Transparency Paths - Documenting the Diversity of User Perceptions

FAUSTO GIUNCHIGLIA, Department of Information Engineering and Computer Science, University of Trento, IT

STYLIANI KLEANTHOUS, Open University of Cyprus & CYENS Centre of Excellence, CY

JAHNA OTTERBACHER, Open University of Cyprus & CYENS Centre of Excellence, CY

TIM DRAWS, Web Information Systems, Department of Software Technology, TU Delft, NL

We are living in an era of global digital platforms, eco-systems of algorithmic processes that serve users worldwide. However, the increasing exposure to *diversity* online – of information and users – has led to important considerations of *bias*. A given platform, such as the *Google* search engine, may demonstrate behaviors that deviate from what users expect, or what they consider fair, relative to their own context and experiences. In this exploratory work, we put forward the notion of *transparency paths*, a process by which we document our position, choices, and perceptions when developing and/or using algorithmic platforms. We conducted a self-reflection exercise with seven researchers, who collected and analyzed two sets of images; one depicting an everyday activity, "washing hands," and a second depicting the concept of "home." Participants had to document their process and choices, and in the end, compare their work to others. Finally, participants were asked to reflect on the definitions of bias and diversity. The exercise revealed the range of perspectives and approaches taken, underscoring the need for future work that will refine the transparency paths methodology.

1 INTRODUCTION

In the last two decades, the Internet, globalization, and the emergence of "social" and global digital platforms have transformed our lives and transcended geographical and cultural borders. It is now easy to connect in virtually zero time to anybody in the world, who lives in a different culture and social context, with different habits and speaking a different language. The amount of *diversity* we have access to has increased exponentially. We are exposed daily to a seemingly unbounded amount of diversity, most of which is unexpected and represents those *unknown unknowns* which many find scary. Yet, evolution operates on a slower timescale than that of technology. We have basically the same instruments and skills to deal with diversity that our grandparents had 50 years ago. In other words, technology has provided us with increased access to diversity, but has fallen short of providing the instruments which would allow us to cope with the *social challenges* that arise with diversity.

Let us consider, for example, the global platforms of the sharing economy (e.g., *Facebook*, *Google*, *Twitter*). Their ability to connect disparate needs by individuals with spare resources provided by other individuals is the key factor in

their success and ability to scale globally. In the current sharing economy, however, a major bottleneck arises when people need to interact. For instance, interaction may take place via direct online social interactions with other people (most often unknown to them) or, much more frequently, indirectly, by reading or reusing online material (e.g., for information retrieval) developed independently by others. This is when and where our current inability to deal with the unknown diversities raises many difficulties.

Contemporary to the increased access to diversity, we have also witnessed an increased growth on bias. Bias is now pervasive on the Web [3, 16, 18, 19] and the effects of online bias are much more far reaching than those that we usually witness in the physical world, where interactions involve a smaller number of people collocated in a given context [1, 5, 6, 15, 20]. Unless a solution is found that will allow us to deal with these phenomena, diversity and bias are bound to grow together following the increase in pervasiveness of the Web. Namely, we will have all the positive effects of diversity negatively balanced by the constant growth of bias (e.g., resulting in a decrease of user trust [4]).

This paper is a first attempt to understand how individuals perceive the relationship between *diversity and bias*. The first contribution is to answer the question of whether this parallel growth is coincidental or correlated. A preliminary answer is that the two phenomena are highly correlated up to the point that, under certain contextual conditions, the presence of diversity is the primary cause for the insurgence of bias. The second contribution is to work towards a general method for dealing with the problems of diversity and bias, having established that they are interrelated. The proposed approach is to identify *the modality by which bias can emerge from the presence of diversity*. To this end, we introduce the notion of *transparency paths*, a systematic methodology and supporting annotations by which it is possible to keep track of all the steps by which diversity and bias arise when observing a certain world phenomenon.

