An Analysis of Content Gaps versus User Needs in the Wikidata Knowledge Graph

David Abián$^{(\boxtimes)[0000-0002-4405-3652]}$,
Albert Meroño-Peñuela$^{[0000-0003-4646-5842]}$, and
Elena Simperl$^{[0000-0003-1722-947X]}$

King's College London, London, United Kingdom
`{david.abian,albert.merono,elena.simperl}@kcl.ac.uk`

Abstract. Content gaps in knowledge graphs impact downstream applications. Semantic Web researchers have studied them mainly in relation to data quality or ontology evaluation, for instance by proposing frameworks to capture various quality dimensions or methods to assess these dimensions, such as completeness, accuracy, or consistency. Less work has been done in framing these gaps in the context of user needs. This limits our ability to design processes and tools to help knowledge engineers tackle such gaps effectively. We propose a framework that: (i) captures core types of content gaps, informed by a literature review on peer-production systems; and, in the areas with such gaps, (ii) quantitatively compares the imbalances in the work on the knowledge graph with the imbalances in users' information needs to clarify the origin of the gaps. We operationalize the framework with gender, recency, geographic, and socio-economic gaps, and apply it to Wikidata by comparing edit metrics with Wikipedia pageviews between 2018 and 2021. We did not find gender or recency gaps endogenous to Wikidata's production. Only exceptionally, Wikidata editors work on under-represented entities (e.g. people from countries with lower Human Development Index) less than they should according to the volume of requests. We hope this study will provide a foundation for knowledge engineers to explore the causes of content gaps and address them if and when needed.

Keywords: knowledge graphs · content gaps · Wikidata · data quality.

1 Introduction

Content gaps in data sources are missed opportunities to meet information needs and achieve greater impact. They can also create or reinforce biases, for instance in artificial intelligence systems that rely heavily on data [49, 43, 65, 27]. Ontologies [36] and knowledge graphs (KGs) such as Wikidata [54] show content imbalances that researchers have documented as harmful.

The Semantic Web community has studied content gaps in relation to data quality or ontology evaluation. Researchers have proposed ways to capture and assess various quality dimensions such as completeness, accuracy, or consistency

[76, 21, 52]. There are also similar approaches for ontologies [69, 42]. More often than not, these approaches deliver quantified accounts of quality, but struggle to put them in context, e.g. what does it mean that the completeness of a dataset has reached 80%? Is it worth aiming for more or is this a good enough result? Studying the evolution of quality indicators can help by putting numbers in perspective but is not enough to determine whether a dataset or ontology is *fit for use*, the litmus test to which most literature in this space refers [76].

We propose a framework that (i) captures core types of content gaps, informed by a literature review on peer-production systems; and, in the areas with such gaps, (ii) quantitatively compares the imbalances in the work on the KG with the imbalances in users' information needs. This is valuable to clarify whether such gaps are endogenous to the KG and, therefore, whether they represent a fitness-for-use problem. We operationalize the framework with gender, recency, geographic, and socio-economic gaps, and apply it to Wikidata by comparing contribution metrics with Wikipedia pageviews between 2018 and 2021. We choose these gaps as they are among those that have attracted the most interest in the literature and, at the same time, are relevant to Wikidata [56]. We collect a representative random sample of each set of instances under study in the KG, enrich the samples with the contribution metrics and Wikipedia pageviews, and analyse the data to answer the following research questions:

> **RQ1:** *Does the contribution to Wikidata show a gender gap that is misaligned with information needs?*
>
> **RQ2:** *Does the contribution to Wikidata show a recency gap that is misaligned with information needs?*
>
> **RQ3:** *Does the contribution to Wikidata show a geographic and socio-economic gap that is misaligned with information needs?*

We study the statistical significance and effect sizes of differences and correlations based on gender, years of birth and death, population of settlements, and several human development indicators with three different metrics of contribution and datasets about three different classes of entities (people, settlements, and countries). We find no evidence of gender or recency gaps in the contribution to Wikidata to a greater extent than in users' information needs or Wikipedia, which suggests that these gaps are exogenous to Wikidata's production processes. Only exceptionally, Wikidata editors work on under-represented entities (e.g. people from countries with lower Human Development Index) less than they should according to the volume of requests.

We hope this study will provide a foundation for knowledge engineers to explore the causes of content gaps and address them more effectively. In applying our framework to Wikidata, our findings also contribute towards Wikimedia's strategic goals to address content gaps [56] by pointing to potential biases in contribution patterns that impact the KG's quality.

2   Background and Related Work

Our work sits at the intersection between KGs and online peer-production systems. Consequently, we first explore related work on the quality of Wikidata, undertaken mainly by the Semantic Web community, and then we give an overview of literature from related fields (CSCW, social computing, computational social sciences) that have studied the relationship between digital artefacts (e.g. Wikipedia, OpenStreetMap) and their socio-technical ecosystems.

2.1   Wikidata Quality

Data quality is multidimensional and often conceived as the *fitness for use* for a task or application [76]. Researchers have proposed ways to capture and assess various quality dimensions, but have worked much less on framing these dimensions considering the actual use of the data. [52] surveyed 28 publications until 2018 about Wikidata quality, noting a prevalence of methods and tools for data completeness or accuracy. According to their data quality dimensions, our study addresses *completeness* and *timeliness*.

Wikidata completeness. There are many ways to get a sense of the completeness of a KG. [2, 8] compared similar Wikidata items to spot those missing information. [23] generated completeness assertions using rules, while [18, 53, 17] annotated Wikidata with completeness metadata and reasoned about the completeness of query results. [48] interpreted metrics on the coverage of scholarly literature on Wikidata, and [58] compared artefacts in Wikidata with other KGs. [40] estimated the completeness of a class in relation to a schema or ontology. [71] compared attribute completeness between different sets of items defined by other attributes (e.g. how complete the attribute "date of birth" is comparing between male computer scientists and female physicists), while [20] used visualizations and dimensional reduction to identify and explore subsets of items missing the same attributes. In 2019, Wikidata implemented Shape Expressions (ShEx)[12] [12, 61], which allows checking the completeness of the data against a schema. Despite these developments, incompleteness remains an issue today.

Wikidata timeliness. As [52] noted, Wikidata allows more frequent updates than other KGs because it is peer-produced. However, the literature on timeliness on Wikidata is limited. [21] studied three timeliness criteria, which Wikidata satisfied: timeliness frequency of the graph, specification of the modification date of statements, and specification of the validity period.

Our work complements these and compares content changes with users' information needs to understand the topics people ask for that the KG may not cover well enough. Furthermore, our framework allows exploring whether content gaps are endogenous to the socio-technical environment where the KG is produced or driven by externalities such as the requirements of the consumers of the graph.

---

[1] `https://www.mediawiki.org/wiki/Extension:EntitySchema`
[2] `https://www.wikidata.org/wiki/Wikidata:Schemas`

2.2   Content Gaps

Gaps are common in online peer production due to a multitude of reasons, including the motivations and interests of the participants, the ways tasks are allocated to participants and the degree to which they coordinate, as well as the technologies they use to contribute [19, 13, 28, 56]. Researchers have documented several types of gaps in online peer production and explored how they come about. For example, [56] compiled a taxonomy that distinguishes between gaps based on characteristics of the contributing community, the users of the peer-produced artefact, and the artefact itself. Our study only addresses the latter, which we refer to as *content gaps*. The Wikimedia Foundation and the communities of Wikipedia and Wikidata have expressed concern about such gaps and have agreed to address them as a strategic priority [56, 44].

In the following, we elaborate on the three types of gaps we address in the current implementation of our framework. We choose these as they are among those that have attracted the most interest in the literature and, at the same time, are relevant to Wikidata [56]. We provide an overview of prior studies of these gaps in the context of Wikipedia and Wikidata and, to a lesser extent, other popular systems.

RQ1: Gender Gap. The fact that Wikipedia and Wikidata cover more and better males than females is well documented [56]. However, its causes and the ways to mitigate it are still subject to ongoing discussions. [24] found significant gender differences in metadata, language, and network structure that partly attributed to the editors. In 2016, [70] concluded that Wikipedia articles about females were slightly more notable than their male counterparts. Furthermore, [75] found a systematic over-representation of men when comparing the labour market with the proportions of males with Wikipedia articles, redirects, images, and mentions. [78] suggested that the quality of Wikidata items on females was similar to the quality of those on males, and that Wikidata's proportions of females within each occupation were aligned with the professional societies' notability assessments. [37] found the creation of more articles on females (65.6%) than on males on the English Wikipedia, and [38] noted that the ratios of articles on females were rising exponentially. Finally, [74] commented that Wikipedia editors had over-corrected the content based on gender to the point of biases against males.

