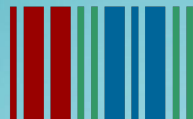


Discussion: Matching new Wikipedia articles to Wikidata items

[[User:Mike_Pee]]

& you!



WIKIDATA
DATA QUALITY
DAYS 2022

Background

- This is a discussion session!
- I have a few slides here to give some background, to make sure we're all on the same page
- Then we'll go into open discussion
- (I might change and then come back to these slides at the end!)
- Etiquette:
 - Please don't interrupt others
 - Raise your hand if you want to speak
 - Feel free to also chat by text alongside voice chat

User:Pi_Bot

- **Matches biographies against Wikidata items**
 - Adopted from Emijrpbob
 - Uses name, year of birth to match items
 - Also imports basic information
- **Creates new items**
 - Follows GZWDer's bot that used to be doing this
 - Delay of 14 days since creation
 - Avoids recently edited articles
 - Does create duplicates that need merging on Wikidata...
 - Opt-in for different wikis: currently running on en, simple, pt, es - add yours here?

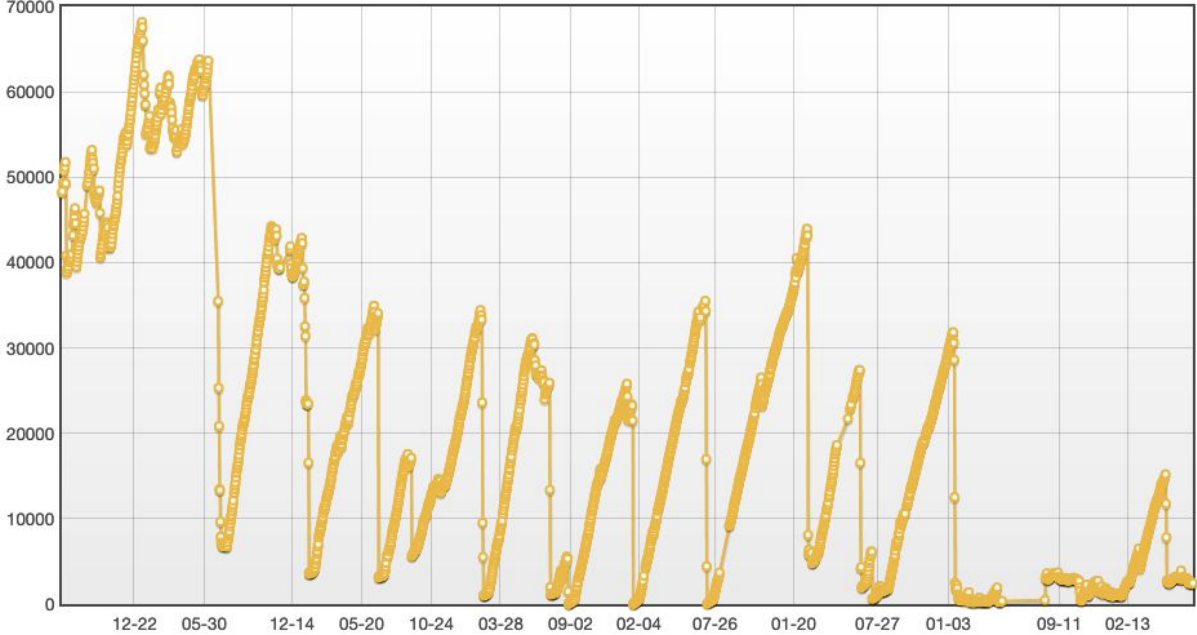
Example: enwiki

Duplicity

Tools Git Talk



Number of articles without Wikidata item on [enwiki](#)



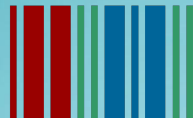
Tools

- Duplicity: <https://wikidata-todo.toolforge.org/duplicity.php>
- ...?

Thoughts for the future / discussion

- **Wikidata Game for matching new articles?**
 - Exists already, but seems broken?
 - Could easily create a new one with Pi bot
 - Would anyone play it?
- **Ways to better integrate with Wikipedia curation?**
 - Technical changes / improvements by WMDE or WMF?
 - Link with Wikidata-powered infoboxes: increase spontaneous linking?
- **Case study of dewiki?**
 - Links articles quickly - how?
- **(Commons - controversial?)**
 - (How do we link the rest of the Commons categories to Wikidata?)

Please discuss!



WIKIDATA
DATA QUALITY
DAYS 2022