Developing Metrics for the Humanities

Alesia Zuccala

Royal School of Library & Information Science

University of Copenhagen

Presentation:

1. Educate humanities scholars and involve them in the process of data curation and metric developments

2. The 'humanistic' focus is on the return to the 'biblio' in bibliometrics: books and measures of books

R. Bod, *A New History of the Humanities: The Search for Principles and Patterns from Antiquity to the Present*, Oxford University Press, 2013.

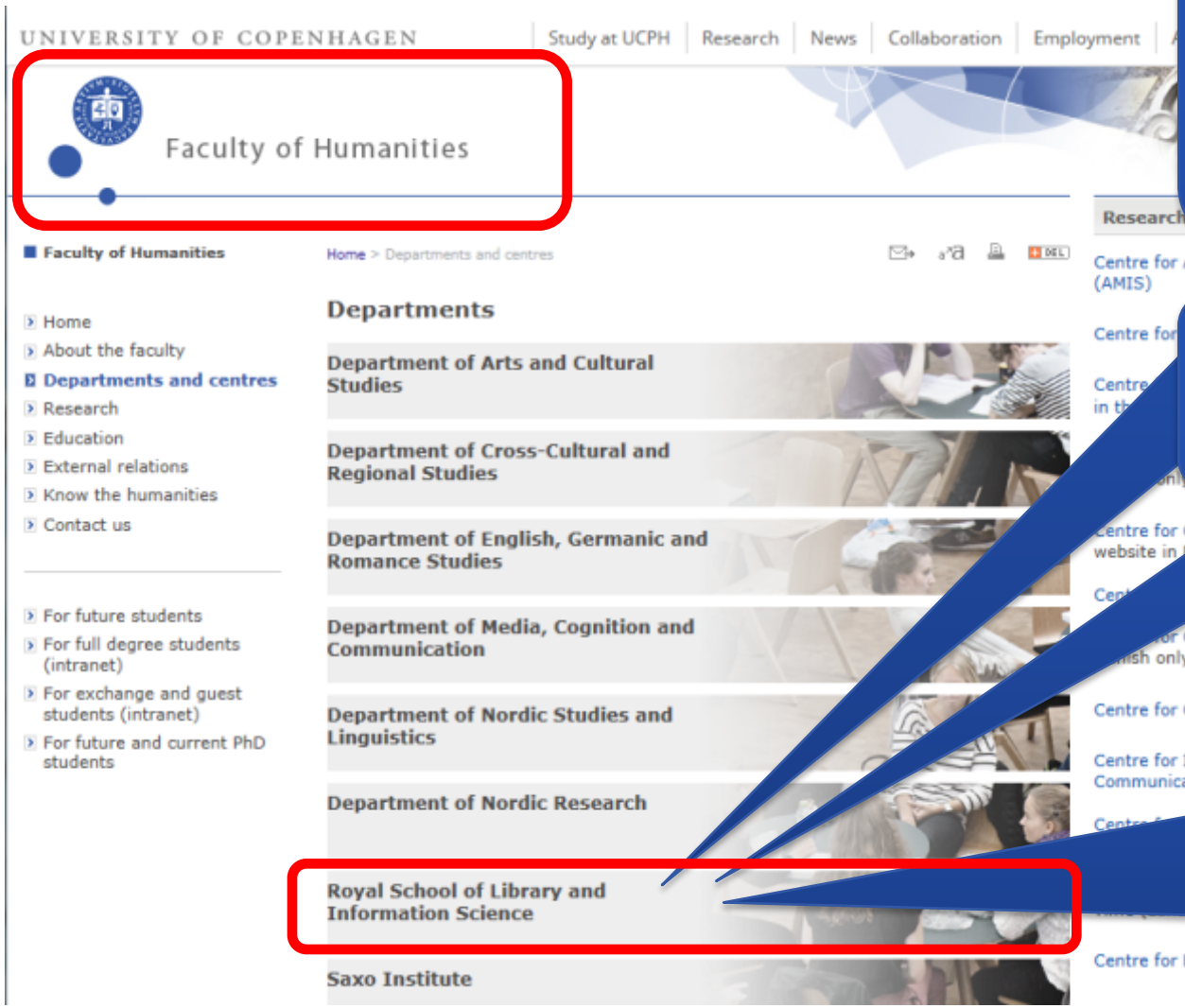
The Humanist

- In practice, the humanist may be defined as a scholar of philosophy, philology, theology, history, archaeology, media studies, and literature.

- Intellectually, they belong to a unique 'tribe' of specialists dedicated to *examining, critiquing, and highlighting patterns of relationship across products of the human mind* (Bod, 2013).

R. Bod, *A New History of the Humanities: The Search for Principles and Patterns from Antiquity to the Present*, Oxford University Press, 2013.

Faculty of Humanities:



"Interplay between people, information and technology"

Undergraduates to PhDs

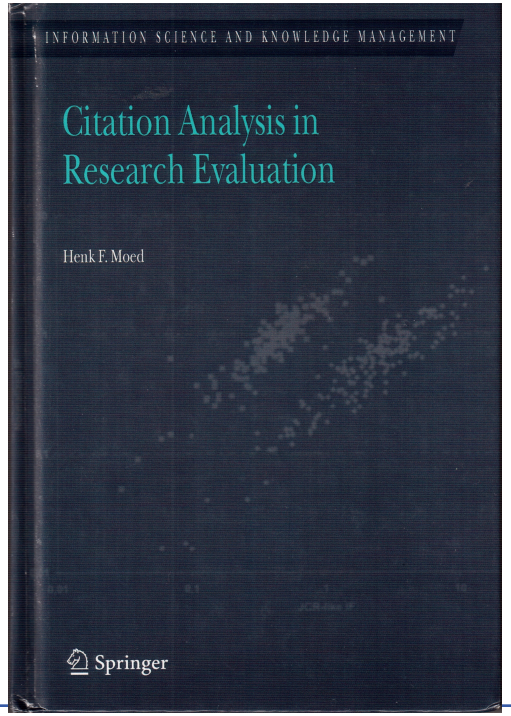
Libraries, Public Administration & Private Companies

Bibliometric Education for Humanists?

