Assessing Knowledge Organization Systems from a gender perspective: Wikipedia Taxonomy and Wikidata Ontologies

Journal:	*Journal of Documentation*
Manuscript ID	JD-11-2023-0230.R1
Manuscript Type:	Article
Keywords:	Inspection, Heuristic Methods, Gender perspective, Wikidata, Wikipedia, Ontology, Taxonomy, Knowledge Organization System (KOS)

Journal of Documentation

Assessing Knowledge Organization Systems from a gender perspective: Wikipedia Taxonomy and Wikidata Ontologies

Abstract

Purpose

Develop a comprehensive framework for assessing the knowledge organization system (KOS), including the taxonomy of Wikipedia and the ontologies of Wikidata, with a specific focus on enhancing management and retrieval from a non-binary gender perspective.

Design/methodology/approach

This study employs heuristic and inspection methods for assessment. A method is designed to inspect Wikipedia's Knowledge Organization Systems (taxonomy), ensuring their compliance with established specifications and international standards. Additionally, an evaluation is conducted to gauge the effectiveness and efficiency of retrieving articles related to women and non-masculine genders using the Catalan category scheme (taxonomy of Viquipèdia), with a focus on identifying its limitations. Furthermore, the research includes a novel evaluation of the Wikidata ontologies in terms of their structure and coverage of gender-related properties and classes. This evaluation includes a comparative analysis with the Wikipedia category scheme (taxonomy) to discern the advantages and enhancements it offers.

Findings

This study evaluates Wikipedia's taxonomy and Wikidata's ontologies, establishing evaluation criteria for gender-based categorization and exploring their structural effectiveness. The evaluation process suggests that Wikidata ontologies may offer a viable solution to address Wikipedia's categorization challenges.

Originality/value

The assessment of Wikipedia categories (taxonomy) based on Knowledge Organization System standards leads to the conclusion that there is ample room for improvement, not only in matters concerning gender identity but also in the overall knowledge organization system to enhance search and retrieval for users. These findings bear relevance for the design of tools to support information retrieval on knowledge-rich websites, as they assist users in exploring topics and concepts.

Keywords:
Knowledge Organization System (KOS); Taxonomy; Ontology; Wikipedia; Wikidata; Gender perspective; Heuristic Methods; Inspection Methods; Knowledge Organization Standards; Comparative Analysis; Gender-Based Knowledge Organization System; Information Retrieval

1. Introduction

Wikipedia is a widely used educational resource with billions of readers in numerous languages, created through open collaboration. Despite its achievements, Wikipedia suffers from a persistent gender bias with a low percentage of content on women and few female editors (Hinnosaar, 2019; Wagner et al., 2016). This gender bias is exacerbated in some Wikipedia editions, such as the Italian or Catalan versions, due to decisions about gender-related categories that should provide access and visualization of content related to gender identities. In these cases, categories like "woman" or "non-binary person" are prohibited for the organization of content and thus information retrieval. These community based decisions lead to some dysfunctions, which are particularly critical in languages that use grammatical gender, such as Catalan and Italian. Addressing this bias is important for providing equitable information retrieval and knowledge representation.

In the digital age, Knowledge Organization Systems (KOS) encompass a range of critical tools such as classification systems, thesauri, lexical databases, ontologies, gazetteers, and taxonomies. These KOS have assumed an increasingly pivotal role in the realm of information management and diverse applications. Their primary function is to meticulously convey semantics, accomplishing a multifaceted array of functions.

First and foremost, KOSs are indispensable for representing and indexing information and documents. They provide a structured framework that aids in the organization and retrieval of information. Furthermore, KOS act as knowledge-based assistants for information seekers, guiding them through the intricacies of data. They serve as semantic guides across various domains and fields, facilitating a deeper understanding of complex subject matter. In addition, KOSs function as communication tools, furnishing a conceptual framework that bridges the gap between experts and non-experts, ensuring a common language for effective communication. Moreover, they offer a foundational structure for knowledge-driven systems, enabling the seamless integration of data and knowledge in various applications (Zeng and Mayr, 2018).

KOSs are pivotal in structuring and classifying vast amounts of information in our digital age. Prominent examples of these systems can be found in Wikipedia and Wikidata. However, evaluating these knowledge organization structures, known as taxonomies in Wikipedia and ontologies in Wikidata, remains a complex challenge. There is currently no established methodology for determining the optimal indicators and metrics required for the comprehensive assessment of these structures. The creation of these metrics often relies on the specific context of the study, which can introduce subjectivity and inconsistency into the assessment process.

This academic paper conducts an in-depth examination of taxonomies and ontologies in Wikipedia and Wikidata. The primary objective is to establish a methodology for evaluating these systems, quantifying categorization issues in Wikipedia, and assessing Wikidata's suitability. It also aims to reduce the gender gap on Wikipedia by visualizing gender diversity from Wikidata. While Wikipedia has limited gender categories, Wikidata provides a broader range, including agender, intersex, non-binary, transgender, and more. The connection between Wikipedia and Wikidata is notable.

Wikidata faces a unique challenge in structuring gender data. While Wikipedia confines itself to male and female categories (in some editions, only the male category), Wikidata's property 21 encompasses a wide array of gender classes, including agender, female, male, intersex, non-binary, transgender female, and transgender male, among others. A pre-existing connection exists between Wikipedia and Wikidata, with Wikidata serving as an integral component of Wikipedia's infrastructure. Furthermore, the utilization of ontologies to enhance information organization and retrieval in Wikipedia is evident in specific cases, such as the management of the "living people" category.

In this challenge about gender data, an essential discussion concerning gender and sex, particularly regarding Property talk:P21 (Wikidata, 2024) has surfaced. Concerns have arisen regarding the conflation between sex and gender within a single category on Wikidata. There is a call for distinct properties to differentiate sex, gender, and potentially gender identity, similar to having separate properties for height and weight. The text highlights issues related to the vague and unclear classification of terms like male, female, man, and woman. It is suggested to have a separate property and values for gender identity, distinct from biological sex, with clear and unambiguous definitions to avoid intentional conflation that causes problems with dataset clarity and unambiguous representation. Furthermore, it discusses similar concerns in official contexts, such as the discussion initiated by the UK government in 2018 regarding managing gender or sex statements, indicating parallel challenges faced by both Wikidata and the government.

Additionally, unresolved situations related to the assignment of property values are pointed out, including issues with assigning "male" to someone who is biologically female, questioning the differentiation between human males and non-human males, confusion between transsexualism and Gender Identity Disorder (GID), and the need for more accurate representation of values such as "intersex" and "transgender." Furthermore, it is noted that special situations, such as assigning gender to anthropomorphic nonhumans and dealing with unknown gender, have not been adequately resolved. The necessity of incorporating a "citation needed" constraint to the property, requiring at least one reference for value assignment, is also analysed.

On a related note, the inappropriate addition of sex or gender statements for living individuals via Quickstatements or bots on Wikidata, leading to harmful misgendering and potential privacy violations, is brought up. Proposed solutions to prevent future harm include disallowing bots and Quickstatements from affecting more than ten items at a time and discouraging the use of labels and given names as references for sex and gender statements. These proposals aim to ensure more careful handling of sex and gender statements to avoid harm and privacy violations, reflecting the community's concerns with promoting responsible and ethical practices on Wikidata.

Finally, this paper evaluates the Knowledge Organization System (KOS) using the Catalan Wikipedia as a case study on the gender gap. It seeks to improve gender identity visualization and accessibility through Wikidata ontologies. It acknowledges potential biases in Wikidata and Wikipedia and their capacity to perpetuate real-world biases. Furthermore, it is essential to acknowledge that Wikidata's potential biases are no greater than those present in the real world (Zhang and Terveen, 2021). Additionally, some authors argue that Wikipedia mirrors real-world biases (Eckert and Steiner, 2013) with the platform having the

capacity to perpetuate and exacerbate gender gaps, shaped not only by editors but also by infrastructural logics (Ford and Wajcman, 2017).

The objective is to evaluate Wikipedia's taxonomy and Wikidata's ontologies to enhance gender diversity visibility. The paper synthesizes theories and insights to establish comprehensive evaluation criteria. The ultimate aim is to provide an objective approach to assess knowledge organization systems in Wikipedia and Wikidata and quantify their structural effectiveness. Subsequent sections will detail the evaluation process and findings, addressing Wikidata's potential as a solution for Wikipedia's categorization challenges.

2. Literature Review

In this section, we provide a comprehensive review following the SALSA framework (Grant and Booth, 2009) to examine the gender gap in Wikipedia and Wikidata. Academic research has extensively investigated the gender gap in both platforms. The Wikipedia appraisal stage involved 97 articles, and the Wikidata appraisal involved 34. A total amount of 21 articles were used to assess Wikipedia (Ferran-Ferrer et al., 2023), and 19 were used to evaluate Wikidata.

2.1. Gender gap in Wikipedia

The gender gap on Wikipedia has been the subject of extensive academic research, with numerous studies exploring biases in content, participation, reading, and potential strategies to address this gap. These studies emphasize the importance of recognizing and addressing biases and barriers to create a more diverse and inclusive Wikipedia community (Ferran-Ferrer et al., 2023).

The under-representation of women as editors and as subjects of biographical coverage is a widely recognized issue in the academic field (Hube, 2017; Falenska et al., 2021). Some articles discuss how gender bias intersects with race, sexuality, security, and marginalization on Wikipedia (Lam et al., 2011; Ju and Stewart, 2019; Tripodi, 2023). Various factors, such as the demographics of editors, platform structure, and cultural values, contribute to these biases, which have significant social implications, affecting the visibility and participation of women and perpetuating existing disparities (Ford and Wajcman, 2017).

Regarding the gender gap in content, research reveals that women are underrepresented among the main figures in all language editions of Wikipedia (Miquel-Ribe and Laniado, 2021). Articles for deletion is a possibility within the decision-making process in Wikipedia article editing. It is the process that determines what constitutes knowledge and what does not in the encyclopedia. Biographies of women and LGBTQ+ individuals are often subject to deletion, resulting in a higher proportion of biographies of women nominated for deletion compared to biographies available about men (Morgan et al., 2013; Hollink et al., 2018; Tripodi, 2023). While there are indications of bias, some authors conclude that there is no clear bias resulting from deletion activity (Worku et al., 2020).

Studies also identify significant gender differences in Wikipedia content, such as biographies of women featuring more prominent family, gender, and relationship themes (Wagner et al., 2016). Linguistic bias in terms of language abstraction and positivity can be observed, along

with structural differences in metadata and hypertext links. In addition, citation practices reveal that female authors are cited less than expected, suggesting a preference for citing male publications (Zheng et al., 2022). These biases may further marginalize female authors, especially in non-Anglophone countries. The gender gap in content creation and participation on Wikipedia perpetuates an unbalanced coverage of topics, creating a cycle where the lack of diversity in content fails to attract and engage different editors, thus exacerbating the existing gender gap (Konieczny and Klein, 2018).

Research on the gender gap in editing and participation highlights various barriers that hinder women's involvement on Wikipedia. These barriers include negative reputation, lack of recognition, fear of deletion, rejection, and alienation. Often, research suggests that women lack confidence in their abilities, feel uncomfortable with editing, and face negative responses to constructive feedback (Collier and Bear, 2012). Factors such as the digital skills gap (Gardner, 2011) and the availability of time for editing (Gruwell, 2015) also contribute to the gender gap. However, visible female editors and constructive comments can help mitigate the gap, as the presence of visible female peers promotes collaborative editing (Evans et al., 2015). Some authors have investigated the gender gap in Germany and suggested a proactive approach to training and educating women to enhance their motivation for writing (Buchem and Kloppenburg, 2013). It has also been highlighted the impact of family responsibilities on women's ability to write, so efforts may need to focus on addressing gender disparities in domestic work (Ferran-Ferrer et al., 2021).

The gender gap extends beyond editing and includes the underrepresentation of individuals. Female participation varies by topic, with a greater presence in gender studies or feminism categories, reflecting traditional gender stereotypes. Generic site restrictions limit the digital credibility and authority of women, hindering their contributions. The complex relationship between the gender gap and harassment requires better understanding, and it is important to create a safe environment for women on and off Wikipedia. Feminist interventions, such as exclusive edit-a-thons for women, have proven effective in countering gender inequality on the platform.

2.2. Gender and Wikidata

The gender gap on Wikidata has been extensively explored in academic research. We can delineate three main categories of studies. A first set of research has delved into the gender gap within Wikidata, presenting diverse methodologies, findings, and recommendations to address this disparity. Meanwhile, a second set aims to quantitatively assess the biographical gender gap in Wikipedia, across various language editions, leveraging Wikidata's multilingual support to facilitate this cross-cultural research. Lastly, a third set of studies emphasise the advocacy and visibility of content pertaining to women in industries traditionally dominated by men, utilising Wikidata for this purpose.

Regarding the initial group of discussions aimed at presenting diverse methodologies, findings, and recommendations to address this disparity, Zhang and Terveen (2021) delved into the gender content gap in Wikidata, seeking to uncover the source of bias. Through a quantitative case study, they examined how individuals were represented in Wikidata compared to existing gender biases. Their findings revealed a prevalence of male-dominated

professions among the most frequently represented categories, closely mirroring real-world gender distribution.

Similarly, Abián, Meroño-Peñuela and Simperl (2022) sought to understand the impact of content gaps in knowledge graphs on downstream applications, with a particular focus on gender disparities within Wikidata. To achieve this, they introduced a framework that compared edit metrics with Wikipedia pageviews, facilitating a quantitative evaluation of discrepancies between knowledge graph content and user needs. As a result, they identified no inherent gender or recency gaps within Wikidata's production, with only a few under-represented entities standing out. A group of articles has focused on analysing gender bias on Wikidata concerning occupations or professional domains. In this line, Das et al. (2019) conducted a holistic analysis of bias measurement on the knowledge graph, specifically focusing on biases in Wikidata across different demographics selected from seven continents. They utilised extensive experiments on a wide range of occupations sampled from various demographics, examining the impact of algorithm bias on the measurement of biassed occupations. Results indicated that the inherent data bias in Wikidata can be influenced by specific algorithm bias and underscored the importance of understanding biases based on socio-cultural differences across demographics. Within this same field, there are three works that concentrate on specific occupations or professional domains:

Lemus-Rojas and Lee (2019) in the STEM fields, Zhu et al. (2023) in Chinese culture and heritage, and Conroy (2023) in French and Francophone literature. The outcomes align with the conclusions observed in the aforementioned comprehensive studies. In the first two cases, Wikidata is highlighted as a critical collection for enhancing the visibility of women. Conroy (2023) found that the gender gap in both subsets closely resembles the global average, with a higher-than-average representation of writers of other genders.

Finally, Pellissier and Suchanek (2019) and Bourli and Pitoura (2020) analysed gender bias on Wikidata through advanced automated processing techniques. Pellissier and Suchanek (2019) proposed a system to index changes in the Wikidata graph and enable users to answer complex SPARQL queries regarding historical changes, while Bourli and Pitoura (2020) introduced measures for identifying bias in the dataset, tested methods for amplifying bias in embeddings, and introduced a debiasing approach. A special case is Mandiberg and Sarıoğlu (2022), who aimed to address the challenges associated with defining a dataset to analyse changes in Wikipedia's gender gap for articles about visual art. The dataset is constructed from the intersection between Wikipedia and Wikidata. The researchers describe the process of using a topic model algorithm to identify a dataset by analysing the words within each article and grouping articles into topics. Their aim was to create a dataset that more closely reflects visual artists' articles on English Wikipedia, addressing potential systemic biases. The topic model algorithm provided a dataset that encompassed a majority of the two WikiProject datasets and the Wikidata sets, while adding additional art-related individuals. It was found to be superior to other options, offering a detailed list of articles about visual arts that mitigated Wikipedia's existing imbalances. The study also highlighted challenges in Wikidata's taxonomies and called for further research on systemic biases reflected in taxonomy systems.

A second set of articles addresses the application of Wikidata, capitalizing on its multilingual capabilities to facilitate comprehensive cross-cultural research, for measuring gender bias in

Wikipedia editions and for resolving this issue. Three of these studies feature contributions from Maximilian Klein and Piotr Konieczny. Klein and Konieczny (2015) and Konieczny and Klein (2018) introduce the Wikipedia Gender Inequality Indicator (WIGI) developed from Wikidata. WIGI calculates, for each country, a score based on the ratio of female and nonbinary gendered biographies to the total number of biographies. This Wikipedia-derived indicator is correlated with four contemporary, widespread gender inequality indices (GDI, GEI, GGGI, and SIGI). Through analysing methodologies and the relationship with Wikipedia data, evidence suggests that the bias in Wikipedia's biographical coverage is aligned with gender bias in socially powerful positions. Concerning the results, Klein and Konieczny (2015) find that the strongest correlations are with individuals born around 1910, indicating that Wikipedia's representation may more accurately reflect current rather than historical gender statuses. The same authors Konieczny and Klein (2018) utilise cultural clusters to highlight how gender inequality can be examined through diverse cultural perspectives.

Klein et al. (2016) delves deeper into the gender bias of content, focusing on women's biographies on Wikipedia. The article underscores the importance of precisely measuring the gender content gap and the critical examination of initiatives intended to mitigate this disparity. The team formulates the Wikidata Human Gender Indicators (WHGI), a robust, longitudinal dataset to monitor gender disparities. It monitors biographical data across multiple facets—such as time, geography, culture, occupation, and language—providing an extensive instrument for elucidating and quantifying the gender bias in Wikipedia's content. The research signals a changing representation of women in 11 dimensions utilising WHGI. Validations against three external datasets back the indicator's accuracy, and reassessment of Wikipedia's gender bias with WHGI suggests that it could enhance depth and impact in future research on the subject.

In a similar line of work, Hollink et al. (2018) tackles the challenge of measuring gender inequalities on Wikipedia, especially when considering multiple languages. The difficulty in finding objective methods to measure and compare gender inequality is underlined, and the potential differences across language editions of Wikipedia are acknowledged. Their methodology focuses on comparing coverage of male and female Members of the European Parliament (MEP) across various Wikipedia language editions using open data. This approach allows for a fair comparison due to the MEPs' notable actions in the real world, and it examines gender discrepancies in both the coverage on Wikipedia and the content within Wikidata entries. An analysis of Wikidata entries for male and female MEPs reveals equal amounts of property-value pairs, contradicting earlier studies that found Wikipedia content related to women emphasised family and relationships. Differences related to real-world disparities suggest that the structured data of Wikidata might be less prone to bias. Moreover, aggregation of data from various Wikipedia language editions might contribute to a more diversified and equitable dataset in Wikidata.

