

The Functional Requirements for Bibliographic Records (FRBR) framework for indexing monographs: Implications for the Book Citation Index™ and metric evaluations

Mads Breum, madsbreum.d.1@gmail.com

Kasper Bruun, kasbru@gmail.com

Alesia A. Zuccala, Associate Professor, University of Copenhagen, a.zuccala@hum.ku.dk

Bernd T. Wunsch, bernd@wunsch.dk

Purpose

Initially this study’s aim was to assess the impact of non-english monographies versus their translated editions. This proved to be difficult since the current practice when indexing scientific and scholarly literature in citation indexes and national repositories do not live up to the IFLA defined Functional Requirements for Bibliographic Records (FRBR) standards. Modelling how the current practice influences bibliometric analysis and looking at possible solutions became the new focus of our study.



Figure 1: Current practice for indexing and citing journal type publications

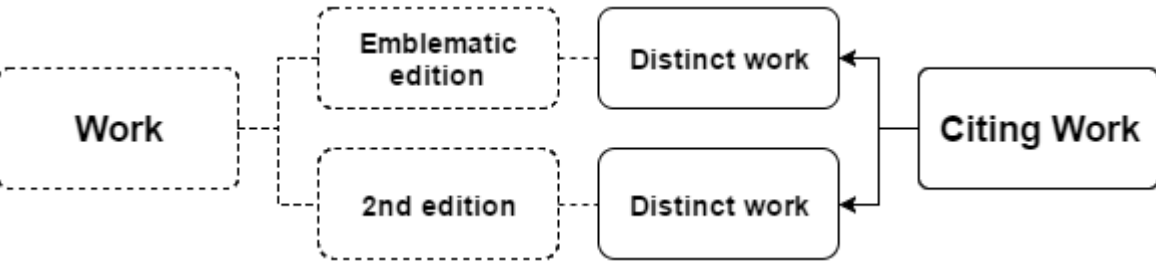


Figure 2: Current practice for indexing and citing book type publications

Method

Based on an initial list of ISBNs from Thomson Reuter’s Book Citation Index (BKCI), we retrieved additional relevant metadata from FRBR compliant datasources; specifically OCLC WorldCat (OCLC) and GoodReads (GR). From an initial dataset of 16,392 ISBNs we ended up with a final dataset of 47,295 ISBNs. Using the relational metadata from OCLC and GR, we identified how each ISBN belonged to a work-level bibliographic entity. Operationalization of the expression-level bibliographic entity was done by grouping ISBN within each work which shared language and publication year metadata.

Results

Of our full dataset, 47,295 ISBNs were identified as being either already indexed or related to an ISBN indexed in BKCI, divided on 16,311 works. Of these ISBNs, 34,236 had the prerequisite metadata for expression identification leading to 20,284 expressions. From these bibliographic entities, it became apparent that 21.5% of all expressions identified, have no representation in BKCI.

	The Book Citation Index
1	Number of ISBNs crawled
2	Extra related ISBNs found in OCLC-WorldCat and Goodreads
3	total unique ISBNs in the dataset under study
4	ISBNs with distinct language and publication year
5	Total Expressions
6	Total Works

