Seeing How the Sausage is Made: Data Storytelling as Means and Method in a Computer Science Writing Course

Michelle Trim
University of Massachusetts Amherst
USA
mtrim@cs.umass.edu

Erin Butler
University of Massachusetts Amherst
USA
ebutler@umass.edu

Christina Suttcliffe
University of Massachusetts Amherst
USA
csutcliffe@umass.edu

Abstract

As data corpus-driven tools and technologies increasingly push users to passively search for an answer, rather than search to understand, we believe that technical and computing disciplinary writing courses have a duty to teach the process of responsible data storytelling. While students can grasp that generative AI makes mistakes, hallucinates, and perpetuates bias, they can need help understanding the antecedent causes of those difficulties. All algorithmically driven decision-making or recommending software have in common a large data set that has been labeled, either by users or by the system itself. The origins of that data and the reasonable applications/deductions and conclusions possible for any given dataset have everything to do with why some tools help and some tools perpetuate harms. By starting at the very beginning and asking students to make sense of data, students can more easily see how purpose and audience impact analysis of any given collection of data. Once those opportunities for rhetorical choice making are known, students become ready to understand the connection between data and complex A.I. systems and some of the ways that bias and other kinds of harm can result if designers are not careful. Combining instruction in a technical coding environment with basic data literacy lessons such as 'the seven data stories,' [14] we developed and delivered a three-week writing unit designed around responsible data exploration and storytelling. In this experience report, we provide the assignment we used, and the scaffolded activities we employed to bring students through the process, remarking on what worked well and what we want to improve. We provide attendees with a link to an R-based notebook with a walk-through lesson on data exploration commands, and the rubric used to assess students' texts, notebooks with code and commentary and results, all existing in a referential context. We provide the survey results of students' perception of learning from this activity. Early findings demonstrate that students internalized lessons about the non-objective nature of data analysis and of specific responsible data storytelling practices required by anyone seeking to ethically represent answers within and limitations of any dataset.

SIGDOC '24, October 20–22, 2024, Fairfax, VA, USA

ACM ISBN 979-8-4007-0519-9/24/10
https://doi.org/10.1145/3641237.3691673

CCS Concepts

• **Social and professional topics** → Professional topics; Computing education.

Keywords

Data storytelling, Data Visualization, Pedagogy, Technical communication

ACM Reference Format:
Michelle Trim, Erin Butler, and Christina Suttcliffe. 2024. Seeing How the Sausage is Made: Data Storytelling as Means and Method in a Computer Science Writing Course. In *The 42nd ACM International Conference on Design of Communication (SIGDOC '24), October 20–22, 2024, Fairfax, VA, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3641237.3691673

1 Introduction

Teaching writing in a computer science context requires attention to technical communication best practices and to long held tenets of WID (writing in the disciplines) pedagogy. The intersection, and overlap, of these two bodies of scholarly practice invite a teaching approach that is simultaneously attentive to the rhetorical needs of a given context, to the ethical obligations to the user, and to the demystification of disciplinary genres and conventions. Our teaching challenges are augmented by our disciplinary writing class requiring attention to the social impacts of computing, a content area that is often at odds with the prevailing cultural and epistemological trends in computer science [26]. Computer science students can struggle with writing purposes and concepts, notably explanation and awareness of audience, and especially in programs with limited integration of writing instruction, a phenomenon at odds with the ACM/IEEE curriculum recommendations [1]. Student struggles in writing for audiences different from themselves seem to coincide with the challenges that computer science sometimes faces as a field in cultivating ethical design practices for diverse and inclusive communities [11]. There is intentionally a lack of instruction in specific genres, so our approach is somewhat at odds with traditional technical communication textbooks which tend to march students through memos, reports, and various proposal forms, sometimes reflecting a more traditional conception of the workplace. Based on a previous study of computing professionals that was conducted in 2022, we know that tech employers have bespoke genres of writing that vary across companies. Most of these genres share underlying expectations for concision, attention to detail, precision in language, clarity, and audience-aware explanation.

