

Wikipedia como objeto de investigación

Wikipedia as an object of research

Jesús Tramullas

Tramullas, Jesús (2015). "Wikipedia como objeto de investigación". *Anuario ThinkEPI*, v. 9, pp. 223-226.

<http://dx.doi.org/10.3145/thinkepi.2015.50>

Publicado en *IweTel* el 16 de marzo de 2015



Resumen: Se analiza *Wikipedia* como objeto de investigación científica. Para ello se revisan varios trabajos científicos que analizan la popular enciclopedia colaborativa. Se demuestra que *Wikipedia* es una manifestación de la producción y consumo colaborativo de conocimiento, y un objeto válido de investigación.

Palabras clave: *Wikipedia*; Publicación científica; Investigación; Revisiones.

Abstract: This short paper analyzes *Wikipedia* as an object of scientific research, contrasting various studies dealing with that popular encyclopedia. The conclusion is that *Wikipedia*, as a manifestation of collaborative production and consumption of knowledge, is a valid subject of scientific research.

Keywords: *Wikipedia*; Scientific publication; Research; Reviews.

Wikipedia como disrupción

Cuando Jimmy Wales y sus colegas lanzaron *Wikipedia* en 2001, la iniciativa para elaborar y compilar una enciclopedia universal online, en acceso abierto, redactada y corregida por voluntarios, pareció a muchos un disparate y un proyecto que iba a tener un escaso recorrido. Catorce años después, oscila entre el sexto y el séptimo lugar en el listado de sitios más visitados de internet, ofrece versiones en 237 idiomas, y el número de páginas consultadas y de visitantes se cuenta en millones por hora. A su estela han ido surgiendo proyectos más especializados como *Commons*, *Wikidata*, *Wiktionary*... con más o menos éxito, animados por el mismo planteamiento de creación y acceso libre al conocimiento (Reagle, 2011).

<http://www.wikipedia.org>

<https://commons.wikimedia.org/wiki/Portada>

<https://www.wikidata.org/wiki/Wikidata:Portada>

<https://www.wiktionary.org>

Como era de esperar, la disrupción del modelo enciclopédico tradicional que supuso *Wikipedia* generó controversias, en especial sobre la calidad

del contenido informativo que ofrece a los usuarios (Saorín, 2012). La impresión en este aspecto, que todavía encuentra eco en la prensa generalista, era que *Wikipedia* carecía de fiabilidad y de calidad en sus contenidos. Esta errónea generalización fue refutada en el trabajo de Giles (2005) publicado en *Nature*, una revista poco sospechosa de sesgo, que comparaba artículos similares de *Wikipedia* y de *Encyclopaedia Britannica*. Sin embargo, cada cierto tiempo los medios se hacen eco de trabajos que demuestran una pretendida falta de fiabilidad de *Wikipedia*.

"Para poder estudiar *Wikipedia* es necesario establecer un marco teórico y metodológico adecuado al contexto científico, social y cultural de este producto de la sociedad-red"

Realmente, si se aborda *Wikipedia* desde una perspectiva crítica cuyos fundamentos culturales y académicos se encuentran en el modelo de

enciclopedia vigente en los siglos XIX y XX, pueden identificarse debilidades. La utilización de este enfoque no es adecuado: no parece lo más recomendable valorar y analizar un producto de información del siglo XXI, un nodo resultado de la sociedad-red, siguiendo criterios de la estructura socio-cultural de los siglos anteriores. Para poder estudiar *Wikipedia* es necesario establecer un marco teórico y metodológico adecuado al contexto científico, social y cultural de *Wikipedia* (Jemenialk, 2014; Fichman; Hara, 2014), lo mismo que para otros productos de la sociedad-red.

Wikipedia en la bibliografía científica

En primer lugar cabe plantear si *Wikipedia* es un fenómeno social, tecnológico y cultural que puede ser sometido al análisis y escrutinio que supone la aplicación de los métodos científicos provenientes de varias disciplinas. La revisión de la publicación científica desde 2001 dice que indudablemente lo es:

- *Scopus* recoge que desde 2006 hasta febrero de 2015 se han publicado 4.932 trabajos sobre la cuestión (atendiendo a contenidos de título, palabras clave y resumen);
- *Web of science* los reduce a 2.275, entre 2002 y 2014;
- *Google scholar*, con una cobertura bastante más amplia, ofrece 4.790 resultados que incluyen el término "*Wikipedia*" en el título del propio trabajo:
<http://scholar.google.es/scholar?q=allintitle%3Awikipedia>
- *Scirus*, antes de su desaparición, ofrecía 7.032 resultados.

Park (2011), tomando datos de *Web of science* y de *Scopus*, llevó a cabo un estudio sobre publicaciones, autores, instituciones, revistas, campos académicos, visibilidad y trabajos más citados.

