

Otevřený přístup k vědeckým výsledkům a datům

(digitalizace a přístup k vědeckým poznatkům)

Jiří Zlatuška
Masarykova Univerzita

Akademické fórum
Praha 21. 3. 2013

Knihovny jako centra vědy

- Kombinace velké knihovny s komplexem vědeckých institucí je známa z cca 300leté historie Alexandrijské knihovny
- Až do nedávné doby byly knihovny výhradně úložištěm v minulosti akumulovaných publikací a vědomostí uložených v nich
- Podpora hierarchicky uspořádaného systému vědění, rozvoj vědění v jednotlivých disciplínách, akademický kontext tvorby nových poznatků (spojeno s tzv. módem 1 produkce vědění)

Standardizace komunikace vědeckých znalostí od 17. století

- Vynález knihtisku a nástup osvícenectví
- Učené společnosti diskutující nové myšlenky
- 1655 založeny časopisy *Journal des sçavans* ve Francii a *Philosophical Transactions of the Royal Society* v Anglii (v Royal Society včetně užití recenzního řízení na bázi peer review)
- Standardy vědecké výměny a testování, Boyle používá v roce 1645 termín „Invisible College“ pro vazby mezi vědci jdoucí mimo jejich instituce a místa působení
- Do konce 18. století začala vycházet víc než tisícovka vědeckých časopisů (Derek de Solla Price uvádí nutnost vzniku časopisů věnovaných abstraktům v roce 1830)
- Rozvoj konceptu autorských práv

Koncentrace

- Budování univerzitních center
- Knihovny spojené těmito centry nebo tvořící spojovací článek mezi jednotlivými prvky (Oxford)
- V americkém prostředí symbol identity podobný divadlu u nás (Library of Congress vs. Národní divadlo)
- Knihovní katalogy jedním z prvních příkladů uchování a zpracování objemných datových souborů (MARC v 60. letech)

Kontext a decentralizace ve 20. století

- Třetí vlna (A. Toffler)
- Komunikace
- Síťové struktury
- „adhokracie“
- Rozbíjení hierarchických struktur

„New Invisible College“ (Wagnerová, Brookings Inst.)

- **Sítě** vazeb mezi vědci, institucemi a projekty; konference a pracovní setkávání, nová dynamika vystupující decentralizovaně ze vzájemných spoluprací
- **Vznik** nových myšlenek kombinací a rekombinací lidí a vědění, samoorganizace spolupracujících vědců do skupin, informace, nová spojení a nové příležitosti vytvářejí ekosystém (nikoli organizovanou entitu) ve kterém vzniká nové vědění

Nové metodologie vědy

- Tradiční vědecké metody provádění *experimentu* a *teoretického bádání*
- Výpočetní technika doplňuje *simulaci* jako legitimní metodologii doplňující klasické dvě; objektem provádějící závěry jsou i programy provádějící simulaci
- Analýza a předzpracování masivně velkých datových souborů umožňuje čtvrté paradigma *syntézy hypotéz* ze závislostí nalezených v datech, namísto obvyklého formování hypotézy a jejího testování; objektem zde jsou primární data, jejich zpracování a následná analýza
- Obě „nové“ metodologie výrazně mění způsob, obsah i rozsah prezentace podkladových informací k výsledkům jejich pomocí dosažených

„Staré“ a „nové“

- Trendy změn je možné nahlížet jako evoluci paradigmat, ale také jako důsledek technických nástrojů, které jsou k dispozici
- Dynamizace, adaptace, zplošťování struktur
- Komunikační sítě, digitalizace a nástroje zpracování informací

Digitální knihovna

- Jedná se o systém přístupu k digitálním datům, nikoli místo nebo symbol moci či vlivu.
- Přístup a vyhledávání uložených objektů:
 - textů, knih, časopisů
 - audiovizuálních materiálů,
 - výzkumných dat,
 - softwaru,
 - vazeb mezi objekty (například článkem a daty).
- Uspořádání podle ad hoc kritérií
- Přístup bez geografických omezení
- Digitální knihovna převyšuje „starou“ v podobném smyslu jako „učící se“ poskytující služby převyšuje „industriální“ továrnu

