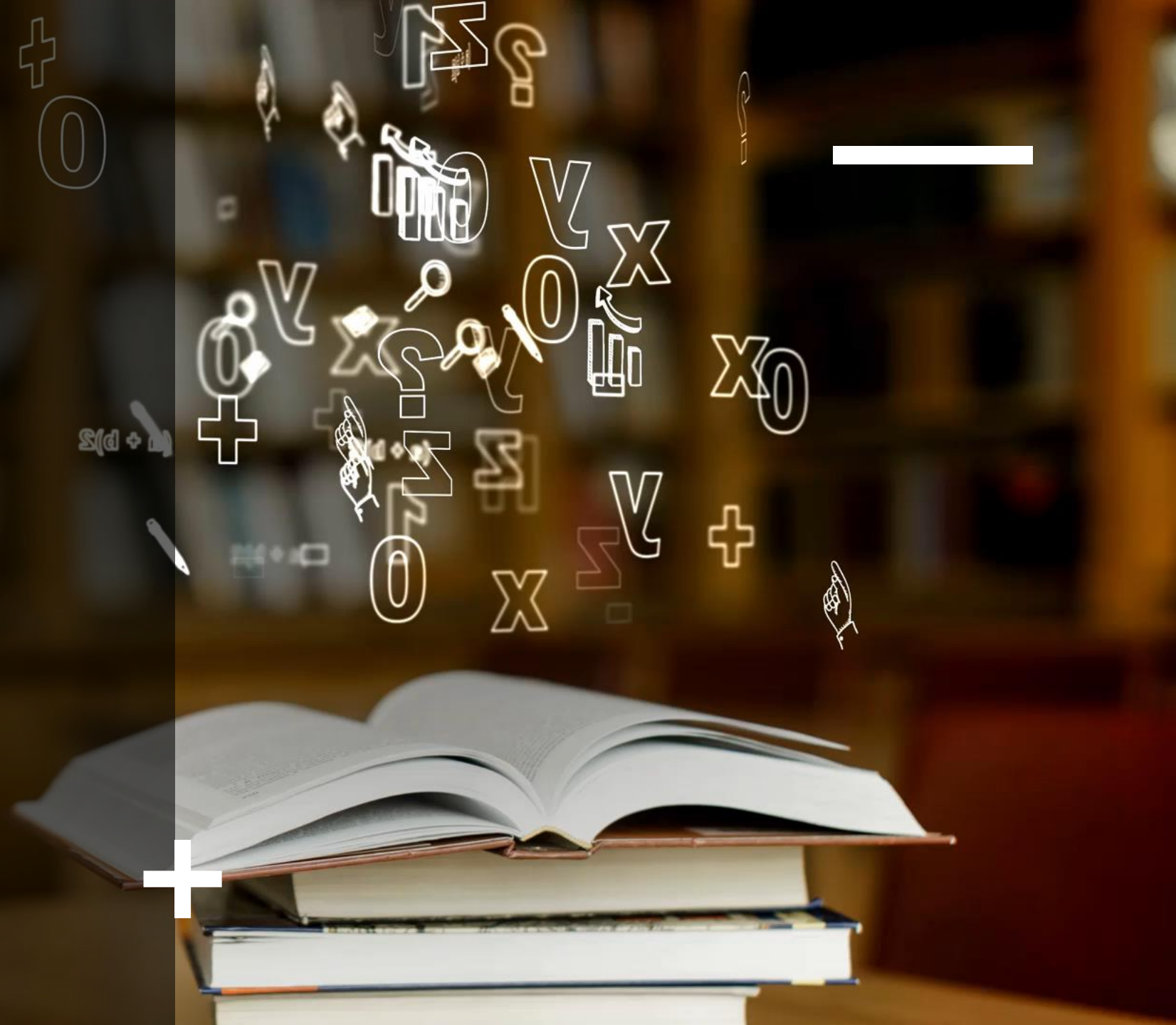


Wikipedia y la aplicación práctica de la equidad de conocimiento

Dra. María Sefidari Huici

Wikinvisibles 2023



Introducción

Wikipedia en el año 2030

- Equidad de conocimiento
- Problemática estructural existente

Preguntas de investigación

- ¿Es posible medir cuantitativamente la calidad del conocimiento de Wikipedia, en base a la equidad de conocimiento y la presencia o ausencia de sesgos, en el contenido destacado en la Portada?
 - Género
 - Raza
 - País de procedencia
- ¿Es posible establecer una evaluación de base para determinar si se está cumpliendo el objetivo estratégico de Wikimedia 2030 de equidad de conocimiento en la Portada?
- ¿La Portada de Wikipedia genera y perpetúa sesgos sistémicos?