Bar-Ilan y **Aharony** (2014) han revisado la producción anual presente en *Scopus* dividiendo la misma entre enfoques tecnológico y social, señalando que parece descender lentamente desde 2012.

Además de la cobertura en las bases de datos, han emergido congresos y conferencias especializados sobre *Wikipedia*, que se celebran con regularidad:

- *The international symposium on open collaboration (OpenSym, previamente denominado WikiSym)* es un congreso anual dedicado a la investigación sobre temas de colaboración abierta, apoyado por la *Association for Computing Machinery (ACM)*, y que da una amplia cobertura a trabajos sobre aspectos de *Wikipedia*:
<http://www.opensym.org>

- *Wikipedia* suele gozar de sesiones en las *ACM Conference on computer supported cooperative work*:
<http://cscw.acm.org>
- *Wikimania*, patrocinada por la *Wikimedia Foundation* desde 2005, es la conferencia anual oficial de la misma, donde se presentan proyectos, estudios y desarrollos:
<https://es.wikipedia.org/wiki/Wikimania%20C3%ADa>
<https://wikimediafoundation.org/wiki/Home>
- *GLAM-Wiki Conference* es el evento bienal dedicado a proyectos *GLAM (galleries, libraries, archives, museums)* en *Wikipedia*, que se ha celebrado en 2013 y 2015, con apoyo de *Europeana* y de la *Wikimedia Foundation*.
https://wikimedia.org.uk/wiki/GLAM-WIKI_2013
https://nl.wikimedia.org/wiki/GLAM-WIKI_2015

A lo anterior habría que añadir los números especiales de revista, talleres o sesiones específicas en el marco de congresos más amplios, etc. Todo ello permite afirmar que *Wikipedia* es un objeto de estudio de interés para disciplinas y campos científicos.

- https://en.wikipedia.org/wiki/Academic_studies_about_Wikipedia
- https://en.wikipedia.org/wiki/Wikipedia:Academic_studies_of_Wikipedia

Áreas de investigación sobre Wikipedia

El volumen de trabajos de investigación publicados sobre *Wikipedia* ha hecho posible que, a su vez, puedan encontrarse en la bibliografía especializada trabajos que analizan el contenido y orientación de las publicaciones, con el objetivo de detectar y documentar sobre los frentes de investigación que pueden identificarse.

Okoli (2009), aplicando el modelo de análisis propuesto por **Okoli** y **Schabram** (2009), usando como fuente artículos publicados en revistas *peer-reviewed*, identifica las siguientes áreas:

- Funcionamiento, revisiones generales, motivaciones para contribuir y procesos editoriales.
- Confiabilidad o fiabilidad, incluyendo estudios críticos y análisis epistemológicos.
- Como fuente de datos, atendiendo a las citas que recibe, y como fuente primaria.
- Aplicaciones en educación y economía.

En 2012, el grupo de investigadores coordinado por **Okoli** presentó un desarrollo más extenso de sus primeros trabajos (**Okoli et al.**, 2012). Para ello revisó unos 450 trabajos recogidos de *WikiLit*.
<http://wikilit.referata.com>

Los resultados obtenidos ofrecen una mejor y más detallada aproximación a los temas investigados sobre *Wikipedia*:

- Generalidades: conceptos de enciclopedia, epistemología, ética, como plataforma de investigación.
- Contenido: calidad y tamaño.
- Como corpus textual: recuperación de información, procesamiento de lenguaje natural, ontologías.
- Infraestructura: legal y técnica.
- Participación: actividades de participación, motivación, cultura de colaboración, resultados, herramientas de software.
- Lectura y lectores: herramientas de software, fuente de conocimiento, valoración y popularidad, credibilidad.

Martin (2010), en un trabajo introductorio sobre aplicación de técnicas estadísticas a *Wikipedia*, que no ofrece datos sobre el corpus revisado, señala las siguientes áreas de investigación:

- Calidad, incluye vandalismo y evolución de artículos.
- Confianza.
- Aspectos semánticos.
- Gobernanza y sociedad.

