# Implicit Bias in Crowdsourced Knowledge Graphs

Gianluca Demartini
University of Queensland, Australia
demartini@acm.org

## ABSTRACT

Collaborative creation of knowledge is an approach which has been successfully demonstrated by crowdsourcing project like Wikipedia. Similar techniques have recently been adopted for the creation of collaboratively generated Knowledge Graphs like, for example, Wikidata. While such an approach enables the creation of high quality structured content, it also comes with the challenge of introducing contributors' implicit bias in the generated Knowledge Graph. In this paper, we investigate how paid crowdsourcing can be used to understand contributor bias for controversial facts to be included into collaborative Knowledge Graphs. We propose methods to trace the provenance of crowdsourced fact checking thus enabling bias transparency rather than aiming at eliminating bias from the Knowledge Graph.

**ACM Reference Format:**
Gianluca Demartini. 2019. Implicit Bias in Crowdsourced Knowledge Graphs. In *Companion Proceedings of the 2019 World Wide Web Conference (WWW '19 Companion), May 13–17, 2019, San Francisco, CA, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3308560.3317307

## 1 INTRODUCTION

General-purpose Knowledge Graphs incorporate factual statements that can be accessed by end-users and that can be integrated across different structured datasets [3]. Knowledge Graphs are typically constructed in a bottom-up fashion by making semantics emerge from data like, for example, textual document collections or Web pages [2, 12]. However, Knowledge Graphs can also be created by means of large-scale manual effort like for example the classic Cyc ontology project [13] that required 900 person-years to be created. More recently, the crowdsourcing project Wikidata [21] was kicked-off to generate a general purpose Knowledge Graph by collecting input from thousands of individual contributors who continuously add and edit factual statements in the Knowledge Graph.

The use of crowdsourcing in combination with machine-based algorithms has been often used for different Semantic Web problems [8, 14, 19]. This is typically done thanks to crowdsourcing platforms like Amazon MTurk [6] which enable programmatic access to a crowd of contributors by means of Human Intelligence Tasks (HITs) that are published by requesters and completed by workers in exchange for small monetary rewards. Common crowdsourcing tasks in the Semantic Web domain include schema mapping, entity extraction and linking, relation extraction, and ontology construction and verification.


*WWW '19 Companion, May 13–17, 2019, San Francisco, CA, USA*

ACM ISBN 978-1-4503-6675-5/19/05.
https://doi.org/10.1145/3308560.3317307

While leveraging crowdsourcing for Knowledge Graph construction, human contributors bring their own point of view on the factual statements they add to the Knowledge Graph. For controversial statements like, for example, Catalonia being part of Spain or being an independent country, contributors may bring their own implicit bias into the Knowledge Graph. There are different ways to deal with this issue. On the one hand, it is possible to pre-select crowd contributors and assign them editorial micro-tasks based on user profiles (e.g., including information on contributors' background) and to build a sample of contributors which is representative of the population, thus removing or diminishing any possible bias effect. On the other hand, we can maintain the idea of having an open call for contributions as in classic crowdsourcing projects, but then keep track of information about workers' background as well as about the information seeking process (i.e., which search engine has been used to retrieve the information, which search query, and which result has been used to identify the contributed information) used to reach a certain conclusion thus tracking potential bias in the answers.

In this paper we focus on the latter approach to deal with bias in crowdsourced Knowledge Graphs: We propose a model to measure and track the implicit bias of crowd contributors into the Knowledge Graph. We present the results of a study of crowdsourced controversial fact verification asking the crowd to provide evidence supporting their claims. Based on such supporting evidence we are able to identify the different points of view that exist in the crowd and keep track of bias in the Knowledge Graph together with the factual statements inserted into it.

The main contribution of this paper are the following:

- We perform a large-scale study involving 600 crowd workers to understand what aspects of worker demographics and of the knowledge creation process significantly affect the statements which are stored in Knowledge Graphs.
- We propose a model to keep track of bias information in crowdsourced Knowledge Graphs like Wikidata and demonstrate the benefits of surfacing bias information to end users of applications like, for example, Semantic Search.

