



Recopilación y registro de datos de afiliación de autor más allá de Web de la Ciencia (WoS)

Collection and registration of author affiliation data beyond the Web of Science

Peter Aspeslagh, ECOOM-University of Antwerp

Universidad Andina del Cusco, Cusco, Agosto 12, 2022

Universidad San Marcos, Lima, Agosto 18, 2022



Datos de afiliación de autor más allá de Web of Science en Flandes

- Proporcionar datos para el parámetro 'internacionalización' en el contexto de la distribución de fondos de investigación (Ciencias sociales y humanísticas, basado en publicaciones)
- Publicaciones no incluidas en WoS: ¿cómo recopilar datos de afiliación de autor de forma alternativa?
- Spin-off : base de datos de organización extendida para múltiples propósitos



Belgica

- País con estructura federal

Flandes

- Evaluación de la investigación y distribución de fondos: competencia regional (Gobierno Flamenco)
- Instrumento de distribución de fondos: resultado de investigación en base de datos bibliográfica nacional para publicaciones en Ciencias Sociales y humanísticas (VABB-SHW)

5 universidades

- University of Leuven (KU Leuven)
- University of Antwerp (UAntwerp)
- Ghent University (UGhent)
- Hasselt University (UHasselt)
- Free University of Brussels (VUB)

1. Parámetro 'internacionalización' en el sistema Flamenco de financiación de la investigación basado en el rendimiento (PFRS)



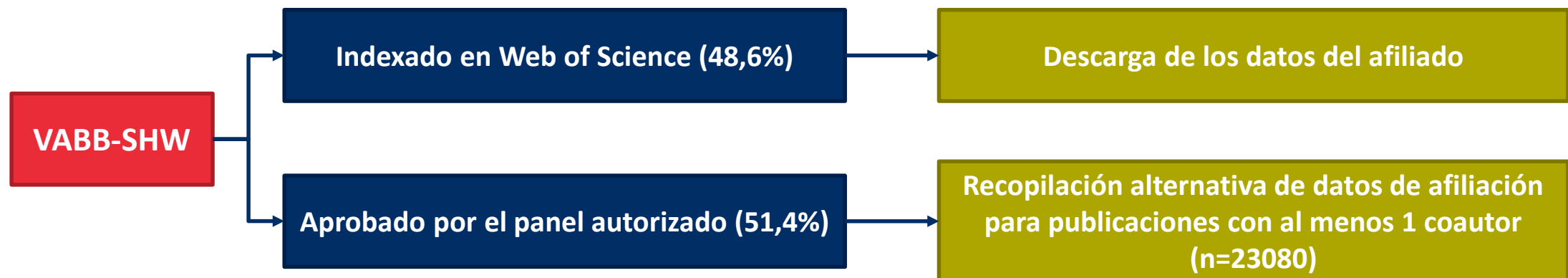
Parámetro 'internacionalización'

- Flandes: sistema de financiación de la investigación basado en el rendimiento (*performance-based research funding system = PFRS*)
- Clave BOF (*Bijzonder OnderzoeksFonds = Fondo Especial de Investigación*) con múltiples parámetros utilizados como criterio para distribuir fondos
- Instrumento: base de datos bibliográfica nacional con resultados de investigación en Ciencias Sociales y Humanísticas (publicaciones – VABB-SHW) de fuente de datos compilados en los cinco universidades por diferente parámetros
- Gestionado en ECOOM-University of Antwerp
- Nuevo parámetro desde 2019: internacionalización
 - Las publicaciones con al menos 1 coautor afiliado a una institución extranjera (es decir, no belga) reciben una mayor ponderación en el cálculo de la distribución de fondos
- Afiliaciones de autor como fuente de datos

1. Parámetro 'internacionalización' en el sistema Flamenco de financiación de la investigación basado en el rendimiento (PFRS)