As our class is charged with teaching the social impacts of computing, we also address the ethos and audience oriented ethical components inherent to technical communication [25]. By focusing on the user of workplace communication in understanding audience, we humanize and complicate students' often abstracted and homogenized conception of users [26]. This user focus seeks to shape a cultural understanding of readers as defined by their information needs and prior knowledge of the technical subject at hand [13]. Shifting students away from *what they want to say* to what *readers need to know* is a powerful step in their maturation as computing professionals. We have found that by disrupting students' notions of data representation as an inherently neutral activity that we can invite more critical thinking about the rhetorical positioning of the data visualizations and data stories that they encounter in their daily lives [16]. This critical thinking is necessary for the ethical interrogation of data driven AI tools and algorithmic systems [21, 27]. With the popularity of AI-assisted technology, and the increasing attention being paid to its potential for unethical implementation and bias-based harm [2], our new unit takes as its project the ethical explanation of data storytelling. By constructing the audience for this unit as a team leader, i.e. a near-peer supervisor, we disrupt some of the uncritical performances of mathematical analysis often ingrained in computer science students. We intentionally construct the analysis as a mundane activity. Our approach enables us to realistically walk students through a rudimentary data exploration that includes audience aware code commenting, explanation, description, reflection and basic data visualization and carries the ethical obligation of explaining the potential and the limitations of their chosen datasets.

This paper shares the content of that assignment, some of the activities used in teaching it, and a report on an IRB-approved study of students' reactions to this new writing unit in our class. We also explain the connection between data storytelling and AI, highlighting the role ethical awareness of audience might play for these future computer scientists in building AI infused tools. Finally, we highlight the ways this assignment with its intense focus on process serves as an example of a pedagogical approach that puts students' technical writing in the center and disincentivizes students' use of AI as a substitute for their own intellectual labor and decision making.

2 Data literacy as foundational knowledge and rhetorical skill

Data literacy is about more than interpreting a chart or graph [6], and as a construct, it describes a set of skills that may be overestimated in STEM students. Considering the prevalence of 'data' as evidence, persuasive construct, and justification for specific civic actions, including the primary driver for algorithmic decision making, it is imperative for students writing about data to "understand randomness, sampling, association, and other concepts that support inferential reasoning" [4]. It is typical for computer science students to take courses that cover statistics with emphasis on probability, machine learning, and other data science practices related to model building. As in other schools, computing students in our program must electively pursue instruction in the rhetorical situatedness of data representation, interpretation and visualization. Even then, it is rare for students to learn how to communicate about data to those with different expertise in their computer science courses [13, 15]. Our writing class is well positioned to address this gap via this new unit focused on developing responsible data communication skills through data storytelling [10].

Because data analysis can be performed without advanced mathematical or statistical knowledge, some may assume that STEM students, particularly juniors and seniors, have this knowledge already [13]. Organizations like the Algorithmic Justice League, Data for Black Lives, and AINow have worked diligently to expose and explain the concept of algorithmic bias to the public. Algorithmic bias can originate from biased data sets and implementations of these systems have adversely impacted marginalized identities worldwide [3, 9]. The concept of bias is not new to students when they enter our class. That said, few students come to our writing class knowing the effects of hasty and improper generalizations and weak inferences from data that originate at the collection stage and lurk in a dataset [13]. Consequently, we consider data literacy to describe a foundational set of skills [19] that include the ability to identify, question, arrange, reason about, and interpret data and its visualizations in supportable and responsible ways [5, 14, 17, 18].

