

Způsoby šíření a využívání elektronických závěrečných prací

Jak nakládat s VŠKP, aby byly přístupné a dohledatelné

Praha, 19. října 2016

Meinhard Kettler, ProQuest

Přehled

- Dostupnost a dohledatelnost elektronických dizertací
 - dohledatelnost - výzvy
 - dostupnost tištěných vs dostupnost online prací
- Význam post-graduálního výzkumu
 - hodnota PhD prací pro studenty a výzkumníky
 - studie užití - London School of Economics
- Text mining elektronických dizertací

Pohled z vnější strany akademické obce:

ProQuest, dlouholetý partner institucí v oblasti šíření a uchování dizertací a VŠKP

Dostupnost a dohledatelnost

- Posun od tištěných k elektronickým
→ mnoho způsobů, jak získat plné texty dizertací a dalších VŠKP.
→ miliony dizertací na webu.
- Nová politika a workflow pro 'born-digital' práce

Ale:

- všechny závěrečné práce z celého světa nejsou dostupné online
- dostupnost ≠ dohledatelnost
nalezení relevantních VŠKP může být obtížné
- jak velká je neviditelná část ledovce?

Proč jsou některé dizertace nevyhledatelné...

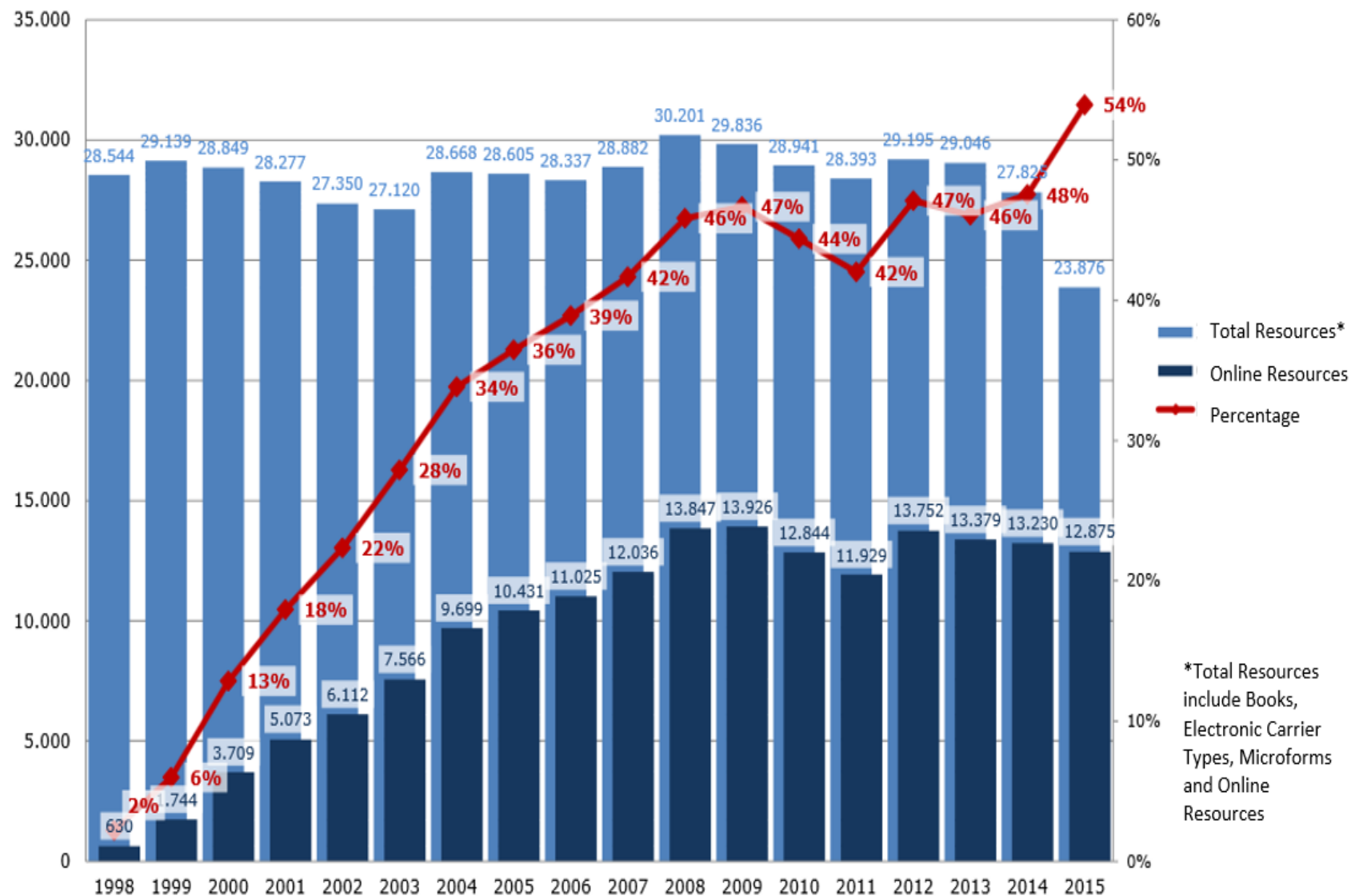
Možné důvody

- Rozhodnutí autora (schválené institucí)
- Časové embargo na zveřejnění publikace dané
 - rozhodnutím autora
 - třetí zainteresovanou stranou
- Monografie publikované pod jiným názvem
- Odevzdané pouze v tištěné verzi, ne v elektronické
- Nejsou uloženy v institucionálním repozitáři ani neexistuje záznam v katalogu