This paper describes a pilot study, conducted within the context of a virtual Winter School on FATE (fairness, accountability, transparency, and ethics), involving researchers working in various areas of computer science. The study involved the case of *image search*. Images provide a direct representation of reality (i.e., a world event or phenomenon) as compared to language, where representations are filtered by the observer. Thus, images allow the observer to make evident their underlying predisposition that later might lead to the perception of diversity and bias. As will be described in detail, participants were asked to collect images relating to two commonly experienced world phenomena, documenting each step of the way their process and their choices made. Data were aggregated and participants were then able to compare their work to that of others. Finally, through responses to a questionnaire, participants were asked to reflect on the relationship between diversity and bias, as well as on the process followed in the study.

The structure of the paper is as follows. In Section 2, we examine some definitions for the concepts of diversity and bias and introduce the notion of *transparency paths*. In Section 3, we detail the methodology followed in our exploratory study. The results from the study are presented in Section 4, while in Section 5 we present the conclusions, arguing in favor of the further development of the transparency paths methodology.

2　DIVERSITY, BIAS AND TRANSPARENCY PATHS

Merriam Webster defines *Diversity* as *"...the condition of having or being composed of differing elements."* However, everything is in fact different from everything else. Genetic diversity allows species to adapt to changes in the environment, production diversity allows economies to adapt to changes in market dynamics, social and cultural diversity fuel progress in the society. Diversity is the key distinguishing feature of life: there will never be two identical moments, two identical places, or two identical individuals. This pervasiveness of diversity is on the basis of the pervasiveness of the search for *similarity*. In a world driven by diversity, learning from cases that are similar to the one at hand, might be the best thing one can do. The relation between similarity and diversity is quite subtle and is

not mutually exclusive. According to Aristotle, the meaning of words, i.e., the objects they denote, should be given in two steps. In the first step this meaning should be defined by the subset of objects that are similar, the so-called *Genus*, while, in a second step, this meaning should be further refined by identifying, within Genus, what makes them different, i.e., their *Differentia* [14]. Thus, for instance, a *triangle* is a *planar figure* (i.e., the Genus that differentiates planar figures from, e.g., solid figures, thus taking the former as similar, if compared to the latter) with *three angles* (i.e., the Differentia that differentiates triangles from, e.g., rectangles).

The key observation is that similarity and diversity are not opposite to each other and are *always relative* to a reference set of assumptions and objects. In the field of Lexical Semantics, this fact has led to a huge amount of work which has allowed to define the meaning of words inside a language [7, 12] and across languages [8]. Recently, in the area of data and knowledge integration, this observation led to the study of the diversity of properties used for describing objects [9] and of the values that these properties take when described by different persons [10]. This previous line of work is the basis on which we advocate a study of bias which should take this notion as being relative and, therefore not absolute and consequently correlated to that of diversity. The majority of the work that is accused of being biased is based on the use of Machine Learning, a technique that exploits the intrinsic similarity of a set of training examples to learn what to do next (see, e.g., [2]). In this context, it is quite implausible to assume that, independently of the technique used as well as of the set of examples used, Machine Learning will generate results that are not unexpected and that many would classify as being biased.

But what is bias? According to Wikipedia, *"... bias is a disproportionate weight in favor of, or against, an idea or thing, usually in a way that is closed-minded, prejudicial, or unfair. People may develop biases for or against an individual, a group, or a belief."* [1]. In its simplicity, the above quote raises a few crucial questions. Who decides the reference point from which we can establish the presence or absence of bias? Who selects the reference "right" idea or thing? Who decides what is disproportionate? It should be immediate to see that the above definition implicitly assumes a priori, a given definition of diversity as well as of similarity.

By taking a closer look at the above definition of bias it is easy to identify four main components:

(1) the existence of an observer who defines a reference point of view and of a set of evaluation criteria as well as measures for those criteria,
(2) an observation of a certain phenomenon for some external motivation,
(3) an evaluation of what is being observed, with respect to the chosen viewpoint,
(4) a value judgement on the results of this evaluation.