The rest of this paper is structured as follows. In the next section we briefly summarize related work about bias in Web Search and on the use of crowdsourcing for Semantic Web data. Then, in Section 3 we present the results of three crowdsourcing studies aimed at understanding which are the dimensions that significantly affect the answer to controversial fact checking HITs to decide what bias information is most useful to keep track of in Knowledge Graphs. In Section 4 we explain how such bias information can be stored in Knowledge Graphs. In Section 5 we discuss the implication of such additional bias data collection for storage and retrieval in Knowledge Graphs like Wikidata. Finally, in Section 6 we draw the main conclusions of this work.

## 2 RELATED WORK

Bias is a well studied concept in Psychology research. A popular example of a large-scale Web study of bias is the Implicit Bias project run by Harvard [9] where reaction tests have been performed to measure implicit bias towards gender, skin color, and other dimensions.

In the area of crowdsourcing, bias has been studied less, but some previous work serves as the basis for this study. Some work has studied how crowdsourcing contributors are biased towards certain attitudes. For example, Otterbacher [16] looked at how certain images lead to diverse annotations in a game-with-a-purpose setting. Results show that images depicting women tend to be described using more subjective adjectives thus confirming the presence of gender-based stereotypes.

Eickhoff [7] looked at the effect of cognitive bias in crowdsourcing. He showed how crowd workers are affected by fellow workers' answer (Bandwagon effect) and by being presented with multiple options (Decoy effect). The existence of the Decoy effect proves that workers judgment is indeed affected by other tasks they have seen before approaching a specific HIT and, more generally, by the their cultural background defined as all the information they have been exposed to before reaching the specific crowdsourcing task. Hube et al. [10] measured the effect of implicit bias in paid crowdsourcing focusing on subjective data labelling tasks. Roitero et al. [17] used crowdsourcing for fact checking presenting an experimental comparison of different judgment scales. In our work we look at the effect of crowd workers' implicit bias on the verification of controversial factual knowledge to be included in a Knowledge Graph.

Crowdsourcing platforms like Amazon MTurk enable access to workers having a large diversity of cultural backgrounds [18]. The two countries where most crowd workers on these platform are from are United States and India [6]. In this paper we focus on these two different groups also because of the different cultural context they are embedded into and how it affects their answers on controversial fact checking HITs.

A common approach to deal with noise and low quality data in crowdsourcing is to aggregate answers collected for the same task from multiple workers by, for example, majority vote or by using more advanced aggregation methods [20]. An alternative approach to aggregating answers is to analyze information about disagreement among workers [1] or to compute confidence measures about the answer [11] by means of appropriate agreement measures [4].

Crowdsourcing has been extensively used for Semantic Web and Linked Data problems. Popular examples of human-in-the-loop solutions using paid micro-task crowdsourcing include the work on schema mapping by Sarasua et al. [19]. Another example demonstrates the use of games-with-a-purpose for entity extraction in text [8]. Mortensen et al. [14] show how crowdsourcing can be used to verify relationships among classes of domain-specific ontologies. They show that even in cases where domain expertise is required to complete the task accurately, crowd workers can perform as good as experts if provided with enough contextual information about the classes they are considering in the HIT. For a comprehensive overview of human-in-the-loop solutions for Web Science problems we refer to [5]. The question we aim to answer in this work is on the effect of crowd workers' implicit bias contributing data to Knowledge Graphs. A strong motivational example for this research question is Wikidata [21] which is a collaboratively created Knowledge Graph to which anyone on the Web can contribute new facts and modify existing ones.

Tracking provenance information in RDF is not a new problem but early solutions have shown to be inefficient by creating an overhead in terms of data to be stored and in indexing resourced required [15]. However, more recently, RDF stores that can efficiently index and query provenance metadata have been created [23]. More than better RDF stores, better ways to represent meta-information about RDF statements (which are more efficient than reification) are being used in practice in Knowledge Graphs like YAGO2 [15]. Such work enables the methods described in our paper by providing a scalable and reliable infrastructure for bias-tracking in Knowledge Graphs.

## 3 CROWDSOURCING CONTROVERSIAL FACT CHECKING

In this section we report the results of an experimental study we performed on top of the Amazon MTurk crowdsourcing platform about how crowd workers can contribute to the verification of controversial facts and on what factors influence their contributions.

