Supporting Responsible Data and Algorithmic Practices in The News Media

Dilruba Showkat

Northeastern University, Boston, MA, USA

Abstract

The journalism discipline has become more data and algorithm-driven than ever before. While the need for transparent algorithmic practices in journalism is widely known, less is known about how to go about doing that in practice. As a result, journalists often face challenges associated with Replicability and Reproducibility (R&R) tasks both within the team and also when checking others' data work. Journalists can be facilitated to practice transparency by providing explicit information about the sources and methodologies – by being responsible dataset and algorithm users both within and outside of the organization. In this work, as a case study, I present a very first responsible dataset and responsible algorithm practices specifically crafted for the domain of journalism, as a step towards motivating and supporting transparent algorithmic practices using a question-driven documentation technique. The outcome of this study is open to critique, adoption, adaptation, and future exploration.

Keywords

responsible journalism, transparent journalism, replicability and reproducibility

1. Introduction

Algorithms are widely used in a variety of application domains ranging from the public and private sector, healthcare, automated hiring systems, to the criminal justice system. Sometimes, these algorithms inherit, reproduce, or even enhance biases against the marginalized population, causing a lack of users' trust in these systems [1, 2, 3, 4]. Moreover, *"models are opinions embedded in mathematics"* [4, p.27], they enable us to focus on only the outcome, predictor variables, and validation data while avoiding anything that promotes an understanding of situations or context [3]. This is problematic, as a result, there is a growing interest in the design of transparent algorithmic systems to make the algorithmic decision making and context more accessible. In a similar vein, there is an increasing focus to produce replicable and reproducible work in Machine Learning (ML) research, data science, and in the healthcare domain among others [5, 6, 7, 8, 9]. Reproducibility also plays a critical role in Journalism (e.g., provenance) [10].

Likewise, the demand for transparent journalism has existed for a long time [10, 11], where journalists are expected to describe what data sources they have used, revealing subjects and data analysis methodology, for verification and reproducibility purposes. *"The essence of journalism is the discipline of verification"* [10, p.79]. There are several limitations which often makes it impractical to implement transparent journalism in practice, for instance, misuse of transparent technology through gaming or manipulation [12], information overload, and others (e.g., cost, presentation). Furthermore, fact-checkers tools (such as politifact) [13, 14] are not informative enough to support journalists' replication tasks – in terms of data and algorithmic analysis. Replicability and Reproducibility (R&R) also plays a significant role in journalism to make sure that journalistic processes are free from biases [15, 16] and the data they put out in the world is accurate – since *"journalism's first obligation is to the truth"* [17, 10]. There is limited research in this space that supports reproducibility tasks within the journalism team. Thus, in this research, building on prior work I will provide a set of question-driven documentation guideline practice to support responsible dataset and algorithm use within journalism team. This work provides implications for making the news story related information (with caution) also available to the public, and implications for related technology design intervention in the journalism context.

2. Related Work

2.1. Data and Algorithmic Practices in Journalism

As data becomes readily available news organizations are increasingly becoming more data-driven than ever

Joint Proceedings of the ACM IUI 2022 Workshops, March 2022, Helsinki, Finland
showkat.d@northeastern.edu (D. Showkat)

before [13, 18]. Journalists work with a variety of datasets and data types (e.g., text, tables, numerical, categorical data [19, 20]) in news storytelling. They use public datasets such as Medicare or Housing datasets [19], they also collect data from other sources such as through interviews, surveys, public websites using various tools and APIs (e.g., ArcGIS). Using public record requests is also very common among them [10, 21]. While some stories are based on a single dataset, others are based on multiple datasets. Journalists can chase original stories or they might also choose to build off of others' work.

