

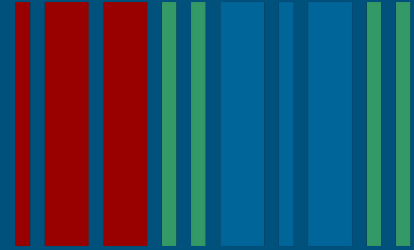
# Wikidata-powered infoboxes on Catalan Wikipedia

[[User:Paucabot]]

Aberystwyth, 05-07-2018

<http://tiny.cc/celticknot>


# What is Wikidata?




- [Types of subatomic particles](#)
- [Descendants of King James I](#)
- [Occupations of women](#)
- [Cities of more than 1 million people](#)
- [Porn actors by sex](#)
- [Films shot in the Balearic Islands](#)

# What is an infobox?


[:en:Help:Infobox](#)

 **Aberystwyth**




**Etimologia** Riu Ystwyth

**Localització**



Wikimedia | © OpenStreetMap

 52°25′00″N, 4°04′00″O

Estat	Regne Unit
Pais	Gal·les
Principal area of Wales	Ceredigion

**Geografia**

Forma part de	Ceredigion
Banyat per	Cardigan Bay i Riu Ystwyth
Limita amb	Clarach Bay

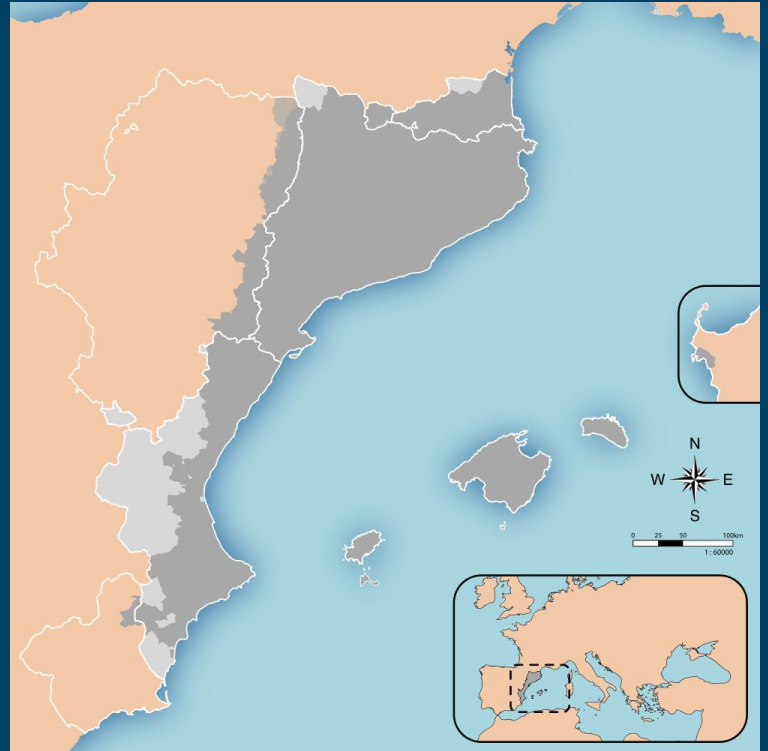
**Indicatiu**

Codi postal	SY23
Prefix telefònic	01970

**Altres dades**

Agermanament	<span>[<a href="#">amaga</a>]</span>
	Kronberg im Taunus
	Saint-Brieuc
	Esquel

# What is Catalan?



Territory	State	Understand <sup>1[48]</sup>	Can speak <sup>2[48]</sup>
 Catalonia	 Spain	6,502,880	5,698,400
 Valencian Community	 Spain	3,448,780	2,407,951
 Balearic Islands	 Spain	852,780	706,065
 Roussillon	 France	203,121	125,621
 Andorra	 Andorra	75,407	61,975
 La Franja (Aragon)	 Spain	47,250	45,000
 Alghero (Sardinia)	 Italy	20,000	17,625
 Carche (Murcia)	 Spain	No data	No data
Total Catalan-speaking territories		11,150,218	9,062,637
Rest of World		No data	350,000
Total		11,150,218	9,412,637

[:en:Catalan language](#)

### Social use [\[ edit \]](#)

Area	At home	Outside home
Catalonia	45	51
Valencian Community	37	32
Balearic Islands	44	41
Roussillon	1	1
Andorra	38	51
Franja Oriental of Aragón	70	61
Alghero	8	4

(% of the population 15 years old and older).

