Linked Data and Holocaust Era Art Markets: Gaps and Dysfunctions in the Knowledge Supply Chain

Laurel Zuckerman[1]

[1] Independent Researcher, Bry sur Marne 94360, FRANCE
Laurel.zuckerman@gmail.com

Abstract. Without robust quality control, insights gained by exploiting linked data will be incomplete, unreliable and misleading. Much of what passes for quality control, however, is not designed to verify complete end-to-end real world processes but only certain isolated steps, leaving gaps for errors which too often are not only tolerated but considered "normal". This paper presents the alarming results of a series of entity reconciliation tests in the domain of Nazi looted art. Only 6% of plundered Jewish collectors were correctly matched to LOD entities. The causes were multiple: missing or confusing Wikimedia content, LOD or authority files, dysfunctions across languages, redirects, problematic labelling, and a bias towards fictional characters over real people. It is not clear who, if anyone, is responsible for verifying and correcting these complex errors stemming from the intersection of multiple jurisdictions, raising the question of governance, in particular where marginalized communities are concerned. Conceived as a first step in an ongoing, iterative process, the paper concludes with some suggestions on how to monitor and improve the usefulness and reliability of linked data, as well as the urgency of doing so.

Keywords: linked data, NER, LOD, best practices, LOD quality, art market, Holocaust, Nazi looted art, erasure, marginalized people, Jewish art collectors

1 Introduction

Linked open data as imagined by Tim Berners-Lee in his seminal paper is a decentralized network knitted together by a common respect of certain conventions.[1] However, if linked data is to fulfill its considerable promise, it needs, as Verborgh and Sande so convincingly argued, to deal with "trivial" problems [2], such as providing the end user with 100% accurate and complete information. To this end, identifying and fixing errors in the overall chain of information is essential. A potentially helpful approach might be end-to-end quality control of an industrial supply chain, with quality monitoring and continuous improvement by domain. This is a particular challenge for crowdsourced information that has no "owner". The temptation to dismiss specific pockets of poor quality as unimportant because statistically small is a dangerous one, as the impact of even a tiny percentage of false information in a specific domain can have such an outsized impact on trust that it casts a cloud of suspicion over the whole.[3] The Holocaust is one such domain where accurate and complete information is considered so important to the well-being of society that laws have been enacted criminalizing the dissemination of certain false information.[4] What a pity it would be then, if the semantic web were, through sloppiness, lack of quality control and diffuse dysfunction, to contribute not to greater knowledge but to greater ignorance.

This paper looks at the small but high stakes domain of linked data for Nazi looted art. The objective is to use successive sets of increasingly targeted tests to identify errors in the linked data knowledge supply chain and to trace them back to their origins so that they may be corrected. The term supply chain, borrowed from industry, is employed to emphasize the focus on real world end-to-end processes as opposed to a theoretical testing environment, as well as a concern for the challenges of quality control, error detection and correction across multiple jurisdictions, involving numerous actors, technologies, and procedures. [5]

2 Method

Named-entity linking (NEL) tests were executed from April to June 2019 and in September 2020 using public data from two sources: “lootedart.com/news”[6], a site specialized in Holocaust –related art cases, and the German Wikipedia *Liste von Restitutionsfällen*, which lists Holocaust-related art restitution cases involving 120 spoliated Jewish collectors.[7]

IBM Bluemix NLP [8] and TextRazor [9], selected from a much larger pool of NEL tools for ease of use, replication and interpretability, were not optimized for the tests; they were untrained, unenhanced with no additional specific domain information. The purpose was not to compare the NLP/NEL systems but rather to see how the underlying linked data resources used in default mode did or did not contribute to providing identification and context and why.

Error verification was manual, involving multiple searches to check whether all entities had been identified and correctly matched and, if not, what had gone wrong. Results from each round of tests determined the research questions and the datasets for the next, narrowing successively like a funnel.

