Categorization – Referencing Knowledge Resources in Linked Open Data

For this assignment, our group worked on the biography of Joseph Ekwe Bilé (1892–1959), whose life story is featured on the Dekoloniale website. Bilé was a prominent activist, intellectual, and cofounder of the anti-colonial group Liga zur Verteidigung der N###rrasse (League for the Defense of the N*gro Race) in Germany. His biography offers insights into anti-colonial resistance, diasporic organization, and the lived experience of Black political life in early twentieth-century Europe.

The aim of the task was to transform this narrative biography into structured, linked open data on Wikidata, thereby contributing to the visibility of marginalized figures in digital knowledge systems. By translating narrative accounts into interoperable data points, our work sought to expose both the epistemic possibilities and limitations embedded in linked data infrastructures (Kitchin, 2022). This aligns with Bowker and Star's argument that digital classification systems both organize and restrict what can be known (Bowker & Star, 2000).

Our group consisted of three members who collaboratively worked on identifying and categorizing data from the Dekoloniale biography. My main role in the project was as one of the categorizers. I focused on extracting details about Bilé's life, such as his political affiliations, occupations, and relationships, and converting this information into data points that could be appropriately represented in Wikidata. This required interpreting the narrative and mapping it within Wikidata's ontology using properties such as "member of," "occupation," "educated at," or "place of residence.". Such interpretive labour reflects how datafication relies on human judgment rather than neutral extraction (Ford & Iliadis, 2023).

I also helped interpret ambiguous data points within Wikidata's ontology. Together, our group crosschecked each data point to ensure consistency and accuracy.

We began by reading the biography aloud as a group, pausing to note any factual statements that could be expressed as structured data. This collaborative reading helped us collectively interpret the text and ensured that no relevant detail was overlooked.

Next, we compared Bilé's existing Wikidata profile with the Dekoloniale biography and found only minimal information: basic dates, nationality, and a few identifiers. Our aim was therefore to supplement it with missing affiliations, locations, and family connections.

We then created thematic categories to organize the extracted data, biographical, professional, social-political, and relational, which helped structure our interpretation (Bowker & Star, 2000). Working through these categories revealed new insights, such as Bilé's connections to other activists and his transnational political involvement. As Bowker and Star note, every act of classification both reveals and conceals aspects of the world; it makes certain narratives legible while obscuring others.

We structured the data using Wikidata properties, linking each point to existing items. Rather than a plain textual note on WWI, we linked him to the corresponding military unit, ensuring interoperability across the linked open data network.

One of the main technical challenges was understanding Wikidata's ontology and selecting the most accurate properties for each relationship. Linking Bilé to family members required creating or identifying those individuals in the database, which involved additional data entry. Similarly, choosing between properties such as "affiliation" and "member of" required careful interpretation of the biography's wording.

We also faced ethical challenges in curating information about a historical figure we did not personally know. Some details, especially regarding his family, were not explicitly stated in the biography. We therefore discussed how to avoid speculation or the risk of misrepresenting Bilé's life. As a result, we adopted a principle of strict verifiability, only including information that could be directly supported by the Dekoloniale text.

These considerations raised important ethical reflections on representation: transforming someone's life into structured data is not a neutral act (Kitchin, 2022). It involves decisions about what counts as relevant information and what remains invisible in the data model. This mirrored discussions in our first assignment, where curatorial choices similarly shaped what the dataset could reveal.

Working on this project revealed that data modeling is not only a technical process but also a form of knowledge production (Ford & Iliadis), 2023, echoing Ford and Iliadis' description of Wikidata as a "more-than-technical project" that depends on social negotiation, collective labor, and interpretive judgment. Our attempts to translate Bilé's life into predefined categories highlighted precisely this tension between narrative richness and infrastructural standardization.

This tension aligns with the course readings on categorization and curating as knowledge production, particularly Bowker and Star's argument that classification systems shape what we come to know about the world. This demonstrates how digital infrastructures do not merely store historical traces but actively shape their legibility and significance. In our case, deciding how to categorize Bilé's activism, whether as occupation, affiliation, or political role, revealed how infrastructures can subtly influence historical legibility.

A concrete example from our work illustrates this more clearly: choosing whether to represent Bilé's political activities as an occupation, an affiliation, or a role revealed how different categorizations foreground different aspects of his historical identity. Such choices directly shape what becomes visible within Wikidata's knowledge structure and what remains backgrounded or obscured.

The process also revealed the limitations of Linked Open Data: while it increases accessibility and connectivity, it can flatten complex human experiences into fixed properties (Bowker & Star, 2000; Ford & Iliadis, 2023). Nevertheless, it provides an important framework for ensuring that marginalized histories, like that of Joseph Ekwe Bilé, are represented and discoverable within global knowledge networks.

Working with Linked Open Data involves both technical precision and ethical responsibility. Making marginalized histories visible requires careful attention to how data is categorized, verified, and contextualized. This aligns with broader discussions in critical data studies about the responsibilities embedded in curatorial practices.

By contributing Bilé's story to Wikidata, we not only made his life more accessible but also participated in a larger process of decolonizing digital archives. This project deepened my understanding of how knowledge infrastructures can both empower and constrain representation, and how critical reflection is essential in curating data responsibly.

Taken together, these reflections underscore that data is never neutral but shaped by the choices, perspectives, and values embedded in its production (Bowker & Star, 2000; Ford & Iliadis, 2023; Kitchin, 2022). Categorizing Joseph Bilé's biography therefore became an instance of how digital documentation participates in historical interpretation and responsibility.

CHARACTERS: 7.303

References:

Bowker, G. C., & Star, S. L. (2000). Sorting things out: Classification and its consequences. MIT Press.

Dekoloniale Project. (n.d.). Joseph Ekwe Bilé (1892–1959). Retrieved January 2025, from

https://www.dekoloniale.de/en/map/stories/joseph-ekwe-bil%C3%A9-1892-1959?view=gallery

Ford, H., & Iliadis, A. (2023). Wikidata as semantic infrastructure: Knowledge representation, data labor, and truth in a more-than-technical project. Social Media + Society, 9(3). https://doi.org/10.1177/20563051231195552

Kitchin, R. (2022). Critical data studies. In The data revolution: Big data, open data, data infrastructures and their consequences (pp. 21–41). SAGE Publications.

Wikidata. (n.d.). Joseph Bilé (Q105336549). https://www.wikidata.org/wiki/Q105336549