When it comes to algorithms, journalists apply a wide range of algorithms; from simple statistical tests (e.g, ANOVA, t-test) to advanced Machine Learning (ML) algorithms (e.g, regression, classification, unsupervised ML) for data analysis [19, 20]. Furthermore, recent work also shows that news are often co-produced using automated tools (e.g., that uses natural language generation and large models [22]) alongside with humans [23, 24], especially to improve efficiency and production. Needless to mention, *"transparency is key. Being able to explain how the stories are created is relevant both in-house and audiences"* [23, p.7]. Previous work also suggests that journalists rely on outlier detection for story idea generation [25, 26, 27, 28], others also report on applying simple spreadsheets manipulation for similar tasks. Regardless of the data analysis methodology or algorithms applied, journalists often perform verification; that is, checking others' data work (verifying either teammate and/or data/charts published by other news organizations) to ensure correctness [29, 19].

Verification enables the journalists to ensure that the data they put out in the world is accurate [10, 30]. Verification often depends on the journalists to remember things, such as the operation (e.g., min, sum) performed (asking a teammate), and the journalists have no way of doing it in a way that is reproducible (for others, even for themselves after a while) without clear documentation [6, 31]. Verification is also challenging even within a team, since people often forget to document methods or how they have arrived at a particular result. Similar to the field of data science (and other related areas [32, 33, 34, 35]) the lack of proper documentation is a common problem in code replication tasks for journalists [20].

Even though journalists makes wide use of data and algorithms in their day-to-day news production, very limited work explored documentation techniques to support journalists' data and algorithmic practices transparent [20, 19, 36]. Previous work also showed close resemblance (also through systematic empirical investigation [28, 27]) between data science and data journalism [37, 20, 38, 39] work practices [34]. Therefore, this work will take inspiration from various explainable methods available in data science and other related areas to propose a responsible data and algorithm practice for journalism – that will improve effective team communication and support transparent journalism.

2.2. Current Trend Towards Explanation

Previous study shows that explaining how/what/why aspects of facebook newsfeed algorithms enhance users awareness of how the system works in the context of social media applications [40]. Prior research also examined explainability in specific domains [41, 42]. For example, Liao et al. [42] applied a question-driven method to facilitate explainable AI user experience (XAI UX) in (adverse) healthcare domain. Existing work also studied fairness and transparency in recommender systems [43]. Others explored the socio-organizational context into explainability [44]. Needless to mention the vast amount of technical work that exists to support technical expert users (e.g., ML engineer, data scientist) explainability needs for explaining black box and white box models [45, 46, 47, 48]. As evidence suggests, less attention has been paid to support transparency and explainability needs in the journalism context. Although journalists apply wide range of data and algorithmic tools in news production [28], applying existing explainable AI (XAI) techniques (even though sophisticated, e.g., [45]) may require prior knowledge in ML or may not be easily adaptable [49] in the news storytelling context. Additionally, in case of diverse teams with differing technical skills (ML users vs. SQL users) using these tools across the team might require extra learning support. As a result, perhaps less technical approach can be suitable for R&R needs in journalism.

Inspired by related works both in data science (documentation based approaches) [50, 51, 42, 44, 52, 1, 53] and transparent data journalism [12, 10, 54, 20, 37, 19, 13], this work will apply a qualitative question-driven documentation technique to support [42] algorithmic transparency in journalism. This approach will require journalists to provide specific data and algorithmic details about the news stories by specifying what/why/how/who/when information at different levels (e.g., individual, organizational, team) of journalistic decision making process to provide richer context [1, 55, 56].

3. Transparent Data and Algorithmic Practices for Journalism

This work will facilitate journalists to properly document, contextualize data and algorithmic decision making in news storytelling, to support the practice of algorithmic transparency. This was achieved through an extensive review of relevant prior work in journalism, data science, and other related areas.

At a very high methodological level, first, factors relevant to the dataset/algorithms use in the journalism domain was categorized using content analysis [57] after synthesizing across prior work (similar to the factors listed in [12]), second, those factors were translated into question-driven explanation (e.g., How, What, Why) following prior work such as [42] and others [40, 44]. Specific details of the methods and processes are provided below.

3.1. Responsible Dataset Practice

Following is the description of the methods that were used to derive the very first question-driven responsible dataset guidelines for journalism (see Figure 1 for detail).