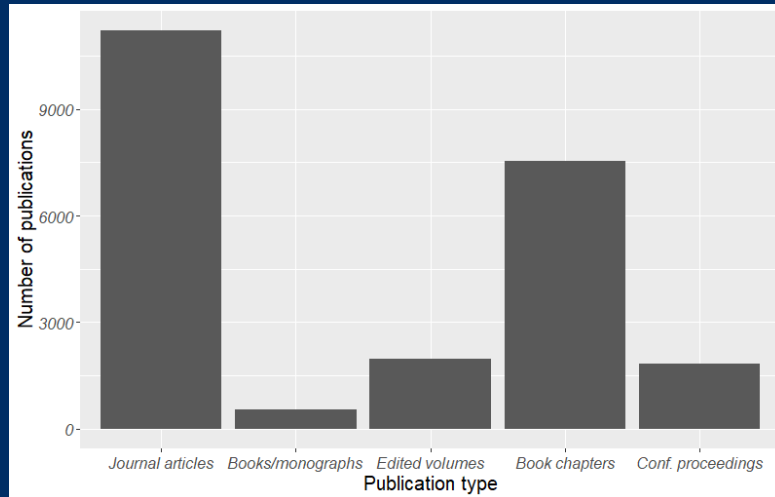
VABB-SHW & datos de afiliación del autor

- La implementación del parámetro 'internacionalización' requiere datos de afiliación del autor para un período de 10 años (a-11 to a-2) (2011-2020)
- Los datos de afiliación del autor no están disponibles en VABB-SHW y deben recopilarse
- Composition of VABB-SHW:
 - El 48,6% de las publicaciones están indexadas en la WoS. **Descarga instantánea** de los datos de afiliados
 - 51,4% aprobado por un panel autorizado (GP). Es necesaria **una recopilación de datos alternativa**
- El 51,5% de las publicaciones del panel autorizado (GP) contienen al menos un coautor. En total, los datos de afiliación del autor deben recuperarse alternativamente para 23080 publicaciones.