To clarify the above characterization of bias, let us consider the following example. Assume, for instance, that a South European Vegan developer trains an algorithm on eating behaviour on a dataset coming from Mongolia, which is well known to have a meat intense cuisine. Most likely she will think that the resulting program is biased, with her value judgement being negative. Now suppose that we have the opposite situation, namely a Mongolian developer training the algorithm on a Vegan dataset. We will most likely have a similar situation, with the twist that the resulting algorithms will be harder to use, this being largely a consequence of the cold weather of Mongolia and the consequent difficulty of growing vegetables. The Mongolian developer might not see the point of using the Vegan dataset, maybe thinking why in the world someone could think of using the Vegan dataset.

The example above should highlight the fact that both the South European Vegan developer and the Mongolian meat-eater developer are well-minded, both with good motivations and good intentions, and that a most profitable

---

[1]https://en.wikipedia.org/wiki/Bias

solution would be to make them understand each other's perspective. But how can we achieve this mutual understanding algorithmically? We build this in two steps. The first is to exploit the definitions of bias and diversity provided in [11]. In this framework we have the following:

(1) We define a *diversity space* as a Cartesian space defined by a set of diversity dimensions and an origin that are used to describe the phenomenon under consideration. In the above example, these diversity dimensions would be the type of food (e.g., vegetables, meat, noodle, etc.) and various other characteristics. In this space one specific point, let us take it to be the origin, could be defined as being what is expected. Even assuming the same coordinates, the origins of the South-European and of the Mongolian would be quite far apart.

(2) Within this diversity space, a certain volume could be defined as being the *bias space*, in other words any point in that space will be defined as being biased.

(3) Any observed phenomenon would be positioned in this diversity space and its distance from what is expected could be computed according to some pre-defined distance metrics. This distance would be a measure of deviation from the expected behavior.

(4) Any observed phenomenon could then be defined as being (un)biased depending on whether it is (out)inside the bias space.

The first observation is that the process above formalizes the process by which bias is generated, as described above. The second is that, while the definition of the diversity space can be somehow motivated by the need to describe what is relevant, the definition of the bias space is motivated by some a priori motivated value judgement. In turn, this value judgement may have strong or weak motivations, can be controversial or not, and so on. But this is irrelevant to this discussion. One could also observe that the definition of the diversity space has itself a major influence on the possible recognition of bias. This is correct and the food example above is a clear example of this. What is bias for a vegan is normal for a meat eater and vice-versa. But how to deal with this problem is outside the goals of this paper.

The second and last step is to build a *transparency path* as the sequence of steps which lead to the identification of diversity and also of bias, as described above. We talk of *diversity transparency path* in the first case and of *bias transparency path* in the second case. These two paths would be nothing else that the sequence of the four decisions listed above, suitably annotated and commented. The construction of these paths provides a major added value in that it provides *transparency* to the process by which certain results have been produced. This transparency is clearly not enough for a final assessment of diversity and bias; in the end, all these evaluations are based on subjective and highly contextual decisions. Still, at least, the construction of transparency paths seem a first crucial step which allows putting any further discussion, assessment, and technology development, on a much more solid foundation.

3 METHODOLOGY

A group of researchers (four female, two male), who work in various areas of computer science, and who were taking part in a workshop on Fairness, Accountability, Transparency and Ethics in AI, took part in this exploratory study. The cultural and linguistic backgrounds of the researchers were varied, with two coming from the Americas (Brazil, USA) and five coming from various regions of Europe (Spain, Italy, Greece, Cyprus, Germany).

The study aimed at developing a methodology for *Transparency Paths*, a means by which each individual can document his or her own context and perspective, while using an algorithmic system. In this case, we developed an exercise having to do with visual depictions of common phenomena / activities. The exercise was meant to stimulate reflection on how global digital platforms such as search engines present a "picture" of the world to their users.

Participants were asked to work individually on a set of two tasks. The facilitators of the workshop were responsible for collecting the data generated through the tasks, and then compiling the data anonymously so that participants could compare their responses to those of others. Finally, through a questionnaire, participants were asked to reflect on issues of diversity and bias, based on the current exercise.