Table 1: Data and metadata retrieval results for the BKCI

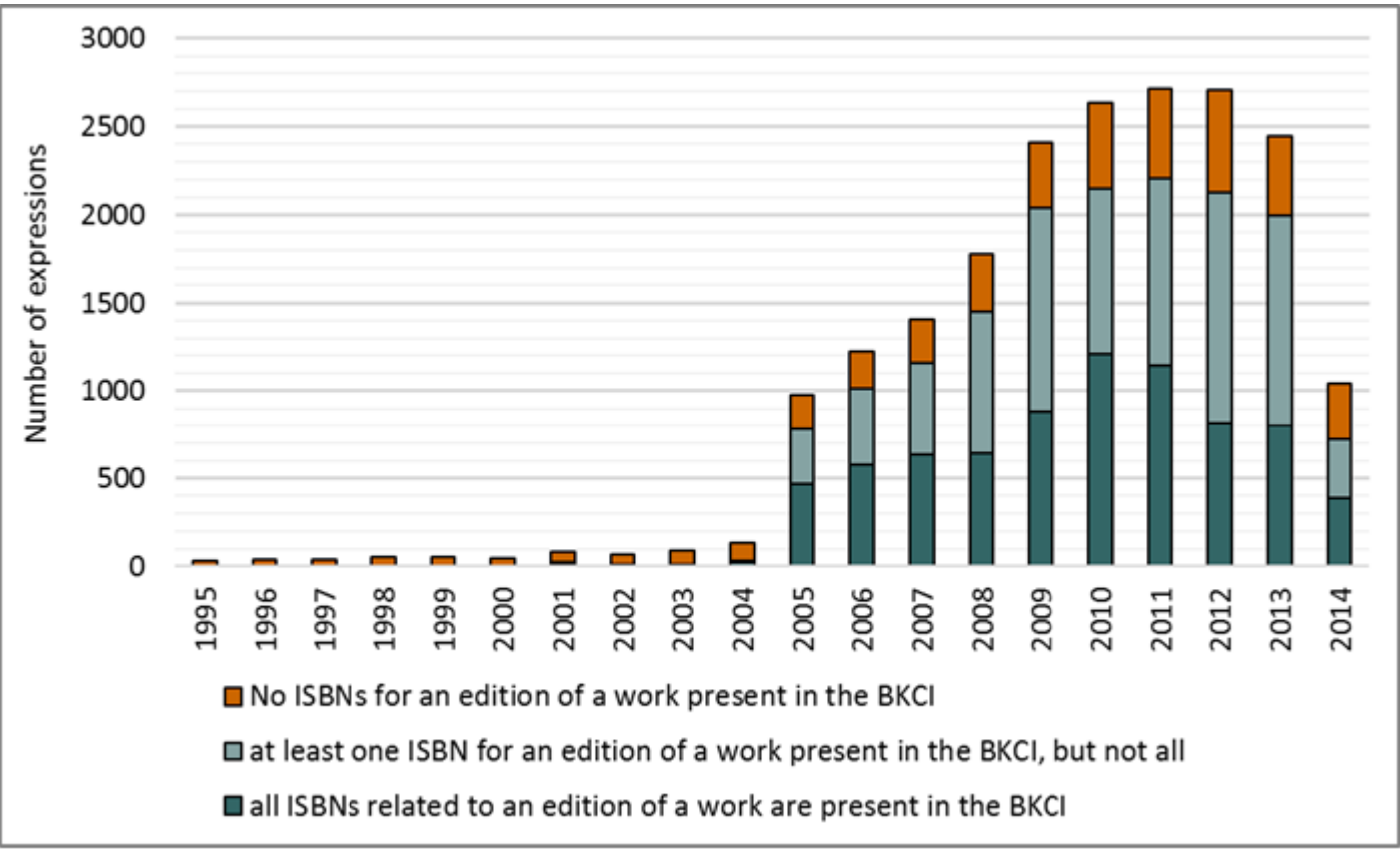


Figure 3: Indexing completeness of expressions found in the BKCI

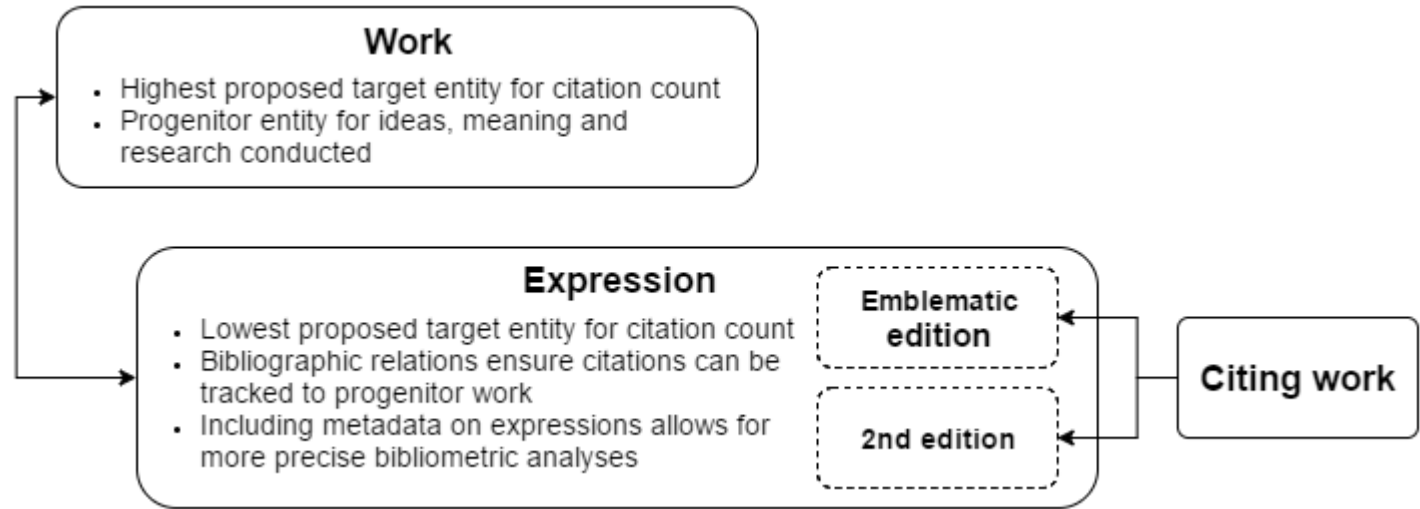


Figure 4: Best practice for indexing and citing publications according to FRBR standards

Conclusions

The current indexing practice and the fact that it does not live up to the FRBR standards have two immediate implications: Research publication databases are biased; with current indexing practice favoring journal type publications when it comes to complete and adequate metadata. Also, monographs are not exhaustively indexed, with one out of every five expressions not being represented in the BKCI.

As we have shown, relatively simple methods and existing metadata can be used to correctly identify relations between different manifestations, expressions, and which work they belong to. Furthermore implementing the FRBR standard could be done without disrupting the current data and metadata infrastructure as it exists today for journal type publications. Since the FRBR is designed to accommodate a wide range of media, this might in the future open up the possibility for including other sources not suited for bibliometric analysis under current practice.