"Women and Wikipedia: Analysis of the Gender Gap in the Co-Production of Knowledge in the Wikipedias of Spain (W&W)" (Ref.PID2020-116936RA-I00).

# ¿Qué se investiga en la brecha de género en Wikipedia y qué soluciones propone la academia?

## Revisión de artículos de investigación publicados entre 2005-2022

Núria Ferran-Ferrer, Juan-José Boté, Julià Minguillón

---

Wikimedia + Gendergap Research Meetup, MediaLab, Madrid, 30 marzo 2023

Universitat Oberta de Catalunya

Universitat de Barcelona
Facultat d'Informació i Mitjans Audiovisuals

Ministerio de Ciencia e Innovación
Agencia Estatal de Investigación

# Proyecto 2022-24 www.ub.edu/wikiwomen




## OBJETIVO PRINCIPAL

¿Cómo las mujeres pueden de forma efectiva acceder, editar y transformar Wikipedia?

## PREGUNTAS DE INVESTIGACIÓN

- **¿Dónde editan las mujeres de las Wikipedias ESP, EUSK, GAL y CAT?**
- **¿Qué diferentes tipos de comportamientos tienen los hombres y las mujeres en WP?**
  - *19h hoy  Julià Minguillón*
- **¿Cómo es el proceso de negociación de los artículos nominados al borrado?**
- **¿Cuales son las fuentes de información que fundamentan el criterio de relevancia (notability)?**
  - *12:30h David Rodríguez-Ordóñez (tesis)*
- ***¿Cómo acceden y persisten las mujeres editoras?***
  - 15h Yessica Macià: mujeres sin categoría
  - 15:45h Miquel Centelles: Wikidata al rescate

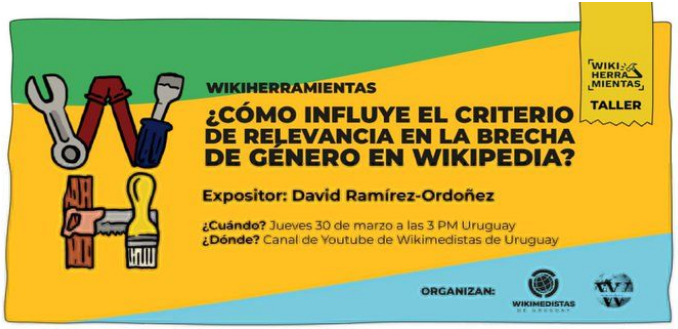
# + EQUIPO



- Laura Fernández, postdoc
  - Portadas WP ESP, ENG, CAT
  - Identidades de género no binarias y WP
- Maite Comalat y Aurora Vall
  - Etnografía jóvenes editoras
- Estudiantes de TFM Master Humanidades Digitales y Master de Contenidos Digitales (mar-jun 2023):
  - Daniella Cardona (UX interacciones),
  - Andrés Bejarano (portadas y interseccionalidad cultural
  - Francisco Kugler (jerarquías y toma de decisiones)
  - Shuwen Tan (categorías)
  - Ran Zhang (fuentes de información e interseccionalidad cultural).

# ACTIVIDADES

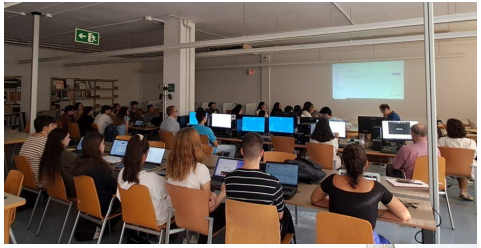












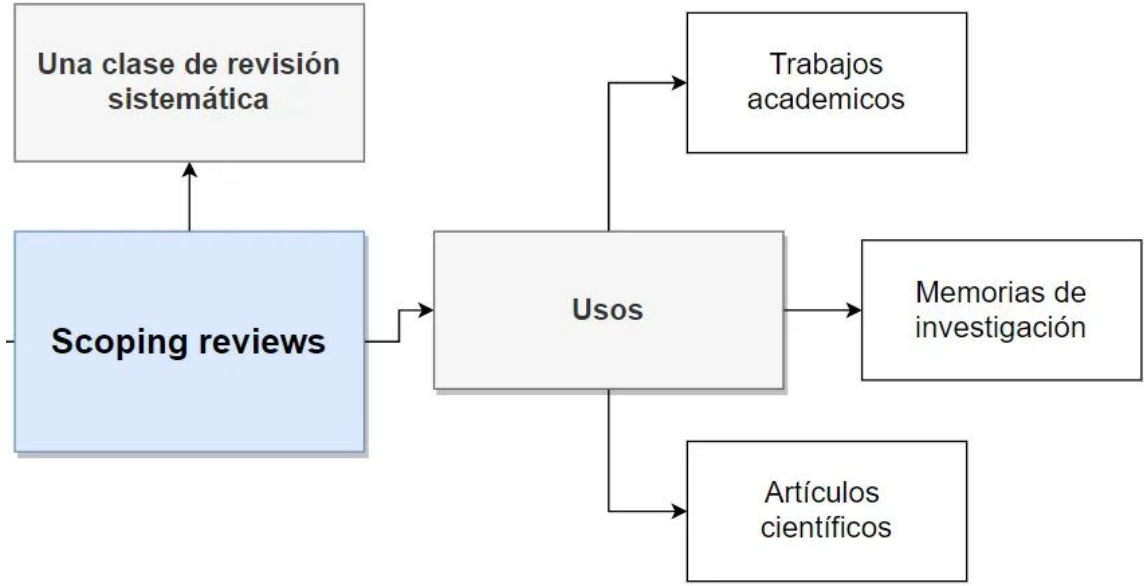

# Tabla de contenidos

Revisión sistemática de la literatura: Scoping Review

- Preguntas de investigación
- Metodología: SALSA framework
  - Búsqueda
  - Evaluación
  - Análisis
  - Síntesis: Resultados
    - Aproximación cuantitativa
    - Aproximación cualitativa



# Revisión sistemática de la literatura: Scoping Review




- Scoping review=revisión sistemática exploratoria
- Síntesis de la evidencia
- Objetivo: explorar de manera amplia (scope) un ámbito de conocimiento de la investigación científica
- El conocimiento analizado es heterogéneo en cuanto a métodos y de disciplinas

Fuentes: Codina, Lluís (2021). Scoping reviews: características, frameworks principales y uso en trabajos académicos
Codina, Lluís (2021). ¿Qué son las scoping reviews y por qué son importantes?