2.1 Data storytelling as technical communication concept

Data storytelling is a translanguaging that exemplifies what the New London Group documented in their concepts of "Multiliteracies" - bringing together the "representational, social, structural, intertextual and ideological" of communication to "account for their purposes." [8]. Put more directly, data storytelling is "an approach that integrates data, visuals, and narratives to create data stories that can help a particular audience to comprehend the key messages underscored by the data with enhanced efficiency and effectiveness" [25]. Selecting, classifying, representing, and inferring the 'key messages' of a group of data are all rhetorical acts that position the writer and their purpose according to the context and the subject of the data itself. Determining which questions can be asked of a given dataset is another layer of rhetorical reasoning that depends on the recognition of social, political and structural features of that data within a particular context [6]. For example, a dataset that charts behaviors related to outdoor activities and uses variables only for 'male' and 'female' is limited in what inferences that it can reasonably support about certain kinds of gender differentiated preferences or behaviors [9]. Not only is the use of a binary classification of gender itself a political act that can lead to inaccurate conclusions, but the potential exists for an underlying assumption of purpose that seeks to generalize about an entire category of people based on what may or may not be a representative sample. Gathering and then explaining data is a common workplace task for technical professionals [17]. Literature across multiple subfields of computer science argue that data storytelling adds a crucial communicative and audience aware component to data visualization [6, 10, 13, 22, 25].

2.2 Technical writing pedagogy in a computers and society context

The relationship between rhetoric, ethics, and technical communication is longstanding. There are many definitions of rhetoric, and one that we find particularly useful for thinking about the 'project' of teaching a disciplinary writing course to computer science students during this era of climate catastrophe and technological disruption comes from technical communication scholar, Dr. Angela Haas. She defines rhetoric "as the negotiation of cultural information—and its historical, social, economic, and political influences—to affect social action (persuade)" [12]. Haas goes on to say that "every culture has its own rhetorical roots, traditions, and practices. Although sometimes forgotten, rhetoric seeks engagement with and participation in effective and responsible civic discourse" [12]. One could argue that one way of thinking about 'responsible civic discourse' includes user-centered and usability centered [24] pedagogical orientations to teaching technical writing as both share a deep sensitivity to understanding and addressing the needs of the audience/user. Lessons in audience awareness in a STEM writing course can humanize the construction of the 'user' and thus facilitate greater understanding of the social impacts of technology [26]. Perhaps it is this attention to responsibility that enables rhetoric (and rhetorically informed teaching) to disrupt traditional or orthodox conceptions of technology, and technical communication by extension. This ability to respond in the moment gives rhetorically informed teaching the opportunity to help students shape new and more useful, more attentive ways of writing with and about technology.

In many ways, technical communication pedagogy is particularly well suited to examining the impacts of computing on society. Whether it is through focuses that demand user participation, involve clients as real audiences, build broad coalitions of technologists and communicators, or integrate multiple disciplines and teach multiple literacies [7], technical communication pedagogy strives to be inclusive, dynamic, adaptive, and situational. Indeed, Cook's layered literacy approach to defining a technical communication pedagogy is particularly well suited to teaching computing students about data representation, as "using visuals appropriately and effectively requires individuals to think about them in terms of their basic, critical, ethical, rhetorical, social, and technological components" [7]. Our assignment requires a basic visualization from students, but our evaluation of their submission targets the choice of visualization and their rationale for choosing it, emphasizing to students their rhetorical decision-making extends to non-alphabetic texts as well.

3 Data Storytelling lesson example

3.1 A data exploration assignment

Our data storytelling unit asked students to select a dataset, develop a research question that their chosen data set could answer, explore the data using some basic *R* programming commands, develop a visualization to either guide their inference or display their findings, and reflect on their process by discussing the limitations of the data set they selected, the extent to which they were able to answer their research question, and what additional information they would need in order to come closer to answering their original or any subsequence research question that guided their exploration. Throughout this process, students were instructed to intersperse code snippets with text blocks, explaining each step of their invention and exploration process. They were also asked to insert code comments that explained the purpose of every command they wrote, e.g. why they used it and what they expected it to do. Students were asked to share their notebooks and datasets in groups for a directed in-class peer review session, followed by a revision plan which they submitted prior to submitting their revised notebooks (with their accompanying datasets) for a grade.