3.1 Task description

Participants were asked to represent each of two phenomena / concepts that are common to people across the world - 1) the activity of *Washing Hands* and 2) the concept of *Home*. Specifically, participants needed to represent each concept/activity by building a collection of visual information artifacts; they were asked to collect 20 images for each concept/activity from two different sources: i) *Dollar Street*, [2] a website featuring a manually curated collection of "photos as data to kill country stereotypes" (see Figure 1), and ii) the *Google* search engine image results. Note that the first source, *Dollar Street*, is not algorithmically mediated, being a hand curated collection of images, while image retrieval using *Google* search is highly mediated by opaque algorithmic processes.
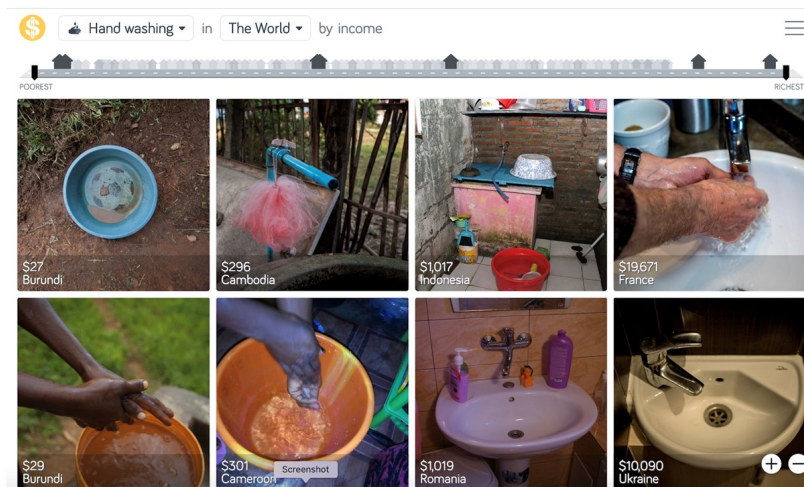

Fig. 1. *Dollar Street* results for Washing Hands

Participants received the following instructions and performed the following tasks:

Step 1 - Construct a dataset (20 images each) of each of the two phenomena and data sources. You may select any images you like, but you must document the criteria you used and explain them.

Step 2 - Using the 40 images for each phenomenon, define a set of key aspects/properties being emphasized in the selected images depicting "washing hands" or "home". Because properties are associated with / describe entities, you need to record the following: i) the entity referenced (e.g., hands), ii) the property (e.g., skin color), iii) property values (i.e., all possible values for the property), iv) dataset for which this property is relevant (or both).

Step 3 - Once the data were compiled, anonymized, and shared with all team members, each team member completed a questionnaire as a means to reflect on the above process as well as one's responses as compared to those of others.

---

[2] https://www.gapminder.org/dollar-street

3.2 Questionnaire

A short, anonymous questionnaire was administered via *Google Forms*. The questionnaire asked the participants to reflect on the following:

(1) The data collection process they followed for each world phenomenon, and how their approach differed from those of others (*Did your data collection process (i.e., Step 1) differ from that of the others, for the phenomenon of "washing hands"? Please explain.*);

(2) Their characterization compared to what others used (*How did your characterization of the images (i.e, Step 2) differ from that of the others, for the phenomenon of "washing hands" / "home'?*);

(3) Whether, in retrospect, they had overlooked some of the facets that characterize the "washing hands" or "home" phenomena, and why they think that happened;

(4) Their self-reflection of whether the datasets/descriptions of datasets they/others have created were biased;

(5) Describing their understanding of the entire process from diversity of bias, explored in this exercise;

(6) Providing their own definition of diversity and bias;

(7) Whether they believe that the inherent diversity in a dataset used to represent a phenomenon in the world (e.g, "washing hands") can become a source of bias.

All six participants responded to the questionnaire. To analyze their responses, we followed a thematic analysis approach [17]. A researcher went through all the anonymized data and coded the participants' responses revealing commonalities and/or differences in the approaches they followed, then shared the results with the participants for their feedback.