Training & Education » CWTS Course for Professionals

CWTS Course 'Measuring Science and Research Performance'

Our one-week course provides participants with the bibliometric knowledge and skills they need to interpret bibliometric statistics properly and usefully.



Bibliometric Education Programs:

Library and Information Science

White, H. & Griffith, B. (1981) Author co-citation: a literature measure of intellectual structure

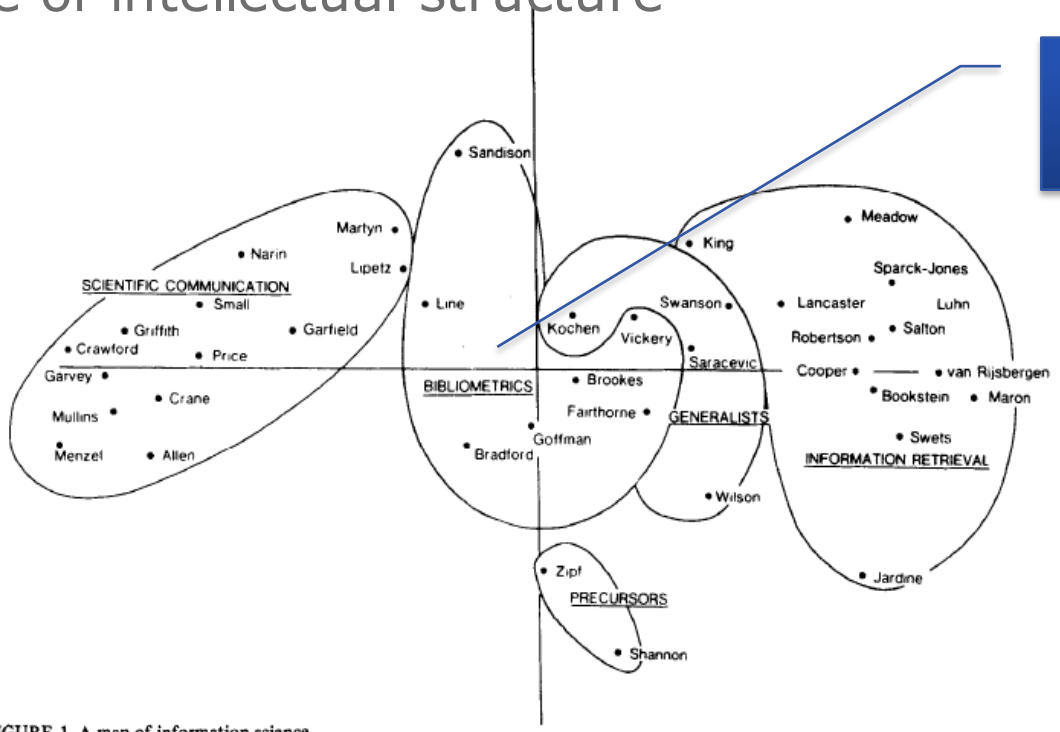
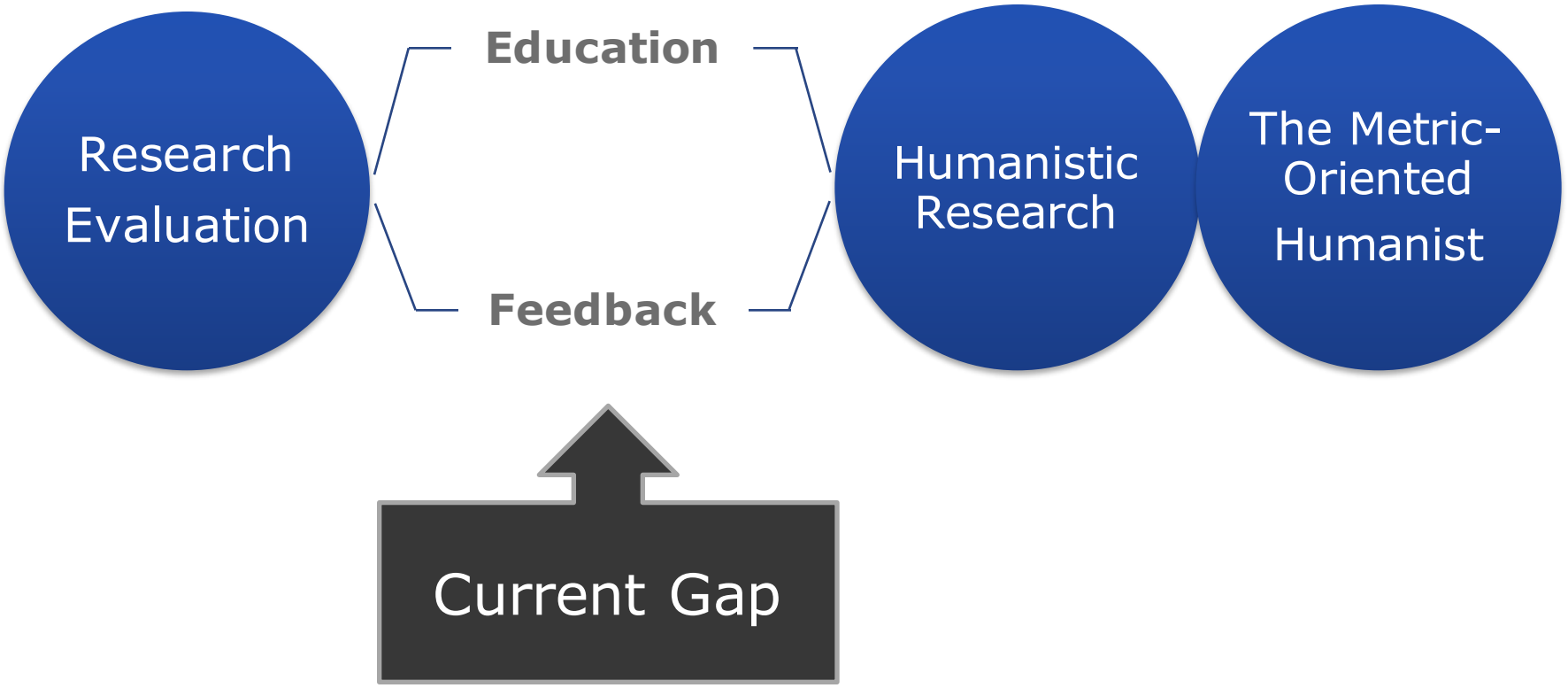


FIGURE 1. A map of information science.