[:en:Catalan language](#)

---

# Catalan Wikipedia

[ca.wikipedia.org](https://ca.wikipedia.org)

- 580k articles
- 20M editions
- 1400 active users
- 25 admins

Source

---



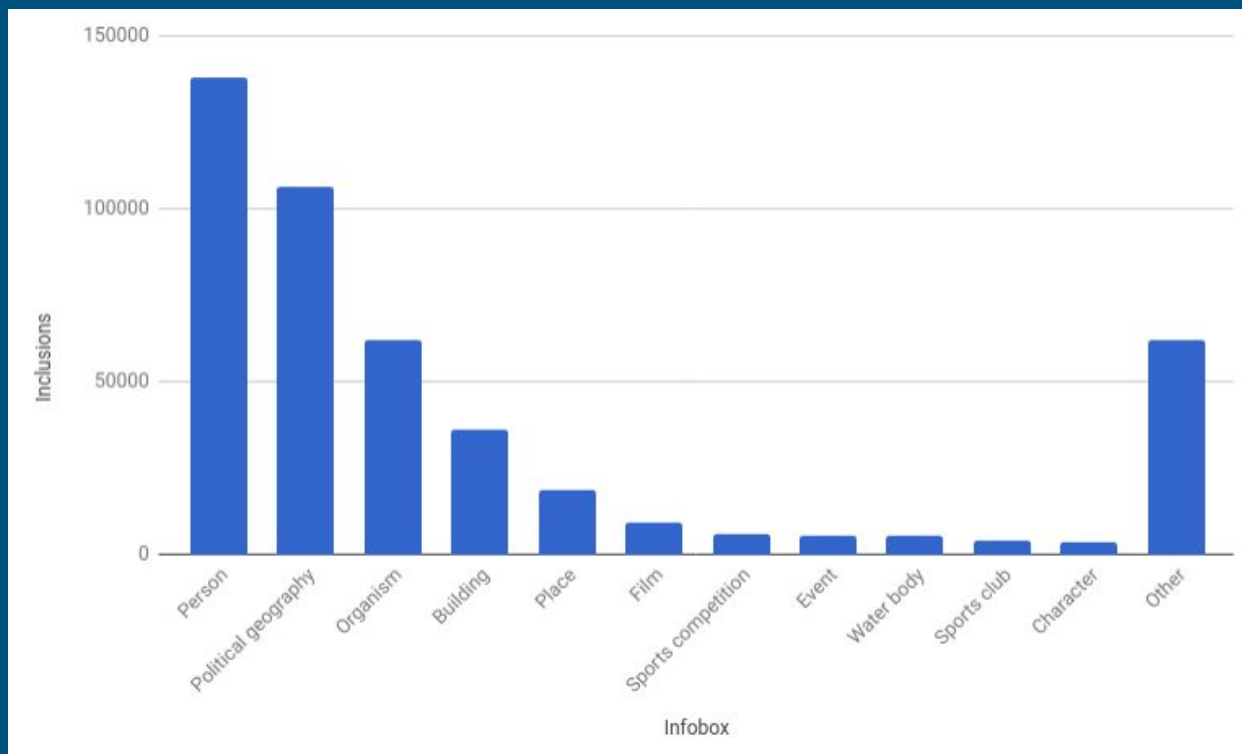
# Statistics





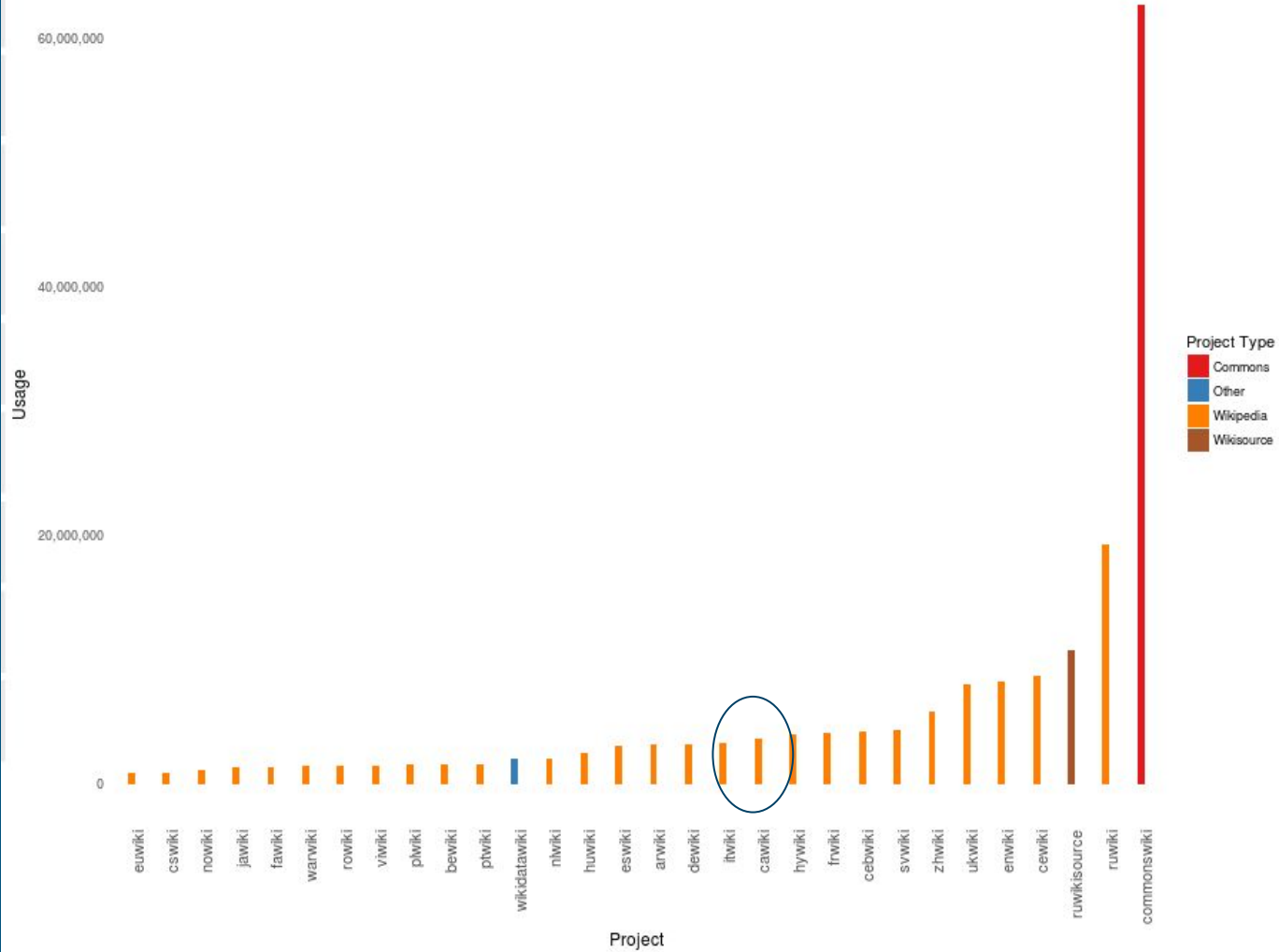
# Statistics

TOTAL ARTICLES	582326
ARTICLES WITH WIKIDATA INFOBOX	458059
<b>78,66%</b>	



# Comparison with other wikipedias

[Source](#)





# Design

---



Biographies

Organizations

Culture

Science & Technology

Geography

Buildings & Structures

Events

Organisms



# Technology



Technology:



We have:

Module:Wikidata ([Q12069631](#))

Module:Infobox ([Q13107716](#))

Template:Infobox ([Q5626735](#))

# The project

- Started in december, 2014
- 10 users involved
- 300+ deprecated infoboxes ([example](#))
- [80 wikidata-powered infoboxes created](#)
- Up to 100 parameters imported by a single infobox
- [Extensive use](#)
- More wikidata-powered infoboxes → more info to Wikidata
- Recommended infoboxes for starting: people, chemical compounds, drugs, organisms, films and videogames
- Some of the infoboxes are very very complex

La guerra dels mons

*The War of the Worlds*

*The WAR of the WORLDS*  
By H. G. Wells  
Author of "Under the Earth," "The Time Machine," etc.



