



Wikipedia's front page ten years evolution: Gender gap and implicit bias in biographies and editorial policies

Núria Ferran-Ferrer ¹

 0000-0002-9037-8837

Miquel Centelles ¹

 0000-0003-1739-4889

Laura Fernández ^{1*}

 0000-0002-0088-5414

¹ Department of Library and Information Science and Audiovisual Communication, Universitat de Barcelona, Barcelona, SPAIN

* Corresponding author: laurafernandez@ub.edu

Citation: Ferran-Ferrer, N., Centelles, M., & Fernández, L. (2025). Wikipedia's front page ten years evolution: Gender gap and implicit bias in biographies and editorial policies. *Online Journal of Communication and Media Technologies*, 15(4), e202535. <https://doi.org/10.30935/ojcm/17483>

ARTICLE INFO

Received: 8 Dec 2024

Accepted: 10 Nov 2025

ABSTRACT

Wikipedia, as one of the world's most visited digital web pages, serves as a global provider of knowledge. However, it is not immune to societal biases, including the gender gap that has been a persistent issue. We conducted a decade-long analysis (2013-2023) of 4,218 main pages from the English and Spanish editions of Wikipedia. Our study explores implicit biases—including language, occupation, ethnic group, religion, country and gender identity—among 22,924 biographies featured on these front pages. Beyond examining the selection of biographical content featured daily on the front page, we analyzed 31 editorial guidelines employed by the gatekeepers who determine which stories gain prominence on these pages. This study underscores the imperative for systemic change within the platform's editorial practices. Recognizing these issues is crucial for mitigating disparities in online knowledge representation and cultivating a more inclusive and diverse digital information landscape.

Keywords: Wikipedia, main page, gender gap, implicit bias, editorial policies, gatekeeping

INTRODUCTION

Information and communication technologies have thrived, promoting diverse expressions and an interconnected public sphere (van Dijk, 2012). Wikipedia, a key player in this contemporary sphere, holds the unique potential to enhance equitable knowledge production through collaborative efforts and virtual spaces for discourse, embodying principles of intellectual democracy (Tkacz, 2014).

Despite transforming information dissemination, Wikipedia still faces gender gaps and implicit biases in editing and content (Antin et al., 2011; Bear & Collier, 2016; Hinnosaar, 2019; Wagner et al., 2016). Following Redi et al.'s (2021) definition, the gender gap is 'the difference between readers of different gender identities in how and how much they access the sites', but the gender gap in Wikipedia extends beyond readers, affecting editors and content. Content-wise, only 20% of biographical articles feature women, highlighting a systemic issue across the nearly 300 Wikipedia language editions (Park & Bridges, 2022). And the platform's ambitious goal of achieving a 25% participation rate by female contributors remains unmet (Leonard & Bond, 2019). A study on the Spanish Wikipedia revealed a mere 13% of female editors (Minguillón et al., 2021). Consequently, a lack of diversity in participation leads to a corresponding lack of diversity in the resulting content. Volunteers' interests skew content, reducing diversity and deterring new editors, worsening the gender gap and perpetuating imbalances in content coverage (Worku et al., 2020).

Recent audits and longitudinal analyses show that these asymmetries are reproduced—and in some cases amplified—on Wikipedia’s main page, where selection functions as a powerful allocator of attention. Classic studies document sharp traffic spikes and lasting downstream effects when items are promoted (Reinoso et al., 2011), while governance work shows that editorial authority is negotiated by a small cadre of experienced contributors (Keegan & Gergle, 2010). More recent evidence indicates that representational imbalances can surface in the biographical content featured there (Sefidari Huici, 2022). Building on this, our decade-long study of the English main page (2011-2024) treats visibility as an outcome in its own right and shows how gendered and intersectional disparities persist over time (Centelles & Ferran-Ferrer, 2024; Ferran-Ferrer et al., 2024). The stakes are considerable: Wikipedia routinely serves ~350 million page views per day, meaning front-page choices shape what hundreds of millions encounter first (Wikipedia, 2025).

To address this vicious cycle, scholars emphasize the need for a comprehensive understanding of Wikipedia’s knowledge production culture (Menking & Erickson, 2015). Reducing the gender gap requires addressing the platform’s deeper logic, embedded in its techno-scientific project (Ford & Wajcman, 2017). This involves recognizing and dismantling exclusionary practices to make Wikipedia more robust, reliable, and transparent (Menking & Rosenberg, 2021).

Communication theories, such as gatekeeping, offer valuable frameworks for understanding Wikipedia and its potential discrimination. The gatekeeping theory, which integrates communication and information science, focuses on the process by which information is filtered for dissemination (Barzilai-Nahon, 2009). Applying gatekeeping theory to Wikipedia, our study focuses on scrutinizing stories selected for the front page—the platform’s most visited page, known as the main page internally, with over 148 million readers in January 2023 (Wikimedia, 2023b). Beyond confirming Wikipedia as a reflective mirror of society skewed by the patriarchy, this study delves into the systemic issues the encyclopedia faces (Ford & Wajcman, 2017). The analysis includes not only content selection on the main pages but also examines the editorial policy guidelines shaping gatekeepers, enhancing understanding, and addressing the systemic problem. We compare the English and Spanish editions, analyzing daily content, editorial guidelines, and insights from volunteer communities through the lens of feminist media studies (Harvey, 2020) and gatekeeping theory (Barzilai-Nahon, 2009).

OBJECTIVES AND RESEARCH QUESTIONS

The primary objective of this research is to comprehensively examine implicit biases within Wikipedia biographies highlighted on the front page, focusing on the daily English, and Spanish language editions spanning a decade (2013-2023). We understand biases as knowledge distortions that reproduce systematic discrimination along various markers of difference, which we analyze from a feminist and intersectional perspective (Collins & Bilge, 2020; Harvey, 2020). We based our understanding on Bernt Rasmussen’s (2020) conceptualization of ‘implicit bias’, which she describes as mental processes beyond our direct control or endorsement that influence our perceptions, judgements and behaviors toward members of socially salient groups. Biases are therefore referred to in terms of language, occupation, ethnicity, religion, country of origin, and gender identity.

Wikipedia’s ‘main page’ functions both as the homepage of a digital platform and as the front page of a digital media, serving as the entry point with features akin to a front page, offering sections like ‘today’s featured article’ and ‘in the news,’ essentially acting as the digital encyclopedia’s daily cover. Through a comparative lens, this research seeks to provide insights into systemic challenges and practices. The ultimate goal is to advocate for positive changes that promote a more equitable and diverse online information environment.

Therefore, the research questions (RQs) that we address are:

- RQ1:** How prevalent are implicit bias based on language, occupation, ethnic group, religion, country and gender identity in the biographical content featured on daily Wikipedia’s front pages?
- RQ2:** How do the editorial guidelines for volunteer gatekeepers influence the selection of front-page stories on Wikipedia?

MATERIALS AND METHODS

Context of the Study

Wikipedia, a top global information source, ranks among the world's ten most visited platforms (Wikipedia, 2023a). Its main page alone garnered over 46.8 billion views as of January 2023, surpassing the combined viewership of Wikipedia's entire top-100 list (Wikipedia, 2023a). Currently, Wikipedia encompasses 336 distinct editions dedicated to various languages and cultures. This research paper examines a decade-long analysis of the front pages from two prominent Wikipedia editions: English, the largest and most researched, and Spanish, among the next largest and extensively studied (Wikipedia, 2024).