4 RESULTS

In analyzing the results, we focused on: i) the data collection process followed by the participant, ii) their approach in characterizing the data and how they would define diversity and bias.

4.1 Data Collection Process

For both datasets ("Washing Hands" and "Home"), five of six participants thought that their approach in collecting the data was different than the other participants mentioning in their responses the *strategy* and the *criteria* they have used to do that. Regarding the *strategy*, they mentioned two approaches. Two out of six collected the first 20 images presented to them, in either the *Dollar Street* collection or the search results page obtained from a *Google* search, describing their strategy as the "typical search process" they use when seeking information sources. In contrast, the remaining participants used a strategy that explicitly focused on constructing a diverse dataset.

In addition to the global strategy they used to seek out representative images, participants referred to some *criteria* they have used. For "Washing Hands", two out of six participants mentioned *socioeconomic aspects* and *demographic characteristics* as criteria for creating a *diverse* dataset, while others focused on *particular objects* depicted on the images (e.g., "image must contain soap"). For developing the "Home" dataset, two out of six focused on creating a *diverse* dataset using *criteria* such us *socioeconomic aspects* and *demographic characteristics*.

4.2 Image Characterization

In characterizing the images collected, participants followed the instructions in **Step 2** above. For both datasets, three of the participants mentioned that their approach was different compared to others, in that they either paid attention

to more or less *detail* on the images when they were describing them. Three of the participants mentioned that their characterizations missed some of the *attributes* that others noted, while five out of six explicitly stated that they overlooked some facets when characterizing the "Washing Hands" dataset and all of them stated they overlooked some facets when characterizing the "Home" dataset. Commenting on why they think they might have overlooked those, two participants commented on the *collective nature* of such tasks, mentioning that this should have not been an individual exercise but rather a collaborative one so we could collectively characterize the images. Four participants referred to how each individual *perceives* the world phenomena in a different way according to their background and experiences, while two participants admitted performing the task in a rather mechanical approach focusing only on the topic rather than the attributes and/or objects depicted in the images.

4.3   Bias and Diversity

Participants were asked to self-evaluate, having compared their own work to that of others. In particular, they indicated on a five-point Likert scale the extent to which the datasets that they created representing "Washing Hands" / "Home" were *biased*. For the "Washing Hands" dataset two participants selected option 1 and 2 indicating their datasets were not biased, three selected option 3 (unsure/neutral) and only one participant selected option 5 indicating that their dataset was highly biased. For the "Home" dataset, three participants selected option 3 (unsure/neutral) and three participants selected option 5 indicating that their dataset was biased.

Next, participants were asked to evaluate others' work. When asked to indicate the extent to which *Others' descriptions of the datasets analyzed for either "Washing Hands" / "Home" were biased*, the majority (four out of six) selected option 3 (unsure/neutral) for the "Washing Hands" dataset, while two selected option 4 and 5 indicating that others' description were biased. Similarly, for the "Home" dataset, three of out of six selected option 3, while two selected options 4 and 5. Only one participant selected option 1 indicating that the descriptions provided by other participants were not biased for the "Home" dataset.

When asked to define the terms *Diversity* and *Bias*, participants looked into different aspects. "*Diversity* has to do with various viewpoints..." as pointed out by one participant. Difference in *perception* and various *viewpoints* were brought up by two other participants. Three participants referred to *equal representation/distribution* of various cases in a dataset as part of a diversity definition (e.g., "Diversity is when all different cases are represented..."). *Complementarity* was mentioned by one participant as being an important part of the diversity definition - "Diversity among men is what makes them set up companies to achieve unattainable goals on their own; in this sense it is synonymous with complementarity." Finally, three participants made a reference to *culture* being an integral part of diversity, in the context of having various cultural backgrounds represented. It is worth mentioning that *culture* and *equal representation/distribution* commonly appeared in the definitions provided by two participants.