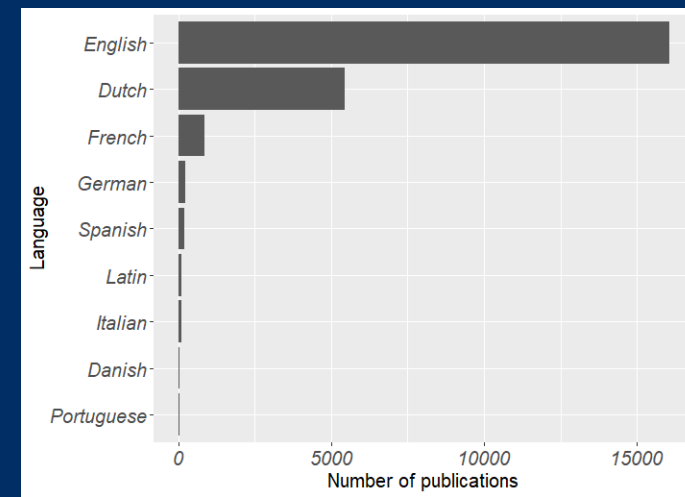


2. Recopilación alternativa de datos de afiliación: pasaporte de publicaciones

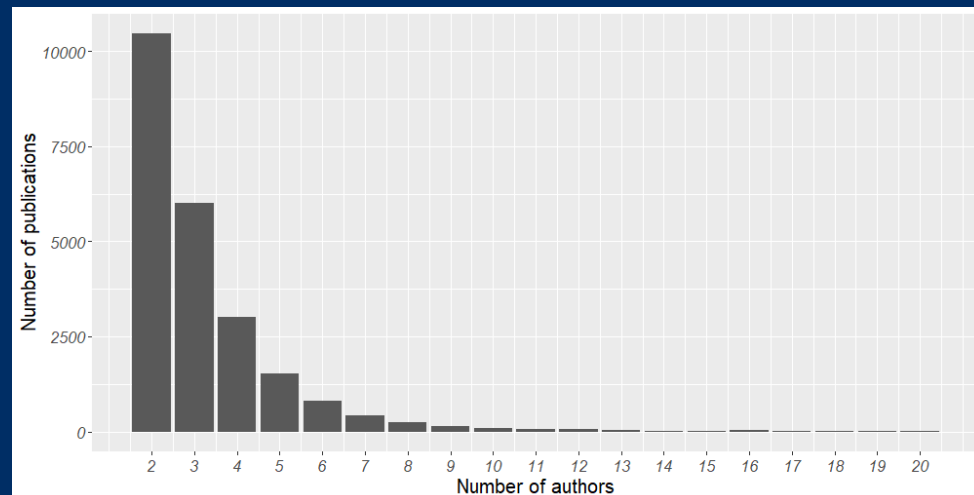
Por tipo de publicación



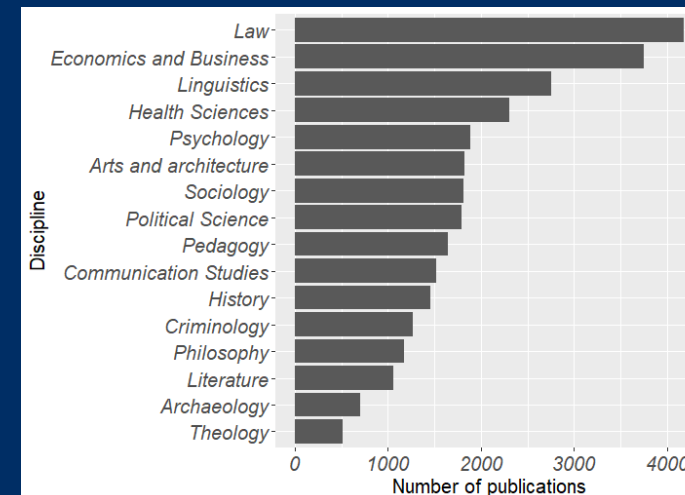
Por idioma



Por número de autores



Por disciplina (clasificación organizacional)

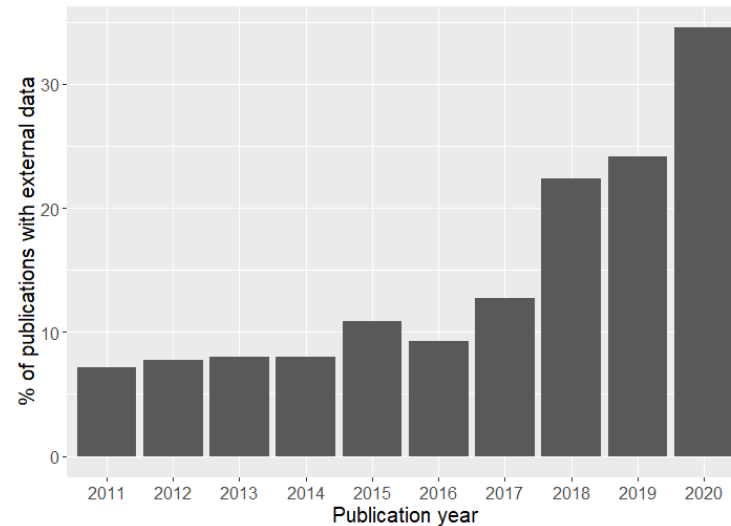


2. Recopilación de datos de afiliados alternativos : quid?

Descargar datos de bases de datos bibliográficas que no sean de WoS (14,9%)



Scopus



- Fuente de datos : Scopus, Crossref, (ESCI, BkCI)*, (Microsoft Academic)
- El número de publicaciones con datos de afiliación externa está aumentando (aunque parte de las publicaciones de 2011-2017 ya se codificaron manualmente antes de que se realizara la comparación con la base de datos externa).
- Países extraídos; organizaciones convertidas en identificadores
- Datos de afiliación a nivel de publicación

Recopilación manual de datos (85,1%)

- Fuente de datos: publicación real o (DOI-redireccionado) página web oficial de la publicación
- Codificación manual mediante aplicación en línea (países y organizaciones)
- Datos de afiliación a nivel de publicación
- Estrategias de búsqueda detalladas

* Sin embargo, algunas publicaciones aprobadas por el panel autorizado (GP) están incluidas en las bases de datos ESCI y BkCI afiliadas a WoS, que se establecieron después del lanzamiento de VABB-SHW. Por lo tanto, pudimos descargar datos afiliados al autor de ESCI y BkCI para un número (limitado) de publicaciones.