Delving into the characteristics and virtues of Wikidata, Hermoso Pulido (2021) discusses how Wikidata has become a significant tool within the Wikimedia ecosystem, improving data linkage and reuse. Specifically, it mentions the adoption of Wikidata in Catalan Wikipedia, noting how its integration with infoboxes and list generation has advanced the project. The article suggests that such technical innovations could be part of the solution in addressing Wikipedia's gender gap. Methodology highlights the use of structured data from Wikidata to evaluate new biographical articles, aiming to encourage user engagement in diversity issues

and track vandalism or errors. This methodology suggests a proactive approach to using structured data for maintaining quality and diversity in biographical content, directly impacting the reduction of Wikipedia's gender gap. Technical challenges are highlighted, such as execution timeouts during SPARQL queries for live data analysis. While some limitations exist for large datasets, initiatives like WCDO show promise in identifying and acting upon content gaps. The article advocates for enhanced cross-collaboration between Wikidata and Wikipedia, suggesting that embedding certain tools could encourage editors to address discrepancies more effectively.

Leveraging the potential of Wikidata, Laouenan et al. (2022) focus on studying different intersectionalities, specifically, they aim to construct a comprehensive and accurate database of notable individuals by cross-verifying the information from various editions of Wikipedia and Wikidata, focusing on specific social science questions about gender, economic growth, urban and cultural development. The researchers collected a significant amount of data from Wikipedia and Wikidata, utilising deduplication techniques and cross-verifying the retrieved information. They found varying degrees of completeness and error rates dependent on notability distribution, classifying the presence of an Anglo-Saxon bias in the English edition of Wikipedia. The strategy resulted in the creation of a cross-verified database of 2.29 million individuals, shedding light on an Anglo-Saxon bias in the English edition of Wikipedia. The study also emphasised the implications of this bias and identified individuals not present in the English edition of Wikipedia.

Finally, the last research strand in this set of papers aims to emphasise the promotion and visibility of content related to women in male-dominated professional spheres through the utilisation of Wikidata. Among these, two articles are authored by Thornton and Seals-Nutt, both affiliated with the Stories Services Collaborative. Thornton and Seals-Nutt (2018) introduce the creation of a web application called Science Stories. This application utilises structured data from Wikidata along with images to narrate compelling science stories, especially focusing on the experiences of women who have contributed to scientific research. The primary goal is to elevate the visibility of these women. The authors illustrate how the use of free software and open standards can lead to the development of visually captivating and interactive science communication experiences. These experiences involve the integration of images with structured statements within a web of interconnected data, all supported by references to published sources. Four articles focus on leveraging Wikidata to promote and illuminate the contributions of women in male-dominated professional fields. In a similar vein, Thornton et al. (2022) delve into how Semantic Web capabilities can consolidate disparate materials to craft narratives, as demonstrated by the WeChangEd research project, which centres on women editors of periodicals in Europe from 1710-1920. The methodology involves developing applications that aggregate data from Wikidata to harness a versatile knowledge graph, facilitating the swift creation of interactive platforms to captivate fresh audiences. The outlined process holds potential value for researchers and cultural heritage institutions seeking web-based avenues for presenting data-driven storytelling.

3.Objectives

The main aim of this research is to explore and compare the effectiveness and efficiency of the KOS of female biographies on Wikipedia and non-male ones. This will be accomplished by evaluating the category structure of the Catalan edition of Wikipedia and the ontology of Wikidata, with the aim of addressing the challenge of visualizing the diversity of gender identities and accessing their content on Wikipedia. We will aim to ascertain whether Wikidata ontologies can offer a more improved means of organizing and representing the information available on Wikipedia regarding the diversity of gender identities.

Therefore, the research questions that we will address are:
QR1: How can a standards inspection method be developed to evaluate the conformance of the KOS in Wikipedia with international specifications and standards established by recognized organizations?
QR2: How does the category scheme of the Catalan edition of Wikipedia impact the effectiveness and efficiency of retrieving articles related to women and non-male genders, and what specific limitations does it present?
QR3: To what extent does the Wikidata ontology facilitate the effective and efficient retrieval of articles concerning women and non-male genders, and what advantages or enhancements does it offer in comparison to the Wikipedia category scheme?

To address these questions, a specific methodology is created and applied for each of the specific objectives (See Table 1):

INSERT TABLE 1

4.Methodology

To explore the nature of Wikipedia as a taxonomy, as opposed to a folksonomy, and provide insights into Wikidata's data model, it is documented in Centelles and Ferran-Ferrer (2024).

4.1. Inspection of standards and guidelines for the evaluation of taxonomy (Wikipedia) and ontologies (Wikidata)

Our study begins by reviewing the most widely accepted standards for the analysis of KOSs, and using them as the basis for designing an evaluation guide tailored to the taxonomic and ontological criteria relevant to Wikipedia and Wikidata. Subsequently, we employed a standards inspection method to assess whether the KOSs of Wikipedia and Wikidata conform to the international specifications and standards defined by recognized organizations.

In the theoretical framework of our study, we draw upon the taxonomic classification proposal of (Souza et al., 2012) and the critical insights of Mazzocchi (2018) into KOS. These foundational works underpin our proposed evaluation criteria for Wikipedia and Wikidata.

Specifically, in the context of Wikipedia, Albuquerque (2017) presents an information architecture framework for the development and management of controlled vocabularies in the context of programming vocabulary projects. Kaplan et al. (2022) introduce an evaluation method for taxonomies, including structural quality criteria such as generality, appropriateness-attainment, and orthogonality, and provide generalized metrics for quantifying generality and appropriateness.

In the domain of ontologies, da Costa et al. (2022) provide an updated review of software architectures, including ontology usage for managing large volumes of data. Wilson et al. (2022) outline a methodology for evaluating ontology quality that considers intrinsic and extrinsic aspects. Amith et al. (2018) offer insights into ontology evaluation within the field of biomedical KOS, which we adapt for evaluating Wikidata. Bolotnikova et al. (2011) propose practical methods for ontology evaluation, especially in automated contexts. Aghaebrahimian et al. (2022) explore the validity of Wikipedia categories for topic labeling, further contributing to the development of our evaluation criteria.

The extrinsic criteria (Kless and Milton, 2010) assess the measurement of external qualities, their application, and the domain, making reference to elements of the outcome as experienced by users. In contrast, quality indicators analyze aspects of structure and domain independently of their use in application contexts.To gain a comprehensive understanding of the efforts to unify the reviewed theories and the proposed methodology for ontology evaluation, see Table 2.

INSERT TABLE 2

4.2. Proposed heuristic evaluation of taxonomies

Heuristic evaluation, aiming to assess whether the taxonomy of Catalan Wikipedia complies with the standards of sound knowledge organization, not only concerning user experience but also formally within the realm of Knowledge Organization Systems (KOS). Based on the theoretical framework, a selection of indicators were selected that have been highlighted in our analysis and achievable with the access and technical resources available to us (See Table 3). When identifying and measuring these indicators, we have considered contributions from specialists and specific standards within the KOS sector, particularly taxonomies, to conduct an inspection analysis of Wikipedia's category scheme.

INSERT TABLE 3

4.3. Analysis of usage logs for the profession case study on gendered professions

For the analysis of logs of the Catalan edition of Wikipedia we have used Pageviews Analysis (https://pageviews.wmcloud.org) which is a suite of eight tools designed for the examination of page views and unique device statistics on Wikimedia Foundation wikis. These tools, namely Pageviews, Langviews, Topviews, Siteviews, Massviews, Redirect Views, Userviews, and Mediaviews, collectively form a comprehensive toolkit for data analysis. The foundation of these tools relies on data sourced from Wikimedia's RESTBase API, which is structured in alignment with the definitions outlined in the Research: Page view

and Research: Unique Devices documentation. Presently, this suite of tools is under the maintenance and stewardship of Community Tech.

To address this analysis, we have chosen the field of professions, and based on state statistical data (INE: Instituto Nacional de Estadística, 2024), we have selected the most masculinized (STEM) and feminized professions (nursing, library science, and teaching) in Spain.

4.4. Heuristic assessment concerning structure and coverage

It is essential to clarify that in Wikidata, property P21 encompasses both gender and sex. However, it is crucial to recognize that these two terms pertain to distinct aspects of human identity and biology. Sex is primarily associated with an individual's physical and genetic characteristics and has historically been classified into two categories: male or female. In contrast, gender is a social and cultural construct that encompasses a broad spectrum of roles, behaviors, expectations, and identities. It extends beyond a binary system, acknowledging that people can identify as male, female, both, neither, or a different gender altogether. It is imperative to comprehend the differentiation between sex and gender, as it is fundamental for fostering inclusivity and honoring the diverse experiences and identities of individuals (García Dauder and Pérez Sedeño, 2017).

Apart from this feature of gender or sex of Wikidata, the members of the Ontology project have identified the limitations that make it not qualify as a proper ontology (Wikimedia, 2022). These limitations can be divided into two groups. The first group was initially identified in WikidataCon 2021, and they are aimed at overcoming barriers to the reuse of data by other services and projects. And the second group is considered to be issues existing in the knowledge representation in Wikidata. In the context of this study, we are primarily interested in the first group, as it identifies elements to overcome if it is to be applied in the categorization of Wikipedia content.

Based on the barriers to reuse formulated by the project members, we present examples related to the classes that make up the range restriction of property P21 (gender or sex). The indicators have been selected considering their relevance and their suitability for the retrieval of gender-related articles, however, this can be extrapolated to other evaluator needs.

INSERT TABLE 3

4.5. Performance of the Wikidata search system

The data from Wikidata can be used for various purposes. Beyond the specific querying of an item or a set of items, Wikidata provides users with methods of data access for linking data without having to download it to another server, for enriching third-party data, or for generating local search services. In all cases, Wikidata's data can be consumed by human users or by automated systems or bots (Wikimedia, 2023b).

In one of the Wikidata guides, "Data Access" (Wikimedia, 2023a), eight methods for accessing Wikidata data are identified and described, three of which are oriented towards direct interaction with users who need to retrieve limited quantities of results (See Table 4).

INSERT TABLE 4

All methods of accessing Wikidata data operate on a foundation formed by the RDF data management system, or RDF repository, Blazegraph (Vrandečić et al., 2023) (See Table 5).

INSERT TABLE 5

Undoubtedly, these figures are impressive and represent the largest open secondary database currently in existence. Nevertheless, in recent years, assessments of the degree of compliance with processes, accessibility, and the use of search services have shown worrisome signs of stagnation. The Wikidata authorities are fully aware of these limitations and, in fact, have set their sights on the need to replace the underlying software of Wikidata, Blazegraph, with one that can better address the challenges of growth and quality.

And, regarding the ontology inconsistencies we mentioned earlier, the evaluation requirements established incorporate the use of more advanced integrity-checking languages than SPARQL functions. Specifically, the WDQS report refers to the Shapes Constraint Language, or SHACL. SHACL allows for graph validation and includes not only the ability to specify a severity level for validation results, but also the possibility of providing suggestions on how to fix the data if a validation result occurs.

The performance assessment of Wikidata follows the overarching evaluation framework introduced by Malyshev et al. (2018). The performance tests cover the period from 2015 to 2022, as specified by the SPARQL query service (Everett, 2015).

5. Results

5.1. Heuristic evaluation of SOK Wikipedia

The heuristic evaluation of the Catalan Wikipedia category scheme has been carried out using the technique of standards inspection, in which a usability expert analyzes whether the interface follows the agreed-upon specifications and the standards defined on an international level. In the case at hand, a set of identified indicators has been generated, particularly based on normative sources. These sources also provide us with methods for obtaining evidence for each indicator, the applied metrics, and, when possible, optimal values.

a) Evaluability

The category schema of Wikipedia is valuable because category creators have various agreed-upon tools for their practice. We highlight the following:

- Categorization guideline (Wikimedia, 2023d) resulting from the discussion and decision-making process specific to the encyclopedia.

- Help for category creators: Help:Category (Wikimedia, 2018) and the "Style Book on Categorization" section in the Categorization guideline (Wikimedia, 2023d).
- Templates for category creators: Category:Maintenance templates for categories (Wikimedia, 2015).
- There is a control over the pages that do not contain categories, for maintenance purposes.

The level of knowledge about these tools has been informally assessed with some individual administrators of the Viquipèdia community, and their lack of awareness regarding them has been conveyed.

b) Reusability

Each category has a unique instance and a single identifier, regardless of its various locations within the schema hierarchies. Wikipedia's schema categories are available for database dumps in three data interchange formats: sql, json, and xml. These formats do not provide semantic information about the concepts and relationships between the concepts in Wikipedia's categorization schema, as the Simple Knowledge Organization System (SKOS) data model could. Consequently, the possibilities for reusing Wikipedia categories in other datasets or information retrieval systems are greatly restricted.

c) Stability

The most notable stability-related metrics during the period 2004-2022 can be seen in Table 6.

INSERT TABLE 6

The annual growth rate remained high during the period 2004-2007, and from 2008 onwards, it experienced a significant decline until 2012. Starting from that year, the rate demonstrates a gradual reduction in the increase, and it stabilizes until 2020 when it experienced a very remarkable increase. This increase may be linked to one of the effects of the Covid-19 pandemic: greater availability of time for contributing to the Catalan edition of Wikipedia.

The most abundant categories are in the fields of science and culture, followed by technology, humanities, and events. The three areas with the fewest categories are biographies, information, and places.

d) Number of categories (concepts)

As of December 31, 2022, Wikipedia's category schema included 102,159 categories. We can compare this size with other similar KOSs that consist of pre-coordinated concepts and aim to represent encyclopedic knowledge.

The List of Subject Headings of the National Catalan Library (LEMAC) contains 112,200 headings, considering both accepted and non-accepted ones. It originates from the translation of the Spanish version of the Library of Congress Subject Headings (LCSH), which was preliminarily published by the Library Services of the Generalitat de Catalunya in 1988.

Making growth rate comparisons between Wikipedia's categorization schema and LEMAC is challenging. Wikipedia's schema is relatively young, still in its first two decades of existence, while LEMAC has been in existence for 35 years, with an even longer history if we consider

its origins. However, the following examples show clear indications of faster growth in Wikipedia's categorization schema.

In 2009, LEMAC experienced a growth of 2,663 new headings, whereas Wikipedia added 8,080 new categories. By 2021, LEMAC's growth amounted to 1,219 new headings, while Wikipedia introduced a staggering 8,616 new categories.

The magnitude of Wikipedia's categories and its growth rate are not comparable to those of Knowledge Organization Systems (KOS) applied to digital encyclopedias. The knowledge tree of the Gran Enciclopèdia Catalana (GEC) comprises 425 categories, and the categories in the Encyclopedia Britannica total 123. In both cases, the hierarchy of the KOS is restricted to two levels.

e) Number of semantic relationships

We lack access to complete data on all hierarchical relationships between supercategories and subcategories. Still, we have a partial count covering the first three levels of the categorization schema, specifically involving main thematic categories and their second-level and third-level subcategories, resulting in 1122 hierarchy relationships.

f) Enrichment or granularity index

In the first three levels of the category schema, there are 143 categories and 1122 hierarchy relationships. The corresponding average enrichment index is 7.8461, significantly surpassing the optimal range, usually between 2 and 5.

g) Degree of precoordination

In the entirety of Wikipedia's category schema, the average number of words that make up category labels is 3.6766, exceeding the maximum value typically recommended by experts, which is between 1.5-2 words. Other data indicating a deviation from this optimal value include the median number of words in category labels, which is three. However, there are exceptions, with some category labels having a maximum of 18 words, such as in the case of "Resolucions del Consell de Seguretat de les Nacions Unides sobre el Tribunal Penal Internacional per a l'antiga Iugoslàvia" (Resolutions of the United Nations Security Council on the International Criminal Tribunal for the former Yugoslavia).

h) Number of levels in hierarchy or depth

For the evaluation of this indicator, we reviewed all hierarchical chains within the main thematic category "Biographies." In all cases, we found more than five levels. Consequently, this exceeds the maximum value recommended by experts

i) Number of categories in the same hierarchy level or breadth

To assess this indicator, we examined the first two levels of subcategorization beyond the eight main thematic categories. In total, this section includes 144 categories, with 130 of them containing subcategories, while the remaining 14 link directly to Wikipedia pages.

Among these 130 categories, there are 17 that have only one subcategory, constituting 13.07% of the assessed section, and in fact, this is the most common case. These instances violate the minimum requirement of two subcategories recommended by experts.

Additionally, there are 35 categories with more than 12 subcategories, making up 26.92% of the evaluated section. These cases exceed the maximum of twelve subcategories suggested by experts. The highest breach of this limit occurs in the "Religion" category, which includes 43 subcategories.

When we sum up violations of both the minimum and maximum subcategory limits, we find that 40% of the assessed categories do not adhere to the optimal values of breadth (52 out of 130 categories).

5.2. Usage logs for the case study Wikipedia

To provide a glimpse of Viquipèdia usage (Catalan Wikipedia), the total number of viewed pages of the Catalan Wikipedia in one month is 49.338.638 and the number of unique devices accesses is 3.480.772 in September 2023 (Wikimedia, 2023c).

And in this section, we present the data derived from our analysis of log entries for feminized professions (such as librarians, nurses, and teachers), juxtaposed with STEM professions (See Table 7). The table encompasses two categories: "Feminized Professions" and "STEM Professions". Each category is further broken down into specific professions, and the corresponding user visualizations statistics are presented. We examine access patterns beginning in June 2023, focusing on the Catalan edition of Wikipedia, and encompassing data from various devices. The table illustrates the engagement and activity levels across different professions within feminized or masculinized professions.

INSERT TABLE 7

The findings from the examination of feminized professions reveal that there is an average of 798 monthly accesses, with a mean of 2 editions monthly. And the outcomes obtained from the examination of STEM professions indicate that there is an average of 1073 monthly accesses, with a mean of 0 editions monthly.

5.3. Heuristic evaluation of Wikidata

These are the results of the heuristic evaluation of the ontologies of Wikidata. Examining Table 8 reveals significant challenges stemming from unproductive class hierarchies in navigation and search. Users are constrained from selecting multiple individuals of the same type within nodes, introducing complexity in search contexts. This limitation hampers quick decision-making, impacts reasoning by disrupting inference and consistency assessments, and impedes automated interoperability in data cooperation. Ascending through upper chains in search contexts confuses users and complicates processes in both search and axiom-based reasoning. The inability to determine generality or specificity transfers the challenge to the search process, hindering statement validation and new knowledge inference. Overall, unproductive class hierarchies present intricate obstacles across various facets of ontology usage.

INSERT TABLE 8

5.4. Performance Wikidata

An in-depth assessment of Wikidata's performance from 2015 to 2022 is presented in Table 9, utilizing the framework established by Malyshev et al. (2018).

INSERT TABLE 9

Table 9 furnishes a comprehensive evaluation of query performance within the Wikidata platform through several key metrics. Under the section denoted as "Query Metrics," distinct trends come to the forefront. Notably, a substantial count of "Good Queries" signifies operations executed with success, contributing significantly to the platform's operational prowess. Conversely, a noteworthy number of "Bad Queries" denotes instances where queries faced issues or failed to deliver intended results, thereby illuminating potential areas for refinement.