Bibliometrics

A Humanistic Approach:



Inciting the Metric Oriented Humanist:

DOI 10.3233/EFI-150969
IOS Press

Inciting the metric oriented humanist: Teaching bibliometrics in a faculty of humanities

Alesia Zuccala
The Royal School of Library and Information Science, University of Copenhagen, Birketinget 6, DK-2300, Copenhagen S, Denmark
E-mail: a.zuccala@hum.ku.dk

In the past few decades the core of bibliometrics has predominantly been 'scientometric' in nature, due to the first commercial citation index having been created for scientific journals and articles. The production of citation indexes for books implies that proper education related to their use is now becoming critical. A new breed of humanistic bibliometrician can emerge successfully if well-trained lecturers are prepared to provide students with a foundation in mathematical indicator construction and a stimulating environment dedicated to problem-oriented learning. We examine some of the key principles and practices associated with teaching bibliometrics to humanists, with the aim of encouraging students to reflect upon new indicators relevant to scholarly research outputs across the humanities. Emphasis is placed on the "biblio" in bibliometrics (i.e., books), digital record-keeping across the humanities, the work of humanists who have sought to explain the development of new knowledge through objective pattern-seeking, and the complementary value of alternative versus traditional indicators of scholarly research performance.

Keywords: Pedagogical philosophy, humanistic bibliometrics, research evaluation for the humanities

1. Introduction

With the introduction of Thomson Reuter's Book Citation Index [1,68] and the addition of thousands of books to Elsevier's Scopus [19], quantitative research assessments are now emerging for the humanities [11,28,44,70]. New or expanded citation indexes; however, cannot be the only starting point for introducing humanists to bibliometrics. Humanities scholars themselves also need to understand the history of this practice and what it can mean for them. Consequently, *there has never been a more critical period for examining the pedagogical challenges associated with teaching bibliometrics to students and scholars involved in the production of humanistic research*.

The purpose of this paper is to propose a pedagogical approach that is both indicative and supportive of the humanist's best interests, particularly in a politically charged academic environment where evaluations are unavoidable. Bibliometricians associated with educational policy have been quick to respond to the general 'evaluation' impetus, but gaps still exist between what humanities scholars value as professionals, where they want their research to have the most influence, and how we might educate them to initiate their own community of bibliometric experts.

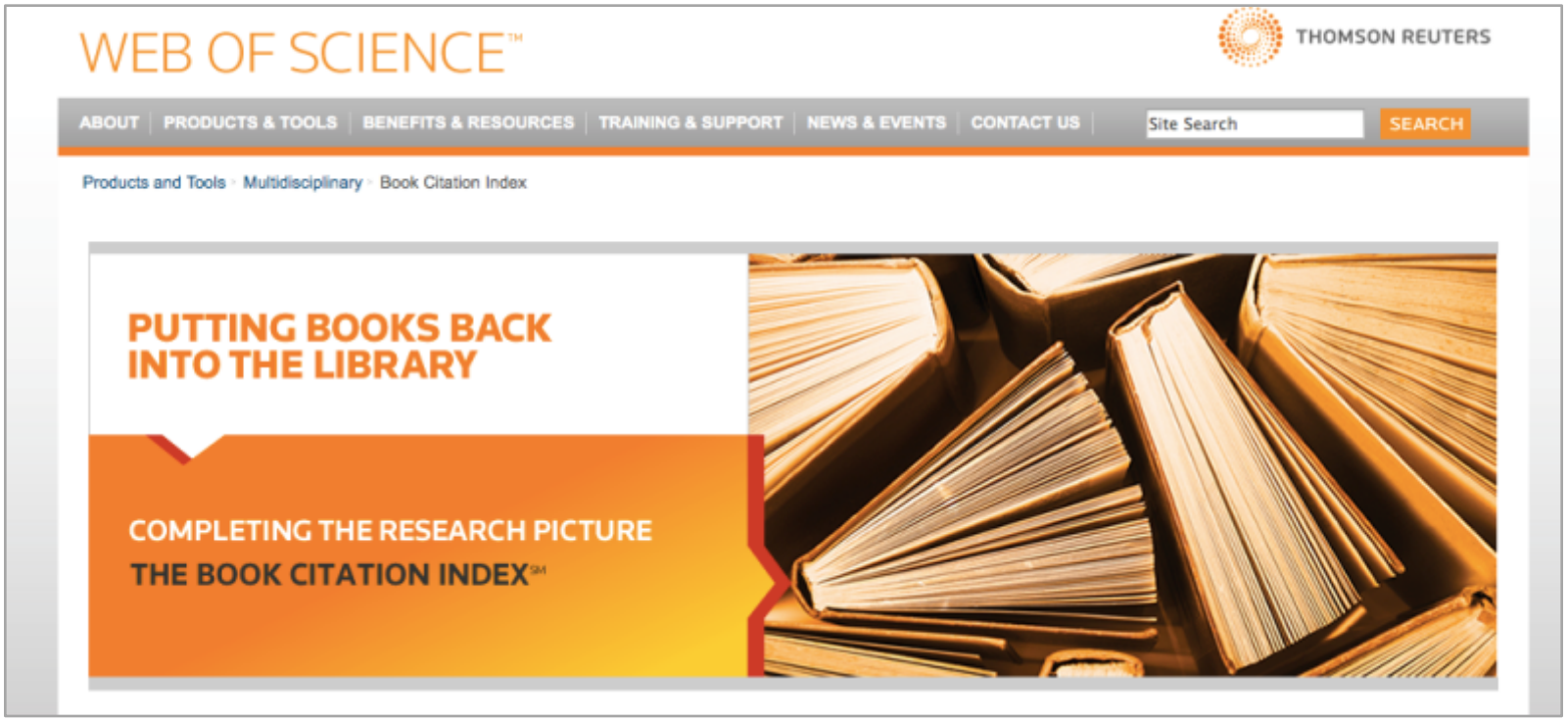
0167-8329/16/$35.00 © 2016 – IOS Press and the authors. All rights reserved