Nielsen (2012; 2015) revisa alrededor de un millar de trabajos, y propone cuatro grandes categorías de investigaciones sobre *Wikipedia*:

- La analizan, incluyendo calidad, errores, cobertura, actualización, fuentes, accesibilidad, análisis de redes, género, vandalismo, edición, autoría, usuarios, factores geográficos, organización, popularidad y motivaciones de usuarios.
- Usan su información, incluyendo utilización de categorías, construcción de tesauros, traducciones, tendencias y predicción, búsqueda y extracción de información para productos derivados.
- Analizan sus extensiones técnicas, incluyendo control de calidad, certificación y valoración, creación automática de contenido y enlaces, bases de datos, *wikis* semánticos, marcado, edición mediante formularios, digitalización de contenidos, geolocalización, extensiones gráficas y de video, edición en tiempo real y programación.
- La usan como recurso de comunicación, incluyendo actitudes

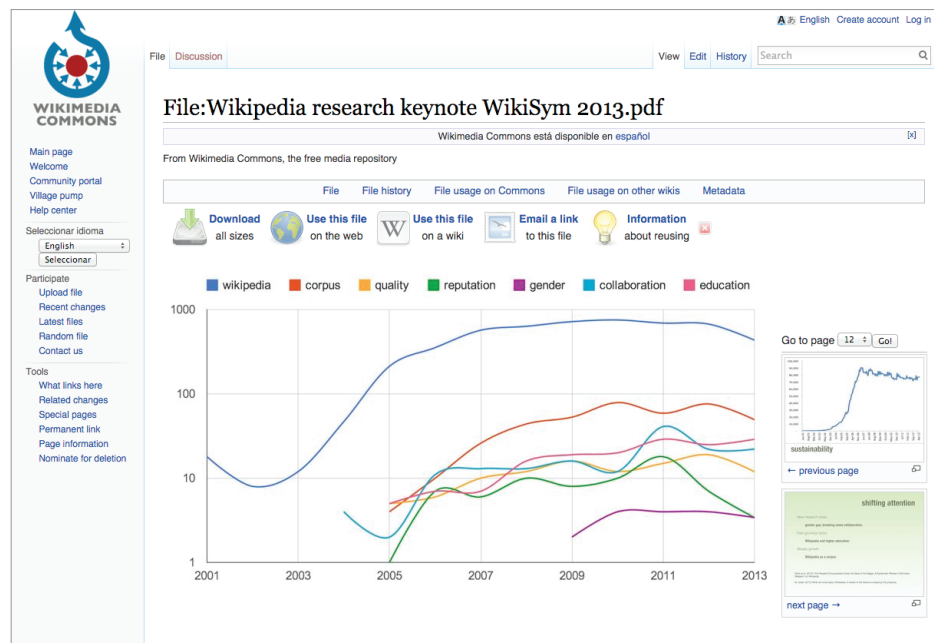
hacia ella, usos, citas, *wikis* especializados, censura, uso como herramienta educativa, comunicación en cursos, y libros de texto y manuales.

Jullien (2012) define *Wikipedia* como un sistema colectivo de producción de conocimiento. Realiza una revisión de los temas tratados en la bibliografía, y para ello considera necesario que previamente se conozca con detalle el funcionamiento de este tipo de comunidades. Su revisión cubre alrededor de 300 trabajos publicados en inglés, francés y español hasta 2011. Los principales temas que identifica corresponden a:

- Estudios sobre la motivación de la participación.
- Procesos o patrones de interacción: actividad y roles de los usuarios, estructura del proyecto, estudio de artículos o conjuntos de ellos, estructura e interacción de los procesos de edición y estructura, organización y gobernanza del proyecto *Wikipedia*.
- Evaluación de calidad, atendiendo a aspectos de calidad de procesos y eficiencia, experiencia de usuario (editor y lector), calidad del producto *Wikipedia* (aspectos de globalidad y calidad de artículos).

Taraborelli (2013) identificó siete grandes frentes de investigación sobre *Wikipedia*, que plasmó en un gráfico (figura 1).

- como entidad, aspectos generales;
- como corpus textual;
- calidad de la información;
- reputación;
- cuestiones de género;



Líneas de investigación sobre *Wikipedia* detectadas por Dario Taraborelli
https://commons.wikimedia.org/w/index.php?title=File:Wikipedia_research_keynote_WikiSym_2013.pdf&page=12

- colaboración;
- educación.

Mesgari et al. (2015), también del grupo de **Okoli** y **Nielsen**, han publicado la más reciente revisión de bibliografía. Su objetivo ha sido analizar la investigación sobre el contenido de *Wikipedia*, restringiéndose a la *Wikipedia* en inglés. Han revisado 110 trabajos académicos, también seleccionados de *WikiLit*, y detectado dos corrientes principales:

- Calidad del contenido.
- Tamaño.

Los autores aseveran:

“Confiabilidad o precisión ha sido el núcleo de las discusiones sobre la calidad del contenido de *Wikipedia*. Según nuestra revisión, los estudios han demostrado resultados diversos, aunque las evaluaciones positivas de *Wikipedia* son más numerosas que las negativas”.