Moreover, the metric detailing the "Total Query Execution Time" offers a panoramic view of Wikidata's efficiency, encapsulating the cumulative time required for executing the entire array of queries. This temporal dimension serves as a pivotal indicator of the platform's responsiveness. In tandem, the metric revealing the "Total Result Rows" speaks volumes about the sheer magnitude of information generated across the spectrum of queries conducted on Wikidata. This voluminous outcome underscores the platform's extensive capacity in producing relevant and diverse information.

6. Discussion

In 2017, the Wikimedia Movement adopted a new strategic plan for 2030, which establishes the goal of "providing knowledge as a service" (becoming a platform that offers open knowledge to the world through interfaces and communities), with a focus on "knowledge equity" (directing our efforts towards knowledge and communities that have been marginalized by power structures and privileges... Breaking social, political, and technical barriers that hinder people from accessing and contributing to free knowledge). This collaborative strategic document places two core principles at its recommendations: inclusivity and a people-centered approach (understood as attending to people's needs). It sets the goal for 2030 as closing the gender gap and focusing on the inclusion of underrepresented groups.

The knowledge organization proposal presented in this paper is fully aligned with the new strategic direction. To increase the visibility and access to the knowledge of Wikipedia, particularly that related to and about marginalized gender groups (women, non-binary individuals, intersex, trans men, and trans women), a dual solution is proposed. On the one hand, a technical solution involving the use of Wikidata ontologies as a knowledge organization system to facilitate information search and retrieval in Wikipedia, without increasing biases existing in reality. Wikidata has shown to be more aligned with the gender perspective than Wikipedia, as demonstrated in in-depth studies (Zhang and Terveen, 2021). On the other hand, a social, cultural, and political solution that involves working with the Wikipedia community to accept internationally recognized standards in the field of knowledge organization and embracing two principles considered strategic by the Wikipedia movement:

inclusivity to avoid the discrimination of marginalized groups and, above all, the principle of people-centered service with a special focus on their informational needs.

With the technical solution, by opting for Wikidata ontologies as the knowledge organization system for Wikipedia's content, alignment with international standards of knowledge organization would be achieved, and it would empower Wikipedia users (readers and editors) to search for and retrieve encyclopedia content according to their needs. Empowerment would occur during the search and navigation process, as users would decide on search elements closer to their needs, unlike the current categories, which are proposed based on the worldview of those who created, classified, and indexed them, without allowing the combination of search elements.

Delving into the details of the proposal in this academic paper, it advocates the use of ontologies as a knowledge organization system and opts for Wikidata because its original purpose is to store data (properties and relationships) from content present in Wikipedia articles in any language. If taken to its fullest potential, Wikidata could become the knowledge organization system for Wikipedia. In this sense, some current Wikipedia categories are already directly constructed from Wikidata ("living people"). Wikipedia already has different sections linked to Wikidata, acting as a knowledge organization system through common examples like InfoBoxes or authority records.

In the evaluation carried out by experts and the heuristics, most assessments of Wikipedia's analysis variables are low. In contrast, for Wikidata, there is room for improvement, and the Wikidata community has already identified these and included them in the agenda for improvement and tool development to address these issues. Two interesting contributions are the entity schema and the backend, which bring substantial benefits to the ontology, such as aspects related to the organization of data related to sex or gender. This demonstrates the potential for improvement. It is also important to recognize the need to revise the name of the "sex or gender" property, which mixes biological characteristics with individual definition and social construction in its label.

In the case of Wikidata, decision-making processes for improvements are made within a smaller community with a more respectful perspective of accepted and recognized international standards in the field of knowledge organization. In contrast, in the case of Wikipedia, the gender bias existing in society is exacerbated by opposing positions on gender diversity expressed with strong ideological arguments.

At the same time, opting for Wikidata ontologies as the knowledge organization system for Wikipedia's content would empower users (readers or editors) to meet their informational needs. We agree that Wikipedia's categorization system (category schemes) is easy for users to understand and closely aligned with their vocabulary and natural language. However, it also has disadvantages, as discussed in this proposal, related to cultural, social, or political biases and imbalances that any controlled vocabulary entails. On the other hand, category schemes, are pre-coordinated thesauri that combine concepts, classes, or terms from a controlled vocabulary at the time of their construction or indexing. This means that there could be a category like "Catalan doctors from the south" created by the Wikipedia community, and a person from anywhere in the world should be able to understand (deduce) that it may include "female doctors living in the southern part of Catalonia." In contrast, the use of ontologies to

organize knowledge would provide a better representation of Wikidata's content because each property represents a single dimension (attribute) of an entity or a set of entities. It is the user who, at the time of the search, combines the attributes that best respond to their need to retrieve the relevant entity or entities. Currently, a Wikipedia category like "Catalan doctors from the south" links two different dimensions of a person: their profession and their origin. However, in an ontology, one could choose which elements to combine (profession, year of birth, place of residence, or any other) independently and in combination with Boolean operators, so that the search would align much more with the user's information needs (human or machine). The knowledge organization system proposed by Wikidata is post-coordinated, empowering users through the system by allowing them to combine predefined attributes. Therefore, Wikidata is the ontology that could bring organization and a better representation of what is known in Wikipedia. In fact, Wikipedia already generates categories that arise from Wikidata, such as "living people."

The Wikidata ontology can be a solution for applying a gender perspective and overcoming the lack of scientific-technical foundation. However, the need for a cultural change within the community, which should accept the process of substitution to align with international technology consensus and promote more equitable access to content diversity, cannot be ignored.

7.Conclusions

This study has delved into the complex issue of gender bias within Wikipedia's knowledge organization system, specifically within its taxonomy categories. By synthesizing the theoretical framework, an extensive review of academic literature, and a detailed analysis of Wikipedia's content structure, we have arrived at the following comprehensive conclusions.

Throughout our research, it has become abundantly clear that gender bias persists within Wikipedia, as attested by a consistent body of academic literature. This bias is not confined to the content itself but also extends to the limited diversity among volunteer contributors across various languages. Our own analysis further reinforces this finding, substantiating the existence of gender bias within Wikipedia's content organization system.

One notable revelation is that Wikipedia's encyclopedic nature tends to prioritize the perspectives of content indexers and categorizers over the needs of its users. Consequently, the categories within Wikipedia have often been designed to facilitate the work of editors, collecting pages under specific concepts, while falling short in terms of enabling effective information retrieval and meeting user information needs.

Furthermore, we have identified that Wikipedia's system of categories (KOS) frequently falls short of established quality standards. For instance, hierarchy depth often exceeds the recommended maximum of 12 levels, and the breadth, indicated by the number of subcategories within a category, deviates from the recommended range of 2 to 5 subcategories.

From the perspective of gender and intersectional analysis, it becomes evident that there is no objective basis for excluding gender identities, such as "women" or "non-binary individuals," as categorization criteria on Wikipedia. This finding highlights the presence of inconsistencies, such as the use of female-gendered first-level categories ("midwives" or "bearded women"), which persist within the platform.

Our study concludes by asserting that Wikipedia's categories hold significant potential for improvement, not only in addressing issues related to gender identity but also in enhancing the overall knowledge organization system for more effective user information retrieval. This recommendation stems from the observation that the vast majority of Wikipedias, with a few exceptions like the Catalan and Italian versions, have seamlessly incorporated gender identity categories into their organizational systems, thereby aiding content search and retrieval for users.

In light of these conclusions, we strongly recommend a comprehensive reevaluation of Wikipedia's content organization system. This reevaluation should focus on inclusivity, equity, and the fulfillment of users' information needs. Acknowledging the potential for the integration of gender identity as a valid classification criterion, Wikipedia can make substantial strides toward aligning its knowledge organization practices with contemporary principles of information access and inclusion.

Shifting our attention to the analysis of Wikidata, our investigation has focused on the technological aspects involved in the organization and retrieval of gender-diverse content within this platform. From this examination, several key conclusions have emerged.

First, our analysis has revealed that Wikidata exhibits a commendable level of sensitivity toward gender diversity, notably seen in the inclusion of a variety of gender categories under the P21 property.

Second, we recommend that Wikidata makes a clear distinction between properties related to biological sex and properties tied to gender identity, as the existing disjunctive labeling of the P21 property conflates these two distinct concepts.

Third, in contrast to Wikipedia, which often grapples with socio-cultural influences in decision-making processes, our analysis shows that Wikidata effectively mirrors real-world gender diversity without exacerbating existing biases as evidences by the research conducted by authors such as Zeng and Treviers (2023). The findings presented in this paper illustrate that Wikidata offers a richer array of tools to represent the diversity of gender identities.

Fourth, the Wikidata community tends to emphasize technical and data-centric arguments in its decision-making processes, diverging from Wikipedia's debates that often involve socio-cultural considerations, particularly regarding gender categories.

Lastly, the linguistic diversity of Wikidata poses unique challenges, particularly in languages where gender differentiation is significant. The debate over gender-neutral labelling in languages like Catalan underscores the importance of linguistic and cultural sensitivity in maintaining the dataset.

In conclusion, this analysis has predominantly delved into the technological aspects of enhancing the representation of gender diversity within Wikidata. However, it is imperative to recognize that a comprehensive solution necessitates a harmonious blend of technological enhancements and cultural considerations in the decision-making processes governing the organization of content in this vital knowledge-sharing platform. This convergence of technology and culture is paramount in fostering inclusivity and equity in the representation of gender diversity in the digital realm.

8.Research funding

This research received support from the Spanish Ministerio de Innovación, Ciencia y Universidades (MCIN) and the Agencia Estatal de Investigación [Grant ref. PID2020-116936RA-I00]. Additionally, it was funded by the Xarxa Vives d'Universitats, comprising 21 universities across Andorra, France, Italy, and Spain, in the Catalan language domain.

9.Bibliography

Abián, D., Meroño-Peñuela, A., and Simperl, E. (2022). An Analysis of Content Gaps Versus User Needs in the Wikidata Knowledge Graph. In Sattler U., Hogan A., Keet M., Presutti V., Almeida J.P.A., Takeda H., Monnin P., Pirrò G., and d'Amato C. (Eds.), *Lect. Notes Comput. Sci.: Vol. 13489 LNCS* (pp. 354–374). Springer Science and Business Media Deutschland GmbH; Scopus. https://doi.org/10.1007/978-3-031-19433-7_21

Acey, Camille E., Bouterse, Siko, Ghoshal, Sucheta, Menking, Amanda, Sengupta, Anasuya, and Vrana, Adele G. (2021). Decolonizing the Internet by Decolonizing Ourselves: Challenging Epistemic Injustice through Feminist Practice. *Global Perspectives*, *2*(1). https://doi.org/10.1525/gp.2021.21268

Aghaebrahimian, Ahmad, Stauder, Andy, and Ustaszewski, Michael. (2022). Testing the validity of Wikipedia categories for subject matter labelling of open-domain corpus data. *Journal of Information Science*, *48*(5), 686–700. https://doi.org/10.1177/0165551520977438

Albuquerque, Fernando Antônio de Araújo Chacon de. (2017). *Arcabouço de arquitetura da informação para ciclo de vida de projeto de vocabulário controlado: Uma aplicação em Engenharia de Software* [Fernando Antônio de Araújo Chacon de]. https://repositorio.unb.br/handle/10482/31288

Amith, Muhammad, He, Zhe, Bian, Jiang, Lossio-Ventura, Juan Antonio, and Tao, Cui. (2018). Assessing the practice of biomedical ontology evaluation: Gaps and opportunities. *Journal of Biomedical Informatics*, *80*, 1–13. https://doi.org/10.1016/j.jbi.2018.02.010

Antin, J., Yee, R., Cheshire, C., and Nov, O. (2011). *Gender differences in Wikipedia editing*. 11–14. https://doi.org/10.1145/2038558.2038561

Bolotnikova, E. S., Gavrilova, T. A., and Gorovoy, V. A. (2011). To a method of evaluating ontologies. *Journal of Computer and Systems Sciences International*, *50*(3), 448–461. https://doi.org/10.1134/S1064230711010072

Bourli, S., and Pitoura, E. (2020). Bias in Knowledge Graph Embeddings. In Atzmuller M., Coscia M., and Missaoui R. (Eds.), *Proc. IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Min., ASONAM* (pp. 6–10). Institute of Electrical and Electronics Engineers Inc.; Scopus. https://doi.org/10.1109/ASONAM49781.2020.9381459

Buchem, Ilona, and Kloppenburg, Julia. (2013). *Gender – Diversität – Wikipedia: Vielfalt gemeinsam gestalten*. Beuth Hochschule für Technik Berlin; Wikimedia Deutschland. https://www.bht-berlin.de/fileadmin/oe/gutz/Sonstige_Veroeffentlichungen/Arbeitspapier_Gender-Diversity-Wikipedia.pdf

Centelles, Miquel, and Ferran-Ferrer, Núria. (2024). Taxonomies and Ontologies in Wikipedia and Wikidata: An In-Depth Examination of Knowledge Organization Systems. [Manuscript submitted for publication]. *Hypertext.Net*, *27*.

Collier, Benjamin, and Bear, Julia. (2012). Conflict, Criticism, or Confidence: An Empirical Examination of the Gender Gap in Wikipedia Contributions. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 383–392. https://doi.org/10.1145/2145204.2145265

Conroy, M. (2023). Quantifying the Gap: The Gender Gap in French Writers' Wikidata. *Journal of Cultural Analytics*, *8*(2). Scopus. https://doi.org/10.22148/001c.74068

da Costa, Tiago Vinícius Remígio, Cavalcante, Everton, and Batista, Thais. (2022). Big Data

Software Architectures: An Updated Review. In Osvaldo Gervasi, Beniamino Murgante, Eligius M. T. Hendrix, David Taniar, and Bernady O. Apduhan (Eds.), *Computational Science and Its Applications – ICCSA 2022* (pp. 477–493). Springer International Publishing. https://doi.org/10.1007/978-3-031-10522-7_33

Das, Maitraye, Hecht, Brent, and Gergle, Darren. (2019). The Gendered Geography of Contributions to OpenStreetMap: Complexities in Self-Focus Bias. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3290605.3300793

Eckert, Stine, and Steiner, Linda. (2013). (Re)triggering backlash: Responses to news about Wikipedia's gender gap. *Journal of Communication Inquiry*, *37*(4), 284–303. https://doi.org/10.1177/0196859913505618

Evans, Sian, Mabey, Jacqueline, and Mandiberg, Michael. (2015). Editing for Equality: The Outcomes of the Art+Feminism Wikipedia Edit-a-thons. *Art Documentation*, *34*(2), 194–203. https://doi.org/10.1086/683380

Everett, Nikolas. (2015, March 5). Wikidata Query Backend Update (take two!). *Wikidata-Tech*. https://lists.wikimedia.org/hyperkitty/list/wikidata-tech@lists.wikimedia.org/message/VPQ226NBQ5D2ZCNUOHJL3X223Z4HUNJF/

Falenska, A, Cetinoglu, O, and Assoc Computat Linguist. (2021). *Assessing Gender Bias in Wikipedia: Inequalities in Article Titles* (WOS:000694722900009). 75–85.

Farzan, R., Savage, S., and Saviaga, C.F. (2016). *Bring on board new enthusiasts! A case study of impact of wikipedia art + feminism edit-a-thon events on newcomers: Vol. 10046 LNCS* (p. 40). Scopus. https://doi.org/10.1007/978-3-319-47880-7_2

Ferran-Ferrer, Núria, Castellanos-Pineda, Patricia, Minguillón, Julià, and Meneses, Julio. (2021). The gender gap on the Spanish Wikipedia: Listening to the voices of women editors. *Profesional de La Información*, *30*(5), Article 5. https://doi.org/10.3145/epi.2021.sep.16

Ferran-Ferrer, Núria, Centelles, Miquel, Macià, Yessica, Boté-Vericad, Juan-José, and Minguillon, Julià. (2023). *Dones de categoria: Anàlisi del biaix de gènere a les*

categories de Viquipèdia: Informe de diagnosi tècnica, posicionament acadèmic i proposta de millora del sistema d'organització del coneixement de Viquipèdia (p. 131). Programa d'Igualtat de Gènere de la Xarxa Vives d'Universitats.

Ford, Heather, and Wajcman, Judy. (2017). 'Anyone can edit', not everyone does: Wikipedia's infrastructure and the gender gap. *Social Studies of Science*, *47*(4), 511–527. https://doi.org/10.1177/0306312717692172

García Dauder, Silvia, and Pérez Sedeño, Eulalia. (2017). *Las "Mentiras" científicas sobre las mujeres*. Los Libros de la Catarata.