- Obtížná rozeznatelnost kvůli jazyku
- Limitované vyhledávání, nefunkční odkazy
- ...a další

Změny v pozadí

- Ne všude jsou definovány celostátní nebo alespoň univerzitní standardy (akademické svobody)
- Různé legislativní problémy – copyright, autorem nastavené embargo, (plagiátorství)
- Ani OA publikace nejsou vždy dobře k nalezení: je nutné indexovat a harvestovat obrovské a stále rostoucí množství prací
- Stále vysoké procento dizertací odevzdaných pouze v tištěné verzi

Tištěné vs Online - Německo

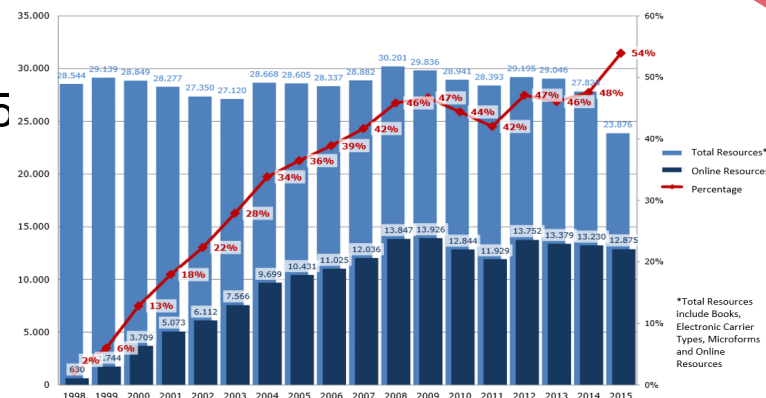


Online Resources Share of Total PhD Dissertations and Habilitations in German National Library Collection by Publication Year (updated 2 March, 2016). Deutsche Nationalbibliothek 2016, <http://www.dnb.de/DE/Wir/Kooperation/dissonline/dissonlineStatistik.html>

Tištěné vs Online – příklady

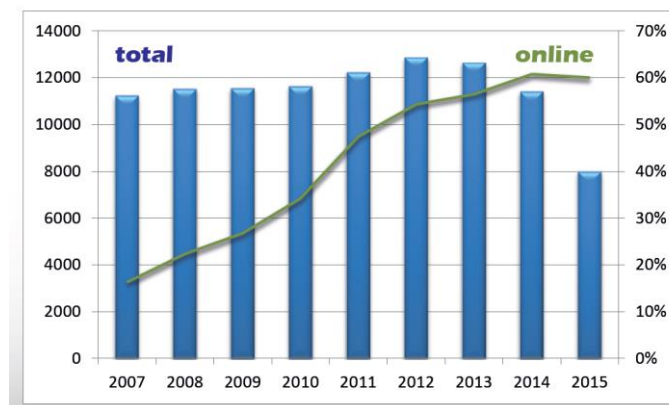
- Německo

- 54 % elektronických dizertací v roce 2015
- Pomalu rostoucí podíl



- Francie

- kolem 60 %
- Díky vyhlášce by se mělo přejít na odevzdávání dizertací pouze v elektronické podobě povinné do roku 2018



Prost, H el ene et al. French Electronic Theses and Dissertations in Europe – A Scientometric Approach.
<https://etd2016.sciencesconf.org/98998>

- USA

- 90+ % odevzdaných dizertací je v elektronické podobě (zdroj obrázků - ProQuest)

Hodnota a význam elektronických dizertací

Dizertace a další VŠKP ...