La equidad de conocimiento como marco teórico

- Gestión del conocimiento en el mundo empresarial
- Tres etapas (Koenig, 2018): Tecnología de la información, recursos humanos, y contenidos y taxonomías
- Taxonomía de Baskerville y Dulipovici (2006), basada en procesos
- Un problema persistente: la medición de la calidad del conocimiento. Surgen tres marcos: gestión de calidad, benchmarking y la equidad de conocimiento
- El valor del conocimiento (Glazer, 1998)



Inequidades digitales y sesgos

- Ciberfeminismo
- Akane Kanai (2020): dinámicas digitales generan culturas de conocimiento para nuevas formas de activismos
- Sesgos implícitos: raza y género
- Relación con Wikipedia



Wikipedia

- Bomis, Nupedia
- Versiones internacionales
- La Portada
- Sección 230



Objetivos y dimensiones de análisis

- Aplicar de forma práctica la teoría de la equidad de conocimiento, tomando cada artículo destacado que aparece en Portada como un objeto de conocimiento, y examinar el contexto circundante para poder determinar qué conocimiento emerge y qué valor podemos darle al mismo.
- Determinar si este conocimiento resultante puede valorarse como sesgado en base a las siguientes dimensiones de análisis: género, raza, país de procedencia, fecha simbólica
- Dos tipos de análisis: 2019 y 8M



Hipótesis

- Para el análisis de 2019: (a) se prioriza el contenido sobre los hombres (b) se invisibiliza el contenido sobre las mujeres aunque exista, (c) se prioriza contenido de personas blancas, (d) se prioriza el contenido de Estados Unidos y Europa y (e) no se cumple el objetivo de knowledge equity de Wikimedia 2030.
- Para el análisis del 8 de marzo: (a) no se habrá utilizado la fecha del 8 de marzo consistentemente para visibilizar la lucha de las mujeres de todo el mundo por la igualdad de derechos (b) en los años que sí se haya hecho, el contenido seleccionado como destacado priorizará a mujeres estadounidenses o británicas (c) apenas habrá contenido sobre mujeres no blancas, perpetuando su invisibilización histórica.



Metodología

- Análisis cuantitativo
- Wayback Machine y archivos
- 2019: 9 Wikipedias
- 8M:
 - Wikipedia en inglés: 19 años
 - Wikipedia en español: 16 años



Resultados: 2019. Wikipedia en inglés

- 96 biografías en 2019
- 77 de hombres (80,21%) y 19 de mujeres (19,79%)
- 86 de raza blanca (89,58%), 2 de raza negra (2,08%), 1 de las Islas del Pacífico (1,04%), 1 aborigen de Australia (1,04%), y 6 de raza inconclusa/no se pudo determinar (6,25%)
- Profesiones: en los hombres hay una mayoría de biografías de políticos (13) y en las mujeres una mayoría de biografías de Escritoras (3), Músicas (3) y Actrices (3)
- 77% de las biografías de hombres y el 79% de las biografías de mujeres procedían de Europa y Norteamérica. No hubo ninguna biografía de hombres de Asia, y ninguna biografía de mujeres de Oriente Medio.



