

Disseminating, Usage and Impact Tracking of ETDs

Making Dissertations & Theses accessible and discoverable

Prague, 19 October, 2016

Meinhard Kettler, ProQuest

Agenda

- Availability and Discoverability of ETDs
 - Challenges of Discoverability
 - Online vs Print availability
- Meaning and Impact of Graduate Research
 - Value of PhD theses for students and researchers
 - Usage study London School of Economics
- Text mining of ETDs

View from outside the academic community:

ProQuest, a long-standing partner of institutions regarding dissemination and preservation of dissertations and theses

Availability and Discoverability

- Shift from print to digital
 - Variety of access points to retrieve content in dissertations and theses.
 - Millions of theses on the web.
- New policies and workflows for 'born-digital' submissions

But:

- Not all the world's graduate output is available online
- Availability \neq Discoverability
 - Finding relevant theses can be difficult
- How big is the invisible part of the iceberg?

Why some dissertations are not found...

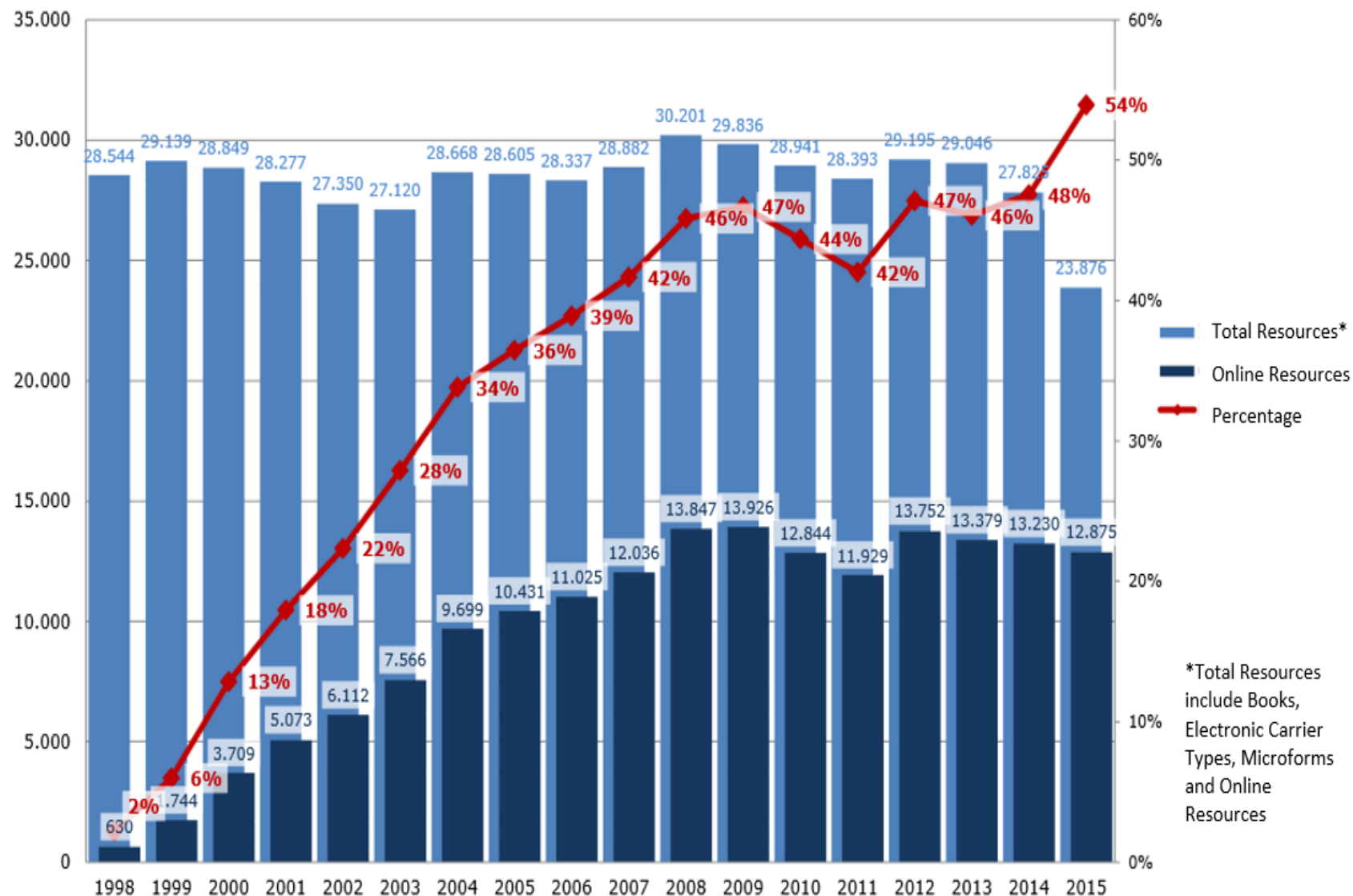
Possible reasons

- Author's decision (accepted by the institution)
- Publication embargo
 - Author
 - Third party copyrights
- Monograph published under a different title