# Preguntas de investigación




- Las *scoping reviews* tienen preguntas de investigación amplias, en nuestro caso:

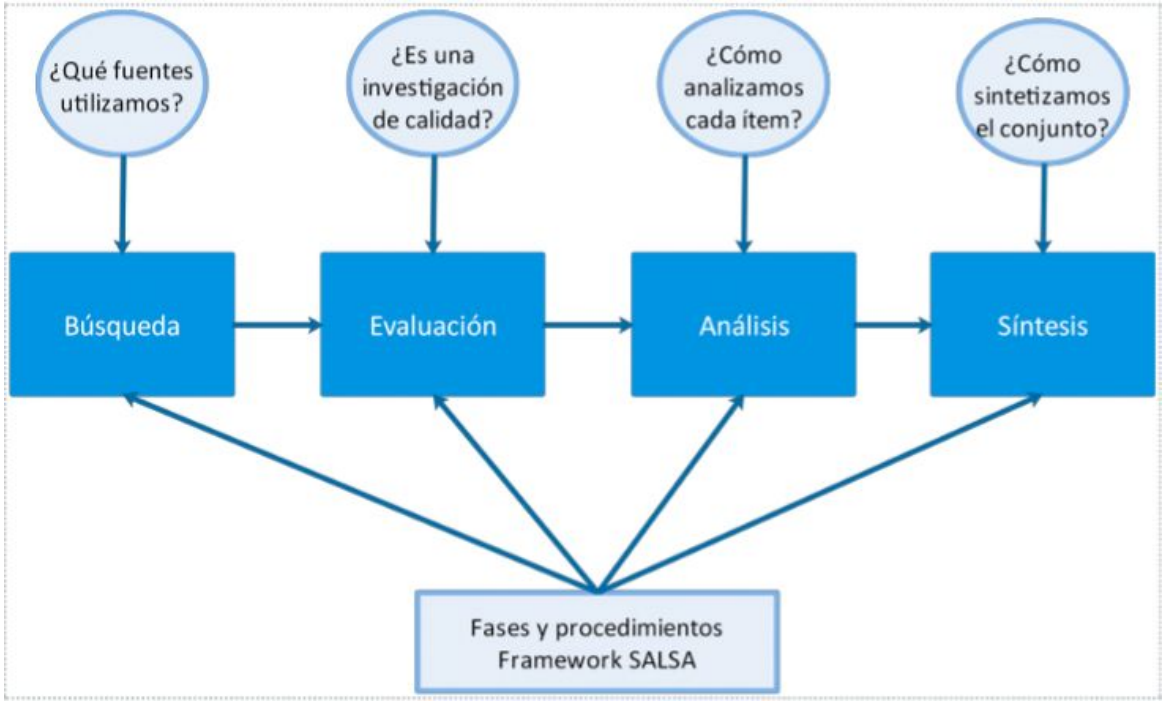
  - ¿Cuál es la **naturaleza** de la investigación **en el ámbito de conocimiento de la brecha de género en Wikipedia?**
    - Métodos de investigación
    - Países
    - Disciplinas
    - Entidades financiadoras
    - Perspectiva género (binaria o no-binaria)
    - …

Aproximación cuantitativa

  - ¿Cuáles son las **razones** de la brecha de género en WP según la academia?
  - ¿Qué **soluciones** prevé la academia para la brecha de género en WP?
  -

Aproximación cualitativa

# Selección de la metodología. SALSA framework




Fuente: Lopezosa, Carlos; Codina, Lluís; Ferran-Ferrer, N. (2023). ChatGPT como apoyo a las systematic scoping reviews: integrando la inteligencia artificial con el framework SALSA. CRICC: UB

**Protocolo SALSA**: permite obtener de manera rigurosa la base de la evidencia

**SALSA: S**earch, **A**ppraisa**L**, **A**nalysis y **S**ynthesis

**Las cuatro preguntas de chequeo y fases de una revisión sistemática**

Una revisión de la literatura se puede enfocar como una investigación que incluye **toma de datos** porque las fases están bien definidas y disponen de **procedimientos rigurosos y transparentes para la selección y análisis de la base de la evidencia**

# Búsqueda (Search)




- La fase de búsqueda se debe resolver mediante la **identificación y selección de las fuentes** que van a proporcionar la base de la evidencia.
- Establecer claramente los métodos de obtención de documentos deben quedar bien establecidos.

- Fuentes: Bases de datos académicas **WoS y Scopus**
- Base de la evidencia: **artículos de investigación** en revistas académicas y **capítulos** de libro
- Método: **ecuaciones de búsqueda** que se ajustan a los objetivos de la investigación

```
TITLE-ABS-KEY ( wikipedia AND ( gender* OR female* OR male* OR woman OR man OR women OR men OR masculi* OR femini* OR sex OR non-binary OR intersex OR transgen* OR agender OR queer* ) )
```

**Des de inicio hasta 2022 todo incluido. Búsqueda realizada 13/01/2022**

```
Num docs Scopus: 420
Num docs WoS: 399
330 duplicates

488 unique items
```

# Evaluación (AppraisaL)

- Esta fase de evaluación se requiere para asegurarnos de que la base de la evidencia responde a los requerimientos del estudio.

- Se realiza a través de una aplicación de criterios de inclusión y exclusión, lo que suele incluir la temática, el rango de fechas de publicación, la metodología y en definitiva la verificación de la calidad de cada documento identificado en la fase anterior.