The English edition, launched in 2001, is the largest with around 7 million pages (Wikimedia, 2023a) and over 46 million registered users (Wikipedia, 2023b). It dominates page views, averaging 97.2 billion monthly, six times more than the next largest edition (Anderson et al., 2016).

Quantitative Analysis of the Biographical Articles of the Front Pages

Wikipedia's main pages display dynamic content daily, featuring current events, highlighted content, and various useful links (Table 1). Managed by volunteers, these pages evolve in format and content. This study identifies consistently present sections on Wikipedia's main page, focusing on biographical entries about individuals from March 2013 to February 2023 across all language editions.

Table 1. Wikitexts and types of articles on the Spanish Wikipedia's main page

Wikitext	Description	Type of article
{{Historia artículo AB}}	Current good article	Good articles
{{Historia artículo ABA}}	Former good article	
{{Historia artículo AXB}}	Current good annex	
{{Historia artículo AXBA}}	Former good annex	
{{Historia artículo AD}}	Current and never-discussed featured article	Featured articles
{{Historia artículo RAD}}	Current and discussed featured article	
{{Historia artículo RADR}}	Former, discussed and no longer featured article	

We began by retrieving URLs of articles linked to Wikipedia's main pages over the specified decade. Filtering focused on articles about individuals, with data gathered from Wikidata for gender and other intersectional aspects. OpenRefine managed this pipeline effectively, handling tasks like cleaning and enriching datasets across multiple language editions. Retrieving historical main page content involved direct retrieval due to unavailable dump URLs.

The English edition of Wikipedia archives its main pages and their content from January 1, 2011, onwards, accessible at 'Wikipedia: main page history' (https://en.wikipedia.org/wiki/Wikipedia:Main_Page_history). Our task involved extracting articles from specific sections between March 2013 and February 2023: 'today's featured article', 'did you know ...', 'in the news', 'on this day ...', and 'today's featured picture'. Over this period, there were 4,218 main pages recorded, with some days having multiple editions and others with no main page published. Each main page URL followed a consistent pattern, and we compiled a list of these URLs for access throughout the entire time frame.

In contrast, the Spanish edition of Wikipedia lacks a dedicated archive for main pages. Instead, we utilized a template called {{historia artículo}} found in the good articles ('artículos buenos-AB') and featured articles ('artículos destacados-AD') sections. This template's usage was tracked through its wikitext on the page 'páginas que enlazan con «plantilla artículo»' revealing 7,018 instances between March 2013 and February 2023. These instances are discussed in threads and include wikitexts associated with good and featured articles, detailed in Table 1 along with their meanings and article types.

From these discussions, URLs of the final articles were reconstructed based on titles included in the discussion page. For instance, the discussion's URL <https://es.wikipedia.org/w/index.php?title=Discusión:Thalía&action=edit> was adjusted to form <https://es.wikipedia.org/wiki/Thalía>. A list of 849 article's URLs from the good and featured article sections throughout the considered period was compiled.

Table 2. Definitions and justifications of Wikidata properties for the analysis of biography content

Wikidata Code	Property	Definition	Justification
P1412	Languages	Language(s) that a person or a people speaks, spoken, written, writes, or signs, including the native or signed language(s).	Linguistic differences can result in inequalities in the visibility of certain groups (Pennycook & Otsuji, 2015).
P103	Native language	Language or languages a person has learned from early childhood.	One's mother tongue can highlight inequalities in representing ethnic and linguistic groups (Spack, 1997).
P106	Occupation	Occupation of a person. Also known as: profession, job, work, career, employment, craft, vocation.	Occupations are often subject to cultural and social stereotypes. Some groups may be stigmatized or marginalized in certain professions, which affects their visibility and representation on platforms such as Wikipedia (Kalleberg, 2018).
P172	Ethnic group	Subject's ethnicity. In general, it means (1) the subject claims it themselves, or (2) it is widely agreed on by scholars, or (3) is fictional and portrayed as such.	Individuals belonging to underrepresented ethnic groups may face inequalities in online representation, which can lead to stereotypes and biases (Carroll, 2017).
P140	Religion or worldview	Religion of a person, organization, or religious building, or associated with this subject. Also known as: religious affiliation, faith, life stance, pantheon, belief system, etc.	Religion can be an important factor in a person's identity and can lead to stigmatization or biases (Knott & Poole, 2016).
P495	Country of origin	Country of origin of this item (creative work, food, phrase, product, etc.).	These properties are related to nationality and migration. People from different countries can experience varying levels of representation and attention (Anthias, 2013).
P27	Country of citizenship	The object is a country that recognizes the subject as its citizen.	
P21	Sex or gender	Sex or gender identity of human or animal. For human: male, female, non-binary, intersex, transgender female, transgender male, agender.	

Both sets of URLs (English and Spanish editions) were processed separately in OpenRefine. Using the 'add column by fetching URLs' feature, we retrieved the HTML code from web services for each listed URL. This generated a new column in OpenRefine containing the HTML code of each URL. For the English edition, this column captured the entire main page HTML, while for the Spanish edition, it retrieved HTML related to good and featured articles. Although Wikipedia's API was an option, we chose OpenRefine due to its reliable performance across the entire data handling process.

Next, the HTML code underwent content extraction using OpenRefine's parsing capabilities, utilizing GREL scripting and Jsoup Java HTML parser. For the English version, URLs of referenced articles in the five specified sections and Wikidata QIDs were extracted. In the case of the Spanish version, only QIDs were extracted.

The subsequent step involved using Wikidata to identify the type of entity each article referred to, focusing on those related to human beings. In OpenRefine, this process utilizes two functionalities sequentially. First, 'reconciling with unique identifiers' allowed simultaneous queries of QIDs to the Wikidata reconciliation service, linking with entity properties. Second, using 'data augmentation,' specific properties like 'instance of' (P31) were selected to filter articles categorized as 'human'. This isolated Wikipedia biographies of human beings. Additional values related to potential implicit biases were then obtained using the same 'data augmentation' functionality, extracting properties from Wikidata (refer to [Table 2](#) for details).

The pipeline ended by downloading two datasets: one containing 22,075 biographies from the main pages of the English edition, and another with 849 biographies from the main pages of the Spanish edition (please see the Data Availability Statement for more information on this).

These biographies originate from seven different sections of the English and Spanish Wikipedia front page (refer to [Table 3](#)).

