

Autores: Mauricio V. Genta y Giselle Bordoy **Institución:** Asociación Civil Wikimedia Argentina

Dirección: Av. Rivadavia 1545 - 3°H.

E-mail: comunicacion@wikimedia.org.ar, <a href="mailto:mailto

Título: Estructurando el conocimiento libre en los proyectos de Wikimedia.

Durante la ponencia se dará una introducción a los diferentes proyectos de la Fundación Wikimedia, a escala global y regional. Se hablará principalmente de los dos proyectos más involucrados con las instituciones culturales, Wikimedia Commons y Wikidata.

Wikimedia Commons, o Commons, es una mediateca que alimenta a todos los proyectos de Wikimedia, como lo es Wikipedia. Contiene más de 41 millones de archivos libres, y actualmente se encuentra en un proceso de estructurar su metadata.

Wikidata es el proyecto más joven de la fundación, y es una base de datos estructurada, multilingüe, libre y colaborativa. La intención es ordenar la información que hay en todos los proyectos, de forma sistemática y que está pueda ser usada y editada, no solo por humanos, si no también por computadoras.

A lo largo de esta ponencia se tdara un breve vistazo por las potencialidades que tienen estos dos proyectos, cuál puede ser su vínculo con la comunidad de gestores de la información, como pueden ser utilizadas en el trabajo diario y cuál es el impacto que pueden tener tanto para los usuarios de los proyectos Wikimedia como para las instituciones culturales que deciden sumarse a estas iniciativas.

Adicionalmente se hablará de proyectos que son de impacto para el mundo de los datos: cómo fue la absorción de Freebase, y cómo se convirtió la principal fuente de información del Knowledge Graph de Google. También un vistazo a RDF (Resource Description Framework), SPARQL (Protocol and RDF Query Language) lenguajes utilizados por Wikidata para recuperar información y visualizarla, la licencia CC-Zero, y Querys federadas.

Tabla de Contenidos

```
Introducción
   Wikimedia Argentina
   Proyectos
      #1lib1ref
      Wikicite
      Wikipedia Library
   Licencias Creative Commons
Wikipedia
   <u>Introducción</u>
   Referencias/Citas y Control de Autoridades
      Control de Autoridades
      CitoID
Wikimedia Commons
   <u>Introducción</u>
      ¿Por qué las instituciones culturales deberían utilizar Wikimedia Commons?
      ¿Qué se permite subir en Wikimedia Commons?
      ¿Qué no se permite subir en Wikimedia Commons?
Wikisource
   Introducción
Wikidata
   Introdución
      ¿Qué es Wikidata?
      ¿Qué quiere decir esto?
      Marco de Descripción de Recursos:
      Datos enlazados abiertos:
      Web semántica:
   Knowledge Graph y Wolfram Alpha
   SPARQL
Bibliografía
```

Introducción

Wikimedia Argentina

Wikimedia Argentina (WMAR) es el capítulo local de la **Fundación Wikimedia**. Esta es una organización no gubernamental y sin fines de lucro que opera Wikipedia (entre otros proyectos).

La asociación civil Wikimedia Argentina tiene como misión brindar apoyo a la comunidad organizando actividades e impulsando proyectos locales para la difusión del contenido libre y la cultura wiki. Es en este marco que se trabaja en el programa de promoción cultural que tiene como objetivo mejorar el contenido y el conocimiento sobre la historia y cultura de Argentina mediante la cooperación con instituciones y organizaciones culturales del país.

Proyectos

#1lib1ref

Esta iniciativa surgió en 2015 con el objetivo de promover la participación de nuevos editores dentro de Wikipedia, principalmente **bibliotecarios o personas relacionados con las bibliotecas y centros de documentación**. Su objetivo es mejorar la cantidad y calidad de la referencias que aparecen en cada artículo, como así también extender el uso de nuestros proyectos en esta comunidad.

