipedia: In: *Dryad Repository – Governance, history and funding* [online]. [cit. 2017-11-17]. Dostupné z: [https://en.wikipedia.org/wiki/Dryad\_(repository)](https://en.wikipedia.org/wiki/Dryad_%28repository%29)

[7] DSpace: In: *About DSpace* [online]. [cit. 2017-11-17]. Dostupné z: http://www.dspace.org/introducing

[8] Wikipedia: In: *Section 508 Amendment to the Rehabilitation Act of 1973* [online]. [cit. 2017-11-28]. Dostupné z: <https://en.wikipedia.org/wiki/Section_508_Amendment_to_the_Rehabilitation_Act_of_1973>

[9] Wikipedia: In: *Section 508 Amendment to the Rehabilitation Act of 1973* [online]. [cit. 2017-11-28]. Dostupné z: <https://en.wikipedia.org/wiki/Section_508_Amendment_to_the_Rehabilitation_Act_of_1973>

[10] Creative Commons: In: *Public Domain Dedication* [online]. [cit. 2017-11-28]. Dostupné z: <https://creativecommons.org/publicdomain/zero/1.0/>

[11] Wikipedia: In: *Dublin Core* [online]. [cit. 2017-12-4]. Dostupné z: <https://cs.wikipedia.org/wiki/Dublin_Core>

[12] Greenberg, Jane. In: *DataDryad.org and the interoperability continuum* [online]. [cit. 2017-11-28]. Dostupné z: <http://slideplayer.com/slide/8766910/>

[13] Data Dryad: In: *Stats* [online]. [cit. 2017-11-28]. Dostupné z: <http://datadryad.org/>

[14] Data Dryad: In: *FAQ* [online]. [cit. 2017-11-28]. Dostupné z: <http://datadryad.org/pages/faq#depositing-howto-citedata>

[15] Greenberg, Jane. In: *DataDryad.org and the interoperability continuum* [online]. [cit. 2017-11-28]. Dostupné z: <http://slideplayer.com/slide/8766910/>

[16] Data Dryad: In: *Accessibility of Dryad contents* [online]. [cit. 2017-11-28]. Dostupné z: <http://datadryad.org/pages/repository>

# METADATA:

<dc:creator>Alena Rinaglová</dc:creator>

<dc:title>DRYAD repository </dc:title>

<dc:subject>Mezinároní repozitář výzkumných dat</dc:subject>

<dc:description>Esej pojednává o neziskové organizaci a digitálním datovém úložišti Dryad, který je zároveň členskou organizací, která se specializuje na volné šíření a bezpečné ukládání vědeckých publikací a článků od akademiků v oblasti biologie a medicíny. <dc:description>

<dc:identifier><http://datadryad.org/><dc:identifier>

<dc:keywords> repository, datacite, creative commons, science commons, digital libraries<dc:keywords>

<dc:type>Text</dc:type>

<dc:format>Public</dc:format>

<dc:language>CZ</dc:language>

<dc:date>2017-11-28</dc:date>

1. výše grantu se mi nepodařila zjistit, záznam z workshopu již není dostupný [↑](#footnote-ref-1)
2. s příjmy nad 12 200 USD [↑](#footnote-ref-2)
3. jedná se o standard, celý název: *Section 508 Amendment to the Rehabilitation Act of 1973* [↑](#footnote-ref-3)
4. do seznamu patří například: metoda, která povoluje uživateli přeskakovat opakující se navigační odkazy; podpora uživatele používat asistivní technologie pro elektronické formy informací dostupných on-line… atd.

(VPAT Dryadu k náhledu zde: <http://datadryad.org/themes/Mirage/docs/Dryad-VPAT.pdf> ) [↑](#footnote-ref-4)
5. pravděpodobně jim můj e-mail nepřišel nebo spadl do spamu, protože jsem nedostala odpověď [↑](#footnote-ref-5)