



# Contribution to Wikidata

Béatrice MARKHOFF, Arnaud SOULET, Léo-Paul RENARD

# Outlines

- **Reminder**
- **Process of adding data**
- **Test**
- **Conclusion**



- Knowledge base
- Presented in 2012 and opened in 2013
- 13 Mds triples, 94M elements, 9 000 properties
- Enriches various related projects (Wikipedia, Wikimedia Commons)

Complex information source for **humans** and  
**machines**

## 3 ways to add data

```
graph TD; A[3 ways to add data] --> B[By independent contributors]; A --> C[By automated robots]; A --> D[By data donations];
```

### By independent contributors

Data verification

### By automated robots

More strict :  
Robot policy

### By data donations

More strict  
and conscientious :  
Contact, exchange, work  
with the community

**All the elements** of Wikidata must be **notable** and therefore correspond to at least 1 of the 3 following cases :

- There is a **corresponding page** on a Wikimedia site
- There is a **clearly identifiable conceptual or material entity** that can be described using serious and publicly available references
- It fills a structural need

Test subject : Kenny McCormick (Q47386)

Test property : cause of death (P509)

## Add via Sparql query

```
INSERT DATA  
{wd:Q47386 wdt:P509 "parka".}
```

Query is malformed: Encountered " "insert" "INSERT "" at line 2, column 1.

**/?\** Blocking by the endpoint ? **/?\**

Test subject : Kenny McCormick (Q47386)

Test property : cause of death (P509)

## Add manually

*Addition of "parka" as cause of death*

The screenshot shows the Wikidata interface for the property 'cause de la mort'. On the left, a blue sidebar contains the text 'cause de la mort'. The main area displays a list of values for this property: 'peste', 'suicide collectif', and 'Mort de rire'. Each value has a 'modifier' button and a link to '1 référence'. A dropdown menu is open for the 'parka' value, showing a list of suggestions: 'parka', 'anorak (parka) manteau à capuche', 'parka vêtement', and 'Parka genus of algae'. To the right of the dropdown, there are buttons for 'publier', 'annuler', and a question mark icon. Below these buttons are links to '+ ajouter un qualificatif', '+ ajouter une référence', and '+ ajouter une valeur'.

**Auto-completion** to find already existing elements

**!/\ We cannot add an element that is not present in the list** and therefore not linked to a clearly identifiable conceptual entity. **!/\**

Test subject : Kenny McCormick (Q47386)  
Test property : cause of death (P509)

## Add manually

We can see that this one appears well during a query

The screenshot shows the Wikidata Query Service interface. The query is: `1 select (count(distinct ?o) as ?nbOfDths)  
2 where {  
3 wd:Q47386 <http://www.wikidata.org/prop/direct/P509> ?o.  
4 SERVICE wikibase:label { bd:serviceParam wikibase:language "fr". }  
5  
6 }`. The results table has one column, `nbOfDths`, and one row with the value `21`. The column and row are highlighted with a red box.

nbOfDths
21

*Before adding*

The screenshot shows the Wikidata Query Service interface after adding a new result. The query is the same as in the previous screenshot. The results table has one column, `nbOfDths`, and two rows: the first row has the value `21` and the second row has the value `22`. The column and the second row are highlighted with a red box.

nbOfDths
21
22

*After adding*



Test subject : kenny mccormick (new element)  
Test property : cause of death (P509)

## Add manually

Add a new element,  
similar to an existing one

Élément Discussion

**kenny mccormick** (Q107361535)

personnage de la série TV South Park [modifier](#)

Plus de langues

Langue	Libellé	Description	Également connu comme
français	kenny mccormick	personnage de la série TV South Park	
anglais	Pas de libellé défini	Aucune description fournie	
basque	Pas de libellé défini	Aucune description fournie	
espagnol	Pas de libellé défini	Aucune description fournie	

Déclarations

cause de la mort [parka](#) [0 référence](#) [modifier](#)

+ ajouter une référence

+ ajouter une valeur

+ ajouter une déclaration

Verification of its creation

Wikidata Query Service

```
1 select ?oLabel
2 where {
3   wd:Q107361535 <http://www.wikidata.org/prop/direct/P509> ?o.
4   SERVICE wikibase:label { bd:serviceParam wikibase:language '[AUTO_LANGUAGE]*'. }
5 }
6 }
```

oLabel

parka

## Add manually

Added reference to give **credibility** to the statement

### Information concerning the statements

▶ : Suggestion to improve statement

⚠ : There are potential problems with the instruction

**Efficient manual addition process**

- The problem that may arise is **not semantic** but on **the assertions** made
- Human bias involved
- **Reliability of the automated robot is more important** because it relies on the knowledge and reasoning/deductions made from the ontologies but it relies on the users' insertions too so it can bias its work

We can have a **fragility of the data** and their veracity because **no effective and regular verification** on the addition of these

/?\ The largest amount of added data comes from **data donations** and **reasoning via bots** which have a more protocol policy so wouldn't manual addition be omitted given the percentage that this technique represents? /?\

# Bibliography

- [https://www.wikidata.org/wiki/Wikidata:Main\\_Page](https://www.wikidata.org/wiki/Wikidata:Main_Page)
- <https://www.wikidata.org/wiki/Wikidata:Contribute/f>
- <https://www.wikidata.org/wiki/Wikidata:Bots/fr>
- <https://fr.wikipedia.org/wiki/Wikidata>