3.1.1. Methods for Responsible Dataset Guideline

The proposed responsible dataset use guideline was heavily inspired by and built upon previous work described in [50, 58, 51, 12, 10, 42], and adapted specifically to be used in journalism. More precisely, Bender and Friedman [51] proposed data statements for text data (though it can be applied more broadly) to alleviate bias and exclusion against certain groups of people in Natural Language Processing (NLP) technology. Gebru et al. [50] also developed datasheets for datasets – a documentation practice to enable accountability and transparency among dataset creators and consumers in the ML community. Diakopoulos and Koliska [12] proposed several factors important for achieving algorithmic transparency in the news media. In this work, I have adapted, refined, and integrated these data documentation practices for journalism. The final prototype is shown in Figure 1, and the specific feature selection criteria are described below:

Major Categories: Consistent with prior work described in Gebru et al. [50], journalists are required to document information for each of the major categories (Blue text in Figure 1): *Motivation*, *Composition*, *Collection*, *Preprocessing*, *Uses*, *Distribution*, and *Maintenance*. To ensure exhaustiveness and thorough characterization of the datasets (e.g., campaign finance, crime investigation [15]), factors in each of the aforementioned categories are further updated based on work of transparent journalism [12, 10, 54, 59] and data science [51, 42, 44, 60] due to their relatedness in data work practices [37, 19, 38, 20]. Furthermore, factors related to data reported in Diakopoulos and Koliska [12] are now carefully incorporated in each of the categories where they logically make sense. For simplicity, I show factors related to only two major categories as follows:

- *Composition* included the following factors: attributes/feature definitions/description, labeled/unlabeled data, data format (e.g., mp4, csv), sample size, missing data/completeness, data category (healthcare), data language (en-us), train/test split, raw vs cleaned data, errors/redundancy, describe sensitive/anonymous/ground truth data
- *Preprocessing* included the following factors: which data was discarded? why? tools used or done manually? Manual/automated labeling, process? annotator/curator demographics (race, class, gender), data transformation, bias handling.

And these factors were then carefully converted into explanation questions. Dataset characterization questions for each of these categories were directly incorporated into the guideline.

Different Journalism Roles and Demographics: Following the work from Bender and Friedman [51], the proposed guideline also enforced journalists to document important demographics (e.g., age, gender, class) features for different journalism roles (e.g., data annotator, speaker, data curator, data collector, scripter, editor, data analyst, presenter, director) to provide transparency against inadvertent biases. These roles' definitions are informed and combined from prior research in [51, 20, 61] to cover a broad range of journalism roles. Some of these roles may have overlapping (data analysis) functions and responsibilities across different organizations [28].

Dataset Explanation in the News Storytelling: In the proposed guideline, journalists should provide context for any dataset used by documenting Who, What, When, Why, and How [40, 44, 1] related questions. For example, journalists were asked to provide context associated with a particular dataset in the *Motivation* category. Together with demographic information across different journalism roles and subjects, it is easy to demystify "WHY" a certain dataset was used in a story. This can also provide an indication of any pre-existing biases that have gone unnoticed. The individual/organization/team, "WHO" worked on the story can be found by combining information from *Motivation*, *Collection*, *Preprocessing*, *Uses*, and *Maintenance* categories. "WHAT" aspect or feature description and other related information for any dataset are covered in the *Composition*, *Collection*, and *Preprocessing* categories. Similarly, "WHEN" information is tracked through *Maintenance*. "HOW" aspect of a dataset is included in the *Collection*, and *Preprocessing* categories. Please note that How, What, Why, Who, When characteristic aspects of a dataset in these categories may not be exclusive, however, they provide all the factors necessary (to the best of my knowledge) for responsible R&R dataset practice. These pieces of information collectively provide sufficient context and insights from individual and organizational decision making perspectives in the news storytelling [44, 1].