RQ2: Recency Gap. There is more content on Wikipedia and Wikidata on more recent events [60, 56, 34, 35, 11]. This *recentism*[3] significantly grew on Wikipedia throughout the 2000s [34]. Breaking news, such as incidents, crises, and deaths, quickly lead to a surge in edits [35, 11, 34]. Some researchers link this to users' information needs [56] rather than other factors endogenous to the Wikipedia ecosystem, based on engagement data that shows that e.g. references about recent events are more frequently hovered/clicked [51], or dates of birth in Wikidata and the historical human population are significantly correlated [38].

RQ3: Geographic and Socio-economic Gap. On a wide range of websites, including Twitter, Flickr, Foursquare, Wikipedia, and OpenStreetMap,

---

[3] https://en.wikipedia.org/wiki/Wikipedia:Recentism

people tend to document urban and artificial entities earlier, better, and more often than rural, semi-natural, and natural entities, which are also more likely to be generated by bots rather than people interested in local topics [29, 31, 7, 6, 56]. The literature also notes Eurocentric, US-centric, pro-Western, and pro-Global North gaps on Wikipedia and OpenStreetMap [26, 30, 60, 64, 10]. These gaps are highly correlated with socio-economic factors such as wealth, literacy, and human development in general [38, 81, 82, 30, 64], so the geographic gap and socio-economic gap partially overlap. These global gaps are reported to be greater than the inequality in the global distribution of wealth [10], although Wikipedia editors have reduced them over time: Europe had 20 times more geo-tagged Wikipedia articles than Africa in 2010, but four times more than Africa in 2017 [25]. It is unclear in which cases these differences in content are linked to varying users' information needs.

3   Methods

We define a framework for quantitatively comparing the imbalances in the work on a KG with the imbalances in users' information needs. This allows us to clarify whether or not the content gaps studied are particular to the KG. We operationalize this framework for Wikidata and apply it to understand three families of potential gaps.

3.1   Framework of Analysis

Design Considerations. We want to: (r1) measure quantitatively, to understand the importance of each gap; (r2) measure imbalances introduced or maintained in a given period, to understand their evolution and be able to draw conclusions specific to the period of interest; (r3) measure imbalances based on any type of attribute, whether categorical (e.g. gender) or numerical (e.g. population), so that the framework applies to many domains; (r4) measure information needs in several representative languages, to avoid bias.

   Dimensions and Metrics. We consider two families of metrics: proxies for the contribution to the KG, and proxies for the information needs, against which the former are compared. Metrics of information needs can be very diverse and context-specific. In contrast, from a comprehensive review of the literature on online peer production, we learned that contribution is mainly characterized and measured in four categories, which we will refer to as $CAPT$: (C) contributions (as a countable noun; e.g. edits in Wikipedia, edits and change-sets in OpenStreetMap); (A) artefacts (e.g. Wikipedia articles, Wikidata items); (P) participants (e.g. Wikipedians or editors in Wikipedia, mappers in Open-StreetMap); and (T) time. A *contribution* is a documented change undertaken by a *participant* by applying a create, update or delete action to an *artefact* at a certain *time*. A participant or a unit of time can have any number of associated contributions, including zero, whereas an artefact should have one or more associated contributions. Contribution metrics can be calculated with filtering and

aggregation operations on $CAPT$ entities. The most common aggregation operation is counting $CAPT$ entities of a certain type; the most common filtering operation, selecting a single value or $CAPT$ entity. The number of contributions is the most widely used contribution metric in the literature on online peer production [3, 63, 46, 33, 14, 9]. The terminology may vary, including names such as "edit count" [57, 47, 39, 32, 4], "number of edits" [79, 80, 73, 62, 50, 16], "quantity of edits" [63], "number of revisions" [72, 67, 59] and "number of user activities" [22], among others.

Measurement Criteria. We want to filter artefacts based on the attributes of interest (e.g. gender) and obtain metrics per artefact, so we can count three other types of $CAPT$ entities as the simplest contribution metrics: contributions, participants, and units of time. In several peer-production systems, most artefacts and participants have hardly any contributions, and most contributions are associated with a few artefacts and participants [41, 72, 62, 5, 46]. Participants can quickly add up large numbers of contributions by making many minor changes in a short time. This makes contributions a noisy metric because, in these cases, many contributions do not mean more value produced or more effort invested. To complement the number of contributions, it is possible to count units of time (e.g. hours, days, months) with contributions or consider the number of different participants with contributions. We also consider metrics linking contribution and information needs: the return on investment (ROI) ratios. We can calculate one of these ratios for each possible combination of contribution (c) metric and information need (n) metric by applying $n/(c+1)$. We assume that the potential content gaps against those artefacts with higher ROI ratios are more likely to be misaligned with information needs.

3.2   Operationalization for Wikidata

As per (r4), we decide to use the *pageviews from users* (not spiders or bots) of the Wikipedias corresponding to the top ten most spoken languages in the world in 2021 according to ethnologue.com[4]: English, Mandarin Chinese, Hindi, Spanish, French, Standard Arabic, Bengali, Russian, Portuguese, and Urdu. Pageviews are considered the "most important content consumption metric"[5] on Wikipedia, and studies and tools use it to identify "concepts with significant increase of the interest from the public" [15]. As each Wikidata item about an entity is linked to the titles of the Wikipedia articles about the same entity, it is possible to automatically enrich a dataset that contains identifiers of Wikidata items with their corresponding Wikipedia pageviews.

We choose the *number of contributions*, the *number of days with contributions* (operationalizing the number of units of time), and the *number of human participants* (operationalizing the number of participants). We consider only human participants for the latter metric because, according to [47], bots make around 85% of contributions to Wikidata items, but we do not discern noise generated by

---

[4] `https://www.ethnologue.com/guides/ethnologue200`
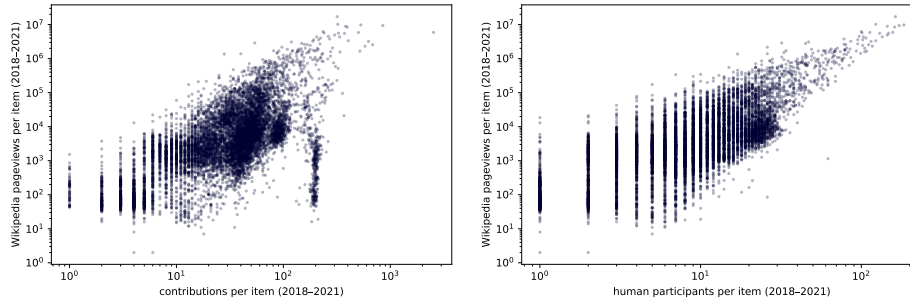[5] `https://meta.wikimedia.org/wiki/Research:Page_view`

Fig. 1. Random sample of Wikidata items in two charts as a function of Wikipedia pageviews and two different contribution metrics: on the left, number of (manual and automatic) contributions, revealing clusters of similar entities that received the same automatic treatment (e.g. vertical line around 200 contributions); on the right, number of human editors, which does not provide this insight but better quantifies the actual effort invested and therefore better correlates with pageviews.

bots with the other two metrics. During our exploratory analysis (Figure 1) we confirm that the combination of contribution metrics chosen is more informative and useful to operationalize contribution than any of the metrics individually.

For Wikidata we should measure the contribution made over at least a few years to avoid an excess of zeros [62]. At the same time, we seek to draw conclusions about the KG's current or most recent socio-technical context. Therefore, we set our study period to be the four years prior to the year of analysis: from 2018 to 2021.

