

Enriching the VABB using external sources

Raf Guns and Peter Aspeslagh
ECOOM UAntwerp

10 September 2018

Overview

- Introduction: VABB
- Initiatives to enrich the database
- Conclusions

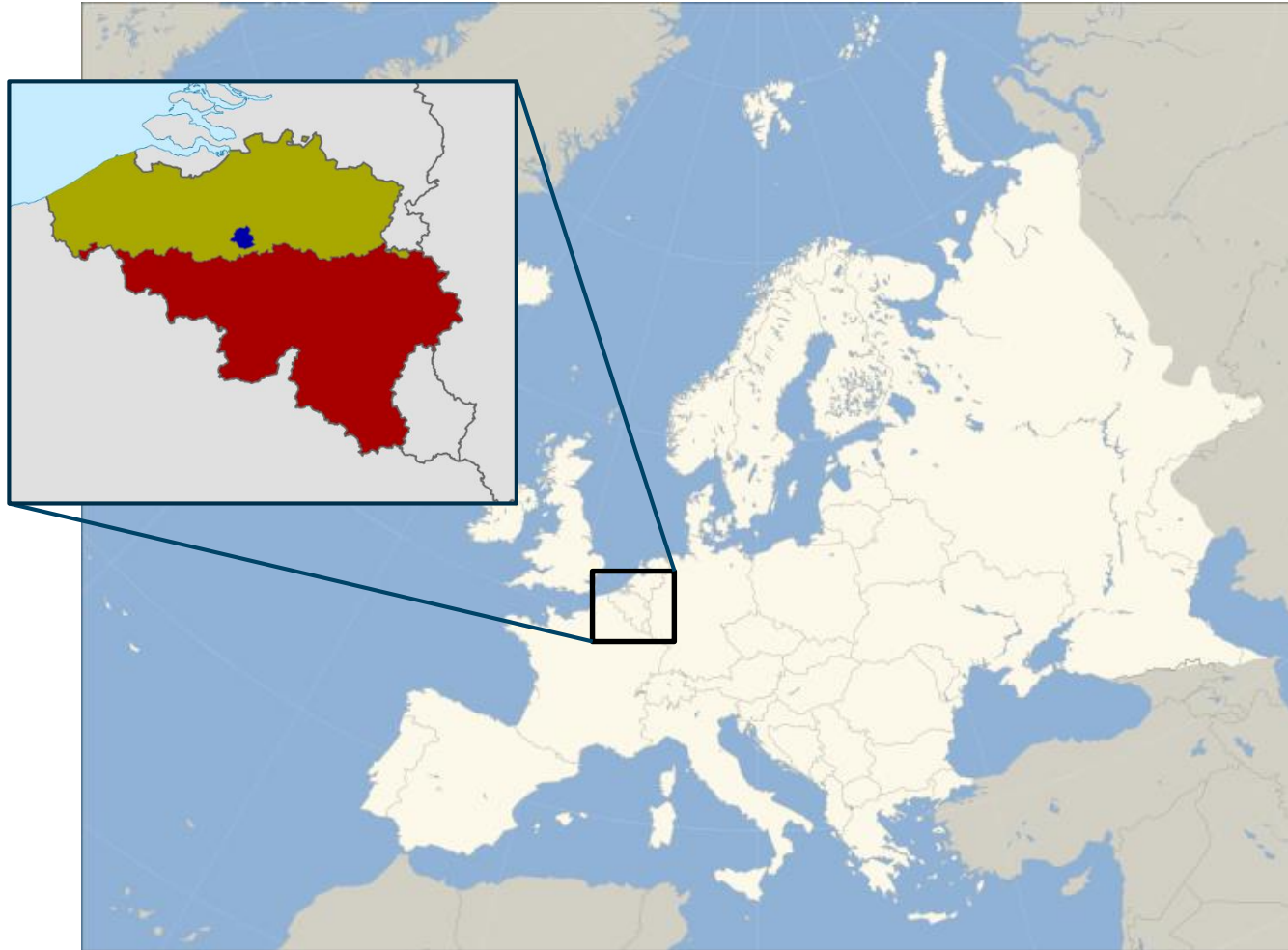
Introduction: VABB

Verleysen, F. T., Ghesquière, P., & Engels, T. C. E. (2014). The objectives, design and selection process of the Flemish Academic Bibliographic Database for the Social Sciences and Humanities (VABB-SHW). In W. Blockmans et al. (Eds.), *Bibliometrics. Use and Abuse in the Review of Research Performance* (pp. 117–127). London: Portland Press.