Digitální knihovna pružnost funkcí

- Vytváření znalostí
- Zveřejňování
- Archivace
- Předávání
- Použití
- Podobně jako přechod z módu 1 do módu 2 znamená digitální knihovna mnohostrannou pružnost a setření hranic mezi tradičními rolemi.
- Klientem přestává být pouhý „čtenář“, stává se jím výzkumník, kterému je poskytována potenciálně mnohem komplexnější služba.

Role knihoven v digitální éře

- Služba ve výzkumu, výuce i publikačních činnostech (vytváření fondů, poskytování nástrojů, zpřístupňování všeho na pracovní stůl, technická expertíza v publikačních projektech, podpora infrastruktury)
- Služba potřebám studentů (nové fyzické i virtuální prostory k práci jednotlivců i skupin, asistence studentům v jejich proměně ve vědecké pracovníky)
- Úschova fondů v digitální podobě
- Snižování nákladů spojených s akademickými činnostmi (alternativní metody vydávání, asistence při sjednávání smluv)
- Podpora akademické komunikace (robustní fondy přístupné on-line, výzkumné prostředí podporující potenciál akademiků, umožňování snižovat cenu publikování výsledků a mít nad ním větší kontrolu, rozvoj infrastruktury vedoucí k využití multimédií)
- Směs činností tradičně přisuzovaných knihovnám s činnostmi řazenými do činnosti vydavatelství

University Publishing in a Digital Age, The Ithaka Report, 2007

Sítě a otevřené vztahy

- Paradigma sítí
- Otevřenost a mnohost možných interakcí
- Kombinováno s požadavky transparence
- Minimalizace komerčních zájmů vydavatelů publikací šířících znalosti veřejného zájmu a pořízené za veřejné prostředky

Otevřený přístup

- Otevřený přístup k vědeckým publikacím lze chápat jako logický důsledek trendů stavějících na vztazích, komunikaci a propojování
- Významným faktorem jsou finance a oddělování tržního prostředí od netržního
- Preference vědeckých poznatků jako veřejného statku
- Ovlivnění přístupu k systému, ve kterém se odehrává vědecká komunikace
- Ovlivnění vlastnictví a práva nakládání s vědeckou produkcí

Budapeštská iniciativa otevřeného přístupu

- Budapest Open Access Initiative (BOAI) – na konferenci Open Society Institute v prosinci 2001
- Definice: *By "**open access**" to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.*

Open Access klíčovým prvkem ERA (ERA Communication - červenec 2012)

5. Optimal circulation, access to and transfer of scientific knowledge including via digital ERA

Objective: to guarantee access to and uptake of knowledge by all.

- **Open Access to publications and data** are central to ERA
- **Digital ERA** needed to maintain Europe as hub of excellence

• So far, libraries and SMEs have difficulties subscribing to journals due to **high costs** and budgetary constraints

Examples of proposed actions for Member States, Stakeholders and Commission:

- To Promote Open Access
- To Foster Knowledge Transfer

"Scientific Information in the Digital Age" (2012)