Despite the existence of tools that allow querying Wikidata's edit history [66, 32], we use simple random samples (see section 4) instead of full sets of instances because: (a) the samples are sufficient to obtain conclusive results from the statistical analysis; and (b) we enrich the data with contribution metrics, but also with Wikipedia pageviews, so we have to combine two metadata sources,[6] the query of which would hardly scale to the full sets of instances.[7] Due to (r1), we choose hypothesis testing to confirm or reject a relationship between contribution and information needs with statistical significance for each potential gap studied. We quantify these relationships with the effect sizes. The unit of analysis is not the artefact, but the combination of the artefact and the attribute under study, as the latter may be multivalued. As per (r3), we choose two types of tests depending on the type of attribute under study: correlations, for numerical attributes (e.g. year of birth); and differences between groups, for categorical attributes (e.g. gender). To compare differences in numerical values between groups we use Mann-Whitney U tests; to check correlations between two numerical variables, Spearman's rank-order correlations. We choose these tests because they are non-parametric, are based on ranks, and do not assume

---

[6] The APIs `https://www.wikidata.org/w/api.php` and `https://wikimedia.org/api/rest_v1/metrics/pageviews/`, respectively. See supplemental material.

[7] We will extend the analysis to the full sets in future work.

normal distributions in the data. The literature has confirmed that contribution on peer-production websites, including Wikidata, does not follow normal distributions; instead, these distributions tend to be highly skewed, concentrated on a few participants and artefacts [41, 72, 62, 5, 46]. Our samples also show this property. In line with [55], we consider that a Spearman's rank-order correlation has a *negligible* strength and, therefore, we conclude that there is no correlation between variables, when $|\rho| < 0.1$. Similarly, we consider that a Mann-Whitney U test shows a *negligible* effect size, and conclude that there is no significant difference between groups, when $|r| < 0.1$. We reject the null hypothesis when $p \leq 0.01$.

3.3   Hypotheses about Wikidata

Note that we consider three contribution metrics, so we perform groups of three tests for comparing contribution and groups of three tests for comparing ROI ratios. Table 1 shows all the hypotheses tested, the tests used for each of them, and their correspondence to our research questions. For **H2–H4** and **H8–H11**, we consider both the set of all items and only those with links to articles in the Wikipedias studied, and both annually (2018, 2019, 2020, and 2021) and over the entire study period (2018–2021). For **H5** and **H6**, we consider *events* as births and as deaths. For **H7**, we consider both the set of all items and only those with links to articles in the Wikipedias studied, both annually (2018, 2019, 2020, and 2021) and over the entire study period (2018–2021), and considering *events* as births and as deaths.

4   Data

We generate and analyse three tabular datasets from Wikidata[8]: (a) 50,000 random items on people with, where defined, sex or gender, year of birth, year of death, occupation, and country of citizenship; (b) 50,000 random items on human settlements with population and, where defined, coordinates, continent, and country; and (c) all 374 items defined as instances of sovereign states.

For each item in each dataset we retrieve and include all the metrics described in Section 3.2, both from 2018 to 2021 and by year. We also include the pageviews broken down by Wikipedia and the corresponding title of the article, if any. The pageviews are quantified as zero for each Wikipedia without an article associated with a given item. We enrich all the datasets with the Human Development Index (HDI) per country and its base indicators according to [45]: Life Expectancy at birth (LE; in years); Expected Years of Schooling (EYS); Mean Years of Schooling (MYS); and Gross National Income per capita (GNIpc; in PPP \$).

We analyse the gender and recency gaps based on the dataset about items on people, and the geographic and socio-economic gap based on the three datasets.

---

[8] The Python modules and SPARQL queries used to generate the datasets are available on https://github.com/davidabian/wikidata-gaps-vs-needs.

Table 1. Research questions, tests used, and hypotheses tested on Wikidata.

RQ1: GENDER GAP (TWO-SIDED FISHER'S EXACT TEST)	
H1	The proportion of items with links to articles in the Wikipedias studied to all items is significantly different between the set of items on males and the set of items on females.
RQ1: GENDER GAP (MANN–WHITNEY U TESTS)	
H2	($c_{male} > c_{female}$): The contribution metrics per item tend to be higher for items on males than for items on females.
H3	($pv_{male} > pv_{female}$): The pageviews per item tend to be higher for items on males than for items on females.
H4	($roi_{female} > roi_{male}$): The ROI ratios per item tend to be higher for items on females than for items on males.
RQ2: RECENCY GAP (MANN–WHITNEY U TESTS)	
H5	The years of the events in items with links to articles in the Wikipedias studied tend to be different than those in items without links to articles in the Wikipedias studied.
RQ2: RECENCY GAP (SPEARMAN'S ρ RANK-ORDER CORRELATIONS)	
H6	The last 500 years interpreted as numbers (1522–2021) are correlated with the proportions of Wikidata items that have the events in those years and links to articles in the Wikipedias studied relative to the Wikidata items that have the events in those years but no links to articles in the Wikipedias studied.
H7	The contribution metrics, pageviews, and ROI ratios per item are correlated with the years of the events.
RQ3: GEOGR. AND SOCIO-ECONOMIC GAP (SPEARMAN'S ρ RANK-ORDER CORRELATIONS)	
H8	The contribution metrics, pageviews, and ROI ratios per item on a settlement are correlated with the average of its population values.
H9	The contribution metrics, pageviews, and ROI ratios per item on a person are correlated with the Human Development Index of the countries of citizenship.
H10	The contribution metrics, pageviews, and ROI ratios per item on a settlement are correlated with the Human Development Index of its country.
H11	The contribution metrics, pageviews, and ROI ratios per item on a state are correlated with its Human Development Index.

5 Results

In this section we present the results of the tests specified in Section 3.3 and Table 1, together with contextual information such as statistics and data visualizations. Many results of the tests based on metrics per item (H2–H4, H7–H11) are synthesized in Table 2 and not repeated in text.

5.1 RQ1: Gender Gap

- Around 36.4% of the items on males and 25.9% of the items on females had links to articles in the Wikipedias studied.
- ROI ratios were not higher for items on females than for items on males.
- An item on a male tended to receive more contribution than an item on a female in 2018, but this was no longer the case in 2021.

As of 30 January 2022, Wikidata had 9,608,862 items on people (instances of Q5), most of them (79.81%, 7,668,492) with some sex or gender defined. These

Table 2. Synthesis of the results of the Mann–Whitney U tests and Spearman's ρ rank-order correlations for hypotheses H2–H4 and H7–H11 based on contribution metrics (c), Wikipedia pageviews (pv), ROI ratios (roi), average of population figures in the item (pop), the Human Development Index of the country (HDI), and its core indicators Life Expectancy at birth (LE; in years), Expected Years of Schooling (EYS), Mean Years of Schooling (MYS), and Gross National Income per capita (GNIpc; in PPP $). ✓ represents the acceptance of the hypothesis shown with a non-negligible effect size ($r \geq 0.1$) according to a metric; =, a conclusive result with a negligible effect size ($r < 0.1$); ?, an inconclusive result ($p > 0.01$); +, a positive correlation ($\rho \geq 0.1$); -, a negative correlation ($\rho \leq -0.1$); and 0, no correlation ($|\rho| < 0.1$ or $p > 0.01$).

RQ1: GENDER GAP

items	tests	2018–21	2018	2019	2020	2021
all people	H2: $c_{male} > c_{female}$	✓==	✓✓✓	✓==	=??	???
	H3: $pv_{male} > pv_{female}$	✓	✓	✓	✓	=
	H4: $roi_{female} > roi_{male}$	???	???	???	???	???
people linked to Wikipedia	H2: $c_{male} > c_{female}$	???	???	???	???	???
	H3: $pv_{male} > pv_{female}$	?	?	?	?	?
	H4: $roi_{female} > roi_{male}$	===	✓✓✓	===	===	===

RQ2: RECENCY GAP (YEARS OF BIRTH AND DEATH)

items	tests	2018–21	2018	2019	2020	2021
all people	H7: corr(birth, c)	000	+00	++0	000	000
	H7: corr(birth, pv)	+	0	+	+	+
	H7: corr(birth, roi)	+++	000	+++	+++	+++
	H7: corr(death, c)	++0	+++	++0	000	000
	H7: corr(death, pv)	+	+	+	+	+
	H7: corr(death, roi)	+++	++0	+++	+++	+++
people linked to Wikipedia	H7: corr(birth, c)	000	000	000	000	000
	H7: corr(birth, pv)	+	0	+	+	+
	H7: corr(birth, roi)	+++	000	000	+++	+++
	H7: corr(death, c)	000	000	000	000	000
	H7: corr(death, pv)	0	0	0	0	0
	H7: corr(death, roi)	000	000	000	000	++0