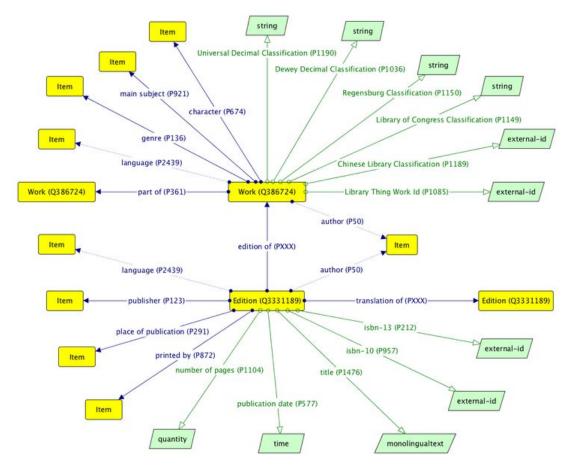
Wikicite

Es indiscutible que la parte más importante del **conocimiento abierto**, son las fuentes y referencias. Wikicite es una iniciativa que junto al proyecto Wikidata, busca mejorar técnicamente el sistema de citas o referencias dentro de nuestros proyectos, entre sus objetivos está el de desarrollar una **base de datos bibliográfica centralizada**. Es apoyado tanto por la Fundación Wikimedia, como Wikimedia Alemania, y las fundaciónes Alfred P. Sloan, Gordon and Betty Moore, y Crossref.

Wikipedia Library

Este proyecto ayuda a editores a acceder a **recursos y fuentes primarias fiables** para mejorar Wikipedia, y a profesionales del conocimiento a compartir sus colecciones con el público. Se ofrece acceso a diferentes bases de datos y colecciones, para que los wikipedistas puedan trabajar con ellas y finalmente utilizarlas como referencias. Entre ellas se pueden mencionar: Cambridge University Press, EBSCO, JSTOR, MIT, Oxford University Press.

APA, Edinburgh University Press, World Bank, Elsevier ScienceDirect, Royal Society Journal, etc.



Modelo de datos para libros propuesto durante Wikicite 2017.

Licencias Creative Commons

Las licencias Creative Commons son unas herramientas que permiten otorgar permisos sobre el derecho de autor (o copyright) al trabajo realizado por individuos o instituciones, ofreciendo un marco legal internacional (desde su versión 4.0). Permiten que el trabajo pueda ser reutilizado, compartido, modificado y copiado , dependiendo de la licencia elegida y dentro de un marco legal.

Las licencias están conformadas por una estructura de tres capas: el código legal, una versión legible por humanos y otras por máquinas. Aunque el conjunto de licencias se divide en libres, y no-libres; solo trabajaremos con las libres ya que son las únicas aptas para un trabajo cultural libre.

Un obra o trabajo cultural libre, debe contener estas cuatro características:

- 1. Libertad de utilizar el trabajo para lo que uno desee, esto va desde realizar copias en diferentes formatos, usarlas para fines políticos, comerciales o religiosos.
- 2. Utilizar la información contenida, esto quiere decir poder utilizar lo aprendido en por ejemplo, textos educativos, diagramas, diseños o papers, en tu propios trabajos.
- 3. Libertad de realizar y distribuir copias, para el propósito que sea mediante el medio que sea, esto quiere decir que puede ser utilizado dentro de un libro, una publicación, etc; tanto la expresión como la información que contiene.
- 4. Libertad para realizar cambios y obras derivadas, esto quiere decir que pueden realizarse traducciones, versiones alternativas, mejoras, etc.

Aunque en muchos casos, ciertas limitaciones o protecciones son impuestas:

- Atribución: Para proteger la integridad de la obra original, o el derecho moral de su autor, se debe reconocer al autor o autores originales de la obra de forma apropiada.
- Compartir Igual: Esto quiere decir, que el trabajo derivado deberá seguir utilizando la misma licencia o una compatible, esto permite que se mantengan el estado de obra o trabajo cultural libre aún con las modificaciones que se le realicen.

Otra opción que surge de estas licencias, es la **CC0** o Dedicación de Dominio Público, donde el autor renuncia a todos los derechos patrimoniales y conexos, que la ley local permite. Nótese, que el derecho moral nunca se pierde, esto quiere decir que nadie puede atribuirse la autoría de la obra más que su autor original.

En los proyectos de Wikimedia se utilizan por lo general (Wikipedia, Wikisource, Wikiespecies, Wikicionario, etc.) una licencia CC-BY-SA 3.0 (Atribución Compartir Igual 3.0),

salvo el caso de Wikidata donde los datos que hacen a la plataforma están en una licencia CC0 de "Dominio Público".

Wikipedia

Introducción

Wikipedia es una enciclopedia **libre, políglota y editada colaborativament**e. Es administrada por la Fundación Wikimedia, una organización sin ánimo de lucro cuya financiación está basada en donaciones. Sus más de **39 millones de artículos en 289 idiomas** han sido redactados conjuntamente por voluntarios de todo el mundo, totalizando más de

2000 millones de ediciones, y prácticamente cualquier persona con acceso al proyecto puede editarlos.