The unit was attempted in seven different sections of our junior/senior disciplinary writing course, with approximately 140 students completing the 2.5-week f`5unit. Of the four instructors who deployed it, only the course director had experience and prior training in using R in a data science context, and some of us had not really coded much before at all. The course director led a workshop which gave an orientation to the tools and goals for the unit, and each instructor developed their own course activities to teach the underlying rhetorical and technical writing concepts embedded in the unit.

3.2 Using Colab/Jupyter notebooks to emphasize process

Because our university has a contract with Google, all students already have access to Google Colab, a sharable coding environment that facilitates text mark-up. Since the product runs entirely in the browser, it is accessible to all students in our classes, and requires no installation or technical support to operate. Within the tool, users can code in R or Python programming language, can share their notebook easily with others, and their work is automatically saved while they write. The code boxes are individually executable, facilitating good coding practices already familiar to advanced computing students such as coding small and debugging as they work. Jupyter Notebooks are a similar popular set of tools that one could use and access for free, however that tool requires a local installation on the user's machine, which can add a layer of logistical complexity that may not fit every teaching environment.

3.3 Evaluating Students' Work

Because this was our first time trying out this new unit, and because the four of us leading the assignment had different prior exposure to programming, we focused the evaluation of the assignment on the presence of required components and the sufficiency of explanation of the student's process in carrying out their data exploration and constructing their data story. We did not grade their code nor did we give extra credit to students whose code reflected advanced understanding of data science analysis using *R*. In fact, few students had used this *R* before, leveling the field a bit between the writing trained instructors and the computer science students. We used the following point assignment rubric to render a grade, and focused our in-text feedback on helping students grasp when and how their explanations feel short of bringing the reader along through their process. Students were invited to resubmit their work for a revised grade following another round of reflection and revision. Our grading rubric had six criteria, with the data exploration code and accompanying comments, explanation of process, discussion of limitations of the dataset, and discussion of potential future questions all weighted equally. Rationale for dataset selection (and its relationship to the question guiding the analysis) and the visualization itself were weighted about half of one of the other categories. In our evaluation, we focused our feedback on the quality of the explanation itself, with emphasis on clarity, thoroughness, audience-appropriate language choices, and coherence.

One of the responsible data practices presented to students in this unit focused on reusability. If aspects of the data exploration process lacked sufficient explanation such that a new person couldn't enter the notebook and reuse code chunks or extend analyses, then the communication was ineffective. Likewise, replicability of analyses is an important scientific concept. Code and process explanations had to be sufficiently explanatory such that a fellow student or co-worker could set up the same analysis with a different data set and be able to evaluate the results in comparison. These two ideas of replicability and reuse are common to computing professionals' work activities but they are not often taught as a function of audience awareness. By emphasizing the necessity of audience aware code comments and explanations, we add an important ethical consideration to the work students do for this unit [28].

4 Examples of in-class activities to support the lesson

An instructor explained that one common side effect of data amalgamation is that specific data is lost in the presentation. Those specific data points can mean the loss of representation of an entire demographic so students were told to look at their data not just for the information given, but also for the information that is hidden. This hidden data can be missing data, but it can also be data that has been generically labeled. Barabara Regenspan, in her book on educational exclusion, called this "haunting" data - because she noted that sometimes the generic labels can label casualties as "collateral damage" when the bias of friend or foe is a factor [23]. Looking at data for what may be purposefully (or accidentally) omitted is a crucial component in data storytelling. One instructor had the students find a graph that was representative of a global population. Students were asked to study the graph in small groups and to report back on the following data: What data was represented? What data was not represented? Where did this data come from? How was it presented and did the method (colors, graph type, etc.) evoke any alternate meanings for your group?

