Narrative Datasets through the Lenses of NLP and HCI

Sharifa Sultana
Cornell University
ss3634@cornell.edu

Renwen Zhang
National University of Singapore
r.zhang@nus.edu.sg

Hajin Lim
Seoul National University
hajin@snu.ac.kr

Maria Antoniak
Cornell University
maa343@cornell.edu

Abstract

In this short paper, we compare existing value systems and approaches in NLP and HCI for collecting narrative data. Building on these parallel discussions, we shed light on the challenges facing some popular NLP dataset types, which we discuss these in relation to widely-used narrative-based HCI research methods; and we highlight points where NLP methods can broaden qualitative narrative studies. In particular, we point towards contextuality, positionality, dataset size, and open research design as central points of difference and windows for collaboration when studying narratives. Through the use case of narratives, this work contributes to a larger conversation regarding the possibilities for bridging NLP and HCI through speculative mixed-methods.

1 Introduction

Human beings are myth-makers; we use stories and imagination to create communities and make sense of the world and our place in it (Bamberg and Georgakopoulou, 2008). Narratives are powerful modes of expression, with physical, emotional, and social benefits for both the narrator and the audience (Pennebaker and Beall, 1986; Pennebaker, 1997; Merz et al., 2014; Oh and Kim, 2016; Tangherlini, 2000). They can also be powerful methods for understanding human behavior and beliefs (Golsteijn and Wright, 2013).

Crucially, narratives are *situated*; they are told and take place in specific social contexts (Piper et al., 2021). Natural language processing (NLP) methods can analyze patterns across large datasets, putting stories into context. But narrative datasets in NLP are often removed from the narratives' original contexts (e.g., scraped internet datasets) or are designed without any explicit context or social grounding (e.g., short and artificial stories).

In contrast, contextuality is of the utmost importance in qualitative human-computer interaction (HCI) approaches to narrative. HCI researchers frequently borrow social science methods including surveys, interviews, focus groups, and ethnography for closer investigations that address the diversity of human life and experiences (Bruner, 1987; Golsteijn and Wright, 2013). Qualitative HCI methods are often constrained to small sample sizes and susceptible to observer biases, but narrative research and portraiture methods enable creative and holistic engagement with participants' experiences and meaning-making processes (Williams, 1984; Wright and McCarthy, 2004; Bardzell et al., 2012).

These differences make narrative datasets an useful case study when considering tensions and possible collaborations between NLP and HCI. Both disciplines face challenges in their study and analysis of narrative. While NLP datasets contain a high volume of data points, their labels are constrained to a specific task; in contrast, smaller HCI datasets, in particular data collected through qualitative methods such as ethnography and interview, are open-ended in research scope but situated in a particular context. Combining these methods can contribute to designing multifaceted datasets while not losing the sight of individual experiences and perspectives in a large volume of stories.

In the following sections, we outline dominant framings of narrative and narrative dataset collection in NLP and HCI. Placing these framings side-by-side highlights a set of tensions—including dataset size, contextuality and positionality, and dataset design—that we finally consider as material for synthesis and mixed methods approaches to narrative data.

2 NLP Framings of Narrative

In a recent overview of NLP and humanist approaches to "narrative understanding," Piper et al. (2021) formulate narrativity as a scalar construct rather than a binary class; texts can include some or all narrative features (e.g., narrator, audience, sequential actions). Most NLP narrative tasks focus on building **abstractions** from narratives by extracting these features and measuring relationships among them. These tasks include extracting narrative **structure**, like scripts, plot units, or narrative arcs (Schank and Abelson, 1977; Lehnert, 1981; Chambers and Jurafsky, 2008, 2009; Goyal et al., 2010; Reagan et al., 2016); modeling **connections** between characters (Bamman et al., 2013; Iyyer et al., 2016; Lukin et al., 2016); **generating** new stories or summaries (Goldfarb-Tarrant et al., 2020; Guan et al., 2020; Akoury et al., 2020); answering **questions** about the story (Richardson et al., 2013), and identifying a correct story **ending** (Chambers and Jurafsky, 2008; Mostafazadeh et al., 2016).