The proposed responsible dataset prototype (Figure 1) is dataset type (e.g., healthcare, finance, housing) agnostic, meaning that the journalists could describe any dataset types used in a story with the help of this prototype. The journalists must also conform to the privacy and anonymity of their news sources such as anonymous data sources [10, 59]. It is also important to note that all personal demographics should be published only after receiving user consent [54]. Describing and characterizing algorithmic information together with datasets will further facilitate journalists' verification [10] and R&R needs [6, 51, 62], discussed below.

3.2. Responsible Algorithmic Use Practice

The methods used for developing the responsible algorithm use guideline to support journalists verification need is provided below.

3.2.1. Methods for Responsible Algorithm Use Guideline

Previously, Mitchell et al. [52] proposed model cards for explaining Machine Learning (ML) models. The very initial prototype of responsible algorithm use for journalism is designed based on taking inspiration from this and other similar works described in [12, 42, 10, 20, 54, 52]. The final responsible algorithm use guideline is shown in Figure 2.

Major Categories: The information for the responsible algorithm/model use was organized in the following categories (Blue text in Figure 2): *Model/Algorithm Used*, *Parametes/Features*, *Tools/Editor*, *Programming Language and Code*, *Hardware*, *Verification*, *Story Narrative Related.* These categories were carefully assembled and informed by previous research in such a way that it covers all the algorithm related details needed for the journalist replication task without being redundant [12, 54, 10].

Algorithmic Explanation in the News Storytelling: Explanation regarding algorithm use is required to be documented in the aforementioned categories, for example, "WHAT" model/algorithm was used should be documented in the *Model/Algorithm Used* category; "WHAT" parameters were chosen and "WHY" should be documented in the *Parameters/Features* category; "WHO" wrote the code, including code/data verification related information should be described in *Programming Language and Code* and in *Verification* category. These features cover specific information to allow journalists to replicate data analysis done by others (even for themselves for later reference), to make sure that journalist's (and their teams) does not have to reproduce code blindly when checking existing data work. Factors related to news story was included in the *Story Narrative Related* informed by the work in Kovach and Rosenstiel [10] consists of specific story related facts such as quotes, names, date-time information. All these factors collectively enable journalists to verify facts/numbers when an error goes unnoticed after publication, by thorough and careful documentation throughout the life-cycle of a story [28, 20].

In the above paragraphs, I described the methods for responsible datasets and algorithm use guidelines in the context of journalism. Responsible dataset practice has the ability to prevent or reveal unforeseen biases (e.g., pre-existing, emergent) in journalistic data work practices. Journalists (with caution and if they are willing) can take certain level of accountability in their dataset use and attain users trust through responsible dataset and algorithmic practices (with caution by revealing what they know and how they know it).

4. Conclusion and Future Work

As journalists become more reliant on data and algorithms, it is important that they become responsible dataset and algorithm users. Therefore, this work proposed a question-driven responsible datasets and algorithm documentation guideline to support journalists' replicability and reproducibility (R&R) needs – as a way to facilitate transparent algorithmic practices in the news media. The proposed guideline requires journalists to document and/or summarize datasets and al-