- Submission in print without electronic copy
- No Institutional Repository – or even no catalogue records
- Limited discoverability due to language
- Limited search functionality, broken links
- ...and more

Challenges in the Background

- Country-wide or university standards not everywhere defined (Academic freedom)
- Inconsistency due to legal issues –
Copyright, author-driven embargoes, (plagiarism)
- OA publications not always findable either:
Need to index and harvest a vast and growing amount of content
- Print submission rate still high in many countries / institutions

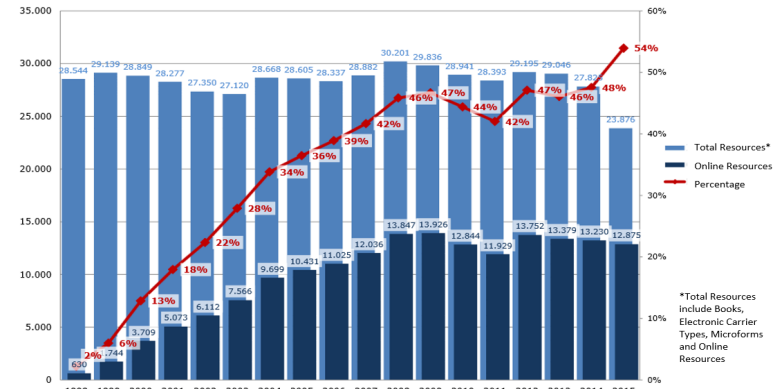
Print vs Online - Germany



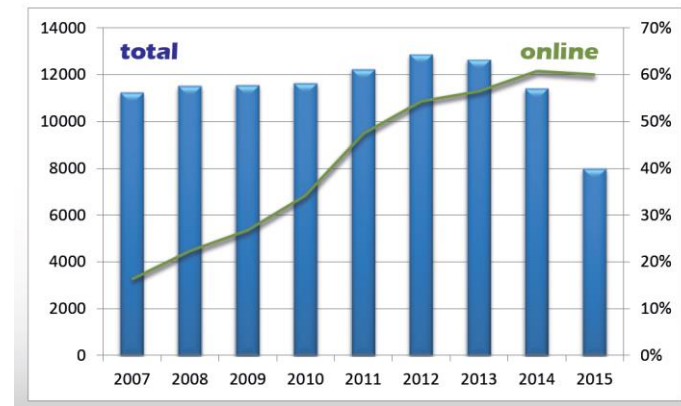
Online Resources Share of Total PhD Dissertations and Habilitations in German National Library Collection by Publication Year (updated 2 March, 2016). Deutsche Nationalbibliothek 2016, <http://www.dnb.de/DE/Wir/Kooperation/dissonline/dissonlineStatistik.html>

Print vs Online – Some examples

- Germany
 - 54 % ETDs in 2015
 - Slowly growing share



- France
 - About 60 % ETDs
 - Decree to transition to mandatory ETD submission until 2018



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

- USA
 - 90+ % of submissions are ETDs (ProQuest figures)

Value and Importance of ETDs

Dissertations and theses ...