Gardner, Sue. (2011, February 20). Nine Reasons Women Don't Edit Wikipedia (in their own words). *Sue Gardner's Blog*. https://suegardner.org/2011/02/19/nine-reasons-why-women-dont-edit-wikipedia-in-their-own-words/

Grant, Maria J., and Booth, Andrew. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, *26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Gruwell, L. (2015). Wikipedia's politics of exclusion: Gender, epistemology, and feminist rhetorical (in)action. *Computers and Composition*, *37*, 117–131. https://doi.org/10.1016/j.compcom.2015.06.009

Hermoso Pulido, T. (2021). Simple Wikidata Analysis for Tracking and Improving Biographies in Catalan Wikipedia. *Web Conf. - Companion World Wide Web Conf., WWW*, 582–583. Scopus. https://doi.org/10.1145/3442442.3452344

Hinnosaar, Marit. (2019). Gender inequality in new media: Evidence from Wikipedia. *Journal of Economic Behavior & Organization*, *163*, 262–276. https://doi.org/10.1016/j.jebo.2019.04.020

Hollink, Laura, Van Aggelen, Astrid, and Van Ossenbruggen, Jacco. (2018). Using the Web of Data to Study Gender Differences in Online Knowledge Sources: The Case of the European Parliament. *Proceedings of the 10th ACM Conference on Web Science*, 381–385. https://doi.org/10.1145/3201064.3201108

Hube, C. (2017). *Bias in wikipedia*. 717–721. Scopus.

https://doi.org/10.1145/3041021.3053375

INE: Instituto Nacional de Estadística. (2024). INE. https://www.ine.es/

Ju, Boryung, and Stewart, Brenton. (2019). "The right information": Perceptions of information bias among Black Wikipedians. *Journal of Documentation*, *75*(6), 1486–1502. https://doi.org/10.1108/JD-02-2019-0031

Kaplan, Angelika, Kühn, Thomas, Hahner, Sebastian, Benkler, Niko, Keim, Jan, Fuchß, Dominik, Corallo, Sophie, and Heinrich, Robert. (2022). Introducing an Evaluation Method for Taxonomies. *Proceedings of the 26th International Conference on Evaluation and Assessment in Software Engineering*, 311–316. https://doi.org/10.1145/3530019.3535305

Karczewska, A., and Kukowska, K. (2021). *Cultural dimension of femininity: Masculinity in virtual organizing knowledge sharing*. 414–422. Scopus. https://doi.org/10.34190/EKM.21.135

Klein, M., Gupta, H., Rai, V., Konieczny, P., and Zhu, H. (2016). Monitoring the gender gap with Wikidata human gender indicators. *Proc. Int. Symp. Open Collab., OpenSym*. Proceedings of the 12th International Symposium on Open Collaboration, OpenSym 2016. Scopus. https://doi.org/10.1145/2957792.2957798

Klein, Max, and Konieczny, Piotr. (2015). Wikipedia in the world of global gender inequality indices: What the biography gender gap is measuring. *Proceedings of the 11th International Symposium on Open Collaboration*, 1–2. https://doi.org/10.1145/2788993.2789849

Kless, Daniel, and Milton, Simon. (2010). Towards Quality Measures for Evaluating Thesauri. In Salvador Sánchez-Alonso and Ioannis N. Athanasiadis (Eds.), *Metadata and Semantic Research* (pp. 312–319). Springer. https://doi.org/10.1007/978-3-642-16552-8_28

Konieczny, P., and Klein, M. (2018). Gender gap through time and space: A journey through Wikipedia biographies via the Wikidata Human Gender Indicator. *New Media and Society*, *20*(12), 4608–4633. Scopus. https://doi.org/10.1177/1461444818779080

Konieczny, Piotr. (2018). Volunteer Retention, Burnout and Dropout in Online Voluntary Organizations: Stress, Conflict and Retirement of Wikipedians. In Patrick G. Coy (Ed.), *Research in Social Movements, Conflicts and Change* (Vol. 42, pp. 199–219). Emerald Publishing Limited. https://doi.org/10.1108/S0163-786X20180000042008

Lam, Shyong (Tony) K., Uduwage, Anuradha, Dong, Zhenhua, Sen, Shilad, Musicant, David R., Terveen, Loren, and Riedl, John. (2011). WP:clubhouse?: An exploration of Wikipedia's gender imbalance. *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, 1–10. https://doi.org/10.1145/2038558.2038560

Laouenan, M., Bhargava, P., Eyméoud, J.-B., Gergaud, O., Plique, G., and Wasmer, E. (2022). A cross-verified database of notable people, 3500BC-2018AD. *Scientific Data*, *9*(1). Scopus. https://doi.org/10.1038/s41597-022-01369-4

Lemus-Rojas, M., and Lee, Y.Y. (2019). Using wikidata to provide visibility to women in STEM. *Proc. Int. Conf. Dublin Core Metadata Appl.*, 126–131. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85088230329&partnerID=40&md5=e37fc07992e9f29aa3de487bb6252e36

Malyshev, Stanislav, Krötzsch, Markus, González, Larry, Gonsior, Julius, and Bielefeldt, Adrian. (2018). Getting the Most Out of Wikidata: Semantic Technology Usage in Wikipedia's Knowledge Graph. In Denny Vrandečić, Kalina Bontcheva, Mari Carmen Suárez-Figueroa, Valentina Presutti, Irene Celino, Marta Sabou, Lucie-Aimée Kaffee, and Elena Simperl (Eds.), *The Semantic Web – ISWC 2018* (Vol. 11137, pp. 376–394). Springer International Publishing. https://doi.org/10.1007/978-3-030-00668-6_23

Mandiberg, M., and Sarıoğlu, D. (2022). Clowns in the Visual Artists: Topic Modeling Wikipedia and Wikidata. *Art Documentation*, *41*(1), 20–37. Scopus. https://doi.org/10.1086/719999

Mazzocchi, Fulvio. (2018). Knowledge Organization System (KOS): An Introductory Critical Account. *Knowledge Organization*, *45*(1). https://doi.org/10.5771/0943-7444-2018-1-54

Miquel-Ribe, M, and Laniado, D. (2021). The Wikipedia Diversity Observatory: Helping communities to bridge content gaps through interactive interfaces. *JOURNAL OF INTERNET SERVICES AND APPLICATIONS*, *12*(1). https://doi.org/10.1186/s13174-021-00141-y

Mora-Cantallops, Marçal, Sánchez-Alonso, Salvador, and García-Barriocanal, Elena. (2019). A systematic literature review on Wikidata. *Data Technologies and Applications*, *53*(3), 250–268. https://doi.org/10.1108/DTA-12-2018-0110

Morgan, J.T., Bouterse, S., Stierch, S., and Walls, H. (2013). Tea & sympathy: Crafting positive new user experiences on wikipedia. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, 839–848. Scopus. https://doi.org/10.1145/2441776.2441871

Pellissier Tanon, T., and Suchanek, F. (2019). Querying the Edit History of Wikidata. In Hitzler P., Kirrane S., Hartig O., de Boer V., Schlobach S., Vidal M.-E., Maleshkova M., Hammar K., Lasierra N., Stadtmüller S., Hose K., and Verborgh R. (Eds.), *Lect. Notes Comput. Sci.: Vol. 11762 LNCS* (pp. 161–166). Springer Science and Business Media Deutschland GmbH; Scopus. https://doi.org/10.1007/978-3-030-32327-1_32

Souza, Renato Rocha, Tudhope, Douglas, and Almeida, and Maurício Barcellos. (2012). Towards a Taxonomy of KOS: Dimensions for Classifying Knowledge Organization Systems. *Knowledge Organization*, *39*(3), 179–192. https://doi.org/10.5771/0943-7444-2012-3-179

Thornton, K., and Seals-Nutt, K. (2018). Science stories: Using IIIF and wikidata to create a linked-data application. In Srinivas K., Fortuna C., Atre M., van Erp M., and Lopez V. (Eds.), *CEUR Workshop Proc.* (Vol. 2180). CEUR-WS; Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85055351166&partnerID=40&md5=2a141ac8b64f5e4eb0048f128ed06c3b

Thornton, K., Seals-Nutt, K., Van Remoortel, M., Birkholz, J.M., and De Potter, P. (2022). Linking women editors of periodicals to the Wikidata knowledge graph. *Semantic*

Web, *14*(2), 443–455. Scopus. https://doi.org/10.3233/SW-222845

Tripodi, Francesca. (2023). Ms. Categorized: Gender, notability, and inequality on Wikipedia. *New Media & Society*, *25*(7), 1687–1707. https://doi.org/10.1177/14614448211023772

Vrandečić, Denny, Pintscher, Lydia, and Krötzsch, Markus. (2023). Wikidata: The Making Of. *Companion Proceedings of the ACM Web Conference 2023*, 615–624. https://doi.org/10.1145/3543873.3585579

Wagner, Claudia, Graells-Garrido, Eduardo, Garcia, David, and Menczer, Filippo. (2016). Women through the glass ceiling: Gender asymmetries in Wikipedia. *EPJ DATA SCIENCE*, *5*. https://doi.org/{10.1140/epjds/s13688-016-0066-4}

Wikidata. (2024). *Property talk:P21*. https://www.wikidata.org/wiki/Property_talk:P21

Wikimedia. (2015, November 17). *Categoria:Plantilles de manteniment per a categories*. https://ca.wikipedia.org/w/index.php?title=Categoria:Plantilles_de_manteniment_per_a_categories&oldid=16026819

Wikimedia. (2018, November 24). *Ajuda:Categoria*. https://ca.wikipedia.org/w/index.php?title=Ajuda:Categoria&oldid=20513864

Wikimedia. (2022). *Wikidata:WikiProject Ontology/Classes*. https://www.wikidata.org/wiki/Wikidata:WikiProject_Ontology/Classes

Wikimedia. (2023a). *Wikidata:Accés a les dades*. https://www.wikidata.org/wiki/Wikidata:Data_access/ca

Wikimedia. (2023b). *Wikidata:Bots*. https://www.wikidata.org/wiki/Wikidata:Bots

Wikimedia. (2023c). *Wikimedia Statistics—Catalán Viquipèdia*. https://stats.wikimedia.org/#/ca.wikipedia.org

Wikimedia. (2023d, October 23). *Wikipedia:Categorization*. https://en.wikipedia.org/w/index.php?title=Wikipedia:Categorization&oldid=1181497476

Wilson, R. Shyama I., Goonetillake, Jeevani S., Ginige, Athula, and Indika, Walisadeera Anusha. (2022). Ontology Quality Evaluation Methodology. In Osvaldo Gervasi,

Beniamino Murgante, Eligius M. T. Hendrix, David Taniar, and Bernady O. Apduhan (Eds.), *Computational Science and Its Applications – ICCSA 2022* (pp. 509–528). Springer International Publishing. https://doi.org/10.1007/978-3-031-10522-7_35

Worku, Zena, Bipat, Taryn, McDonald, David W., and Zachry, Mark. (2020). Exploring Systematic Bias through Article Deletions on Wikipedia from a Behavioral Perspective. *Proceedings of the 16th International Symposium on Open Collaboration*, 1–22. https://doi.org/10.1145/3412569.3412573

Zeng, Marcia Lei, and Mayr, Philipp. (2018). Knowledge Organization Systems (KOS) in the Semantic Web: A Multi-Dimensional Review. *International Journal on Digital Libraries*, 1–22.

Zhang, Charles Chuankai, and Terveen, Loren. (2021). Quantifying the Gap: A Case Study of Wikidata Gender Disparities. *17th International Symposium on Open Collaboration*, 1–12. https://doi.org/10.1145/3479986.3479992

Zheng, Xiang, Chen, Jiajing, Yan, Erjia, and Ni, Chaoqun. (2022). Gender and country biases in Wikipedia citations to scholarly publications. *Journal of the Association for Information Science and Technology*, *74*(2), 219–233. Scopus. https://doi.org/10.1002/asi.24723

Zhu, L., Xu, A., Deng, S., Heng, G., and Li, X. (2023). Entity Management Using Wikidata for Cultural Heritage Information. *Cataloging and Classification Quarterly*, *61*(1), 20–46. Scopus. https://doi.org/10.1080/01639374.2023.2188338

Journal of Documentation

Assessing Knowledge Organization Systems from a gender perspective: Wikipedia Taxonomy and Wikidata Ontologies

Centelles, Miquel & Ferran-Ferrer, Núria

Abstract

Purpose
Develop a comprehensive framework for assessing the knowledge organization system (KOS), including the taxonomy of Wikipedia and the ontologies of Wikidata, with a specific focus on enhancing management and retrieval from a non-binary gender perspective.
Design/methodology/approach
This study employs heuristic and inspection methods for assessment. A method is designed to inspect Wikipedia's Knowledge Organization Systems (taxonomy), ensuring their compliance with established specifications and international standards. Additionally, an evaluation is conducted to gauge the effectiveness and efficiency of retrieving articles related to women and non-masculine genders using the Catalan category scheme (taxonomy of Viquipèdia), with a focus on identifying its limitations. Furthermore, the research includes a novel evaluation of the Wikidata ontologies in terms of their structure and coverage of gender-related properties and classes. This evaluation includes a comparative analysis with the Wikipedia category scheme (taxonomy) to discern the advantages and enhancements it offers.
Findings
This study evaluates Wikipedia's taxonomy and Wikidata's ontologies, establishing evaluation criteria for gender-based categorization and exploring their structural effectiveness. The evaluation process suggests that Wikidata ontologies may offer a viable solution to address Wikipedia's categorization challenges.
Originality/value
The assessment of Wikipedia categories (taxonomy) based on Knowledge Organization System standards leads to the conclusion that there is ample room for improvement, not only in matters concerning gender identity but also in the overall knowledge organization system to enhance search and retrieval for users. These findings bear relevance for the design of tools to support information retrieval on knowledge-rich websites, as they assist users in exploring topics and concepts.

Keywords:
Knowledge Organization System (KOS); Taxonomy; Ontology; Wikipedia; Wikidata; Gender perspective; Heuristic Methods; Inspection Methods; Knowledge Organization Standards; Comparative Analysis; Gender-Based Knowledge Organization System; Information Retrieval

1. Introduction

Wikipedia is a widely used educational resource with billions of readers in numerous languages, created through open collaboration. Despite its achievements, Wikipedia suffers from a persistent gender bias with a low percentage of content on women and few female editors (Hinnosaar, 2019; Wagner et al., 2016). This gender bias is exacerbated in some Wikipedia editions, such as the Italian or Catalan versions, due to decisions about gender-related categories that should provide access and visualization of content related to gender identities. In these cases, categories like "woman" or "non-binary person" are prohibited for the organization of content and thus information retrieval. These community based decisions lead to some dysfunctions, which are particularly critical in languages that use grammatical gender, such as Catalan and Italian. Addressing this bias is important for providing equitable information retrieval and knowledge representation.

In the digital age, Knowledge Organization Systems (KOS) encompass a range of critical tools such as classification systems, thesauri, lexical databases, ontologies, gazetteers, and taxonomies. These KOS have assumed an increasingly pivotal role in the realm of information management and diverse applications. Their primary function is to meticulously convey semantics, accomplishing a multifaceted array of functions.

First and foremost, KOSs are indispensable for representing and indexing information and documents. They provide a structured framework that aids in the organization and retrieval of information. Furthermore, KOS act as knowledge-based assistants for information seekers, guiding them through the intricacies of data. They serve as semantic guides across various domains and fields, facilitating a deeper understanding of complex subject matter. In addition, KOSs function as communication tools, furnishing a conceptual framework that bridges the gap between experts and non-experts, ensuring a common language for effective communication. Moreover, they offer a foundational structure for knowledge-driven systems, enabling the seamless integration of data and knowledge in various applications (Zeng and Mayr, 2018).

KOSs are pivotal in structuring and classifying vast amounts of information in our digital age. Prominent examples of these systems can be found in Wikipedia and Wikidata. However, evaluating these knowledge organization structures, known as taxonomies in Wikipedia and ontologies in Wikidata, remains a complex challenge. There is currently no established methodology for determining the optimal indicators and metrics required for the comprehensive assessment of these structures. The creation of these metrics often relies on the specific context of the study, which can introduce subjectivity and inconsistency into the assessment process.

This academic paper conducts an in-depth examination of taxonomies and ontologies in Wikipedia and Wikidata. The primary objective is to establish a methodology for evaluating these systems, quantifying categorization issues in Wikipedia, and assessing Wikidata's suitability. It also aims to reduce the gender gap on Wikipedia by visualizing gender diversity from Wikidata. While Wikipedia has limited gender categories, Wikidata provides a broader range, including agender, intersex, non-binary, transgender, and more. The connection between Wikipedia and Wikidata is notable.

Wikidata faces a unique challenge in structuring gender data. While Wikipedia confines itself to male and female categories (in some editions, only the male category), Wikidata's property 21 encompasses a wide array of gender classes, including agender, female, male, intersex, non-binary, transgender female, and transgender male, among others. A pre-existing connection exists between Wikipedia and Wikidata, with Wikidata serving as an integral component of Wikipedia's infrastructure. Furthermore, the utilization of ontologies to enhance information organization and retrieval in Wikipedia is evident in specific cases, such as the management of the "living people" category.

In this challenge about gender data, an essential discussion concerning gender and sex, particularly regarding Property talk:P21 (Wikidata, 2024) has surfaced. Concerns have arisen regarding the conflation between sex and gender within a single category on Wikidata. There is a call for distinct properties to differentiate sex, gender, and potentially gender identity, similar to having separate properties for height and weight. The text highlights issues related to the vague and unclear classification of terms like male, female, man, and woman. It is suggested to have a separate property and values for gender identity, distinct from biological sex, with clear and unambiguous definitions to avoid intentional conflation that causes problems with dataset clarity and unambiguous representation. Furthermore, it discusses similar concerns in official contexts, such as the discussion initiated by the UK government in 2018 regarding managing gender or sex statements, indicating parallel challenges faced by both Wikidata and the government.

Additionally, unresolved situations related to the assignment of property values are pointed out, including issues with assigning "male" to someone who is biologically female, questioning the differentiation between human males and non-human males, confusion between transsexualism and Gender Identity Disorder (GID), and the need for more accurate representation of values such as "intersex" and "transgender." Furthermore, it is noted that special situations, such as assigning gender to anthropomorphic nonhumans and dealing with unknown gender, have not been adequately resolved. The necessity of incorporating a "citation needed" constraint to the property, requiring at least one reference for value assignment, is also analysed.

On a related note, the inappropriate addition of sex or gender statements for living individuals via Quickstatements or bots on Wikidata, leading to harmful misgendering and potential privacy violations, is brought up. Proposed solutions to prevent future harm include disallowing bots and Quickstatements from affecting more than ten items at a time and discouraging the use of labels and given names as references for sex and gender statements. These proposals aim to ensure more careful handling of sex and gender statements to avoid harm and privacy violations, reflecting the community's concerns with promoting responsible and ethical practices on Wikidata.

Finally, this ~~T~~he paper evaluates the Knowledge Organization System (KOS) using the Catalan Wikipedia as a case study on the gender gap. It seeks to improve gender identity visualization and accessibility through Wikidata ontologies. It acknowledges potential biases in Wikidata and Wikipedia and their capacity to perpetuate real-world biases. Furthermore, it is essential to acknowledge that Wikidata's potential biases are no greater than those present in the real world (Zhang and Terveen, 2021). Additionally, some authors argue that Wikipedia mirrors real-world biases (Eckert and Steiner, 2013) with the platform having the

Commented [1]: CITA?

capacity to perpetuate and exacerbate gender gaps, shaped not only by editors but also by infrastructural logics (Ford and Wajcman, 2017).

The objective is to evaluate Wikipedia's taxonomy and Wikidata's ontologies to enhance gender diversity visibility. The paper synthesizes theories and insights to establish comprehensive evaluation criteria. The ultimate aim is to provide an objective approach to assess knowledge organization systems in Wikipedia and Wikidata and quantify their structural effectiveness. Subsequent sections will detail the evaluation process and findings, addressing Wikidata's potential as a solution for Wikipedia's categorization challenges.

2. Literature Review~~Theoretical Framework~~

~~The gender gap on Wikipedia has been the subject of extensive academic research, with numerous studies exploring biases in content, participation, reading, and potential strategies to address this gap. These studies emphasize the importance of recognizing and addressing biases and barriers to create a more diverse and inclusive Wikipedia community (Ferran-Ferrer et al., 2023).~~

In this section, we provide a comprehensive review following the SALSA framework (Grant and Booth, 2009) to examine the gender gap in Wikipedia and Wikidata. Academic research has extensively investigated the gender gap in both platforms. The Wikipedia appraisal stage involved 97 articles, and the Wikidata appraisal involved 34. A total amount of 21 articles were used to assess Wikipedia (Ferran-Ferrer et al., 2023), and 19 were used to evaluate Wikidata.

2.1. Gender gap in Wikipedia

The gender gap on Wikipedia has been the subject of extensive academic research, with numerous studies exploring biases in content, participation, reading, and potential strategies to address this gap. These studies emphasize the importance of recognizing and addressing biases and barriers to create a more diverse and inclusive Wikipedia community (Ferran-Ferrer et al., 2023).