Resultados: 2019. Wikipedia en español

- 18 biografías en 2019
- 15 (83,3%) fueron de hombres y 3 (16.7%) de mujeres
- 11 fueron de raza blanca (61,1%), 1 de raza negra (5.6%), 1 mestiza (5,6%) y 5 inconclusas (27,8%)
- Las profesiones de los hombres tendieron a corresponderse con exploración, música y artes y ciencias, mientras que las mujeres fueron una impostora y una militar.
- 73,3% de las biografías de hombres y el 66,6% de las biografías de mujeres proceden de Europa. Un 20% de las biografías de hombres proceden de Norteamérica.
- No hubo ninguna biografía de hombres de Asia, Oceanía, Centroamérica o Sudamérica, y ninguna biografía de mujeres de Asia, África, Oceanía, Norteamérica o Centroamérica.



Resultados: 2019. Wikipedia en catalán

- 66 biografías en 2019
- 60 de hombres (90,9%) y 6 de mujeres (9,1%)
- 62 de raza blanca (93,9%), 2 de raza negra (3%), 1 asiática (1,5%) y 1 inconclusa (1,5%). Todas las mujeres fueron de raza blanca.
- Profesiones: hombres tendieron a corresponderse con artistas, músicos, arquitectos y escritores, mientras que las mujeres predominantemente fueron nobles y artistas
- 88,8% de las biografías de hombres y el 83,8% de las biografías de mujeres proceden de Europa. No hubo ninguna biografía de hombres de Centroamérica o Sudamérica, y no hubo ninguna biografía de mujeres de Asia, África, Oceanía, Centroamérica o Sudamérica.



Resultados: 2019. Wikipedia en euskera

- 4 biografías en 2019
- 3 de hombres (75%) y 1 de mujeres (25%)
- 100% personas de raza blanca
- Profesiones: Todos los hombres fueron escritores, y la mujer una política
- Todos los hombres fueron de Europa, y la única mujer de Sudamérica



Resultados: 2019. Wikipedia en gallego

- 2 biografías en 2019
- 1 hombre (50%) y una mujer (50%)
- Físico teórico y geógrafa exploradora
- 100% personas de raza blanca
- 100% procedentes de Norteamérica



Resultados: 2019. Wikipedia en asturiano

- 4 biografías en 2019
- 3 de mujeres (75%) y 1 de hombres (25%)
- 100% personas de raza blanca
- 100% personas de Europa



Resultados: 2019. Wikipedia en occitano

- 10 biografías en 2019
- 100% hombres
- 8 de raza blanca (80%) y 2 inconclusa (20%)
- Profesiones: poeta (en dos ocasiones) y con una aparición cosmonauta, político, general, emperador, escritor, artista, lingüista y profeta
- 8 de Europa (80%), 1 de África (10%) y 1 de Oriente Próximo /Asia Occidental (10%)



Resultados: 2019. Wikipedia en aragonés y Wikipedia en extremeño

– Un artículo en 2019, cero biografías



Resultados: 8M. Wikipedia en inglés

- Seis mujeres en 19 años (2004 a 2022, incl.)
- 4 estadounidenses y 2 británicas
- 5 de raza blanca
- 3 eventos relacionados con la lucha de las mujeres



Resultados: 8M. Wikipedia en español

- Cero mujeres en 16 años (2007 a 2022, incl.)
- Dos hombres
- Dos consolas de videojuegos y un videojuego
- Cinco artículos relacionados con Japón, cuatro con España y tres con Estados Unidos



Patrones observados

- Mayoría abrumadora de hombres en 2019, excepto gallego y asturiano. Occitano no muestra ninguna mujer
- Mayoría abrumadora de personas de raza blanca. Euskera, occitano y asturiano sólo muestran personas de raza blanca
- Wikipedia en inglés roza el 90% de biografías de personas blancas. Ninguna Wikipedia supera el 6% de biografías de personas negras, con la Wikipedia en inglés mostrando un 2.08% de biografías de personas negras en 2019, la Wikipedia en español un 5.6%, y un 3% la Wikipedia en catalán.
- Mayoría abrumadora de personas de Europa (especialmente) y Norteamérica. Euskera sólo seleccionó personas europeas en 2019



