

# DATI FANTASTICI E DOVE TROVARLI

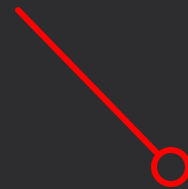
## Wikidata, SIRBeC e OpenRefine

ItWikiCon 2022, Verbania  
1º ottobre, Sala Zeda

Elena Martellotta – Elemar (WMIT)  
Mattia Minelli – Hitrandil (WMIT)



# DATI REGIONE LOMBARDIA



**Wikidata**

**OpenRefine**

**QuickStatements**

**2021**

Analisi dell'openess del  
patrimonio culturale italiano



Progetto pilota replicabile

## Architetture

## Architetture della Regione Lombardia

NUM_...	TIPO_...	DENOMINAZIONE	COMPLESSI...	DENOMINAZIONE	LOC
BS270-001...	palazzo	Palazzo Fisogni - complesso	bene complesso	Palazzo Fisogni - c	
CO250-00...	casa	Stallazzo a Villa Incino - complesso	bene complesso	Stallazzo a Villa In	
CO260-00...	chiesa	Chiesa di S. Fereolo	bene individuo		
1A060-001...	torre	Torre Via Misericordia	bene individuo		
BS290-001...	casa	Casa Via Volta	bene individuo		
CO180-00...	campanile	Campanile della Chiesa di S. Eusebio	bene compone...	Chiesa di S. Euseb	
BS280-000...	villa	Villa Sgroj	bene individuo		
CO190-00...	chiesa	Chiesa di S. Paolo	bene compone...	Chiesa di S. Paolo	
MN360-00...	cascina	Corte Renolfa - complesso	bene complesso	Corte Renolfa - co	
1r050-000...	casa	Corpo Sud di Casa Glazel	bene compone...	Casa Glazel - com	
LC120-000...	campanile	Campanile della Chiesa di S. Antonio Abate	bene compone...	Chiesa di S. Antonio Abate - co...	
CO250-00...	villa	Villa Coletti di Parravicino	bene individuo		CO Erba

**STATO DELL'ARTE  
INIZIALE:**→ [SIRBeC](#)

→ dataset Architetture

→ 17.833 record in CC0

**OBIETTIVO:**

→ Inserimento in Wikidata

Okay, tutto  
molto bello...  
ma come?



**OpenRefine**

## Pro

- ❖ Modifiche massive
- ❖ Estensione su Wikidata
- ❖ Integrazione con Quickstatements

## Contro

- ❖ Lavoro collaborativo sincrono
- ❖ Riconciliazioni



# Pro

# Contro

## ❖ Modifiche massive

**Trasformazione di testo personalizzata sulla colonna location**

Espressione Linguaggio General Refine Expression Language (GREL) ▾

```
value.replace("(", "").replace(")", "")
```

Nessun errore sintattico.

**Anteprima** [Storia](#) [Preferiti](#) [Aiuto](#)

2700.	(45.95412091299003, 9.927470034113344)	45.95412091299003, 9.927470034113344
4201.	(45.86310342219098, 8.780412417259734)	45.86310342219098, 8.780412417259734
5044.	(45.16361673078002, 10.790668975239042)	45.16361673078002, 10.790668975239042
5596.	(45.65968379407625, 9.258594342240817)	45.65968379407625, 9.258594342240817
6326.	(45.69858210744174, 10.073700884196363)	45.69858210744174, 10.073700884196363
7022.	(45.53252952351208, 9.228857199127201)	45.53252952351208, 9.228857199127201
7298.	(45.47971231807529, 9.197519898483087)	45.47971231807529, 9.197519898483087

In caso di errore  mantieni l'originale  Ritrasforma fino a  volte o finchè non ci sia nessun cambiamento