In defining *Bias*, two participants mentioned the relativity of bias to diversity (e.g., "Degree of divergence from diversity."). Two other participants mentioned that defined bias as the *deviation from ground truth* - "I prefer a statistical definition of bias - it is a measured deviation from the current point." Similar to defining *Diversity*, three participants mentioned the *Equal representation of cases/reality* as a way of defining bias - "Bias exists when only one end is represented heavily and probably an observer could suspect who is behind this representation."

4.4   Can Diversity Become a Source of Bias in a Dataset?

To better understand how the participants perceived this exercise overall, and to prompt them to reflect on the existence/lack of diversity in the datasets created individually, we asked them *"Can the inherent diversity in a dataset*

used to represent a phenomenon in the world (e.g, "Washing Hands") become a source of bias? Explain your answer." All participants believed that the answer is yes, referring to the need for *awareness of the alternative* views, and bias being the *absence of diversity*. In the same vein, two participants also referred to the unequal *distribution of representations* of cases in a dataset. Interestingly, only one person made a reference to the definition of bias being *context dependent*.

4.5 Overall Reflection in this Process

Finally, we asked participants to reflect on the overall process of this exercise. With respect to constructing their own datasets from two different sources on two different phenomena, participants thought that we should not have used an individual process. They stressed the importance of different viewpoints when such tasks are taking place - *"we have approached the concept of bias on an individual basis and that in my opinion is a mistake. Bias is never individual, even if it is executed individually, our choices are socially constructed (our focus from a professional background, our geographical location, etc) we cannot understand our perception of the world without taking into account our place in the world.".*

Another important aspect relates to the assumptions that we as humans make - *"I realized that from the beginning we decided on how to select photos based on some assumptions. Then when we were characterizing the images we "saw" only those characteristics that we were looking for and not the complete picture. This is how probably biased datasets are built.".* A different reflection is that the exercise could be a task for understanding how bias and diversity might be transferred in AI enabled systems through datasets - *"I think of this exercise as a self-reflection exercise. It enabled us to reflect on concepts like diversity and bias, and also become more aware of the fact that any dataset aimed at visual depictions of world phenomena that we would create would somehow be biased.".*

5 CONCLUSIONS

Without a doubt, *bias* in data, information, and algorithms has become a key concern for those working in AI and data science [13]. In this work, we related the notion of bias, to that of diversity, as we believe that the increasing trend towards global digital platforms has brought these issues to the forefront. We therefore operationalized a notion of bias as consisting of diversity (in information, perspectives, and processes) along with an observer's negative evaluation of this diversity. To this end, we advocated for the need for *transparency paths*, a generalized methodology by which we might keep track of all of the steps by which diversity and bias arise when observing a certain world phenomenon (e.g., while using a search engine to retrieve images depicting the concept of "home").

We presented an exploratory study – a first attempt at developing a method for transparency paths. In particular, we carried out an exercise with fellow researchers, who are interested in / working on FATE, as a first step towards developing our process. Obviously, we need to refine our methods and perform the exercise with a range of different groups and stakeholders, in order to develop a robust methodology for transparency paths. In the longer term, we are hopeful that through the use of transparency paths, this can be a means to mitigate bias and misunderstanding, since one will be able to document and communicate their context and perspective to others.

REFERENCES

[1] Ahmed Allam, Peter Johannes Schulz, and Kent Nakamoto. 2014. The impact of search engine selection and sorting criteria on vaccination beliefs and attitudes:Two experiments manipulating google output. *Journal of Medical Internet Research* 16, 4 (2014), e100. https://doi.org/10.2196/jmir.2642

[2] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. Machine bias. *ProPublica, May* 23 (2016), 2016.