Starting with the careers website for our school, we asked students to critically examine the charts indicating top employers, percentage of graduates they hired, and which careers were most common for last year's graduates. The results of this activity demonstrated that the students were able to look beyond the chart and start to think of the individuals represented (and those that weren't) - telling the story that was / was not within the data. Questions emerged slowly as the instructor modeled one analytical approach by asking students to define the population represented in the graph. Students noticed that the graph seemed to be comprehensive, but then realized there was only a 44% response rate on the survey that drove it. Others observed that Masters students were lumped in with the four-year graduates, requiring the user to click a specific tab to examine the populations separately. Another student noted that the scale of the graph gave an initial impression that lots of students were being hired by a particular company until they realized that the representation was of the number of hires compared to how many new graduates the other companies had hired. It became clear that companies with only one hire were the most common and that all of them were not listed on the chart. Soon students were ready to identify the 'story' embedded in this careers page and students were able to see that while true, the story they identified had a definite slant. Students were then asked to reflect on their sense of the site's ethicality – why or why not the slant they had perceived was still within the bounds of what counts as 'responsible' data analysis. Another group, who found a chart on Population (by Age and Gender) on the US Census Bureau site were more critical, asking "why are men in blue and women in what appears to be pink?" Another student asked: "What about non-binary residents?" These activities with existing, real data visualizations were foundational in helping the students to appreciate how data is rhetorically presented and thus, what their ethical responsibility is in providing accurate reporting in their own digital storytelling.

5 Effectiveness of the Data Storytelling unit

5.1 Observations of student learning based on work submission

We were delighted and surprised to find that most students were enthusiastic about the data storytelling project, and they were diligent in looking for data that aligned with their interests. One education-minded student looked at K-12 school attendance data from the state of Connecticut, hypothesizing that the recent "food security measures in the state" would have a positive impact on "the rate of school attendance." While they instead found that overall attendance continues to drop in recent years, they were undeterred, wondering if "the pattern would differ in other states or across the country" and "what about over a longer time period?" All of us felt that if we had allowed the students to continue this project into another unit - many of them would have continued their research to discover the "why's" of the stories they were finding through their data explorations. We saw their desire to know more about the stories hidden in their data to be an important awakening. Rather than seeing one visualization and assuming a surface interpretation, our students were beginning to accept the assumption that there was always more going on than they could see in any one view. This recognition is a key precursor to their understanding of the rhetorical representation of data and the rhetoric driving data story telling.

In their submitted assignments, students excelled in two key areas: identifying limitations of their chosen datasets and determining what kinds of questions could be reasonably and ethically answered by the datasets. As part of the assignment guidelines, we asked students to select real datasets about a topic of their choice, rather than providing curated datasets. Inherent in this task was the challenge of identifying the limitations and potential problems that are common in real-world datasets. By inspecting the metadata associated with their datasets, as well as the data itself, most students were able to identify where their data was inconsistent, unreliable, incomplete, or otherwise questionable. They articulated these limitations within their final assignments. For example, one student noted that his dataset's metric of "deaths per 100,000" was unclear because of ambiguity in the metadata. Another student pointed out that the data they examined came from a survey in which only households with an associated email address or phone number were eligible to participate, which automatically excluded certain groups. After articulating the limitations of their datasets, students were able to temper or otherwise qualify the conclusions they drew in their data stories, enacting the ethical reasoning we were hoping to teach.

Similarly, many students successfully identified the kinds of questions their datasets could or could not reasonably answer, often limiting themselves to narrower questions so as not to generalize. During class discussions, they resonated with stories of hastily drawn conclusions and questionable generalizations taken from current event data reporting scandals, and this seemed to contribute toward a more discerning consideration of their own data. Overall, students were able to recognize when questions were beyond the scope of their dataset, particularly when such questions required additional data or research. This felt important to us as recognizing what one cannot reasonably support with a given dataset speaks directly to the ethical use of evidence, and to the shaping of ones ethos through their manner of argument support.

Consistent with the trends we have observed during other units and assignments, we found that students struggled with clear explanations of the datasets themselves, as well as explanations of their process and reasoning. There was often a gap between what students understood and what they explicitly communicated. We are interested in spending more time on this unit to address the learning challenge students faced in constructing explanations that worked for their readers. As this aspect of the assignment is key to our class, we want to add another peer review to the unit, one where we ask readers to ask questions raised by their peers' work, illuminating those gaps where writers' explanations were insufficient.