imposta a blank

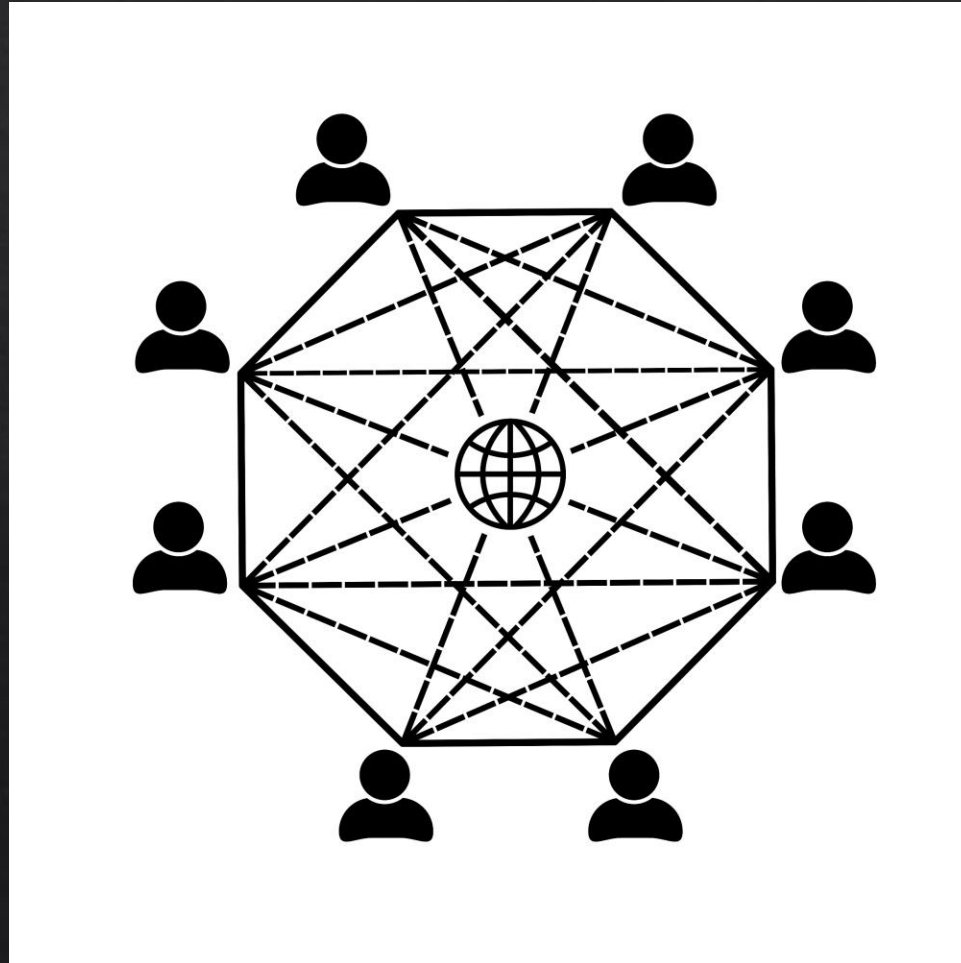
salva l'errore

## Pro

❖ Modifiche massive

## Contro

❖ Lavoro collaborativo sincrono





# Pro

- ❖ Modifiche massive
- ❖ Estensione su Wikidata

# Contro

- ❖ Lavoro collaborativo sincrono

The screenshot shows the Wikidata edit interface for the property `NUM_SCHEDA_SIRBEC`. At the top, a navigation bar lists various properties: `NUM_SCHEDA_SIRBEC`, `scheda sirbec`, `TIPO_ARCHITETTURA`, `DENOMINAZIONE`, `COMPLESSITA_DEL_BENE`, `DENOMINAZIONE_BENE_COMPLESSO`, `NUM_SCHEDA_BENE_COMPLESSO`, `PROVINCIA`, `COMUNE`, `LOCALITA`, `INDIRIZZO`, `UBICAZIONE_NON_VIABILISTICA`, `AUTORI`, `STILE_SCUOLA`, `DATAZIONE`, `STATO_CONSERV`, `USO_ATTUALE`, `USO_STORICO`, `lng`, `lat`, and `location`.

The main content area is titled `NUM_SCHEDA_SIRBEC` and includes a `rimuovi` button. It is divided into three sections:

- Termini**: Contains two entries. The first has the label `Denominazione` and the description `TIPO_ARCHITETTURA`. The second has the label `TIPO_ARCHITETTURA`. Each entry includes a `rimuovi` button and a checkbox for `sovrascrivi, se esistente`.
- Dichiarazioni**: Contains three entries. The first is `identificativo Lombard` with the value `NUM_SCHEDA_SIRBEC`. The second is `Paese` with the value `Italia`. The third is `unità amministrativa` with the value `COMUNE`. Each entry includes a `rimuovi` button, a `+ aggiungi qualificatore` button, a `> 0 note` link, and `+ aggiungi fonte` and `+ aggiungi valore` buttons.

