

Human Values in Stories: Narrative-informed Directions for Value Discovery

Alexandra Barancová, Giovanni Sileno, Sennay Ghebreab
Informatics Institute, University of Amsterdam, The Netherlands
{a.barancova, g.sileno, s.ghebreab}@uva.nl

Abstract

This paper examines ways in which narrative inquiry might inform natural language processing methods to explore human values. Considering the centrality of storytelling to human sense-making and the far reach of stories across today’s digital infrastructures, we argue that a narrative lens could be fruitful for computational approaches to value analysis. In this paper, we bring together existing work on narrative understanding and value detection, identifying and elaborating on three focal points for NLP task design: narratives, corpora, and audiences.

1 Introduction

Narratives surround us, shaping what we find important and how we relate to each other and to the world. Although this is not new, recent years have seen an increase in the sheer amount and reach of stories, a so-called ‘story boom,’ manifested by the online story economy, the emergence of story consultants, or all the places—online and offline—where you are prompted to “share your story” (Mäkelä and Meretoja, 2022). Research too has witnessed a ‘narrative turn,’ establishing the relevance and urgency of studying narratives beyond the domain of fiction, into matters of public and ethical concern such as values (Korthals Altes and Moenandar, 2024). This attention stems from real-world phenomena that narratives co-shape, like misinformation or polarisation, and from the sociotechnical conditions they circulate in, such as platforms or machine learning datasets.

This contribution revolves around the question: *in what ways can narrative inquiry inform and guide NLP methods for exploring human values?* With this short paper, we (1) argue that a narrative approach could be a fruitful direction for NLP work on human values expressed through natural language, and (2) identify several starting points enabling this line of research. In doing so, we connect several existing research strands and offer a

conceptual foundation for further developments. These two objectives are closely linked to the issue of positive impact; a narrative approach can contribute to a more nuanced and contextualised understanding of human values. At the same time, it is important to acknowledge that value interpretation, much like intention recognition, may be abused in certain circumstances.¹ Nevertheless, shared notions of public good and democratic systems require a diversity of views and positions in society, as those expressed in narratives, to be given adequate visibility. With this aim, we complement and expand the existing body of NLP work that examines normative concerns through concepts such as bias, fairness, or representational harm.

1.1 Research gap(s)

Although NLP research has begun to address ethical concepts, the breadth and depth of this area of study remains limited (Blodgett et al., 2020; Vida et al., 2023), including when it comes to dealing with values. For one, this gap presents in terms of limited linguistic forms that are studied. For instance, to date many NLP approaches to detecting and identifying human values have come from the tradition of argument mining (Kiesel et al., 2022).² While this has catalysed attention to values, these approaches are predominantly centred around a rationalised, justification-focused framing thereof; data-wise this translates to language that comes in premise-stance-conclusion formats. Although still fundamental, this leaves a large part of how people express themselves and shape their views untouched—meandering, in-talk, at times contradictory, based on personal stories and accounts of lived and indirect experience.³

¹We address this further in the Ethical and societal considerations Section.

²See for example entries submitted to the Human Value Detection task (Kiesel et al., 2022) in 2023 and 2024.

³In spontaneous conversations, narratives may occupy up to 40% of conversation time (Eggs and Slade, 1997).

Second, much existing work on values is based on psychology-inspired frameworks, such as Schwartz’s (2012) theory of personal values and the Moral Foundations Theory (Graham et al., 2013), which, in practice, have largely been operationalised as closed taxonomy lists (a multiclass classification problem). This leaves little room for any domain-tailored or situated understanding of values, called for by e.g. empirical ethics (Helm and Gerlek, 2026), data feminism (D’Ignazio and Klein, 2020), or even “neighbouring” disciplines like HCI that have long stressed the need for open-ended value discovery (Le Dantec et al., 2009).

The ways in which we express and develop our moral stances make use of a broad set of dialogical tools and processes (see e.g. Boyd et al., 2021). To account for this, it is crucial that NLP research considers this diversity. Guided by the evocative metaphor of values being ‘storied’ (JafariNaimi et al., 2015)⁴ and given the rich field of narrative ethics (see Section 2.3), we see narrative analysis as one promising avenue forward. While computational narrative understanding (see Section 2.1) has established a practice of narrative framing in NLP, we see potential in extending this to questions surrounding human values; engaging not only with who and what stories are about (their entities, character relations, event chains, emotional arcs, etc.), but also what the normative implications of what it said or written might be.