Because the assignment asked students to submit their datasets along with their Colab notebooks, but not the accompanying metadata for their chosen dataset, their data stories were most effective when they included descriptions of the key metadata necessary to understand their datasets. This was a learning experience for us as we realized we should have required that students explicitly report that information. A significant proportion of data stories lacked these kinds of descriptions, with students' text seeming to indicate that this metadata was self-evident. We found that this was an important stumbling block in the chain of communication; for example, if the purpose of each row in a dataset was unclear, it was then difficult to interpret the rest of the dataset and its accompanying data story. Students needed to be told that they have to explain how the data set works/what its structure is and what types of information are contained in its cells.

Some students also struggled to explain their processes of data exploration. This issue took several forms. In some instances, students moved from one calculation or visualization to another without clear reasoning for the progression (e.g. noticing an outlier and then comparing that one column or record to the rest without explaining that the goal was to better understand if the outlier was really an outlier). In other instances, students neglected to explain their choice of visualization either wholly or in part. Subsequent revisions indicated that many students found their visualizations to be obvious choices that did not require explanation. While this somewhat makes our argument for us in that students need instruction in visual rhetoric, it illustrates the challenges faced in trying to teach multiple concepts in a new way. In still other cases, students omitted code comments entirely, opting to prioritize summative comments for an audience that they perceived as being unlikely to read granular explanations of coding choices. This observation also supports our thinking that another peer review focused exclusively on effective awareness of audience will help.

5.2 A study of student perceptions of the data storytelling unit

We invited students to complete an IRB-approved, optional and anonymous survey following the submission of their final project. Our response rate of 58% yielded 81 completed surveys that largely supported our individual observations. When asked what concepts were new to them, their responses were somewhat unexpected. Nearly half of the students (46.9%) responded that they had never written before about how they might take an analysis further and two thirds (67.9%) indicated that they had never been asked to find their own data set to analyze or to write about the limitations of a data set before. When we asked which of these concepts they felt most confident in performing following the completion of the unit, again, their answers surprised us. Just over half (50.6%) expressed confidence in selecting a data set, discussing the limitations of their dataset (63%) and writing about their analysis process (69.1%), and these three activities were new to a similar proportion of respondents at the outset of the unit. Based on point assignments, their confidence in their explanatory skills may have been a bit overly optimistic, but by and large, students did well on this assignment overall.

5.3 Analysis of overall results

We also asked a series of open-ended questions, including asking students what they think they most learned from the unit. A significant number of students expressed sentiments like these: "The biggest takeaway for me was learning how to be more skeptical of data stories and be more curious about the metadata instead of just taking things at face value" and "The importance of being skeptical in the sources that we use and considering their impact, limitations, and future analyses." One student remarked that "I learned that things might not be as they seem. I had very surprising results when I completed my project that I was not expecting, especially based on what I have learned in my own education. I think that this can be extended to any field that employs data collection, as this data can bring things to light that might not always be extremely obvious." Similarly, another student responded, "My hypothesis was proven wrong, so this project showed me the importance of being flexible to all sides of the story while you work with data. Moreover, I now appreciate that not every story is "complete"; there are always missing variables and other factors to be aware of." Many students also gave responses that were some variations of this one: "I think that I gained a better understanding of how to communicate the limitations of a dataset in my data analysis." Taken together, we see these responses and the corresponding performance by students on this assignment as strong support that the use of a digital notebook that interspersed coding with process-focused explanation connected in many students' minds the intersecting ideas of responsible data analysis, audience-aware communication, and the rhetorical nature of data storytelling. We also see areas where we can improve our assignment instructions (requiring inclusion of metadata about the dataset) and increase student opportunities for reflection on their learning using an additional peer review session. Because significant portions of the process work for this assignment were completed in class, because students became genuinely interested in their findings, and because each students' analysis was specific to the context of our class and their selected dataset, we believe there were few to no incidences of unapproved tools, such as text and code generators. This observation supports our claim that a focus on process work disincentivizes the use of generative AI as a substitute for students' own critical thinking and writing.