Motivation	1. Why was this dataset used/created/collected? 2. Who was the dataset creator (news team, internal/external collaboration organization, reporter information) or collector? 3. Was there any funding source? then provide relevant information. 4. What was the inspiration for this story? 5. What values (e.g., privacy, trust) are protected/violated and for whom (at the organizational/individual level)? 6. Who could be impacted by the dataset misuse?
Composition	1. Document dataset attributes/feature information. 2. Was the data labeled /unlabeled? 3. What was the data format (mp4, CSV)? What was the dataset language (en-us)? 4. What is the dataset/sample size? Is the dataset train/test split available? 5. Are there any missing values or data points? How was the case of missing values resolved? 6. What was the data category (e.g., healthcare)? 7. Was the data raw/cleaned data? any errors/redundancy in the data? 8. Describe specifics for sensitive/anonymous/ground truth data.
Collection	1. How the data was collected (interview, scraping with API)? 2. What were data collector/speaker demographics (e.g., age, class, race, gender). 3. What software/hardware tools were used (if any)? 4. Please provide consent and recruitment (site) information for the subjects involved. 5. How many sources were interviewed? Interview language. 6. How is anonymous source privacy protected? Sampling method (probabilistic/random).
Preprocessing	1. Which data (point) was discarded? why? 2. Data curator demographics? 3. Any tools used or were done manually? why? 4. Manual or automated labeling? What was the labeling process (for ground truth data)? 5. What is annotator demographics (e.g., race, class, gender)? 6. What data transformation was applied? 7. How were biases handled?
Uses	1. Were the dataset used in the past by other reporters/news organizations (story link, use context)? 2. Are there any limitations in the dataset? 3. Any dataset consumer information (who can/cannot use this dataset)? 4. How the dataset should/should not be used? 5. Any stakeholders who might be impacted directly/indirectly/excluded by use of the dataset? 6. What are the potential misuse or harm?
Distribution	1. Dataset provenance (sources, private/public, license, copyright, history, owner)? 2. Document any ink if available for public dataset (with personally identifiable information). 3. Document any Github repository for code, scrapper. 4. Clarify special consideration for the anonymous datasets.
Maintenance	1. When was the last time the public dataset was published/last updated? 2. Document private dataset collection date. 3. Personally identifiable information should be published with consent.

Figure 1: Responsible Dataset Use Guideline Questions for journalism.

gorithm related information by answering several key questions regarding news storytelling. The questions were derived and informed by relevant prior work from transparent journalism [12, 10, 19, 13, 20, 54] and data science among others [50, 42, 32, 34, 33, 51, 58, 44, 52, 38]. The proposed responsible documentation guideline is specifically crafted for journalism (or journalists team internal use), but perhaps maybe with caution or upon request can be made available for the citizens as well.

There are several ways this work can be extended in the future, first, this work should be evaluated with journalists and other stakeholders to understand diverse (critical) user information needs [63] (e.g., what information is safe to reveal and to whom). Secondly, the factors reported in the initial guideline, though, was meant to be exhaustive, but likely it is not because, in the real-world journalism practice, things might change due to various factors outside of data and algorithms (e.g., resource/timing constraints, legal, profit vs non-profit); as a result, new questions might emerge and add up. Lastly, as newsrooms are increasingly adopting automated news production, thus, how the proposed method will scale is an open line of inquiry.

5. Acknowledgments

I would like to thank anonymous reviewers for their valuable comments and feedback.

Model/Algorithm	1. What supervised/unsupervised algorithm/model was used? 2. Linear/non-linear model? 3. What statistical test was applied? why? 4. Data analyst demographics (e.g., race, class, gender). 5. What specific data operations (max/min/avg, outliers) were used? 6. Which outlier detection method was used? why?
Parameters / Features	1. Which features/target variables were used/not used for prediction? Why (for decision-making context)? 2. What were the feature settings/range/threshold value and weightning? 3. Accuracy, precision, recall, confidence, F1 score, AUC/ROC metrics. 4. Error analysis or uncertainty. 5. What features or feature sets were identified as interesting? Why?
Tools / Editor	1. What software editor or coding platform, (e.g., spreadsheet, jupyter notebook) was used? 2. Any data visualization tool used?
Programming Language / Code	1. Which language (Python, C/C++, SQL, R) was used? 2. Sample database query, publish source code or pseudocode for reproducibility, code used for visualization, citations/link. 3. Summary of ML models or statistics. 4. What were programmer/coder demographics? 5. Time/date information.
Hardware	1. Desktop/laptop/mobile/tablet devices, scanner, computer-related information. 2. Any special considerations for proprietary algorithms/devices/software? how they were accessed or audit information?
Verification/Replication	1. Are story drafts shared with all subjects and they were given a chance to talk? 2. How was verification done in the case of anonymous sources? 3. Reverse-engineered code/algorithm/results? details? 4. Who are the external or domain experts contacted? 5. Team members responsible for checking numbers from data analysis? 6. How code replication/verification was done? 7. How conflict was resolved in case of disagreement? 8. How were biases checked? 9. Demographics of editor, reporters who verified code/data.
Story Narrative Related	1. Demographics of scripter. 2. All facts such as quotes, age, contact address, name, titles, links, time-date, facts/numbers discovered or obtained from outside sources double-checked? 3. Story publication/start/update date and details.