Table 3. The analyzed sections on the English and Spanish Wikipedia's main page

Section	Description	Equivalent in the Spanish edition
<i>Today's featured article</i>	This section features the daily 'featured article' from Wikipedia, showcasing diverse high-quality content on a daily rotating basis.	<i>Artículo destacado</i>
<i>In the news</i>	This section provides summaries of recent news events and links to Wikipedia articles related to those events.	<i>Actualidad</i>
<i>Did you know</i>	A section featuring interesting and often lesser-known facts about various topics, accompanied by links to relevant articles.	This section does not exist in the Spanish edition
<i>On this day</i>	A historical section that highlights significant events that occurred on the current date throughout history.	<i>Efemérides</i>
<i>From today's featured list</i>	This section highlights a featured list of notable people, places, or other items.	<i>Artículo bueno</i>
<i>Today's featured picture</i>	A section that showcases a high-quality image from Wikimedia commons.	<i>Recurso destacado</i>
This section does not exist in the English edition	A portal is a Wikipedia page that serves as a cover page for articles in the same topic or area. Portals are intended for both editors and readers of Wikipedia.	<i>Portales</i>

Table 4. Front page guidelines: English and Spanish Wikipedia editions

WP edition	General guides for each section	Specific guides for each section	Total number of guides
English	Main page		24
		Featured article candidates	
		Featured list candidates	
		Featured picture candidates	
		Featured article criteria	
		Featured list criteria	
		Featured picture criteria	
		Featured and good topic candidates	
		Featured and good topic criteria	
		Today's featured article/ requests	
Spanish	Featured article	Picture of the day	7
		About today's featured article	
	In the news		
	Did you know /guidelines (DYK)		
	Selected anniversaries		
	Good articles		
	Syndication		
	Información de 'Wikipedia discusión: Portada'		
	¿Qué es un artículo destacado?		
	Candidatos a recursos destacados		
Spanish	Plantilla: Efemérides		7
	Selección de artículos buenos		
	Portal		
	Tutorial VECAD		

Analysis of the Editorial Guidelines

Each section of the front page provides volunteers access to 24 support pages in English and 7 in Spanish, aiding content selection, writing, and editing with precise editorial guidelines on procedural steps and style preferences crucial for Wikipedia contributors.

We carried out a content analysis of each of the editorial guidelines for each section of the front pages, of both the English and Spanish editions of Wikipedia. These editions have implemented structured editorial practices aimed at guiding volunteers in their role as gatekeepers for six of the sections of the front page. See [Table 4](#) for the list of editorial guidelines analyzed. Each of the guides for its corresponding section were codified using a qualitative data analysis software according to an *ad hoc* codebook (see [Table 5](#)).

Table 5. Codebook for the analysis of Wikipedia's editorial guidelines about the front page

Categories	Descriptions
Sections	Titles or links to pages on Wikipedia that outline editorial policies and guidelines
Responsible for the section	Person/bot
Interaction: Feedback	Option of readers and Wikipedians to suggest, send a message because there is an error
Interaction: Appropriation	Option for the Wikipedians to appropriate the content of a given section
Guidelines	Eligibility criteria
Number abbreviations	Total number of abbreviations of names of pages, names of rules, names of guidelines, etc. that appears in the document for coding
Code	Total number of instructions of how to code a page in Wiki language that appears in the document for coding Values: - Intelligible: codes that are self-explanatory - Unintelligible: codes that require subjective interpretation or knowledge
Quality criteria	The criteria of quality explained in the guidelines and policies & redirection of content in relation to the Wikimedia movement strategy 2023
Notability	The criteria to define notability

RESULTS

Implicit Bias in the Content on Wikipedia's Front Pages (RQ1)

In this section, we present the findings of our analysis using a feminist media studies lens, examining the attributes of the prominent individuals featured on the front pages of both English and Spanish Wikipedia.

Languages (P1412 and P103)

In the English edition of Wikipedia, an analysis of the 'language' attribute reveals a striking dominance of the English language. The term 'English' occurs 6,836 times, surpassing the second most frequent language, 'French,' by a factor of four, with 1,590 occurrences. The dataset encompasses a total of 17,323 values across 294 different languages. Notably, the top four languages collectively account for 61.73% of all occurrences; these are four languages official of Western European countries with historical connections as languages of former colonies or ex-colonies.

English, as a native language, is highly represented and occurs three times under different variations (English, American English, and British English). Furthermore, the data indicates that the six official languages of the UN constitute over 62% of the values represented on the English Wikipedia main pages during the specified period, with English being the most frequently repeated, accounting for nearly 40% of these occurrences.

A parallel pattern emerges in the Spanish edition of Wikipedia, where the 'language' attribute reveals that 'English' (261 occurrences) surpasses 'Spanish' (173 occurrences) by a substantial margin. It is important to remember that the analysis focuses on the Spanish Wikipedia edition. The dataset contains 824 values spread across 76 languages, with the top four languages collectively representing 69.44% of all occurrences. These four languages, as in the English edition, share the characteristics of being official languages in Western European countries and languages of former colonies or ex-colonies.

In the 'native language' attribute, similar to the English edition, 'English' occurs three times (as English, American English, and British English). Interestingly, the data highlights that the six official languages of the UN account for a significant 72.05% of the values represented on the Spanish Wikipedia front page during the specified period, with English being the most frequently repeated as a native language, constituting 61% of these occurrences.

Occupation (P106)

In the English edition of Wikipedia, an examination of the 'occupations' categories shows a notable bias in favor of politicians. The category 'politician' appears with 4,835 occurrences, surpassing the second-ranked category, 'writer,' which has 2,447 occurrences, nearly doubling its frequency. The dataset comprises a total

of 52,334 recorded values across the front pages, and the top 20 categories, as presented here, out of a total of 1,944 distinct occupations, collectively represent 37.11% of the data.

Further analysis of this data reveals a concentration of occupations that are predominantly linked to the realm of politics and writing. This concentration might indicate a bias toward these specific occupational categories in the English edition, raising questions about the diversity of representation and the potential influence of these biases on knowledge dissemination.

A similar pattern is observed in the Spanish edition of Wikipedia, where the 'occupations' categories exhibit a parallel bias. The category 'politician' also holds the top position with 172 occurrences, and 'writer' ranks second with 126 occurrences, mirroring the English Wikipedia's arrangement. The dataset in the Spanish edition includes 2,675 recorded values for this attribute across the main pages, with the top 20 categories, as presented, collectively representing 40.71% of the 503 unique occupations identified.

Notably, the data highlights that half of the mentioned occupational categories in the top 20 are closely related to the realm of cultural enterprises. This pattern reflects a concentration of occupations in certain fields, potentially influencing the diversity of knowledge and representation within the Spanish edition of Wikipedia.

Ethnic group (P172)

An examination of the 'ethnic group' attribute in the English edition of Wikipedia reveals a notable pattern. The term 'African Americans' appears 288 times, sharing this frequency with 183 other values. The top 20 values within this dataset collectively account for 25.9% of all occurrences. In contrast, Europe's representation is comparatively low at 8.3%, encompassing 33 values, while Latin America's representation is even lower, at just 0.96%, with 17 occurrences out of a total of 1,761. These findings suggest that, within the context of Wikipedia, 'white-European-Caucasian' may be perceived as the default or the absence of an ethnic group. This discrepancy in representation raises questions about the extent to which Wikipedia's coverage of ethnic groups is equitable and reflective of the world's diversity.

In the Spanish edition of Wikipedia, a similar pattern is observed in the 'ethnic group' attribute. The term 'Afroamericanos' (Afro-Americans) occupies the first position, mirroring the English Wikipedia's representation. Additionally, the top 20 values in this dataset collectively account for 42.07% of all occurrences. Notably, English-speaking ethnic groups, such as Indian migrants in England, Italian Americans, as well as British and Irish groups, constitute over a third of this category. This data suggests that the representation of ethnic groups on the Spanish Wikipedia front pages is also influenced by English-speaking cultures and may not fully encompass the breadth of global diversity.

Religion (P140)

A prominent trend emerges in the English edition of Wikipedia, revealing a notable bias towards Abrahamic religions. Among the top 20 values, 15 are associated with Abrahamic religions. Particularly striking is the dominance of Christianity, with all values collectively representing 63.14% of the dataset, equivalent to 12,216 occurrences out of a total of 19,347. This dominance is in stark contrast to the global distribution of major religious groups, where Christianity holds 30.1% of global adherents, and Islam claims 25.1% (Religious Information by Country, 2024). This disparity in the representation of religious affiliations within the English edition of Wikipedia raises questions about the diversity of religious perspectives and the potential influence of these biases on knowledge dissemination.