2. Recopilación de datos de afiliación alternativa (manual): validez y estrategias de búsqueda

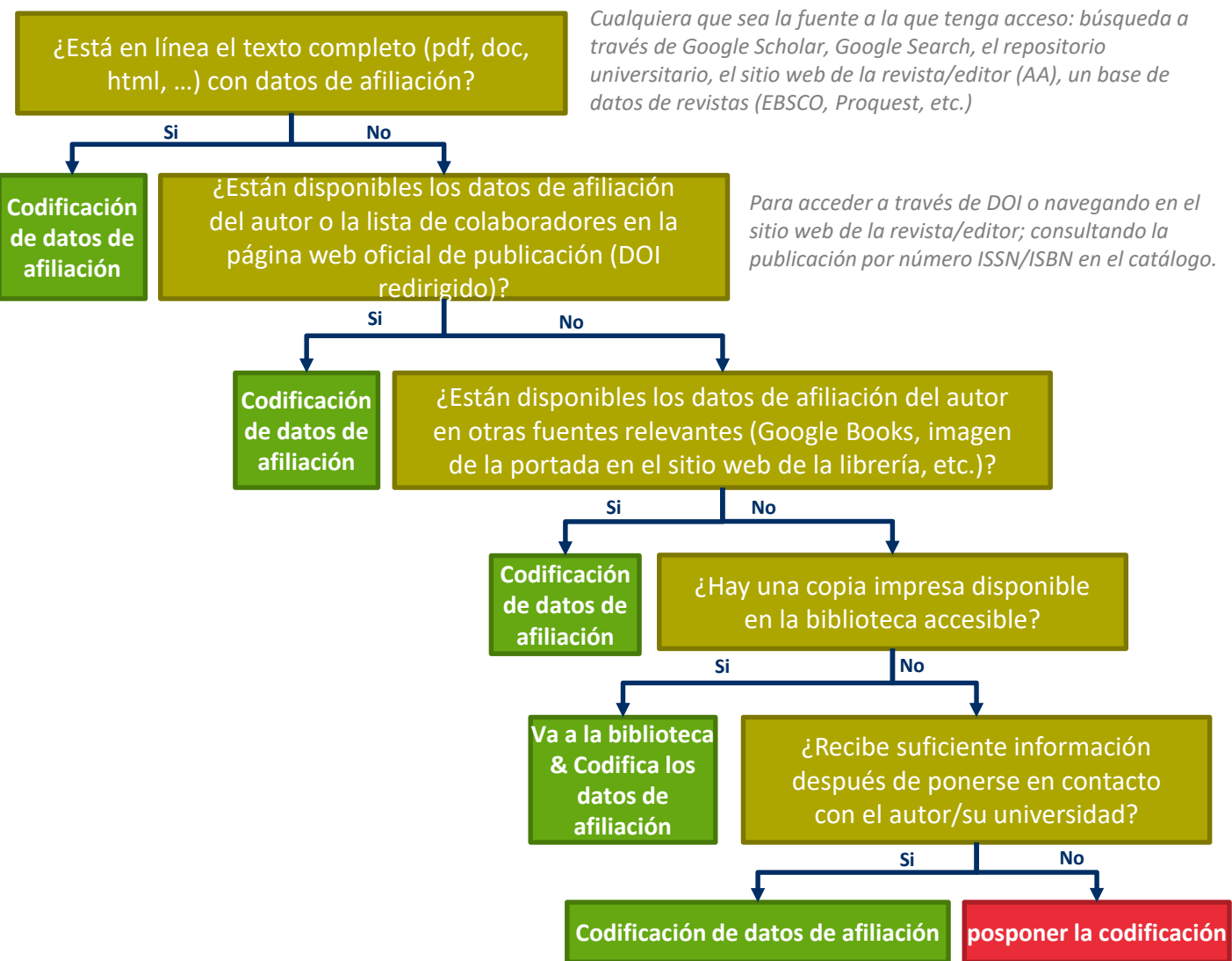
¿Qué son los datos de afiliación válidos?