Referencias/Citas y Control de Autoridades

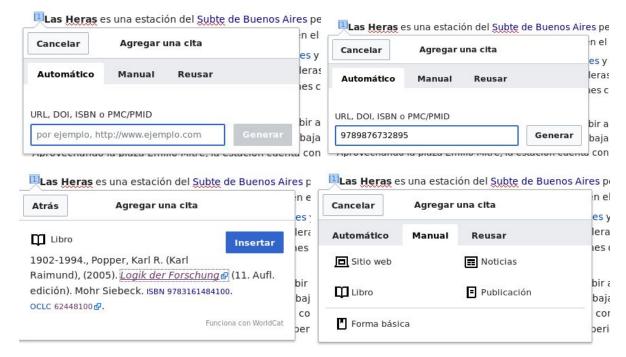
Control de Autoridades

Wikipedia junto a Wikidata, provee al final de cada artículo una plantilla de **control de autoridades**, la cual enlaza a las principales bibliotecas y servicios de este tipo alrededor del mundo. Se alimenta directamente de la información que le provee Wikidata.

```
WorldCatむ・VIAF: 88919448&・ISNI: 0000 0001 2142 9031&・Dialnet: 157452&・BNE: XX1720293&・BNF: cb11892985q& (data)&・CANTIC: a10431366&・GND: 118513532&・LCCN: n79007035&・NDE: 00433878&・NKC: jn19931000310&・NLA: 35100025&・SUDOC: 026739917&・ULAN: 500006476&・SELIBR: 207839&・BIBSYS: 90070625&・ICCU: IT\/CCU\/CFIV\007696&・Open Library: 0L18928A&・BVMC: 2946&
```

CitoID

Es un servicio que genera de forma automática o semi-automática las referencias a partir de un enlace web, número DOI, ISBN, PMID o PMCID, se encuentra disponibles en todas las Wikipedias, ya sea con el editor de texto o visual.



Interfaz de la herramienta.

Wikimedia Commons

Introducción

Con más de **41 millones** de archivos, Wikimedia Commons es uno de los repositorios de archivos multimedia en línea más grandes. Construido con las obras compartidas por miles de voluntarios, **Commons aloja archivos de imágenes, videos y audio**

que son usados para por ejemplo ilustrar Wikipedia y también el resto de los proyectos de la Fundación Wikimedia. Todas las obras en Commons tienen una "licencia libre". Eso significa que cualquiera puede usarlas, compartirlas y/o comercializarlas, simplemente siguiendo las reglas de la licencia – usualmente dando crédito al autor y preservando la licencia para que los demás también puedan compartir la obra.

¿Por qué las instituciones culturales deberían utilizar Wikimedia Commons?

Wikimedia Commons le da la posibilidad a diferentes instituciones culturales de generar un reservorio online de sus archivos. De esta manera, la institución que decide compartir parte de su patrimonio cultural a Commons tiene como ventaja que usuarios de todo el mundo puedan visitar su catálogo en línea y descargarlo con fines educativos. A su vez al digitalizar el contenido también se lo resguarda frente al deterioro del tiempo.

¿Qué se permite subir en Wikimedia Commons?

Puedes subir obras que tú has creado completamente. Eso incluye fotos y videos de:

- Paisajes naturales, animales, plantas
- Personalidades públicas y personas fotografiadas en lugares públicos (sólo con el consentimiento en algunos casos, según las normas locales)
 - Objetos utilitarios o no artísticos.
 - También puedes subir gráficos, mapas, diagramas y audio, siempre que sean originales.

¿Qué no se permite subir en Wikimedia Commons?

Commons no puede aceptar obras creadas o basadas en las obras de otros – incluyendo

fotografías de las mismas. Por defecto, no puedes subir las obras de los demás. Esto incluye materiales como:

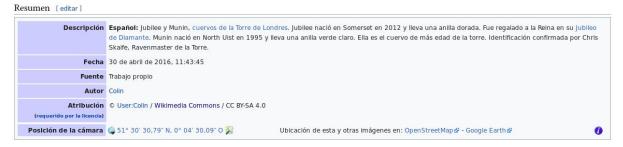
- Logos.
- Tapas de discos (vinilos, CD, DVD, Bluray).
- Fotos promocionales.
- Capturas de pantalla de programas de televisión, películas, DVD y programas informáticos.
- Dibujos de personajes de libros, televisión o películas incluso si los dibujaste tú.
- La mayoría de las imágenes publicadas en Internet.
- Fotos tuyas que no fueron tomadas por vos.
- Obras contempladas y protegidas por la ley 11.723.