# Pro

- ❖ Modifiche massive
- ❖ Estensione su Wikidata

# Contro

- ❖ Lavoro collaborativo sincrono
- ❖ Riconciliazioni

**Riconcilia la colonna "NUM\_SCHEDA\_SIRBEC"** [Access Service API](#)

Riconcilia ogni cella con una entità tra questi tipi Inoltre, usa informazioni rilevanti da altre colonne

<input type="radio"/> casa Q3947	<input type="checkbox"/> SIRBEC_BACKUP	<input type="text"/>
<input type="radio"/> complesso di edifici Q1497375	<input type="checkbox"/> MANCANTI	<input type="text"/>
<input type="radio"/> chiesa Q16970	<input type="checkbox"/> TIPO_ARCHITETTURA	<input type="text"/>
<input type="radio"/> cascina Q1169748	<input type="checkbox"/> istanza2_architettura	<input type="text"/>
<input type="radio"/> torre Q12518	<input type="checkbox"/> DENOMINAZIONE	<input type="text"/>
<input type="radio"/> torre campanaria Q200334	<input type="checkbox"/> COMPLESSITA_DEL_BENE	<input type="text"/>
<input type="radio"/> basilica minore Q120560	<input type="checkbox"/> DENOMINAZIONE_BENE_COMPLESSO	<input type="text"/>
<input type="radio"/> palazzo Q16560	<input type="checkbox"/> NUM_SCHEDA_BENE_COMPLESSO	<input type="text"/>
	<input type="checkbox"/> PROVINCIA	<input type="text"/>
	<input checked="" type="checkbox"/> COMUNE	<input type="text"/>
	<input type="checkbox"/> LOCALITA	<input type="text"/>
	<input type="checkbox"/> INDIRIZZO	<input type="text"/>

Riconcilia sul tipo

Riconcilia senza specificare un tipo

Matcha automaticamente i candidati con alta confidenza

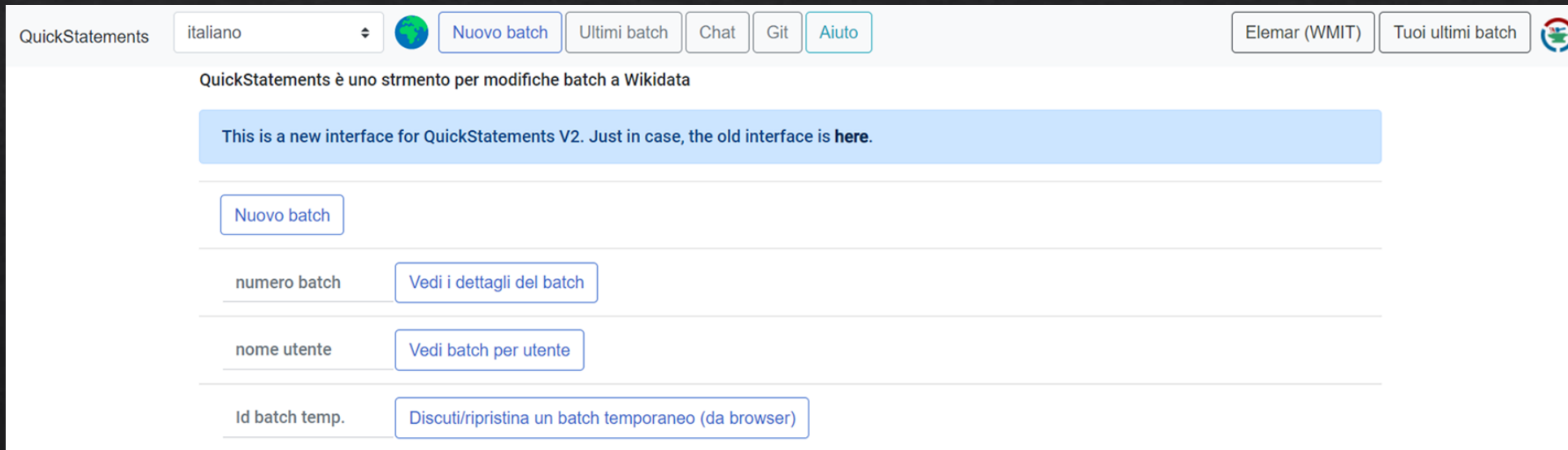
Maximum number of candidates to return [TODO - translate]

# Pro

- ❖ Modifiche massive
- ❖ Estensione su Wikidata
- ❖ Integrazione con Quickstatements

# Contro

- ❖ Lavoro collaborativo sincrono
- ❖ Riconciliazioni



QuickStatements

italiano



Nuovo batch

Ultimi batch

Chat

Git

Aiuto

Elemar (WMIT)

Tuoi ultimi batch



### Batch su Wikidata per Elemar (WMIT) [Batch]

Stato:



0% (0) di 42 fatti

1	init	Pietra de' Giorgi castle [Q93451513]	<b>ADD</b>	Alias	<i>it</i> : "Castello di Pietra de' Giorgi"			
2	init	Pietra de' Giorgi castle [Q93451513]	<b>ADD</b>	Statement	country [P17]: Italy [Q38]			
3	init	Pietra de' Giorgi castle [Q93451513]	<b>ADD</b>	Statement	located in the administrative territorial entity [P131]: Pietra de' Giorgi [Q39880]			
4	init	Pietra de' Giorgi castle [Q93451513]	<b>ADD</b>	Sources to	located in the administrative territorial entity : Pietra de' Giorgi [P131] [Q39880]	stated in [P248]	:SIRBeC [Q111178821]	
5	init	Pietra de' Giorgi castle [Q93451513]	<b>ADD</b>	Statement	street address [P6375]: <i>it</i> : "Via del Castello"			
6	init	Pietra de' Giorgi castle [Q93451513]	<b>ADD</b>	Sources to	street address [P6375]: <i>it</i> : "Via del Castello"	stated in [P248]:SIRBeC [Q111178821]		
7	init	Pietra de' Giorgi castle [Q93451513]	<b>ADD</b>	Statement	instance of [P31]: castle [Q23413]			
8	init	Pietra de' Giorgi castle [Q93451513]	<b>ADD</b>	Sources to	instance of [P31]: castle [Q23413]	stated in [P248]:SIRBeC [Q111178821]		
9	init	Pietra de' Giorgi castle [Q93451513]	<b>ADD</b>	Statement	heritage designation [P1435]: Italian national heritage [Q26971668]			
10	init	Pietra de' Giorgi castle [Q93451513]	<b>ADD</b>	Sources to	heritage designation [P1435]: Italian national heritage [Q26971668]	stated in [P248]	:SIRBeC [Q111178821]	

All  errori  Iniz. 10

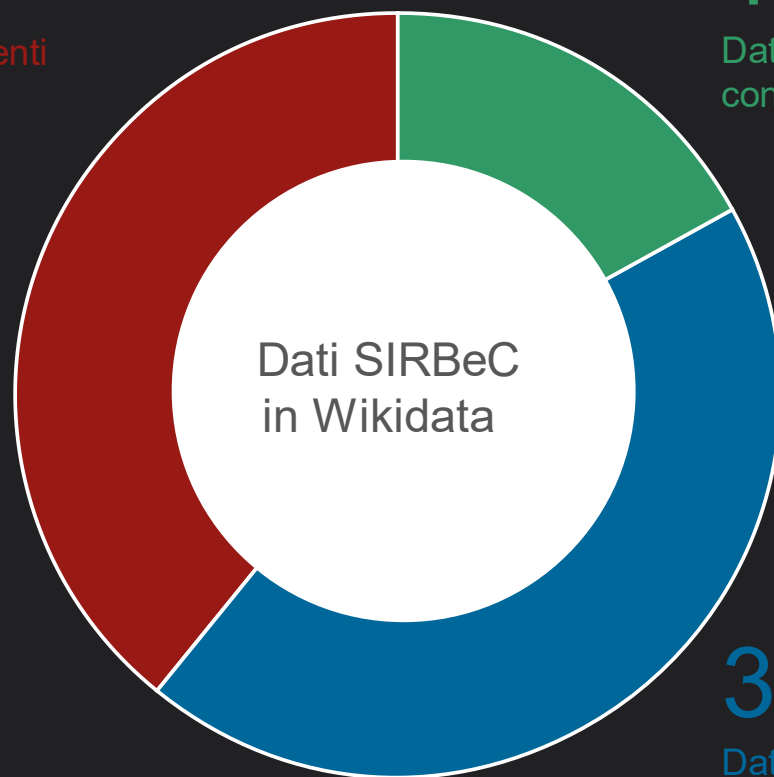
# QuickStatements

44 %

Dati non esistenti  
su Wikidata

17%

Dati su Wikidata  
con SIRBeC id



Dati SIRBeC  
in Wikidata

39 %

Dati su Wikidata  
senza SIRBeC id

# Follow up

## Confronto con la **community**

- violazioni
- comuni soppressi
- monumenti complessi e componenti
- coordinate

## Dialogo con la **Regione Lombardia**

- correzione discrepanze prossime alla pubblicazione
- work in progress per immagini dal portale in CC0



[elena.martellotta@wikimedia.it](mailto:elena.martellotta@wikimedia.it)

[mattiamine00@gmail.com](mailto:mattiamine00@gmail.com)