COMM + Recommendation

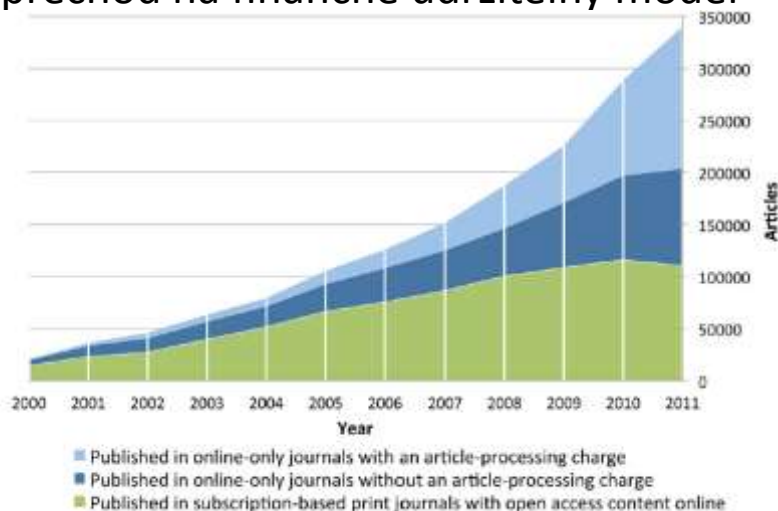
Z dokumentů popisujících směr vývoje European Research Area

The Commission will:

- **Establish open access to scientific publications as a general principle for all EU funded projects in Horizon 2020.** For research data, develop a flexible approach that takes into account different scientific areas and business-related interests.

(http://ec.europa.eu/research/era/optimal-circulation_en.htm)

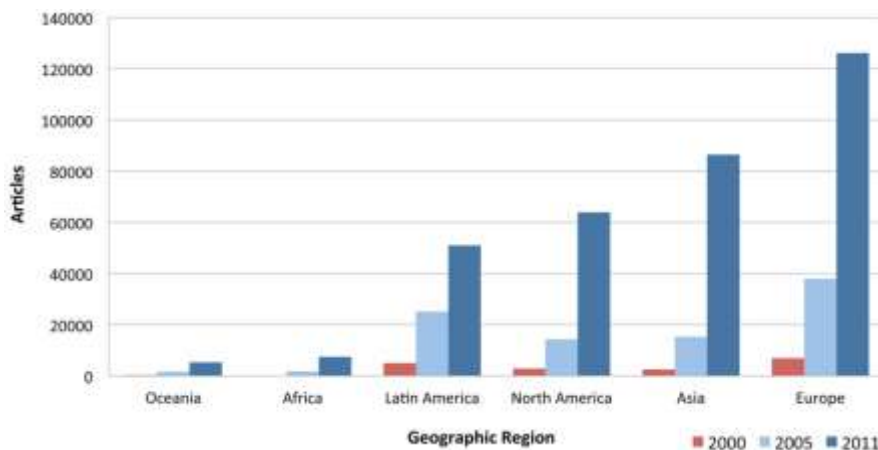
Výrazná akcelerace Open Access publikací a přechod na finančně udržitelný model



Annual volumes of articles in full immediate open access journals, split by type of open access journal.

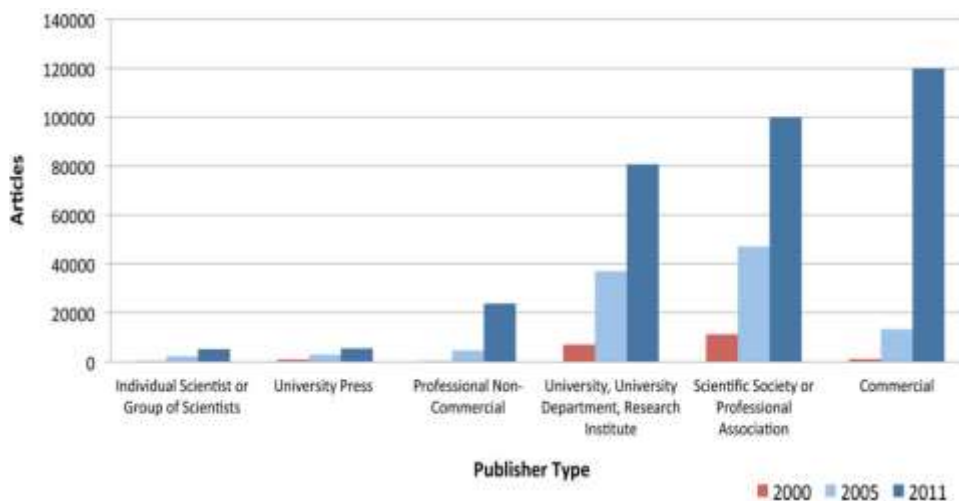
Zdroj: Laakso and Björk *BMC Medicine* 2012, **10**:124
<http://www.biomedcentral.com/1741-7015/10/124>