### 3.1 Crowdsourcing Setting

We analyzed three different scenarios including two controversial fact verification tasks and one fake news verification task. For each use case, we recruited 200 crowd workers (100 from US and 100 from India) for a total of 600 workers involved in the study run in January 2018. We rewarded them $0.05 for controversial fact HITs and $0.10 for the fake news verification HIT. For each task we asked a binary question (yes/no or true/false) about a certain statement or news video. For each answer they provided, we also asked workers to justify it: We requested to provide information about the search engine used to find the answer, which search query they used, which URL contains information in support to their answer, and in which position the URL was ranked in the search engine result page (SERP). We additionally collect information about workers' age and gender.

### 3.2 Study I: Catalonia Independence

The first controversial fact HIT we run asked crowd workers to answer the question and provide evidence on whether Catalonia is an independent country or not. We hypothesize this being a controversial fact given the independence referendum held in October 2017 in Catalonia which has been deemed unlawful by Spanish courts. Out of 200 workers, 116 answered 'yes' and 84 answered 'no' to the question on Catalonia independence, showing how the question is indeed controversial and a clear truth is not obviously obtainable by aggregating such answers. This is consistent with claims from previous work [1] in support of not aggregating crowd answers but rather maintaining the diversity of opinions expressed by the crowd. When comparing the two crowd populations (US and India) we observe that in both cases the majority of respondents answered 'yes' but with a smaller majority in US (47 vs 46) as compared to

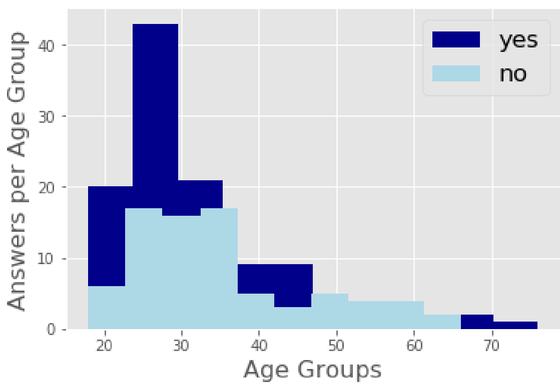



**Figure 1: Age distribution of workers responding 'yes' and 'no' to the question on whether Catalonia is an independent country.**

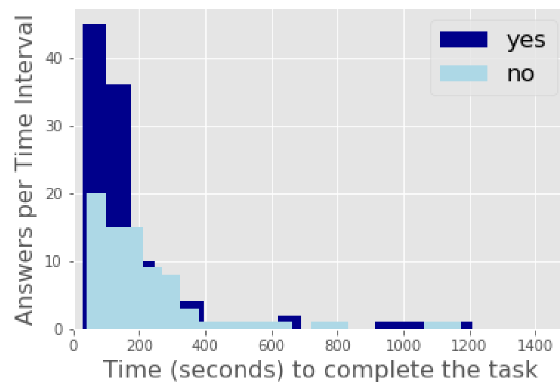



**Figure 2: Task completion time distribution for workers responding 'yes' and 'no' to the question on whether Catalonia is an independent country.**

India (62 vs 33). A Chi-squared test shows a not significant effect of the location of workers (p-value=0.058).

*Age Bias.* The age distribution of respondents is presented in Figure 1. We can see that a large group of workers who provided positive answers towards Catalonia independence tends to be younger than average. Indeed, the median age of workers who responded 'yes' is 28 as compared to the median age of those who responded 'no' which is 32.5. The effect of age on the answer for this HIT is significant (t-test $p < 0.05$) indicating that workers who answered 'yes' are significantly younger. This is another demonstration that based on the crowd population who provides factual statements to a Knowledge Graph, different information could be encoded in it.

*Gender Bias.* When looking at the distribution of answers over different gender groups we see that 35 male and 73 female workers answered 'yes' while 46 male and 33 female workers answered 'no'. We can observe a clear pattern in terms of gender bias among the respondents who provided their gender information: Female contributors achieved a strong majority for 'yes' while male contributors had a majority for 'no'. A Chi-square test shows a significant effect ($p < 0.01$) of gender on the answer.

*Task Duration Bias.* Figure 2 shows the distribution of the time taken to answer the questions in the task. In terms of time taken to complete the task, workers who responded 'yes' took a median time of 110 seconds to complete the task, while workers who answered 'no' took longer (median=180s). The difference in time is not significant showing that the time taken to complete the task had no impact on the actual answer.