In the Spanish edition of Wikipedia, a similar pattern emerges, highlighting a bias in favor of Abrahamic religions. Of the top 20 values, 15 are related to Abrahamic religions. Moreover, all values linked to Christianity collectively account for 67.62% of the dataset, comprising 236 values out of a total of 349. This data suggests that the Spanish edition of Wikipedia also exhibits a bias towards Abrahamic religions, particularly Christianity. It raises concerns about whether this representation fully reflects the religious diversity of the Spanish-speaking world.

Country (P495 and P27)

The analysis of the 'country of origin' attribute in the front pages of the English edition of Wikipedia reveals a significant bias, with the USA accounting for a predominant 49.06% of the appearances on English Wikipedia covers. The top 10 values, including the UK, Japan, France, Italy, India, Dutch East Indies, Spain, Canada, Germany, and the USA, collectively make up an overwhelming 86.15% of the total, indicating the concentration of these countries. Furthermore, Africa and Latin America are notably underrepresented, with only 7 out of the 87 countries from these two regions being mentioned in this attribute.

Similarly, in the case of 'country of citizenship' in the English edition, the USA maintains a significant presence, accounting for 24.01% of the total values. The UK, Great Britain, and Ireland follow with 12.35%. Notably, very few countries outside the Global North manage to achieve triple-digit figures, highlighting the dominance of a select group of countries, including Argentina, Brazil, India, China, Indonesia, and Mexico. The geographic bias observed in both attributes implies that the representation of countries on the English Wikipedia front pages is skewed towards a select group, which may not fully reflect the global diversity of countries and citizenships.

A similar pattern of geographic bias is observed in the Spanish edition of Wikipedia. In the case of 'country of origin,' the USA maintains its predominance, accounting for 52.39% of the appearances on Spanish Wikipedia covers. Additionally, 14 current European countries collectively occupy 20.99% of a total of 99 territories, both current and former.

Spanish-speaking countries or regions represent only 13.55% of the total values, with Spain contributing to half of this representation. This data reveals the dominance of certain countries, particularly from Europe and the USA, and a limited representation of Spanish-speaking regions.

In the 'country of citizenship' attribute on the Spanish Wikipedia edition, the USA, again, holds a significant presence, accounting for 18.47% of the total values. Spain follows with 9.23%, and the UK stands at 7.19%. Notably, only Argentina, with 50 values, has a double-digit representation, while the rest are from the Global North. Furthermore, it is noteworthy that Africa and Asia are not represented in this attribute, raising concerns about the underrepresentation of countries from these continents.

Gender identities (P21)

The analysis of gender representation on the front page of the English edition of Wikipedia uncovers a significant gender imbalance. Among the 22,075 records analyzed, men dominate the representation at a staggering 70.8%, while women account for just 29%, and individuals with other gender identities are exceedingly rare, making up only 0.2%.

Furthermore, a noticeable trend has emerged over the years, indicating that the gender gap has not significantly decreased. The year with the most balanced gender representation was 2017, with a gap of 630 values. However, in 2022, the gap had increased to 1,191, almost doubling the difference. Despite global initiatives to increase the visibility of women on Wikipedia, this trend has not reversed, highlighting persistent gender disparities. Women have experienced an increase in representation, with 317 appearances on the English Wikipedia cover in 2014, peaking at 1,035 entries in 2021, marking more than three times the representation compared to 2014.

Similarly, the analysis of gender representation on the front pages of the Spanish edition of Wikipedia reveals a significant gender disparity. Out of a total of 849 records for this property on the Spanish Wikipedia front page, men are overwhelmingly represented at 81.62%, women at 18.13%, and other gender identities at a minuscule 0.23%. The data indicates that the gender disparity is even more pronounced in the Spanish edition, with a higher percentage of men and a lower representation of women and other gender identities compared to the English edition.

The analysis of gender representation on the front page of Wikipedia in both English and Spanish editions reveals notable imbalances. Men significantly outnumber women, underscoring persistent gender disparities despite global initiatives to increase visibility. The trend over the years indicates a lack of substantial improvement. Efforts to address and rectify these gender disparities on Wikipedia remain crucial for fostering a more inclusive representation of diverse voices and perspectives.

Editorial Guidelines (RQ2)

In our analysis of the editorial guidelines for English and Spanish Wikipedia, we reviewed a total of 31 webpages (see [Table 3](#)) used as policies and guidelines by the editor community, including 17 for English Wikipedia and 7 for the Spanish. From this sample, we detected a total of 210 external references to new sections that provide users with a guideline to edit Wikipedia or expand information about any of its sections. These sections typically appear as hyperlinks embedded within the guidelines. In English Wikipedia, we identified 175 such references, while in Spanish Wikipedia, there were 35 hyperlinks leading to new sections.

The web pages outlining editorial policies for editing Wikipedia's front page emphasize that editing is open to everyone. Daily editing responsibilities primarily rest with a community of volunteer editors. Each section's responsibility varies, largely handled by Wikipedians rather than bots. For instance, only users with the role of administrators in English Wikipedia can manage last-minute changes to the main page and oversee moving content from the 'did you know' section template. Notably, only two bots are mentioned: one in English Wikipedia for the 'featured articles' section and one in Spanish Wikipedia, which updates the 'efemérides' section template daily.

In Spanish Wikipedia, edit count is the primary criterion for various tasks: for voting on featured resources, a user needs a minimum of 500 contributions and six months since their first edit; 100 contributions are required for CAD or VECAD reviews, and 50 edits qualify users to update article statuses on Wikidata. While all Wikipedia editors' contributions are valued, experience, management roles, and higher edit counts play a significant role in determining consensus for tasks like selecting Featured Lists in English Wikipedia or approving/rejecting good articles in Spanish Wikipedia.

Regarding biased content on the Wikipedia front page, there is a template aimed at 'minimizing bias effects.' However, its effectiveness hinges on the topics editors contribute to. Other templates, like those in the 'did you know' and 'candidates featured list' sections, also acknowledge implicit biases. Wikipedia encourages global user participation and feedback at different stages, though tasks such as voting for good articles require prior editing history and administrator roles for main page edits. We identified 17 prompts for reader and editor suggestions, questions, error reports, or feedback (12 in English guidelines, 5 in Spanish). These prompts often refer to Wikipedia's code of conduct, emphasizing constructive input and interaction, though such messages are relatively few.

Regarding the formal aspects of the editorial guidelines, the vocabulary often includes specialized language with numerous abbreviations. We identified over 90 abbreviations in the guidelines, with English Wikipedia using them much more extensively than Spanish Wikipedia, which had only 14. The most common abbreviations in English were the name of certain sections (e.g., DYK: did you know; ITN: in the news; POTD: picture of the day) and frequently used concepts for each section (e.g., FT: featured topic; GT: good topic; FTQ: featured topic question; FPC: featured picture candidates; AfD: articles for deletion, or QPQ: quid pro quo, which is a requirement for nomination). In the Spanish guidelines, the most repeated were AB (artículo bueno) and AD (artículo destacado), including certain formulas that are frequent in a given guideline (e.g., CAD: candidato a artículo destacado).