Datos válidos de afiliación del autor: la mención explícita de una organización (universidad, organización sin fines de lucro, agencia gubernamental, empresa, etc.) a la que el autor estaba afiliado en el momento de la publicación. El mero nombre de una ciudad es insuficiente.

Como **prueba de los datos** que se deben proporcionar a las autoridades de financiación, un documento pdf del texto completo debe estar disponible o los datos de afiliación en el sitio web de la revista/editor. En el caso de Google Books o portadas de libros, se deben tomar capturas de pantalla del documento.



¿Cómo puedo recuperar los datos?



2. Recopilación alternativa (manual) de datos de afiliación: registro



- Una vez que se encuentran los datos correctos de afiliación del autor, los datos se registran a través de un formulario web
- Los datos de afiliación se pegan en un cuadro de texto. Posteriormente se codifica por país y organización (seleccionado de una lista de más de 100.000 organizaciones – cf. infra)
- Una pestaña permite el almacenamiento correcto de pdf
- Al recuperar datos de afiliación, se pueden agregar/modificar DOI y resúmenes faltantes o incorrectos (enriquecimiento de datos)
- Se requiere una revisión y validación final
- El conjunto de publicaciones a codificar se enumeran alfabéticamente, por ISSN o por prefijo ISBN.

Challenges to comparative oral histories of cinema audiences [journal article] [Download fulltext \(pdf\)](#)

Retrieve data 1: Affiliation data 2: DOI 3: Save fulltext 4: Abstract 5: Summary & finish

Retrieve data Validate Completed Step 2: check DOI

1. Affiliation data (manually added or external sources) [cvabb:495526]

You can register the affiliation data in this form. First, paste the affiliation data as available on the fulltext or on the website of the publication on the left side of you screen (A). Indicate if the points below apply. Subsequently, code the affiliation data by distilling the countries and affiliated organizations from the text you pasted (B). Finally, validate the data by pushing the 'Validate' button on the top right of your screen if authorized by the ECOM staff.

A. Paste affiliation data & notifications

Save pasted affiliation data and notifications

Paste affiliation data (textual)

Pierluigi Ercole (1)
Lies Van de Vijver (2,3)
Daniela Treveri Gennari (3)

1 - De Montfort University (Leicester, UK)
2 - Ghent University
3 - Oxford Brookes University

Source: Actual publication

No affiliation data
No affiliation data due to missing page in fulltext
Partial affiliation data (at least one foreign)

B. Code affiliation data

Assign country: -- Belgium -- Add

Assign institution [GROUP]: Add organization... Add

Countries

Belgium	Delete
United Kingdom	Delete

Organizations

Organization	Department	Delete	Update
Ghent University		Delete	Update
De Montfort University		Delete	Update
Oxford Brookes University		Delete	Update

2. Recogida de datos de afiliación alternativa: situación actual

- El 92,88% de las 23080 publicaciones que no están en WoS y con al menos un coautor están codificadas
- Investigadores flamencos de Ciencias Sociales y Humanidades colaboran con colegas de 141 países y 5566 * organizaciones diferentes
- La primera fase retroactiva está casi completa; cada año habrá que codificar unas 3000 publicaciones nuevas.
- A partir del próximo ciclo, la codificación se realizará a nivel de autor de publicación.

* Debido a la conversión de GRID a ROR, se omitió menos del 1% del número total de organizaciones por razones prácticas.



2. Recogida de datos de afiliación alternativa: primeras conclusiones

- Falta una gran parte de los datos de afiliación de los autores cuando solo se concentra en las bases de datos bibliográficas comunes. En particular, se pierden datos sobre producción de investigación local y publicaciones no anglosajonas.
- Muchos datos de afiliación de autores están disponibles en línea, pero no están estructurados y no son interoperables.
- Los datos de afiliación, al menos en la lista de colaboradores, idealmente siempre deberían incluirse en las vistas previas, en los sitios web de revistas/editores, en servicios como Google Books, etc.
- Sensibilizar a los investigadores sobre la importancia de los datos de afiliación. Los datos a menudo se olvidan o están incompletos.