- cover a huge variety of research subjects including exotic topics
- cover topics more comprehensively (than journal articles)
- serve as sources of methodological and theoretical approaches
- help to identify strategic gaps in the literature
- provide recommendations for future research
- are used as sources of up-to-date literature reviews
- serve as examples regarding level of detail and writing style

= represent valuable content for researchers and PhD students

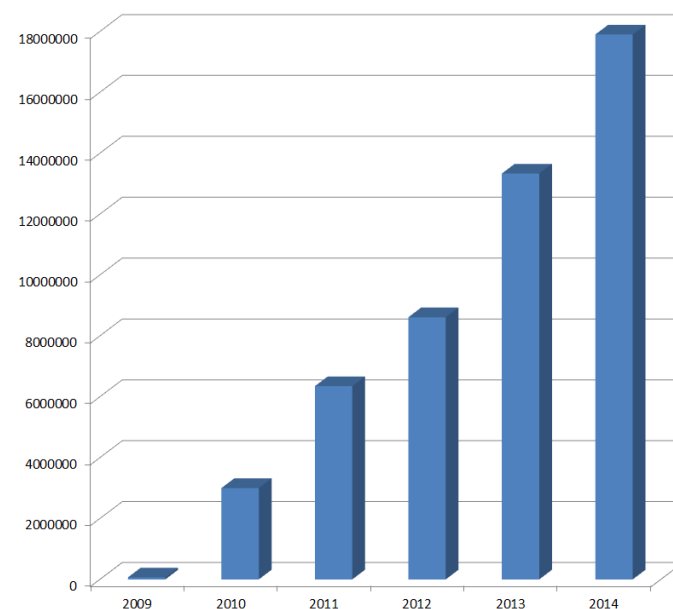
- seem to be neglected sometimes
- should get higher visibility to help authors and institutions

ETD Distribution Channels

- 1,790 Institutional Repositories (Open DOAR)
- Open ETD portals, e.g.
 - DART Europe <http://www.dart-europe.eu/basic-search.php>
 - NDLTD ETD Search <http://www.ndltd.org/>
- PQDT Global
ProQuest Dissertations and Theses

Success story from Finland: Theseus.fi

- Shared DSpace repository of 25 Finnish universities of applied sciences, hosted by the National Library
- Mainly bachelor-level theses
- Students have to upload theses as part of the graduation process
- After 6 years, Theseus contained 90.000+ theses –
15.000 theses added each year
- Almost 18 million full-text downloads from Theseus in 2014 – more than the total number of downloads from all of the other Finnish repositories



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>

Usage of digitized PhD theses – London School of Economics

- Digitization project 2014 (9,000 items)
- Download numbers nearly tripled from 2014 to 2015 with newly added digitized content
- Downloads per item reached the same level as before
- Increased overall traffic from social media platforms, including for some of the digitized (old) theses
- Take down policy instead of individual author permission
- Correlation of usage and citations could not be proved



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>

Usage and Impact

Challenges for analysing usage on Open Access repositories:

- Non-human access
- Missing standards

Initiatives to create reliable and comparable metrics, e.g.

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

No impact tracking of theses (?)



ETDs for Text and Data Mining

ETDs receive attention as resource for Text and Data mining

- High number of works
- Extensive content
- Allow for
 - Language analysis
 - Subject specific studies
 - Studies on graduate students, etc.
- Material covering longer periods needed
- Full text indexing needed

Example 2016:

Linking between dissertations and peer-reviewed articles in Psychology was examined at Univ. Kansas using ProQuest data

→ 80% of Psychology dissertations remain unpublished after 7 years.

Conclusion

- ETDs continue to have an high value
- Need to be presented more systematically
- Usage should be monitored more closely with comparable figures
- Impact in citations
- ETDs represent a treasure for TDM

Any questions?

Thank You!

Meinhard Kettler
meinhard.kettler@proquest.com