As in other areas of NLP, some narrative research falls into *shared tasks*, where **artificial** story datasets are often (though not always) used for testing a particular technical ability of a system. These datasets are sometimes created and often labeled by crowdworkers, and they include brief scenarios not explicitly connected to broader social contexts and narratives. For example, one of the widely used corpora for testing performance on the Story Cloze task is ROCStories dataset which is a collection of 100,000 crowdsourced "five-sentence commonsense stories" (Mostafazadeh et al., 2016).

Narrative research in NLP also includes *corpus-based* studies, where researchers use narrative models to learn about a particular dataset and its authors. Corpus-based studies depend on **curated** datasets that range widely, e.g., fictional works (e.g., novels, fairytales) (Jans et al., 2012; Iyyer et al., 2016), news stories (Chambers and Jurafsky, 2008), biographies (Bamman and Smith, 2014), and personal stories shared orally or on social media (Gordon and Swanson, 2009; Ouyang and McKeown, 2014; Antoniak et al., 2019). These curated datasets were authored in social contexts separate from the NLP research study and are gathered afterwards. Curated datasets can also be used for shared tasks, e.g., coreference resolution (Bamman et al., 2020), story generation (Akoury et al., 2020).

There are a small number of **naturalistic** NLP narrative datasets that lie outside of the above categories. For example, Sap et al. (2020) collected autobiographical stories and retellings of these stories from crowdworkers; this data was shared as part of the research study but was also grounded in the authors' personal experiences.

And finally, many modern NLP methods for narratives rely on large, pretrained models (Devlin et al., 2019). These models are trained on **massive** and (mostly) undocumented datasets, containing a mixture of documents from unrelated domains to generalize to other domains and tasks (after fine-tuning). These pretraining datasets, like the aptly-named Pile (Gao et al., 2020), are too large for full datasheet descriptions (Gebru et al., 2021) and can encode human biases (Bender et al., 2021).

3 HCI Framings of Narrative

Four key themes are associated with HCI's sensibility of narrative: (a) fact (universal/objective truth) (b) experience (global, local, and day-to-day experiences) (c) interpretation (perceived understanding of a and b) (d) fiction (imaginations and cultural value-system based storytelling) (Bruner, 1990; Sterling, 2009; Golsteijn and Wright, 2013). HCI researchers often ask questions to understand problems better and care about accuracy, legitimacy, and materiality (i.e., why and how certain issues are important) of information. Many subdomains of HCI refer to and build on users' experiences regarding narratives (Feuston and Piper, 2019). HCI practitioners' and designers' interactions with social settings and/or professional environments frequently influence their experiences in a given time and situation, and so they consciously refrain from making generalizable statements and encourage the mention of **contextuality**, which is strongly associated with the narratives (Golsteijn and Wright, 2013). Experience-centered research also values **empathy** to understand the researchers' orientation to the user, and whether they are motivated to empathize with the users' needs and emotional responses (Wright and McCarthy, 2008).

HCI uses both qualitative and quantitative techniques to gather and examine narratives. Both quantitative techniques (e.g., surveys and computational analysis of social media data) and qualitative approaches (e.g., interviews, observation, and focus groups) and artistic techniques are frequent in HCI. In this paper, we focus on qualitative HCI methods for narratives, as they differ from NLP approaches in terms of ontology and epistemology, representing two distinct worldviews (Slevitch, 2011).

In **surveys**, researchers conduct statistical analyses and evaluate the responses based on standard tests and sets of metrics. Using qualitative text coding in cases of free-text responses within the survey is also common. More specifically, HCI scholars have conducted surveys to investigate people's motivations, goals, and challenges with regard to posting narratives on social media (Sannon et al., 2019), as well as factors that influence their decision-making when it comes to online disclosure (Bazarova et al., 2015; Zhang et al., 2021) These approaches are motivated by understanding a specific set of people in a specific setting and social context.