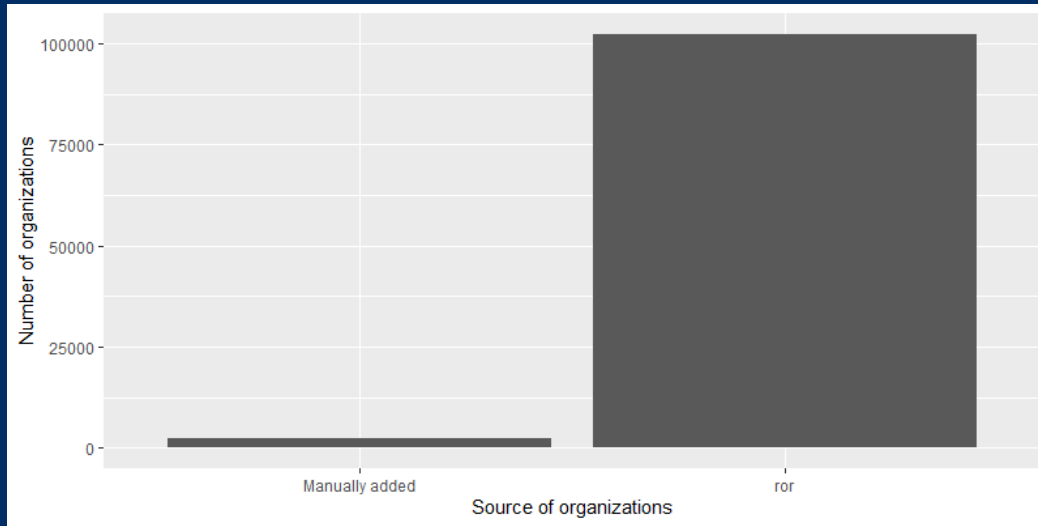
3. Base de datos de organizaciones : Spin-off



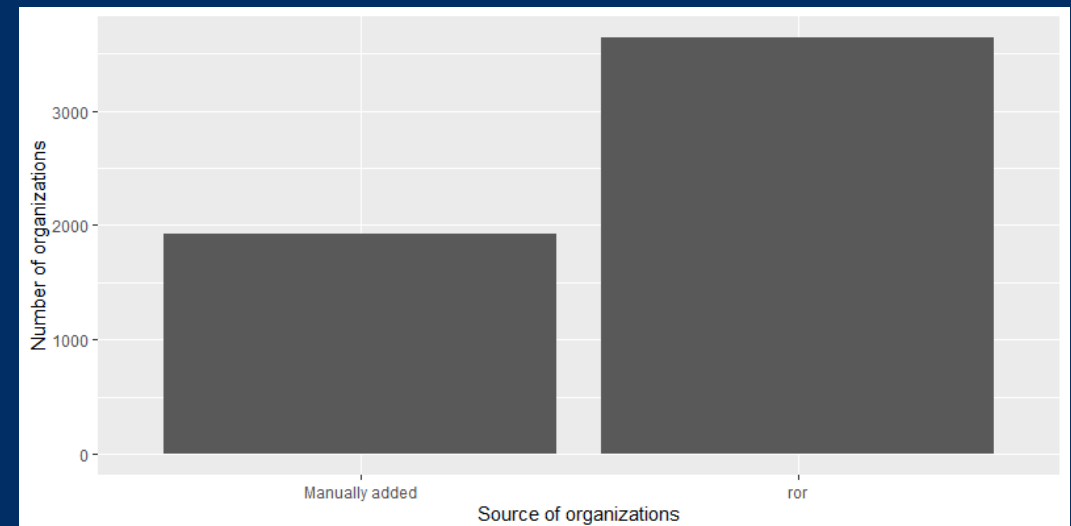
- Registro de organizaciones afiliadas: conectar identificador de una organización derivado de una lista de más de 100.000 organizaciones
- Fuente de datos de las organizaciones: Registro de Organizaciones de Investigación (ROR) (anteriormente: GRID)
- ROR no captura todas las organizaciones afiliadas; necesidad de agregar manualmente nuevas organizaciones (no ROR)
- Resultado: base de datos extendida de organizaciones que se puede utilizar para múltiples propósitos. “ROR+” como resultado del proyecto de recopilación de datos de afiliación, que abarca también a las organizaciones locales
- Las organizaciones nuevas/locales se pueden codificar de acuerdo con el modelo de datos ROR+
- Uso potencial: herramientas administrativas, proyectos de investigación, ...
- No tenemos la ambición de reemplazar ROR, sino ampliar el conocimiento sobre las organizaciones y compartirlo en algún momento.

