

IV Jornadas Predoctorales (ESI)

WikiPapers, una recopilación colaborativa de literatura sobre wikis usando MediaWiki y su extensión semántica

Emilio José Rodríguez Posada

Tutor: Juan Manuel Doderó Beardo

Universidad de Cádiz

Diciembre de 2012

¿Qué es un wiki?

- Un **wiki** es un **sitio web**
- Que permite a los visitantes:
 - Consultar los contenidos
 - Y también ¡**modificarlos!**
- El primer wiki se inventó en 1996
- Y en 2001 se fundó Wikipedia:
 - Wikipedia es el ejemplo más conocido
 - Pero **existen miles de wikis** por toda la red



Ward Cunningham, inventor del primer wiki.
Foto: Carrigg Photography (CC BY-SA 3.0)

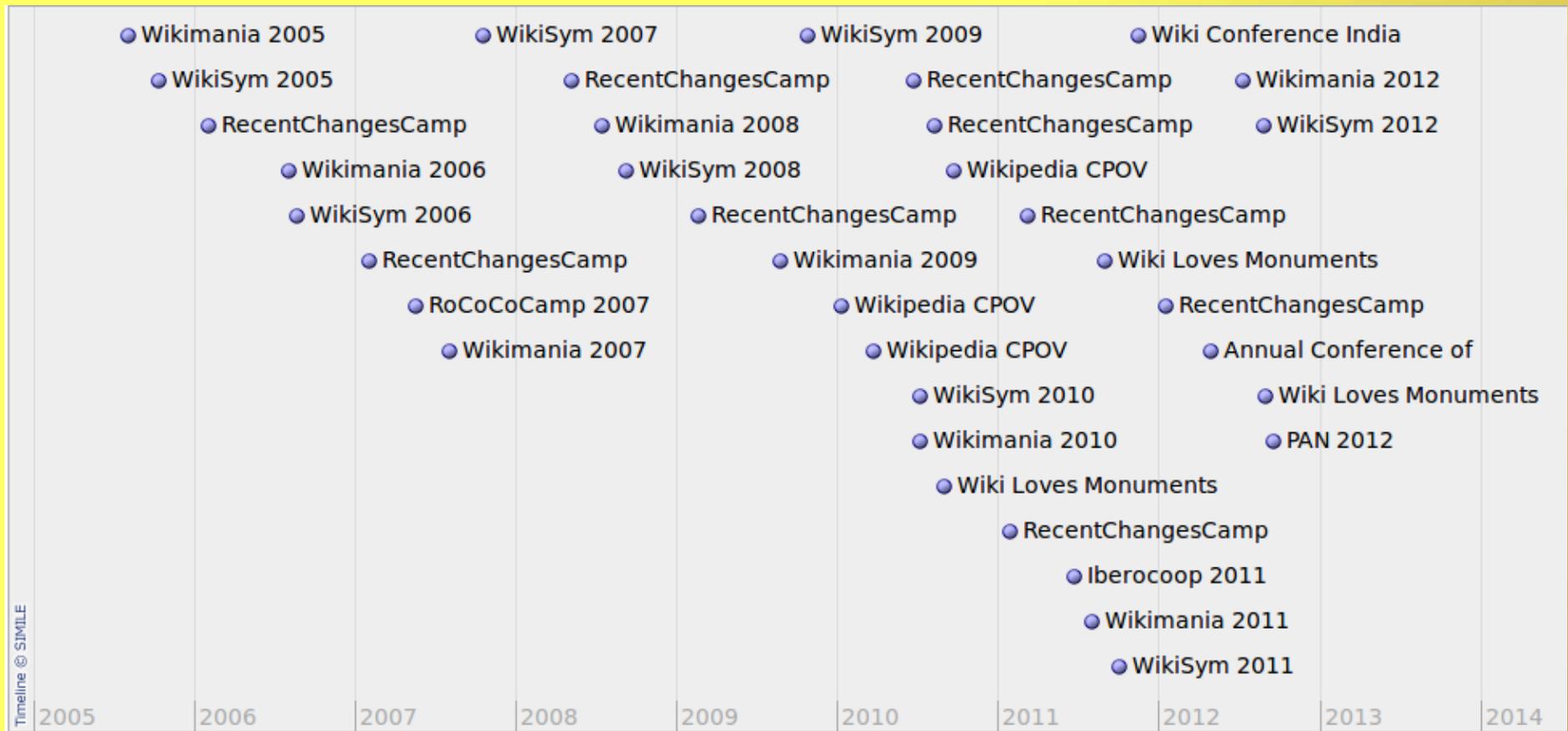
¿Por qué investigar los wikis?