1.2 Definitions

Human values Values are a concept that has long been studied in a range of domains and described using various definitions, frameworks, and corresponding measures. In terms of function, many definitions converge around the idea that ‘values serve as guiding principles of what people consider important in life’ (Cheng and Fleischmann, 2010: 2).⁵ In discourse, values are typically anchored to abstract concepts—such as ‘care’—which require contextual situating to acquire concrete meaning. For our working definition, we build on Friedman’s (1997) approach that emphasises values dealing with human welfare and justice. While distinctions between, e.g., moral, conventional, and personal domains of social knowledge have been used to

demarcate values in psychology, this framing opts for the ‘broader term “human values” instead of simply moral values to highlight the complexity of social life and to provide the basis for analyses wherein personal and conventional values can become morally implicated’ (Friedman, 1997: 5). Values, then, are defined jointly with value dimensions deemed relevant by the observer or speaker/writer. Situated events either promote or demote certain values, intentionally or not.

Narratives Although with varying prerequisites, many narrative theorists concur that events can be seen as the fundamental units of narrative (Abbott, 2021). In its minimal form, a narrative can therefore be understood as a representation of one or multiple events. Scholars like Ryan (2001) further suggest that narrativity should be treated as a “fuzzy set”, assessing the presence of various features as a question of degree. Importantly, narratives too are situated; along with the *story* that the events illustrate, they are always co-constituted by manner in which they are told, the narrative *discourse*, shaped by the context or conversation.⁶ Narratives are mutable and take on many forms: spoken, written, texted, filmed, streamed, drawn, but also linear, branching, immersive, interactive.

2 Background

Here, we bring together a slice of existing research grouped by the three pairwise intersections of narratives, computational methods, and human values.

2.1 Studying narratives at scale

Coming up with ways to study narratives at scale has been the concern of the growing field of computational narrative understanding (see, e.g., Piper, 2023 for an overview). Much of the work in this domain developed from structural approaches to narrative (e.g. Barthes and Duisit, 1975; Greimas and Schleifer, 1983) and domains like narratology (e.g. Genette, 1983; Bal and Boheemen, 2009), which frame narrative in terms of functional and formalised units that lend themselves to modelling.

⁶Broken down into an elementary level, a narrative can be seen as an instance when ‘[s]omeone tells someone somewhere that someone did something(s) [to someone] somewhere at some time for some reason’ (Piper, 2023: 3), where the first subclause refers to the components that make up the frame of telling (heterodiegetic, in narratological terms), and what follows is in the frame of the storyworld (diegetic). These elements can be framed as NLP tasks. For instance, the very first ‘someone’ could be studied through point of view identification or dialogue attribution.

⁴Following this thesis, ‘values function as hypotheses to draw out what the situation is, and what action it demands’ (JafariNaimi et al., 2015: 103).

⁵This is based on a meta-inventory by Cheng and Fleischmann (2010), which maps the theories and definitions associated with values across various disciplines.

While alone this line of work may well be challenged as being somewhat closed and prescriptive, increasingly various complementary “bottom up” approaches have emerged. For instance, unsupervised methods can help surface domain-specific narrative features, e.g., inferring canonical character tropes based on large corpora of film summaries (Bamman et al., 2013). Participatory methods foreground the diversity of human perception when it comes to the interpretation of narratives, e.g., demonstrating that non-experts pay attention to different things than literary scholars when judging the similarity of two stories (Nguyen et al., 2014). Similarly, emphasis is increasingly being placed on looking to a wider range of sites in which storytelling takes place—beyond literature—to daily exchanges in online communities (Antoniak et al., 2024) as well as through different media, like memes (Fharook et al., 2022).