- Empiricism in the Humanities
- Eschew false assessment ideals
- A return to the 'biblio' (books)
- Digital humanities versus humanistic bibliometrics
- Record keeping for a humanistic programme
- Explaining mathematical indicators (e.g. normalization)
- Problem-based learning

A Return to the "Biblio"

Biblio + *Metron* =

Books and Measures



Denmark: BFI Performance Points

How BFI-points are calculated

BFI-points are calculated at the University level based on the following formula:

- BFI = B x C x F, where
  - B = Basic points
  - C = collaboration factor (1,25 if any external authors)
  - F = Local authors / All authors

Basic BFI-points:

Publication type	Level 1 *	Level 2 **
Monograph with publisher	5	8
Monograph in book series	5	8
Article in series (journal, book series, conference serie)	1	3
Contribution to anthology with publisher	0,5	2
Doctoral thesis	5	
Patent	1	

Level 1 – Normal
Level 2 – Prestigious

SOURCE: http://www.phdontrack.net/evaluation-and-ranking/bibliometric-funding-denmark/

The Book Citation Index:

Recent assessments of the BKCI point to numerous indexing problems, which can lead to over-counting or under-counting of books (e.g., Gorraiz et al., 2013; Leydesdorff & Felt, 2013; Torres-Salinas et al., 2013).

Metrics for Books as 'Families' of Works

Regardless of where or how publication counts or citations of books are collected:

It is essential to recognize that 'books' often belong to bibliographic families

Metrics for Books as 'Families' of Works

We can examine bibliographic families both theoretically and empirically:

1. The family-oriented entity-relationship model: The Functional Requirements for Bibliographic Records (FRBR)

1. Collect and assess data from the Book Citation Index, Danish Pure Registry, OCLC-WorldCat, and Goodreads

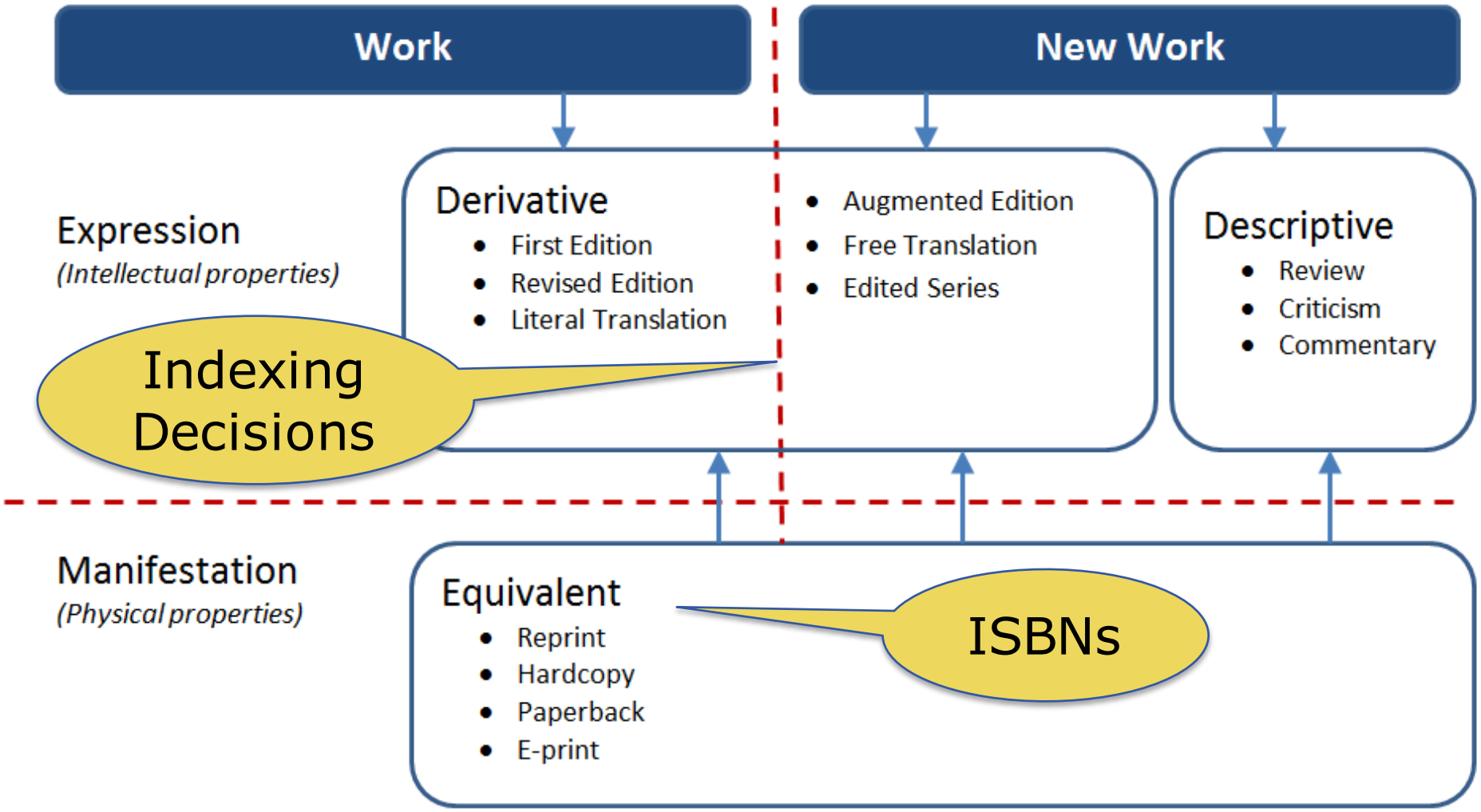
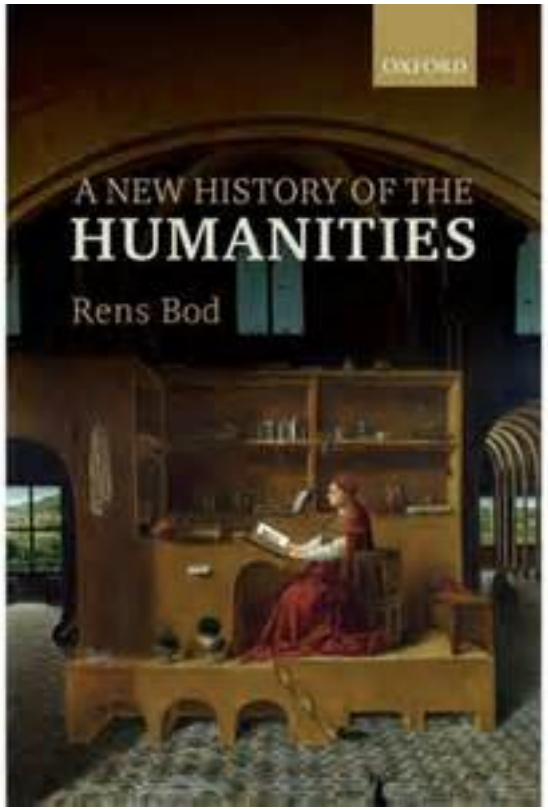