Regarding formal aspects, the language used in the guidelines includes references to computer code. We identified a total of 53 intelligible codes referenced: 22 for the Spanish edition and 31 for the English edition. Intelligible codes are those that any reader of Wikipedia guidelines, even without coding background, could understand, such as `{{citation needed}}`. Unintelligible codes appeared more often, particularly in the English guidelines: we found 88 codes, 26 for the Spanish case and 62 for the English case. We refer to unintelligible code when the guidelines refer to more complex code that includes a lot of abbreviations and/or requires training to understand, for example: `& feedformat = atom or <s> ... </s>`.

Both English and Spanish Wikipedia editions adhere to fundamental quality pillars like notability and neutrality in their editorial policy guidelines. While they share similar criteria, each edition has unique sections shaped by their respective communities. For instance, English Wikipedia's featured articles guidelines define articles as 'well written, comprehensive, factually accurate, neutral, stable, and compliant with Wikipedia's copyright policy, free of plagiarism or too-close paraphrasing.' Featured articles on Wikipedia are distinguished by professional prose, an engaging lead that defines scope and inclusion criteria, comprehensiveness, and a summary style. Nominations consider the time since a similar article was featured

and the article's relevance. Good articles, on the other hand, must be well-written, contain accurate and verifiable information, provide broad coverage, maintain a neutral viewpoint, stability, and include suitable images with proper copyright licenses whenever possible. Featured pictures must meet high technical standards, have a free-content license, and enhance relevant articles. Picture quality is judged by its display in high resolution and its status as an exemplary representation of its subject. The 'did you know' section emphasizes hooks based on established facts that are unlikely to change and should be neutral and relevant beyond novelty. Articles featured here need not be the best but should offer new and improved content. Additionally, there's a guideline on representativeness, aiming for a balance between US and biography topics in updates. In the 'in the news' section, it's stated that the inclusion of topics is often contentious. The primary criteria are that the event must be current or have occurred recently and must have consensus for posting.

In the Spanish edition, the editorial guidelines are mostly provided for the 'artículo destacado' ('featured article') and the 'efemérides' ('on this day') section. The featured article must meet quality criteria (verifiable sources, neutrality, clarity, comprehensiveness, stability) and include images, tables, or multimedia from Commons. In the Spanish 'efemérides' section, guidelines specify years of inclusion, content quality (references, images from commons, Wikimedia project inclusion), and relevance to the Spanish-speaking world. When selecting between multiple events or individuals, guidelines provide: 'If after applying all of the above several options remain without notable differences, preference is given to women, little known countries, unusual professions, biographies of interest, exceptional events, etc., and a balance is sought between recent and old dates, and between the countries represented.'

Regarding content redirection, which is part of the Wikimedia movement strategy 2030, there's no direct guidance on how to balance content. In the 'in the news' sections, there are no strict rules; the key criteria are that the content must be current and recently covered in news sources. Editorial guidelines emphasize avoiding topics driven by commercial or political interests, especially during events like elections or product launches. Articles must also relate specifically to events that occurred on the publication date.

Editorial policies for supplementary content, like photographs, offer practical recommendations such as specifying minimum pixel dimensions and prioritizing images where individuals face left to guide reader attention naturally.

In Spanish Wikipedia, there's concern about implicit biases in content, especially in sections like 'efemérides' ('on this day'), where guidelines prioritize minorities in terms of gender, origin, and professions when options are similar.

The selection process for front-page articles, such as 'article of the day,' involves criteria like edit count, article age, and adherence to Wikipedia's quality standards. The process spans four detailed pages outlining how to propose articles.

DISCUSSION AND CONCLUSIONS

The analysis of Wikipedia's front page biographical content, focusing on gender and intersectional disparities, reveals significant implicit biases across various attributes during the decade 2013-2023. In both the English and Spanish editions, English language dominance is evident, reflecting historical and colonial influences. Additionally, the representation of native languages mirrors the dominance of English, with implications for cultural inclusivity and representation. A significant bias towards certain countries, notably the USA, is evident in both editions, reflecting a skewed representation of global diversity.

Regarding the occupation bias, a notable bias towards politicians and writers is observed in both editions, given that these professions are not the ones with the largest presence in Wikipedia as a whole. This data raises concerns about the diversity of representation and its influence on knowledge dissemination. The popular press has highlighted disparities in coverage between male-associated topics, for instance, Kleeman (2015) compared the list of pornographic actresses to the list of female poets in *The New Statesman*. The list of female porn stars has attracted over 1,000 editors, with more than 2,500 edits, while the list of poets has been edited 600 times by nearly 300 editors (Kleeman, 2015). Surprisingly, despite women's main occupations on Wikipedia being related to the world of entertainment and adult cinema, these occupations are not featured on the front pages. Instead, women who appear on the front pages are mainly politicians (9.24% in

the English Wikipedia and 6.43% in the Spanish Wikipedia) and writers (4.68% in the English Wikipedia and 4.71% in the Spanish Wikipedia).

In terms of ethnicity, there is a notable underrepresentation of certain ethnic groups, particularly from Africa and Latin America, suggesting a default perception of 'white-European-Caucasian' as the norm, as well as an ethnicity-free identity. This raises questions about Wikipedia's equitable coverage of ethnic diversity: the most represented ethnicities are non-white, which reveals a colonial/ethnic bias based on the understanding of whiteness as the norm and the absence of racialization. Viewing gaps on Wikipedia through the lens of a systemic issue (Ford & Wajcman, 2017) reveals that they stem from the fact that the majority of editors share similar social and cultural characteristics (Miquel-Ribé & Laniado, 2021), as well as from the culture, dynamics, and values of online communities. This can result in a content gap but also a contributor representation disparity. These events focus on creating and editing feminist content while addressing misogynistic language, particularly to improve the representation of African women (Bear & Collier, 2016; Ukwoma et al., 2021) and serve as educational experiences (Hood & Littlejohn, 2018), empowering participants to challenge implicit biases and power dynamics on Wikipedia.

Concerning the religion bias, Abrahamic religions, particularly Christianity, dominate the representation, indicating biases in religious perspectives and knowledge dissemination.

In reference to gender identities, men significantly outnumber women on Wikipedia's front pages in both editions, highlighting persistent gender disparities despite efforts to increase visibility. Women's presence on the platform remains limited, with only 13.2% to 22.5% of biographies covering females, as noted by Tripodi (2023), accounting for only 19% of total coverage. This disparity permeates various aspects of Wikipedia, including article deletion rates, writing length, lexical choices, classification methods, and network positions. Moreover, disparities in multilingual notability underscore gender discrepancies in the dissemination of biographies across different languages, potentially influencing international audiences' perspectives. Several studies have shed light on the manifestations of gender bias on Wikipedia, examining its intersections with race and sexuality and the resulting experiences of safety and marginalization on the platform (Ju & Stewart, 2019; Lam et al., 2011; Toupin, 2020; Tripodi, 2023).

In relation to the analysis of editorial guidelines for volunteer gatekeepers in English and Spanish Wikipedia sheds light on their influence on front-page story selection.

Although about thirty guides were analyzed, ultimately, hundreds of additional references were linked from the guides, expanding on the information contained in the guides, especially in English Wikipedia, which is much more detailed and delves deeper into details by associating web pages with the initial guides.

The guidelines place all responsibility for selecting the content for the main page on the participating volunteer editor. Bots play a limited role in updating specific sections, but the full responsibility for updating and managing different sections largely falls on Wikipedians. In the English Wikipedia, administrators have the authority to make last-minute changes to the front page and manage certain sections. This responsibility, which falls exclusively on volunteers who must work on content publication for the main page every day of the week, leads volunteers, driven by their interests, to often prioritize specific topics. This undermines diversity (Worku et al., 2020), deepening the gender gap.

