



¿Cómo se organiza el contenido a través de las categorías en Wikipedia?

Yessica Macià, Miquel Centelles y Núria Ferran-Ferrer

Wikimedia + Gendergap Research Meetup
Madrid - 01/04/2023



Universitat
Oberta
de Catalunya



UNIVERSITAT DE
BARCELONA

Facultat d'Informació
i Mitjans Audiovisuais



MINISTERIO
DE CIENCIA
E INNOVACIÓN



Índice de contenidos



1. Antecedentes: las deliberaciones de C&G en Viquipèdia
2. Objeto y objetivos de nuestro trabajo
3. Folksonomías, la forma de categorización de Wikipedia
4. Metodología de nuestro estudio
5. ¿Cómo afecta esto a la visibilidad de géneros?
6. Nuestra propuesta
7. Conclusiones

Antecedentes: las deliberaciones sobre categorías



- Dones amb categoria



- Objetivos de esta categorización:

- Necesidad de visibilización las identidades de género marginalizadas en la Viquipèdia
- Importancia de la Wikipedia como consulta de información, búsqueda de referentes
- Brecha de género evidente: <20% de biografías de mujeres, <20% de editores son mujeres
- Precedentes de debates y discusiones en torno al tema

- Metodología cualitativa y cuantitativa:

- Análisis del discurso de los argumentos empleados a partir de Judie Schneider y su equipo

Antecedentes: las deliberaciones sobre categorías



Análisis de las deliberaciones en la Viquipèdia:

- 'Categoria Dones': discussió de 2014
- 'Categorització segregada per gèneres, altre cop' (2015-2017)
- 'Eliminació de la Categoria:Persones no binàries' (2022)

Y, asimismo, la votación final de la categorización por género (2018)

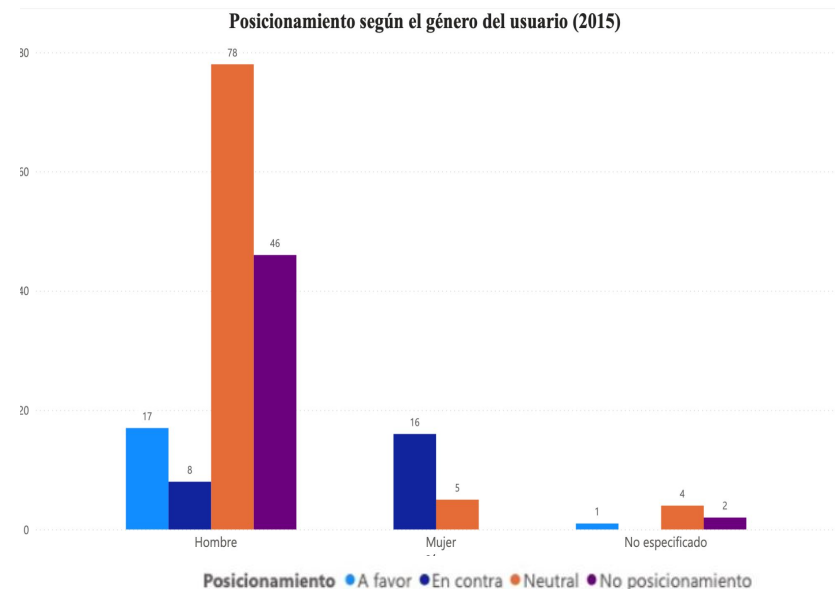
¿La invisibilización de las mujeres en Viquipèdia responde a un origen tecnológico o ideológico?

Antecedentes: las deliberaciones sobre categorías



Datos obtenidos en el estudio

- Media de 100 participaciones en cada discusión
- Mayor tasa de participación masculina (77%-89%)
- Negativa mayoritaria por parte de las usuarias
- Mayor número de argumentos de valores por la parte 'en contra' de la categorización por género ("peor calidad")