Figure 1. Modified FRBR model (adapted from Tillet, 2001) of 'bibliographic' families for a scholarly work.

Model of a ’Complex’ Work:



I would consider the English [expression] as a kind of improved edition of the Dutch book. When the Dutch work was translated into English, I sent it to OUP [Oxford University Press] and incorporated the comments by the 5 OUP reviewers into the English version…. (R. Bod, personal communication, June 16, 2016)

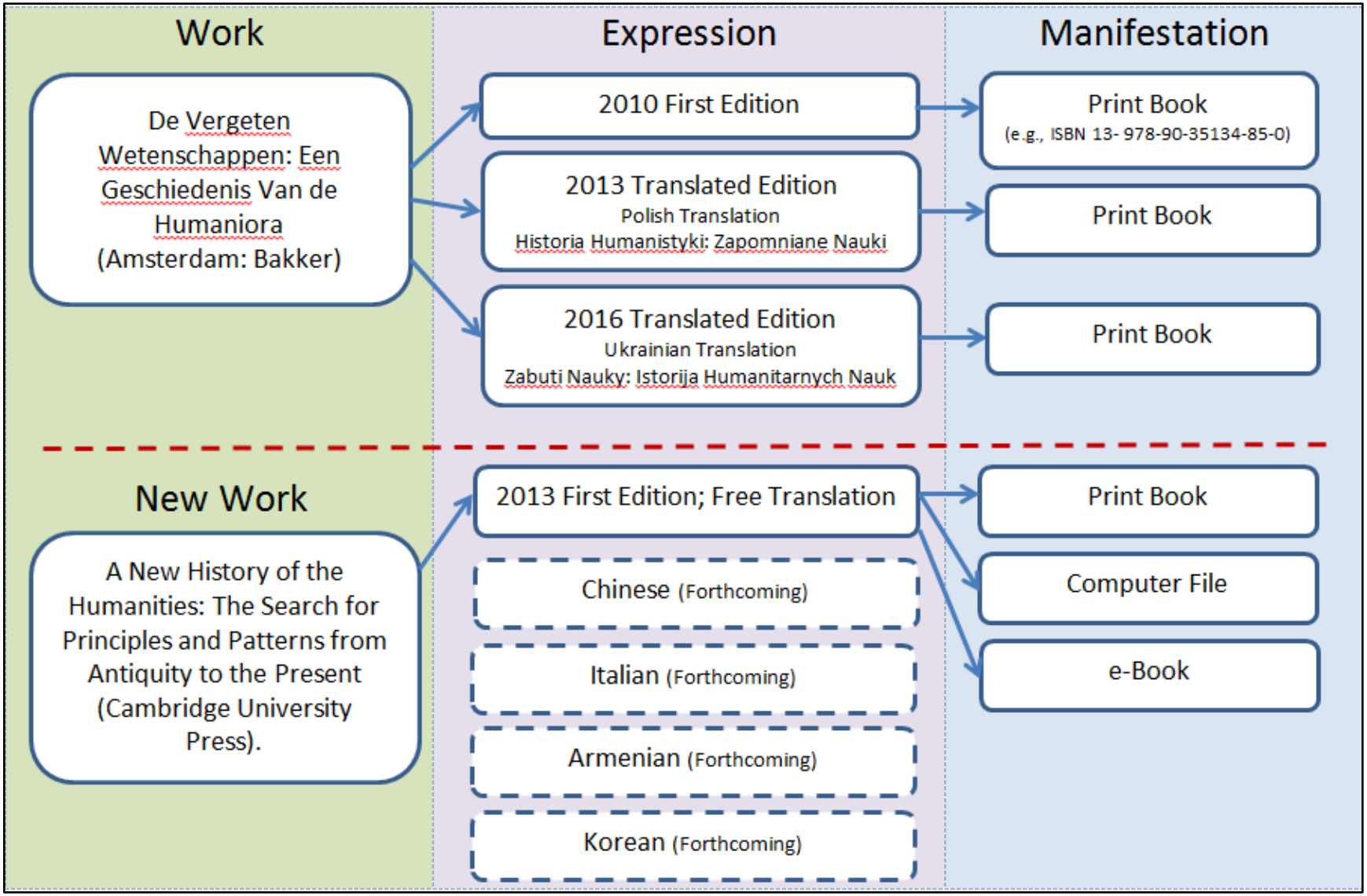


Figure 2. Model of a complex 'work' with expressions and manifestations of a 'new work'

Research:

- *Do books currently indexed in the Book Citation Index (BKCI) have adequate metadata and data designed to reflect inherent familial components and relationships?*

Method:

Extracted ISBNs for monographs (over 16,000):

- from the BKCI
- From the Danish Pure Repository

Matched with equivalent ISBNs and 'family' related ISBNs in:

- OCLC-WorldCat
- Goodreads

OCLC™
The world's libraries. Connected.