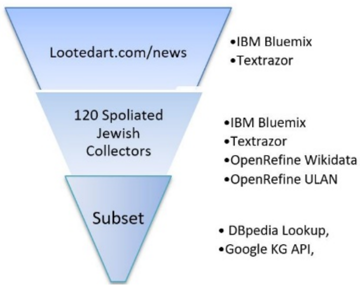

Fig. 1.

3 Entity Linking Tests

3.1 Lootedart.com news articles

Lootedart.com/news is published by the Central Registry of Information on Looted Cultural Property 1933-45. Created in 2001 to provide a "central repository of information on Nazi looting and contemporary efforts to research and resolve all outstanding issues", the Central Registry website aggregates news about Holocaust-related art cases on lootedart.com/news. Lootedart.com/news articles have four characteristics that deserve special mention in the context of these entity recognition and linking tests. First, though the news articles themselves are recent, they contain references to people, organizations, places, and events situated in the past. Second the deceased individuals mentioned in the articles tend to appear in a variety of other sources, including other news articles, books, catalogues, legal cases, scholarly documents, museum websites, government archives and authority files. Third, the lootedart.com/news articles concern the Holocaust and Holocaust research, which European authorities have declared exempt from both GDPR and Right to Be Forgotten.[10] And fourth, while lootedart.com is an English language site, it aggregates news from other languages, which means the people mentioned may be considered notable in languages other than English.

Method: News articles from lootedart.com/news were pasted into the demos for IBM BlueNLP- entities and Textrazor entities, and the output was analysed to see what had been identified, how it was classified ("persons", "organizations", "places", etc) and which LOD entities it had been reconciled to. IBM Bluemix linked to DBpedia while Textrazor linked to Wikipedia, Freebase and Wikidata.

For each entity in the news article, the following questions were asked: Was the entity recognized? Was it correctly classified as person/organization/place/event/etc? Was it linked to DBpedia (Bluemix; DBpedia Spotlight), or Wikipedia/Wikidata/Freebase (TextRazor) or anything else? Was this matching correct? If the LOD match was incorrect, which wrong entity was returned? What was the likely cause of the error? Are there errors for which the cause cannot be identified?

Could any patterns be detected? What (and where) might be the fix?

Results: Missing entities and mismatches for Jewish art collectors

Nearly all visual artists, media outlets, museums, legal terms (including arcane vocabularies) and most Holocaust-related institutions mentioned in the news articles were successfully recognized and reconciled by both Bluemix and Textrazor. However, results were very poor for art collectors and dealers mentioned in the same articles. Textrazor, which links entities not only to Wikipedia but also to Wikidata, did better than Bluemix which linked only to DBpedia, but both tended to omit the persecuted Jewish collectors despite the hundreds of news article published about them.

In several cases, the names of persecuted Jewish art collectors were incorrectly classified as "places", even though the first and last names were in the text.

Books and films about the Holocaust victims or claimants, and the actors and directors involved in created them tended to be successfully recognized and linked to Dbpedia (Bluemix) or Wikipedia/Freebase/Wikidata (Textrazor) at high rates, while in the very same texts, the people whose stories are dramatized tend NOT to be recognized in DBpedia (Bluemix) with some links to Wikipedia (TextRazor: Maria Altman).

Questions raised by the differences in entity linking and selection of a data source for a second round of entity linking tests

Why did the entity linking software fail to find so many persecuted Jewish art collector and art dealers? Even when Wikipedia, Wikidata or Viaf entities exist? Why are published results for entity linking so much better than the results observed here? Is it a question of the texts used or the method or something else?

Differences in spelling (ü vs ue; ö vs oe), languages other than English, as well as missing, badly created or confusing authority files or Wikimedia entrees coincided with many of the dysfunctions observed. (Emil Georg Bührle was linked to DBpedia and Wikipedia, but Emil Buehrle was not; Wikpedia labels that included parentheses performed worse than labels without parentheses). The failure of entity linking for entities that existed in Wikipedia or Wikidata was of particular concern.