<http://www.portlandpress.com/pp/books/online/wg87/087/0117/0870117.pdf>



You are here



VABB

Database of **peer-reviewed** publications from **social sciences and humanities** (SSH)

- Either indexed in Web of Science,
- Or selected by the Authoritative Panel

Period:

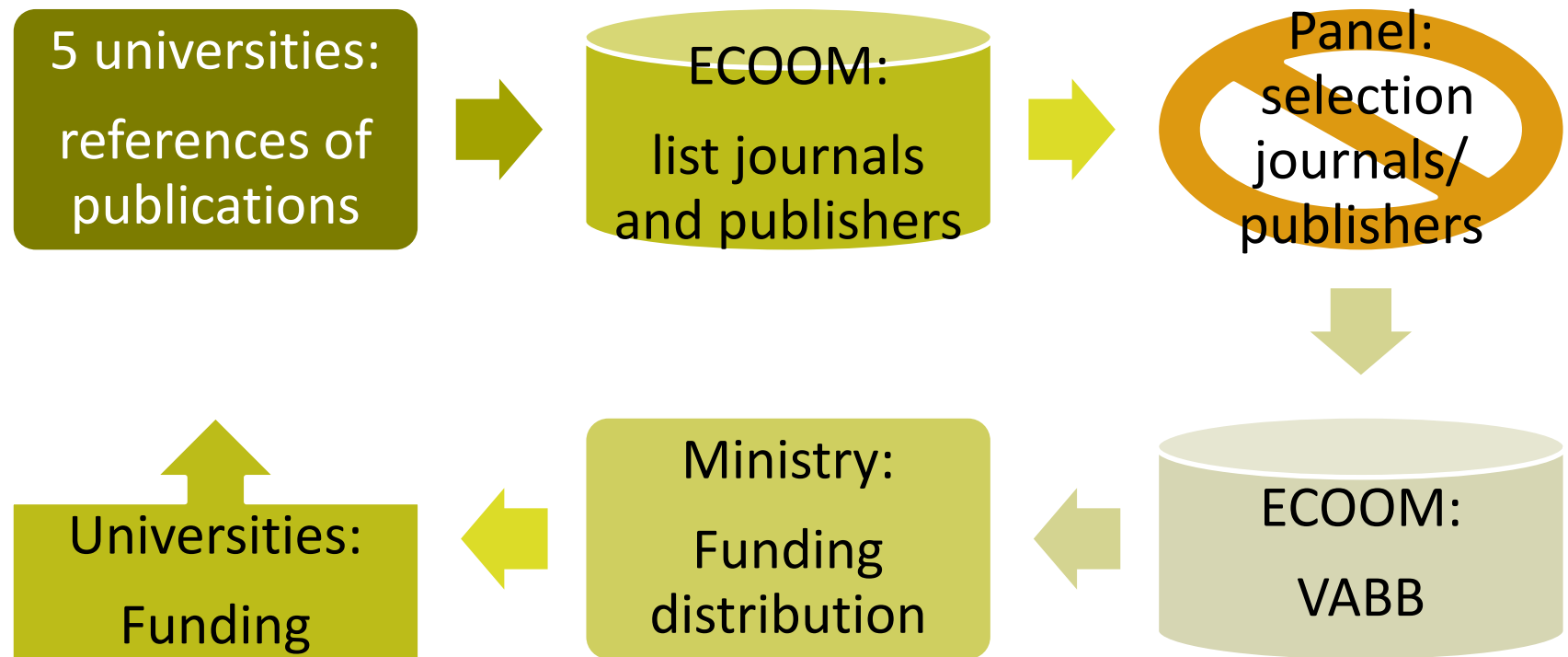
- 2000—2016

<https://www.ecoom.be/en/vabb>

Publication types:

- articles, monographs, edited books, book chapters, proceedings

Players in yearly VABB update



Uses of the database

Parameter in performance-based research funding system

- Main reason for existence

Services to government

- bibliometric screening of panel members for Research Foundation Flanders

Monitoring and reporting

- www.vlaamsindicatorenboek.be

Research

- Publication patterns, international comparison, impact, cognitive distance...