- Exlusion language (ENG y ESP): 0
- Exclusion non article or chapters: 6
- Exclusion non academic articles: 17
- Exclusion out of period analysis: 1
- Exclusion full text not available: 1
- Duplications: 7
- Thematic exclusion screening process
  - No gender perspective: 31
  - No WP: 14
  - WP No object of study: 352

Initial num docs: 488
Exclusion: 390

**Total included docs: 98**

# Análisis (Appraisal)

- Consiste en el diseño de los **esquemas de extracción de datos** con los cuales van a ser analizados de forma sistemática todas y cada una de las referencias que forman la base de la evidencia.

  - Discipline
  - Autor: number, gender, affiliation and country
  - Topic: participation/readership/editing
  - Object of the study
  - Objectives
  - Questions/hypotheses
  - Research Strategy
  - Methodology
  - Results
  - Further Research
  - Limitations
  - Gender perspective
  - Reasons Gender Gap
  - Solutions Gender Gap

Uso Atlas.TI para hacer el análisis del contenido.

# Síntesis (Syntesis)




- En esta fase se presentan los **resultados principales** a través de tablas o matrices, así como de resúmenes narrativos

- El objetivo es presentar una síntesis que sea algo más que la suma de las partes, para lo cual se requiere una presentación de resultados que no sea meramente agregativa, sino que incluya configuración e interpretación de resultados

- En particular, se espera la identificación de patrones, la de temas emergentes y la de huecos y oportunidades de investigación

- Trial ChatGPT: *Lopezosa, Carlos; Codina, Lluís; Ferran-Ferrer, N. (2023). ChatGPT como apoyo a las systematic scoping reviews: integrando la inteligencia artificial con el framework SALSA. CRICC: UB*

# Limitaciones expuestas




En cualquier búsqueda de documentos siempre existe la probabilidad de algún sesgo:

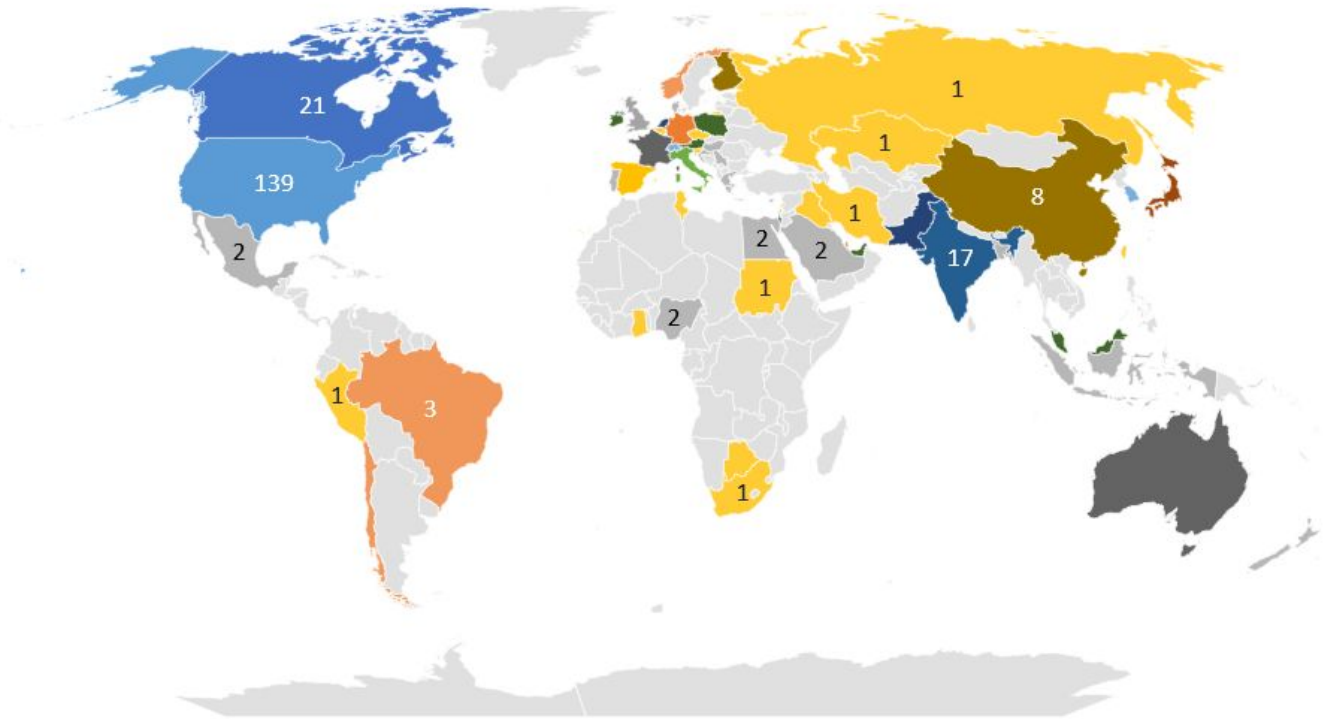
- No abarcar todos los artículos posibles
- Descuidar términos o elementos booleanos en la ecuación de búsquedas
- Diferencias de criterio a la hora de seleccionar documentos entre el personal investigador
- No considerar otras bases de datos




RESULTADOS:

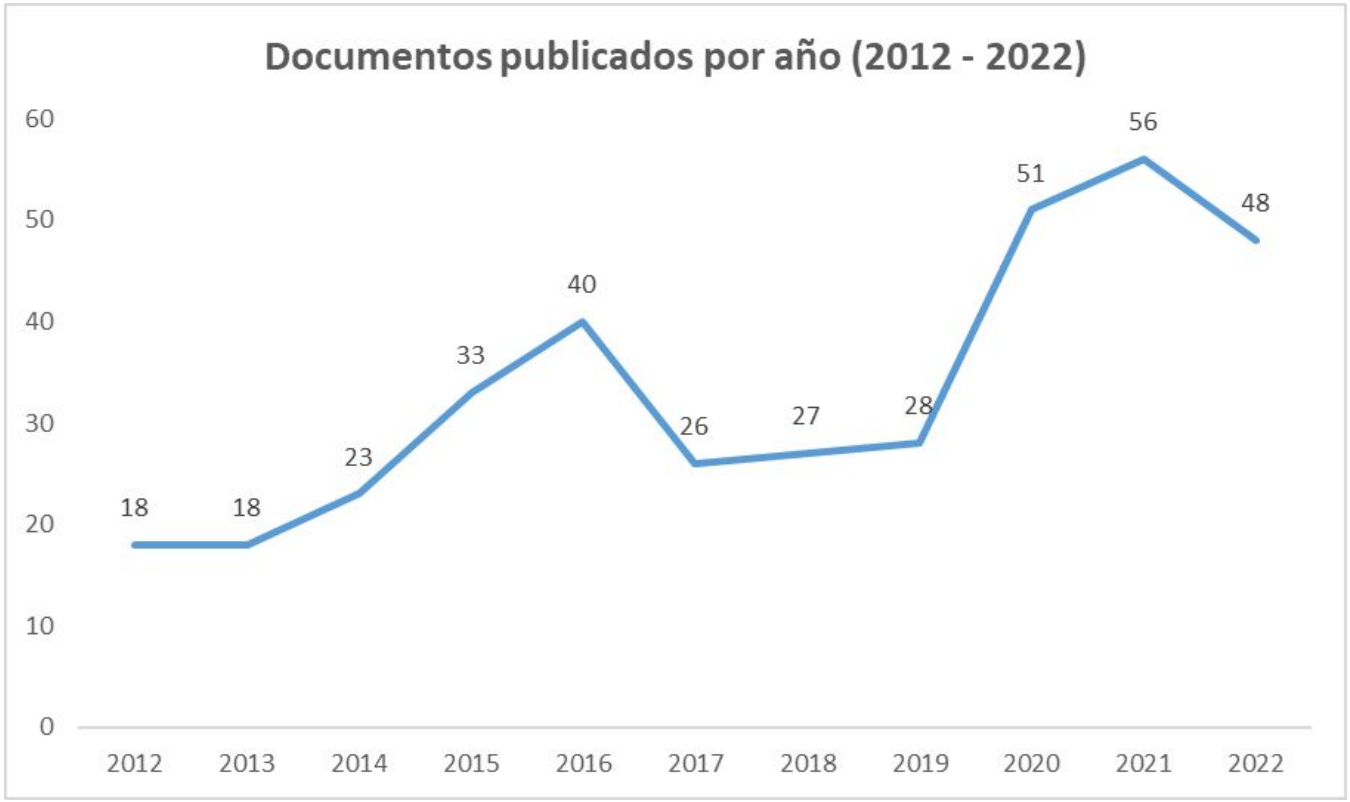
¿Cuál es la naturaleza de la investigación en el ámbito de conocimiento de la brecha de género en Wikipedia?

# Producción por paises en Scopus (2012 - 2022)




Con tecnología de Bing
© Australian Bureau of Statistics, GeoNames, Geospatial Data Edit, Microsoft, Navinfo, OpenStreetMap, TomTom, Wikipedia, Zenrin

# Documentos publicados por año




2002-12: 150 artículos
**2012-22: 270**

2001 Nace Wikipedia

2007: 1r artículo gender Art+Feminism

2008 WMF encuesta: 13% editoras mujeres

2009 Primera revisión literatura de Okoli et al. sin perspectiva de género

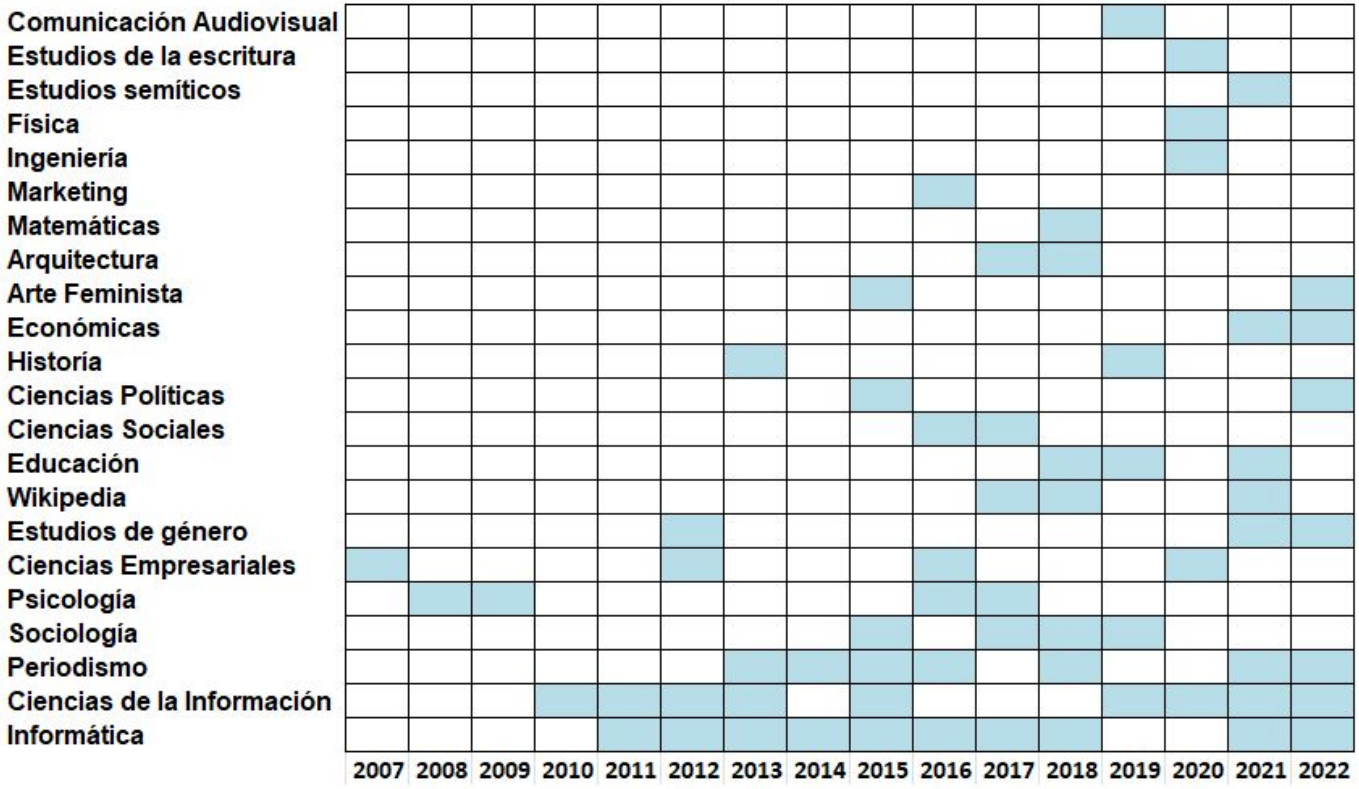
**2011 Sue Gadner Executive Dr WMF official efforts female-friedly environment**