Tipus llibre

**Fitxa tècnica**

**Autor** H.G. Wells

**Llengua** anglès

**Publicació** Regne Unit, 1898

**Editorial** Heinemann

**Edicions** Q41530078

**Detalls de l'obra**

**Gènere** novel·la de ciència-ficció i ciència-ficció

**Nominacions i premis**

**Premis** NPR Top 100 Science Fiction and Fantasy Books

**Sèrie**

← L'home invisible

The Sleeper Awakes, Edison's Conquest of Mars i  
Fighters from Mars →

# Before

[:ca:Pau Casals i Defilló](#)

# After

```
{{Infotaula de persona
|imatge      = Pablocasals.jpg
|data_naixement  = {{Data naixement|1876|12|29}}
|data_defuncio   = 22-10-1973
|sepultura      = Cementiri d[[el Vendrell]]
|pares         = [[Carles Casals i Riba]]<br/>[[Pilar
Defilló Amiguet]]
|instrument     = [[Violoncel]], violí, piano, flauta
|ocupacio      = [[Pedagog]], [[director
d'orquestra|director]] i [[compositor]]
|lloc_web      = http://www.paucasals.org
}}
```

```
{{Infotaula de persona}}
```

---





At what cost?



# Before

{{Infobox  
| name = Infobox number

| title = {{PAGENAME}}

| image = [[File:{{{image|}}} |size=250x100px]]

| caption = {{{caption|}}}

| label1 = Type

| data1 = {{{type|}}}

| label2 = Eponym

| data2 = {{{eponym|}}}

...

```
{{Infobox
| name    = Infobox number

| title   = {{PAGENAME}}

| image   = [[File:{{#invoke:Wikidata |claim |property=P18 |value={{image}}} | list=false }}
|size=250x100px]]
| caption = {{#invoke:Wikidata | claim | property=P18 |qualifier=P2096 | formatting=ca }} }}

| label1  = Type
| data1   = {{#invoke:Wikidata | claim | property=P31 | value={{type}}} }}

| label2  = Eponym
| data2   = {{#invoke:Wikidata | claim | property=P138 | value={{eponym}}} }}

...
```

... and after

# Let's analyze it:

```
{{#invoke:Wikidata | claim | property=P31 | value={{type}}}}
```

Segur <https://www.wikidata.org/wiki/Q712986>

### Statements

instance of	natural number ...	<a href="#">edit</a>
	followed by 64 ...	
	follows 62 ...	
	0 references	<a href="#">+ add reference</a>
	odd number ...	<a href="#">edit</a>
	followed by 65 ...	
	follows 61 ...	
	0 references	<a href="#">+ add reference</a>
	centered octahedral number ...	<a href="#">edit</a>
	0 references	<a href="#">+ add reference</a>
	22-gonal number ...	<a href="#">edit</a>
	0 references	<a href="#">+ add reference</a>
	composite number ...	<a href="#">edit</a>
	0 references	<a href="#">+ add reference</a>
		<a href="#">+ add value</a>



## 7 Seixanta-tres

<b>Tipus</b>	nombre natural, nombre senar, nombre octaèdric centrat, nombre 22-gonal i nombre compost
<b>Propietats</b>	
<b>Valor</b>	63
<b>Factorització</b>	$3^2 \times 7$
<b>Altres numeracions</b>	
<b>Numeral romà</b>	LXIII
<b>Binari</b>	111111 <sub>2</sub>
<b>Hexadecimal</b>	3F <sub>16</sub>
<a href="#">← seixanta-dos</a> <a href="#">seixanta-quatre →</a>	

(Q712986)

:ca:Seixanta-tres



What properties do we  
import?



# Example

Common name, scientific name, image (+caption), audio (+caption), video (+caption), means of locomotion, gestation period, color, main food source, this taxon is source of, life expectancy, habitat, host, generation time, length, height, mass, wingspan **MICROORGANISMS** -> ICTV virus genome composition, discoverer, illness, Gram staining **GENETICS** -> ploidy, genome size **MUSHROOMS** -> hymenium type, mushroom cap shape, hymenium attachment, stipe character, spore print color, mushroom ecological type, edibility **PLANTS** -> is pollinated by, is pollinator of, flower color, has fruit type **FOSSILS** -> temporal range start, temporal range end, IUCN conservation status **NOMENCLATURE** -> parent taxon, taxon name, taxon rank, taxon author, author citation, botanist author abbreviation, date of taxon name publication, taxonomic type, nomenclatural status, original spelling, named after, basionym, this zoological name is coordinate with, taxon synonym, replaced synonym, ex taxon author, gender of a scientific name of a genus, original combination **DISTRIBUTION** -> taxon range map image (+caption), endemic to.