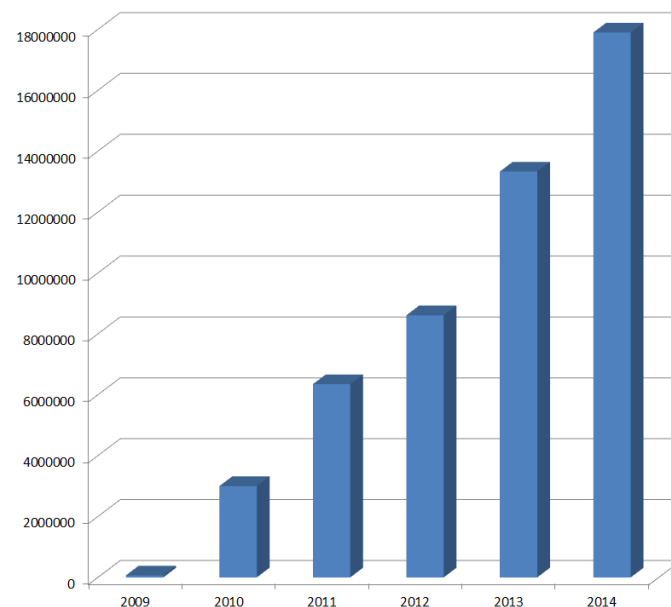
- pokrývají širokou škálu výzkumných témat včetně neobvyklých témat
 - na rozdíl od článků pokrývají témata komplexněji
 - slouží jako zdroj metodologických a teoretických přístupů
 - pomáhají identifikovat strategické mezery v literatuře
 - poskytují doporučení pro budoucí výzkum
 - slouží jako aktuální přehledy literatury
 - slouží jako příklad odpovídající úrovně detailnosti a stylu psaní
- = představují hodnotný obsah pro výzkumníky a doktorandy
- zdají se být někdy opomíjeny
 - zasloužily by si získat větší pozornost (pomoc autorům a institucím)

Distribuční kanály

- 1,790 institucionálních repozitářů (Open DOAR)
- Otevřené portály pro dizertace, např.
 - DART Europe <http://www.dart-europe.eu/basic-search.php>
 - ND LTD ETD Search <http://www.ndltd.org/>
 - Base (Univ. Bielefeld) <https://www.base-search.net>
- PQDT Global
ProQuest Dissertations and Theses

Úspěšný příklad z Finska: Theseus.fi

- Společný repozitář 25 finských univerzit aplikovaných věd v systému Dspace, provozovaný Národní knihovnou
- Převážně práce na bakalářské úrovni
- Studenti musí svou práci poskytnout do repozitáře, aby splnili podmínky závěrečných zkoušek
- 6letý provoz, Theseus obsahuje 90.000+ VŠKP - roční přírůstek 15.000
- Téměř 18 milionů stažených plných textů ze systému Theseus v roce 2014 – vyšší než celkový počet stažených dokumentů ze všech ostatních finských repozitářů



Ilva, Jyrki. Repositories and ETDs – a success story from Finland. Presentation at Open Repositories, Indianapolis, June 10, 2015. <http://urn.fi/URN:NBN:fi-fe2015061110223>

Užití digitalizovaných PhD dizertací – London School of Economics

- Digitalizační projekt 2014 (2,000 položek)
- Počet stažení se v období 2014 – 2015 téměř ztrojnásobil po přidání zdigitalizovaných dokumentů
- Počet stažení v přepočtu na položku se nezměnil
- Zvýšilo se množství lidí přicházejících ze sociálních sítí, a to i na záznamy některých digitalizovaných (starších) dizertací
- Neuzavírají se individuální smlouvy s autory, na požádání se dokumenty mažou
- Nebyla prokázána korelace mezi užitím a citovaností



Bennett, Linda & Flanagan, Dimity. Measuring the impact of digitized theses: a case study from the London School of Economics. *Insights*. 2016, 29(2): 111–119;
<http://dx.doi.org/10.1629/uksg.300>

Užívání a význam

Výzvy pro analyzování využití v otevřených repozitářích:

- strojový přístup
- chybějící standardy

Iniciativy k vytváření spolehlivých a srovnatelných metrik, např.

- VB: IRUS-UK
<http://irus.mimas.ac.uk/>
- Německo: DINI
<https://dini.de/english/>

Žádné sledování dopadu dizertací
(?)



VŠKP, text mining a data mining

VŠKP si získávají pozornost jako zdroj pro text a data mining

- velké množství prací
- rozsáhlý obsah
- vhodné pro
 - jazykovou analýzu
 - oborově zaměřené studie
 - průzkumy zaměřené na post-graduální studium
- materiály pokrývající delší časové období
- nutná fulltextová indexace

Příklad 2016:

s použitím dat z ProQuestu bylo prošetřeno, jak často se dizertace z oboru psychologie na univerzitě v Kansasu dočkají publikování v podobě recenzovaných článků (prolinkování mezi dizertacemi a recenzovanými články)

→ 80% dizertací z oboru psychologie zůstalo nepublikovaných po 7 letech.

Závěr

- Dizertace nabývají na hodnotě
- Musí být prezentovány více systematicky
- Jejich využívání by mělo být monitorováno more closely with comparable figures
- Dopad v citovanosti
- VŠKP jsou pokladem pro text a data mining

Otázky?

Děkuji!

Meinhard Kettler
meinhard.kettler@proquest.com