2.2 Computational, language-focused approaches to values

Previous work has shown that values lend themselves to study through natural language, offering a complementary—and for some purposes perhaps arguably better (Boyd et al., 2021)—approach to established methods such as self-report questionnaires. In line with this intuition, tasks like value detection or elicitation have been approached at scale using, e.g., topic modelling (Boyd et al., 2021), BERT-based models (Kiesel et al., 2022), and increasingly various generative LLMs (Ye et al., 2025). This has been accompanied by hand-annotated datasets, e.g., the Human Value Detection dataset (Kiesel et al., 2022) or the Moral Foundations Reddit Corpus (Trager et al., 2022).⁷

Alongside access to human written and spoken language data, recent advances in language models with corresponding efforts to understand, align, and monitor these models have formed a substantial chunk of NLP work on values. In this area, one line of work focuses on inferring value judgments and moral stances of LLMs based on generated textual outputs, often adopting questionnaires and taxonomies developed for individual human assessments, like the Moral Foundations Theory (e.g. Bulla et al., 2024) or Political Compass Test (e.g. Wright et al., 2024). As Röttger et al. (2024) show, such approaches are still far from robust, with models demonstrating substantially different

stances depending on slight changes in formulation or format.

Turning to narratives, although computational narrative research has not, to the best of our knowledge, explicitly focused on human values as such, we nevertheless do find examples that clearly engage normative concerns. On the one hand, this involves narratives as the object of study; data that are narrative in form. Work on story morals, for example, has explored narratives from a range of genres including including folktales, novels, films, personal stories from social media and news items (Hobson et al., 2024; Piper, 2025). On the other hand, we also see examples of narrative schemas used for (value-laden) NLP tasks. For example, Sourati et al. (2024) approach analogy detection on a dataset of texts grouped by proverbs using a breakdown of the source stories into narrative components. Antoniak et al. (2019) map power relations in a corpus of personal accounts of giving birth, which can be seen as an example combining both narrative data and analysis.

2.3 Narrative and ethics

Despite existing normative work in NLP, the interconnection between narrative and ethics has rarely been developed in this field. However, in the humanities (and some social sciences), narrative ethics has been extensively theorised and examined. Within the hermeneutic tradition for example, Kearney and Williams (1996) identified three areas for the ethical potential of narrative: *testimonial* (bear witness to the past, remember), *empathic* (identify with others), *critical/utopian* (challenge official stories, present counter-narratives).⁸ This foregrounds at least three different corresponding ways in which computational work that integrates narratives—be that as data, method, output, or a combination—is ethically implicated.

Importantly, work on the ethics of storytelling points out that narratives shape what we value and highlights the interconnectedness of stories (Meretoja, 2018).⁹ The notion of ‘collective narratives’ mentioned above likewise connects individual narratives to more aggregated forms thereof (similar to those others have also termed e.g., ‘deep stories’ or ‘meta-narratives’), foregrounding the political and normative function that they can come to assume.

⁸Kearney and Williams (1996) build on the work of scholars like Paul Ricoeur, Hannah Arendt, Martha Nussbaum.

⁹As Meretoja (2018) puts it, ‘[t]he stories we tell are never entirely our own’ (p.18).

⁷Annotation is currently mainly either done by small teams of trained annotators or outsourced to crowd workers.

Addressing the importance of collective narratives for values and beliefs about social reality, [Bliuc and Chidley \(2022\)](#) offer a typology of functions that they can have for groups of people, depending on whether these are shared within or between groups and ultimately tend to consensus (unifying, empowering) or dissent (divisive, factionalizing). Such models provide a glimpse of both the ethical potential and importance of narratives, including in relation to values.

3 Towards narrative-informed computational methods

The ethical implications of narrative, briefly introduced above, set a rich stage for computational work on both narrative understanding and human values. Notably, narratives offer one way to bridge the gap between what people think and what people do, as stories often weave aspects of both together.

Building on narrative theory, we plot a few focal points that can help us move towards scalable narrative-inspired approaches to value discovery. We illustrate each of these with concrete questions and tasks for a guiding application: studying the values discussed in online communities using social media data.¹⁰ The overall challenge invites a holistic and situated approach to natural language corpora, building on the various dimensions relevant to narrativity; this includes the content, the context, and the conversation.