On Wikipedia, roles are generally assigned based on the number of edits and contributions, giving more influence to experienced users and those in managerial roles. The platform's governance is shaped by power dynamics established since its inception. The underlying infrastructures of Wikipedia are rooted in foundational epistemologies that often marginalized women and underrepresented groups whose knowledge diverges from established standards. Wikipedia's foundation is built upon existing infrastructures, including those of traditional encyclopedic projects and the open Internet (Star, 1999). Scholars such as Ford and Wajcman (2017) have shed light on the emerging and often overlooked sources of gender disparity within Wikipedia's infrastructure.

The use of abbreviations and references to computer code vary between the English and Spanish guidelines, reflecting differences in language conventions and technical aspects. Overall, the guidelines shape the selection process by delineating roles, criteria, and community engagement practices. Our analysis of editorial guidelines aligns with previous research pointing to the ambiguous nature of Wikipedia's policies, which leads to the reconsideration of the nature of 'neutrality' practiced on Wikipedia (Matei & Dobrescu,

2011) as well as with previous research that demonstrates that internet skills are gendered (Hargittai & Shaw, 2015). The analysis of editorial guidelines highlights the need for improvements in accessibility related to the main page. Simplifying the guidelines by reducing abbreviations, technical language, and the volume of pages, sections, and hyperlinks could make them more accessible to novice editors and those unfamiliar with Wikipedian jargon. Extending participation requirements in discussions could further democratize the editing process, potentially diversifying contributors and broadening editing interests while reducing gaps (Worku et al., 2020).

Community input is also considered, with calls for suggestions and feedback, although these interactions are relatively limited compared to the entire set of guidelines. As indicated in the guidelines, in order to be able to participate in the decision-making process, editors must have a minimum number of editions and certain years of experience editing Wikipedia. In particular, editors need to have sufficient time to invest in discussions about article deletions. The academic literature on the gender gap in Wikipedia notices that time distribution is also gendered and societal gender differences, such as care burdens, results in important variations in discretionary leisure time (Eckert & Steiner, 2013; Gruwell, 2015; Ferran-Ferrer et al., 2021).

Considering the implicit bias present in the Wikipedia biographies, we consider a feminist revision of the 'notability' parameter is necessary. 'Notability' on Wikipedia determines whether a topic deserves its own article by requiring significant coverage in reliable, independent sources external to the subject. This criterion ensures that articles meet Wikipedia's standards of quality, reliability, relevance, and verifiability. However, for items selected for the main page, editorial guidelines prioritize article quality, thorough citations, recent updates, neutrality, comprehensiveness, stability (free from edit wars), and compliance with Wikipedia standards, rather than sorting by importance or significance alone.

Wikipedia's focus on article quality over significance differs significantly from journalistic norms for front-page content. Additionally, the reliance on secondary sources to maintain neutrality can indirectly disadvantage underrepresented groups. Media coverage of issues affecting underrepresented groups and subjects facing systematic discrimination is disproportionately lower compared to more privileged profiles (Ramírez-Ordóñez & Ferran-Ferrer, 2023). This lack of coverage limits their visibility on Wikipedia, affecting their chances of having a biography published. This disparity is particularly pronounced for women, who struggle to attract mainstream media attention and are underrepresented, with only 24% of news sources featuring women. Moreover, topics where women are most visible receive the least coverage (Macharia, 2020). The case of Spain is even worse, women are represented as the main character at just 7.3% in television news (Matud et al., 2021).

In summary, this research holds that implicit bias in Wikipedia main pages are perpetuated or even exacerbated: on the one hand, if the existing majority content considered most notable is about biographies of men or topics of traditionally male interests (Worku et al., 2020), this is what prevails on the main page as well. For this over-representation to take place, it is worth noting the influence of phenomena such as a large majority of male editors (Minguillón et al., 2021) and the processes of fast deletion of women's biographies (Tripodi, 2023; Ramírez-Ordóñez et al., 2022). The barriers preventing women's biographies or other content from becoming good or featured articles are compounded by their lack of visibility on the front page, and vice versa. However, it's important to note that the theory attributing gender bias on Wikipedia solely to reflecting offline biases is not entirely accurate (Wagner et al., 2015, 2016). Indeed, when participation lacks diversity, the resultant content reflects this absence (Hinnosaar, 2019). Editing volunteers, guided by their own interests, tend to prioritize specific topics, thereby diminishing diversity (Worku et al., 2020), which in turn deters new editors and exacerbates the gender gap. This bias extends into reality, surpassing even that of Wikidata (Zhang & Terveen, 2021).

Diversifying contributions becomes imperative while acknowledging that Wikipedia's hierarchical collaboration model amplifies dominant voices, often overlooking non-dominant perspectives (Ukwoma et al., 2021). Effectively bridging the gender gap entails, scholars argue, a deeper comprehension of Wikipedia's knowledge production culture (Menking & Erickson, 2015), embedded within its encyclopedic and open ethos, which might conceal exclusionary practices (Ford & Wajcman, 2017). Wikipedia's gender bias is not solely shaped by editors but also by infrastructural logics (Ford & Wajcman, 2017). The objective extends beyond bias reduction to transforming Wikipedia into a more resilient, trustworthy, and transparent platform for

knowledge generation (Menking & Rosenberg, 2021). Embracing this comprehensive approach is crucial for fostering enduring change.

Our research is limited by its reliance on Wikidata for front-page content analysis. The ongoing debate on Wikidata, particularly around the interchangeable use of 'sex' and 'gender' in property (P21), underscores the need for clear differentiation. We recommend that Wikidata distinguishes properties related to biological sex from those associated with gender identity for more accurate analysis.

Looking ahead, our future research interest lies in studying more linguistic editions of Wikipedia and exploring the evolving role of bots and artificial intelligence in main page configurations. While their current influence is limited, monitoring their potential to democratize the process is crucial, given the opaque nature of algorithmic systems and the values embedded within them (Geiger, 2017). In future research, we aim to combine the gatekeeping communication theory (Barzilai-Nahon, 2009) with other communication approaches such as the agenda-setting theory (McCombs & Shaw, 1972) in relation to Wikipedia. We plan to employ qualitative and participatory methodologies to investigate the role of gatekeepers in organising main page content and the Wikipedia main page potential influence and effects on the media landscape and the public reception. This insider view will inform measures to address the gender gap and other implicit biases on Wikipedia, and the agenda-setting approach will let us compare Wikipedia's main page content and its potential effects with those offered by other traditional and new media.

Author contributions: **NF-F:** methodology, investigation, resources, writing – original draft, writing – review & editing, visualization, conceptualization, software, validation, formal analysis, data curation, supervision, project administration, funding acquisition; **MC:** methodology, investigation, resources, writing – original draft, writing – review & editing, visualization, conceptualization, software, validation, formal analysis, data curation; **LFA:** methodology, investigation, resources, writing – original draft, writing – review & editing, visualization, formal analysis. All authors approved the final version of the article.

Funding: This article was supported by the Spanish Ministerio de Innovación, Ciencia y Universidades (MCIN) and the Agencia Estatal de Investigación with the project Women and Wikipedia [PID2020116936RA-I00]. Additionally, it was funded by Wikimedia Foundation with the project Cover Women [G-RS-2402-15223]. The Juan de la Cierva Postdoctoral contract is part of the grant FJC2021-046861-1, funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR.