Ethnography, observations, interviews, focus group discussions, and story-telling are common methods used by qualitative HCI researchers (Sultana and Ahmed, 2019). Other than transcription-based qualitative text coding, image, audio and video coding are also used for analysis (Andalibi et al., 2017; Sharma and De Choudhury, 2015). In this regard, not only the texts, images, audio, and videos of the participants are used, but also pictures of the environment, background noise, and even reasons and results of unintended interruptions can contribute to the richness of narratives.

HCI researchers also use curated **social media data** and conduct statistical analyses and qualitative text coding on this data to understand certain research problems. HCI researchers in some domains also use online ethnographic techniques on social media and collaborative platforms (e.g., gaming, e-commerce) to grow deeper understandings of these communities (Mim and Ahmed, 2020).

Many HCI researchers and participants use **artistic techniques** like drawings and installations for addressing research queries (Sturdee et al., 2021). It is quite common for artists to conduct an auto-ethnography with themselves, in which their creation of art is their narrative. In many cases, such narratives are symbolic and contextual.

Finally, many HCI researchers adopt **mixed methods** where they use both qualitative and quantitative approaches to grow a wider and deeper understanding of their research problems. For example, a recent feminist-HCI design used a survey to understand the spread of gender harassment on social media and also conducted interviews and focus group discussions for participatory design and user evaluation of *Unmochon* (Sultana et al., 2021).

4 Interdisciplinary Tensions

Volume and depth of narratives. On one hand, NLP techniques can analyze more narratives (and often more normalized) with reduced researcher workload, though at the loss of qualitative detail. On the other hand, qualitative methods in HCI offer a deeper understanding of narratives based on a more limited sample size. Despite smaller datasets, HCI often depends on theoretical saturation of narratives, in which all important themes are represented, while in NLP, even if the number of datapoints is greater, researchers interested in a particular community often rely on an extracted sample decided by someone else (i.e., *curated* datasets) which might not capture all relevant themes.

4.1 Abstraction and Contextuality

While gaining holistic understandings of an individual in HCI, researchers care about participants' life experiences, social relationships, and observable artifacts surrounding them. In NLP research, it is often impossible to glean such relevant and detailed information from individuals in a large dataset, where abstraction rather than situatedness is the goal. This contrast also pertains to the *agency and privacy* of the narrative sources—whether authors are informed about and consent to the inclusion of their narrative in the dataset for research purposes—as well as the uncertain *representation* of different groups. Narrative data in NLP often lacks explicit context (artificial datasets) or is used out of context (curated and massive datasets); naturalistic datasets generated specifically for the NLP study are more rare (Sap et al., 2020), unlike in HCI. For example, it is common in NLP to scrape data passively that was written in a different context than the research study, as opposed to interview studies in HCI, where researchers explicitly collect stories for the current study. However, NLP datasets are often designed tasks that model abstractions, like common narrative arcs, where simplified datasets can help researchers tackle specific tasks.

4.2 Closed and Open Dataset Design

HCI's emphasis on contextuality opens rather than constrains research possibilities: when narratives are collected in HCI, the emphasis is on high-level and open-ended research goals, focusing on discovering things that have not been explored sufficiently or that might even be in conflict with the researchers' assumptions. Similarly, when labeling themes in narratives, multiple HCI researchers are involved in an open coding process, in which independent coders develop their own themes before combining and refining these themes to ensure the validity and reliability of their interpretations. In contrast, shared narrative tasks in NLP rely on labeled datasets intended for one task, whose data and labels are meant to model one specific concept (like story conclusions) that the researchers already hold, even if multiple annotators are involved.

5 Towards a Mixed Methods Approach

We argue that mixed methods research, drawing from both NLP and HCI, could allow for richer narrative datasets and more holistic understandings of narrative and the social impacts of narrative. This *triangulation* of methods not only minimizes the biases of researchers and enhances the validity of the findings, but also reveals different dimensions of the phenomenon being investigated (McQueen and Knussen, 2002). While prior work has suggested mixed methods in many other NLP contexts (e.g., grounded topic modeling (Baumer et al., 2017)), narratives are particularly well-suited because of the strong research interest on both sides and the tensions enumerated above.