Patrones observados

- Uso inconsistente de la Wikipedia en inglés de la fecha del 8M para reivindicar mujeres o la lucha por sus derechos. 6 mujeres, todas británicas o estadounidenses, todas blancas excepto una.
- La Wikipedia en español no ha mostrado nunca a una mujer ese día. Sí aparecen hombres. La mayoría de los artículos seleccionados para esta fecha están relacionados con Japón, España y Estados Unidos. No aparece ningún artículo relacionado con los derechos de las mujeres, pero sí varios relacionados con el mundo de los videojuegos



Conclusiones

- Objetivo: Aplicación práctica de la teoría de equidad de conocimiento al contenido de Wikipedia
- Propuesta:
 - Se puede valorar la calidad del contenido en función de los sesgos sistémicos
 - Estos sesgos son medibles cuantitativamente
 - La presencia de sesgos afecta negativamente la calidad del conocimiento en su conjunto



Conclusiones

– Hallazgos más significativos:

- Independientemente del idioma, el tamaño o el contenido existente, todas las Wikipedias examinadas muestran sesgos con respecto a las biografías que destacan en Portada
- Las Wikipedias con más problemas de sesgos en este sentido serían la Wikipedia en catalán y la Wikipedia en occitano, seguidas de cerca por la Wikipedia en inglés y la Wikipedia en español.
- Las Wikipedias en gallego y en asturiano son las únicas que gestionan la dimensión de género con éxito (teniendo en cuenta que muestran muy pocas biografías cada una), pero en las dimensiones de raza y país de procedencia obtienen los peores resultados



Conclusiones

- Hallazgos más significativos:

- Hipótesis iniciales:

- Siete Wikipedias destacaron biografías en Portada en 2019, y cinco superaban el 75% de biografías de hombres, con cuatro superando el 80% (y una mostrando únicamente hombres).
 - Siete Wikipedias mostraron mayoritariamente a personas de raza blanca, con seis mostrando un 80% o más de biografías de personas de raza blanca, y tres mostrando únicamente biografías de personas de raza blanca.
 - Siete Wikipedias mostraron mayoritariamente biografías de personas de Europa y Norteamérica
 - 8 de marzo: la Wikipedia en inglés realiza un uso inconsistente de dicha fecha para conmemorar a 81 mujeres o la lucha por sus derechos, dando preferencia a mujeres blancas y de Estados Unidos o Reino Unido. La Wikipedia en español nunca ha destacado a una mujer en esta fecha (pero sí a hombres)



Conclusiones

- Preguntas de investigación:
 - (a) hemos demostrado que es posible medir cuantitativamente la presencia de sesgos en el contenido destacado de la Portada de Wikipedia, a lo largo de las dimensiones de género, raza y país de procedencia
 - (b) nuestra forma de medir esta presencia de sesgos puede constituir una evaluación de base para determinar el progreso en el cumplimiento del objetivo de Wikimedia 2030 de equidad del conocimiento en la Portada, ya que se puede replicar midiendo los años siguientes (e incluso los anteriores si se quiere tener una visión más global), en cualquier Wikipedia que use este sistema de selección de contenido destacado para la Portada, y ver si los sesgos aumentan, se mantienen o disminuyen, e incluso puede inspirar la búsqueda de distintas fechas simbólicas para examinar en mayor profundidad otras dimensiones de análisis
 - (c) la Portada de Wikipedia efectivamente tiene la capacidad de generar y perpetuar sesgos sistémicos.



Conclusiones

- Oportunidades y limitaciones:
 - Ventajas del análisis cuantitativo: permite agrupar datos con dimensiones de análisis
 - Queda por explorar el contenido no biográfico
 - Posibilidad de profundizar en otras dimensiones de análisis en las biografías
 - Efecto traducción
 - Artículos que pierden la condición de destacado



Conclusiones

– Medidas de acción:

- Intencionalidad del contenido a destacar en Portada: eliminar los sesgos derivados de la selección de contenido destacado en Portada es algo que puede hacer la comunidad existente (y no las comunidades infrarrepresentadas), que además puede ir más allá e identificar fechas simbólicas para visibilizar contenido actualmente invisibilizado
- Vuelta a los orígenes de la Portada
- Posibilidades educativas y de divulgación de una Portada que destaca conocimiento sin sesgos



Muchas gracias