The verification had to be manual because it required detailed domain knowledge backed up by extensive research – which poses the question: on what is automated quality evaluation of entity linking based? [11]

NEL results for persecuted Jewish collectors in the news articles were so poor (Bluemix made *more erroneous links than correct links* to persons involved in Nazi art looting) and possible causes so numerous that the lootedart.news tests were cut short and testing shifted to a targeted subset representing these entities to try to identify the sources of error.

3.2 *Liste von Restitutionsfällen*: Named-Entity Linking for spoliated art collectors

The German Wikipedia page *Liste Von Restitutionsfällen* was initially created on March 23, 2009 and, at the time of testing, had been updated by 60 separate users, most recently on April 27, 2019. A corresponding page exists in the French Wikipedia but not in the English language Wikipedia.

The data used for the entity linking tests were drawn from column three of the Wikipedia list, which contains, in German, both the name of the Jewish collector and the name of the institution which received the claim. There are 223 rows with 120 names. Some names appear more than once because the families have filed multiple claims for different works of art. The assumption is that these 120 names reflect a crowdsourced consensus of German Wikipedia contributors concerning the most notable Jewish art collectors who were victims of Nazi persecution. There are, in addition, 156 footnotes, most of which identify news articles or books that mention the Jewish collectors and their cases.

All 120 names had been the subject of numerous news articles, in the *New York Times, BBC, NPR, Wall Street Journal, Guardian, Spiegel, Süddeutsche, Le Monde, Le Figaro*, and other contributing publications. Some had Wikipedia entries in several languages, some only in one language, some no Wikipedia page at all. More than half were referenced in both Viaf and Wikidata.

Method: The names of 120 Jewish collectors were copied and pasted, as one list, into the demo for the entity linking tools IBM Bluemix and Textrazor or via the API. The results on screen and in JSON were analysed, using Google searches and authority files to verify whether each match was correct.[1]

Results: IBM Bluemix failed to link 94% of the spoliated art collectors

The results confirm the observations made in the initial round of tests on lootedart.com/news items. Less than 6% of the names were correctly matched to DBpedia. Of 120 Jewish collectors mentioned in the German Wikipedia list of Restitution cases, only seven were linked to the correct person in DBpedia by Bluemix. These were Jacques Goudstikker, Alfred Hess, Max Silberberg, Sophie Lissitzkuy-Küppers, Richard Semmel, Max Emdem, Eduard Einschlag, and Leo Bendel. Due to technical problems in the German DBpedia, however, none of the correct DBpedia links actually functioned; the identity was verified by consulting the corresponding German Wikipedia pages.

Bluemix failed to match the other 113 Jewish collectors to DBpedia. Additional tests and verifications were needed to understand why.

In three cases – for Paul Stern, Arthur Goldschmidt, and Arthur Feldmann – Bluemix found a DBpedia entity, but it was for the wrong person. (The correct person had no Wikipedia page.)

Table 1. Results *Table BlueMix Dbpedia entity linking: spoliated collectors*

Quantity	IBM BLUEMIX NEL results for the 120 spoliated art collectors in the German Wikipedia *Liste von Restitutionsfällen*	%
7	Correctly linked to Dbpedia entities	**5,83%**
3	Incorrectly linked to wrong Dbpedia entities	**2,50%**
113	**NOT correctly linked by IBM Bluemix to DBpedia**	**94,17%**
data	https://de.wikipedia.org/wiki/Liste_von_Restitutionsf%C3%A4llen	
IBM Bluemix NLP	https://natural-language-understanding-demo.ng.bluemix.net/	

[1] https://docs.google.com/spreadsheets/d/1ua3lv2KljSEipg1oauS9bEEDYY4AP8OcnMCtrTxZqSw/edit?ts=5f649400#gid=0

Questions raised: What was the cause of the 94% Named-Entity Linking failure in Bluemix?

Was the problem due to a glitch in disambiguation? A confusion between several similar-looking entities?

Was the problem due to language? Issues with spelling differences had been noted in the first round of tests. What else might be at work?