2011 Lam et al.: brecha de género entre Wikipedistas

2012: Nace Wikidata

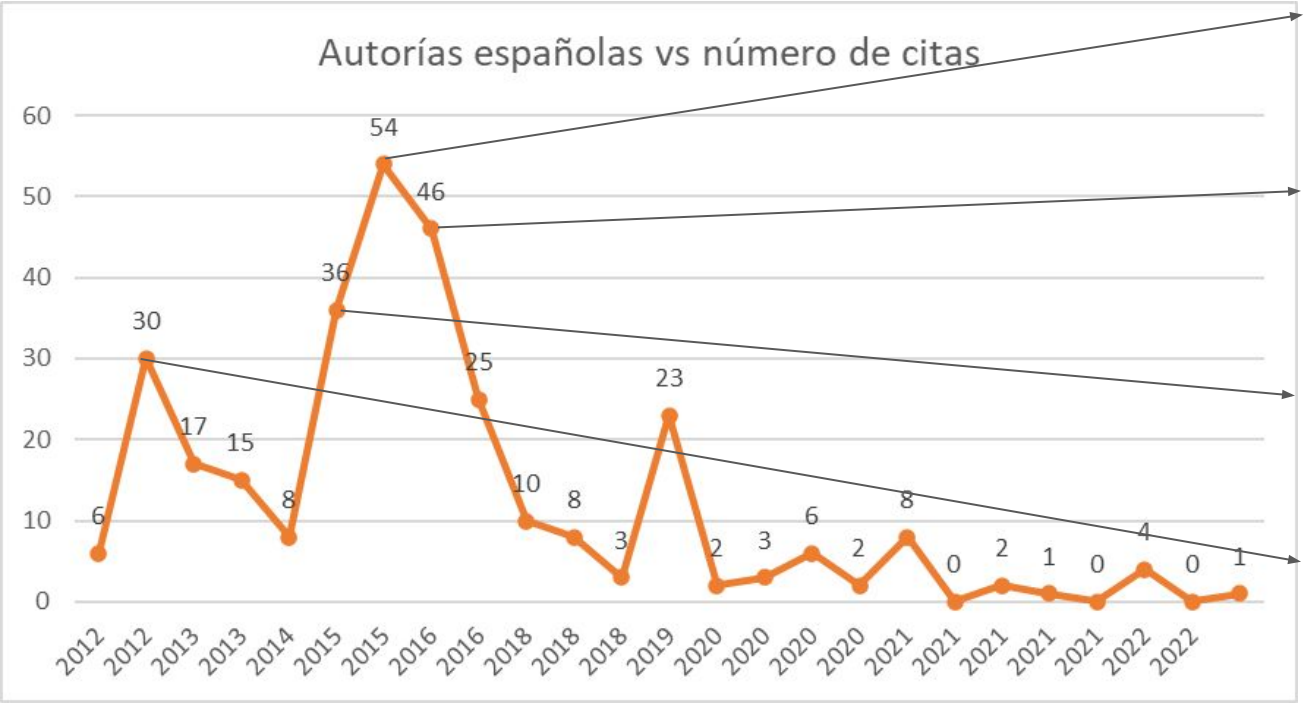
2015: WP no consigue 25% contenidos de mujeres

# Disciplinas




Se omiten múltiples publicaciones en un mismo año

# Documentos publicados por año.
# Autorías españolas




Graells-Garrido, Eduardo, Mounia Lalmas, and Filippo Menczer. ‘First Women, Second Sex: Gender Bias in Wikipedia’. In Proceedings of the 26th ACM Conference on Hypertext & Social Media - HT ’15, 165–74. Guzelyurt, Northern Cyprus: ACM Press, 2015. https://doi.org/10.1145/2700171.2791036. (54 citas)

Eom, Young-Ho, Pablo Aragón, David Laniado, Andreas Kaltenbrunner, Sebastiano Vigna, and Dima L. Shepelyansky. ‘Interactions of Cultures and Top People of Wikipedia from Ranking of 24 Language Editions’. Edited by Zhong-Ke Gao. PLOS ONE 10, no. 3 (4 March 2015): e0114825. https://doi.org/10.1371/journal.pone.0114825. (46 citas)

Iosub, Daniela, David Laniado, Carlos Castillo, Mayo Fuster Morell, and Andreas Kaltenbrunner. ‘Emotions under Discussion: Gender, Status and Communication in Online Collaboration’. Edited by Pablo Brañas-Garza. PLoS ONE 9, no. 8 (20 August 2014): e104880. https://doi.org/10.1371/journal.pone.0104880. (36 citas)

Laniado, David, Andreas Kaltenbrunner, Carlos Castillo, and Mayo Fuster Morell. ‘Emotions and Dialogue in a Peer-Production Community: The Case of Wikipedia’. In Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration, 1–10. Linz Austria: ACM, 2012. https://doi.org/10.1145/2462932.2462944. (30 citas)