Figure 2: Responsible Algorithm use Guideline to facilitate verification and reproduciblity in News Storytelling.

References

[1] C. D'Ignazio, L. F. Klein, Data feminism, MIT Press, 2020.

[2] A. Christin, Algorithms in practice: Comparing web journalism and criminal justice, Big Data & Society 4 (2017) 2053951717718855.

[3] V. Eubanks, Automating inequality: How high-tech tools profile, police, and punish the poor, St. Martin's Press, 2018.

[4] C. O'neil, Weapons of math destruction: How big data increases inequality and threatens democracy, Crown, 2016.

[5] E. Raff, A step toward quantifying independently reproducible machine learning research, Advances in Neural Information Processing Systems 32 (2019) 5485–5495.

[6] M. Feinberg, W. Sutherland, S. B. Nelson, M. H. Jarrahi, A. Rajasekar, The new reality of reproducibility: The role of data work in scientific research, Proceedings of the ACM on Human-Computer Interaction 4 (2020) 1–22.

[7] B. Haibe-Kains, G. A. Adam, A. Hosny, F. Khodakarami, L. Waldron, B. Wang, C. McIntosh, A. Goldenberg, A. Kundaje, C. S. Greene, et al., Transparency and reproducibility in artificial intelligence, Nature 586 (2020) E14–E16.

[8] S. Chattopadhyay, I. Prasad, A. Z. Henley, A. Sarma, T. Barik, What's wrong with computational notebooks? pain points, needs, and design opportunities, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–12.

[9] J. A. Tullis, B. Kar, Where is the provenance? ethical replicability and reproducibility in giscience and its critical applications, Annals of the American Association of Geographers 111 (2021) 1318–1328.

[10] B. Kovach, T. Rosenstiel, The elements of journalism: What newspeople should know and the pub-

lic should expect, Three Rivers Press (CA), 2014.

[11] T. Aitamurto, M. Ananny, C. W. Anderson, L. Birnbaum, N. Diakopoulos, M. Hanson, J. Hullman, N. Ritchie, Hci for accurate, impartial and transparent journalism: Challenges and solutions, in: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–8.

[12] N. Diakopoulos, M. Koliska, Algorithmic transparency in the news media, Digital Journalism 5 (2017) 809–828.

[13] K. McBride, T. Rosenstiel, The new ethics of journalism: Principles for the 21st century, CQ Press, 2013.

[14] politifact, Politifact: The poynter institute, 2022. URL: https://www.politifact.com/, accessed: 2022-1-1.

[15] S. M. Julia Angwin, Jeff Larson, L. Kirchner, Machine bias, 2016. URL: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing, accessed: 2020-07-08.

[16] S. U. Noble, Algorithms of oppression: How search engines reinforce racism, nyu Press, 2018.

[17] I. Shapiro, Evaluating journalism: Towards an assessment framework for the practice of journalism, Journalism Practice 4 (2010) 143–162.

[18] N. H. Riche, C. Hurter, N. Diakopoulos, S. Carpendale, Data-driven storytelling, CRC Press, 2018.

[19] J. Gray, L. Chambers, L. Bounegru, The data journalism handbook: How journalists can use data to improve the news, " O'Reilly Media, Inc.", 2012.

[20] F. Chevalier, M. Tory, B. Lee, J. van Wijk, G. Santucci, M. Dörk, J. Hullman, From analysis to communication: Supporting the lifecycle of a story, in: Data-Driven Storytelling, AK Peters/CRC Press, 2018, pp. 169–202.

[21] H. De Burgh, Investigative journalism, Routledge, 2008.