Zvětšující se podíl evropských vydavatelů v režimu Open Access



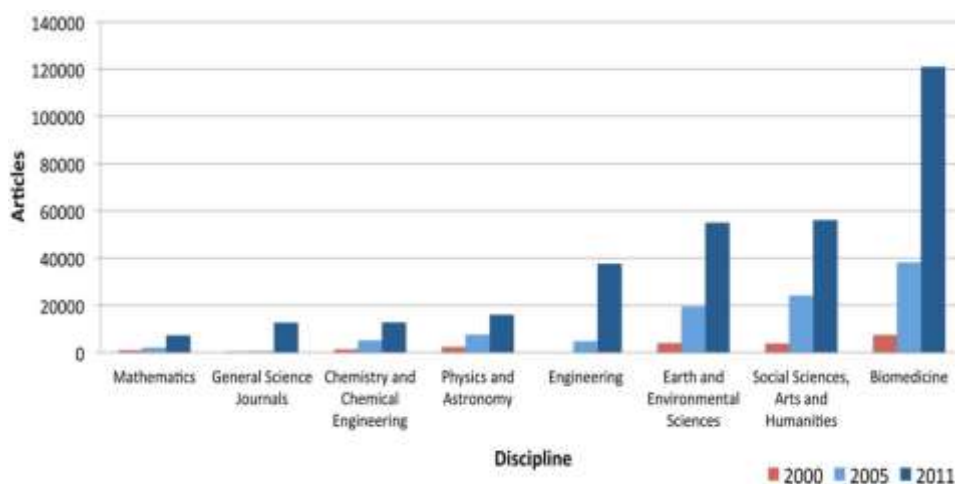
Zdroj: Laakso and Björk *BMC Medicine* 2012, **10**:124
<http://www.biomedcentral.com/1741-7015/10/124>

Podíl komerčních vydavatelů v režimu Open Access



Zdroj: Laakso and Björk *BMC Medicine* 2012, **10**:124
<http://www.biomedcentral.com/1741-7015/10/124>

Nárůst publikací Open Access se nevyhýbá žádným oborům



Zdroj: Laakso and Björk *BMC Medicine* 2012, **10**:124
<http://www.biomedcentral.com/1741-7015/10/124>

„Science as an Open Enterprise“ Report (The Royal Society, červen 2012)

- Sdílení a zpřístupňování vědeckých informací
- Vypracování metod nakládání s vědeckými informacemi tak, aby v maximální míře sloužily inovativnímu a produktivnímu výzkumu, který odráží veřejný zájem
- Případové studie užívání vědeckých dat
- Užívání repozitářů vědeckých publikací a dat

Nové způsoby publikování vědeckých výsledků a související úloha knihoven

- Zpřístupňování plných textů vědeckých prací, zdrojových dat a algoritmů zpracování v repozitářích (viz např. „Executable paper Grand Challenge“, <http://www.executablepapers.com>)
- Obecně netriviální problém vzhledem k potřebě anonymizace osobních nebo dokonce citlivých údajů, které mohou být součástí např. medicínských dat
- Reprodukce výsledků a usnadnění procesu peer review (hovoří se i o dvojitém slepém posuzování u elektronických publikací)
- Článek ve vědeckém časopise může být jen formou reklamy, upoutávkou na výsledky, které jsou v repozitáři