RQ3: GEOGRAPHIC AND SOCIO-ECONOMIC GAP (POPULATION)

items	tests	2018–21	2018	2019	2020	2021
all settlements	H8: corr(pop, contr)	+++	+++	+++	+++	+++
	H8: corr(pop, need)	+	+	+	+	+
	H8: corr(pop, ROI)	+++	+++	+++	+++	+++
settlements linked to Wikipedia	H8: corr(pop, contr)	+++	+++	+++	+++	+++
	H8: corr(pop, need)	+	+	+	+	+
	H8: corr(pop, ROI)	+++	+++	+++	+++	+++

RQ3: GEOGRAPHIC AND SOCIO-ECONOMIC GAP (HDI), 2018–21

items	tests	HDI	LE	EYS	MYS	GNIpc
all people	H9: corr(contr, _)	+++	000	+++	+++	+++
	H9: corr(need, _)	0	–	0	0	0
	H9: corr(ROI, _)	000	---	000	000	000
people linked to Wikipedia	H9: corr(contr, _)	+++	+++	+++	+++	+++
	H9: corr(need, _)	0	0	0	0	0
	H9: corr(ROI, _)	000	---	000	000	000
all settlements	H10: corr(contr, _)	+++	+++	+++	+++	+++
	H10: corr(need, _)	+	+	+	+	+
	H10: corr(ROI, _)	+++	+++	+++	+++	+++
settlements linked to Wikipedia	H10: corr(contr, _)	+++	+++	+++	+++	+++
	H10: corr(need, _)	+	+	+	+	+
	H10: corr(ROI, _)	+++	+++	+++	+++	+++
all states	H11: corr(contr, _)	+++	+++	+++	+++	+++
	H11: corr(need, _)	+	+	+	+	+
	H11: corr(ROI, _)	+++	+++	+++	+++	+++
states linked to Wikipedia	H11: corr(contr, _)	+++	+++	+++	+++	+++
	H11: corr(need, _)	+	+	+	+	+
	H11: corr(ROI, _)	+++	+++	+++	+++	+++

RQ3: GEOGRAPHIC AND SOCIO-ECONOMIC GAP (HDI)

items	tests	2018–21	2018	2019	2020	2021
all people	H9: corr(HDI, contr)	+++	++0	0++	+++	000
	H9: corr(HDI, need)	0	0	0	0	0
	H9: corr(HDI, ROI)	000	000	000	000	000
people linked to Wikipedia	H9: corr(HDI, contr)	+++	++0	+++	+++	000
	H9: corr(HDI, need)	0	0	0	0	0
	H9: corr(HDI, ROI)	000	000	000	000	000
all settlements	H10: corr(HDI, contr)	+++	+++	+++	+++	+++
	H10: corr(HDI, need)	+	+	+	+	+
	H10: corr(HDI, ROI)	+++	+++	+++	+++	+++
settlements linked to Wikipedia	H10: corr(HDI, contr)	+++	+++	+++	+++	+++
	H10: corr(HDI, need)	+	+	+	+	+
	H10: corr(HDI, ROI)	+++	+++	+++	+++	+++
all states	H11: corr(HDI, contr)	+++	+++	0++	+++	+++
	H11: corr(HDI, need)	+	+	+	+	+
	H11: corr(HDI, ROI)	+++	+++	+++	+++	+++
states linked to Wikipedia	H11: corr(HDI, contr)	+++	+++	0++	+++	+++
	H11: corr(HDI, need)	+	+	+	+	+
	H11: corr(HDI, ROI)	+++	+++	+++	+++	+++

included 5,819,674 items on males and 1,847,850 items on females, that is, 3.15 times more items on males than on females. Males and females accounted for 99.99% of all items with some sex or gender defined, with the remaining 0.01% of the values, in order of number of occurrences, being transgender female, non-binary, transgender male, eunuch, intersex, genderfluid, genderqueer, agender, transgender person, cisgender female, and many others.

H1. Out of 30,231 items on males and 9,540 items on females in the sample, 10,993 (36.36%) and 2,467 (25.86%) had links to articles in the Wikipedias studied, respectively. As hypothesized, these differences were statistically significant (Fisher's exact test, two-sided $p < .001$).

H2–H4, 2018–2021, all the items. Small effect sizes ($r = 0.1$) for the numbers of human participants and pageviews.

H2–H4, 2018–2021, items with links to the Wikipedias studied.
None of the hypotheses were accepted.

H2–H4, metrics per year, all the items. H2 ($c_{\text{male}} > c_{\text{female}}$) was only accepted for 2018, with small effect sizes, $r \in [0.1, 0.2]$; and for 2019, with a small effect size for the number of activity days, $r = 0.1$. Throughout the entire study period there was a negative monotonic evolution of effect sizes and a positive monotonic evolution of p-values. The differences in the average values of contribution between items on males and items on females also evolved monotonically (Figure 2), starting in 2018 with higher average values of contribution to items on males than items on females and ending in 2021 with lower ones. H3 ($pv_{\text{male}} > pv_{\text{female}}$) was accepted for 2018, 2019, and 2020, with small effect sizes, $r = 0.1$.
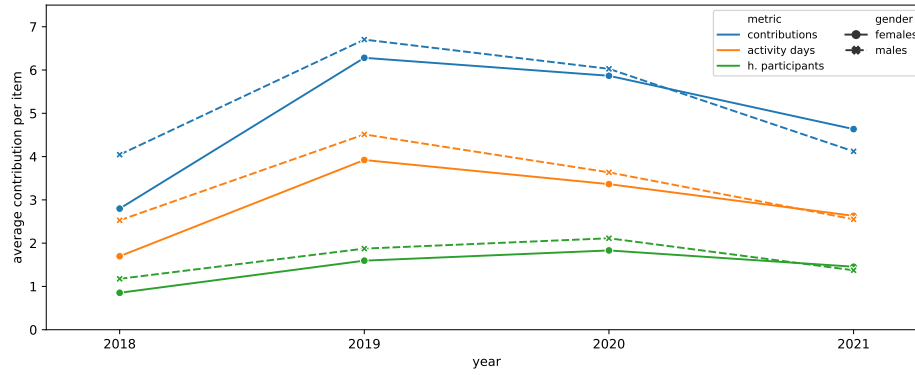
Fig. 2. Evolution of the average values of contribution metrics per item by gender.

H2–H4, metrics per year, items with links to the Wikipedias studied. H4 ($roi_{\text{female}} > roi_{\text{male}}$) was accepted for 2018 with small effect sizes, $r = 0.1$.

5.2   RQ2: Recency Gap

- People with items linked to articles in the Wikipedias studied tended to have more recent years of birth and death than the rest of the people with items.
- The item on a person tended to have more pageviews and higher ROI ratios associated with it the more recent the years of birth and death were.

H5. As hypothesized, there were significant differences between the years of birth in items linked to articles in the Wikipedias studied ($n = 11963, \text{Med} = 1942$) and those that were not ($n = 15098, \text{Med} = 1927$), two-sided $p < .001, r = .13$. There were also significant differences between the years of death in items linked to articles in the Wikipedias studied ($n = 5843, \text{Med} = 1962$) and those that were not ($n = 7706, \text{Med} = 1942$), two-sided $p < .001, r = .15$.
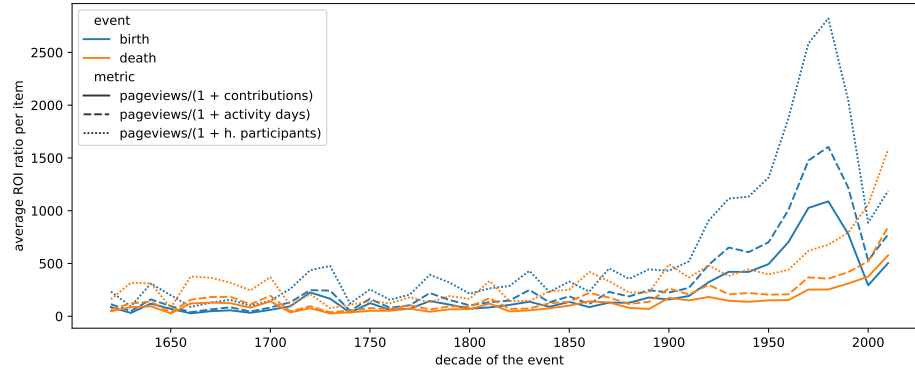
Fig. 3. Average ROI ratios per item by decade of the event.