*Search Result Rank Bias.* Figure 3 shows the ranking of the URL used to support the answer about Catalonia being an independent country as presented in the SERP used to identify the answer and relative supporting evidence. In terms of ranking, we can observe that the median ranking of URLs that support 'yes' answers is 1.62 which is significantly different (t-test $p < 0.01$) from the rank of URLs used as support by workers who answered 'no' (median=2.32).

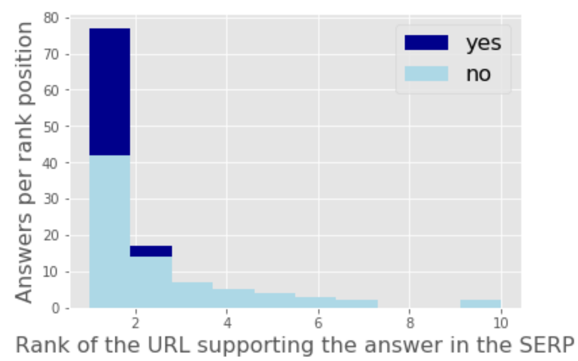



**Figure 3: Search result ranking of the URL used as supporting evidence by workers responding 'yes' and 'no' to the question on whether Catalonia is an independent country.**

This indicates that workers who go lower in the SERP to find the supporting evidence tend to answer 'no'.

*Top-level Domain of Supporting Evidence.* When looking at the most used domains for the supporting evidence URLs, we observe that the most popular sources of evidence are 'en.wikipedia.org ' (49 times), 'www.independent.co.uk' (43), 'www.google.com'[1] (20), 'www.telegraph.co.uk' (13), and 'www.bbc.com' (10). When differentiating the source of supporting evidence for 'yes' and 'no' answers, we obtain the results presented in Table 1. We can observe how the top domains for each answer are quite different showing a prevalence of 'www.independent.co.uk' for workers who answered 'yes'.

## 3.3 Study II: Israel Capital

A second experiment we run on Amazon MTurk asked 200 workers (100 from US and 100 from India) which is the capital city of Israel

[1]The URLs having as domain www.google.com are SERPs indicating how workers picked their answers based on the snippets of information provided in the SERP without clicking further away to any retrieved result.

**Table 1: Top-3 domains of URLs supporting 'yes' and 'no' answers to whether Catalonia is an independent country.**

| Domain supporting Yes | Count | Domain supporting No | Count |
|---|---|---|---|
| www.independent.co.uk | 35 | en.wikipedia.org | 26 |
| en.wikipedia.org | 23 | www.bbc.com | 10 |
| www.google.com | 18 | www.telegraph.co.uk | 9 |

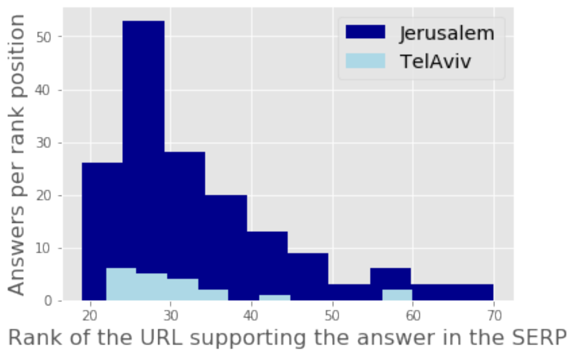



**Figure 4: Age distribution of workers responding to the question on which is the capital of Israel.**

giving the options 'Jerusalem', 'Tel Aviv', and 'Other' as possible answers. We hypothesize this being a controversial fact given the decision to move the US embassy from Tel Aviv to Jerusalem and the following discussion that took place at the United Nations. Out of 200 workers, 164 answered 'Jerusalem', 20 answered 'Tel Aviv', and 10 answered 'Other' to the question, showing how the question is less controversial than the previous one about Catalonia independence. Even in this case we can perform an analysis of the supporting evidence provided by the crowd workers to understand where the minority point of view comes from and aim at not discarding it by simply aggregating the data and taking the most popular answer.