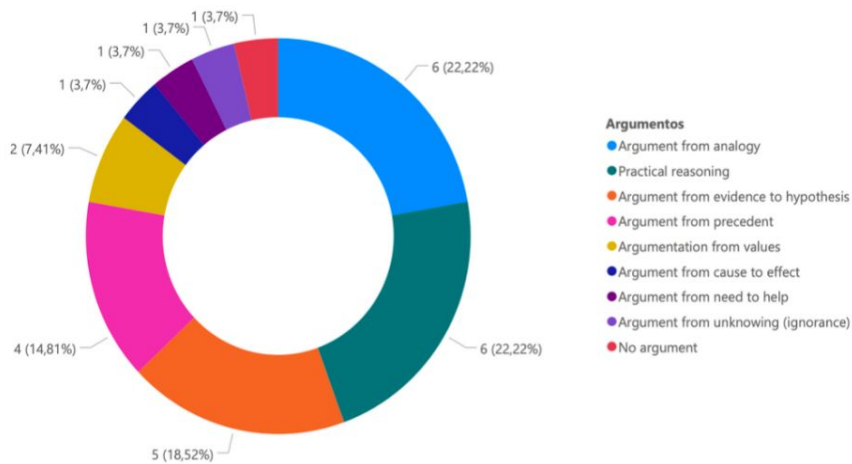
Ejemplo de participación en la deliberación de 2015.
Gráfico de elaboración propia

Antecedentes: las deliberaciones sobre categorías

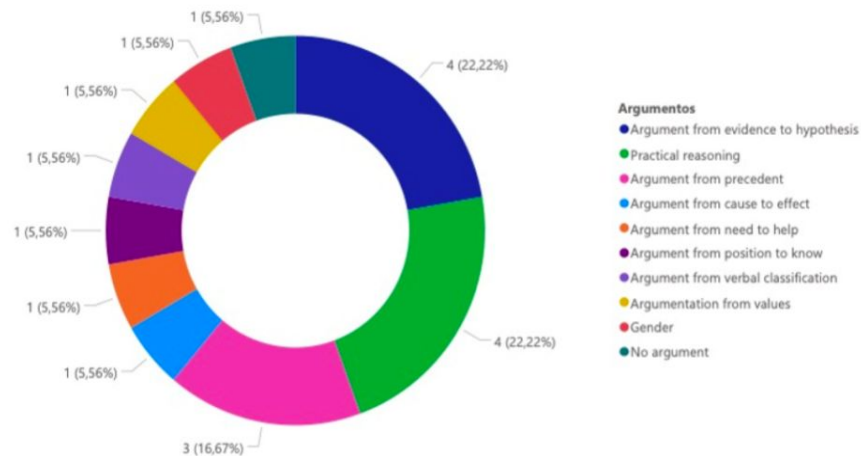


Índice de argumentos empleados según su tipología ('a favor')

Argumentos a favor (2014)



Argumentos a favor (2015)

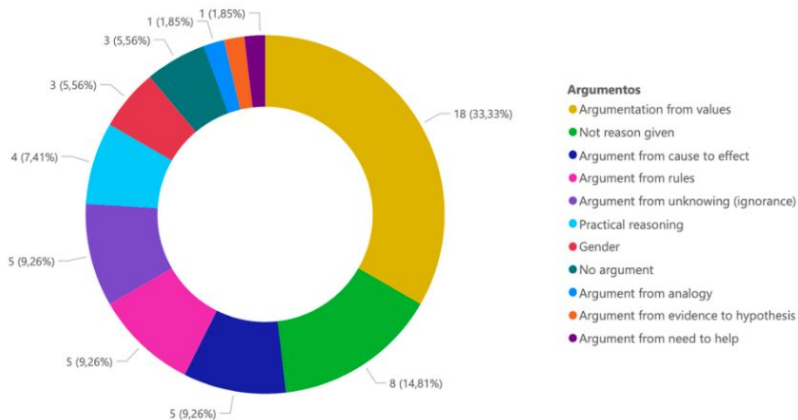


Antecedentes: las deliberaciones sobre categorías

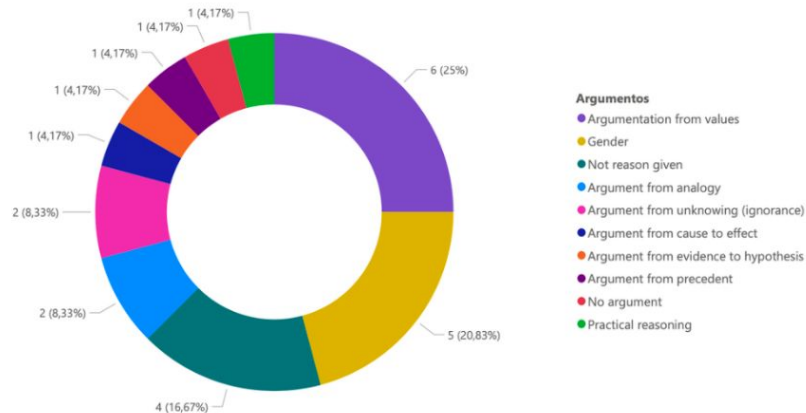


Índice de argumentos empleados según su tipología ('en contra')