Narrative instances The first level of analysis concerns the stories themselves. For a large part, this extends what the examples included in Section 2.1 are already doing. For example, many structural models used to analyse narratives that focus on character roles (e.g. Hero, Villain, Victim, etc.) or Greimas' actants are fundamentally axiological ([Korthals Altes and Moenandar, 2024](#)); directing our attention to how and why certain entities fit these roles within a narrative can be a fruitful way to understand the values that shape a story. Crucially, as the focus moves to examining storytelling in domains that do not necessarily follow literary conventions, new models may be needed. Small stories research ([Georgakopoulou, 2006](#)), for instance, could lend itself to dealing with posts and comments exchanged in online communities.

¹⁰We chose this starting point given existing NLP research and datasets concerning both values (e.g. [Trager et al., 2022](#)) and narratives (e.g. [Antoniak et al., 2024](#)) in this domain.

Corpora The broader context within (/against) which individual narratives sit is shaped by collective narratives. Although this idea has been widely echoed (see Section 2.3 for several examples), there are few methodological pointers to study the phenomenon of so-called meta-narratives. NLP could help tackle this empirically; for example in the form of methods for aggregating individual narratives, identifying patterns such as recurring event chains, developing macrostrategies for coding different points of view, or finding characteristic turning points over time. This could bring into view the values underpinning dominant narratives in particular corpora, as well as those intended to challenge them.

Audiences Often the glue between narratives and collections rests in audiences; be they imagined, lurking, engaged, oppositional, or involved, social relationships are integral to narratives. As [Bamberg \(2006\)](#) puts it, it is 'in and through talk' that speakers establish both the content and form of social interaction that come together in a (narrative) engagement. This third direction orients us to understanding stories as parts of unfolding conversations, a process that simultaneously shapes the values at stake. Therefore, in addition to narrative features, structures, conventions, etc., it is important to examine the (social) dynamics that accompany them; who listens, reads and reacts? How are meaning and value attributed, shaping the interpretation ([Korthals Altes and Moenandar, 2024](#))? This invites an investigation of, e.g., emotional arcs unfolding across conversations, as well as within singular narratives; a consideration of metadata that might hint at reactions triggered by data points in a corpus; or mapping how particular or prototypical stories travel and spread within or between groups.

4 Conclusion and future work

This paper reorganises existing insights from narrative ethics, narratology, and computational narrative understanding as a departure point for developing computer-supported approaches to value discovery. We argue that a narrative framing can be fruitful for NLP work aiming to explore human values and identify three focal points relevant for task definition in this domain: series of (individual) narratives, (aggregated) corpora, and (interactions of) audiences. New analytical tools developed along these lines would support a deeper social understanding than current methods.

Future work could examine what aspects of narrativity tend to be value-laden, or whether certain narrative features convey value judgments more often than others. Another line of study concerns the dynamics exhibited in discourse, linking individual narratives to evolving positions of different stakeholders in conversations. Importantly, we see value frameworks as neither fixed nor universally applicable, so determining which one(s) to use will remain a context-dependent choice.

From a practical standpoint, future works following the proposed research lines could focus, for one, on conceptual aspects; e.g., introducing dedicated schemas for value elicitation in narrative instances and drafting associated codebooks to guide annotation. Second, on new resources, which will need to be assembled and tested; e.g., curating datasets to operationalise the identification of meta-narratives, or studying recurrent narrative patterns that emerge from the lived experiences of a particular group of people. Third, on the various associated methodological challenges; e.g., integrating the analysis of narrative components with network studies, such as opinion dynamics, or finding ways to capture (dis)similarity, dissonance, and flag ambiguities, which are inevitable when working with subjective, multifaceted concepts like human values.

Ethical and societal considerations

In terms of ethical and societal considerations, it is crucial to first stress that narrative communication poses its own risks; while it can inform, it can equally misinform audiences, often in compelling, immersive, or difficult to falsify/assess ways (Piper, 2023). For this reason, it is important to remain critical when studying narratives (Mäkelä et al., 2021; Mäkelä and Meretoja, 2022). In the process of modelling human values through a narrative lens, it is also important to remain vigilant of the risk that narratives can become subsumptive (Meretoja, 2023), as a result of which that they can overshadow or suppress other voices and positions.