6 CONCLUSION

The new field of data science [3] is fast becoming that which undergirds ubiquitous technologies ranging from recommendation systems in social media and shopping to model building for insurance cost structures to decision systems to generative artificial intelligence [20]. There is no field protected from these technological developments, no corner of civic life that is absent from their influence. Technical communication educators have an imperative as among the best qualified professionals to lead the charge in infusing data literacy and ethical data story telling into technical communication instruction. To explain data is to write about data, and that activity is most definitely in the domain of technical writing professionals. Our assignment showed us that despite some of us having no formal training in computing or even prior experience with coding, that with a little mutual support and guidance, we were able to effectively administer a new learning unit on data story telling for upper-level computing majors. We learned that we can improve our teaching of this unit through more explicit instructions regarding including the metadate of datasets, and the conscious awareness of the need to read and understand that metadata, and by spending even more time practicing the explanatory elements needed to make a data exploration reusable and replicable by real audiences. We also learned that provoking student curiosity can be an effective way to keep students engaged and willing to undertake a new learning experience.

References

[1] ACM 2023. ACM/IEEE 2023 Curriculum. Retrieved from https://csed.acm.org/
[2] Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim code. Polity.
[3] Berman, F. *et al.*, "Realizing the potential of data science," Commun. ACM, vol. 61, no. 4, pp. 67–72, Mar. 2018
[4] Bobkowski, P. S., & Etheridge, C. E. (2023). Spreadsheets, Software, Storytelling, Visualization, Lifelong Learning: Essential Data Skills for Journalism and Strategic Communication Students. Science Communication, 45(1), 95-116. https://doi.org/10.1177/10755470221147887
[5] Brumberger, E. (2022, July). Re-thinking the Role of the Visual in Professional Communication: the Case for a Liberal Arts Core. In 2022 IEEE International Professional Communication Conference (ProComm) (pp. 364-365). IEEE.
[6] Chen, Lujie Karen. Jiaqi Gong, and Louise Yarnall. 2024. Foundational Tools for Coaching Data Storytelling. In Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 2 (SIGCSE 2024). Association for Computing Machinery, New York, NY, USA, 1894. https://doi.org/10.1145/3626253.3633432
[7] Cook, K. C. (2002). Layered Literacies: A Theoretical Frame for Technical Communication Pedagogy. Technical Communication Quarterly, 11(1), 5–29. https://doi.org/10.1207/s15427625tcq1101_1
[8] Cope, B. and Mary Kalantzis. 2009. "Multiliteracies" New Literacies, New Learning. Pedagogies: An International Journal, 4(3), pp. 164-195
[9] D'ignazio, C., & Klein, L. F. (2023). Data feminism. MIT press.
[10] Echeverria, V., Roberto Martinez-Maldonado, and Simon Buckingham Shum. 2017. Towards data storytelling to support teaching and learning. In Proceedings of the 29th Australian Conference on Computer-Human Interaction (OzCHI '17). Association for Computing Machinery, New York, NY, USA, 347–351. https://doi.org/10.1145/3152771.3156134
[11] Garcia, Rosalinda, Patricia Morreale, Pankati Patel, Jimena Noa Guevara, Dahana Moz-Ruiz, Sabyatha Satish Kumar, Prisha Velhal, Alec Busteed, and Margaret Burnett. 2024. Beyond "Awareness": If We Teach Inclusive Design, Will Students Act On It? In Proceedings of the 2024 ACM Conference on International Computing Education Research (ICER '24), Vol. 1. ACM, New York, NY, USA, 434–451. https://doi.org/10.1145/3632620.3671101