Was the problem due to a technical issue, either with the NPL product used or with the entity knowledge base?

Was the entity knowledge base outdated, so that information about the names, while present live, were not in the old download used?

Was there a problem due to recent data protection laws?

Did the Google Knowledge Graph identify them?

Bluemix reconciles entities to DBpedia. Was there a specific issue with DBpedia for these datasets?

3.3 Targeted tests of components to pinpoint defects

There are many places for failure in the long and complex supply chain that underlies entity linking. Bluemix and Textrazor, like many NLP tools, rely on linked data assembled by DBpedia and the Wikimedia foundations, as well as many other components.

Method: To try to diagnose the problem – to see where the "break" had occurred, a new round of small, targeted tests were performed using four tools that frequently play an important role in the knowledge supply chains of cultural heritage and news organizations.

- OpenRefine: Wikidata [12],
- OpenRefine ULAN [13],
- DBpedia Lookup[2][14]
- Google Knowledge Graph API [15],

OpenRefine is widely used by cultural heritage institutions for data cleaning and entity linking, notably with Wikidata and more recently, the Union List of Artists Names (ULAN). TMS, the leading museum management software in the USA, offers integration with ULAN for entity lookup for collections management. [16]

Tests focused on identifying patterns in the errors of the 120 names and identifying a small, representative subset that could be used for quality testing and monitoring.

[2] DBpedia graciously offered assistance in understanding how datasets are indexed and filtered.

Table 2. Summary of Errors, Causes, Possible Solutions[3] [17]

<table>
<tr><th>Defect observed</th><th>Possible cause</th><th>Where in process?</th><th>Who can fix?</th><th>Proposed solution</th></tr>
<tr><td rowspan="5">1. Wrong entity reconciled</td><td>Homonym AND no Wikipedia for spoliated collector</td><td>content</td><td>content - non tech</td><td>Create Wikipedia in English for spoliated collector</td></tr>
<tr><td>Homonym AND Wikipedia in language other than English</td><td>content</td><td>content - non tech</td><td>Create Wikipedia in English for spoliated collector</td></tr>
<tr><td>Homonym in Wikidata,(Openrefine)</td><td>content</td><td>content, non tech</td><td>improve description labels, add birth, death</td></tr>
<tr><td>Redirect to incorrect name</td><td>content</td><td>content</td><td>Delete redirect</td></tr>
<tr><td>Missing, badly coded or misleading ULAN</td><td>early-content</td><td>Content Getty</td><td>Create missing entities in ULAN, correct type errors</td></tr>
<tr><td rowspan="4">2.Failure to recognize entity in text</td><td>spelling error in content</td><td>late - NLP</td><td>tech or content</td><td rowspan="4">improve NLP NEL tech</td></tr>
<tr><td>long, multi-part name (Dr., Baron, Prince, von, de,)</td><td>late - NLP</td><td rowspan="3">tech</td></tr>
<tr><td>special characters like ü or ë</td><td>late - NLP</td></tr>
<tr><td>confusing female name "Alice and John Dupont"</td><td>late - NLP</td></tr>
<tr><td rowspan="9">3. Entity recognized but failure to reconcile to LOD Dbpedia or Wikidata</td><td>No Wikipedia in any language</td><td>early - content</td><td>content, non tech</td><td>create Wikipedia entity in English</td></tr>
<tr><td>Old datafile used by NEL tool, omits newer entities</td><td>selection of dataset</td><td>tech-procedure</td><td>Show dates and filters of datafile used by NEL tool</td></tr>
<tr><td rowspan="2">Wikipedia in German but not in English</td><td>late - NLP</td><td>tech</td><td>Index and use all languages, in NLP tools (DBpedia)</td></tr>
<tr><td>early - content</td><td>content, non tech</td><td>If no tech fix, create in English Wikipedia</td></tr>
<tr><td>Labels with ()</td><td>NLP</td><td>tech</td><td>Verify, improve NEL</td></tr>
<tr><td rowspan="2">Wikipedia exists in English but different spelling</td><td>late - NLP</td><td>tech</td><td>improve NLP NEL tech</td></tr>
<tr><td>Early, content</td><td>content, non tech</td><td>content redirect in Wikipedia</td></tr>
<tr><td>Homonym creates confusion, prevents selection of entity</td><td>late - NLP</td><td>tech</td><td>NLP improvement, using additional info about entity</td></tr>
<tr><td>Wikidata exists but has been filtered out of LOD dataset</td><td>middle, late - NLP</td><td>tech, procedure</td><td>Don't filter Wikidata. Better integrate with NLP NEL tools</td></tr>
</table>