- Los **usuarios** al participar en los wikis **generan muchos datos** de forma indirecta:
 - Estos datos se quedan guardados en el historial
 - Las conversaciones de los usuarios son visibles
 - Y también existen registros de:
 - Visitas, bloqueos, ataques, etc
- Todos estos datos son **públicos**:
 - Y generalmente tienen licencia libre (CC, GFDL, ...)
- De forma que se facilita acceder y analizar estos datos:
 - Se puede estudiar cómo funcionan, sus dinámicas, sus estructuras, el comportamiento de los usuarios, etc

¿Hay actividad investigadora? (I)

- Desde el año 2005 ha ido creciendo el interés por los wikis. Existen numerosos:
 - **Congresos:** WikiSym, CLEF/Pan Lab
 - **Workshops:** WikiAI, SemWiki, MathWikis
 - **Conferencias:** Wikimanía, WikiCon, SMWCon, Wiki Conference India, Wikipedia Academy, Wikipedia CPOV Conference
 - **Competiciones:** WikiViz (visualización datos)

¿Hay actividad investigadora? (II)



Se observa como cada año que pasa hay más eventos

Mucha actividad = Muchos *papers*

- Ha habido **varios intentos de recopilar toda la literatura** sobre wikis. Pero han fallado por el camino:
 - Han resultado ser **incompletos**
 - Han sido **abandonados** con el tiempo y se han perdido los datos
 - No explotan la potencia de los metadatos de las publicaciones
 - Por ejemplo con web semántica
 - Y hay escasa o **nula generación de gráficas**, tablas, estadísticas, etc
- Algunos **ejemplos fallidos** son: grupos de Zotero o CiteULike, recopilaciones en webs personales, etc. Pero todo estaba muy disperso e incompleto.
- Para investigar sobre wikis **necesito conocer el estado del arte**:
 - **Esto me lleva a inventar mi propio sistema: WikiPapers**

<http://wikipapers.referata.com>

¿Por qué hacerlo en un wiki?

- Es fácilmente editable
 - No requiere registro
 - **Es colaborativo**
- Es enriquecido con **semántica**
 - Permite búsquedas, filtrados
 - Inferir conocimiento
- Generación de listados:
 - En base a metadatos
- Generación de **estadísticas**:
 - Gráficas
 - Tablas



The screenshot shows the WikiPapers website interface. It features a navigation sidebar on the left with links like 'Main page', 'Publications', and 'Tools'. The main content area includes a 'Main Page' header, a 'WELCOME TO WikiPapers' message, and several sections: 'WikiPapers' (description), 'How can I help?' (volunteer information), 'How can I create a new page?' (content creation forms), 'License' (copyright information), 'Export' (data export options), and 'Download' (XML dump information). A 'Publications distribution by year' bar chart is visible on the right, showing an increasing trend from 2007 to 2012. The chart data is as follows:

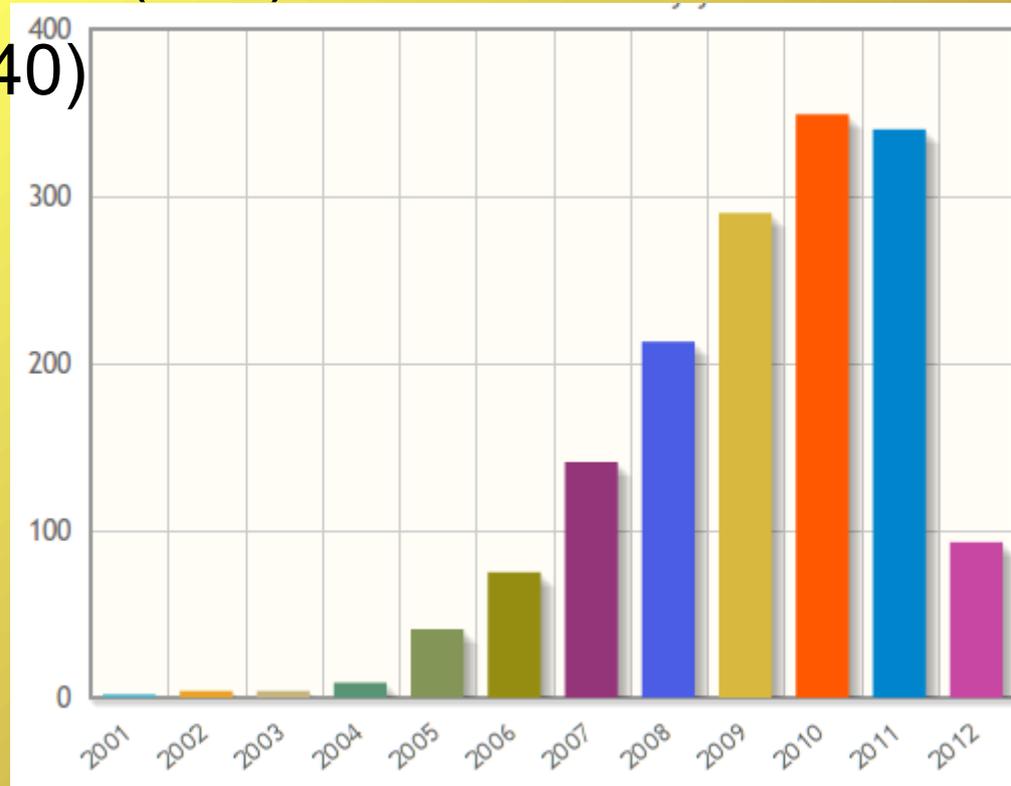
Year	Publications
2007	0
2008	0
2009	0
2010	10
2011	40
2012	130
2013	200
2014	250
2015	250
2016	250
2017	250
2018	250
2019	250
2020	250
2021	250
2022	250