3. Base de datos de organizaciones: composición

Número total de organizaciones por fuente de datos (ROR y adición manual)



Número total de organizaciones utilizadas durante la codificación de datos de afiliación por fuente de datos (ROR y adición manual)



- La base de datos de la organización consiste en un total de 104652 organizaciones únicas (102424 en ROR (97,87%), 2228 agregadas manualmente (2,13%))*
- Si solo seleccionamos las organizaciones únicas que se utilizaron durante la codificación de los datos de afiliación (n=5566)*, las proporciones cambian: el 65,47% estaban cubiertos por ROR; 34,53% se agregaron manualmente
- La adición de nuevas organizaciones es necesaria debido a la naturaleza de la base de datos de publicaciones: una gran cantidad de publicaciones no anglosajonas y afiliaciones locales posteriores

* Debido a la conversión de GRID a ROR, se omitió menos del 1% del número total de organizaciones por razones prácticas.

3. Base de datos de organizaciones: enriquecimiento de datos y modelo de datos extendido

Enriquecimiento de la base de datos de organizaciones: agregar información nueva/modificada al registros organizacionales ROR/agregados manualmente (identificadores, etiquetas, acrónimos, línea de tiempo, archivos, ...) en la aplicación.

The screenshot displays the ROR application interface for the University of Antwerp. The interface is divided into several sections:

- General data:** Name: University of Antwerp, Extra information, External database: ror, External database ID: <https://ror.org/008x57b05>, Location on Google Maps.
- Contact data:** [Contact 1] [Edit], Municipality: Antwerp (Geonames) [Google Maps], Country: BE.
- ROR Relations & ECOOM Relations:** Overview, Child(ren) of (Antwerp Management School, VIB-UAntwerp Center for Molecular Neurology), Other related items of (Related) (Antwerp University Hospital).
- Related publications due to affiliations:** 4537 publication(s).
- Timeline:** Overview, Date, Event.
- ROR Data & ECOOM Enriched data:**
 - General:** Established: 2003, Type [Overview]: Education.
 - Labels [Overview] [+]:** Table with columns: Type, Language, Label, Source.

Type	Language	Label	Source
acronym	eng	UA	ECOOM
acronym	dut	UA	ECOOM
alias	NA	UAntwerp	ROR
label	spa	Universidad de Amberes	ECOOM
label	dut	Universiteit Antwerpen	ROR
label	fre	Université d'Anvers	ROR
label	ger	Universität Antwerpen	ROR
 - Links [Overview]:** Table with columns: Title, Link.

Title	Link
University of Antwerp	https://www.uantwerpen.be/en/
 - External identifiers [Overview] [+]:** Table with columns: Source, Identifier.

Source	Identifier
Commercial Register BE (KBO)	0257216482
FundRef	501100007660
GRID	grid.5284.b
ISNI	0000 0001 0790 3681
OrgRef	844674
Wikidata	Q175241
Wikipediaurl	University_of_Antwerp

3. Base de datos de organizaciones: enriquecimiento de datos y modelo de datos extendido

Adición de archivos sobre la organización

Subir pdf, docx, xlsx, imagen etc.

The screenshot shows the left sidebar of the ROR interface for the 'University of Antwerp' profile. The 'Files' menu item is highlighted with a yellow box. Other menu items include Functions, Pages, Timeline, Categories, Connections, Contacts, References, Metadata, ID Management, Labels, Websites, Map, Relations, and Publications with affiliations.