[3] See results https://docs.google.com/spreadsheets/d/e/2PACX-1vSexB-fi-DOrW4Ce8I6EVgFMfHNCQYnDflJDtVvXh0sxCgt4mIZx7t7cLaKBOgxOd0jaHdggjh_lJm62/pubhtml

Discussion: The table above lists dysfunctions observed. It is a mix of technical, user content and integration problems in different parts of the process for the specific domain of spoliated collectors. Experience in complex logistics supply chains suggests that these issues can be resolved and quality improved in this domain by applying an end-to-end approach involving all the actors in the process.[18]

Legacy choices concerning which language an entity was created in (content) and how it was indexed or filtered by the LOD data aggregators like DBpedia used by NEL tools (tech) had a major impact on end results. Entities created in the German Wikipedia but not in the English Wikipedia were not picked up by DBpedia Lookup, or Google KG. The spoliated collector Walter Westfeld, for example, is referenced in the German Wikipedia, Wikidata, VIAF, GND and ISNI, as well as hundreds if not thousands of newspaper articles, books and scholarly papers, but Westfeld was still invisible to IBM Bluemix, Textrazor, Google Knowledge Graph, and ULAN OpenRefine reconciliation.

More than half (68) of the 120 spoliated collectors (including Walter Westfeld) were successfully matched to Wikidata entities using OpenRefine reconciliation suggestions. However these links were not picked up in earlier tests with Textrazor which performed entity linking using Wikidata, which suggests that Textrazor's NEL demo may be using a smaller or less recent subset of Wikidata.

OpenRefine ULAN linking was successful for only nine out of 120 names, less than 10%. Manual verification confirmed that names of the spoliated collectors simply did not exist in ULAN. This has major implications both cultural heritage institutions that perform data cleaning and reconciliation using OpenRefine, as well as users of the TMS collections management system and Europeana, both of which use ULAN for entity linking. [19]

Including languages other than English in key DBpedia indexes and extractions could solve a significant portion of the problems. Performance may be improved by translating information contained in foreign languages into English.[20] Alternatively, a content based approach could ensure that all Wikipedia pages exist also in English – perhaps in the framework of a Wikipedia project like Women Scientists.[21]

4 Lessons Learned and Recommendations

On the cumulative impact of English-language bias in a multilingual knowledge supply chain for linked data

Concern about multilingual entity linking is not new [22]; however it would appear that the cumulative effect of a bias favoring English in all stages of the knowledge supply chain may be underappreciated.

At every stage – the creation of Wikipedia content [23], data cleaning[4], data selection, filtering, indexing, NEL development and product marketing – English is treated differently from other languages, which can find themselves excluded from datasets. Even inclusive, multilingual knowledge platforms like Wikidata can be subject to retroactive filtering according to criteria that favor English, such as existence in Wikipedia or number of links.[24]

Each small action, taken independently and with no malicious intent, results in a winnowing of knowledge stored in languages other than English. The impact of such an exclusionary process becomes problematic in dealing with knowledge, like that related to the Holocaust, which is stored in languages other than English, because it occurred in places where people spoke languages other than English.

Knowledge is specific to a domain, and NEL testing practices ignore this fact

Global recall scores are meaningless when it comes to assessing the quality of entity linking in a specific domain. The mass of sports and celebrity and artist-related information, which can be correctly linked, obscures, when aggregated, a devastating failure to recognize and link entities in a specialized domain like art collectors spoliated during the Holocaust.