H6. There was no conclusive (Spearman's ρ) rank-order correlation between years of birth between 1522 and 2021 and the proportions of Wikidata items with those years of birth that had links to articles in the Wikipedias studied, $\rho(493) = .10$, $p = .026$. The correlation was conclusive and positive for id. years of death $\rho(493) = .16$, $p < .001$.

H7, 2018–2021, all the items. For years of birth, there were positive Spearman's correlations with pageviews and ROI ratios per item, with small effect sizes, $\rho \in [0.1, 0.2]$. For years of death, there were positive Spearman's correlations with activity days, human participants, pageviews, and ROI ratios per item, with small effect sizes, $\rho \in [0.1, 0.2]$. See Figure 3.

H7, 2018–2021, items with links to the Wikipedias studied. For years of birth, there were positive correlations with pageviews and ROI ratios per item. For years of death, there were no correlations.

H7, metrics per year, all the items. Considering years of birth and Spearman's correlations with contribution per item, there were one positive correlation in 2018 and two in 2019; with pageviews per item, positive correlations in 2019, 2020, and 2021; and with ROI ratios, the three positive correlations in each of the years 2019, 2020, and 2021. Considering years of death and Spearman's correlations with contribution per item, there were the three positive correlations in 2018 and two in 2019; with pageviews per item, positive correlations for all years; and with ROI ratios, two positive correlations in 2018, and the three positive correlations in each of the years 2019, 2020, and 2021.

H7, metrics per year, items with links to the Wikipedias studied. Considering years of birth and Spearman's correlations, there were positive correlations with pageviews per item in 2019, 2020, and 2021; and with ROI ratios, in each of the years 2020 and 2021. Considering years of death and Spearman's correlations with ROI ratios, there were two positive correlations in 2021.

5.3 RQ3: Geographic and Socio-economic Gap

- The more populated a settlement, the more contribution and pageviews, and the higher ROI ratios.
- The higher the Human Development Index of a country or the country of a settlement, the more contribution and pageviews, and the higher ROI ratios.
- The higher the Human Development Index of a person's country, the more contribution, but not the more pageviews or the higher ROI ratios.
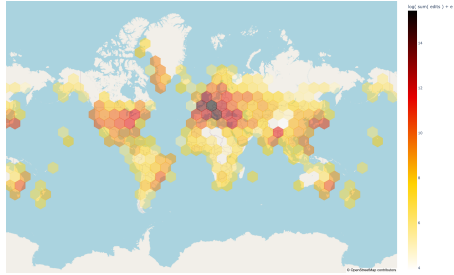

Fig. 4. Sum of contributions to items on settlements per area in logarithmic scale.
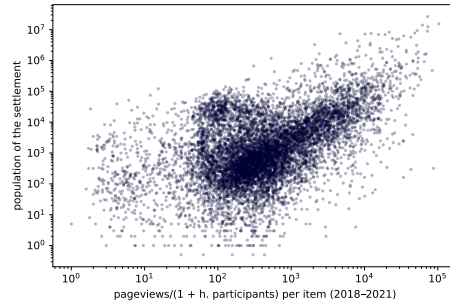
Fig. 5. Population of a settlement (y-axis, log) and the ROI ratio based on Wikipedia pageviews and number of human participants of the item (x-axis, log).

H8 (population). For the entire period, the strength of the correlations (ρ) was 0.31–0.37 with contribution metrics and 0.55–0.58 with pageviews and ROI ratios (Figure 5). Considering only those items linked to Wikipedia articles, the correlations with pageviews showed an increasing monotonic evolution, from 0.55 in 2018 to 0.60 in 2021; and ROI ratios as well, from 0.48–0.52 in 2018 to 0.60–0.61 in 2021.

H9 (HDI, people). There were positive correlations between the Human Development Index (HDI) of the countries of citizenship and the contribution metrics per item considering the entire study period (2018–2021), $\rho \in [0.1, 0.2]$. All HDI base indicators were positively correlated, except Life Expectancy at birth (LE) when considering all items. There was no correlation between HDI and pageviews or ROI ratios in any case. There were weak negative correlations between Life Expectancy at birth (LE) and ROI ratios when considering only items linked to Wikipedia articles, and between Life Expectancy at birth (LE) and both pageviews and ROI ratios when considering all items, $\rho \in [-0.1, -0.2]$.

H10 (HDI, settlements). There were positive Spearman's ρ rank-order correlations between the Human Development Index (HDI) of the countries of settlements and the contribution metrics, pageviews, and ROI ratios per item.

The strength of the correlations (ρ) was 0.4–0.5 with contribution metrics, 0.4 with pageviews, and 0.3 with ROI ratios. When considering only items linked to Wikipedia articles, the strength was 0.5–0.6 with contribution metrics, 0.5 with pageviews, and 0.3 with ROI ratios. See also Figure 4.

H11 (HDI, countries). There were positive Spearman's correlations between the Human Development Index (HDI) of the countries and the contribution metrics, pageviews, and ROI ratios per item. For the entire period, the strength of the correlations (ρ) was 0.31–0.51 with contribution metrics, 0.60 with pageviews, and 0.56–0.58 with ROI ratios.

6   Discussion

From the effect sizes we found that the influence of gender, time, and socio-economic factors on the contribution and information needs per item on a person was subtle. In contrast, the influence of geographic and socio-economic factors on the contribution and information needs associated with settlements and countries was considerable. We found no evidence of gender, recency, or urban imbalances in the contribution to Wikidata to a greater extent than in users' information needs or Wikipedia. This finding suggests that these content gaps are not endogenous to Wikidata, something consistent with previous literature. [1] documented that birth dates in DBpedia, sourced from Wikipedia, tended to be more recent than in Wikidata; and [75] found a systematic over-representation of males in Wikipedia compared to the labour market, whereas [78] found that Wikidata's representation of males was comparable to the professional societies' notability assessments. We did find a slightly larger socio-economic gap in the contribution to Wikidata than in users' information needs based on the development indices of the countries of citizenship of the people represented in Wikidata. This was not found in other classes of instances, such as settlements or the countries themselves. In summary, **the only content gaps we found that may be endogenous to Wikidata were subtle and related to socio-economic aspects of the people represented, whereas famous gaps such as gender and recency gaps could be explained by users' information needs, perhaps in conjunction with external systems** adjacent to Wikidata, e.g. Wikipedia and web search engines.

We argue that a KG's fine granularity and structure can act as equalizers of content differences between traditionally over- and under-represented groups. In Wikidata, **ontological properties (e.g. about people: place of birth/death, father, mother, etc.) leverage this fine granularity and can be understood as placeholders for information that is needed for the graph to be complete**, making missing information more explicit, and therefore helping to avoid gaps and biases. In contrast, it is not necessary, and generally not aligned with Wikipedia's policies, to create a Wikipedia article about e.g. a female just to mention her in the article about a male. [37] already considered that every Wikidata item on a "human without a Wikipedia article" was a "structural item" and exemplified that "a member of royalty without a Wikipedia article [...]

is needed to make a family tree complete". The presence of pre-established properties and constraints (e.g. ShEx [12, 61], Wikidata property constraints) with which to include the data, initially for making the KG ontologically predictable for software agents, can also help avoid bias, as it lets editors easily identify incompleteness at the KG entity level, and therefore solve it.

Meeting information needs is usually the purpose of a data source, and contribution is the form of resource investment through which a collaborative KG meets these needs. Therefore, the distance between the two is relevant and should be monitored. Nonetheless, distributing contribution solely on the basis of recorded information needs may not necessarily be the best decision. First, with this approach we estimate past needs and the extent to which contribution was aligned with them, but it would be preferable to determine which contribution will be able to meet future needs. Second, our metadata on information needs can only reflect the needs of the people who use the metadata source, Wikipedia in our case. However, not everyone consults Wikipedia for every information need, nor in the same cases, nor with the same frequency; in fact, access to Wikipedia has been banned or limited in entire countries [77, 68]. Finally, not all information needs may be equally pressing, but our framework does not make such a distinction.