[22] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, et al., Language models are few-shot learners, arXiv preprint arXiv:2005.14165 (2020).

[23] C.-G. Lindén, H. Tuulonen, A. Bäck, N. Diakopoulos, M. Granroth-Wilding, L. Haapanen, L. Leppänen, M. Melin, T. Moring, M. Munezero, et al., News automation: The rewards, risks and realities of 'machine journalism' (2019).

[24] N. Diakopoulos, Automating the news, Harvard University Press, 2019.

[25] M. Broussard, Artificial intelligence for investigative reporting: Using an expert system to enhance journalists' ability to discover original public affairs stories, Digital Journalism 3 (2015) 814–831.

[26] A. Jain, B. Sharma, P. Choudhary, R. Sangave, W. Yang, Data-driven investigative journalism for connectas dataset, arXiv preprint arXiv:1804.08675 (2018).

[27] D. Showkat, E. P. S. Baumer, Outliers: More than numbers? (2020).

[28] D. Showkat, E. P. S. Baumer, Where do stories come from? examining the exploration process in investigative data journalism, Proceedings of the ACM on Human-Computer Interaction 5 (2021) 1–31.

[29] B. Lee, M. Brehmer, P. Isenberg, E. K. Choe, R. Langner, R. Dachselt, Data visualization on mobile devices, in: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, 2018, pp. 1–8.

[30] D. G. Johnson, N. Diakopoulos, What to do about deepfakes, Communications of the ACM 64 (2021) 33–35.

[31] D. Showkat, Determining newcomers barrier in software development: An it industry based investigation, in: Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing, 2018, pp. 165–168.

[32] I. Steinmacher, T. Conte, M. A. Gerosa, D. Redmiles, Social barriers faced by newcomers placing their first contribution in open source software projects, in: Proceedings of the 18th ACM conference on Computer supported cooperative work & social computing, 2015, pp. 1379–1392.

[33] A. X. Zhang, M. Muller, D. Wang, How do data science workers collaborate? roles, workflows, and tools, Proceedings of the ACM on Human-Computer Interaction 4 (2020) 1–23.

[34] M. Muller, I. Lange, D. Wang, D. Piorkowski, J. Tsay, Q. V. Liao, C. Dugan, T. Erickson, How data science workers work with data: Discovery, capture, curation, design, creation, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–15.

[35] A. Y. Wang, D. Wang, J. Drozdal, X. Liu, S. Park, S. Oney, C. Brooks, What makes a well-documented notebook? a case study of data scientists' documentation practices in kaggle, in: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, 2021, pp. 1–7.

[36] N. Diakopoulos, Algorithmic accountability reporting: On the investigation of black boxes

