Everything you need to become a data quality hero on Wikidata!

WIKIDATA

Why does this matter?

Wikidata is a central piece of infrastructure in today's technology world. Our data is used in the Wikimedia Projects and far beyond to help give more people more access to more knowledge. We want to provide verifiable, reliable and up-to-date information to everyone.
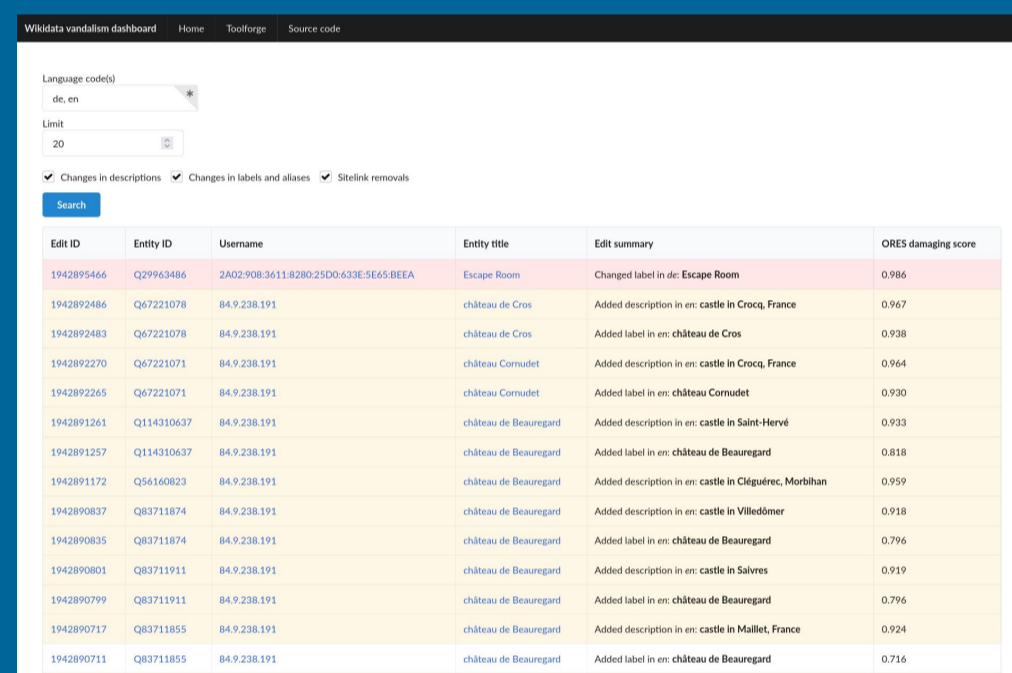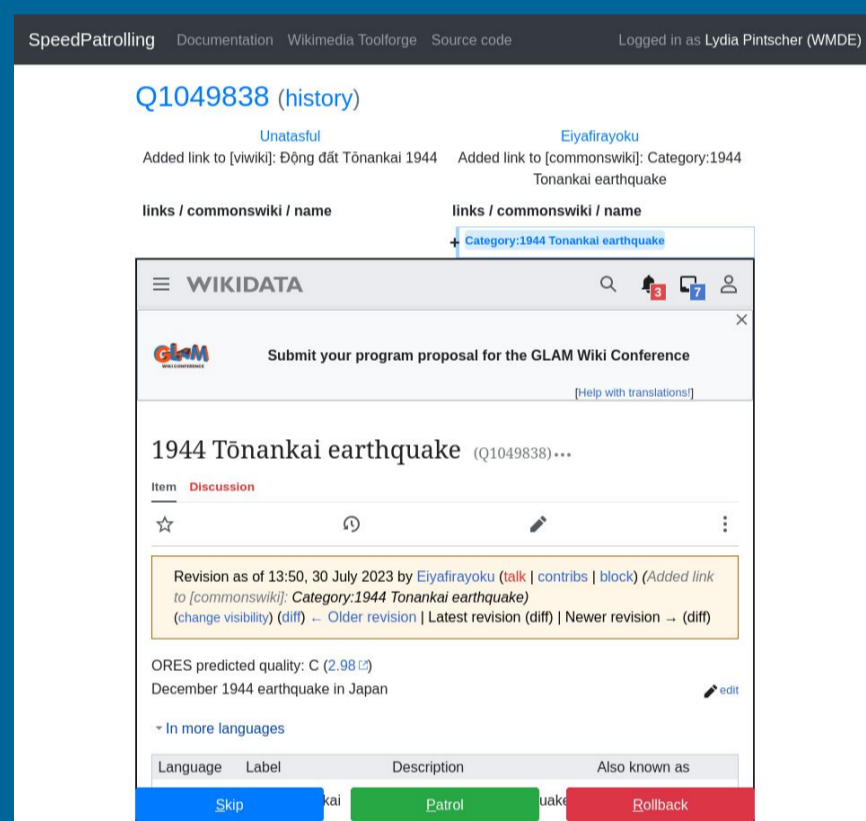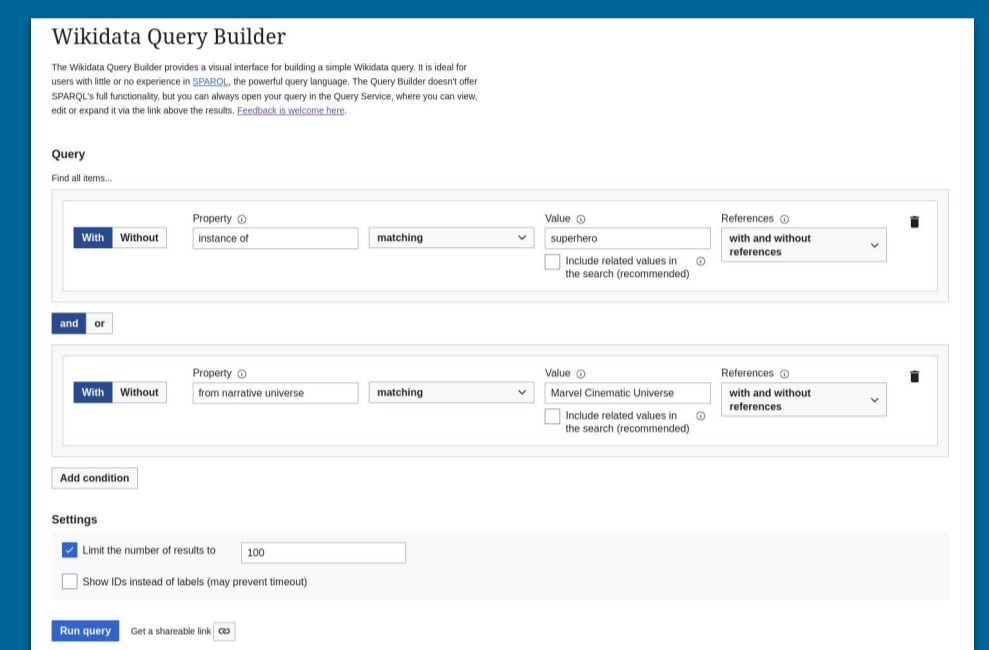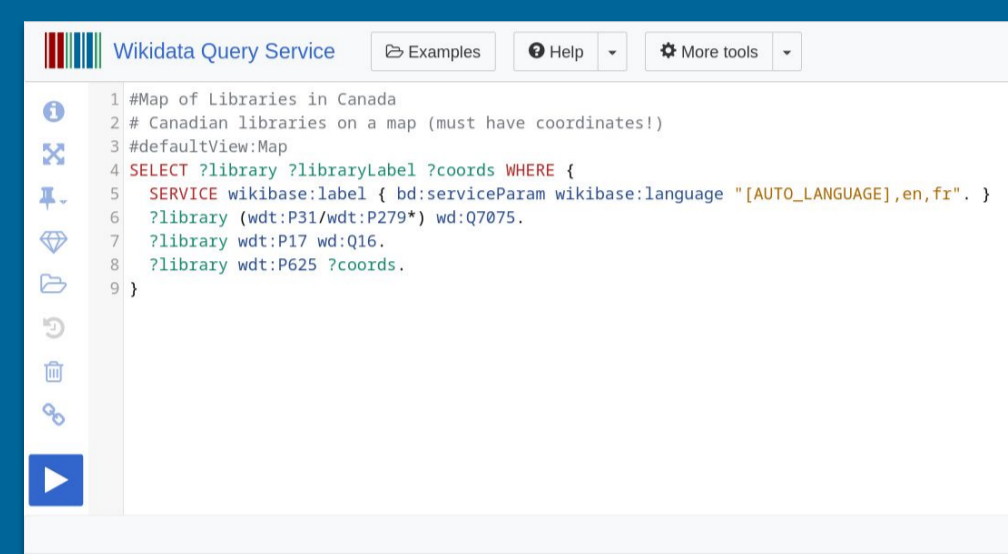
What types of data quality issues are there?