7   Conclusions

Despite the Semantic Web community's interest and progress in measuring and improving the quality of KGs such as Wikidata, differences in content coverage persist for unclear reasons. It is possible to learn more about the socio-technical grounds of such differences by comparing the imbalances in the work on the KG with the imbalances in information needs considering the problematic attributes (e.g. gender). In this work we have defined a quantitative framework to achieve this and applied it to gender, recency, and geographic and socio-economic gaps in Wikidata. Our results suggest that, in general, these gaps are not endogenous to Wikidata's production, although exceptions are possible, e.g. based on development indices of people's countries of citizenship.

We plan to continue analysing content gaps in KGs. With a greater investment of resources, we will use the full sets of instances in Wikidata instead of samples. We will analyse more attributes and classes of instances, which could reveal or rule out more content gaps with respect to information needs. It would also be helpful to implement software solutions to monitor possible content gaps in KGs with the proposed approach, probably considering a shorter period. Finally, imbalances in contribution and information needs separately are also relevant and impact downstream applications, so monitoring them and warning users about their existence would be beneficial, as well as completing the KGs based on these insights.

Supplemental Material Statement: The datasets, SPARQL queries, and code are available on `https://github.com/davidabian/wikidata-gaps-vs-needs`.

References

1. Abián, D., Bernad, J., Trillo-Lado, R.: Using contemporary constraints to ensure data consistency. In: Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing. pp. 2303–2310. ACM, Limassol, Cyprus (Apr 2019). https://doi.org/10.1145/3297280.3297509, https://dl.acm.org/doi/10.1145/3297280.3297509

2. Ahmeti, A., Razniewski, S., Polleres, A.: Assessing the Completeness of Entities in Knowledge Bases. In: Blomqvist, E., Hose, K., Paulheim, H., Ławrynowicz, A., Ciravegna, F., Hartig, O. (eds.) The Semantic Web: ESWC 2017 Satellite Events. pp. 7–11. Lecture Notes in Computer Science, Springer International Publishing, Cham (2017). https://doi.org/10.1007/978-3-319-70407-4_2

3. Anthony, D., Smith, S.W., Williamson, T.: Reputation and Reliability in Collective Goods the Case of the Online Encyclopedia Wikipedia. Rationality and Society **21**(3), 283–306 (Aug 2009). https://doi.org/10.1177/1043463109336804

4. Arazy, O., Lifshitz-Assaf, H., Balila, A.: Neither a Bazaar nor a cathedral: The interplay between structure and agency in Wikipedia's role system. Journal of the Association for Information Science and Technology **70**(1), 3–15 (Jan 2019). https://doi.org/10.1002/asi.24076

5. Arazy, O., Lifshitz-Assaf, H., Nov, O., Daxenberger, J., Balestra, M., Cheshire, C.: On the "How" and "Why" of Emergent Role Behaviors in Wikipedia. Assoc Computing Machinery, New York (2017)

6. Arsanjani, J.J., Helbich, M., Bakillah, M., Loos, L.: The emergence and evolution of OpenStreetMap: a cellular automata approach. International Journal of Digital Earth **8**(1), 74–88 (Jan 2015). https://doi.org/10.1080/17538947.2013.847125

7. Arsanjani, J.J., Mooney, P., Helbich, M., Zipf, A.: An Exploration of Future Patterns of the Contributions to OpenStreetMap and Development of a Contribution Index. Transactions in Gis **19**(6), 896–914 (Dec 2015). https://doi.org/10.1111/tgis.12139

8. Balaraman, V., Razniewski, S., Nutt, W.: Recoin: Relative Completeness in Wikidata. In: Companion Proceedings of the World Wide Web Conference 2018 (www 2018). pp. 1787–1792. Assoc Computing Machinery, New York (2018). https://doi.org/10.1145/3184558.3191641

9. Begin, D., Devillers, R., Roche, S.: Contributors' Withdrawal from Online Collaborative Communities: The Case of OpenStreetMap. Isprs International Journal of Geo-Information **6**(11), 340 (Nov 2017). https://doi.org/10.3390/ijgi6110340

10. Beytía, P.: The Positioning Matters: Estimating Geographical Bias in the Multilingual Record of Biographies on Wikipedia. In: Companion Proceedings of the Web Conference 2020. pp. 806–810. WWW '20, Association for Computing Machinery, New York, NY, USA (Apr 2020). https://doi.org/10.1145/3366424.3383569

11. Bubendorff, S., Rizza, C., Prieur, C.: Construction and dissemination of information veracity on French social media during crises: Comparison of Twitter and Wikipedia. Journal of Contingencies and Crisis Management **29**(2), 204–216 (Jun 2021). https://doi.org/10.1111/1468-5973.12351

12. Candela, G., Escobar, P., Sáez, M., Marco-Such, M.: A Shape Expression approach for assessing the quality of Linked Open Data in libraries. Semantic Web **Preprint**(Preprint), 1–21 (2021). https://doi.org/10.3233/SW-210441

13. Cho, H., Chen, M., Chung, S.: Testing an Integrative Theoretical Model of Knowledge-Sharing Behavior in the Context of Wikipedia. Journal of the American Society for Information Science and Technology **61**(6), 1198–1212 (Jun 2010). https://doi.org/10.1002/asi.21316

14. Ciampaglia, G.L., Taraborelli, D.: MoodBar: Increasing New User Retention in Wikipedia through Lightweight Socialization. Assoc Computing Machinery, New York (2015). https://doi.org/10.1145/2675133.2675181

15. Ciglan, M., Nørvåg, K.: WikiPop: personalized event detection system based on Wikipedia page view statistics. In: Proceedings of the 19th ACM international conference on Information and knowledge management. pp. 1931–1932. CIKM '10, Association for Computing Machinery, New York, NY, USA (Oct 2010). https://doi.org/10.1145/1871437.1871769

16. Clement, M., Guitton, M.J.: Interacting with bots online: Users' reactions to actions of automated programs in Wikipedia. Computers in Human Behavior **50**, 66–75 (Sep 2015). https://doi.org/10.1016/j.chb.2015.03.078

17. Darari, F., Prasojo, R.E., Razniewski, S., Nutt, W.: COOL-WD: A Completeness Tool for Wikidata. In: SEMWEB (2017)

18. Darari, F., Razniewski, S., Prasojo, R.E., Nutt, W.: Enabling Fine-Grained RDF Data Completeness Assessment. In: Bozzon, A., CudreMauroux, P., Pautasso, C. (eds.) Web Engineering (icwe 2016). vol. 9671, pp. 170–187. Springer-Verlag Berlin, Berlin (2016). https://doi.org/10.1007/978-3-319-38791-8_10

19. Das, M., Hecht, B., Gergle, D.: The Gendered Geography of Contributions to OpenStreetMap: Complexities in Self-Focus Bias. In: Chi 2019: Proceedings of the 2019 Chi Conference on Human Factors in Computing Systems. Assoc Computing Machinery, New York (2019). https://doi.org/10.1145/3290605.3300793

20. Destandau, M., Fekete, J.D.: The missing path: Analysing incompleteness in knowledge graphs. Information Visualization **20**(1), 66–82 (Jan 2021). https://doi.org/10.1177/1473871621991539

21. Faerber, M., Bartscherer, F., Menne, C., Rettinger, A.: Linked Data Quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO. Semantic Web **9**(1) (2018)

22. Fraga, B.S., Couto da Silva, A.P., Murai, F.: Online Social Networks in Health Care: A Study of Mental Disorders on Reddit. Ieee, New York (2018). https://doi.org/10.1109/WI.2018.00-36

23. Galarraga, L., Razniewski, S., Amarilli, A., Suchanek, F.M.: Predicting Completeness in Knowledge Bases. In: Wsdm'17: Proceedings of the Tenth Acm International Conference on Web Search and Data Mining. pp. 375–383. Assoc Computing Machinery, New York (2017). https://doi.org/10.1145/3018661.3018739

24. Graells-Garrido, E., Lalmas, M., Menczer, F.: First Women, Second Sex: Gender Bias in Wikipedia. In: Proceedings of the 26th ACM Conference on Hypertext & Social Media. pp. 165–174. HT '15, Association for Computing Machinery, New York, NY, USA (Aug 2015). https://doi.org/10.1145/2700171.2791036