The under-representation of women as editors and as subjects of biographical coverage is a widely recognized issue in the academic field (Hube, 2017; Falenska et al., 2021). Some articles discuss how gender bias intersects with race, sexuality, security, and marginalization on Wikipedia (Lam et al., 2011; Ju and Stewart, 2019; Tripodi, 2023). Various factors, such as the demographics of editors, platform structure, and cultural values, contribute to these biases, which have significant social implications, affecting the visibility and participation of women and perpetuating existing disparities (Ford and Wajcman, 2017).

Regarding the gender gap in content, research reveals that women are underrepresented among the main figures in all language editions of Wikipedia (Miquel-Ribe and Laniado, 2021). Articles for deletion is a possibility within the decision-making process in Wikipedia article editing. It is the process that determines what constitutes knowledge and what does not in the encyclopedia. Biographies of women and LGBTQ+ individuals are often subject to deletion, resulting in a higher proportion of biographies of women nominated for deletion compared to biographies available about men~~women~~ (Morgan et al., 2013; Hollink et al.,

2018; Tripodi, 2023). ~~However, content that may be of interest to men appears to be more likely to be nominated for speedy deletion.~~ While there are indications of bias, some authors conclude that there is no clear bias resulting from deletion activity (Worku et al., 2020).

Studies also identify significant gender differences in Wikipedia content, such as biographies of women featuring more prominent family, gender, and relationship themes (Wagner et al., 2016). Linguistic bias in terms of language abstraction and positivity can be observed, along with structural differences in metadata and hypertext links. In addition, citation practices reveal that female authors are cited less than expected, suggesting a preference for citing male publications (Zheng et al., 2022). These biases may further marginalize female authors, especially in non-Anglophone countries. The gender gap in content creation and participation on Wikipedia perpetuates an unbalanced coverage of topics, creating a cycle where the lack of diversity in content fails to attract and engage different editors, thus exacerbating the existing gender gap (Konieczny and Klein, 2018).

Research on the gender gap in editing and participation highlights various barriers that hinder women's involvement on Wikipedia. These barriers include negative reputation, lack of recognition, fear of deletion, rejection, and alienation. Often, research suggests that women lack confidence in their abilities, feel uncomfortable with editing, and face negative responses to constructive feedback (Collier and Bear, 2012). Factors such as the digital skills gap (Gardner, 2011) and the availability of time for editing (Gruwell, 2015) also contribute to the gender gap. However, visible female editors and constructive comments can help mitigate the gap, as the presence of visible female peers promotes collaborative editing (Evans et al., 2015). Some authors have investigated the gender gap in Germany and suggested a proactive approach to training and educating women to enhance their motivation for writing (Buchem and Kloppenburg, 2013). It has also been~~They have~~ highlighted the impact of family responsibilities on women's ability to write, so efforts may need to focus on addressing gender disparities in domestic work (Ferran-Ferrer et al., 2021).

The gender gap extends beyond editing and includes the underrepresentation of individuals. Female participation varies by topic, with a greater presence in gender studies or feminism categories, reflecting traditional gender stereotypes. Generic site restrictions limit the digital credibility and authority of women, hindering their contributions. The complex relationship between the gender gap and harassment requires better understanding, and it is important to create a safe environment for women on and off Wikipedia. Feminist interventions, such as exclusive edit-a-thons for women, have proven effective in countering gender inequality on the platform.

~~with non-heterosexual orientations (Miquel-Ribe and Laniado, 2021). Access to reliable citations that comply with Wikipedia policies presents challenges in the editing process. Female administrators play a significant role in fostering an inclusive environment on Wikipedia, and there are notable differences in their approach compared to male administrators (Farzan et al., 2016; Karczewska & Kukowska, 2021). Female participation varies by topic, with a greater presence in gender studies or feminism categories, reflecting traditional gender stereotypes. Generic site restrictions limit the digital credibility and authority of women, hindering their contributions. The complex relationship between the gender gap and harassment requires better understanding, and it is important to create a safe environment for women on and off Wikipedia. Feminist interventions, such as exclusive~~

~~edit-a-thons for women, have proven effective in countering gender inequality on the platform.~~

2.2. Gender and Wikidata

The gender gap on Wikidata has been extensively explored in academic research. We can delineate three main categories of studies. A first set of research has delved into the gender gap within Wikidata, presenting diverse **methodologies, findings, and recommendations** to address this disparity. Meanwhile, a second set aims to **quantitatively assess the biographical gender gap** in Wikipedia, across various language editions, leveraging Wikidata's multilingual support to facilitate this cross-cultural research. Lastly, a third set of studies emphasise the **advocacy and visibility of content pertaining to women in industries traditionally dominated by men**, utilising Wikidata for this purpose.

Regarding the initial group of discussions aimed at presenting diverse methodologies, findings, and recommendations to address this disparity, Zhang and Terveen (2021) delved into the gender content gap in Wikidata, seeking to uncover the source of bias. Through a quantitative case study, they examined how individuals were represented in Wikidata compared to existing gender biases. Their findings revealed a prevalence of male-dominated professions among the most frequently represented categories, closely mirroring real-world gender distribution.

Similarly, Abián, Meroño-Peñuela and Simperl (2022) sought to understand the impact of content gaps in knowledge graphs on downstream applications, with a particular focus on gender disparities within Wikidata. To achieve this, they introduced a framework that compared edit metrics with Wikipedia pageviews, facilitating a quantitative evaluation of discrepancies between knowledge graph content and user needs. As a result, they identified no inherent gender or recency gaps within Wikidata's production, with only a few under-represented entities standing out. A group of articles has focused on analysing gender bias on Wikidata concerning occupations or professional domains. In this line, Das et al. (2019) conducted a holistic analysis of bias measurement on the knowledge graph, specifically focusing on biases in Wikidata across different demographics selected from seven continents. They utilised extensive experiments on a wide range of occupations sampled from various demographics, examining the impact of algorithm bias on the measurement of biassed occupations. Results indicated that the inherent data bias in Wikidata can be influenced by specific algorithm bias and underscored the importance of understanding biases based on socio-cultural differences across demographics. Within this same field, there are three works that concentrate on specific occupations or professional domains:

Lemus-Rojas and Lee (2019) in the STEM fields, Zhu et al. (2023) in Chinese culture and heritage, and Conroy (2023) in French and Francophone literature. The outcomes align with the conclusions observed in the aforementioned comprehensive studies. In the first two cases, Wikidata is highlighted as a critical collection for enhancing the visibility of women. Conroy (2023) found that the gender gap in both subsets closely resembles the global average, with a higher-than-average representation of writers of other genders.

Finally, Pellissier and Suchanek (2019) and Bourli and Pitoura (2020) analysed gender bias on Wikidata through advanced automated processing techniques. Pellissier and Suchanek

(2019) proposed a system to index changes in the Wikidata graph and enable users to answer complex SPARQL queries regarding historical changes, while Bourli and Pitoura (2020) introduced measures for identifying bias in the dataset, tested methods for amplifying bias in embeddings, and introduced a debiasing approach. A special case is Mandiberg and Sarıoğlu (2022), who aimed to address the challenges associated with defining a dataset to analyse changes in Wikipedia's gender gap for articles about visual art. The dataset is constructed from the intersection between Wikipedia and Wikidata. The researchers describe the process of using a topic model algorithm to identify a dataset by analysing the words within each article and grouping articles into topics. Their aim was to create a dataset that more closely reflects visual artists' articles on English Wikipedia, addressing potential systemic biases. The topic model algorithm provided a dataset that encompassed a majority of the two WikiProject datasets and the Wikidata sets, while adding additional art-related individuals. It was found to be superior to other options, offering a detailed list of articles about visual arts that mitigated Wikipedia's existing imbalances. The study also highlighted challenges in Wikidata's taxonomies and called for further research on systemic biases reflected in taxonomy systems.

A second set of articles addresses the application of Wikidata, capitalizing on its multilingual capabilities to facilitate comprehensive cross-cultural research, for measuring gender bias in Wikipedia editions and for resolving this issue. Three of these studies feature contributions from Maximilian Klein and Piotr Konieczny. Klein and Konieczny (2015) and Konieczny and Klein (2018) introduce the Wikipedia Gender Inequality Indicator (WIGI) developed from Wikidata. WIGI calculates, for each country, a score based on the ratio of female and nonbinary gendered biographies to the total number of biographies. This Wikipedia-derived indicator is correlated with four contemporary, widespread gender inequality indices (GDI, GEI, GGGI, and SIGI). Through analysing methodologies and the relationship with Wikipedia data, evidence suggests that the bias in Wikipedia's biographical coverage is aligned with gender bias in socially powerful positions. Concerning the results, Klein and Konieczny (2015) find that the strongest correlations are with individuals born around 1910, indicating that Wikipedia's representation may more accurately reflect current rather than historical gender statuses. The same authors Konieczny and Klein (2018) utilise cultural clusters to highlight how gender inequality can be examined through diverse cultural perspectives.

Klein et al. (2016) delves deeper into the gender bias of content, focusing on women's biographies on Wikipedia. The article underscores the importance of precisely measuring the gender content gap and the critical examination of initiatives intended to mitigate this disparity. The team formulates the Wikidata Human Gender Indicators (WHGI), a robust, longitudinal dataset to monitor gender disparities. It monitors biographical data across multiple facets—such as time, geography, culture, occupation, and language—providing an extensive instrument for elucidating and quantifying the gender bias in Wikipedia's content. The research signals a changing representation of women in 11 dimensions utilising WHGI. Validations against three external datasets back the indicator's accuracy, and reassessment of Wikipedia's gender bias with WHGI suggests that it could enhance depth and impact in future research on the subject.

In a similarly line of work, Hollink et al. (2018) tackles the challenge of measuring gender inequalities on Wikipedia, especially when considering multiple languages. The difficulty in finding objective methods to measure and compare gender inequality is underlined, and the

potential differences across language editions of Wikipedia are acknowledged. Their methodology focuses on comparing coverage of male and female Members of the European Parliament (MEP) across various Wikipedia language editions using open data. This approach allows for a fair comparison due to the MEPs' notable actions in the real world, and it examines gender discrepancies in both the coverage on Wikipedia and the content within Wikidata entries. An analysis of Wikidata entries for male and female MEPs reveals equal amounts of property-value pairs, contradicting earlier studies that found Wikipedia content related to women emphasised family and relationships. Differences related to real-world disparities suggest that the structured data of Wikidata might be less prone to bias. Moreover, aggregation of data from various Wikipedia language editions might contribute to a more diversified and equitable dataset in Wikidata.

Delving into the characteristics and virtues of Wikidata, Hermoso Pulido (2021) discusses how Wikidata has become a significant tool within the Wikimedia ecosystem, improving data linkage and reuse. Specifically, it mentions the adoption of Wikidata in Catalan Wikipedia, noting how its integration with infoboxes and list generation has advanced the project. The article suggests that such technical innovations could be part of the solution in addressing Wikipedia's gender gap. Methodology highlights the use of structured data from Wikidata to evaluate new biographical articles, aiming to encourage user engagement in diversity issues and track vandalism or errors. This methodology suggests a proactive approach to using structured data for maintaining quality and diversity in biographical content, directly impacting the reduction of Wikipedia's gender gap. Technical challenges are highlighted, such as execution timeouts during SPARQL queries for live data analysis. While some limitations exist for large datasets, initiatives like WCDO show promise in identifying and acting upon content gaps. The article advocates for enhanced cross-collaboration between Wikidata and Wikipedia, suggesting that embedding certain tools could encourage editors to address discrepancies more effectively.

Leveraging the potential of Wikidata, Laouenan et al. (2022) focus on studying different intersectionalities, specifically, they aim to construct a comprehensive and accurate database of notable individuals by cross-verifying the information from various editions of Wikipedia and Wikidata, focusing on specific social science questions about gender, economic growth, urban and cultural development. The researchers collected a significant amount of data from Wikipedia and Wikidata, utilising deduplication techniques and cross-verifying the retrieved information. They found varying degrees of completeness and error rates dependent on notability distribution, classifying the presence of an Anglo-Saxon bias in the English edition of Wikipedia. The strategy resulted in the creation of a cross-verified database of 2.29 million individuals, shedding light on an Anglo-Saxon bias in the English edition of Wikipedia. The study also emphasised the implications of this bias and identified individuals not present in the English edition of Wikipedia.

Finally, the last research strand in this set of The last set of six papers aims to emphasize is focused highlight the promotion and visibility of content related to women in male-dominated professional spheres through the utilisation of Wikidata. Among these, two articles are authored by Thornton and Seals-Nutt, both affiliated with the Stories Services Collaborative. Thornton and Seals-Nutt (2018) introduce the creation of a web application called Science Stories. This application utilises structured data from Wikidata along with images to narrate compelling science stories, especially focusing on the experiences of women who have

contributed to scientific research. The primary goal is to elevate the visibility of these women. The authors illustrate how the use of free software and open standards can lead to the development of visually captivating and interactive science communication experiences. These experiences involve the integration of images with structured statements within a web of interconnected data, all supported by references to published sources. Four articles focus on leveraging Wikidata to promote and illuminate the contributions of women in male-dominated professional fields. In a similar vein, Thornton et al. (2022) delve into how Semantic Web capabilities can consolidate disparate materials to craft narratives, as demonstrated by the WeChangEd research project, which centres on women editors of periodicals in Europe from 1710-1920. The methodology involves developing applications that aggregate data from Wikidata to harness a versatile knowledge graph, facilitating the swift creation of interactive platforms to captivate fresh audiences. The outlined process holds potential value for researchers and cultural heritage institutions seeking web-based avenues for presenting data-driven storytelling.

Finally, the last research strand in this set of papers aims to emphasize the promotion and visibility of content related to women in male-dominated professional spheres through the utilisation of Wikidata. Among these, two articles are authored by Thornton and Seals-Nutt, both affiliated with the Stories Services Collaborative. Thornton and Seals-Nutt (2018) introduce the creation of a web application called Science Stories. This application utilises structured data from Wikidata along with images to narrate compelling science stories, especially focusing on the experiences of women who have contributed to scientific research. The primary goal is to elevate the visibility of these women. The authors illustrate how the use of free software and open standards can lead to the development of visually captivating and interactive science communication experiences. These experiences involve the integration of images with structured statements within a web of interconnected data, all supported by references to published sources. Four articles focus on leveraging Wikidata to promote and illuminate the contributions of women in male-dominated professional fields. In a similar vein, Thornton et al. (2022) delve into how Semantic Web capabilities can consolidate disparate materials to craft narratives, as demonstrated by the WeChangEd research project, which centres on women editors of periodicals in Europe from 1710-1920. The methodology involves developing applications that aggregate data from Wikidata to harness a versatile knowledge graph, facilitating the swift creation of interactive platforms to captivate fresh audiences. The outlined process holds potential value for researchers and cultural heritage institutions seeking web-based avenues for presenting data-driven storytelling.

~~Efforts to address the gender gap and marginalization on Wikipedia involve various interventions and considerations. Female mentoring is identified as a crucial factor in promoting women's inclusion. However, it is also important to address gender disparities among influential editors who shape policies and perform higher-level tasks. While progress has been made in attracting more women to Wikipedia, achieving greater gender parity among influential editors is necessary (Antin et al., 2011).~~ ~~Female administrators play a significant role in fostering an inclusive environment on Wikipedia, and there are notable differences in their approach compared to male administrators (Farzan et al., 2016; Karczewska & Kukowska, 2021).~~

~~The gender gap on Wikipedia has broad and significant consequences that go beyond the digital world. It reinforces existing gender stereotypes, perpetuating the idea that men's contributions are more valuable and notable than women's. This biased representation not only distorts our understanding of history but also hinders progress toward gender equality. By erasing women's contributions from historical records, we reinforce the notion that women have not played a significant role in shaping our societies, thus perpetuating social inequality in the process.~~

~~The gender gap on Wikipedia is a matter of concern due to its far-reaching consequences beyond online representation. This gap not only perpetuates gender stereotypes but also erases women's contributions from historical records and reinforces social inequality (Acey et al., 2021). Biased algorithms further exacerbate the problem by making it difficult to find information about women, undervaluing their achievements in the process (Wagner et al., 2016). To effectively address this gender gap, it is essential to allocate sufficient resources to include diverse perspectives and ensure fair knowledge representation on the platform. In this way, we can actively fight against discrimination and promote a more equitable production of knowledge on Wikipedia.~~

~~To address this gender gap, it is crucial to allocate sufficient resources to include diverse perspectives and ensure equitable representation on the platform. This involves actively promoting women's participation as editors and expanding content related to women. It is also necessary to critically examine algorithms and classification practices to ensure that they do not perpetuate gender biases. A collective effort is needed to address this gap and work toward a more equitable and representative knowledge production on Wikipedia.~~

~~Regarding the gender gap in Wikidata, there are few studies about it. In fact, the database emerged in late 2012, and the first authors to approach the topic were Klein et al. (2016) and Konieczny (2018). These latter authors pointed out the difficulty of measuring the representation or absence of women's biographies on the encyclopedia and, motivated by this concern, they generated the "Wikidata Human Gender Indicators" (WHGI) based on Wikidata data, which provides a longitudinal, biographical dataset across time and space, cultures, languages, and professions, among 11 total study elements (Konieczny and Klein, 2018; Zhang and Terveen, 2021). This research confirms, with data, that the gender gap is a phenomenon that has occurred throughout history and across cultures. Several authors also agreed that Wikidata offers solutions to the study of the gender gap on Wikipedia (Mora-Cantallops et al., 2019; Zhang and Terveen, 2021). Thus, Zhang and Terveen worked to quantify the gender gap and were able to specify that the representation of women in Wikidata were more aligned with reality than Wikipedia and did not exacerbate the gender bias.~~

~~Citar article d'Hipertext.net Taxonomies and Ontologies (Centelles and Ferran-Ferrer, 2024). This article delves into the knowledge organization systems (KOS) of both Wikipedia and Wikidata, scrutinizing their structures, functions, and relationships, with a focus on gender-related content classification in the Catalan edition of Wikipedia.~~

3.Objectives

The main aim of this research is to explore and compare the effectiveness and efficiency of the KOS of female biographies on Wikipedia and non-male ones. This will be accomplished by evaluating the category structure of the Catalan edition of Wikipedia and the ontology of Wikidata, with the aim of addressing the challenge of visualizing the diversity of gender identities and accessing their content on Wikipedia. We will aim to ascertain whether Wikidata ontologies can offer a more improved means of organizing and representing the information available on Wikipedia regarding the diversity of gender identities.

Therefore, the research questions that we will address are:
QR1: How can a standards inspection method be developed to evaluate the conformance of the KOS in Wikipedia with international specifications and standards established by recognized organizations?
QR2: How does the category scheme of the Catalan edition of Wikipedia impact the effectiveness and efficiency of retrieving articles related to women and non-male genders, and what specific limitations does it present?
QR3: To what extent does the Wikidata ontology facilitate the effective and efficient retrieval of articles concerning women and non-male genders, and what advantages or enhancements does it offer in comparison to the Wikipedia category scheme?