**Orangutan de Borneo**

*Pongo pygmaeus*



**Estat**

Extinta | Amenaçada | Risc mitjà

EX EW **CR** EN VU NT LC

En perill greu

**Taxonomia**

Super-regne	Eukaryota
Regne	Animalia
Fílum	Chordata
Classe	Mammalia
Ordre	Primates
Família	Hominidae
Gènere	<i>Pongo</i>
Espècie	<b><i>Pongo pygmaeus</i></b>

LINNAEUS, 1760

**Distribució**





## Diamant



**Fórmula química** C  
**Epònim** Invincible

### Classificació

**Categoria** Mineral  
**Nickel-Strunz 10a ed.** 1.CB.10a  
**Nickel-Strunz 9a ed.** 1.CB.10a  
**Nickel-Strunz 8a ed.** I/B.02  
**Dana** 1.3.6.1

### Propietats

**Hàbit cristal·li** Octàedre  
**Simetria** cubic-hexoctahedral  
**Grup espacial** space group 227  
**Massa molar** 12,01 grams per mol  
**Exfoliació** {111}  
**Fractura** fractura concoïdal  
**Duresa** 10  
**Color de la ratlla** blanc  
**Densitat** 3,51 i 3,515  
**Índex de refracció** 2,418  
**Estatut IMA** mineral heretat (G)

## Coreà

한국어 i 조선말

### Parlants

73.233.270

nadius



### Classificació lingüística

llengües humanes

llengües coreanes

### Característiques

**Sistema d'escriptura** Hangul i Hanja

**Dialectes** dialecte de Seül, llengua jeju, Chungcheong dialect, Gyeongsang dialect, Korean language in China, Koryo-mar, Zainichi Korean language i Pyongan dialect

**Institució de normalització** Institut Nacional de Llengua Coreana

**Estudiat per** estudis coreans

### Codis

**ISO 639-1** ko

**ISO 639-2** kor

**ISO 639-3** kor🔗

**Glottolog** kore1280🔗

**Linguasphere** 45-AAA-a🔗

**IETF** ko

More samples



# Pros and cons





# Pros

- Easy to add
- Help standardize articles
- Take advantage of Wikidata
- It really helps improve Wikidata

# Cons

- There is some (technical) work to create them
  - There are some parameters not handled by Wikidata
  - There is some tweaking required for plural, capital letters, feminine forms ...
-

# Next steps



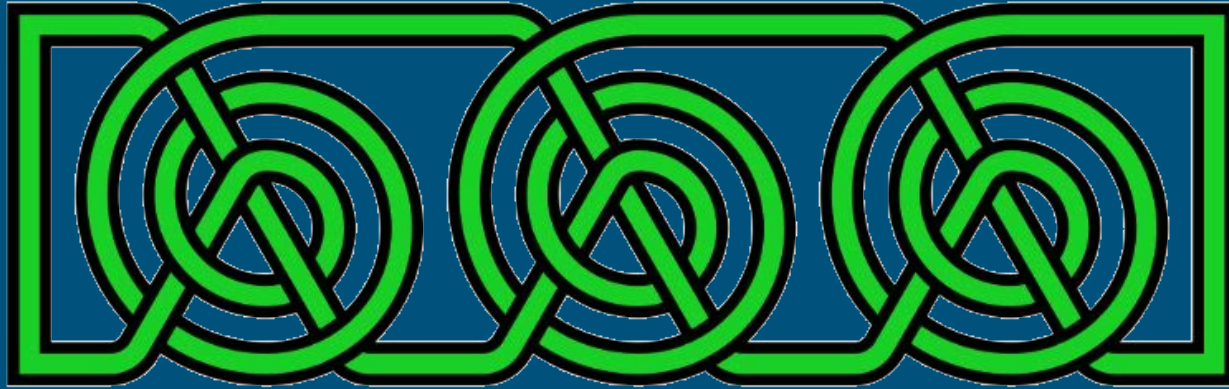
- Multilingual infoboxes [People](#) , [Building](#)
- en wikipedia
- More wikidata-powered templates
- Less wikidata-powered templates
- Cross-wiki infoboxes
- Less parameters?
- Substitute automatic infoboxes for manual ones?

# Different approaches

---

Here

# What about Celtic and indigineous languages?



How can we continue?

<https://labur.eus/ohWTFs>

Thanks to

---

[[User:Mike Peel]] + [[User:Amadalvarez]]  
+ [[User:Theklan]] + [[User:Vriullop]]

# Questions?