Added all ISBNs and metadata related to monographs to an SLQ database

Table 1. ISBN matching and retrieval results for total manifestations, expressions, and works.

65% extra ISBNs found

		The Book Citation Index	∞	The Danish Pure repository
1	Number of ISBNs crawled	16,392 (35%)		8,604 (81%)
2	Number of overlapping ISBNs		35 (0.41%)	
3	Extra related ISBNs found in OCLC-WorldCat and Goodreads	30,903 (65%)		2,042 (19%)
4	Total unique ISBNs in the dataset under study	47,295 (100%)		10,646 (100%)
5	ISBNs with distinct language and publication year	34,236		8,362
6	Total Expressions	20,284		7,844
7	Total Works	16,311		8,195

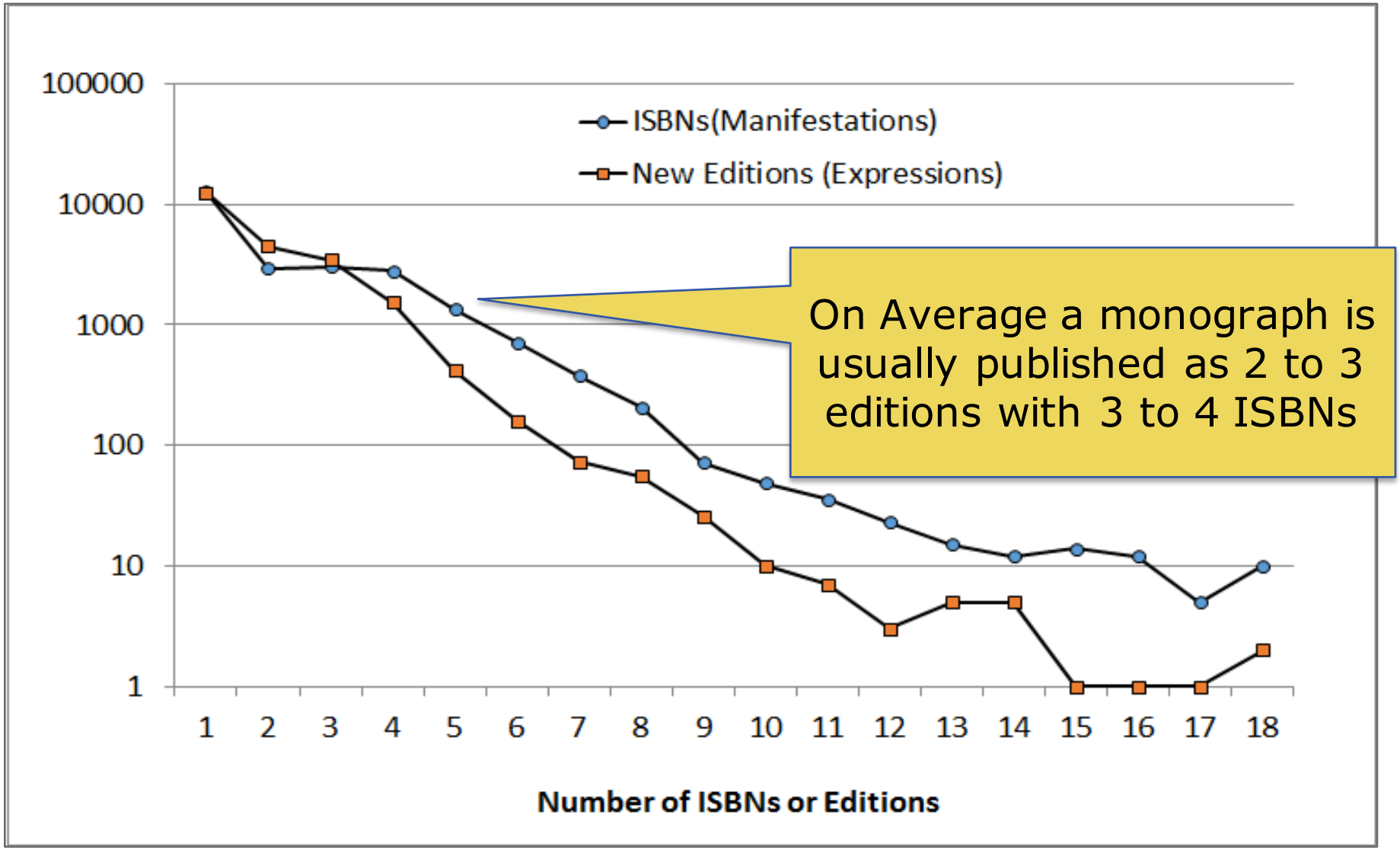


Figure 3. Frequency distribution of works with one or more ISBNs and published as one or more edition.