Descripción de la organización

Información básica: Fecha de constitución, fecha de cese

The screenshot shows the ROR profile page for the University of Antwerp. Several sections are highlighted with yellow boxes: 'General data' (Name: University of Antwerp, External database: ror, External database ID: https://ror.org/008x57b05, Location on Google Maps), 'Contact data' (Municipality: Antwerp, Country: BE), 'ROR Relations & ECOOM Relations' (Child(ren) of: Antwerp Management School, VIB-UAntwerp Center for Molecular Neurology; Other related items of (Related): Antwerp University Hospital), 'Related publications due to affiliations' (4537 publication(s)), and 'Timeline' (Date, Event).

Datos de contacto: 1-n-relación

Una organización puede tener múltiples direcciones. Oportunidad de registrarlos. Indicando dirección principal

Extendiendo las relaciones

Relaciones con y entre organizaciones adicionales
Posibles nuevos tipos de relaciones

Añadir la cronología

Mapeo de la historia de la organización por fecha de inicio y (opcional) fecha de finalización

Exportación/intercambio de datos/API

URL con metadatos en JSON, datos de la organización accesibles a través de ROR ID, identificador interno, nombre

The screenshot shows the left sidebar of the ROR interface. The 'JSON Export URL' menu item is highlighted with a yellow box. Other menu items include JSON Export URL Aff-Pub.

3. Base de datos de organizaciones: enriquecimiento de datos y modelo de datos extendido

Copy Ext ID Copy Sys ID Add keyword... Add

ROR Data & ECOOM Enriched data

General

Established: 2003

Type [Overview]: Type
Education

Labels [Overview] [+]:

Type	Language	Label	Source
acronym	eng	UA	ECOOM
acronym	dut	UA	ECOOM
alias	NA	UAntwerp	ROR
label	spa	Universidad de Amberes	ECOOM
label	dut	Universiteit Antwerpen	ROR
label	fre	Université d'Anvers	ROR
label	ger	Universität Antwerpen	ROR

Links [Overview]:

Title	Link
University of Antwerp	https://www.uantwerpen.be/en/

External identifiers [Overview] [+]

Source	Identifier
Commercial Register BE (KBO)	0257216482
FundRef	501100007660
GRID	grid.5284.b
ISNI	0000 0001 0790 3681
OrgRef	844674
Wikidata	Q175241
Wikipediaurl	University_of_Antwerp

Adición de palabras clave

- Mapeo más detallado de la organización
- Conjunto general de palabras clave o dependiente del proyecto

Categorías adicionales

- ROR(GRID) tiene un conjunto limitado de categorías; el número de categorías se puede ampliar aquí.

Siglas, etiquetas y alias adicionales

- Acrónimos, etiquetas y alias integrados en una tabla, distinguidos por tipo de etiqueta
- Fecha de inicio y fecha de finalización para cada acrónimo, etiqueta y alias
- Variable de idioma también disponible para acrónimos
- Indicando la fuente de datos (adicionales) (ROR, ECOOM, ...)

Enlaces: 1-n-relación

- Las organizaciones pueden tener un sitio web multilingüe; se pueden agregar enlaces a múltiples páginas de inicio

Identificadores externos adicionales

- Tipos adicionales de identificadores
- Ej.: número de registro mercantil

Observaciones finales

- Se requieren datos de afiliación de autor de publicaciones que no sean de WoS para el parámetro de distribución de fondos de investigación
- Recopilación alternativa de datos: otras bases de datos bibliográficas; recopilación manual de datos en particular
- La recuperación de documentos precisos requiere estrategias de búsqueda específicas
- Hay muchos datos de afiliación disponibles, pero no estructurados.
- El registro de los datos generó un spin-off: una base de datos extendida de organizaciones (de investigación)
- Desarrollo de la base de datos “ROR+” como proyecto separado para múltiples propósitos

Muchas Gracias!

E-mail: peter.aspeslagh@uantwerpen.be

ORCID: 0000-0002-2753-3144

ResearchGate: <https://www.researchgate.net/profile/Peter-Aspeslagh>