When comparing workers from India and US we did not observe a significant difference (chi-square $p > 0.05$) in their answers with a strong majority for 'Jerusalem' as the answer in both populations.

*Age Bias.* The age distribution of respondents is presented in Figure 4. We can see that the distribution of the age from the two groups is similar with median values of 30 for the group answering 'Jerusalem' and 28.5 for the group answering 'Tel Aviv'. No statistically significant effect of age has been observed.

*Gender Bias.* Figure 5 shows the gender bias in the answers. Again, we can see similar distributions across the two groups and no significant effect of gender on the answer has been observed.

*Task Duration Bias.* Figure 6 shows the distribution of the time taken to answer the questions in the task. In terms of time taken to complete the task, we can observe a similar distribution for both answers (the median time of 'Jerusalem' respondent was 140.5s; the median time of 'Tel Aviv' respondent was 143.5s). Indeed, a statistical test shows no significant differences in the time taken to answer.

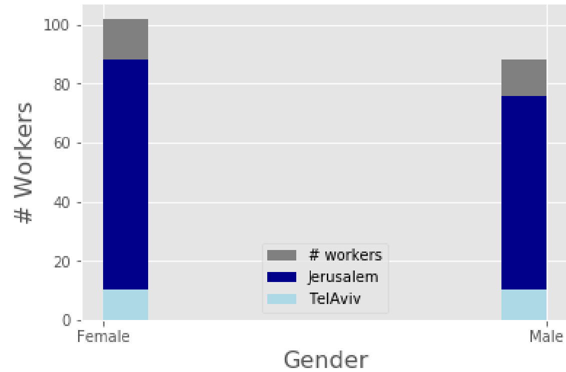



**Figure 5: Gender distribution of workers responding to the question on which is the capital of Israel.**

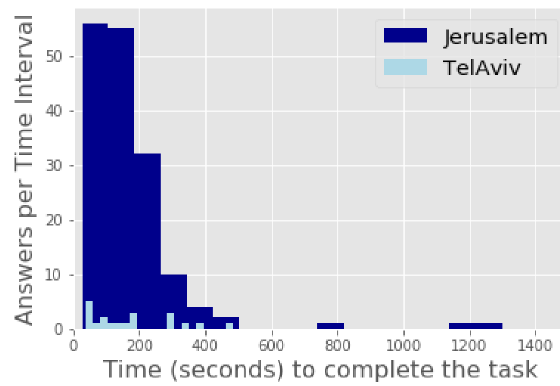



**Figure 6: Task completion time for workers responding to which is the capital of Israel.**

*Search Result Rank Bias.* Figure 7 shows the ranking of the URL used to support the answer about the capital of Israel as presented in the SERP used to identify the answer and the relative supporting evidence. In terms of ranking, we can observe very similar patterns with the median rank of supporting URLs at 2.3 for 'Jerusalem' and at 2.35 for 'Tel Aviv' with no statistically significant difference.

## 3.4 Study III: Pope News

Finally, we crowdsourced a task where we asked crowd workers whether a video[2] showing the Pope performing a trick was real or fake, again, also asking to provide supporting evidence for their claim. In this case the fact was not controversial but rather meant to check the veracity of the news which was not real but made up for a TV show.

Out of the 200 crowd workers involved in the study, 107 said that the video news story is real while 93 said that it is fake. In this case, despite the objectivity of the fact being false, we observed high disagreement levels.

[2] http://www.youtube.com/embed/MYjeWGH8Z1c

When comparing the two crowd population (US and India) we observe that the majority of Indian workers believed the news being real (66 vs 34) while the majority of workers from US (59 vs 41) believed the news was fake. The fake video was indeed created for a US-based TV show justifying the higher awareness of US workers. A Chi-squared test shows a significant difference among the answering pattern of the Indian and US-based worker populations ($p < 0.01$).

*Age Bias.* Figure 8a shows the distribution of workers' age based on their answer to the fake news identification task. We can observe that workers who believe the news being authentic are younger (median=28) as compared to those believing the news being fake (median=32). A t-test shows a statistically significant effect of workers' age on believing or not to the news ($p < 0.01$) demonstrating how younger contributors are more prone to being mislead by this fake news.