Initiatives to enrich the database

Classification: organizational vs. cognitive

Extending Literary Darwinism

Culture and alternatives to adaptation

Jan Verpooten

Research Centre of Marketing and Consumer Science,
University of Leuven, Belgium

Literature

Economics & business

Scientific Study of Literature 3:1 (2013), 19–27. DOI 10.1075/ssol.3.1.05ver
ISSN 2210-4372 / E-ISSN 2210-4380 © John Benjamins Publishing Company

Cognitive classification

= classification based on
journal/book *content*

Only for peer-reviewed
publications so far

Used for research and
services, NOT for funding

<https://doi.org/10.1007/s11192-018-2775-x>

Scientometrics
<https://doi.org/10.1007/s11192-018-2775-x>



A comparison of cognitive and organizational classification of publications in the social sciences and humanities

Raf Guns¹ · Linda Sile¹ · Joshua Eykens¹ · Frederik T. Verleysen¹ · Tim C. E. Engels¹

Received: 15 January 2018
© Akadémiai Kiadó, Budapest, Hungary 2018

Abstract We study the discrepancy between two ways of classifying publications in the social sciences and humanities (SSH): on the basis of the contents of publications and publication channels (cognitive classification) or on the basis of the organizational structure of departments, faculties etc. (organizational classification). Using data for the period 2000–2015 from 76,076 peer reviewed publications (co-authored by SSH researchers affiliated to a Flemish university, we compare the organizational classification of the publications with a cognitive classification based on OECD Fields Of Science. In all fields except one, the cognitive discipline with the most publications is the one that most closely matches the organizational discipline, although there are substantial differences between disciplines in terms of overlap and concentration. At a higher aggregation level, we find that 73% of publications from the humanities are published in humanities publication channels, while this is only the case for 59% of publications from the social sciences. Social sciences are shown to have important links to medicine and health sciences. The spread of publications over cognitive disciplines, including non-SSH disciplines, is only partially related to the assignment of some publication channels to multiple FOS fields as

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11192-018-2775-x>) contains supplementary material, which is available to authorized users.

✉ Raf Guns
raf.guns@uantwerpen.be
Linda Sile
linda.sile@uantwerpen.be
Joshua Eykens
joshua.eykens@uantwerpen.be
Frederik T. Verleysen
frederik.verleysen@uantwerpen.be
Tim C. E. Engels
tim.engels@uantwerpen.be

¹ Centre for R&D Monitoring (ECOOM), Faculty of Social Sciences, University of Antwerp, Antwerp, Belgium

Published online: 19 May 2018

Springer



Cognitive classification

Data source	Used for	Classification scheme(s)	N
Web of Science	6,474 journals	Web of Science Subject Categories	37,376
Scopus	7,208 journals	All Science Journal Classification (ASJC)	46,785
Norwegian Register for Scientific Journals	4,030 journals	NPI scientific fields	41,570
ISSN.org	945 journals	- Dewey Decimal Classification (DDC) - Universal Decimal Classification (UDC)	2,123
OCLC Classify	7,935 books and proceedings	- Dewey Decimal Classification (DDC) - Library of Congress Classification (LCC)	14,149

DOIs

Given DOI, we can obtain **rich metadata**:

- Add missing fields
- Include new information: e.g. citations, OA availability...
- Check our own metadata for errors

Where to **retrieve** a publication's DOI?

- DOI is delivered as part of metadata
- CrossRef API <https://github.com/CrossRef/rest-api-doc>



Process:

- Scraping CrossRef (via DOI)
- Matching (meta)data with VABB
- Store enriched (meta)data in collaborative tool via identifier

The same process will be repeated for matching via title field (partial string matching)

Example

```
## [vabb_citecheck_at_vabb:tabres[1491]]
Classes 'data.table' and 'data.frame': 1 obs. of 119 variables:
 $ Loi                : Factor w/ 39838 levels "c:vabb:135542",...: 35767
 $ total_results      : int NA
 $ search_terms       : chr NA
 $ start_index        : int NA
 $ items_per_page     : int NA
 $ alternative.id     : chr NA
 $ container.title    : chr "psychologie in de gezondheidszorg"
 $ created            : chr "2012-06-30"
 $ deposited          : chr "2012-06-30"
 $ DOI                : chr "10.1007/978-90-313-9259-9_19"
 $ indexed            : chr "2018-05-05"
 $ ISSN              : chr NA
 $ issue              : chr NA
 $ issued             : chr "2002"
 $ license_date       : chr NA
 $ license_URL        : chr NA
 $ license_delay.in.days : chr NA
 $ license_content.version: chr NA
 $ member             : chr "297"
 $ page               : chr "323-340"
 $ prefix             : chr "10.1007"
 $ publisher          : chr "Bohn Stafleu van Loghum"
 $ reference.count    : chr "0"
 $ score              : chr "60.850758"
 $ source             : chr "Crossref"
 $ title              : chr "De psychologische behandeling van dove cliënten"
 $ type               : chr "book-chapter"
 $ URL                : chr "http://dx.doi.org/10.1007/978-90-313-9259-9_19"
 $ volume             : chr NA
 $ author             : List of 1
 .. $ :classes 'tbl_df', 'tbl' and 'data.frame': 1 obs. of 3 variables:
 .. .. $ given      : chr "M."
 .. .. $ family     : chr "Broesterhuizen"
 .. .. $ sequence   : chr "first"
 $ link               : List of 1
 .. $ :classes 'tbl_df', 'tbl' and 'data.frame': 1 obs. of 4 variables:
 .. .. $ URL        : chr "http://www.springerlink.com/index/pdf/10.1007/978-90-313-9259-9_19"
 .. .. $ content.type : chr "unspecified"
 .. .. $ content.version : chr "vor"
 .. .. $ intended.application: chr "similarity-checking"
```