Acknowledgments: The authors would like to thank Andrés Bejarano for his dedicated work on this research during his master's thesis in digital humanities at the University of Barcelona and Wikipedian volunteers across Catalan, English, and Spanish editions, including those from Wikidones, Wikimujeres, WikiEsfera, and Mujeres de Portada. The authors would also like to thank Pablo Aragon, research scientist at Wikimedia Foundation.

Ethics declaration: The research project Women and Wikipedia was approved by the Ethics Committee of the Universitat Oberta de Catalunya on 20 January 2021 and then the project was transferred to the University of Barcelona and this ethical certification was considered to be valid.

Declaration of interest: The authors declared no competing interest.

Data availability: The dataset generated during the current study is not publicly available yet due to restrictions from our university repository, which only makes the dataset publicly available after the results are published in an academic article. However, the dataset is available from the corresponding author on reasonable request and will be available after publication in the CORA repository from the university of Barcelona at the following link: <https://doi.org/10.34810/data1427>. Preliminary draft available at: <https://dataverse.csuc.cat/dataset.xhtml?persistentId=doi%3A10.34810%2Fdata1427&version=DRAFT>.

REFERENCES

- Anderson, M., Hitlin, P., & Atkinson, M. (2016). Wikipedia at 15: Millions of readers in scores of languages. *Pew Research Center*. <https://www.pewresearch.org/short-reads/2016/01/14/wikipedia-at-15/>
- Anthias, F. (2013). Intersectional what? Social divisions, intersectionality and levels of analysis. *Ethnicities*, 13(1), 3-19. <https://doi.org/10.1177/1468796812463547>
- Antin, J., Yee, R., Cheshire, C., & Nov, O. (2011). Gender differences in Wikipedia editing. In F. Ortega, & A. Forte (Eds.), *Proceedings of the 7th International Symposium on Wikis and Open Collaboration* (pp. 11-14). ACM. <https://doi.org/10.1145/2038558.2038561>
- Barzilai-Nahon, K. (2009). Gatekeeping: A critical review. *Annual Review of Information Science and Technology*, 43(1), 1-79. <https://doi.org/10.1002/aris.2009.1440430117>

- Bear, J. B., & Collier, B. (2016). Where are the women in Wikipedia? Understanding the different psychological experiences of men and women in Wikipedia. *Sex Roles*, 74(5), 254-265. <https://doi.org/10.1007/s11199-015-0573-y>
- Berndt Rasmussen, K. (2020). Implicit bias and discrimination. *Theoria*, 86(6), 727-748. <https://doi.org/10.1111/theo.12227>
- Carroll, T. W. (2017). Intersectionality and identity politics: Cross-identity coalitions for progressive social change. *Signs: Journal of Women in Culture and Society*, 42(3), 600-607. <https://doi.org/10.1086/689625>
- Centelles, M., & Ferran-Ferrer, N. (2024). Assessing knowledge organization systems from a gender perspective: Wikipedia taxonomy and Wikidata ontologies. *Journal of Documentation*, 80(7), 124-147. <https://doi.org/10.1108/JD-11-2023-0230>
- Collins, P. H., & Bilge, S. (2020). *Intersectionality*. John Wiley & Sons.
- Eckert, S., & Steiner, L. (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>
- Ferran-Ferrer, N., Castellanos-Pineda, P., Minguillón, J., & Meneses, J. (2021). The gender gap on the Spanish Wikipedia: Listening to the voices of women editors. *El Profesional de La Información*, 30(5), Article e300516. <https://doi.org/10.3145/epi.2021.sep.16>
- Ferran-Ferrer, N., Centelles, M., Macià, Y., Boté-Vericad, J.-J., & Minguillón, J. (2024). *Dones de categoria: Anàlisi del biaix de gènere a les categories de Viquipèdia* [Women in category: Analysis of gender bias in Wikipedia categories]. Xarxa Vives d'Universitats; Centre de Recerca en Informació, Comunicació i Cultura (CRICC). <https://www.vives.org/book/dones-de-categoria-analisi-del-biaix-de-genere-a-les-categories-de-viquipedia/>
- Ferran-Ferrer, N., Fernández, L., & Centelles, M. (2024). Behind the front page: A comparative gender gap study on Wikipedia's main page through gatekeeping and agenda-setting theories. *10th European Communication Conference, ECREA 2024, Communication and Social (Dis)order*. Faculty of Social Sciences, University of Ljubljana, Slovenia, September 24-26.
- Ford, H., & Wajcman, J. (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>
- Geiger, R. S. (2017). Beyond opening up the black box: Investigating the role of algorithmic systems in Wikipedia organizational culture. *Big Data & Society*, 4(2), 205395171773073. <https://doi.org/10.1177/2053951717730735>
- 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>
- Hargittai, E., & Shaw, A. (2015). Mind the skills gap: The role of Internet know-how and gender in differentiated contributions to Wikipedia. *Information, Communication & Society*, 18(4), 424-442. <https://doi.org/10.1080/1369118X.2014.957711>
- Harvey, A. (2020). *Feminist media studies*. Polity Press.
- Hinnosaar, M. (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>
- Hood, N., & Littlejohn, A. (2018). Hacking history: Redressing gender inequities on Wikipedia through an Editathon. *The International Review of Research in Open and Distributed Learning*, 19(5), 203-217. <https://doi.org/10.19173/irrodl.v19i5.3549>
- Ju, B., & Stewart, B. (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>
- Kalleberg, A. L. (2018). Job insecurity and well-being in rich democracies. *The Economic and Social Review*, 49(3), 241-258.
- Keegan, B., & Gergle, D. (2010). Egalitarians at the gate: One-sided gatekeeping practices in social media. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (pp. 131-134). ACM. <https://doi.org/10.1145/1718918.1718943>
- Kleeman, J. (2015). The Wikipedia wars: Does it matter if our biggest source of knowledge is written by men. *New Statesman*. <https://www.newstatesman.com/long-reads/2015/05/wikipedia-has-colossal-problem-women-dont-edit-it>
- Knott, K., & Poole, E. (2016). *Media portrayals of religion and the secular sacred: Representation and change*. Routledge. <https://doi.org/10.4324/9781315594514>