*Task Duration Bias.* Figure 8b shows the distribution of the time taken to complete the HIT. While the median time to answer the task was higher (267.5s) for contributors who stated that news is not real as compared to contributors who said the news is real (187.5s) a t-test shows no significant differences on the time taken to answer based on the answer.

*Search Result Rank Bias.* Figure 9 shows the position of the supporting evidence identified by workers in the SERP for the fake news identification task. The median rank of the supporting evidence for respondents who believe the news is real was 2 while for respondent who believe the news is fake was 2.3. While workers who did go deeper in the ranking managed to identify the non-truthfulness of the news, no statistically significant difference has been observed based on the rank of the evidence.

## 3.5 Result Discussion

From the results presented above we understand that while certain facts may be more controversial than others (with the 'Catalonia independence' task showing more controversy than the 'Israel capital'

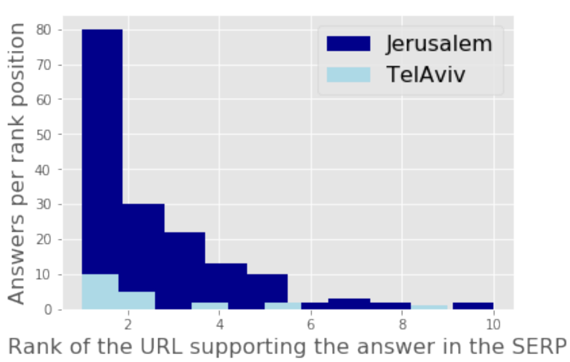
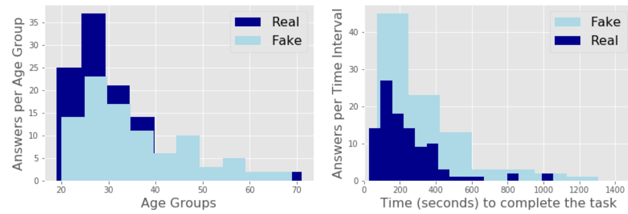



**Figure 7: Search result ranking of the URL used as supporting evidence by workers responding to the question on which is the capital of Israel.**




(a) (b)

**Figure 8: (a) Age of crowd workers responding to the task of fake news identification. (b) Time taken to answer to the fake news task by crowd workers.**

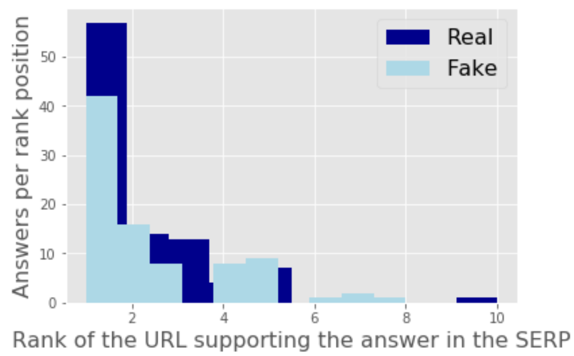



**Figure 9: Search result ranking of the URL used as supporting evidence by workers responding to the fake news identification task.**

task), we have observed the importance of keeping track of information provenance (both contributors' background and the source of evidence) when creating a crowdsourced Knowledge Graph.

A recent attempt to keep track of provenance in Knowledge Graphs is the concept of *reference* in Wikidata. For example, the statement 'Tom Cruise' - 'date of birth' -'3 July 1962' has 6 references among which one is from English Wikipedia[3]. While this is useful information to support facts stored in Knowledge Graphs, it lacks meta-information about how this was identified as a source. In the experiments we presented above, we have shown the effect of how different contributors may end up providing different statements by using different support evidences as reference for the statement. We claim it is then important to incorporate provenance metadata about contributors' implicit bias as well as the statement source of evidence.

# 4 TRACKING BIAS INFORMATION IN KNOWLEDGE GRAPHS

In this section we present a model to track bias information for controversial statements in Knowledge Graphs by adding annotations about information provenance and about how supporting evidence has been identified by contributors. Specifically, we propose to give

[3] https://www.wikidata.org/wiki/Q37079


**Figure 10: Knowledge Graph entity with controversial facts.**

the option for each fact stored in a crowdsourced Knowledge Graph like Wikidata to keep information about the Search Engine used to retrieve the reference supporting the fact, the search query used to retrieve the reference, and the position in the ranked list of results in the SERP in which the reference was found.