25. Graham, M., Dittus, M.: Geographies of Digital Exclusion: Data and Inequality. Pluto Press (Jan 2022)

26. Graham, M., Hogan, B., Straumann, R.K., Medhat, A.: Uneven Geographies of User-Generated Information: Patterns of Increasing Informational Poverty. Annals of the Association of American Geographers **104**(4), 746–764 (Jul 2014). https://doi.org/10.1080/00045608.2014.910087

27. Gumnishka, I.: Gender and racial bias in computer vision (Apr 2021), `https://thegoodai.co/2021/04/14/gender-and-racial-bias-in-computer-vision/`

28. Hecht, B., Gergle, D.: Measuring self-focus bias in community-maintained knowledge repositories. In: Proceedings of the fourth international conference on Communities and technologies. pp. 11–20. C&T '09, Association for Computing Machinery, New York, NY, USA (Jun 2009). https://doi.org/10.1145/1556460.1556463

29. Hecht, B., Stephens, M.: A Tale of Cities: Urban Biases in Volunteered Geographic Information. Proceedings of the International AAAI Conference on Web and Social Media **8**(1), 197–205 (May 2014), `https://ojs.aaai.org/index.php/ICWSM/article/view/14554`

30. Herfort, B., Lautenbach, S., de Albuquerque, J.P., Anderson, J., Zipf, A.: The evolution of humanitarian mapping within the OpenStreetMap community. Scientific Reports **11**(1), 3037 (Feb 2021). https://doi.org/10.1038/s41598-021-82404-z

31. Johnson, I.L., Lin, Y., Li, T.J.J., Hall, A., Halfaker, A., Schöning, J., Hecht, B.: Not at Home on the Range: Peer Production and the Urban/Rural Divide. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. pp. 13–25. CHI '16, Association for Computing Machinery, New York, NY, USA (May 2016). https://doi.org/10.1145/2858036.2858123

32. Kaffee, L.A., Endris, K.M., Simperl, E.: When Humans and Machines Collaborate: Cross-lingual Label Editing in Wikidata. Assoc Computing Machinery, New York (2019). https://doi.org/10.1145/3306446.3340826

33. Kamptner, E., Kessler, F.: Small-scale crisis response mapping: comparing user contributions to events in OpenStreetMap. Geojournal **84**(5), 1165–1185 (Oct 2019). https://doi.org/10.1007/s10708-018-9912-1

34. Keegan, B., Gergle, D., Contractor, N.: Hot Off the Wiki: Structures and Dynamics of Wikipedia's Coverage of Breaking News Events. American Behavioral Scientist **57**(5), 595–622 (May 2013). https://doi.org/10.1177/0002764212469367

35. Keegan, B.C., Brubaker, J.R.: 'Is' to 'Was': Coordination and Commemoration on Posthumous Wikipedia Biographies. Assoc Computing Machinery, New York (2015). https://doi.org/10.1145/2675133.2675238

36. Keet, C.M.: An exploration into cognitive bias in ontologies. In: Joint Ontology Workshops 2021 Episode VII: The Bolzano Summer of Knowledge, JOWO 2021. vol. 2969 (2021), `https://pubs.cs.uct.ac.za/id/eprint/1474/`

37. Klein, M., Gupta, H., Rai, V., Konieczny, P., Zhu, H.: Monitoring the Gender Gap with Wikidata Human Gender Indicators. Assoc Computing Machinery, New York (2016). https://doi.org/10.1145/2957792.2957798

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

39. Lin, Y., Chen, Y.: Do less active participants make active participants more active? An examination of Chinese Wikipedia. Decision Support Systems **114**, 103–113 (Oct 2018). https://doi.org/10.1016/j.dss.2018.08.002

40. Luggen, M., Difallah, D., Sarasua, C., Demartini, G., Cudre-Mauroux, P.: Non-parametric Class Completeness Estimators for Collaborative Knowledge Graphs-The Case of Wikidata. In: Ghidini, C., Hartig, O., Maleshkova, M., Svatek, V., Cruz, I., Hogan, A., Song, J., Lefrancois, M., Gandon, F. (eds.) Semantic Web - Iswc 2019, Pt I, vol. 11778, pp. 453–469. Springer International Publishing Ag, Cham (2019). https://doi.org/10.1007/978-3-030-30793-6_26

41. Ma, D., Sandberg, M., Jiang, B.: Characterizing the Heterogeneity of the OpenStreetMap Data and Community. Isprs International Journal of Geo-Information **4**(2), 535–550 (Jun 2015). https://doi.org/10.3390/ijgi4020535

42. McDaniel, M., Storey, V.C.: Evaluating Domain Ontologies: Clarification, Classification, and Challenges. ACM Computing Surveys **52**(4), 70:1–70:44 (Sep 2019). https://doi.org/10.1145/3329124

43. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., Galstyan, A.: A Survey on Bias and Fairness in Machine Learning. ACM Computing Surveys **54**(6), 115:1–115:35 (Jul 2021). https://doi.org/10.1145/3457607

44. Miquel Ribé, M., Vaidla, K., Fort, F., Torres, A.: Wikimedia 2030 movement strategy : How an inclusive open strategy process has placed people at the centre. BiD: textos universitaris de biblioteconomia i documentació (47) (Dec 2021). https://doi.org/10.1344/BiD2021.47.05
45. Miščević, N.: United nations development programme, human development report 2020. the next frontier human development and the anthropocene. Croatian Journal of Philosophy **21**(1 (61)), 231–235 (2021)
46. Mooney, P., Corcoran, P.: Analysis of Interaction and Co-editing Patterns amongst OpenStreetMap Contributors. Transactions in Gis **18**(5), 633–659 (Oct 2014). https://doi.org/10.1111/tgis.12051
47. Mueller-Birn, C., Karran, B., Lehmann, J., Luczak-Roesch, M.: Peer-production system or collaborative ontology engineering effort: What is Wikidata? Assoc Computing Machinery, New York (2015). https://doi.org/10.1145/2788993.2789836
48. Nielsen, F.Å., Mietchen, D., Willighagen, E.: Scholia, scientometrics and wikidata. In: Blomqvist, E., Hose, K., Paulheim, H., Ławrynowicz, A., Ciravegna, F., Hartig, O. (eds.) The Semantic Web: ESWC 2017 Satellite Events. pp. 237–259. Springer International Publishing, Cham (2017)
49. Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M.E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., Broelemann, K., Kasneci, G., Tiropanis, T., Staab, S.: Bias in data-driven artificial intelligence systems—An introductory survey. WIREs Data Mining and Knowledge Discovery **10**(3), e1356 (2020). https://doi.org/10.1002/widm.1356, `https://onlinelibrary.wiley.com/doi/abs/10.1002/widm.1356`
50. Park, H., Park, S.J.: Communication behavior and online knowledge collaboration: evidence from Wikipedia. Journal of Knowledge Management **20**(4), 769–792 (2016). https://doi.org/10.1108/JKM-08-2015-0312
51. Piccardi, T., West, R., Redi, M., Colavizza, G.: Quantifying Engagement with Citations on Wikipedia. Nauchnye I Tekhnicheskie Biblioteki-Scientific and Technical Libraries (10), 77–86 (2020). https://doi.org/10.33186/1027-3689-2020-10-63-76
52. Piscopo, A., Simperl, E.: What we talk about when we talk about Wikidata quality: a literature survey. Assoc Computing Machinery, New York (2019). https://doi.org/10.1145/3306446.3340822
53. Prasojo, R.E., Darari, F., Razniewski, S., Nutt, W.: Managing and Consuming Completeness Information for Wikidata Using COOL-WD. In: Proceedings of the 7th International Workshop on Consuming Linked Data (COLD) (2016), `http://ceur-ws.org/Vol-1666/paper-02.pdf`
54. Radstok, W., Chekol, M.W., Schaefer, M.T., et al.: Are knowledge graph embedding models biased, or is it the data that they are trained on? In: Wikidata Workshop 2021 co-located with the 20th International Semantic Web Conference (ISWC 2021) (2021)
55. Rea, L.M., Parker, R.A.: Designing and Conducting Survey Research: A Comprehensive Guide. Jossey-Bass Publishers (1992)
56. Redi, M., Gerlach, M., Johnson, I., Morgan, J., Zia, L.: A Taxonomy of Knowledge Gaps for Wikimedia Projects (Second Draft). arXiv:2008.12314 [cs] (Jan 2021), `http://arxiv.org/abs/2008.12314`
57. Restivo, M., van de Rijt, A.: No praise without effort: experimental evidence on how rewards affect Wikipedia's contributor community. Information Communication & Society **17**(4), 451–462 (Apr 2014). https://doi.org/10.1080/1369118X.2014.888459