To address these questions, a specific methodology is created and applied for each of the specific objectives (See Table 1):

INSERT TABLE 1

4.Methodology

To explore the nature of Wikipedia as a taxonomy, as opposed to a folksonomy, and provide insights into Wikidata's data model, it is documented in Centelles and Ferran-Ferrer (2024).

4.1. Inspection of standards and guidelines for the evaluation of taxonomy (Wikipedia) and ontologies (Wikidata)

Our study begins by reviewing the most widely accepted standards for the analysis of KOSs, and using them as the basis for designing an evaluation guide tailored to the taxonomic and ontological criteria relevant to Wikipedia and Wikidata. Subsequently, we employed a standards inspection method to assess whether the KOSs of Wikipedia and Wikidata conform to the international specifications and standards defined by recognized organizations.

In the theoretical framework of our study, we draw upon the taxonomic classification proposal of (Souza et al., 2012) and the critical insights of Mazzocchi (2018) into KOS. These foundational works underpin our proposed evaluation criteria for Wikipedia and Wikidata.

Specifically, in the context of Wikipedia, Albuquerque (2017) presents an information architecture framework for the development and management of controlled vocabularies in the context of programming vocabulary projects. Kaplan et al. (2022) introduce an evaluation method for taxonomies, including structural quality criteria such as generality, appropriateness-attainment, and orthogonality, and provide generalized metrics for quantifying generality and appropriateness.

In the domain of ontologies, da Costa et al. (2022) provide an updated review of software architectures, including ontology usage for managing large volumes of data. Wilson et al. (2022) outline a methodology for evaluating ontology quality that considers intrinsic and extrinsic aspects. Amith et al. (2018) offer insights into ontology evaluation within the field of biomedical KOS, which we adapt for evaluating Wikidata. Bolotnikova et al. (2011) propose practical methods for ontology evaluation, especially in automated contexts. Aghaebrahimian et al. (2022) explore the validity of Wikipedia categories for topic labeling, further contributing to the development of our evaluation criteria.

The extrinsic criteria (Kless and Milton, 2010) assess the measurement of external qualities, their application, and the domain, making reference to elements of the outcome as experienced by users. In contrast, quality indicators analyze aspects of structure and domain independently of their use in application contexts.

To gain a comprehensive understanding of the efforts to unify the reviewed theories and the proposed methodology for ontology evaluation, see Table 2.

INSERT TABLE 2

4.2. Proposed heuristic evaluation of taxonomies

Heuristic evaluation, aiming to assess whether the taxonomy of Catalan Wikipedia complies with the standards of sound knowledge organization, not only concerning user experience but also formally within the realm of Knowledge Organization Systems (KOS). Based on the theoretical framework, a selection of indicators were selected that have been highlighted in our analysis and achievable with the access and technical resources available to us (See Table 3). When identifying and measuring these indicators, we have considered contributions from specialists and specific standards within the KOS sector, particularly taxonomies, to conduct an inspection analysis of Wikipedia's category scheme.

INSERT TABLE 3

4.3. Analysis of usage logs for the profession case study on gendered professions

For the analysis of logs of the Catalan edition of Wikipedia we have used Pageviews Analysis (https://pageviews.wmcloud.org) which is a suite of eight tools designed for the examination of page views and unique device statistics on Wikimedia Foundation wikis. These tools, namely Pageviews, Langviews, Topviews, Siteviews, Massviews, Redirect Views, Userviews, and Mediaviews, collectively form a comprehensive toolkit for data analysis. The foundation of these tools relies on data sourced from Wikimedia's RESTBase

API, which is structured in alignment with the definitions outlined in the Research: Page view and Research: Unique Devices documentation. Presently, this suite of tools is under the maintenance and stewardship of Community Tech.

To address this analysis, we have chosen the field of professions, and based on state statistical data (*INE: Instituto Nacional de Estadística*, 2024), we have selected the most masculinized (STEM) and feminized professions (nursing, library science, and teaching) in Spain.

4.4. Heuristic assessment concerning structure and coverage

It is essential to clarify that in Wikidata, property P21 encompasses both gender and sex. However, it is crucial to recognize that these two terms pertain to distinct aspects of human identity and biology. Sex is primarily associated with an individual's physical and genetic characteristics and has historically been classified into two categories: male or female. In contrast, gender is a social and cultural construct that encompasses a broad spectrum of roles, behaviors, expectations, and identities. It extends beyond a binary system, acknowledging that people can identify as male, female, both, neither, or a different gender altogether. It is imperative to comprehend the differentiation between sex and gender, as it is fundamental for fostering inclusivity and honoring the diverse experiences and identities of individuals (García Dauder and Pérez Sedeño, 2017).

Apart from this feature of gender or sex of Wikidata, the members of the Ontology project have identified the limitations that make it not qualify as a proper ontology (Wikimedia, 2022). These limitations can be divided into two groups. The first group was initially identified in WikidataCon 2021, and they are aimed at overcoming barriers to the reuse of data by other services and projects. And the second group is considered to be issues existing in the knowledge representation in Wikidata. In the context of this study, we are primarily interested in the first group, as it identifies elements to overcome if it is to be applied in the categorization of Wikipedia content.

Based on the barriers to reuse formulated by the project members, we present examples related to the classes that make up the range restriction of property P21 (gender or sex). The indicators have been selected considering their relevance and their suitability for the retrieval of gender-related articles, however, this can be extrapolated to other evaluator needs.

INSERT TABLE 4

Commented [2]: INSERT TABLE 4

~~Table 3. Method for assessing the quality of the Wikidata ontology~~

~~Indicator~~	~~Description~~
~~Instances used as classes~~	~~The "instance of" (P31) property only accepts classes as values, as indicated by its type "Wikidata property for the relationship of the element to its class" (Q28326730).~~

Disarray at the Upper Levels of the Ontology	The top-level of the ontology should feature highly general classes (e.g., Time, Space, Event) independent of specific domains. These concepts must be mutually exclusive and collectively cover the knowledge domains of the ontology.
Semantic Deviation	An entity is seen from multiple perspectives, with distinct properties in each, but these merge into a single class. While individual subclass relationships are correct, their combined configuration is not.
Cycles or Loops in the "subclass de" (P279) Property	Class A has a subclass B, and class B is also a subclass of A, either directly or indirectly.
Redundant Generalization	Class A is both a subclass of B and a subclass of B's direct or indirect subclass.
Inconsistent Modeling	Differential treatment of two classes in terms of the number and types of classes they are linked to.
Repetition of Classes	The same class is defined multiple times.

4.5. Performance of the Wikidata search system

The data from Wikidata can be used for various purposes. Beyond the specific querying of an item or a set of items, Wikidata provides users with methods of data access for linking data without having to download it to another server, for enriching third-party data, or for generating local search services. In all cases, Wikidata's data can be consumed by human users or by automated systems or bots (Wikimedia, 2023b).

In one of the Wikidata guides, "Data Access" (Wikimedia, 2023a), eight methods for accessing Wikidata data are identified and described, three of which are oriented towards direct interaction with users who need to retrieve limited quantities of results (See Table 4).

INSERT TABLE 54

All methods of accessing Wikidata data operate on a foundation formed by the RDF data management system, or RDF repository, Blazegraph (Vrandečić et al., 2023) (See Table 5).

INSERT TABLE 65

Undoubtedly, these figures are impressive and represent the largest open secondary database currently in existence. Nevertheless, in recent years, assessments of the degree of compliance with processes, accessibility, and the use of search services have shown worrisome signs of stagnation. The Wikidata authorities are fully aware of these limitations and, in fact, have set their sights on the need to replace the underlying software of Wikidata, Blazegraph, with one that can better address the challenges of growth and quality.

And, regarding the ontology inconsistencies we mentioned earlier, the evaluation requirements established incorporate the use of more advanced integrity-checking languages than SPARQL functions. Specifically, the WDQS report refers to the Shapes Constraint Language, or SHACL. SHACL allows for graph validation and includes not only the ability to specify a severity level for validation results, but also the possibility of providing suggestions on how to fix the data if a validation result occurs.

The performance assessment of Wikidata follows the overarching evaluation framework introduced by Malyshev et al. (2018). The performance tests cover the period from 2015 to 2022, as specified by the SPARQL query service (Everett, 2015).

5. Results

5.1. Heuristic evaluation of SOK Wikipedia

The heuristic evaluation of the Catalan Wikipedia category scheme has been carried out using the technique of standards inspection, in which a usability expert analyzes whether the interface follows the agreed-upon specifications and the standards defined on an international level. In the case at hand, a set of identified indicators has been generated, particularly based on normative sources. These sources also provide us with methods for obtaining evidence for each indicator, the applied metrics, and, when possible, optimal values.

a) Evaluability

The category schema of Wikipedia is valuable because category creators have various agreed-upon tools for their practice. We highlight the following:

- Categorization guideline (Wikimedia, 2023d) resulting from the discussion and decision-making process specific to the encyclopedia.
- Help for category creators: Help:Category (Wikimedia, 2018) and the "Style Book on Categorization" section in the Categorization guideline (Wikimedia, 2023d).
- Templates for category creators: Category:Maintenance templates for categories (Wikimedia, 2015).
- There is a control over the pages that do not contain categories, for maintenance purposes.

The level of knowledge about these tools has been informally assessed with some individual administrators of the Viquipèdia community, and their lack of awareness regarding them has been conveyed.

b) Reusability

Each category has a unique instance and a single identifier, regardless of its various locations within the schema hierarchies. Wikipedia's schema categories are available for database dumps in three data interchange formats: sql, json, and xml. These formats do not provide semantic information about the concepts and relationships between the concepts in Wikipedia's categorization schema, as the Simple Knowledge Organization System (SKOS) data model could. Consequently, the possibilities for reusing Wikipedia categories in other datasets or information retrieval systems are greatly restricted.

c) Stability

The most notable stability-related metrics during the period 2004-2022 can be seen in Table 76.

INSERT TABLE 6

The annual growth rate remained high during the period 2004-2007, and from 2008 onwards, it experienced a significant decline until 2012. Starting from that year, the rate demonstrates a gradual reduction in the increase, and it stabilizes until 2020 when it experienced a very remarkable increase. This increase may be linked to one of the effects of the Covid-19 pandemic: greater availability of time for contributing to the Catalan edition of Wikipedia.

The most abundant categories are in the fields of science and culture, followed by technology, humanities, and events. The three areas with the fewest categories are biographies, information, and places.

d) Number of categories (concepts)

As of December 31, 2022, Wikipedia's category schema included 102,159 categories. We can compare this size with other similar KOSs that consist of pre-coordinated concepts and aim to represent encyclopedic knowledge.

The List of Subject Headings of the National Catalan Library (LEMAC) contains 112,200 headings, considering both accepted and non-accepted ones. It originates from the translation of the Spanish version of the Library of Congress Subject Headings (LCSH), which was preliminarily published by the Library Services of the Generalitat de Catalunya in 1988.

Making growth rate comparisons between Wikipedia's categorization schema and LEMAC is challenging. Wikipedia's schema is relatively young, still in its first two decades of existence, while LEMAC has been in existence for 35 years, with an even longer history if we consider its origins. However, the following examples show clear indications of faster growth in Wikipedia's categorization schema.

In 2009, LEMAC experienced a growth of 2,663 new headings, whereas Wikipedia added 8,080 new categories. By 2021, LEMAC's growth amounted to 1,219 new headings, while Wikipedia introduced a staggering 8,616 new categories.

The magnitude of Wikipedia's categories and its growth rate are not comparable to those of Knowledge Organization Systems (KOS) applied to digital encyclopedias. The knowledge tree of the Gran Enciclopèdia Catalana (GEC) comprises 425 categories, and the categories in the Encyclopedia Britannica total 123. In both cases, the hierarchy of the KOS is restricted to two levels.

e) Number of semantic relationships

We lack access to complete data on all hierarchical relationships between supercategories and subcategories. Still, we have a partial count covering the first three levels of the categorization schema, specifically involving main thematic categories and their second-level and third-level subcategories, resulting in 1122 hierarchy relationships.

f) Enrichment or granularity index

In the first three levels of the category schema, there are 143 categories and 1122 hierarchy relationships. The corresponding average enrichment index is 7.8461, significantly surpassing the optimal range, usually between 2 and 5.

g) Degree of precoordination

In the entirety of Wikipedia's category schema, the average number of words that make up category labels is 3.6766, exceeding the maximum value typically recommended by experts, which is between 1.5-2 words. Other data indicating a deviation from this optimal value include the median number of words in category labels, which is three. However, there are exceptions, with some category labels having a maximum of 18 words, such as in the case of "Resolucions del Consell de Seguretat de les Nacions Unides sobre el Tribunal Penal Internacional per a l'antiga Iugoslàvia" (Resolutions of the United Nations Security Council on the International Criminal Tribunal for the former Yugoslavia).

h) Number of levels in hierarchy or depth

For the evaluation of this indicator, we reviewed all hierarchical chains within the main thematic category "Biographies." In all cases, we found more than five levels. Consequently, this exceeds the maximum value recommended by experts

i) Number of categories in the same hierarchy level or breadth

To assess this indicator, we examined the first two levels of subcategorization beyond the eight main thematic categories. In total, this section includes 144 categories, with 130 of them containing subcategories, while the remaining 14 link directly to Wikipedia pages.

Among these 130 categories, there are 17 that have only one subcategory, constituting 13.07% of the assessed section, and in fact, this is the most common case. These instances violate the minimum requirement of two subcategories recommended by experts.

Additionally, there are 35 categories with more than 12 subcategories, making up 26.92% of the evaluated section. These cases exceed the maximum of twelve subcategories suggested by experts. The highest breach of this limit occurs in the "Religion" category, which includes 43 subcategories.

When we sum up violations of both the minimum and maximum subcategory limits, we find that 40% of the assessed categories do not adhere to the optimal values of breadth (52 out of 130 categories).

5.2. Usage logs for the case study Wikipedia

To provide a glimpse of Viquipèdia usage (Catalan Wikipedia), the total number of viewed pages of the Catalan Wikipedia in one month is 49.338.638 and the number of unique devices accesses is 3.480.772 in September 2023 (Wikimedia, 2023c).

And in this section, we present the data derived from our analysis of log entries for feminized professions (such as librarians, nurses, and teachers), juxtaposed with STEM professions (See Table 7). The table encompasses two categories: "Feminized Professions" and "STEM Professions." Each category is further broken down into specific professions, and the

corresponding user visualizations statistics are presented. We examine access patterns beginning in June 2023, focusing on the Catalan edition of Wikipedia, and encompassing data from various devices. The table illustrates the engagement and activity levels across different professions within feminized or masculinized professions.

INSERT TABLE 8~~7~~

The findings from the examination of feminized professions reveal that there is an average of 798 monthly accesses, with a mean of 2 editions monthly. And the outcomes obtained from the examination of STEM professions indicate that there is an average of 1073 monthly accesses, with a mean of 0 editions monthly.

5.3. Heuristic evaluation of Wikidata

These are the results of the heuristic evaluation of the ontologies of Wikidata. Examining Table 9 reveals significant challenges stemming from unproductive class hierarchies in navigation and search. Users are constrained from selecting multiple individuals of the same type within nodes, introducing complexity in search contexts. This limitation hampers quick decision-making, impacts reasoning by disrupting inference and consistency assessments, and impedes automated interoperability in data cooperation. Ascending through upper chains in search contexts confuses users and complicates processes in both search and axiom-based reasoning. The inability to determine generality or specificity transfers the challenge to the search process, hindering statement validation and new knowledge inference. Overall, unproductive class hierarchies present intricate obstacles across various facets of ontology usage.

INSERT TABLE 9~~8~~

5.4. Performance~~Performace~~ Wikidata

The following table provides a comprehensive evaluation of Wikidata's performance over the period from 2015-22, using Malyshev et al. (2018) framework.

INSERT TABLE 10~~9~~

Table 9 furnishes a comprehensive evaluation of query performance within the Wikidata platform through several key metrics. Under the section denoted as "Query Metrics," distinct trends come to the forefront. Notably, a substantial count of "Good Queries" signifies operations executed with success, contributing significantly to the platform's operational prowess. Conversely, a noteworthy number of "Bad Queries" denotes instances where queries faced issues or failed to deliver intended results, thereby illuminating potential areas for refinement.

Moreover, the metric detailing the "Total Query Execution Time" offers a panoramic view of Wikidata's efficiency, encapsulating the cumulative time required for executing the entire array of queries. This temporal dimension serves as a pivotal indicator of the platform's responsiveness. In tandem, the metric revealing the "Total Result Rows" speaks volumes

about the sheer magnitude of information generated across the spectrum of queries conducted on Wikidata. This voluminous outcome underscores the platform's extensive capacity in producing relevant and diverse information.

6. Discussion

In 2017, the Wikimedia Movement adopted a new strategic plan for 2030, which establishes the goal of "providing knowledge as a service" (becoming a platform that offers open knowledge to the world through interfaces and communities), with a focus on "knowledge equity" (directing our efforts towards knowledge and communities that have been marginalized by power structures and privileges... Breaking social, political, and technical barriers that hinder people from accessing and contributing to free knowledge). This collaborative strategic document places two core principles at its recommendations: inclusivity and a people-centered approach (understood as attending to people's needs). It sets the goal for 2030 as closing the gender gap and focusing on the inclusion of underrepresented groups.

The knowledge organization proposal presented in this paper is fully aligned with the new strategic direction. To increase the visibility and access to the knowledge of Wikipedia, particularly that related to and about marginalized gender groups (women, non-binary individuals, intersex, trans men, and trans women), a dual solution is proposed. On the one hand, a technical solution involving the use of Wikidata ontologies as a knowledge organization system to facilitate information search and retrieval in Wikipedia, without increasing biases existing ~~biases~~ in reality. Wikidata has shown to be more aligned with the gender perspective than Wikipedia, as demonstrated in in-depth studies (Zhang and Terveen, 2021). On the other hand, a social, cultural, and political solution that involves working with the Wikipedia community to accept internationally recognized standards in the field of knowledge organization and embracing two principles considered strategic by the Wikipedia movement: inclusivity to avoid the discrimination of marginalized groups and, above all, the principle of people-centered service with a special focus on their informational needs.

With the technical solution, by opting for Wikidata ontologies as the knowledge organization system for Wikipedia's content, alignment with international standards of knowledge organization would be achieved, and it would empower Wikipedia users (readers and editors) to search for and retrieve encyclopedia content according to their needs. Empowerment would occur during the search and navigation process, as users would decide on search elements closer to their needs, unlike the current categories, which are proposed based on the worldview of those who created, classified, and indexed them, without allowing the combination of search elements.