Such additional metadata can be stored in Knowledge Graphs using existing techniques like, for example, *reification* in RDF [15]. This allows us to generate statements about statements thus making it possible to provide references for statements (e.g., like in Wikidata) and to annotate subgraphs with metadata about them (e.g., supporting URLs). While reification has been shown to generate a large volume of data if used extensively, alternative approaches to deal with the overhead introduced by reification also exist. For example, in [22] authors assign temporal validity information to facts by means of custom properties to create additional statements about statements:

- #3: DavidBeckham playsFor RealMadrid
- #3 since 2003
- #3 until 2007

In a similar way we propose to trck provenance and information about implicit bias which can be efficiently stored and retrieved from custom knowledge bases like, for example, TripleProv [23].

In the case of the Catalonia use case, we would store the following statements:

- #1: Catalonia instanceOf 'autonomous community of Spain'
- #1 supportedBy 42%
- #1 mainSourceOfEvidence en.wikipedia.org
- #1 avgRankOfEvidence 2.32
- #1 medianAgeOfContributors 32.5
- #2: Catalonia instanceOf 'sovereign state'
- #2 since 2017
- #2 supportedBy 58%
- #2 mainSourceOfEvidence independent.co.uk
- #2 avgRankOfEvidence 1.62
- #2 medianAgeOfContributors 28

## 5 DISCUSSION

Modern data models for Knowledge Graphs like the one used by Wikidata already allow for the use of qualifiers (e.g., start and end


**Figure 11: Entity card in SERP proposing to the end user alternative statements from the background Knowledge Graph with bias information.**

date of the validity of a statement like 'President of the United States') and references for statements[4]. The model we propose builds on top of existing data models and existing RDF stores and allows to embed in the Knowledge Graph information about the bias certain statements carry with them from the data creators. Such provenance information can then allow data consumers to be aware of possible existing bias in the data and to deal with it as they deem appropriate (e.g., selecting the majority answer or selecting the answer preferred by the demographics group they belong to).

As the information stored in Knowledge Graphs can be highly influenced by the population of contributors that contributed it, such information would be biased and follow up usage of the Knowledge Graphs (e.g., empowering Semantic Search systems) could lead to biased results returned to end users. Thus, rather than avoiding to store biased information in the Knowledge Graph, we believe it is a better option to keep track of bias (see Figure Figure 10 for an example) in order to better inform end users of potential bias and envision such information to be surfaced to them as, for example, shown in Figure 11 for entity cards in SERPs.

## 6 CONCLUSIONS

In this paper we discussed the need to track bias information into Knowledge Graphs. We presented the results of three crowdsourcing experiments involving 600 crowd workers that show how factual statements stored into Knowledge Graphs can be controversial. We have motivated the need for keeping track of provenance metadata for such controversial statements and shown how different populations of contributors can significantly influence the facts stored in a Knowledge Graphs.

[4] https://www.mediawiki.org/wiki/Wikibase/DataModel/Primer

Future directions for this line of research include the evaluation with end users of the benefits of surfacing bias and provenance information in applications like Semantic Search. More than this, it is important to design automatic methods for the selection of the subset of statements in a crowdsourced Knowledge Graph that would benefit most from bias metadata being preserved due, for example, to their controversial dimension.

*Acknowledgments.* This project is supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 732328.

## REFERENCES


[1] Aroyo, L., Welty, C.: Truth is a lie: Crowd truth and the seven myths of human annotation. AI Magazine 36(1), 15–24 (2015)

[2] Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., Ives, Z.: Dbpedia: A nucleus for a web of open data. The semantic web pp. 722–735 (2007)

[3] Bizer, C.: The emerging web of linked data. IEEE intelligent systems 24(5) (2009)

[4] Checco, A., Roitero, A., Maddalena, E., Mizzaro, S., Demartini, G.: Let's agree to disagree: Fixing agreement measures for crowdsourcing. In: 5th AAAI Conference on Human Computation and Crowdsourcing (HCOMP). pp. 11–20 (2017)

[5] Demartini, G., Difallah, D.E., Gadiraju, U., Catasta, M.: An introduction to hybrid human-machine information systems. Foundations and Trends in Web Science 7(1), 1–87 (2017), https://doi.org/10.1561/1800000025