Conclusión

La revisión de los trabajos anteriores demuestra que *Wikipedia* es un fenómeno tecnológico, social y cultural que no puede desdeñarse, ni tampoco abordarse desde enfoques simplistas o generalistas. La gran cantidad de trabajos publicados¹ sobre los *wikis* y sus aplicaciones en varios campos de actividad es un buen indicador de la importancia y validez del enfoque colaborativo en la creación y diseminación de conocimiento. La cuestión clave radica en el desarrollo de una discusión activa y constructiva sobre la integración de *Wikipedia* en las actividades formativas y en los procesos informacionales de los usuarios.

Nota

1. Véanse los trabajos recopilados en *Wikipapers*, iniciativa de **Emilio J. Rodríguez-Posada**.
http://wikipapers.referata.com/wiki/Main_Page

Bibliografía

Bar-Ilan, Judit; Aharony, Noa (2014). “Twelve years of Wikipedia research”. En: *Procs of the 2014 ACM conf on web science. WebSci’14*, pp. 243-244. New York: ACM Press.
<http://dx.doi.org/10.1145/2615569.2615643>

Fichman, Pnina; Hara, Noriko (eds.) (2014). *Global Wikipedia: International and cross-cultural issues in online collaboration*. Rowman & Littlefield. ISBN: 978 0 8108 9101 2

Giles, Jim (2005). “Internet encyclopaedias go head to head”. *Nature*, v. 438, n. 7070, pp. 900-901.
<http://dx.doi.org/10.1038/438900a>

Jemenialk, Dariusz (2014). *Common knowledge? An ethnography of Wikipedia*. Stanford: Stanford University Press. ISBN: 978 08 0478 944 8

Jullien, Nicolas (2012). “What we know about Wikipedia: A review of the literature analyzing the project(s)”.

SSRN electronic journal, 86 pp.
<http://dx.doi.org/10.2139/ssrn.2053597>

Martin, Owen S. (2011). “A Wikipedia literature review”. *ArXiv*, abs/1110.5863.
<http://arxiv.org/abs/1110.5863>

Mesgari, Mostafa; Okoli, Chitu; Mehdi, Mohamad; Nielsen, Finn-Årup; Lanamäki, Arto (2015). “The sum of all human knowledge’: A systematic review of scholarly research on the content of Wikipedia”. *Journal of the Association for Information Science and Technology*, v. 66, n. 2, pp. 219-245.
<http://spectrum.library.concordia.ca/978618>
<http://dx.doi.org/10.1002/asi.23172>

Nielsen, Finn-Årup (2012). “Wikipedia research and tools: Review and comments”. *SSRN electronic journal*.
<http://dx.doi.org/10.2139/ssrn.2129874>

Nielsen, Finn-Årup (2015). *Wikipedia research and tools: Review and comments*. Working paper.
http://www2.compute.dtu.dk/pubdb/views/edoc_download.php/6012/pdf/imm6012.pdf

Okoli, Chitu (2009). “A brief review of studies of Wikipedia in peer-reviewed journals”. En: *Procs of the 3rd Intl conf on digital society*. Cancun, México, pp. 155-160.
<http://dx.doi.org/10.1109/ICDS.2009.28>

Okoli, Chitu; Schabram, Kira (2009). “Protocol for a systematic literature review of research on the Wikipedia”. *Sprouts: Working papers on information systems*, 9, pp. 1-10.
http://sprouts.aisnet.org/866/1/Wikipedia_lit_review_-_protocol.pdf
<http://dx.doi.org/10.1145/1643823.1643912>

Okoli, Chitu; Mehdi, Mohamad; Mesgari, Mostafa; Nielsen, Finn-Årup; Lanamäki, Arto (2012). “The people’s encyclopedia under the gaze of the sages: A systematic review of scholarly research on Wikipedia”. *SSRN electronic journal*.
<http://dx.doi.org/10.2139/ssrn.2021326>

Park, Taemin-Kim (2011). “The visibility of Wikipedia in scholarly publications”. *First Monday*, v. 16, n. 8.
<http://www.ojphi.org/ojs/index.php/fm/article/view/3492/3031>

Reagle, Joseph-Michael (2011). *Good faith collaboration. The culture of Wikipedia*. Cambridge, MA: MIT Press.
<http://reagle.org/joseph/2010/gfc>

Saorín, Tomás (2012). *Wikipedia de la A a la W*. Barcelona: Colección EPI-UOC. ISBN: 978 84 9029 012 5

Taraborelli, Dario (2013). “Descending mount Everest: Steps towards applied Wikipedia research”. En: *Procs of the 9th Intl symposium on open collaboration WikiSym/OpenSym 2013*, Keynote.
https://commons.wikimedia.org/wiki/File:Wikipedia_research_keynote_WikiSym_2013.pdf

Jesús Tramullas
Universidad de Zaragoza
Departamento de Ciencias de la Documentación
tramullas@unizar.es