

# **Web2Cit**

## **Improve automatic citations in Wikipedia**

**Diegodlh & Scann**

<https://meta.wikimedia.org/wiki/Web2Cit>

[https://commons.wikimedia.org/wiki/File:Web2Cit\\_at\\_LD4\\_Wikidata\\_Affinity.pdf](https://commons.wikimedia.org/wiki/File:Web2Cit_at_LD4_Wikidata_Affinity.pdf)

...

✕ **Add a citation**

**Automatic** Manual Re-use

URL, DOI, ISBN, PMC/PMID, QID, title, or citation

https://www.pagina12.com.ar/365595 **Generate**

...

< **Add a citation**

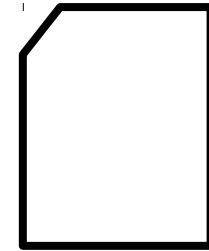
📁 Website **Insert**

Página12 (1630607038). "Mi Reloj Interno", una app que permite mejorar el descanso | Fue desarrollada por científicas argentinas" 🌐. *PAGINA12*. Retrieved 2021-09-13.  
[Check date values in: |date= \(help\)](#)

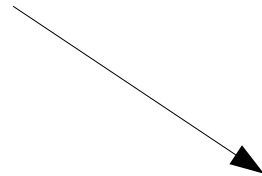
Powered by Zotero



*translator*



web  
source



{ }

← translator

{ }

{ }

← translator

{ }

{ }

← translator

~~{ }~~

{ }

← translator

~~{ }~~

{  }

← translator

~~{ }~~

{ }

← translator

~~{ }~~

## Citoid performance

Andrew Lih & Robert Fernandez'  
Citoid performance for news article citations

Citoid performance for most cited news sites (n=91)

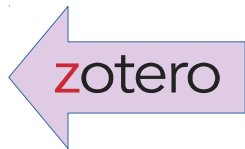


only “31.9% had all four components correct”

Our (preliminary) results:

- 48% correct responses

<https://meta.wikimedia.org/wiki/Web2Cit/Research>



```
if (programmer) {  
  return "OK"  
} else {  
  throw new Error();  
}
```

# Web2Cit

# Web2Cit



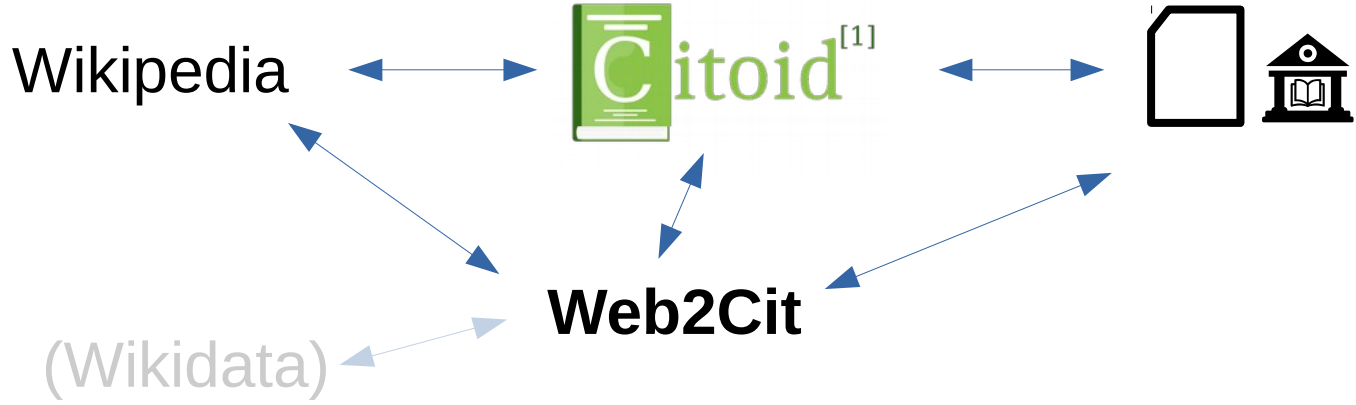
**WIKIMEDIA**  
FOUNDATION

**Diego de la Hera**  
*Project manager*  
*Lead developer*

**Evelin “Scann” Heidel**  
(Wikimedistas de Uruguay)  
*Communications &*  
*Community Manager*

**Gimena del Rio Riande**  
**Nidia Hernández**  
**Romina De León**  
(HD CAICYT LAB – CONICET)  
*Research subproject*  
*members*