58. Ringler, D., Paulheim, H.: One Knowledge Graph to Rule Them All? Analyzing the Differences Between DBpedia, YAGO, Wikidata & co. In: KernIsberner, G., Furnkranz, J., Thimm, M. (eds.) Ki 2017: Advances in Artificial Intelligence, vol. 10505, pp. 366–372. Springer International Publishing Ag, Cham (2017). https://doi.org/10.1007/978-3-319-67190-1_33

59. Rizoiu, M.A., Xie, L., Caetano, T., Cebrian, M.: Evolution of Privacy Loss in Wikipedia. Assoc Computing Machinery, New York (2016). https://doi.org/10.1145/2835776.2835798

60. Samoilenko, A., Lemmerich, F., Weller, K., Zens, M., Strohmaier, M.: Analysing Timelines of National Histories Across Wikipedia Editions: A Comparative Computational Approach. Proceedings of the International AAAI Conference on Web and Social Media **11**(1), 210–219 (May 2017), https://ojs.aaai.org/index.php/ICWSM/article/view/14881

61. Samuel, J.: ShExStatements: Simplifying Shape Expressions for Wikidata. In: Companion Proceedings of the Web Conference 2021. pp. 610–615. WWW '21, Association for Computing Machinery, New York, NY, USA (Apr 2021). https://doi.org/10.1145/3442442.3452349

62. Sarasua, C., Checco, A., Demartini, G., Difallah, D., Feldman, M., Pintscher, L.: The Evolution of Power and Standard Wikidata Editors: Comparing Editing Behavior over Time to Predict Lifespan and Volume of Edits. Computer Supported Cooperative Work-the Journal of Collaborative Computing and Work Practices **28**(5), 843–882 (Sep 2019). https://doi.org/10.1007/s10606-018-9344-y

63. Schott, M., Grinberger, A.Y., Lautenbach, S., Zipf, A.: The Impact of Community Happenings in OpenStreetMap-Establishing a Framework for Online Community Member Activity Analyses. Isprs International Journal of Geo-Information **10**(3), 164 (Mar 2021). https://doi.org/10.3390/ijgi10030164

64. Sheehan, E., Meng, C., Tan, M., Uzkent, B., Jean, N., Burke, M., Lobell, D., Ermon, S.: Predicting Economic Development using Geolocated Wikipedia Articles. Assoc Computing Machinery, New York (2019). https://doi.org/10.1145/3292500.3330784

65. Stanczak, K., Augenstein, I.: A Survey on Gender Bias in Natural Language Processing. arXiv:2112.14168 [cs] (Dec 2021), http://arxiv.org/abs/2112.14168

66. Tanon, T.P., Suchanek, F.: Querying the Edit History of Wikidata. In: Hitzler, P., Kirrane, S., Hartig, O., DeBoer, V., Vidal, M.E., Maleshkova, M., Schlobach, S., Hammar, K., Lasierra, N., Stadtmuller, S., Hose, K., Verborgh, R. (eds.) Semantic Web: Eswc 2019 Satellite Events, vol. 11762, pp. 161–166. Springer International Publishing Ag, Cham (2019). https://doi.org/10.1007/978-3-030-32327-1_32

67. Twyman, M., Keegan, B.C., Shaw, A.: Black Lives Matter in Wikipedia: Collaboration and Collective Memory around Online Social Movements. Assoc Computing Machinery, New York (2017). https://doi.org/10.1145/2998181.2998232

68. Uzun, R.: National Interest vs. Online Freedom of Expression: The Discussions of Internet Users on the Blocking of 'Wikipedia' in Turkey. Etkileşim (5), 10–22 (Apr 2020). https://doi.org/10.32739/etkilesim.2020.5.82, https://dergipark.org.tr/en/pub/usuifade/issue/53630/715436

69. Vrandečić, D.: Ontology Evaluation. In: Staab, S., Studer, R. (eds.) Handbook on Ontologies, pp. 293–313. International Handbooks on Information Systems, Springer, Berlin, Heidelberg (2009). https://doi.org/10.1007/978-3-540-92673-3_13

70. Wagner, C., Graells-Garrido, E., Garcia, D., Menczer, F.: Women through the glass ceiling: gender asymmetries in Wikipedia. Epj Data Science **5**, 5 (Mar 2016). https://doi.org/10.1140/epjds/s13688-016-0066-4

71. Wisesa, A., Darari, F., Krisnadhi, A., Nutt, W., Razniewski, S.: Wikidata Completeness Profiling Using ProWD. Assoc Computing Machinery, New York (2019). https://doi.org/10.1145/3360901.3364425

72. Wolfer, S., Mueller-Spitzer, C.: How Many People Constitute a Crowd and What Do They Do? Quantitative Analyses of Revisions in the English and German Wiktionary Editions. Lexikos **26**, 347–371 (2016)

73. Yasseri, T., Kertesz, J.: Value Production in a Collaborative Environment Sociophysical Studies of Wikipedia. Journal of Statistical Physics **151**(3-4), 414–439 (May 2013). https://doi.org/10.1007/s10955-013-0728-6

74. Young, A.G., Wigdor, A.D., Kane, G.C.: The Gender Bias Tug-of-War in a Co-creation Community: Core-Periphery Tension on Wikipedia. Journal of Management Information Systems **37**(4), 1047–1072 (Oct 2020). https://doi.org/10.1080/07421222.2020.1831773

75. Zagovora, O., Floeck, F., Wagner, C.: "(Weitergeleitet von Journalistin)": The Gendered Presentation of Professions on Wikipedia. Assoc Computing Machinery, New York (2017). https://doi.org/10.1145/3091478.3091488

76. Zaveri, A., Rula, A., Maurino, A., Pietrobon, R., Lehmann, J., Auer, S.: Quality assessment for Linked Data: A Survey. Semantic Web **7**(1), 63–93 (2016). https://doi.org/10.3233/SW-150175

77. Zhang, A.F., Livneh, D., Budak, C., Robert, L., Romero, D.: Shocking the Crowd: The Effect of Censorship Shocks on Chinese Wikipedia. Proceedings of the International AAAI Conference on Web and Social Media **11**(1), 367–376 (May 2017), `https://ojs.aaai.org/index.php/ICWSM/article/view/14895`

78. Zhang, C.C., Terveen, L.: Quantifying the Gap: A Case Study of Wikidata Gender Disparities. In: 17th International Symposium on Open Collaboration. pp. 1–12. OpenSym 2021, Association for Computing Machinery, New York, NY, USA (Sep 2021). https://doi.org/10.1145/3479986.3479992

79. Zhang, X., Wang, C.: Network Positions and Contributions to Online Public Goods: The Case of Chinese Wikipedia. Journal of Management Information Systems **29**(2), 11–40 (2012). https://doi.org/10.2753/MIS0742-1222290202

80. Zhu, H., Kraut, R.E., Kittur, A.: Effectiveness of Shared Leadership in Wikipedia. Human Factors **55**(6), 1021–1043 (Dec 2013). https://doi.org/10.1177/0018720813515704

81. Zia, M., Seker, D.Z., Cakir, Z.: Spatial Evolution of Openstreetmap Dataset in Turkey. In: Isikdag, U., Rahman, A.A., Castro, F.A., Karas, I.R. (eds.) 3rd International Geoadvances Workshop. vol. 42-2, pp. 169–172. Copernicus Gesellschaft Mbh, Gottingen (2016). https://doi.org/10.5194/isprs-archives-XLII-2-W1-169-2016

82. Zia, M., Cakir, Z., Seker, D.Z.: Turkey OpenStreetMap Dataset - Spatial Analysis of Development and Growth Proxies. Open Geosciences **11**(1), 140–151 (Jan 2019). https://doi.org/10.1515/geo-2019-0012