(2014).
[37] K. Kirkpatrick, Putting the data science into journalism, 2015.
[38] P. Guo, Data science workflow: Overview and challenges, OCTOBER 30, 2013. URL: https://cacm.acm.org/blogs/blog-cacm/169199-data-science-workflow-overview-and-challenges/fulltext, accessed: 2020-08-22.
[39] P. Bradshaw, The inverted pyramid of data journalism, JULY 07, 2011. URL: https://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/, accessed: 2020-05-22.
[40] E. Rader, K. Cotter, J. Cho, Explanations as mechanisms for supporting algorithmic transparency, in: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, ACM, 2018, p. 103.
[41] T. Kulesza, M. Burnett, W.-K. Wong, S. Stumpf, Principles of explanatory debugging to personalize interactive machine learning, in: Proceedings of the 20th international conference on intelligent user interfaces, 2015, pp. 126–137.
[42] Q. V. Liao, M. Pribić, J. Han, S. Miller, D. Sow, Question-driven design process for explainable ai user experiences, arXiv preprint arXiv:2104.03483 (2021).
[43] N. Sonboli, J. J. Smith, F. C. Berenfus, R. Burke, C. Fiesler, Fairness and transparency in recommendation: The users' perspective, arXiv preprint arXiv:2103.08786 (2021).
[44] U. Ehsan, Q. V. Liao, M. Muller, M. O. Riedl, J. D. Weisz, Expanding explainability: Towards social transparency in ai systems, arXiv preprint arXiv:2101.04719 (2021).
[45] H. Nori, S. Jenkins, P. Koch, R. Caruana, Interpretml: A unified framework for machine learning interpretability, arXiv preprint arXiv:1909.09223 (2019).
[46] M. 2020, What is responsible machine learning? (preview), 2020. URL: https://docs.microsoft.com/en-us/azure/machine-learning/concept-responsible-ml, accessed: 2020-11-12.
[47] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi, A survey of methods for explaining black box models, ACM computing surveys (CSUR) 51 (2018) 1–42.
[48] A. Abdul, J. Vermeulen, D. Wang, B. Y. Lim, M. Kankanhalli, Trends and trajectories for explainable, accountable and intelligible systems: An hci research agenda, in: Proceedings of the 2018 CHI conference on human factors in computing systems, 2018, pp. 1–18.
[49] H. Kaur, H. Nori, S. Jenkins, R. Caruana, H. Wallach, J. Wortman Vaughan, Interpreting interpretability: Understanding data scientists' use of interpretability tools for machine learning, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–14.
[50] T. Gebru, J. Morgenstern, B. Vecchione, J. W. Vaughan, H. Wallach, H. Daumé III, K. Crawford, Datasheets for datasets, arXiv preprint arXiv:1803.09010 (2018).
[51] E. M. Bender, B. Friedman, Data statements for natural language processing: Toward mitigating system bias and enabling better science, Transactions of the Association for Computational Linguistics 6 (2018) 587–604.
[52] M. Mitchell, S. Wu, A. Zaldivar, P. Barnes, L. Vasserman, B. Hutchinson, E. Spitzer, I. D. Raji, T. Gebru, Model cards for model reporting, in: Proceedings of the conference on fairness, accountability, and transparency, 2019, pp. 220–229.
[53] E. M. Bender, T. Gebru, A. McMillan-Major, S. Shmitchell, On the dangers of stochastic parrots: Can language models be too big?, in: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 2021, pp. 610–623.
[54] N. Diakopoulos, Transparency, in: The Oxford Handbook of Ethics of AI, 2020.
[55] C. D'Ignazio, L. F. Klein, Seven intersectional feminist principles for equitable and actionable covid-19 data, Big data & society 7 (2020) 2053951720942544.
[56] M. Kogan, A. Halfaker, S. Guha, C. Aragon, M. Muller, S. Geiger, Mapping out human-centered data science: Methods, approaches, and best practices, in: Companion of the 2020 ACM International Conference on Supporting Group Work, 2020, pp. 151–156.
[57] M. Vaismoradi, J. Jones, H. Turunen, S. Snelgrove, Theme development in qualitative content analysis and thematic analysis (2016).
[58] E. M. Bender, A typology of ethical risks in language technology with an eye towards where transparent documentation can help, in: Future of Artificial Intelligence: Language, Ethics, Technology Workshop, 2019.
[59] N. Diakopoulos, Ethics in data-driven visual storytelling, in: Data-Driven Storytelling, AK Peters/CRC Press, 2018, pp. 233–248.
[60] S. Holland, A. Hosny, S. Newman, J. Joseph,

K. Chmielinski, The dataset nutrition label: A framework to drive higher data quality standards, arXiv preprint arXiv:1805.03677 (2018).
[61] B. Lee, N. H. Riche, P. Isenberg, S. Carpendale, More than telling a story: Transforming data into visually shared stories, IEEE computer graphics and applications 35 (2015) 84–90.
[62] M. Broussard, Big data in practice: Enabling computational journalism through code-sharing and reproducible research methods, Digital Journalism 4 (2016) 266–279.
[63] J. Kemper, D. Kolkman, Transparent to whom? no algorithmic accountability without a critical audience, Information, Communication & Society 22 (2019) 2081–2096.