[6] Difallah, D.E., Catasta, M., Demartini, G., Ipeirotis, P.G., Cudré-Mauroux, P.: The dynamics of micro-task crowdsourcing: The case of amazon mturk. In: 24th International Conference on World Wide Web. pp. 238–247. WWW '15 (2015)

[7] Eickhoff, C.: Cognitive biases in crowdsourcing. In: Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. pp. 162–170. ACM (2018)

[8] Feyisetan, O., Luczak-Rösch, M., Simperl, E., Tinati, R., Shadbolt, N.: Towards hybrid NER: A study of content and crowdsourcing-related performance factors. In: 12th European Semantic Web Conference, ESWC. pp. 525–540 (2015)

[9] Greenwald, A.G., McGhee, D.E., Schwartz, J.L.: Measuring individual differences in implicit cognition: the implicit association test. Journal of personality and social psychology 74(6), 1464 (1998)

[10] Hube, C., Fetahu, B., Gadiraju, U.: Understanding and mitigating worker biases in the crowdsourced collection of subjective judgments. In: Proceedings of the 37th Annual ACM Conference on Human Factors in Computing Systems. CHI '19 (2019)

[11] Maddalena, E., Roitero, K., Demartini, G., Mizzaro, S.: Considering assessor agreement in ir evaluation. In: Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval. pp. 75–82. ICTIR, ACM (2017)

[12] Maedche, A., Staab, S.: Ontology learning for the semantic web. IEEE Intelligent systems 16(2), 72–79 (2001)

[13] Matuszek, C., Cabral, J., Witbrock, M.J., DeOliveira, J.: An introduction to the syntax and content of cyc. In: AAAI Spring Symposium: Formalizing and Compiling Background Knowledge and Its Applications to Knowledge Representation and Question Answering. pp. 44–49 (2006)

[14] Mortensen, J.M., Musen, M.A., Noy, N.F.: Crowdsourcing the verification of relationships in biomedical ontologies. In: AMIA Annual Symposium Proceedings. vol. 2013, p. 1020. American Medical Informatics Association (2013)

[15] Nguyen, V., Bodenreider, O., Sheth, A.: Don't like rdf reification?: Making statements about statements using singleton property. In: 23rd International Conference on World Wide Web. pp. 759–770. WWW '14 (2014)

[16] Otterbacher, J.: Crowdsourcing stereotypes: Linguistic bias in metadata generated via gwap. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. pp. 1955–1964. CHI '15, ACM (2015)

[17] Roitero, K., Demartini, G., Mizzaro, S., Spina, D.: How many truth levels? six? one hundred? even more? validating truthfulness of statements via crowdsourcing. In: Proceedings of the 2nd International Workshop on Rumours and Deception in Social Media (2018)

[18] Ross, J., Irani, L., Silberman, M.S., Zaldivar, A., Tomlinson, B.: Who are the crowdworkers?: Shifting demographics in mechanical turk. In: CHI '10 Extended Abstracts on Human Factors in Computing Systems. pp. 2863–2872 (2010)

[19] Sarasua, C., Simperl, E., Noy, N.F.: Crowdmap: Crowdsourcing ontology alignment with microtasks. In: International Semantic Web Conference. pp. 525–541 (2012)

[20] Venanzi, M., Guiver, J., Kazai, G., Kohli, P., Shokouhi, M.: Community-based bayesian aggregation models for crowdsourcing. In: Proceedings of the 23rd International Conference on World Wide Web. pp. 155–164. WWW '14, ACM (2014)

[21] Vrandečić, D., Krötzsch, M.: Wikidata: a free collaborative knowledgebase. Communications of the ACM 57(10), 78–85 (2014)

[22] Wang, Y., Zhu, M., Qu, L., Spaniol, M., Weikum, G.: Timely yago: Harvesting, querying, and visualizing temporal knowledge from wikipedia. In: 13th International Conference on Extending Database Technology. pp. 697–700. EDBT (2010)

[23] Wylot, M., Cudre-Mauroux, P., Groth, P.: Tripleprov: Efficient processing of lineage queries in a native rdf store. In: Proceedings of the 23rd international conference on World wide web. pp. 455–466. ACM (2014)