At the bottom, there are links for 'Other academic wikis' and 'WikiPapers' branding.

Captura de pantalla WikiPapers

WikiPapers: estado actual

- El proyecto **WikiPapers** incluye por ahora:
 - Lista de **publicaciones** sobre wikis (+1700)
 - Lista de **conjuntos de datos** (+90)
 - Lista de **herramientas** (+40)
 - Lista de **autores** (+1000)
- También información de:
 - Revistas, eventos
 - Preguntas abiertas
 - Ejemplos
 - Y mucho más...



Número de publicaciones por año

¿Qué información se recoge? (I)

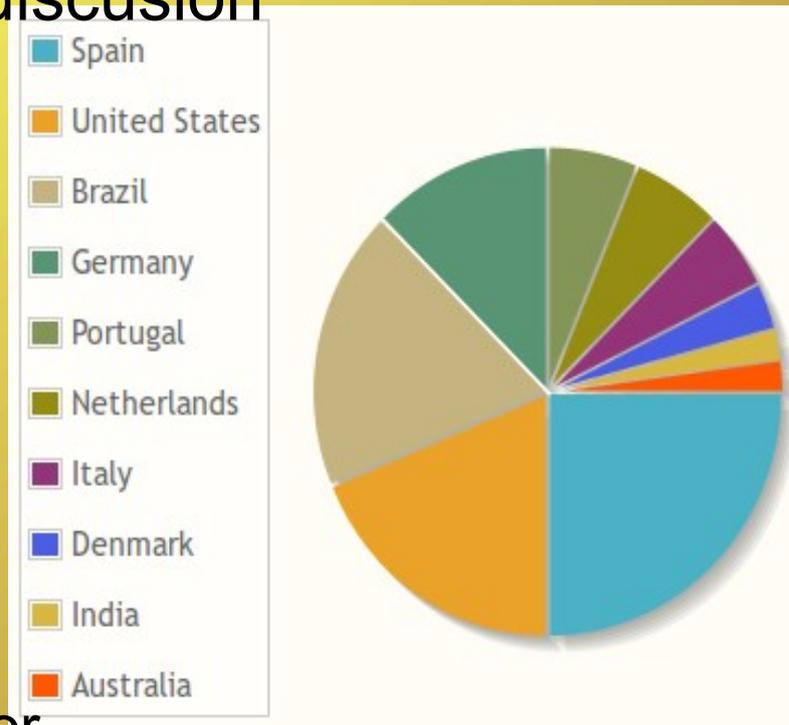
- Sobre las **publicaciones**:

- Título, autores, palabras clave, año, revista o congreso, DOI, idioma, licencia, enlaces al fichero y motores de búsqueda, abstract, las referencias que incluye, las citas que recibe y un espacio de discusión

- Sobre los **autores**:

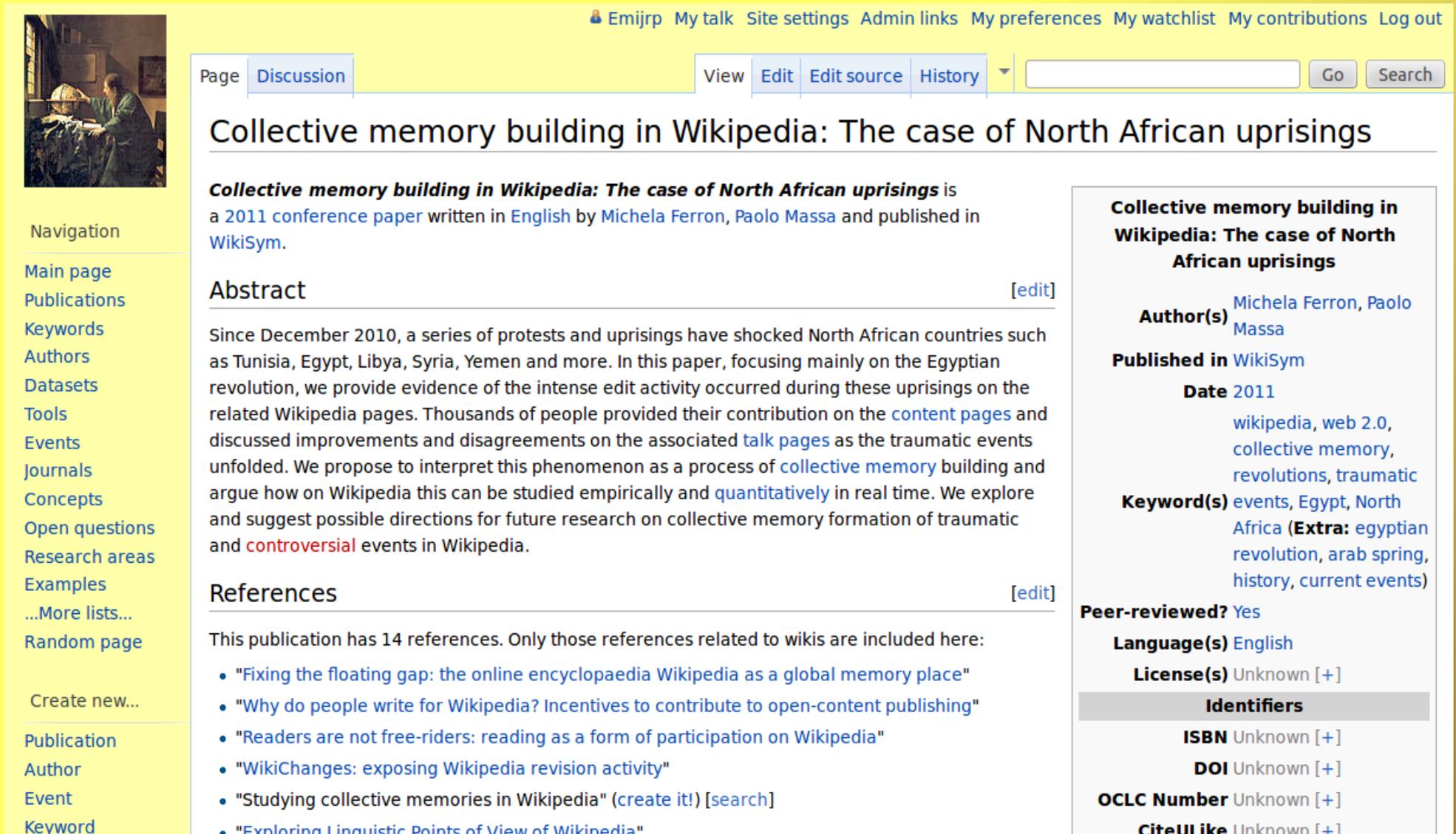
- Nombre, afiliación, país
- Índice de coautores
- Página web
- **Estadísticas**:

- Número de publicaciones
- Y número de citas que recibe el autor



Distribución autores por país

¿Qué información se recoge? (II)



Emijrp My talk Site settings Admin links My preferences My watchlist My contributions Log out

Page Discussion View Edit Edit source History

Collective memory building in Wikipedia: The case of North African uprisings

Collective memory building in Wikipedia: The case of North African uprisings is a 2011 conference paper written in English by Michela Ferron, Paolo Massa and published in WikiSym.