Argumentos en contra (2014)



Argumentos en contra (2015)



Antecedentes: las deliberaciones en torno a C&G



Conclusiones que obtuvimos en este estudio:

- Gran peso de los valores propios de la Wikipedia (Viquipèdia)
- Los valores que definen a cada comunidad son decididos por los administradores, una posición que requiere mucho tiempo de mantener por los requisitos que conlleva.
- Los procesos de toma de decisiones se realizan en subconjuntos de la comunidad que ejercen el poder. Al consenso democrático sólo pueden participar unos pocos.
- La causa predominante de la escasez de biografías femeninas es la falta de referencias bibliográficas, debido a la poca presencia de la mujer en medios y en la esfera pública.

Objeto y objetivos de nuestro trabajo



¿Cómo podemos conseguir paliar la brecha de género en la Wikipedia?

- Si la categorización por género no es aceptada, ¿significa eso que con la categorización actual existe al menos una accesibilidad en los contenidos de mujeres?
 - Evaluación de la implementación de sistemas de organización de conocimiento de Wikipedia y Wikidata
 - Estudiar la calidad de estos instrumentos desde las necesidades de las personas usuarias y desde la perspectiva de género.

¿Cómo se ordena el contenido en Wikipedia y Wikidata?

Folksonomías, categorización colectiva de Wikipedia

- El poder de creación de las categorías lo ejerce cualquier persona que edite la Wikipedia
- Organizar el conocimiento de manera colaborativa y descentralizada.
- Lenguaje pre-coordinado
- Las categorías son resultado de su cosmovisión
- Los usuarios pueden agregar etiquetas o tags a los artículos para categorizarlos y facilitar su acceso
- Estas etiquetas se convierten en una especie de índice temático que permite navegar por los contenidos de la enciclopedia de manera intuitiva.



¿Cómo se ordena el contenido en Wikipedia y Wikidata?

Cómo funcionan las *folksonomías* de Wikipedia

- Las etiquetas o tags utilizadas en las folksonomías de Wikipedia no están predefinidas, cualquier usuario puede crear una nueva etiqueta al momento de editar un artículo
- La categorización se lleva a debate, y seguidamente a votación, en caso de desacuerdo (**la última palabra en manos de los administradores**)
- Posibilidad de duplicidad o ambigüedad de la taxonomía

- ▼ [Infermeria](#) (3 cat., 18 p.)
 - ▼ [Infermers](#) (2 cat., 7 p.)
 - ▶ [Infermers americans](#) (1 cat., 9 p.)
 - ▼ [Infermers europeus](#) (3 cat., 19 p.)
 - ▼ [Infermers de l'àmbit catalanoparlant](#) (1 cat., 6 p.)
 - ▼ [Infermers catalans](#) (1 cat., 15 p.)
 - ▼ [Infermers catalans del sud contemporanis](#) (1 cat., 6 p.)
 - ▶ [Infermers del Barcelonès](#) (1 cat., 2 p.)
 - ▶ [Infermers britànics](#) (17 p.)
 - ▼ [Infermers espanyols](#) (1 cat., 22 p.)
 - ▶ [Infermers catalans del sud contemporanis](#) (1 cat., 6 p.)
 - ▶ [Llevadores](#) (1 cat., 25 p.)
 - ▶ [Primers auxilis](#) (1 cat., 17 p.)

Ejemplo de la redundancia en la categoría "Enfermería" en Viquipèdia

¿Cómo se ordena el contenido en Wikipedia y Wikidata?

Organización ontológica del conocimiento en Wikidata

- Base de datos estructurada ontológicamente: relaciones semánticas que existen entre las entidades.
- A diferencia de Wikipedia, es un lenguaje controlado: utiliza esquemas estandarizados de lenguajes de Web semántica
- Permite que sus datos puedan ser utilizados por otros proyectos de Wikimedia, como Wikipedia, para enriquecer sus contenidos con información precisa y actualizada
- Language post-coordinado: empodera a la persona que busca la información y la búsqueda se rige por sus necesidades de información, su cosmovisión.

¿Cómo se ordena el contenido en Wikipedia y Wikidata?