Example

Publications database

De psychologische behandeling van dove cliënten

Pages

Timeline

Files

Categories 0

Connections 1

Contacts

References

Websites

Structure

Metadata

ID Management

Map

Tags

Print

VABB ▾

+

Databases ▾

Auxiliary modules ▾

Visualisation ▾

Peter ▾

Language ▾

Search item...

Note Add

VABB > Publications database > De psychologische behandeling van dove cliënten

De psychologische behandeling van dove cliënten

Basic data (VABB)

System ID	35920
Loi	c:vabb:177342
Title	De psychologische behandeling van dove cliënten
Authors (1)	<ul style="list-style-type: none">Broesterhuizen, Marcel
Year of Publication	2002
Language	Dutch; Flemish
Omschrijving	De psychologische behandeling van dove cliënten / Broesterhuizen, Marcel ** Houten, 2002 [=c:vabb:177342]
Journal Title (VABB)	
Journal Title (DB)	
Start page	323
End page	340
Pages	
Btitle	Psychologie in de gezondheidszorg: een praktijkoriëntatie
Ptitle	
Bseries	
Type	VABB-4

Basic data (VABB & updates)

System ID	29
Loi	c:vabb:177342
Title	De psychologische behandeling van dove cliënten
Authors (1)	<ul style="list-style-type: none">Broesterhuizen, Marcel
Year of Publication	2002
Language	Dutch; Flemish
Omschrijving	De psychologische behandeling van dove cliënten / Broesterhuizen, Marcel ** Houten, 2002 [=c:vabb:177342]
Journal Title VABB	
Journal Title (DB):	
Start page	323
End page	340
Pages	
Btitle	Psychologie in de gezondheidszorg: een praktijkoriëntatie
Ptitle	Bohn Stafleu van Loghum
Bseries	
Type	VABB-4

Example

☰

VABB ▾

+

Databases ▾

Auxiliary modules ▾

Visualisation ▾

Peter ▾

Language ▾

🔍

Search item...

☰

Publications database

🔍 De psychologische behandeling van dove cliënten

📄 Pages

📅 Timeline

📁 Files

📁 Categories 0

🔗 Connections 1

✉ Contacts

📖 References

🌐 Websites

🌳 Structure

📊 Metadata

Note Add

VABB > Publications database > De psychologische behandeling van dove cliënten

Item

Add a connection

About

-- Role -- ▾

Combine

Add

Select: All

Items

Title	Extra information	📄	🗑
External links			
http://www.springerlink.com/index/pdf/10.1007/978-90-313-9259-9_19		📄	🗑

Example: data correction

Publication

Title	Specific nutritional needs for children: are children small adults?
Author	VANDENPLAS, YVAN
Number	ISSN 1573-4013
Publication	2005
Physical description	1, p. 135-142

Specific Nutritional Needs for Children. Are Children Small Adults?

[Vol. 1 , Issue. 2]

Author(s):

Yvan Vandenplas, Bruno Hauser and Jean De Schepper Pages 135 - 142 (8)

Authors

Gender:

- < Ongoing research project coordinated by Emanuel Kulczycki and Vincent Larivière
- Automation?
 - gender-guesser
 - <https://genderize.io/>

CHRIS	VAN DEN HAUTE	female	mostly_male	either	female
CHRIS	MASUI	female	mostly_male	either	female
CHRIS	BAEKEN	female	mostly_male	either	male
CHRIS	TIMMERMAN	female	mostly_male	either	female
GEERT	VAN OVERWALLE	female	male	boy	female

To conclude

Flemish VABB: designed and built in context of **funding system**

Other uses (esp. research) lead to enrichment and improvement of the database

Importance of unique **identifiers** for connecting to other systems

- Internal
- External: ISSN, ISBN, DOI, ORCID (future)