[3] Ricardo Baeza-Yates. 2018. Bias on the web. *Commun. ACM* 61, 6 (2018), 54–61. https://doi.org/10.1145/3209581

[4] Tim Draws, Zoltán Szlávik, Benjamin Timmermans, Nava Tintarev, Kush R. Varshney, and Michael Hind. 2021. Disparate Impact Diminishes Consumer Trust Even for Advantaged Users. In *International Conference on Persuasive Technology*. Springer, Cham. https://arxiv.org/abs/2101.12715

[5] Tim Draws, Nava Tintarev, Ujwal Gadiraju, Alessandro Bozzon, and Benjamin Timmermans. 2021. This Is Not What We Ordered : Exploring Why Biased Search Result Rankings Affect User Attitudes on Debated Topics. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Virtual Event, Canada) *(SIGIR '21)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3404835.3462851

[6] Robert Epstein and Ronald E. Robertson. 2015. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences of the United States of America* 112, 33 (2015), E4512–E4521. https://doi.org/10.1073/pnas.1419828112

[7] F. Giunchiglia, Khuyagbaatar Batsuren, and Gabor Bella. 2017. Understanding and Exploiting Language Diversity.. In *IJCAI*. 4009–4017.

[8] F. Giunchiglia, K. Batsuren, and A. A. Freihat. 2018. One World–Seven Thousand Languages. In *Proceedings 19th International Conference on Computational Linguistics and Intelligent Text Processing*. 18–24.

[9] Fausto Giunchiglia and Mattia Fumagalli. 2020. Entity Type Recognition – dealing with the Diversity of Knowledge. In *Knowledge Representation Conference (KR)*. Rhodes, Greece.

[10] F. Giunchiglia, V. Maltese, and B. Dutta. 2012. Domains and context: first steps towards managing diversity in knowledge. *Journal of Web Semantics, special issue on Reasoning with Context in the Semantic Web* (2012), 53–63.

[11] Fausto Giunchiglia, Jahna Otterbacher, Styliani Kleanthous, Khuyagbaatar Batsuren, Veronika Bogina, Avital Shulner Tal, and Tsvi Kuflik. 2021. Diversity, Bias and Algorithmic Transparency. *arXiv preprint arXiv:2104.05658* (2021).

[12] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller. 1990. Introduction to WordNet: An on-line lexical database. *International journal of lexicography* 3, 4 (1990), 235–244.

[13] Eirini Ntoutsi, Pavlos Fafalios, Ujwal Gadiraju, Vasileios Iosifidis, Wolfgang Nejdl, Maria Esther Vidal, Salvatore Ruggieri, Franco Turini, Symeon Papadopoulos, Emmanouil Krasanakis, Ioannis Kompatsiaris, Katharina Kinder-Kurlanda, Claudia Wagner, Fariba Karimi, Miriam Fernandez, Harith Alani, Bettina Berendt, Tina Kruegel, Christian Heinze, Klaus Broelemann, Gjergji Kasneci, Thanassis Tiropanis, and Steffen Staab. 2020. Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10, 3 (2020), 1–14. https://doi.org/10.1002/widm.1356

[14] William Thomas Parry and Edward Hacker. 1991. *Aristotelian Logic*. State University of New York Press.

[15] Frances A. Pogacar, Amira Ghenai, Mark D. Smucker, and Charles L.A. Clarke. 2017. The Positive and Negative Influence of Search Results on People's Decisions about the Efficacy of Medical Treatments. In *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval (ICTIR '17)*. Association for Computing Machinery, New York, NY, USA, 209–216. https://doi.org/10.1145/3121050.3121074

[16] Cornelius Puschmann. 2018. Beyond the Bubble: Assessing the Diversity of Political Search Results. *Digital Journalism* (2018), 1–20. https://doi.org/10.1080/21670811.2018.1539626

[17] David R Thomas. 2006. A general inductive approach for analyzing qualitative evaluation data. *American journal of evaluation* 27, 2 (2006), 237–246.

[18] Ryen W. White. 2013. Beliefs and biases in web search. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '13)*. Association for Computing Machinery, New York, NY, USA, 3–12. https://doi.org/10.1145/2484028.2484053

[19] Ryen W. White and Ahmed Hassan. 2014. Content bias in online health search. *ACM Transactions on the Web* 8, 4 (2014), 1–33. https://doi.org/10.1145/2663355

[20] Ryen W. White and Eric Horvitz. 2014. Belief dynamics in web search. *Journal of the Association for Information Science and Technology* 65, 11 (2014), 2165–2178. https://doi.org/10.1002/asi.23128