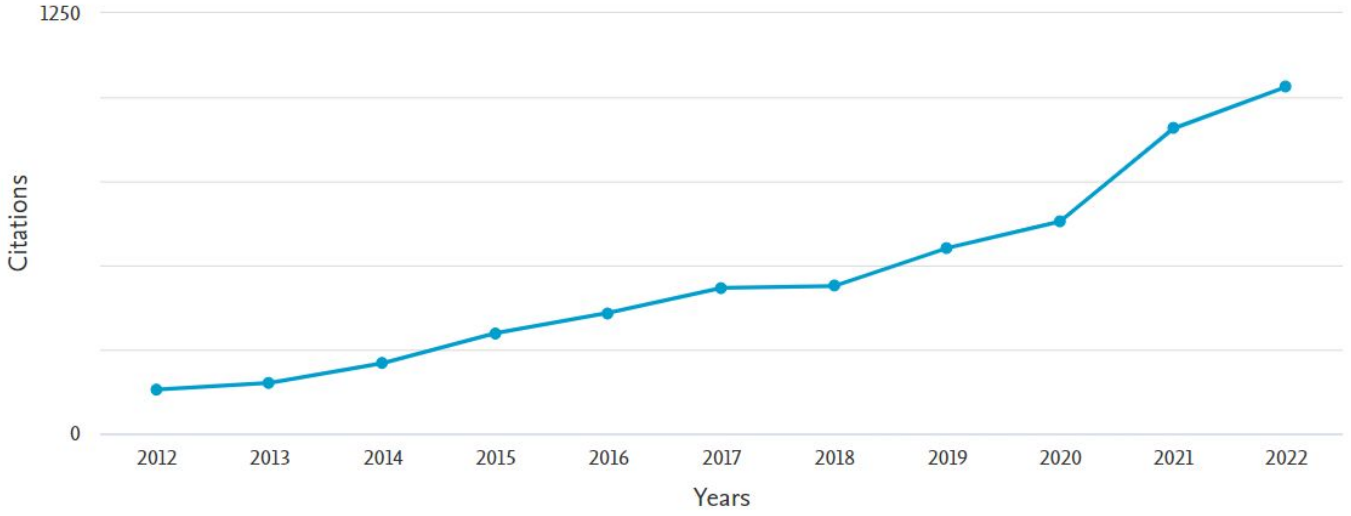
# Citaciones (2012-2022) en Scopus

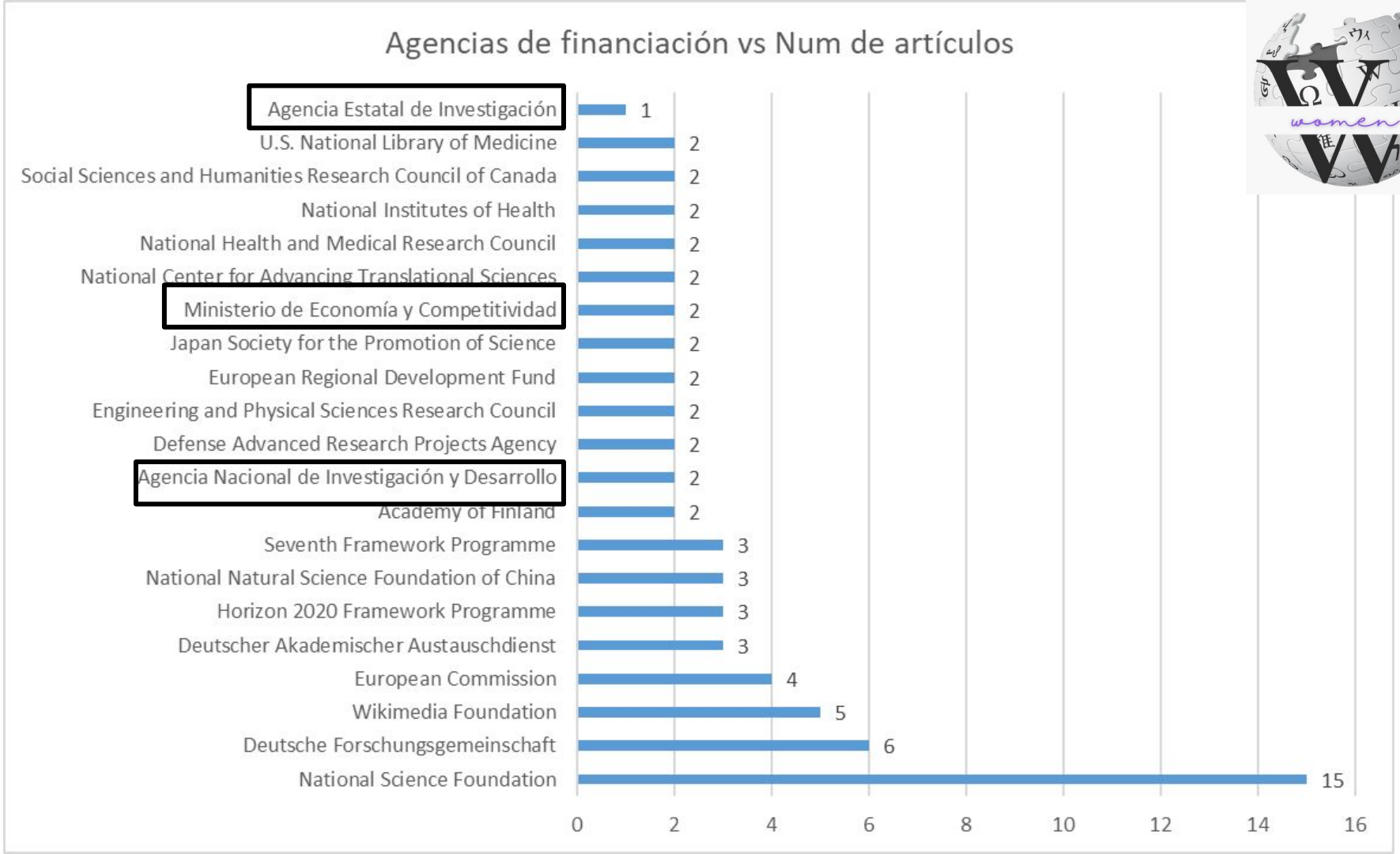
This is an overview of citations for the documents you've selected.