The Essential Role of Human Domain Experts

It should be emphasized that a *link does not mean a correct link*. Human domain experts are needed to identify such errors.[25] The question of how data quality is evaluated is also posed, as no automated system can catch these errors. There is arguably no “acceptable” fail rate given the sensitivity of this data. There is a growing awareness of the importance of data quality in linked data [26], too often quality is thought of as an isolated content management issue not as a supply chain challenge.

[4] How many developers or data wranglers think it is “normal” to remove foreign characters as part of “cleaning”?

4.1 Recommendations

Open the Black Box (and know the contents)

NLP NEL tools are often plugged in, as a kind of blackbox component, by cultural heritage institutions, news organizations, developers and scholars. Yet the completeness and accuracy of NEL depends on filters, selections and rules made in the laboratories of these components which neither developers nor end users know nor understand. In the case of spoliated art collectors, many NEL failures were traced back to a language filtering choice made by DBpedia: some products index only English with the result that entities created only in German are not linked.

Awareness about exactly what is in the black boxes (dates of dataset, filters applied, operations) and how this impacts end-to-end quality in NEL should be encouraged as it makes it easier to avoid problems and find solutions.

For DBpedia Lookup, for example, a solution could be to include German and other key languages in the index (tech solution at the DBpedia level); or users could create the missing entities in the English Wikipedia (user content solution). Another option would be for developers of NEL tools to shift to the multilingual Wikidata.

Domain specific *Data Quality Dashboard*

Quality Monitoring should be regular because even names which have been successfully reconciled in past tests can suddenly fail to reconcile due to the creation of homonyms, duplicate records or other changes. Likewise, every element in this knowledge chain is constantly evolving - the content, the tech, the procedures, and the tools used to test; it is essential to verify that they continue to work together correctly and to spot quickly any new defects in the process. This suggests that NEL should be tested and validated: 1) by specific micro-domain, 2) by end users who know and care about the accuracy and completeness of the information being transmitted, 3) at regular intervals, 4) with a regular test set, and 5) with results published on a public Data Quality Dashboard.

Domain specific Test Datasets

A standard domain specific test set, like that for spoliated collectors, makes results easy to replicate and interpret, facilitating error correction. Test sets could include:

- Names previously linked correctly
- Names previously linked incorrectly or not linked at all
- Information about the status of these names in Wikipedia, Wikidata, Viaf etc
- Texts (news article, provenances, legal documents...) that are known to contain these names.[5]

[5] Due to link rot it is recommended to retrieve the text as well as an internet archive link

5 Concluding Remarks

The entity reconciliation results in this paper point to serious defects in linked data in the domain of looted art. If NEL fails on the best documented names- publicly known spoliated Jewish collectors drawn from a crowdsourced Wikipedia page - what are the implications for the rest of the spoliated Jewish collectors in linked data?

They, like so many marginalized populations, will be invisible in linked data, as if they had never existed at all.

It is at the intersections of content, process and tech, over multiple jurisdictions and long periods of time, that the battle for end-to-end quality is won or lost.

A solid methodology for monitoring LOD erasure will make it easier to identify defects and take corrective action, in a virtuous cycle that continuously improves NEL results and reduces the number of errors that need correction.

The methods explored here can be applied to any domain, in particular those that concern marginalized populations documented in languages other than English.