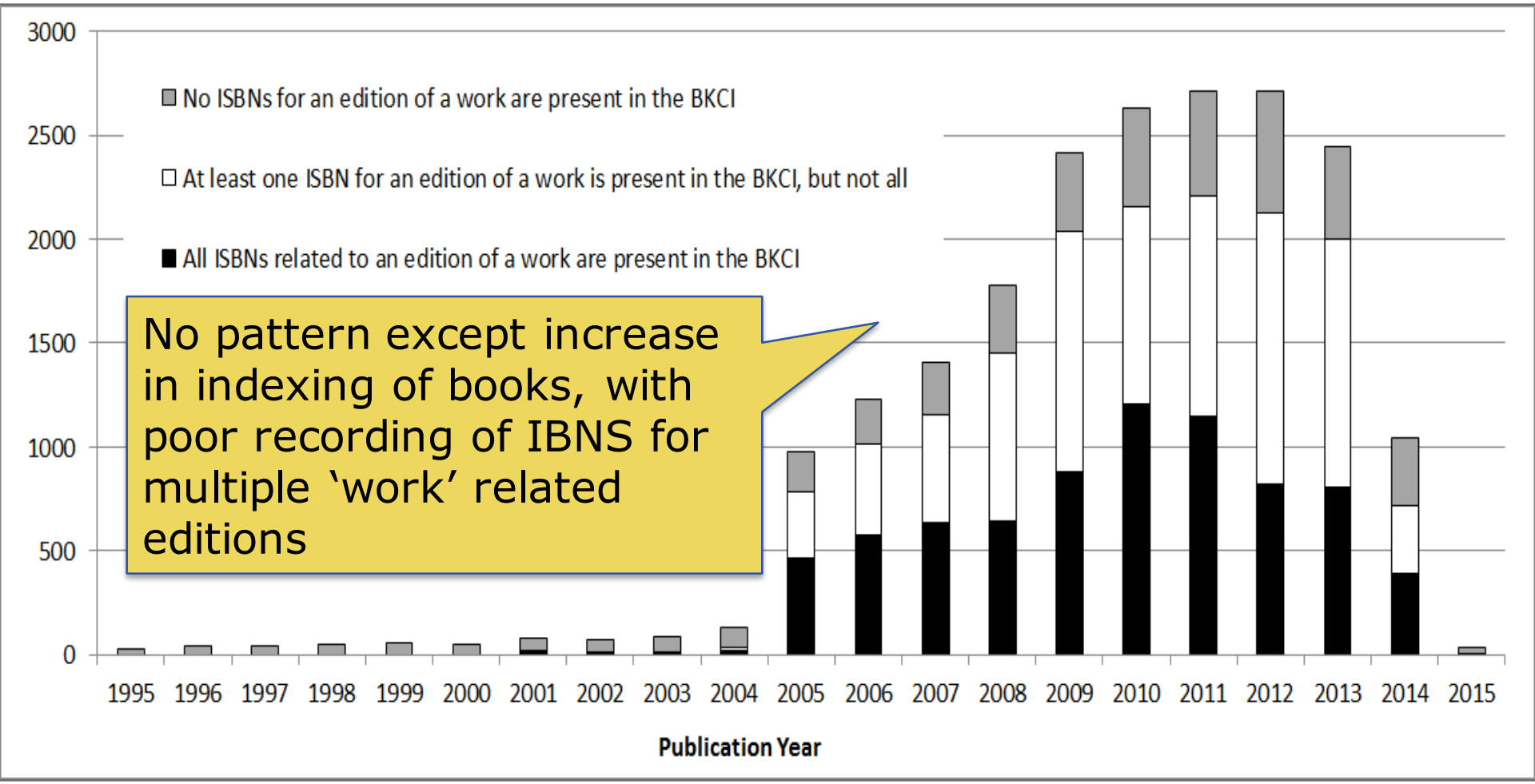


Figure 4. Indexing quality of the BKCI based on ISBNs per edition for publication years 1995-2015.

Many are Danish texts published as one edition, or a translated edition that was not originally published in English by translating author