Abstract [edit]

Since December 2010, a series of protests and uprisings have shocked North African countries such as Tunisia, Egypt, Libya, Syria, Yemen and more. In this paper, focusing mainly on the Egyptian revolution, we provide evidence of the intense edit activity occurred during these uprisings on the related Wikipedia pages. Thousands of people provided their contribution on the [content pages](#) and discussed improvements and disagreements on the associated [talk pages](#) as the traumatic events unfolded. We propose to interpret this phenomenon as a process of [collective memory](#) building and argue how on Wikipedia this can be studied empirically and [quantitatively](#) in real time. We explore and suggest possible directions for future research on collective memory formation of traumatic and [controversial](#) events in Wikipedia.

References [edit]

This publication has 14 references. Only those references related to wikis are included here:

- "Fixing the floating gap: the online encyclopaedia Wikipedia as a global memory place"
- "Why do people write for Wikipedia? Incentives to contribute to open-content publishing"
- "Readers are not free-riders: reading as a form of participation on Wikipedia"
- "WikiChanges: exposing Wikipedia revision activity"
- "Studying collective memories in Wikipedia" (create it!) [search]
- "Exploring Linguistic Points of View of Wikipedia"

Collective memory building in Wikipedia: The case of North African uprisings

Author(s) Michela Ferron, Paolo Massa

Published in WikiSym

Date 2011

Keyword(s) wikipedia, web 2.0, collective memory, revolutions, traumatic events, Egypt, North Africa (**Extra:** egyptian revolution, arab spring, history, current events)

Peer-reviewed? Yes

Language(s) English

License(s) Unknown [+]

Identifiers

ISBN Unknown [+]

DOI Unknown [+]

OCLC Number Unknown [+]

CiteULike Unknown [+]

Ejemplo de *paper* con metadatos a la derecha, abstract y referencias

http://wikipapers.referata.com/wiki/Collective_memory_building_in_Wikipedia:_The_case_of_Nor

¿Qué información se recoge? (III)

Emijrp My talk Site settings Admin links My preferences My watchlist My contributions Log out

Page **Discussion** View Edit Edit source History Go Search

Felipe Ortega

Felipe Ortega is an [author](#) from [Spain](#).

Tools

Tool	Description
WikiXRay	WikiXRay is a robust and extensible software tool for an in-depth quantitative analysis of the whole Wikipedia project.



Felipe Ortega
(Alternative names for this author)



Affiliation Unknown [+]
Country Spain
Antonio J. Reinoso, Daniel Izquierdo-Cortazar, Gregorio Robles, Israel Herraiz, Jesús M. González-Barahona, Joaquín Rodríguez López, Joseph M. Reagle, Luca de Alfaro, **Rut Jesus**
Co-authors
Website <http://felipeortega.net/>

Statistics
Authorship Publications (10), datasets (0), tools (1)
Total (14), average (1.4), median (0),

Navigation
Main page
Publications
Keywords
Authors
Datasets
Tools
Events
Journals
Concepts
Open questions
Research areas
Examples
...More lists...
Random page
Create new...
Publication
Author
Event
Keyword

Ejemplo de autor con metadatos a la derecha, lista de herramientas, vídeo incrustado...
http://wikipapers.referata.com/wiki/Felipe_Ortega

¿Qué información se recoge? (IV)

- También hay información sobre:
 - **Conjuntos de datos (*Datasets*)**
 - Bases de datos de Wikipedia (son públicas)
 - Y también de otros wikis de Internet
 - **Herramientas**
 - *Frameworks* para extraer datos de wikis
 - Procesamiento de datos
 - Estadísticas y visualización
- Todos los **metadatos** de publicaciones, autores, *datasets* y herramientas son **exportables**:
 - En los formatos BibTeX, RDF, CSV y JSON

Conclusiones y trabajo futuro

- **WikiPapers** es un **sistema para recopilar literatura sobre wikis** (ya hay +1700 *papers*, +90 *datasets*, +40 herram., +1000 autores)
- Soluciona los problemas que existían en otros enfoques
 - **Tiene mejoras:** estadísticas, gráficas, tablas, enriquecido con semántica
- WikiPapers es un **estado del arte** en continua actualización:
 - Es actualizado si aparecen nuevos papers
 - A diferencia de los **Systematic Literature Reviews** que lo publicas en una revista y son estáticos
- Ejemplo práctico, las **gráficas** de esta presentación se han extraído de WikiPapers:
 - **Y se actualizan automáticamente cuando hay cambios**
- **El proyecto ha tenido buena acogida** entre los investigadores de este campo:
 - Y algunos ya han participado añadiendo publicaciones

Gracias por su atención



¿Preguntas?

<http://wikipapers.referata.com>