Cómo se relacionan las *folksonomías* de Wikipedia y Wikidata

- Las etiquetas o tags utilizadas en las folksonomías de Wikipedia pueden estar vinculadas a elementos de Wikidata, lo que permite enriquecer la información de los artículos de Wikipedia con datos estructurados y precisos.
- ¿Podemos emplear esto para potenciar la creación de categorías automatizadas basadas en los 7 géneros existentes en Wikidata?

Metodología



Análisis desde tres perspectivas:

- Análisis taxonómico de las categorías de Wikipedia
 - Albuquerque (2017)
 - Kaplan et al. (2022)
 - Kless & Milton (2010)
- Análisis de la ontología de Wikidata
 - Ahmad et al. (2022)
 - Amith et al. (2018)
 - Gangemi et al. (2006)
 - Gervasi et al. (2022)
- Análisis heurístico de la experiencia de usuario en la accesibilidad de los contenidos en ambos escenarios
 - Morales-Vargas (2023)

Metodología



Análisis desde tres perspectivas:

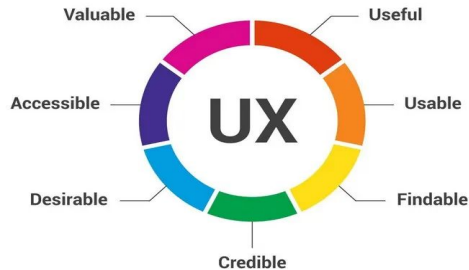
[Taula de criteris d'avaluació.xlsx](#)

Metodología



El análisis heurístico sentenciará la calidad de la accesibilidad, la usabilidad y la arquitectura de la información en Wikidata

Diferentes tests y diferente muestra poblacional:



- Usuarios “novel”: poca experiencia en el manejo de Wikipedia
- Usuarios avanzados: controlan las ediciones y el funcionamiento de los usuarios
 - Activistas
 - Académica/os
 - Usuarios con conciencia de género y sin conciencia de género

¿Cómo afecta esto a la visibilidad de géneros?



Nuestra principal tesis:

- Sesgos en la administración de categorías
- Dificultad de acceder a los contenidos de mujeres y más de otros géneros invisibilizados
- Impedimentos y limitaciones en la participación de usuarios en las deliberaciones finales de la categorización

→ **Demostrar lo caótico de la organización actual y establecer un prototipo que mejore la usabilidad y la accesibilidad**

Nuestra propuesta



→ Propuesta técnica:

- ◆ Folksonomía más estandarizada, sin quitar el derecho de creación comunitaria de los usuarios
- ◆ Una organización que atienda a las necesidades de información de los usuarios
- ◆ Automatizar la creación de categorías que requieran más tiempo de crear a partir de bots o aplicaciones como OntoRefine
- ◆ Posibilitar la categorización y visibilización por género
- ◆ Buscador que cumpla con los criterios de UX para facilitar el acceso de los contenidos

Nuestra propuesta



→ Propuesta cambio cultural:

- ◆ Repensar los procesos de toma de decisiones
- ◆ Fomentar la cultura de la diferencia desde las comunidades
- ◆ Más diversidad en la participación = Menos brecha = Más calidad de los contenidos

Conclusiones

- Establecer los sesgos que existen detrás de las deliberaciones para determinar si existe discriminación por género o de otro tipo
 - ◆ *Hipótesis: la brecha de género se mantiene por la posible existencia de valores sesgados y sexistas*
- Aumentar la visibilidad de los contenidos de mujeres y otros géneros proponiendo una taxonomía más accesible
 - ◆ *Hipótesis: existe una caótica folksonomía, en la que las ambigüedades, los genéricos y las redundancias dificultan entender la localización de los contenidos*
- Potenciar la accesibilidad y la usabilidad de la Wikipedia con un prototipo de buscador desde una perspectiva de UX
 - ◆ *Hipótesis: la dificultad de uso del buscador limita el acceso a contenidos específicos de mujeres y otros géneros*

Futher research

- Realizar un análisis comparativo entre las tomas de decisiones de dos Wikipedias, IT y CAT, que comparten experiencias similares con la categorización por género
- Analizar las consultas de borrado y los argumentos empleados en sus deliberaciones/votaciones respecto a las biografías de mujeres y colectivo LGTBI+ (y su comparativa con lo que ocurre con las de biografías de hombres)
- Analizar la calidad de las referencias bibliográficas que sustentan las biografías de mujeres, como uno de los principales motivos del borrado de las mismas



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/>