5.1 Customizing an Approach

Prior work has suggested various frameworks to select between mixed methods approaches (Heuer and Buschek, 2021; Inie and Derczynski, 2021). These mixed methods may follow different design patterns, including *explanatory*, *exploratory*, *parallel*, and *nested* methods (Creswell and Clark, 2017). Therefore, the choice of study design should be guided by research questions and goals. For example, if researchers aim to understand the structural patterns of a certain type of narrative (e.g., mental health disclosures on social media) and examine its situatedness (i.e., audiences and context), they might consider an explanatory sequential mixed methods design, where researchers first use quantitative methods to analyze scraped data followed by qualitative interviews and selected narratives in social context. Contrarily, to understand how individuals frame a particular event or phenomenon (e.g., the COVID-19 pandemic) and see if that frame can be applied to a larger population, researchers might opt for an exploratory sequential mixed methods design, characterized by an initial qualitative phase of data collection and analysis, followed by quantitative analysis drawing on a larger dataset.

5.2 Contextuality

Because situatedness or contextuality are essential components of narrative, contextuality can act as a bridging frame in these mixed methods designs, to move between the volume of narrative data in NLP and the depth of analysis in qualitative HCI methods. Researchers can move between "zooming in" on specific stories using qualitative methods and "zooming out" to analyze larger patterns across stories (rather than just one or the other). For example, HCI methods can be used to gather qualitative detail about a dataset's context, while research methods and tools from NLP can help HCI researchers situate smaller datasets within their larger-scale, cross-community contexts (Zhang et al., 2017; Lucy and Bamman, 2021). Both sets of methods can also help address how platform design, moderation, and other contextual features shape the sharing of narratives online.

5.3 Positionality in Design and Evaluation

Qualitative HCI methods emphasize reflexivity and positionality. These practices can encourage NLP researchers to recognize the inherent biases in their research questions, datasets, modeling architectures, procedures, and interpretation of results. Narrative tasks are not simple; each instance usually has multiple right answers, and researchers need to be aware of their own biases in evaluation. For example, when selecting appropriate story endings, annotators are not operating with a "view from nowhere" but from particular values and circumstances (Nagel, 1989). The positionality of the researchers should also be considered in relation to the narrative authors; the authors' positions are often lost in NLP datasets, even when those datasets are labeled with internal states (e.g., sentiment) known only to the authors. Classifying internal states carries risks (Stark, 2018), which are compounded in the study of personal narratives, where affect, relationships, and narratival motivations are intertwined. One strategy to address this challenge is to include the narrative authors in the dataset design (Heuer and Buschek, 2021). And NLP methods can be used to explore the authors' and researchers' positionality by comparing biased linguistic patterns (Bolukbasi et al., 2016; Caliskan et al., 2017) contained in narratives and case notes.

5.4 Openness to Discovery and Disagreement

Discovery and disagreement are central components to the open research focus in qualitative HCI methods. When designing labeled datasets and shared tasks, NLP can adopt HCI's open approach; rather than constraining the data and labels to a test a single technical ability, decided a priori, NLP researchers can take an open approach—one that allows for complicated labels that emerge from multiple annotators' interpretation of the data. This opportunity to include labeler disagreements has been noted by a large body of work (Inie and Derczynski, 2021), but given the complexity of narrative data and its many intertwining features (Piper et al., 2021), customized labels (e.g., hierarchical) could more realistically represent narratives than artificial benchmarks with limited utility. On the other side, NLP methods like topic modeling can help surface themes and discourses that are not immediately apparent to qualitative coders (Baumer et al., 2017). NLP methods can also be used to to identify outlier narratives whose structure or framing is unusual for the dataset (Antoniak et al., 2019).