References

[1] CERN, Tim Berners-Lee's proposal, http://info.cern.ch/Proposal.html

[2] Verborgh, R. and Vander Sande, M. (2020), The Semantic Web identity crisis: in search of the trivialities that never were, Semantic Web Journal, IOS Press, Vol. 11 No. 1, pp. 19–27. https://ruben.verborgh.org/articles/the-semantic-web-identity-crisis/

[3] Färber, Michael et all, Linked Data Quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO, Semantic Web 1 (2016) 1–5, IOS Press, http://www.semantic-web-journal.net/system/files/swj1366.pdf

[4] Yad Vashem, Holocaust Denial Laws and Other Legislation Criminalizing Promotion of Nazism, https://www.yadvashem.org/holocaust/holocaust-antisemitism/holocaust-denial-laws.html

[5] Robinson, Carol, Malhotra, v., Defining the concept of supply chain quality management and its relevance to academic and industrial practice, International Journal of Production Economics, Volume 96, Issue 3, 18 June 2005, Pages 315-337, https://doi.org/10.1016/j.ijpe.2004.06.055

[6] Central Registry for Information on Looted Cultural Property, https://lootedart.com/news

[7] German Wikipedia, Liste von Restitutionsfällen, https://de.wikipedia.org/wiki/Liste_von_Restitutionsf%C3%A4llen

[8] IBM Bluemix, https://www.ibm.com/demos/live/natural-language-understanding/self-service/home

[9] Textrazor, https://www.textrazor.com/demo

[10] The International Holocaust Remembrance Alliance (IHRA), https://www.holocaustremembrance.com/stories/reference-holocaust-gdpr

[11] Gruetze, Toni, et all, CohEEL: Coherent and efficient named entity linking through random walks, Journal of Web Semantics, Volumes 37–38, March 2016, Pages 75-89, https://doi.org/10.1016/j.websem.2016.03.001

[12] OpenRefine, https://www.wikidata.org/wiki/Wikidata:Tools/OpenRefine

[13] Getty Research Institute, Getty Vocabularies OpenRefine Tutorial, https://www.getty.edu/research/tools/vocabularies/obtain/getty_vocabularies_openrefine_tutorial.pdf
[14] DBpedia Lookup, https://wiki.dbpedia.org/lookup
[15] Google Knowledge Graph API, https://developers.google.com/knowledge-graph
Dataset file with Wikidata, Viaf and birth, death dates for spoliated collectors.
[16] TMS Collections and eMuseum by Gallery Systems, https://www.canada.ca/en/heritage-information-network/services/collections-management-systems/collections-management-software-vendor-profiles/tms-emuseum-gallery-systems-profile.html
[17] see attached Datasets https://docs.google.com/spreadsheets/d/e/2PACX-1vTh2segiyBrwwIGi53198QFPj1xgwhwlwfMDM1cV9Z1IHI8K6E1Bj-A9P_riodQKJKcFDaLimAhMnIu/pubhtml?gid=1954686210&single=true
[18] Understanding Supply Chain Excellence - Best Practices and Case Studies | MSU Online, https://www.michiganstateuniversityonline.com/resources/supply-chain/supply-chain-excellence/
[19] https://www.slideshare.net/antoineisaac/designing-a-multilingual-knowledge-graph-dcmi2018
[20] Gesese, Genet Asefa, and Alam, M. and Sack, H., Semantic Entity Enrichment by Leveraging Multilingual Descriptions for Link Prediction, ArXiv, abs/2004.10640 (2020), https://arxiv.org/abs/2004.10640
[21] https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Women_scientists.
[22] Muhao et all, Multilingual Knowledge Graph Embeddings for Cross-lingual Knowledge Alignment, Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17), (2017) https://www.ijcai.org/Proceedings/2017/0209.pdf
[23] https://en.wikipedia.org/wiki/Wikipedia:Systemic_bias
[24] Ismayilov, Ali et all, Wikidata through the Eyes of DBpedia, Editor(s): Aidan Hogan, Semantic Web 0 (0) 1–11 IOS Press
[25] Fink, Eleanor E., Overview and Recommendations for Good Practices, American Art Collaborative (AAC) Linked Open Data (LOD) Initiative (2018), https://s3.amazonaws.com/assets.saam.media/files/documents/2020-07/OverviewandRecommendationsAccessible.pdf
[26] Aroyo, Lora, Data excellence: Better data for better AI, ODSC 2020 (2020) https://www.slideshare.net/laroyo/data-excellence-better-data-for-better-ai