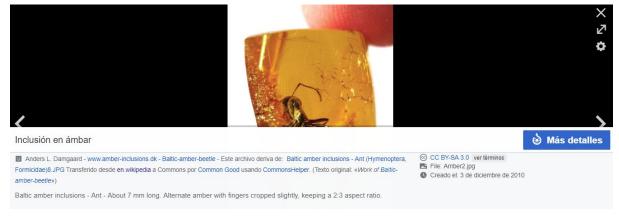
Ficha de metadatos utilizada en obras de arte.



Ficha de metadatos utilizada en libros



Ficha de metadatos utilizada en fotografías.



Visor de imágenes de Wikipedia, mostrando la información del recurso.

Wikisource

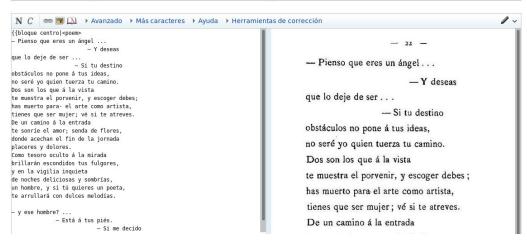
Introducción

Wikisource es la biblioteca o c**ompendio de fuentes primarias** del proyecto, así como Wikidata lo es de datos y Commons de archivos multimedia. Permite trasladar textos, traducirlos, descargarlos en formatos como PDF, ePub, Mobi, etc. Al igual que en el resto de los proyectos, estos textos deben estar con licencia libres o ya en dominio público.

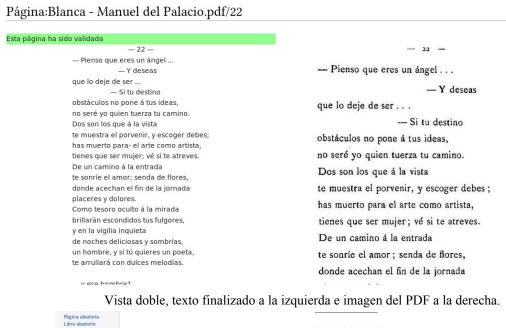
Ejemplos:

Cocina eclectica - Juana Manuela Gorriti Cuentos - Ángel de Estrada A la distancia - Miguel Cané

Edición de «Página:Blanca - Manuel del Palacio.pdf/22»



Vista de edición doble, texto Wiki a la izquierda e imagen del PDF a la derecha.





Vista de un libro finalizada, con índice y las opciones para descargar como ePub, PDF, etc.

Wikidata

Introdución

¿Qué es Wikidata?

Es la base de datos libre, estructurada y multilingüe de Wikimedia. Forma parte de la web semántica, esto quiere decir que está pensada para ser leída y editada tanto por máquinas como humanos. Es posible crear elementos que no tengan un artículo en Wikipedia, pero resulte relevante para los proyectos, o y estén relacionados con uno.

¿Qué quiere decir esto?

Se puede subir el título de un libro por más que este no tenga artículo en Wikipedia,

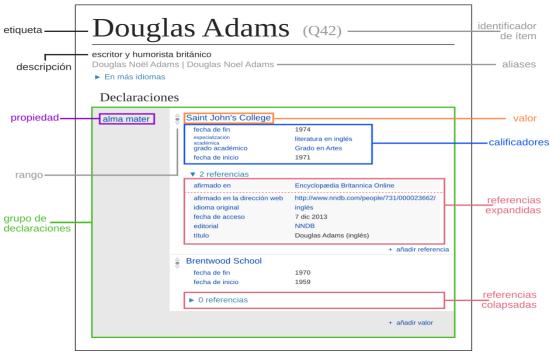
lo importante es que haya un artículo del autor de esa obra o cumpla una necesidad estructural. Se trata de armar una red de datos estructurados que tengan características relacionales entre sí.