**Dennis Tobar**  
*Development subproject*  
*member*



**define and maintain**

**translation procedures and tests**

**simply and collaboratively**

## Web2Cit configuration

- Per **domain** (website)
  - Translation **tests**: path (webpage)
    - Translation **goals**
  - Translation **templates**: path (webpage)
    - Translation **procedures**
      - **Selection** steps
      - **Transformation** steps

<https://www.example.com/news/article1>

<i>domain</i> (website)	<i>path</i> (webpage)
----------------------------	--------------------------



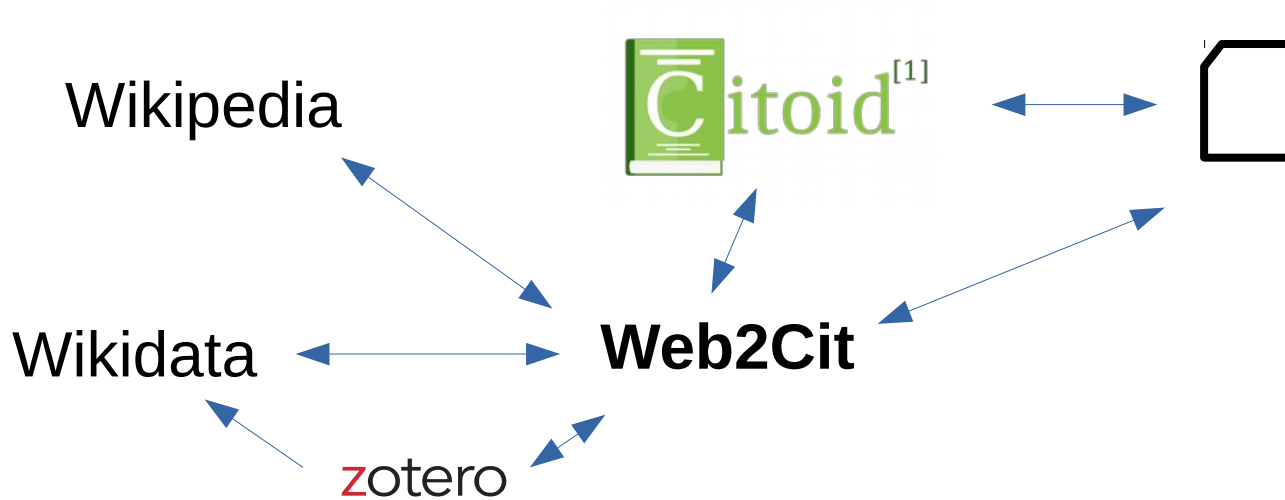
# Demo

- User script
- Web2Cit translation summary
- Translation test
- Translation template
- Other webpage, same domain

**Multiple translation templates**

**Template applicability**

**Translation subgroups**



- Automatic references in Wikidata:
  - Citoid integration (T199197): [demo](#)
  - [CiteTool](#)
- Automated item creation:
  - Direct: *“lets make it dead simple to input bibliographic data on Wikidata”* (Wikicite Telegram)
  - Via Zotero

# Thanks!

## Introduction guide

[https://meta.wikimedia.org/wiki/Web2Cit/User\\_guide](https://meta.wikimedia.org/wiki/Web2Cit/User_guide)

- Videos
  - How Web2Cit works
  - Web2Cit ecosystem
  - How to use Web2Cit
- Links
  - User script
  - Translation fields
  - Selection steps
  - Transformation steps

## Web2Cit research

<https://meta.wikimedia.org/wiki/Web2Cit/Research>

## News & updates

<https://meta.wikimedia.org/wiki/Web2Cit/News>

- Monitor
- Real-time editor & sidebar
- Selection steps
  - URL (T304326)
  - Header (T304333)
  - JSON-LD (T304332)
- Transformation steps
  - Case (T302692)
  - Replace (T302691)
  - Custom (T305883)

## Continue conversation

- Web2Cit discussion:  
<https://meta.wikimedia.org/wiki/Talk:Web2Cit>
- Phabricator
  - #web2cit