



GFS - Global Fact Sync



How is data edited in Wikipedia/Wikidata?

Where does it come from?

And how can we synchronize it globally?

Initial Analysis Results

- Wikipedias contain more than 500 million facts in infoboxes templates (not adopting Wikidata)
- Data in Wikidata and Wikipedia originates externally and is added via:
 - Copying online databases
 - Copy & paste from HTML/PDF
- Study about difficulty of syncing properties:
 - Easy (no change, clear method), e.g. person height, release date
 - Difficult (updated, varying methods), e.g. population count

Methods:

1. Create a tool that captures and presents information about which facts and references are available in the Wikiverse.
2. Include selected external sources to increase available information within the Wikiverse.
3. Adds facts and reference from article text

Introducing the
GlobalFactSync
(GFS) Data Browser



“Human(s)-in-the-loop” synchronization

- Final decision to synchronize a value is made by **an editor who understands consensus and the implications**
- No automatic imports - goal is to drastically reduce the time for editors to research all references for individual facts and **make infobox maintenance faster and more effective**

Contact

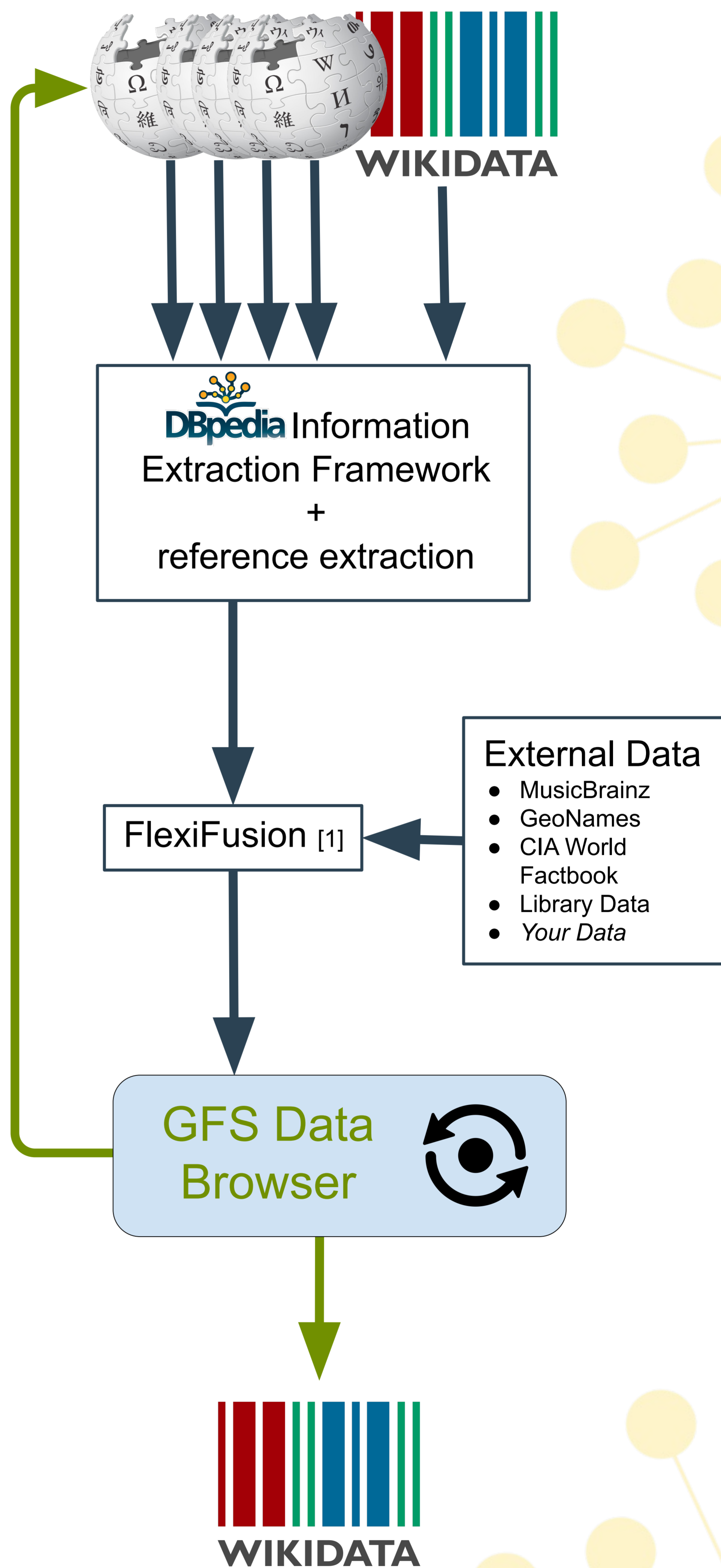
Marvin Hofer

Email: hofer@infai.org

Project email: gfs@infai.org



<https://meta.wikimedia.org/wiki/Grants:Project/DBpedia/GlobalFactSyncRE>



References

[1] Frey, J., Hofer, M., Obraczka, D., Lehmann, J., & Hellmann, S. (2019, October). DBpedia FlexiFusion the Best of Wikipedia> Wikidata> Your Data. In *International Semantic Web Conference* (pp. 96-112). Springer, Cham.