Marco de Descripción de Recursos:

Conocido como **Resource Description Framework** en inglés, es un estándar de la web semántica propuesto por la W3C, y utilizado por Wikidata para presentar sus datos. Su principal función es permitir que la información pueda ser leída tanto por humanos como máquina o computadoras, utilizando un modelo para realizar declaraciones sobre recursos basado en tres declaraciones: Entidad, Atributo y Valor. Un ejemplo sería: **Tren (Entidad) - Color (Atributo) - Rojo (Valor).**

Datos enlazados abiertos:

Es un método para intercambiar datos, mostrarlos y conectarlos, utilizando **identificadores de recursos uniformes**. Surgió en el W3C como un proyecto para interconectar datos utilizando RDF (Marco de Descripción de Recursos), el mismo se encuentra en plena expansión y busca conectar información entre diferentes bases de datos utilizando **Identificadores de Recursos Únicos** (URI , por sus siglas en inglés). En Wikidata, cada elemento (objeto, país, persona, especie, etc) tiene un URI, que es una letra Q procedido de un número, por ejemplo Argentina está identificada por el Q414.



Secciones de un elemento en Wikidata.

Identificadores identificador VIAF 88919448 @ editar ▶ 3 referencias + añadir valor € 11892985q identificador BnF @ editar 1 referencia + añadir valor 118513532 identificador GND @ editar > 2 referencias + añadir valor SUDOC 026739917 1 referencia + añadir valor

Algunos identificadores externos en el elemento de Borges.

Web semántica:

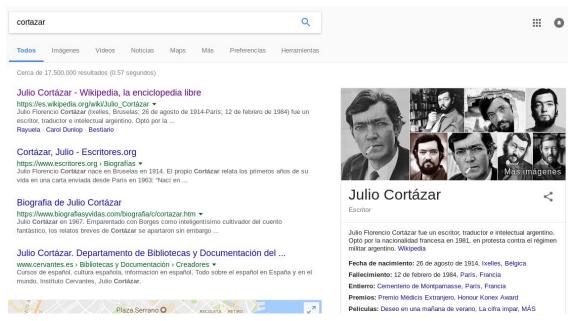
Los últimos dos puntos, como así también Wikidata, forman parte del conjunto de protocolos, reglas y actividades que forman a la Web Semántica. De forma resumida, busca agregar metadatos semánticos y ontológicos, para que la información presentada en la misma esté mejor definida, permitiendo una automaticidad en la descripción llevando a tener una interoperabilidad entre sistemas y convirtiéndolo en un sistema o infraestructura común. Con contenidos, significados y relaciones en un lenguaje universal, resulta más sencillo compartir, procesar y transferir datos. En lo referido a Wikidata se trabajará con OWL (Ontology Web Language), RDF (Resource Description Framework) y SparQL (Protocol and RDF Query Language / Protocolo y lenguaje de consulta de RDF).

Utilizando SPARQL, se pueden obtener resultados personalizados: (Darle al botón Ejecutar o Run para visualizar el mapa)

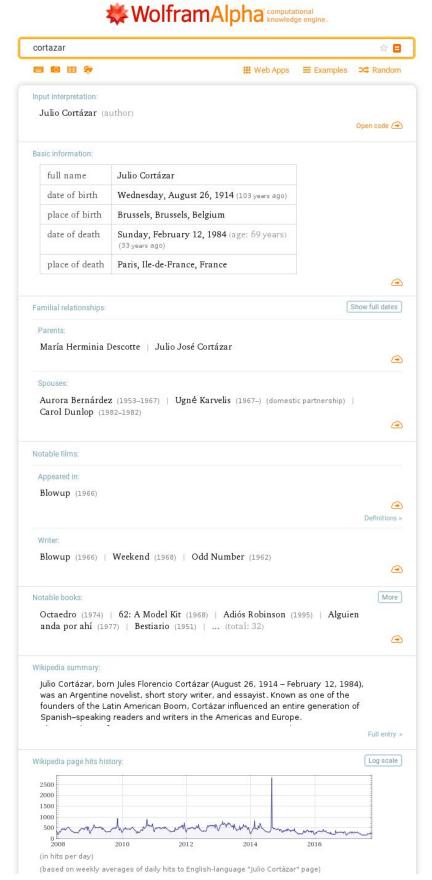
- Parques nacionales alrededor del mundo: http://tinyurl.com/zsyh412
- Localización de las estaciones del Subte H: http://tinyurl.com/zwlnea5
- Poetas según lugar de nacimiento en Argentina: http://tinyurl.com/hl644hm

Knowledge Graph y Wolfram Alpha

Estas dos herramientas, extraen información directamente de Wikidata.



Knowledge Graph de Google.