Delving into the details of the proposal in this academic paper, it advocates the use of ontologies as a knowledge organization system and opts for Wikidata because its original purpose is to store data (properties and relationships) from content present in Wikipedia articles in any language. If taken to its fullest potential, Wikidata could become the knowledge organization system for Wikipedia. In this sense, some current Wikipedia categories are already directly constructed from Wikidata ("living people"). Wikipedia already has different sections linked to Wikidata, acting as a knowledge organization system through common examples like InfoBoxes or authority records.

In the evaluation carried out by experts and the heuristics, most assessments of Wikipedia's analysis variables are low. In contrast, for Wikidata, there is room for improvement, and the Wikidata community has already identified these and included them in the agenda for improvement and tool development to address these issues. Two interesting contributions are the entity schema and the backend, which bring substantial benefits to the ontology, such as aspects related to the organization of data related to sex or gender. This demonstrates the potential for improvement. It is also important to recognize the need to revise the name of the "sex or gender" property, which mixes biological characteristics with individual definition and social construction in its label.

In the case of Wikidata, decision-making processes for improvements are made within a smaller community with a more respectful perspective of accepted and recognized international standards in the field of knowledge organization. In contrast, in the case of Wikipedia, the gender bias existing in society is exacerbated by opposing positions on gender diversity expressed with strong ideological arguments.

At the same time, opting for Wikidata ontologies as the knowledge organization system for Wikipedia's content would empower users (readers or editors) to meet their informational needs. We agree that Wikipedia's categorization system (category schemes) is easy for users to understand and closely aligned with their vocabulary and natural language. However, it also has disadvantages, as discussed in this proposal, related to cultural, social, or political biases and imbalances that any controlled vocabulary entails. On the other hand, category schemes, are pre-coordinated thesauri that combine concepts, classes, or terms from a controlled vocabulary at the time of their construction or indexing. This means that there could be a category like "Catalan doctors from the south" created by the Wikipedia community, and a person from anywhere in the world should be able to understand (deduce) that it may include "female doctors living in the southern part of Catalonia." In contrast, the use of ontologies to organize knowledge would provide a better representation of Wikidata's content because each property represents a single dimension (attribute) of an entity or a set of entities. It is the user who, at the time of the search, combines the attributes that best respond to their need to retrieve the relevant entity or entities. Currently, a Wikipedia category like "Catalan doctors from the south" links two different dimensions of a person: their profession and their origin. However, in an ontology, one could choose which elements to combine (profession, year of birth, place of residence, or any other) independently and in combination with Boolean operators, so that the search would align much more with the user's information needs (human or machine). The knowledge organization system proposed by Wikidata is post-coordinated, empowering users through the system by allowing them to combine predefined attributes. Therefore, Wikidata is the ontology that could bring organization and a better representation of what is known in Wikipedia. In fact, Wikipedia already generates categories that arise from Wikidata, such as "living people."

The Wikidata ontology can be a solution for applying a gender perspective and overcoming the lack of scientific-technical foundation. However, the need for a cultural change within the community, which should accept the process of substitution to align with international technology consensus and promote more equitable access to content diversity, cannot be ignored.

7.Conclusions

This study has delved into the complex issue of gender bias within Wikipedia's knowledge organization system, specifically within its taxonomy categories. By synthesizing the theoretical framework, an extensive review of academic literature, and a detailed analysis of Wikipedia's content structure, we have arrived at the following comprehensive conclusions.

Throughout our research, it has become abundantly clear that gender bias persists within Wikipedia, as attested by a consistent body of academic literature. This bias is not confined to the content itself but also extends to the limited diversity among volunteer contributors across various languages. Our own analysis further reinforces this finding, substantiating the existence of gender bias within Wikipedia's content organization system.

One notable revelation is that Wikipedia's encyclopedic nature tends to prioritize the perspectives of content indexers and categorizers over the needs of its users. Consequently, the categories within Wikipedia have often been designed to facilitate the work of editors, collecting pages under specific concepts, while falling short in terms of enabling effective information retrieval and meeting user information needs.

Furthermore, we have identified that Wikipedia's system of categories (KOS) frequently falls short of established quality standards. For instance, hierarchy depth often exceeds the recommended maximum of 12 levels, and the breadth, indicated by the number of subcategories within a category, deviates from the recommended range of 2 to 5 subcategories.

From the perspective of gender and intersectional analysis, it becomes evident that there is no objective basis for excluding gender identities, such as "women" or "non-binary individuals," as categorization criteria on Wikipedia. This finding highlights the presence of inconsistencies, such as the use of female-gendered first-level categories ("midwives" or "bearded women"), which persist within the platform.

Our study concludes by asserting that Wikipedia's categories hold significant potential for improvement, not only in addressing issues related to gender identity but also in enhancing the overall knowledge organization system for more effective user information retrieval. This recommendation stems from the observation that the vast majority of Wikipedias, with a few exceptions like the Catalan and Italian versions, have seamlessly incorporated gender identity categories into their organizational systems, thereby aiding content search and retrieval for users.

In light of these conclusions, we strongly recommend a comprehensive reevaluation of Wikipedia's content organization system. This reevaluation should focus on inclusivity, equity, and the fulfillment of users' information needs. Acknowledging the potential for the integration of gender identity as a valid classification criterion, Wikipedia can make substantial strides toward aligning its knowledge organization practices with contemporary principles of information access and inclusion.

Shifting our attention to the analysis of Wikidata, our investigation has focused on the technological aspects involved in the organization and retrieval of gender-diverse content within this platform. From this examination, several key conclusions have emerged.

First, our analysis has revealed that Wikidata exhibits a commendable level of sensitivity toward gender diversity, notably seen in the inclusion of a variety of gender categories under the P21 property.

Second, we recommend that Wikidata makes a clear distinction between properties related to biological sex and properties tied to gender identity, as the existing disjunctive labeling of the P21 property conflates these two distinct concepts.

Third, in contrast to Wikipedia, which often grapples with socio-cultural influences in decision-making processes, our analysis shows~~demonstrates~~ that Wikidata effectively mirrors real-world gender diversity without exacerbating existing biases as evidences by the research conducted by authors such as Zeng and Treviers (2023). The findings presented in this paper illustrate that Wikidata offers a richer array of tools to represent the diversity of gender identities.~~.~~

Fourth, the Wikidata community tends to emphasize technical and data-centric arguments in its decision-making processes, diverging from Wikipedia's debates that often involve socio-cultural considerations, particularly regarding gender categories.

Lastly, the linguistic diversity of Wikidata poses unique challenges, particularly in languages where gender differentiation is significant. The debate over gender-neutral labelling in languages like Catalan underscores the importance of linguistic and cultural sensitivity in maintaining the dataset.

In conclusion, this analysis has predominantly delved into the technological aspects of enhancing the representation of gender diversity within Wikidata. However, it is imperative to recognize that a comprehensive solution necessitates a harmonious blend of technological enhancements and cultural considerations in the decision-making processes governing the organization of content in this vital knowledge-sharing platform. This convergence of technology and culture is paramount in fostering inclusivity and equity in the representation of gender diversity in the digital realm.

8.Research funding

This research received support from the Spanish Ministerio de Innovación, Ciencia y Universidades (MCIN) and the Agencia Estatal de Investigación [Grant ref. PID2020-116936RA-I00]. Additionally, it was funded by the Xarxa Vives d'Universitats, comprising 21 universities across Andorra, France, Italy, and Spain, in the Catalan language domain.

9.Bibliograhy

Abián, D., Meroño-Peñuela, A., and Simperl, E. (2022). An Analysis of Content Gaps Versus User Needs in the Wikidata Knowledge Graph. In Sattler U., Hogan A., Keet M., Presutti V., Almeida J.P.A., Takeda H., Monnin P., Pirrò G., and d'Amato C. (Eds.), *Lect. Notes Comput. Sci.: Vol. 13489 LNCS* (pp. 354–374). Springer Science and Business Media Deutschland GmbH; Scopus. https://doi.org/10.1007/978-3-031-19433-7_21

Acey, Camille E., Bouterse, Siko, Ghoshal, Sucheta, Menking, Amanda, Sengupta, Anasuya, and Vrana, Adele G. (2021). Decolonizing the Internet by Decolonizing Ourselves: Challenging Epistemic Injustice through Feminist Practice. *Global Perspectives*, *2*(1). https://doi.org/10.1525/gp.2021.21268

Aghaebrahimian, Ahmad, Stauder, Andy, and Ustaszewski, Michael. (2022). Testing the validity of Wikipedia categories for subject matter labelling of open-domain corpus data. *Journal of Information Science*, *48*(5), 686–700. https://doi.org/10.1177/0165551520977438

Albuquerque, Fernando Antônio de Araújo Chacon de. (2017). *Arcabouço de arquitetura da informação para ciclo de vida de projeto de vocabulário controlado: Uma aplicação em Engenharia de Software* [Fernando Antônio de Araújo Chacon de]. https://repositorio.unb.br/handle/10482/31288

Amith, Muhammad, He, Zhe, Bian, Jiang, Lossio-Ventura, Juan Antonio, and Tao, Cui. (2018). Assessing the practice of biomedical ontology evaluation: Gaps and opportunities. *Journal of Biomedical Informatics*, *80*, 1–13. https://doi.org/10.1016/j.jbi.2018.02.010

Antin, J., Yee, R., Cheshire, C., and Nov, O. (2011). *Gender differences in Wikipedia editing*. 11–14. https://doi.org/10.1145/2038558.2038561

Bolotnikova, E. S., Gavrilova, T. A., and Gorovoy, V. A. (2011). To a method of evaluating ontologies. *Journal of Computer and Systems Sciences International*, *50*(3), 448–461. https://doi.org/10.1134/S1064230711010072

Bourli, S., and Pitoura, E. (2020). Bias in Knowledge Graph Embeddings. In Atzmuller M., Coscia M., and Missaoui R. (Eds.), *Proc. IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Min., ASONAM* (pp. 6–10). Institute of Electrical and Electronics Engineers Inc.; Scopus. https://doi.org/10.1109/ASONAM49781.2020.9381459

Buchem, Ilona, and Kloppenburg, Julia. (2013). *Gender – Diversität – Wikipedia: Vielfalt gemeinsam gestalten*. Beuth Hochschule für Technik Berlin; Wikimedia Deutschland. https://www.bht-berlin.de/fileadmin/oe/gutz/Sonstige_Veroeffentlichungen/Arbeitspapier_Gender-Diversity-Wikipedia.pdf

Centelles, Miquel, and Ferran-Ferrer, Núria. (2024). Taxonomies and Ontologies in Wikipedia and Wikidata: An In-Depth Examination of Knowledge Organization Systems. [Manuscript submitted for publication]. *Hypertext.Net*, *27*.

Collier, Benjamin, and Bear, Julia. (2012). Conflict, Criticism, or Confidence: An Empirical Examination of the Gender Gap in Wikipedia Contributions. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 383–392. https://doi.org/10.1145/2145204.2145265

Conroy, M. (2023). Quantifying the Gap: The Gender Gap in French Writers' Wikidata. *Journal of Cultural Analytics*, *8*(2). Scopus. https://doi.org/10.22148/001c.74068

da Costa, Tiago Vinícius Remígio, Cavalcante, Everton, and Batista, Thais. (2022). Big Data Software Architectures: An Updated Review. In Osvaldo Gervasi, Beniamino Murgante, Eligius M. T. Hendrix, David Taniar, and Bernady O. Apduhan (Eds.), *Computational Science and Its Applications – ICCSA 2022* (pp. 477–493). Springer International Publishing. https://doi.org/10.1007/978-3-031-10522-7_33

Das, Maitraye, Hecht, Brent, and Gergle, Darren. (2019). The Gendered Geography of Contributions to OpenStreetMap: Complexities in Self-Focus Bias. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3290605.3300793

Eckert, Stine, and Steiner, Linda. (2013). (Re)triggering backlash: Responses to news about Wikipedia's gender gap. *Journal of Communication Inquiry*, *37*(4), 284–303. https://doi.org/10.1177/0196859913505618

Evans, Sian, Mabey, Jacqueline, and Mandiberg, Michael. (2015). Editing for Equality: The Outcomes of the Art+Feminism Wikipedia Edit-a-thons. *Art Documentation*, *34*(2), 194–203. https://doi.org/10.1086/683380

Everett, Nikolas. (2015, March 5). Wikidata Query Backend Update (take two!). *Wikidata-Tech*. https://lists.wikimedia.org/hyperkitty/list/wikidata-tech@lists.wikimedia.org/message/VPQ226NBQ5D2ZCNUOHJL3X223Z4HUNJF/

Falenska, A, Cetinoglu, O, and Assoc Computat Linguist. (2021). *Assessing Gender Bias in Wikipedia: Inequalities in Article Titles* (WOS:000694722900009). 75–85.

Farzan, R., Savage, S., and Saviaga, C.F. (2016). *Bring on board new enthusiasts! A case study of impact of wikipedia art + feminism edit-a-thon events on newcomers: Vol. 10046 LNCS* (p. 40). Scopus. https://doi.org/10.1007/978-3-319-47880-7_2

Ferran-Ferrer, Núria, Castellanos-Pineda, Patricia, Minguillón, Julià, and Meneses, Julio. (2021). The gender gap on the Spanish Wikipedia: Listening to the voices of women editors. *Profesional de La Información*, *30*(5), Article 5. https://doi.org/10.3145/epi.2021.sep.16

Ferran-Ferrer, Núria, Centelles, Miquel, Macià, Yessica, Boté-Vericad, Juan-José, and Minguillon, Julià. (2023). *Dones de categoria: Anàlisi del biaix de gènere a les categories de Viquipèdia: Informe de diagnosi tècnica, posicionament acadèmic i proposta de millora del sistema d'organització del coneixement de Viquipèdia* (p. 131). Programa d'Igualtat de Gènere de la Xarxa Vives d'Universitats.

Ford, Heather, and Wajcman, Judy. (2017). 'Anyone can edit', not everyone does:

Wikipedia's infrastructure and the gender gap. *Social Studies of Science*, *47*(4), 511–527. https://doi.org/10.1177/0306312717692172

García Dauder, Silvia, and Pérez Sedeño, Eulalia. (2017). *Las "Mentiras" científicas sobre las mujeres*. Los Libros de la Catarata.

Gardner, Sue. (2011, February 20). Nine Reasons Women Don't Edit Wikipedia (in their own words). *Sue Gardner's Blog*. https://suegardner.org/2011/02/19/nine-reasons-why-women-dont-edit-wikipedia-in-their-own-words/

Grant, Maria J., and Booth, Andrew. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, *26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Gruwell, L. (2015). Wikipedia's politics of exclusion: Gender, epistemology, and feminist rhetorical (in)action. *Computers and Composition*, *37*, 117–131. https://doi.org/10.1016/j.compcom.2015.06.009

Hermoso Pulido, T. (2021). Simple Wikidata Analysis for Tracking and Improving Biographies in Catalan Wikipedia. *Web Conf. - Companion World Wide Web Conf., WWW*, 582–583. Scopus. https://doi.org/10.1145/3442442.3452344

Hinnosaar, Marit. (2019). Gender inequality in new media: Evidence from Wikipedia. *Journal of Economic Behavior & Organization*, *163*, 262–276. https://doi.org/10.1016/j.jebo.2019.04.020

Hollink, Laura, Van Aggelen, Astrid, and Van Ossenbruggen, Jacco. (2018). Using the Web of Data to Study Gender Differences in Online Knowledge Sources: The Case of the European Parliament. *Proceedings of the 10th ACM Conference on Web Science*, 381–385. https://doi.org/10.1145/3201064.3201108

Hube, C. (2017). *Bias in wikipedia*. 717–721. Scopus. https://doi.org/10.1145/3041021.3053375

INE: Instituto Nacional de Estadística. (2024). INE. https://www.ine.es/

Ju, Boryung, and Stewart, Brenton. (2019). "The right information": Perceptions of information bias among Black Wikipedians. *Journal of Documentation*, *75*(6), 1486–

1502. https://doi.org/10.1108/JD-02-2019-0031

Kaplan, Angelika, Kühn, Thomas, Hahner, Sebastian, Benkler, Niko, Keim, Jan, Fuchß, Dominik, Corallo, Sophie, and Heinrich, Robert. (2022). Introducing an Evaluation Method for Taxonomies. *Proceedings of the 26th International Conference on Evaluation and Assessment in Software Engineering*, 311–316. https://doi.org/10.1145/3530019.3535305

Karczewska, A., and Kukowska, K. (2021). *Cultural dimension of femininity: Masculinity in virtual organizing knowledge sharing*. 414–422. Scopus. https://doi.org/10.34190/EKM.21.135

Klein, M., Gupta, H., Rai, V., Konieczny, P., and Zhu, H. (2016). Monitoring the gender gap with Wikidata human gender indicators. *Proc. Int. Symp. Open Collab., OpenSym*. Proceedings of the 12th International Symposium on Open Collaboration, OpenSym 2016. Scopus. https://doi.org/10.1145/2957792.2957798

Klein, Max, and Konieczny, Piotr. (2015). Wikipedia in the world of global gender inequality indices: What the biography gender gap is measuring. *Proceedings of the 11th International Symposium on Open Collaboration*, 1–2. https://doi.org/10.1145/2788993.2789849

Kless, Daniel, and Milton, Simon. (2010). Towards Quality Measures for Evaluating Thesauri. In Salvador Sánchez-Alonso and Ioannis N. Athanasiadis (Eds.), *Metadata and Semantic Research* (pp. 312–319). Springer. https://doi.org/10.1007/978-3-642-16552-8_28

Konieczny, P., and Klein, M. (2018). Gender gap through time and space: A journey through Wikipedia biographies via the Wikidata Human Gender Indicator. *New Media and Society*, *20*(12), 4608–4633. Scopus. https://doi.org/10.1177/1461444818779080

Konieczny, Piotr. (2018). Volunteer Retention, Burnout and Dropout in Online Voluntary Organizations: Stress, Conflict and Retirement of Wikipedians. In Patrick G. Coy (Ed.), *Research in Social Movements, Conflicts and Change* (Vol. 42, pp. 199–219). Emerald Publishing Limited. https://doi.org/10.1108/S0163-786X20180000042008

Lam, Shyong (Tony) K., Uduwage, Anuradha, Dong, Zhenhua, Sen, Shilad, Musicant, David R., Terveen, Loren, and Riedl, John. (2011). WP:clubhouse?: An exploration of Wikipedia's gender imbalance. *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, 1–10. https://doi.org/10.1145/2038558.2038560

Laouenan, M., Bhargava, P., Eyméoud, J.-B., Gergaud, O., Plique, G., and Wasmer, E. (2022). A cross-verified database of notable people, 3500BC-2018AD. *Scientific Data*, *9*(1). Scopus. https://doi.org/10.1038/s41597-022-01369-4

Lemus-Rojas, M., and Lee, Y.Y. (2019). Using wikidata to provide visibility to women in STEM. *Proc. Int. Conf. Dublin Core Metadata Appl.*, 126–131. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85088230329&partnerID=40&md5=e37fc07992e9f29aa3de487bb6252e36