Knihovny provozující repozitáře

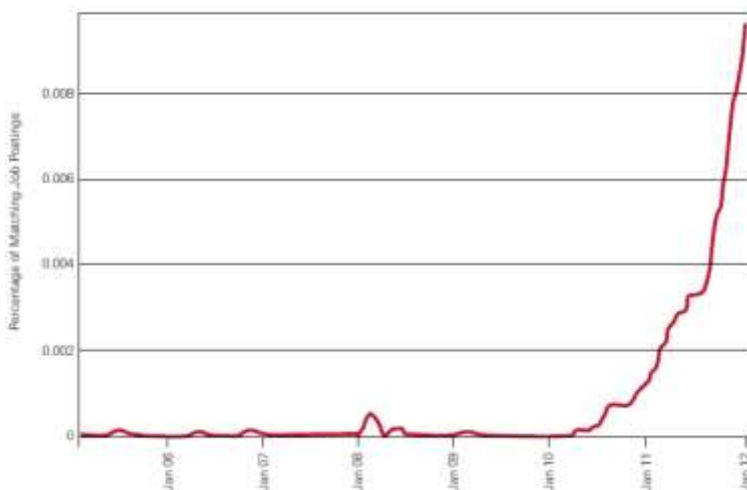
- Správa a zpřístupňování masivních objemů dat
- Specializace „datových vědců“ jako extenze profilu knihovníka, možná i součást obecné „vědecké gramotnosti“?
- Dodržování otevřeného přístupu jako standardu s odlišením nutných výjimek
- Transparentní zásady práce správců repozitářů
- Udržování přístupnosti dat a zřetelného systému zajišťujícího navigaci v ukládaných údajích
- Dodržování omezení plynoucích z ochrany soukromí
- Zásady pro aktivní užití dat, resp. vypracování strategií, které zabraňují ztrátě později ještě využitelných dat

Nové úlohy univerzitních knihoven

- Odpovědnost za provoz repozitářů
- Znalosti pomáhající vědcům data z repozitářů používat
- Podpora efektivní práce s vědeckými informacemi vystavenými on-line
- Zabezpečování uvolňování dat podle institucionální politiky
- „Datoví vědci“ kombinující informatiky a knihovníky – nejen věc reálné poptávky, ale i příležitost nového profilu studia s tímto uplatněním

„Datový vědec“ - nová žádaná profese

b Job adverts using the phrase 'data scientist' on Indeed.com



Zdroj:
Science as
an Open
Enterprise,
RAS 2012

Kvalita v prostředí otevřeného přístupu - práce knihoven významným prvkem

- Pro hodnocení univerzit je doporučení hodnotit otevřenost přístupu k datům se stejným významem jako data o publikovaných článcích
- Práce na standardech dobré praxe
- Podpora kolaborativního výzkumu nad efektivně sdílenými daty
- Nové způsoby přípravy vědců pro práci s těmito daty a souvisejícími nástroji

Nové modely financování publikační a recenzní činnosti

- Alternativy k financování cestou předplatného
- Podíl knihoven na efektivním institucionálním přístupu
- „Obrácený model“?
 - knihovny pečující o zpřístupňování výsledků vlastní instituce, nejen knih nakoupených zvenčí
 - vydavatelé časopisů poskytující službu šíření výsledku a recenze, nikoli primárně prodej tiskovin s výsledky

Závěry pro Českou republiku

- Otevřený přístup není v ČR nijak reflektován v národních politikách
- Politické postoje na úrovni, která by neměla obstát ani u Křováků – stačí srovnat aktuální návrhy omezující obecný přístup k informacím (plus pasáž z projevu ministra Blažka na volebním sněmu ODS) a srovnat s *Open Data White Paper* britské vlády z června 2012, který je nejen kompatibilní s návrhy The Royal Society, ale rozšiřuje je na obecný přístup k informacím (pod heslem „Building Transparent Society“)
- Neexistuje „oficiální“ obdoba doporučení ze zprávy *Accessibility, sustainability, excellence: how to expand access to research publications* pracovní skupiny vedené Janet Finchovou s doporučeními pro financující agentury požadovat open access dosažených výsledků, což má v prostředí, kde nejsou univerzity vedeny leadery, nýbrž účetními sledujícími pokyny vrchnosti, výrazný brzdící efekt