- Lam, S. K., Uduwage, A., Dong, Z., Sen, S., Musicant, D. R., Terveen, L., & Riedl, J. (2011). WP: clubhouse?: An exploration of Wikipedia's gender imbalance. In F. Ortega, & A. Forte (Eds.), *Proceedings of the 7th International Symposium on Wikis and Open Collaboration* (pp. 10-11). ACM. <https://doi.org/10.1145/2038558.2038560>
- Leonard, V., & Bond, S. E. (2019). Advancing feminism online tools, visibility, and women in classics. *Studies in Late Antiquity*, 3(1), 4-16. <https://doi.org/10.1525/sla.2019.3.1.4>
- Macharia, S. (2020). Global media monitoring project (GMMP). In *The international encyclopedia of gender, media, and communication* (pp. 1-6). John Wiley & Sons. <https://doi.org/10.1002/9781119429128.iegmc074>
- Matei, S. A., & Dobrescu, C. (2011). Wikipedia's "neutral point of view": Settling conflict through ambiguity. *The Information Society*, 27(1), 40-51. <https://doi.org/10.1080/01972243.2011.534368>
- Matud, M. P., Espinosa, I., & Rodríguez Wangüemert, C. (2021). Women and men portrayal on television news: A study of Spanish television newscast. *Feminist Media Studies*, 21(2), 298-314. <https://doi.org/10.1080/14680777.2019.1681489>
- McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, 36(2), 176-187.
- Menking, A., & Erickson, I. (2015). The heart work of Wikipedia: Gendered, emotional labor in the world's largest online encyclopedia. In B. Begole, J. Kim, K. Inkpen, & W. Woo (Eds.), *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 207-210). ACM. <https://doi.org/10.1145/2702123.2702514>
- Menking, A., & Rosenberg, J. (2021). WP:NOT, WP:NPOV, and other stories Wikipedia tells us: A feminist critique of Wikipedia's epistemology. *Science Technology and Human Values*, 46(3), 455-479. <https://doi.org/10.1177/0162243920924783>
- Minguillón, J., Meneses, J., Aibar, E., Ferran-Ferrer, N., & Fàbregues, S. (2021). Exploring the gender gap in the Spanish Wikipedia: Differences in engagement and editing practices. *PLOS ONE*, 16(2), Article e0246702. <https://doi.org/10.1371/journal.pone.0246702>
- Miquel-Ribé, M., & Laniado, D. (2021). The Wikipedia diversity observatory: Helping communities to bridge content gaps through interactive interfaces. *Journal of Internet Services and Applications*, 12(1), Article 25. <https://doi.org/10.1186/s13174-021-00141-y>
- Pageviews Analysis. (2023). *Pageviews Analysis*. <https://pageviews.wmcloud.org/pageviews/?project=en.wikipedia.org&platform=all-access&agent=user&redirects=0&range=latest-20&pages=Cat|Dog>
- Park, D. E., & Bridges, L. M. (2022). Meet students where they are: Centering Wikipedia in the classroom. *Communications in Information Literacy*, 16(1), 4-23. <https://doi.org/10.15760/comminfolit.2022.16.1.2>
- Pennycook, A., & Otsuji, E. (2015). *Metrolingualism: Language in the city*. Routledge. <https://doi.org/10.4324/9781315724225>
- Ramírez-Ordóñez, D., & Ferran-Ferrer, N. (2023). Endurance against oblivion: The case of the Articles for Deletion with gender perspective in Wikipedia. *Proceedings of the Wiki Workshop* (10th Edition). https://wikiworkshop.org/2023/papers/WikiWorkshop2023_paper_34.pdf
- Ramírez-Ordóñez, D., Ferran-Ferrer, N., & Meneses, J. (2022). Wikipedia and gender: The deleted, the marked, and the unpolluted biographies. *Proceedings of the Wiki Workshop* (9th Edition). https://wikiworkshop.org/2022/papers/WikiWorkshop2022_paper_27.pdf
- Redi, M., Gerlach, M., Johnson, I., Morgan, J., & Zia, L. (2021). *A taxonomy of knowledge gaps for Wikimedia projects*. arXiv. <http://arxiv.org/abs/2008.12314>
- Reinoso, A. J., González-Barahona, J.M., Muñoz-Mansilla, R., & Herraiz Tabernero, I. (2011). A quantitative examination of the impact of featured articles in Wikipedia. In *Proceedings of the ICSoft 2011-International Conference on Software and Data Technologies*. <https://oa.upm.es/7243/1/86cameraReady.pdf>
- Religious Information by Country. (2024). Religious information by country. *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Religious_information_by_country&oldid=1220768353
- Sefidari Huici, M. (2022). Equidad de conocimiento y sesgos: Un análisis cuantitativo del contenido destacado en la portada de Wikipedia [Knowledge equity and biases: A quantitative analysis of featured content on the Wikipedia homepage]. *IC Revista Científica de Información y Comunicación*, 19, 141-163. <https://doi.org/10.12795/IC.2022.119.07>

- Spack, R. (1997). The rhetorical construction of multilingual students. *TESOL Quarterly*, 31(4), Article 765. <https://doi.org/10.2307/3587759>
- Star, S. L. (1999). The ethnography of infrastructure. *American Behavioral Scientist*, 43(3), 377-391. <https://doi.org/10.1177/00027649921955326>
- Tkacz, N. (2014). *Wikipedia and the politics of openness*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226192444.001.0001>
- Toupin, S. (2020). Feminist peer production. In M. O'Neil, C. Pentzold, & S. Toupin (Eds.), *The handbook of peer production* (pp. 311-321). Wiley. <https://doi.org/10.1002/9781119537151.ch23>
- Tripodi, F. (2023). Ms. categorized: Gender, notability, and inequality on Wikipedia. *New Media & Society*, 25(7), 1687-1707. <https://doi.org/10.1177/14614448211023772>
- Ukwoma, S. C., Osadebe, N. E., Okafor, V. N., & Ezeani, C. N. (2021). Unveiling the veiled: Wikipedia collaborating with academic libraries in Africa in creating visibility for African women through art plus feminism Wikipedia edit-a-thon. *Digital Library Perspectives*, 37(4), 449-462. <https://doi.org/10.1108/DLP-08-2020-0079>
- van Dijk, J. (2012). *The network society*. SAGE.
- Wagner, C., Garcia, D., Jadidi, M., & Strohmaier, M. (2015). It's a man's Wikipedia?: Assessing gender inequality in an online encyclopedia. In *Proceedings of the 9th International AAAI Conference on Web and Social Media* (pp. 454-463). PKP. <https://doi.org/10.1609/icwsm.v9i1.14628>
- Wagner, C., Graells-Garrido, E., Garcia, D., & Menczer, F. (2016). Women through the glass ceiling: Gender asymmetries in Wikipedia. *EPJ Data Science*, 5(1). <https://doi.org/10.1140/epjds/s13688-016-0066-4>
- Wikimedia. (2023a). *List of Wikipedias by edits per article*. https://meta.wikimedia.org/wiki/List_of_Wikipedias_by_edits_per_article
- Wikimedia. (2023b). *Wikimedia statistics: All wikis*. <https://stats.wikimedia.org/#/all-projects>
- Wikipedia. (2023a). *List of most-visited websites*. https://en.wikipedia.org/w/index.php?title=List_of_most-visited_websites&oldid=1189478387
- Wikipedia. (2023b). *Statistics*. <https://en.wikipedia.org/wiki/Special:Statistics?action=raw>
- Wikipedia. (2024). *List of Wikipedias*. https://en.wikipedia.org/w/index.php?title=List_of_Wikipedias&oldid=1229876652
- Wikipedia. (2025). *Wikipedia: Statistics*. <https://en.wikipedia.org/w/index.php?title=Wikipedia:Statistics&oldid=1320603209>
- Worku, Z., Bipat, T., McDonald, D. W., & Zachry, M. (2020). Exploring systematic bias through article deletions on Wikipedia from a behavioral perspective. In G. Robles, K.-J. Stol, & X. Wang (Eds.), *Proceedings of the 16th International Symposium on Open Collaboration* (pp. 1-22). ACM. <https://doi.org/10.1145/3412569.3412573>
- Zhang, C. C., & Terveen, L. (2021). Quantifying the gap: A case study of Wikidata gender disparities. In G. Robles, K.-J. Stol, & X. Wang (Eds.), *Proceedings of the 16th International Symposium on Open Collaboration* (pp. 1-12). ACM. <https://doi.org/10.1145/3479986.3479992>