Malyshev, Stanislav, Krötzsch, Markus, González, Larry, Gonsior, Julius, and Bielefeldt, Adrian. (2018). Getting the Most Out of Wikidata: Semantic Technology Usage in Wikipedia's Knowledge Graph. In Denny Vrandečić, Kalina Bontcheva, Mari Carmen Suárez-Figueroa, Valentina Presutti, Irene Celino, Marta Sabou, Lucie-Aimée Kaffee, and Elena Simperl (Eds.), *The Semantic Web – ISWC 2018* (Vol. 11137, pp. 376–394). Springer International Publishing. https://doi.org/10.1007/978-3-030-00668-6_23

Mandiberg, M., and Sarıoğlu, D. (2022). Clowns in the Visual Artists: Topic Modeling Wikipedia and Wikidata. *Art Documentation*, *41*(1), 20–37. Scopus. https://doi.org/10.1086/719999

Mazzocchi, Fulvio. (2018). Knowledge Organization System (KOS): An Introductory Critical Account. *Knowledge Organization*, *45*(1). https://doi.org/10.5771/0943-7444-2018-1-54

Miquel-Ribe, M, and Laniado, D. (2021). The Wikipedia Diversity Observatory: Helping communities to bridge content gaps through interactive interfaces. *JOURNAL OF INTERNET SERVICES AND APPLICATIONS*, *12*(1). https://doi.org/10.1186/s13174-021-00141-y

Mora-Cantallops, Marçal, Sánchez-Alonso, Salvador, and García-Barriocanal, Elena. (2019). A systematic literature review on Wikidata. *Data Technologies and Applications*, *53*(3), 250–268. https://doi.org/10.1108/DTA-12-2018-0110

Morgan, J.T., Bouterse, S., Stierch, S., and Walls, H. (2013). Tea & sympathy: Crafting positive new user experiences on wikipedia. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, 839–848. Scopus. https://doi.org/10.1145/2441776.2441871

Pellissier Tanon, T., and Suchanek, F. (2019). Querying the Edit History of Wikidata. In Hitzler P., Kirrane S., Hartig O., de Boer V., Schlobach S., Vidal M.-E., Maleshkova M., Hammar K., Lasierra N., Stadtmüller S., Hose K., and Verborgh R. (Eds.), *Lect. Notes Comput. Sci.: Vol. 11762 LNCS* (pp. 161–166). Springer Science and Business Media Deutschland GmbH; Scopus. https://doi.org/10.1007/978-3-030-32327-1_32

Souza, Renato Rocha, Tudhope, Douglas, and Almeida, and Maurício Barcellos. (2012). Towards a Taxonomy of KOS: Dimensions for Classifying Knowledge Organization Systems. *Knowledge Organization*, *39*(3), 179–192. https://doi.org/10.5771/0943-7444-2012-3-179

Thornton, K., and Seals-Nutt, K. (2018). Science stories: Using IIIF and wikidata to create a linked-data application. In Srinivas K., Fortuna C., Atre M., van Erp M., and Lopez V. (Eds.), *CEUR Workshop Proc.* (Vol. 2180). CEUR-WS; Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85055351166&partnerID=40&md5=2a141ac8b64f5e4eb0048f128ed06c3b

Thornton, K., Seals-Nutt, K., Van Remoortel, M., Birkholz, J.M., and De Potter, P. (2022). Linking women editors of periodicals to the Wikidata knowledge graph. *Semantic Web*, *14*(2), 443–455. Scopus. https://doi.org/10.3233/SW-222845

Tripodi, Francesca. (2023). Ms. Categorized: Gender, notability, and inequality on Wikipedia. *New Media & Society*, *25*(7), 1687–1707. https://doi.org/10.1177/14614448211023772

Vrandečić, Denny, Pintscher, Lydia, and Krötzsch, Markus. (2023). Wikidata: The Making Of. *Companion Proceedings of the ACM Web Conference 2023*, 615–624. https://doi.org/10.1145/3543873.3585579

Wagner, Claudia, Graells-Garrido, Eduardo, Garcia, David, and Menczer, Filippo. (2016). Women through the glass ceiling: Gender asymmetries in Wikipedia. *EPJ DATA SCIENCE*, *5*. https://doi.org/{10.1140/epjds/s13688-016-0066-4}

Wikidata. (2024). *Property talk:P21*. https://www.wikidata.org/wiki/Property_talk:P21

Wikimedia. (2015, November 17). *Categoria:Plantilles de manteniment per a categories*. https://ca.wikipedia.org/w/index.php?title=Categoria:Plantilles_de_manteniment_per_a_categories&oldid=16026819

Wikimedia. (2018, November 24). *Ajuda:Categoria*. https://ca.wikipedia.org/w/index.php?title=Ajuda:Categoria&oldid=20513864

Wikimedia. (2022). *Wikidata:WikiProject Ontology/Classes*. https://www.wikidata.org/wiki/Wikidata:WikiProject_Ontology/Classes

Wikimedia. (2023a). *Wikidata:Accés a les dades*. https://www.wikidata.org/wiki/Wikidata:Data_access/ca

Wikimedia. (2023b). *Wikidata:Bots*. https://www.wikidata.org/wiki/Wikidata:Bots

Wikimedia. (2023c). *Wikimedia Statistics—Catalán Viquipèdia*. https://stats.wikimedia.org/#/ca.wikipedia.org

Wikimedia. (2023d, October 23). *Wikipedia:Categorization*. https://en.wikipedia.org/w/index.php?title=Wikipedia:Categorization&oldid=1181497476

Wilson, R. Shyama I., Goonetillake, Jeevani S., Ginige, Athula, and Indika, Walisadeera Anusha. (2022). Ontology Quality Evaluation Methodology. In Osvaldo Gervasi, Beniamino Murgante, Eligius M. T. Hendrix, David Taniar, and Bernady O. Apduhan (Eds.), *Computational Science and Its Applications – ICCSA 2022* (pp. 509–528). Springer International Publishing. https://doi.org/10.1007/978-3-031-10522-7_35

Worku, Zena, Bipat, Taryn, McDonald, David W., and Zachry, Mark. (2020). Exploring

Systematic Bias through Article Deletions on Wikipedia from a Behavioral Perspective. *Proceedings of the 16th International Symposium on Open Collaboration*, 1–22. https://doi.org/10.1145/3412569.3412573

Zeng, Marcia Lei, and Mayr, Philipp. (2018). Knowledge Organization Systems (KOS) in the Semantic Web: A Multi-Dimensional Review. *International Journal on Digital Libraries*, 1–22.

Zhang, Charles Chuankai, and Terveen, Loren. (2021). Quantifying the Gap: A Case Study of Wikidata Gender Disparities. *17th International Symposium on Open Collaboration*, 1–12. https://doi.org/10.1145/3479986.3479992

Zheng, Xiang, Chen, Jiajing, Yan, Erjia, and Ni, Chaoqun. (2022). Gender and country biases in Wikipedia citations to scholarly publications. *Journal of the Association for Information Science and Technology*, *74*(2), 219–233. Scopus. https://doi.org/10.1002/asi.24723

Zhu, L., Xu, A., Deng, S., Heng, G., and Li, X. (2023). Entity Management Using Wikidata for Cultural Heritage Information. *Cataloging and Classification Quarterly*, *61*(1), 20–46. Scopus. https://doi.org/10.1080/01639374.2023.2188338

Table 1. Overview of Methodologies Employed for Individual Research Goals

Specific Objectives	**Methods**
O1. Developing a Standards Inspection Method for Wikipedia KOS	* Inspection of standards and guidelines for the evaluation of taxonomy (Wikipedia) and ontologies (Wikidata)
O2. Evaluating Wikipedia's Catalan knowledge organization system (taxonomy) on Gender-Related Article Retrieval	* Proposal for a heuristic evaluation of the taxonomies * Analysis of logs usage for the case study on gendered professions
O3. Enhancing Gender-Related Article Retrieval with Wikidata Ontologies	* Proposal for a heuristic assessment of the Wikidata ontology concerning structure of gender properties and classes * Analysis of performance in Wikidata

Table 2. Methodological Proposal for Ontology Evaluation

Evaluation criteria				
Extrinsic criteria	Measurement of external qualities	Analysis of external quality (structure)		
	Application	Context Design	Efficiency	
			Accessibility	
			Availability	
			Recoverability	
			Understandability/Clarity	
	Domain		Adaptability	
			Precision	
			Relevance	
			Full Functionality	
			Timeliness/Convenience	Relevance or Currentness
				Volatility
			Credibility	History
				Authority

Intrinsic criteria	Intrinsic domain features	Vocabulary semantics	Conciseness
		Architecture design	Coverage
			External Consistency
			Comprehensibility
	Intrinsic structural qualities	Syntax	Regulatory compliance
		Hierarchy	Complexity
		Architecture design	Internal consistency
			Modularity

Table 3. Proposed Indicators Used in the Standards Inspection Method for the Wikipedia Category Scheme (Taxonomy)

Indicator	Reference	Description	Methodology	Value
Evaluability	Alòs-Moner et al. (2010)	There are evaluation mechanisms in place to determine the levels of quality of the category scheme and to detect deviations over time.	Existence of agreed, approved, and disseminated procedures.	Binary value
Reusability	Alòs-Moner et al. (2010) Fraunhofer ISST and INIT (2009)	The category scheme must be useful in different classification scenarios and for use within Wikipedia, in whole or in part. The degree of reusability in each context will depend, to a large extent, on the requirements for specificity and comprehensiveness of that context. What data exchange format is available for the extraction and implementation of the category scheme.	Existence of agreed, approved, and disseminated procedures. Comparison of procedures with the content of the category scheme.	Binary value
Stability	Alòs-Moner et al. (2010)	The structure and chosen concepts must be long-lasting, unless the requirements of continuous updates recommend the incorporation of changes. In no case will categories requiring temporal updates be included (for example, current budget).	Analysis of temporal data on the creation of categories.	Binary value
Number of categories (concepts)	Alòs-Moner et al. (2010) Stock (2016)	Counting KOS categories (concepts) and comparing them with similar resources, with the	Counts are based on Wikipedia category data dumps, with comparisons to analog-	Comparison

		average number of documents per category as a supplemental dimension indicator.	format library catalogs and encyclopedias.	
Number of semantic relationships	Alòs-Moner et al. (2010) Stock (2016)	Calculation of semantic relationships between categories (concepts) in KOS..	The calculation is performed using data dumps related to Wikipedia categories.	Case study based on database dumps
Enrichment index or granularity	Alòs-Moner et al. (2010) Gil Leiva (2008) Lancaster (2002) Stock (2016)		Average between the total number of relationships and the number of categories. References indicate the maximum number of levels ranging from 2 to 5.	Optimal values
Degree of precoordination	Alòs-Moner et al. (2010) Lancaster (2002) Stock (2016)	Precoordination involves combining concepts at the time of category creation or when using them for categorization, as opposed to postcoordination, which involves users combining concepts during search.	The calculation is based on data dumps of Wikipedia categories, and computes the average between the number of meaningful words (nouns, adjectives, and verbs) in the categories and the total number of categories. References suggest a maximum number of levels ranging from 1.5 to 2.	Case study based on database dumps
Number of levels in the hierarchy or depth	Alòs-Moner et al. (2010) Stock (2016)	This considers categories linked by the hierarchical relationship in the same chain, from the top level to the lowest level.	The average is calculated between the total number of levels and the number of categories. References indicate a maximum number of levels at 5.	Optimal values
Number of categories at the same hierarchy level or breadth	Alòs-Moner et al. (2010) Fraunhofer ISST and INIT (2009) Stock (2016)	This takes into account the subcategories of all categories, from the top level to the one immediately above the lowest level.	The average is calculated between the sum of subcategories and the total number of categories (excluding the last-level categories). References	Optimal values

			indicate a minimum of 2 and a maximum of 12.	

Table 43. Method for assessing the quality of the Wikidata ontology

Indicator	Description
Instances used as classes	The "instance of" (P31) property only accepts classes as values, as indicated by its type "Wikidata property for the relationship of the element to its class" (Q28326730).
Disarray at the Upper Levels of the Ontology	The top level of the ontology should feature highly general classes (e.g., Time, Space, Event) independent of specific domains. These concepts must be mutually exclusive and collectively cover the knowledge domains of the ontology.
Semantic Deviation	An entity is seen from multiple perspectives, with distinct properties in each, but these merge into a single class. While individual subclass relationships are correct, their combined configuration is not.
Cycles or Loops in the "subclass de" (P279) Property	Class A has a subclass B, and class B is also a subclass of A, either directly or indirectly.
Redundant Generalization	Class A is both a subclass of B and a subclass of B's direct or indirect subclass.
Inconsistent Modeling	Differential treatment of two classes in terms of the number and types of classes they are linked to.
Repetition of Classes	The same class is defined multiple times.

Table 54. Approaches to accessing Wikidata

Access point	Description
Search (2023a)	Search in contexts where we can use known entity designations or specify queries based on simple data relationships.
Linked Data Interface with URI	The Linked Data Interface provides access to individual entities via URI: `http://www.wikidata.org/entity/Q???` For contexts where we need to retrieve individual and complete entities that we already know.
Wikidata Query Service (2023b)	In contexts with a known data structure pattern of three components (subject, property, object), it offers two interfaces: one for SPARQL experts and one for assisted query generation.

Table 6. Blazegraph repository statistics (February, 2022)

Indicator		Number
Contributors	Registered	565,000
	Unregistered (different IPs)	1.6 million
	Active per month	46,000
	Bots	359[1]
Elements		101 million
Properties		10,800
	For external identifiers	7,800
Statements		1,440 million
	For external identifiers	206 million
	Average per item	14.3
Editions		1,800 million
	Per day	699,000
Monthly page views	12 months average	420 million
Wikipedia articles using Wikidata	(January 2023)	75-97%
Wikipedia articles using Wikidata (caWiki)	(January 2023) Including article infoboxes, self-categorization, descriptions, and maintenance work indicators	90.6%

[1] Wikidata:Bots (2023b)

Table 7. Stability metrics of the Catalan Wikipedia 2005-22

Indicator	Value
Minimum number of new categories. Year 2005	866
Maximum number of new categories. Year 2021	8616
Median new categories. Year 2014	5621
Average number of new categories per year	1769.57

Table 8. Comparative Analysis of Log Entries: Feminized Professions Versus STEM Professions

Page title **“Category of…”**	**Visualizations**	**Daily Mean**	**Editions**	**Editors**
Teachers	299	1	1	1
Nurses	264	1	0	0
Librarians	235	1	1	1
Total Feminized Professions	798	3	2	2
Scientists	474	1	0	0
Engineers	313	1	0	0
Physicians	286	1	0	0
TOTAL STEM professions	1,073	3	0	0

Table 9. Results of the heuristic evaluation of the ontologies of Wikidata

Indicator	Evidence	Consequences
Instances used as classes	Asunción Estévez (Q115264435) is a subclass of woman (Q467), and Ilya Varlamov English (Q4103885) is a subclass of the law on foreign agents of Russia (Q17071473). In the classes contributing to the range restriction of property P31, there is a tendency to treat twins of individuals from each gender as subclasses of their corresponding gender. Technically, they are not class elements, and they can be subjects of "instance of" or "part of" properties, but not subclasses.	Unproductive class hierarchies will arise in navigation and search, as users will not be able to select more than one individual of the same type within nodes occupied by instances.
Disarray at the Upper Levels of the Ontology	At the ontology's top level, root classes exist, including the top-level class "entity" (Q35120). Among these, there are 201 presumed classes, some of which lack links to any superclass, like "comic" (Q58209506), and instances such as "Civil Code of the Republic of Korea" (Q5124449). Lower down, the class "entity" (Q35120) assumes this role. However, it has an excessive number of direct subclasses (59 as of 22/6/2023), and conceptually, they do not represent fundamental, distinct knowledge domains comprehensively.	*In search contexts requiring starting from the ontology's top level or ascending the hierarchy, the current state is unclear and non-intuitive, hindering quick decision-making. *In reasoning contexts involving ontology-related axioms, the disorder at the upper levels substantially impacts inference and consistency assessments. *In data cooperation and linking contexts with other projects, the absence of a standardized top ontology level hampers automated interoperability.
Semantic Deviation	The class "gender minority" (Q11894636) is seen from two angles: it is an indirect subclass of "collective entity" (Q99527517) and of classes in the intangible realm of representation like "abstract object"	Ascending through the upper chains of a class in search contexts leads to nodes with significant semantic distance, which can confuse users. This consequence can also apply in inference

	(Q7184903). Similarly, the class "māhū" is approached from two perspectives: as an indirect subclass of "collective entity" (Q99527517) and an indirect subclass of "abstract object" (Q7184903).	contexts using the transitivity of the "subclass of" (P279) property.
Cycles or Loops in the "subclasse de" (P279) Property	*"Linguistic unit" (wd:Q11953984) indirectly subclasses "emic unit" (wd:Q5371079), while "unitat èmica" (wd:Q5371079) directly subclasses "linguistic unit" (wd:Q11953984). *"Constituent" (wd:Q1786828) indirectly subclasses "unit" (wd:Q2198779), and "unit" (wd:Q2198779) directly subclasses "constituent" (wd:Q1786828).	In both search and axiom-based reasoning, determining which elements are more general or specific becomes impossible, transferring the loop to the search process.
Redundant Generalization	*"Male organism" class (wd:Q44148) is a subclass of "organism" (wd:Q7239). Additionally, "male" (wd:Q44148) is a subclass of "eukariote" (wd:Q19088), and "eukariote" is a subclass of "organism" (wd:Q7239). *"Identity" class (wd:Q844569) is a subclass of "quality" (wd:Q1207505). Furthermore, "identity" (wd:Q844569) is a subclass of "self-concept" (wd:Q1860557), "self-concept" is a subclass of "concept" (wd:Q151885), "concept" is a subclass of "aptitude" (wd:Q1347367), and "aptitude" is a subclass of "quality" (wd:Q1207505).	This redundancy complicates search and inference processes, hiding potentially crucial intermediate classes when opting for the shorter path.
Inconsistent Modeling	"Intersexual" (Q1097630) is a root class, while "female" (Q6581072)	In search contexts, it steepens the user's learning curve for a new

	and "male" (Q6581097) are embedded in hierarchical chains with up to 12 superclassification levels.	class structure as they cannot compare it to a previously familiar class with shared semantic connections.
Repetition of Classes	Two classes labeled as "physical objects" have two different identifiers: Q61961344 and Q98119401. Two classes labeled as "object" have two different identifiers: Q4406616 and Q488383. Two classes labeled as "folk culture" have two different identifiers: Q4384751 and Q4461766.	Redundancy confuses users during searches and hinders statement validation and new knowledge inference.

Table 10. Wikidata Performance Assessment

Query Metrics	Values
Good queries	5,242,253
Bad queries	157,791
Total query execution time	651,976
Total result rows	7.56 Bil