Document *h*-index : 36 View *h*-graph

427 cited documents + Save to list

Date range: 2012 to 2022 Exclude self citations of all authors Exclude citations from books Update



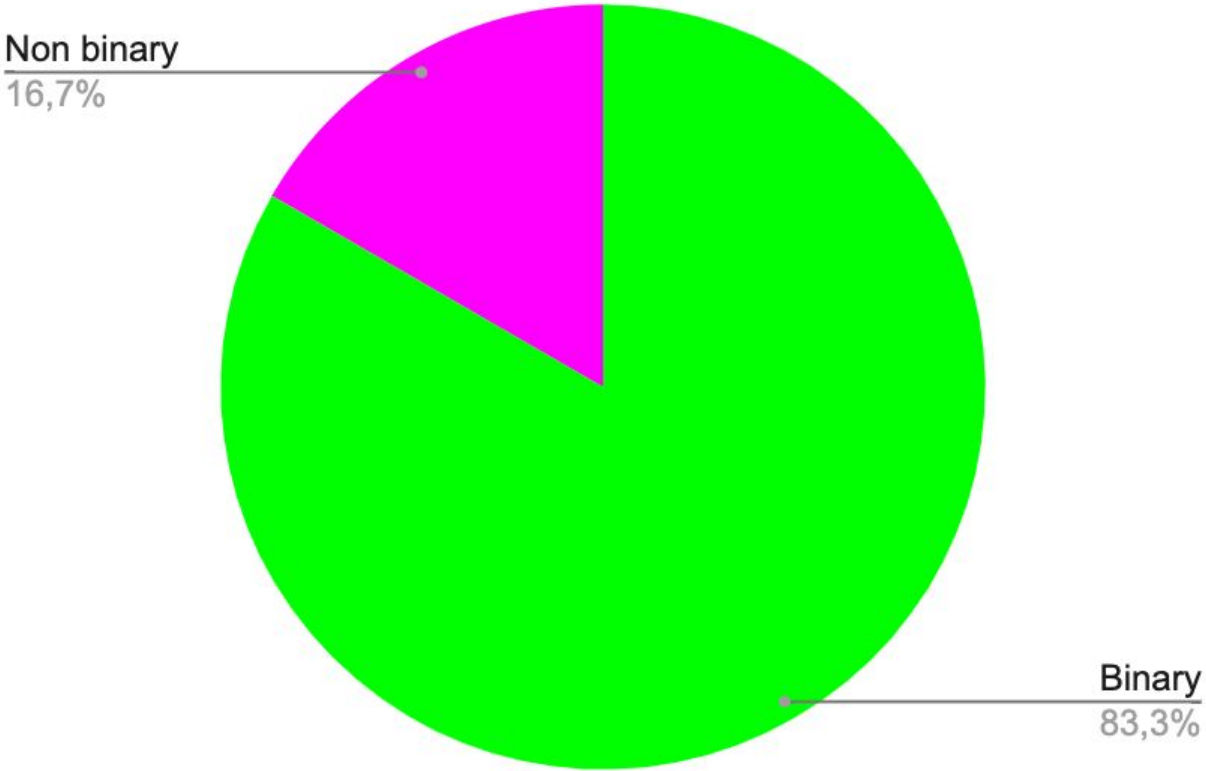

Agencias de financiación vs Num de artículos
Agencia Estatal de Investigación
U.S. National Library of Medicine
Social Sciences and Humanities Research Council of Canada
National Institutes of Health
National Health and Medical Research Council
National Center for Advancing Translational Sciences
Ministerio de Economía y Competitividad
Japan Society for the Promotion of Science
European Regional Development Fund
Engineering and Physical Sciences Research Council
Defense Advanced Research Projects Agency
Agencia Nacional de Investigación y Desarrollo
Academy of Finland
Seventh Framework Programme
National Natural Science Foundation of China
Horizon 2020 Framework Programme
Deutscher Akademischer Austauschdienst
European Commission
Wikimedia Foundation
Deutsche Forschungsgemeinschaft
National Science Foundation
1
2
2
2
2
2
2
2
2
2
2
2
2
3
3
3
3
4
5
6
15
0
2
4
6
8
10
12
14
16

# Principales metodologías

- Informática: minería de datos (API Wikipedia)
  - Wikidata (dataset) aumenta diversidad y reduce brecha género y otros sesgos
- Ciencias sociales: encuestas
  - Estudiantes, EUA
- Etnografía y entrevistas
- Análisis crítico, teorización


- FORD, Heather; WAJCMAN, Judy. ‘Anyone can edit’, not everyone does: Wikipedia’s infrastructure and the gender gap. *Social studies of science*, 2017, 47.4: 511-527.
- SHAW, Aaron; HARGITTAI, Eszter. The pipeline of online participation inequalities: The case of Wikipedia editing. *Journal of communication*, 2018, 68.1: 143-168.
- ZHANG, Charles Chuankai; TERVEEN, Loren. Quantifying the gap: a case study of Wikidata gender disparities. In: *Proceedings of the 17th International Symposium on Open Collaboration*. 2021. p. 1-12.

# Perspectiva de género






# RESULTADOS

¿Cuáles son las razones de la brecha de género en WP según la academia?