- Vandalism
- Ontology issues
- Unintentionally bad edits
- Missing data (especially references)
- ...

Your secret super power: Queries

query.wikidata.org

Many data quality tools rely on queries to help find wrong, missing, etc. data. You can query directly with SPARQL in the **Wikidata Query Service** (e.g. "politicians born within the last century missing a picture"). Or you can use the **Query Builder** for an easier (but slightly less powerful) way to put together a query.
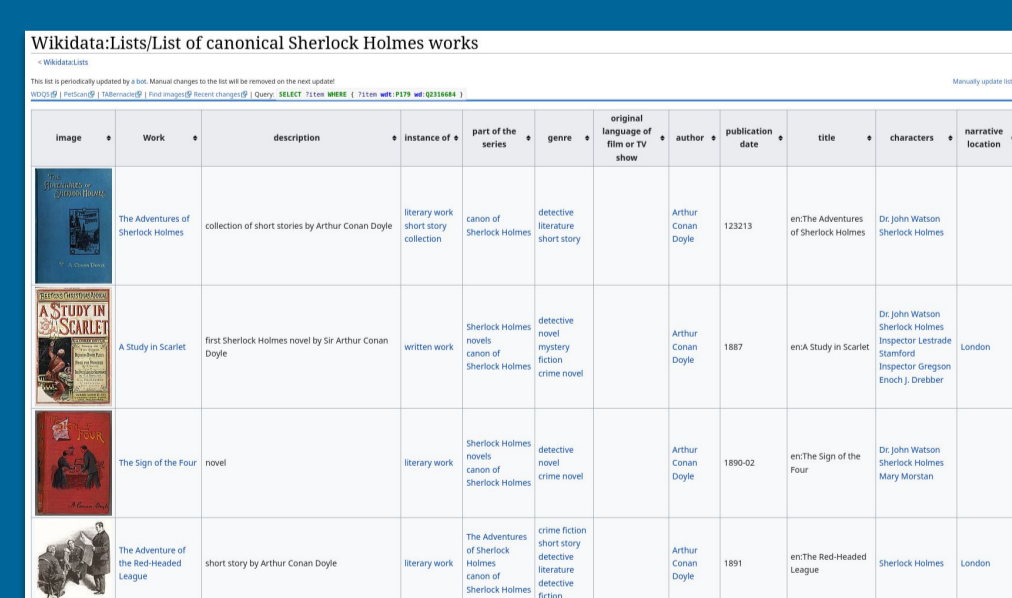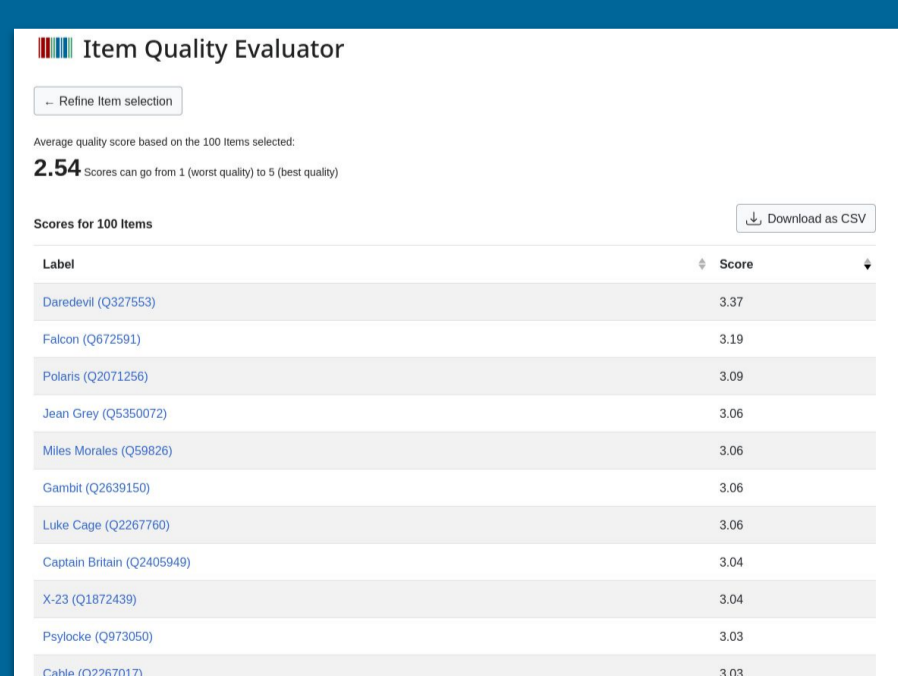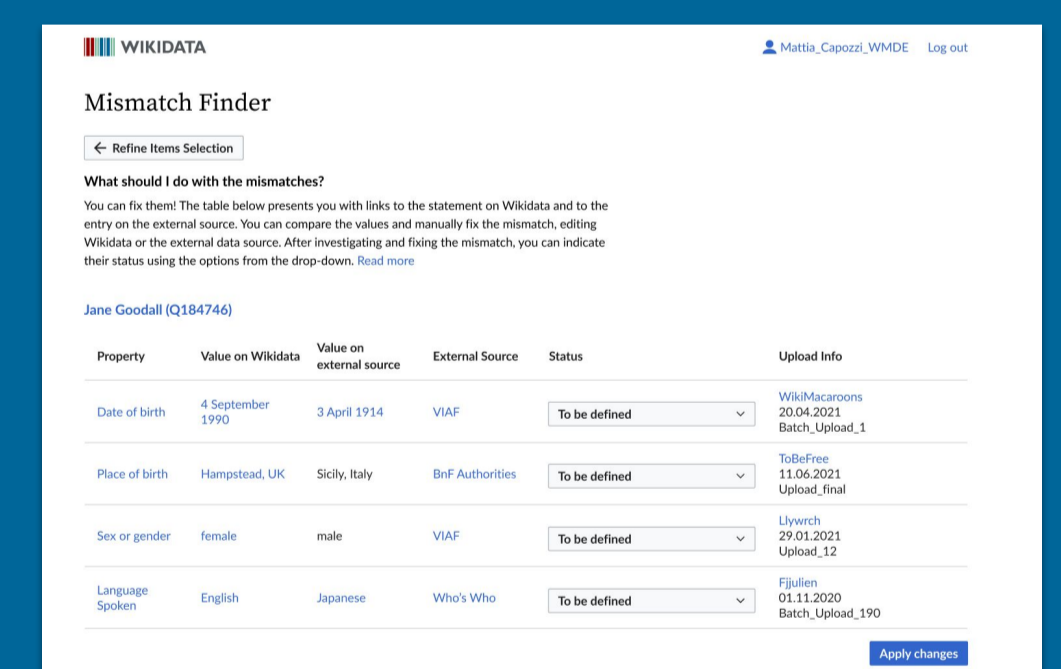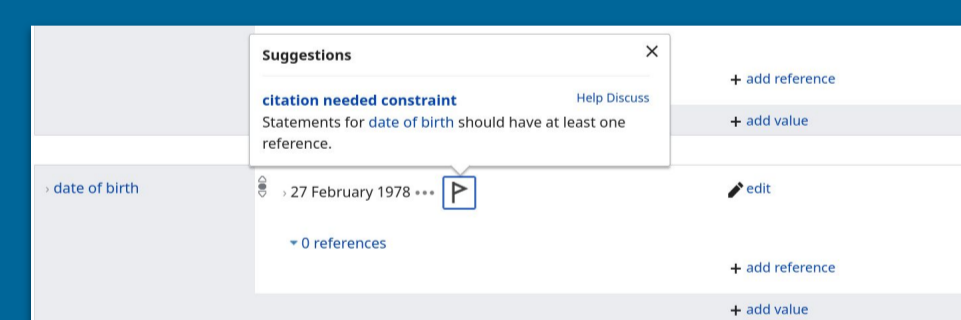
Vandalism fighting