[12] Haas, A. M. (2012). Race, Rhetoric, and Technology: A Case Study of Decolonial Technical Communication Theory, Methodology, and Pedagogy. Journal of Business and Technical Communication, 26(3), 277-310. https://doi.org/10.1177/1050651912439539
[13] Johnson, R. R. (1998). User-Centered Technology: A Rhetorical Theory for Computers and Other Mundane Artifacts. (n.p.): State University of New York Press.
[14] Kang, M.. (2016). Exploring the 7 different types of data stories. Mediashift. http://mediashift.org/2015/06/exploring-the-7-different-types-of-data-stories/
[15] Koby, M. 2020. Data Science Education: Curriculum and pedagogy. In Proceedings of the 2020 ACM Conference on International Computing Education Research (ICER '20). Association for Computing Machinery, New York, NY, USA, 324–325. https://doi.org/10.1145/3372782.3407110
[16] Lindgren, C. A. (2024). A Stasis Network Methodology to Reckon with the Rhetorical Process of Data: How a Data Team Qualified Meaning and Practices. Technical Communication Quarterly, 1-26.
[17] Li, Y., Ye Wang, Yugyung Lee, Huan Chen, Alexis Nicolle Petri, Teryn Cha, Teaching Data Science through Storytelling: Improving Undergraduate Data Literacy, Thinking Skills and Creativity, Volume 48, 2023, 101311, ISSN 1871-1871, https://doi.org/10.1016/j.tsc.2023.101311.
[18] Lindoo, E.. 2022. A Case for Introducing Visual Data Storytelling in CS/CIS Curriculums. J. Comput. Sci. Coll. 38, 5 (November 2022), 42–53.
[19] McDowell, K. and Turk, M.J. (2024), "Teaching data storytelling as data literacy", Information and Learning Sciences, Vol. 125 No. 5/6, pp. 321-345. https://doi.org/10.1108/ILS-06-2023-0068
[20] McKee, H. and James E. Porter. 2020. Ethics for AI Writing: The Importance of Rhetorical Context. In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '20). Association for Computing Machinery, New York, NY, USA, 110–116. https://doi.org/10.1145/3375627.3375811
[21] Noble, S. U. 2018. Algorithms of Oppression: How Search Engines Reinforce Racism. New York University Press.
[22] Oberascher, L., Ploder, C., Spiess, J., Bernsteiner, R., & Van Kooten, W. (2023). Data Storytelling to communicate Big Data internally - a Guide for Practical Usage. European Journal of Management Issues, 31(1), 27-39. https://doi.org/10.15421/192303
[23] Regenspan, B.. 2014. Haunting and the Educational Imagination. Bold Visions in Educational Research. Volume 43. Sense Publishers.
[24] Schneider, S. (2005). Usable Pedagogies: Usability, Rhetoric, and Sociocultural Pedagogy in the Technical Writing Classroom. Technical Communication Quarterly, 14(4), 447–467. https://doi.org/10.1207/s15427625tcq1404_4
[25] Shao, H., Roberto Martinez-Maldonado, Vanessa Echeverria, Lixiang Yan, and Dragan Gasevic. 2024. Data Storytelling in Data Visualisation: Does it Enhance the Efficiency and Effectiveness of Information Retrieval and Insights Comprehension? In Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 195, 1–21. https://doi.org/10.1145/3613904.3643022
[26] Trim, M.. 2017. Increasing ethical awareness in [future] software developers using audience-based writing. In Proceedings of the 35th ACM International Conference on the Design of Communication, Article 30, August 11–13, 2017, Halifax, NS, 5 pages. DOI: https://doi.org/10.1145/3121113.3121219
[27] Trim, M.. 2020. Computing's social obligation. SIGCAS Comput. Soc. 48, 3–4 (September 2019), 13–14. https://doi.org/10.1145/3383641.3383644
[28] Williamson, Bill and Scott J. Kowalewski. 2018. Cultivating a rhetoric of advocacy for usability studies and user-centered design. Commun. Des. Q. Rev 5, 3 (November 2017), 35–47. https://doi.org/10.1145/3188173.3188177