Studie Evropské komise „National Open Access and Preservation Policies in Europe“

(EC/ERA/7FP, doi:10.2777/74027, 2011)

- GA ČR podepsala Berlínskou deklaraci, ale nemá strategii šíření a přístupu
- Navzdory neexistující národní strategii dělá aktivní kroky AV ČR a některé univerzity
- Aktivní fond podpory ustanoven v AV ČR
- Ze zákona existuje pouze odevzdávání povinných výtisků, bez práce s digitálními informacemi
- Dobrým příkladem v ČR je systém **theses.cz**
- Repozitáře řešeny jen na úrovni jednotlivých institucí (AV ČR a některé univerzity), v databázi RIV se neshromažďují plné texty
- Projekt WebArchiv cílí na archivaci části webu, ale je závislý na smlouvách s jednotlivými vydavateli
- Autoři nejsou motivováni využívat svých práv např. formou licence Creative Commons

Studie Evropské komise „National Open Access and Preservation Policies in Europe“

(EC/ERA/7FP, doi:10.2777/74027, 2011)

- Knihovníci neuspěli v tlaku na snížení DPH na elektronické publikace
- Ad hoc podpora přístupu k elektronickým zdrojům, avšak bez participace pražských institucí
- AV ČR má speciální fond podpory publikací s otevřeným přístupem
- Neexistuje obecná podpora šíření výsledků
- Digitalizace kulturního dědictví jen na národní úrovni
- Národní repozitář šedé literatury v NTK (s výhledem zahrnutí kulturních materiálů)
- V autorském zákoně chybí úprava digitálního uchování výzkumných výsledků
- Absence koordinace aktérů na národní úrovni

Výhledy celosvětově

- Ve Spojených státech deklaroval prezident Obama záměr přejít na model otevřeného přístupu k vědeckým výsledkům
- V EU plány na Horizon 2020 předpokládají přístup v režimu Open Access k vědeckým výsledkům a v možné míře i ke zdrojovým datům
- Ve Velké Británii plán zohlednit požadavek Open Access k výsledkům by měl být zohledněn i v hodnocení výzkumu REF
- Výrazný rozpor s komerčními zájmy uvád

Masarykova univerzita v hodnocení dostupnosti publikací v žebříčku Webometrics 2013

Druhé místo celosvětově v parametru **Openess** (*The global effort to set up institutional research repositories is explicitly recognized in this indicator that takes into account the number of rich files published in dedicated websites according to the academic search engine Google Scholar.*)

<u>ranking</u>	<u>University</u>	<u>Det.</u>	<u>Country</u>	<u>Presence Rank*</u>	<u>Impact Rank*</u>
1	Harvard University			51	1
214	Masaryk University in Brno / Masarykova Univerzita v Brně			325	450
19	(2) Universidade de São Paulo USP			31	53
205	Universidade Federal de Santa Catarina UFSC			272	379
45	University of Helsinki / Helsingin yliopisto			66	147

- Univerzity / AV ČR a jejich knihovny jsou zřetelně lépe připraveny než stát (i když např. na MU došlo v posledních měsících pod tlakem PŘF, LF i FI k ústupu od institucionální politiky budování repozitáře)
- Bez národní politiky a zohlednění v přiřazování finančních prostředků bude postup zbytečně pomalý
- Nutná je výrazná podpora nových aktivit knihoven na jednotlivých institucích
- Reálnost podpory závisí mimo jiné na jasné identifikace veřejného přínosu otevřených modelů práce s informacemi a vyloučení omezení, pro která se často vyjadřují např. někteří průmyslníci