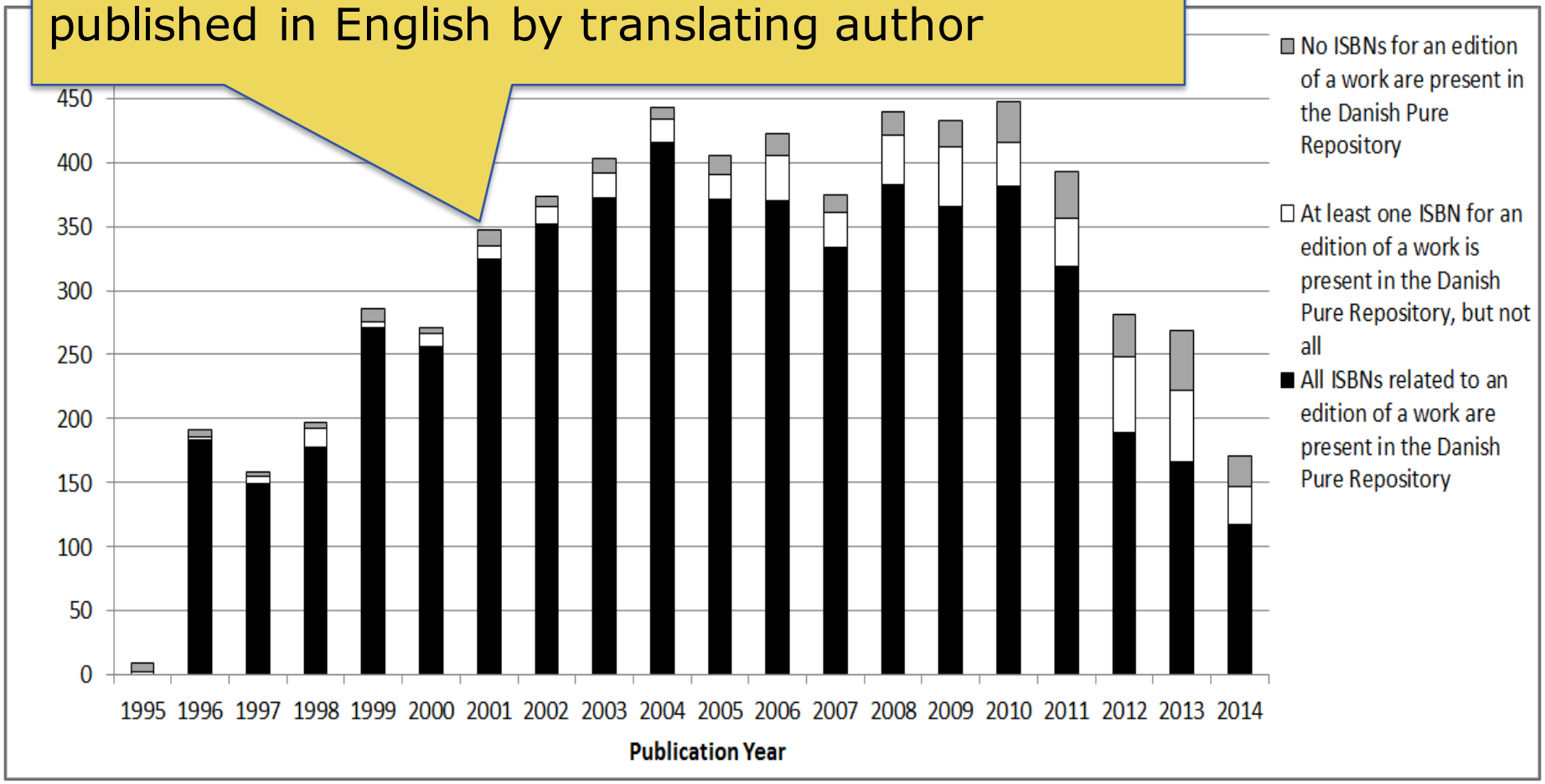
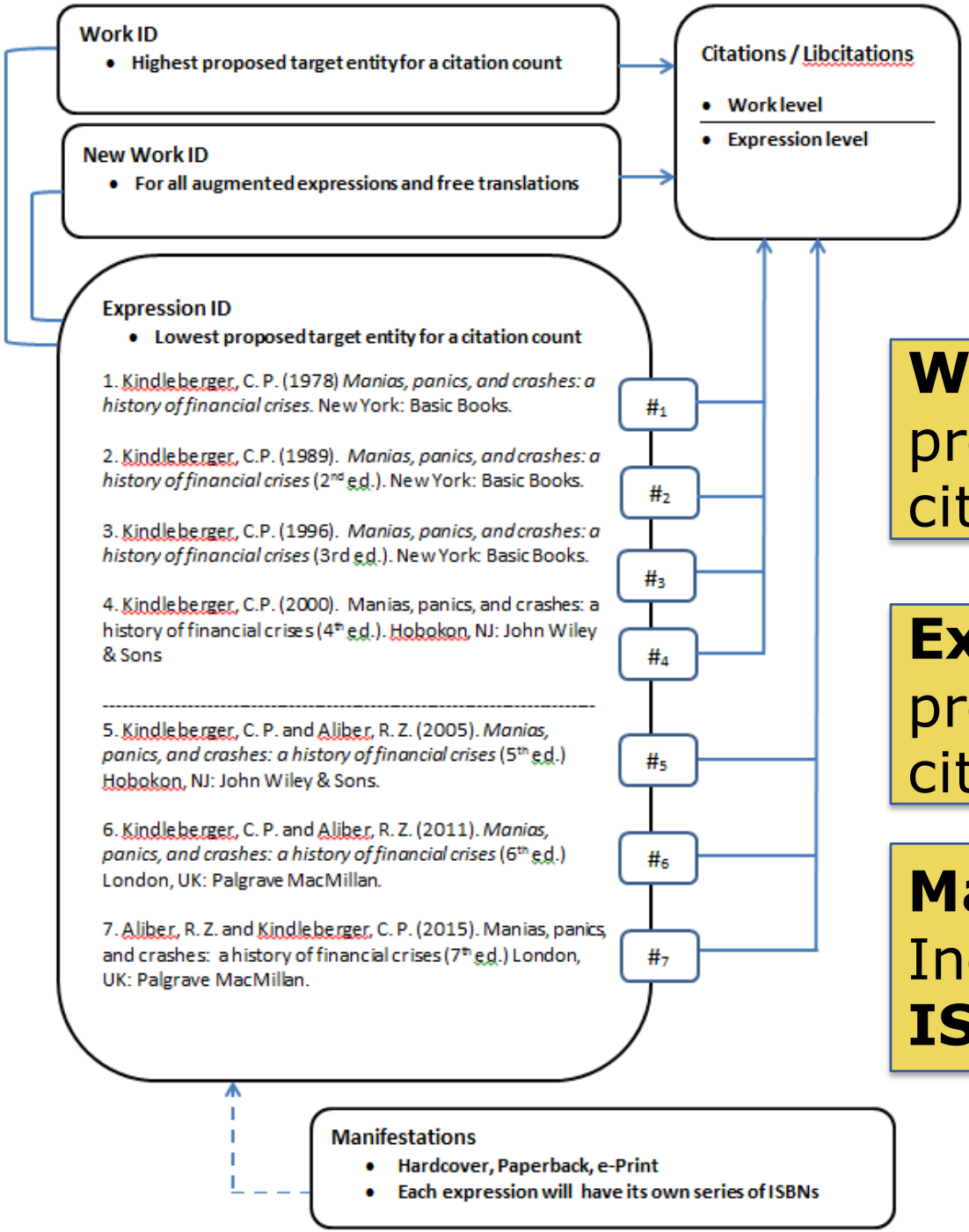


Figure 5. Indexing quality of the Danish Pure Repository based on ISBNs per edition for publication years 1995-2015.

Indexing model for the BKCI for more accurate 'metrics'



Work ID: Highest proposed target for citation count

Expression ID: Lowest proposed target for citation count

Manifestations: Include all series of **ISBNs** per edition

Conclusions:

- Book Citation Index is not an accurate index of monographs as complex works
- National data repositories (as in the Danish case) tend to be more accurate – but little overlap with the BKCI in terms of coverage
- Family-oriented entity-relationship model – the FRBR - is recommended for new indexing approach to mongraphs in the BKCI

Research:

- Zuccala, A. (2016) Inciting the metric oriented humanist: Teaching bibliometrics in a faculty of humanities. Education for Information 32(2), 149-164.

- Zuccala, A., Breum, M., Bruun, K., and Wunsch, B. (2017, forthcoming). Metric assessments of books as 'families' of works. In Journal of the Association for Information Science and Technology.

Thank you kindly for your attention!

E-mail: a.zuccala@hum.ku.dk