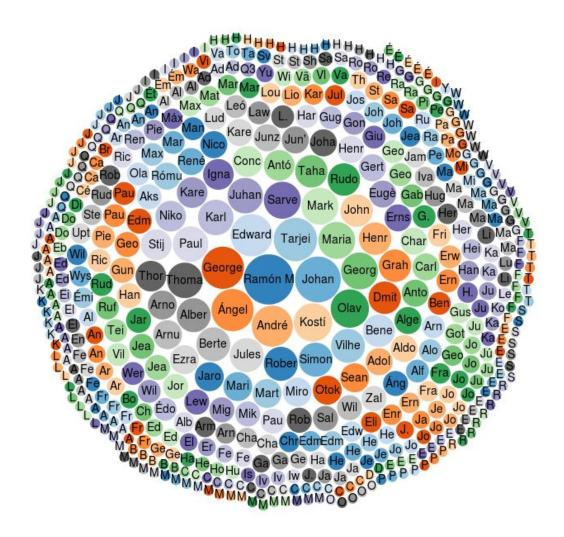


Información provista por Wolfram Alpha sobre Julio Cortazar.

SPARQL

Personas nominadas al premio nobel de literatura: http://tinyurl.com/y8ohweyf

```
#List of authors unsuccessfully nominated for Nobel prize in literature
SELECT ?nominee ?nomineeLabel (SAMPLE(?citizenshipLabel) AS ?country) (COUNT
(DISTINCT ?year) as ?timesNominated)
WHERE
{
    BIND( wd:Q37922 as ?prize )
    ?nominee p:P1411 [ ps:P1411 ?prize; pq:P585 ?time ]
    BIND( year(?time) as ?year )
    OPTIONAL {
        ?nominee wdt:P27 [ rdfs:label ?citizenshipLabel ] filter (lang(?
citizenshipLabel) = "en") .
     }
    FILTER NOT EXISTS { ?nominee wdt:P166 ?prize . }
    SERVICE wikibase:label { bd:serviceParam wikibase:language "es" }
}
GROUP BY ?nominee ?nomineeLabel
ORDER BY DESC(?timesNominated) ?nomineeLabel
```



Bibliografía

- 1. Byrne, J. (2014). New images released are quickly put to use. Wikimedia Foundation.
- Erxleben, F., Günther, M., Krötzsch, M., Mendez, J., & Vrandečić, D. (2014).
 Introducing Wikidata to the Linked Data Web. *The Semantic Web ISWC 2014* Lecture Notes in Computer Science, 50-65. doi:10.1007/978-3-319-11964-9_4
- 3. Ford, H. (2016). Foreword: Wikipedia and the Sum of All Human Information. Nordisk Tidsskrift for Informationsvidenskab og Kulturformidling, årg. 5, nr.
- 4. Hatfield, P. (2013). Canadian Copyright Collection from the British Library on Wikimedia Commons. *Wikimedia Foundation*.
- 5. Kapsalis, E. (2016). The Impact of Open Access on Galleries, Libraries, Museums, & Archives. *Smithsonian Emerging Leaders Development Program*.
- Mitraka, E., Waagmeester, A., Burgstaller-Muehlbacher, S., Schriml, L. M., Su, A. I.,
 & Good, B. M. (2015). Wikidata: A platform for data integration and dissemination
 for the life sciences and beyond. doi:10.1101/031971
- 7. Muzammiluddin, S. (2016). Looking at Wikidata and the future: Gerard Meijssen. *Wikimedia Foundation*.
- 8. Phillips, L., Donaldson, D., Wyatt, Liam, Stierch, S., Hinojo, Alex., (2010). Lessons from Wikipedians in Residence. *GLAM-Wiki US*.
- 9. Stinson, A. (2015). Librarian as public knowledge leader: ways to use Wikipedia. *Wikimedia Foundation*.
- 10. Teplitskiy, M., Lu, G., & Duede, E. (2016). Amplifying the impact of open access: Wikipedia and the diffusion of science. *Journal of the Association for Information Science and Technology*, 68(9), 2116-2127. doi:10.1002/asi.23687
- 11. Vaidya, G., Kontokostas, D., Knuth, M., Lehmann, J., & Hellmann, S. (2015). DBpedia Commons: Structured Multimedia Metadata from the Wikimedia Commons. The Semantic Web - ISWC 2015 Lecture Notes in Computer Science, 281-289. doi:10.1007/978-3-319-25010-6_17