speedpatrolling.toolforge.org   pltools.toolforge.org/rech   wdvd.toolforge.org

- A lot of problematic edits are automatically tagged and you can **filter Recent Changes** by them to quickly find issues to fix.
- **Speed Patrolling** shows you an edit and gives you quick options to mark it as patrolled or revert it. Great for a boring subway ride!
- **reCh** gives you a quick overview of recent edits incl. options to revert problematic ones.
- The **Wikidata Vandalism Dashboard** is especially useful to monitoring edits to labels, descriptions, aliases and sitelinks for a specific language.

Finding ontology issues and other problematic data

Help:Constraints   WD:Schemas   mismatch-finder.toolforge.org

- **Constraint Checks** and **EntitySchemas** are both ways to encode how data should be modeled to make it possible to find statements that do not match the agreed upon modeling. When a Constraint is violated there will be a little warning next to the statement explaining what is wrong.
- The **Mismatch Finder** is there to compare Wikidata's data against other sources. When the data isn't the same you can review it and make changes as necessary.

Monitoring and evaluating

item-quality-evaluator.toolforge.org   integraality.toolforge.org   WD:Listeria

- The **Item Quality Evaluator** gives every Item a quality score. You can create a list of Items to find the lowest-quality ones for improvement. Great for an editathon!
- **Integraality** and **Listeria** are tools to put query results in a wiki page. The wiki page is updated regularly so you can put it on your watchlist and see any changes in the data you care about.

WikiProjects

Here you can find your fellow heroes and coordinate your next mission:
- WD:WikiProject Counter-Vandalism
- WD:WikiProject Data Quality

Contact

Lydia Pintscher
lydia.pintscher@wikimedia.de
@wikidata