6 Conclusion

As both a research tool and as an object of study, narrative datasets have been widely used in both NLP and HCI. This short work is not intended to describe all approaches to narrative in these disciplines, nor is it intended to provide solutions to all the described challenges. Boundaries between disciplines are fluid, especially in regards to stories shared on social media, where platform design, moderation, and many other HCI concepts have shaped the stories studied via computational NLP methods. Many different fields (e.g., literary studies) are concerned with narratives; we have constrained our discussion to datasets in NLP and qualitative HCI because we see room for cross-pollination and conversations. Stories can be powerful tools of persuasion and expression, and richer methods that draw from both NLP and HCI can raise new questions and open up new directions.

7 Acknowledgements

Thank you to our anonymous reviewers, whose comments were very helpful in preparing this paper. We also thank Sharifa Sultana's Facebook Fellowship for supporting this work.

References

Nader Akoury, Shufan Wang, Josh Whiting, Stephen Hood, Nanyun Peng, and Mohit Iyyer. 2020. STORIUM: A Dataset and Evaluation Platform for Machine-in-the-Loop Story Generation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 6470–6484, Online. Association for Computational Linguistics.

Nazanin Andalibi, Pinar Ozturk, and Andrea Forte. 2017. Sensitive self-disclosures, responses, and social support on instagram: The case of depression. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, page 1485–1500, New York, NY, USA. Association for Computing Machinery.

Maria Antoniak, David Mimno, and Karen Levy. 2019. Narrative paths and negotiation of power in birth stories. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW).

Michael Bamberg and Alexandra Georgakopoulou. 2008. Small stories as a new perspective in narrative and identity analysis.

David Bamman, Olivia Lewke, and Anya Mansoor. 2020. An annotated dataset of coreference in English literature. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 44–54, Marseille, France. European Language Resources Association.

David Bamman, Brendan O'Connor, and Noah A. Smith. 2013. Learning latent personas of film characters. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 352–361, Sofia, Bulgaria. Association for Computational Linguistics.

David Bamman and Noah A. Smith. 2014. Unsupervised discovery of biographical structure from text. *Transactions of the Association for Computational Linguistics*, 2:363–376.

Jeffrey Bardzell, Shaowen Bardzell, Carl DiSalvo, William Gaver, and Phoebe Sengers. 2012. The humanities and/in hci. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 1135–1138.

Eric PS Baumer, David Mimno, Shion Guha, Emily Quan, and Geri K Gay. 2017. Comparing grounded theory and topic modeling: Extreme divergence or unlikely convergence? *Journal of the Association for Information Science and Technology*, 68(6):1397–1410.

Natalya N Bazarova, Yoon Hyung Choi, Victoria Schwanda Sosik, Dan Cosley, and Janis Whitlock. 2015. Social sharing of emotions on facebook: Channel differences, satisfaction, and replies. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, pages 154–164.

Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21, page 610–623, New York, NY, USA. Association for Computing Machinery.

Tolga Bolukbasi, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam T Kalai. 2016. Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In *Advances in Neural Information Processing Systems*, pages 4349–4357.

Jerome Bruner. 1987. Life as narrative. *Social research*, pages 11–32.

Jerome Bruner. 1990. Acts of meaning.

Aylin Caliskan, Joanna J Bryson, and Arvind Narayanan. 2017. Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334):183–186.

Nathanael Chambers and Dan Jurafsky. 2008. Unsupervised learning of narrative event chains. In *Proceedings of ACL-08: HLT*, pages 789–797, Columbus, Ohio. Association for Computational Linguistics.

Nathanael Chambers and Dan Jurafsky. 2009. Unsupervised learning of narrative schemas and their participants. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 602–610, Suntec, Singapore. Association for Computational Linguistics.

John W Creswell and Vicki L Plano Clark. 2017. *Designing and conducting mixed methods research*. Sage Publications.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Jessica L Feuston and Anne Marie Piper. 2019. Everyday experiences: small stories and mental illness on instagram. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–14.

Leo Gao, Stella Biderman, Sid Black, Laurence Golding, Travis Hoppe, Charles Foster, Jason Phang, Horace He, Anish Thite, Noa Nabeshima, et al. 2020. The pile: An 800gb dataset of diverse text for language modeling. *arXiv preprint arXiv:2101.00027*.

Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2021. Datasheets for datasets. *Commun. ACM*, 64(12):86–92.

Seraphina Goldfarb-Tarrant, Tuhin Chakrabarty, Ralph Weischedel, and Nanyun Peng. 2020. Content planning for neural story generation with aristotelian rescoring. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 4319–4338, Online. Association for Computational Linguistics.

Connie Golsteijn and Serena Wright. 2013. Using narrative research and portraiture to inform design research. In *IFIP Conference on Human-Computer Interaction*, pages 298–315. Springer.

Andrew Gordon and Reid Swanson. 2009. Identifying personal stories in millions of weblog entries. In *Third International Conference on Weblogs and Social Media, Data Challenge Workshop, San Jose, CA*, volume 46.

Amit Goyal, Ellen Riloff, and Hal Daumé III. 2010. Automatically producing plot unit representations for narrative text. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 77–86, Cambridge, MA. Association for Computational Linguistics.

Jian Guan, Fei Huang, Zhihao Zhao, Xiaoyan Zhu, and Minlie Huang. 2020. A knowledge-enhanced pretraining model for commonsense story generation. *Transactions of the Association for Computational Linguistics*, 8:93–108.

Hendrik Heuer and Daniel Buschek. 2021. Methods for the design and evaluation of HCI+NLP systems. In *Proceedings of the First Workshop on Bridging Human–Computer Interaction and Natural Language Processing*, pages 28–33, Online. Association for Computational Linguistics.

Nanna Inie and Leon Derczynski. 2021. An IDR framework of opportunities and barriers between HCI and NLP. In *Proceedings of the First Workshop on Bridging Human–Computer Interaction and Natural Language Processing*, pages 101–108, Online. Association for Computational Linguistics.

Mohit Iyyer, Anupam Guha, Snigdha Chaturvedi, Jordan Boyd-Graber, and Hal Daumé III. 2016. Feuding families and former Friends: Unsupervised learning for dynamic fictional relationships. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1534–1544, San Diego, California. Association for Computational Linguistics.

Bram Jans, Steven Bethard, Ivan Vulić, and Marie Francine Moens. 2012. Skip n-grams and ranking functions for predicting script events. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 336–344, Avignon, France. Association for Computational Linguistics.

Wendy G. Lehnert. 1981. Plot units and narrative summarization. *Cognitive Science*, 5:293–331.

Li Lucy and David Bamman. 2021. Characterizing English variation across social media communities with BERT. *Transactions of the Association for Computational Linguistics*, 9:538–556.

Stephanie Lukin, Kevin Bowden, Casey Barackman, and Marilyn Walker. 2016. PersonaBank: A corpus of personal narratives and their story intention graphs. pages 1026–1033.

Ronald A McQueen and Christina Knussen. 2002. *Research methods for social science: A practical introduction*. Pearson Education.

Erin L Merz, Rina S Fox, and Vanessa L Malcarne. 2014. Expressive writing interventions in cancer patients: A systematic review. *Health Psychology Review*, 8(3):339–361.

Nusrat Jahan Mim and Syed Ishtiaque Ahmed. 2020. Others' images: Online social media, architectural improvisations, and spatial marginalization in bangladesh. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14.

Nasrin Mostafazadeh, Nathanael Chambers, Xiaodong He, Devi Parikh, Dhruv Batra, Lucy Vanderwende, Pushmeet Kohli, and James Allen. 2016. A corpus and cloze evaluation for deeper understanding of commonsense stories. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 839–849, San Diego, California. Association for Computational Linguistics.

Thomas Nagel. 1989. *The view from nowhere*. oxford university press.

Pok-Ja Oh and Soo Hyun Kim. 2016. The effects of expressive writing interventions for patients with cancer: A meta-analysis. In *Oncology Nursing Forum*, volume 43.