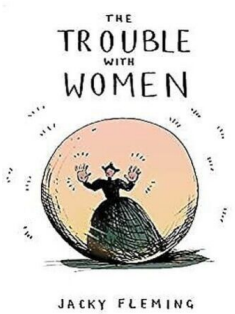
## RAZONES DEL GENDER GAP



# Razones vistas desde la academia




- 267 citas/ 75 artículos
- Simbiosis, sesgos que se refuerzan:  sesgo participación & sesgo contenidos
- Zoom desde la primera capa a las raíces:
  - *Women trouble*
    - la culpa la tiene las características de las mujeres y WP está excusada
    - no es solo adquisición de competencias, estas habilidades están vinculadas a la cultura tecno científica masculina. Para tener éxito el modelo es el masculino.
  - *Mirror effect*
    - la culpa la tiene la sociedad y WP está excusada



# Razones vistas desde la academia




- *Systemic problem*
  - la culpa es de la infraestructura que fundamenta la base de la epistemología
    - herencia de los procesos que provienen de
      - la enciclopedia occidental moderna (con raíces coloniales): all knowledge but not everyone's knowledge
      - el movimiento open source
    - capas ocultas en cuanto a códigos, políticas y lógicas alrededor que afectan los géneros y colectivos marginalizados
      - NPOV (neutral)
      - NOR (no original)
      - Verifiability/Verificabilidad (fuentes de información)
        - Podemos cuestionar estos pilares? Podríamos usar sistema automático para verificar relevancia de una persona y no el tipo o núm. fuentes
    - las infraestructuras producen relaciones de poder. Y la ja jerarquía en WP se crea a partir de habilidades y de conocimiento sobre normas, códigos, políticas.

# Razones vistas desde la academia

*Vivimos en sociedades muy jerarquizadas en las que el acceso al poder por parte de unos sectores incluye determinar cuál es el discurso legítimo y cuál no lo es; quién tiene derecho a exponer su punto de vista y quién se tiene que callar.*

*El sector al que más se le ha negado la palabra sistemáticamente y a lo largo de la historia ha sido precisamente el de las mujeres, que no es una minoría, sino más de la mitad de la población.*

Juliano, Dolores (2017). Tomar la Palabra. Edicions Bellaterra

https://ca.wikiquote.org/wiki/María_Dolores_Juliano_Corregido




RESULTADOS

¿Cómo propone la academia que se podría reducir la brecha de género?

SOLUCIONES



# Soluciones vistas desde la academia

- 138 citas / 60 artículos (61,22% de los artículos
- Mirada a la primera capa:
  - Mejoras cuantitativas
  - Mejoras cualitativas
  - Soluciones prácticas
- Mira más profunda
  - Soluciones sistémicas

# Soluciones de las investigadoras-activistas

- **¿Más mujeres editoras y más biografías de mujeres?**

**“I think it would be great to move away from the percentage thing. It comes with an implication that when we reach a higher percentage of women, our work is done.”**

Wade is incredulous that simply because historically women have not been prominent in chemistry that there should be fewer of them on Wikipedia, calling the notion “absurd.” “Women may not have been famous chemists, but they were authors, activists, and research scientists. They didn’t get titles but did a huge amount for society—and keep doing,” she says.

As a historian, the Wellcome Trust’s White also struggles with the assertion that there is a “correct” number of women to be highlighted. “The ratio of women should be higher,” she says. “ ‘How high?’ is a question often used to distract us from actually getting on with making sure they are better represented,” she says. **“It’s also about more than just the number of women; it’s how they are represented and networked and ensuring that their contributions are fairly valued.”**



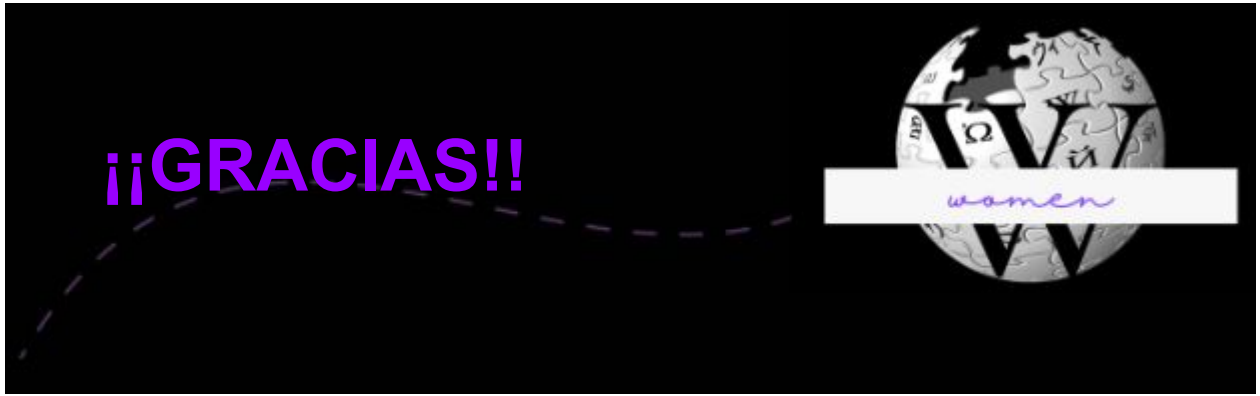
Jass Wade, 2017
Fuente: https://cen.acs.org/careers/women-in-science/Wikipedias-women-scientists/96/i37

# Soluciones vistas desde la academia

- **Mirada más profunda: “el problema sistémico”**
  - Aumentar las contribuciones de las mujeres no será suficiente
  - Hace falta entender la cultura de la producción de conocimiento en WP
    - Tener presente las raíces que esconden prácticas excluyentes
      - Enciclopédicas
      - Cultura abierta (*open source*)
    - Reconocer y superar barreras de la plataforma para la diversidad
      - Lógicas imbricadas en las relaciones de la Infraestructura del proyecto técnico-científico

**La solución: no solo es reducir la brecha:**

**sino hacer WP más robusta, confiable y transparente como espacio de producción de conocimiento**




# **MUJERES** y **WIKIPEDIA**: análisis de la brecha de género en la coproducción de conocimiento en las Wikipedias españolas

**PID2020-116936RA-I00** https://www.ub.edu/wikiwomen/

CC BY SA

Universitat Oberta de Catalunya

Universitat de Barcelona
Facultat d'Informació i Mitjans Audiovisuals

Ministerio de Ciencia e Innovación

Agencia Estatal de Investigación