Jessica Ouyang and Kathy McKeown. 2014. Towards automatic detection of narrative structure. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*, pages 4624–4631, Reykjavik, Iceland. European Languages Resources Association (ELRA).

James W Pennebaker. 1997. Writing about emotional experiences as a therapeutic process. *Psychological Science*, 8(3):162–166.

James W Pennebaker and Sandra K Beall. 1986. Confronting a traumatic event: toward an understanding of inhibition and disease. *Journal of Abnormal Psychology*, 95(3):274.

Andrew Piper, Richard Jean So, and David Bamman. 2021. Narrative theory for computational narrative understanding. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 298–311, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Andrew J. Reagan, Lewis Mitchell, Dilan Kiley, Christopher M. Danforth, and Peter Sheridan Dodds. 2016. The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, 5:1–12.

Matthew Richardson, Christopher J.C. Burges, and Erin Renshaw. 2013. MCTest: A challenge dataset for the open-domain machine comprehension of text. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 193–203, Seattle, Washington, USA. Association for Computational Linguistics.

Shruti Sannon, Elizabeth L Murnane, Natalya N Bazarova, and Geri Gay. 2019. " i was really, really nervous posting it" communicating about invisible chronic illnesses across social media platforms. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–13.

Maarten Sap, Eric Horvitz, Yejin Choi, Noah A. Smith, and James Pennebaker. 2020. Recollection versus imagination: Exploring human memory and cognition via neural language models. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 1970–1978, Online. Association for Computational Linguistics.

Roger C Schank and Robert P Abelson. 1977. Scripts, plans, goals, and understanding: An inquiry into human knowledge structures.

Sanket S Sharma and Munmun De Choudhury. 2015. Measuring and characterizing nutritional information of food and ingestion content in instagram. In *Proceedings of the 24th International Conference on World Wide Web*, pages 115–116.

Lisa Slevitch. 2011. Qualitative and quantitative methodologies compared: Ontological and epistemological perspectives. *Journal of quality assurance in hospitality & tourism*, 12(1):73–81.

Luke Stark. 2018. Algorithmic psychometrics and the scalable subject. *Social Studies of Science*, 48(2):204–231. PMID: 29726810.

Bruce Sterling. 2009. Cover story design fiction. *Interactions*, 16(3):20–24.

Miriam Sturdee, Makayla Lewis, Angelika Strohmayer, Katta Spiel, Nantia Koulidou, Sarah Fdili Alaoui, and Josh Urban Davis. 2021. A plurality of practices: artistic narratives in hci research. In *Creativity and Cognition*, pages 1–1.

Sharifa Sultana and Syed Ishtiaque Ahmed. 2019. Witchcraft and hci: Morality, modernity, and postcolonial computing in rural bangladesh. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–15.

Sharifa Sultana, Mitrasree Deb, Ananya Bhattacharjee, Shaid Hasan, SM Raihanul Alam, Trishna Chakraborty, Prianka Roy, Samira Fairuz Ahmed, Aparna Moitra, M Ashraful Amin, et al. 2021. 'un-mochon': A tool to combat online sexual harassment over facebook messenger. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–18.

Timothy R Tangherlini. 2000. Heroes and lies: Storytelling tactics among paramedics. *Folklore*, 111(1):43–66.

Gareth Williams. 1984. The genesis of chronic illness: narrative re-construction. *Sociology of health & illness*, 6(2):175–200.

Peter Wright and John McCarthy. 2004. *Technology as experience*. MIT Press Cambridge.

Peter Wright and John McCarthy. 2008. Empathy and experience in hci. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 637–646.

Justine Zhang, William Hamilton, Cristian Danescu-Niculescu-Mizil, Dan Jurafsky, and Jure Leskovec. 2017. Community identity and user engagement in a multi-community landscape. *Proceedings of the International AAAI Conference on Web and Social Media*, 11(1):377–386.

Renwen Zhang, Natalya N. Bazarova, and Madhu Reddy. 2021. Distress disclosure across social media platforms during the covid-19 pandemic: Untangling the effects of platforms, affordances, and audiences. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–15.