When reaching for narrative as an analytical approach, the notion of meta-hermeneutics may also be helpful; that is, a reflection on the values that are imbued in the interpretative layer that goes hand in hand with the analytical process (Korthals Altes and Moenandar, 2024). This type of self-reflexive work could be particularly useful in NLP contexts, where as e.g. Birhane et al. (2022) show, engagement with the assumptions involved in the research

process itself are currently limited.

Finally, we believe that it would not be desirable to replace or “automate” qualitative research in this area; the two should remain complementary, as both are needed to address values in a nuanced and situated manner. Similarly, audits, human validation, case studies, bias mitigation and assessment remain crucial, and perhaps even more important if NLP methods are to move into the realm of normative assessment. Representational harm, discrimination, erasure, and dual use are but a few of the dangers that such methods can bring with them (Blodgett et al., 2020; Hovy and Prabhumoye, 2021). This area of research is no exception.

References

- H Porter Abbott. 2021. *The Cambridge introduction to narrative*. Cambridge University Press.
- Maria Antoniak, David Mimno, and Karen Levy. 2019. [Narrative Paths and Negotiation of Power in Birth Stories](#). *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–27.
- Maria Antoniak, Joel Mire, Maarten Sap, Elliott Ash, and Andrew Piper. 2024. [Where Do People Tell Stories Online? Story Detection Across Online Communities](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7104–7130, Bangkok, Thailand. Association for Computational Linguistics.
- Mieke Bal and Christine van Boheemen. 2009. *Narratology: introduction to the theory of narrative*, 3rd ed edition. University of Toronto Press, Toronto.
- Michael Bamberg. 2006. [Stories: Big or small: Why do we care?](#) *Narrative Inquiry*, 16(1):139–147.
- 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.
- Roland Barthes and Lionel Duisit. 1975. [An Introduction to the Structural Analysis of Narrative](#). *New Literary History*, 6(2):237.
- Abeba Birhane, Pratyusha Kalluri, Dallas Card, William Agnew, Ravit Dotan, and Michelle Bao. 2022. [The Values Encoded in Machine Learning Research](#). *arXiv preprint*.
- Ana-Maria Bluc and Alexander Chidley. 2022. [From cooperation to conflict: The role of collective narratives in shaping group behaviour](#). *Social and Personality Psychology Compass*, 16(7):e12670.

- Su Lin Blodgett, Solon Barocas, Hal Daumé Iii, and Hanna Wallach. 2020. [Language \(Technology\) is Power: A Critical Survey of “Bias” in NLP](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5454–5476, Online. Association for Computational Linguistics.
- Ryan Boyd, Steven Wilson, James Pennebaker, Michal Kosinski, David Stillwell, and Rada Mihalcea. 2021. [Values in Words: Using Language to Evaluate and Understand Personal Values](#). *Proceedings of the International AAAI Conference on Web and Social Media*, 9(1):31–40.
- Luana Bulla, Aldo Gangemi, and Misael Mongiovì. 2024. [Do Language Models Understand Morality? Towards a Robust Detection of Moral Content](#). In Nardine Osman and Luc Steels, editors, *Value Engineering in Artificial Intelligence*, volume 14520, pages 98–113. Springer Nature Switzerland, Cham.
- An-Shou Cheng and Kenneth R. Fleischmann. 2010. [Developing a meta-inventory of human values](#). *Proceedings of the American Society for Information Science and Technology*, 47(1):1–10.
- Catherine D’Ignazio and Lauren F. Klein. 2020. *Data Feminism*. Strong Ideas. The MIT Press, Cambridge.
- Suzanne Eggins and Diana Slade. 1997. *Analysing Casual Conversation*. Equinox, London.
- Shaik Fharook, Syed Sufyan Ahmed, Gurram Rithika, Sumith Sai Budde, Sunil Saumya, and Shankar Bilaradar. 2022. Are you a hero or a villain? a semantic role labelling approach for detecting harmful memes. In *Proceedings of the workshop on combating online hostile posts in regional languages during emergency situations*, pages 19–23.
- Batya Friedman, editor. 1997. *Human values and the design of computer technology*. Number no. 72 in CSLI lecture notes. CSLI Publications ; Cambridge University Press, Stanford, Calif. : Cambridge ; New York.
- Gérard Genette. 1983. Narrative discourse: An essay in method. *Cornell University Press*, 3.
- Alexandra Georgakopoulou. 2006. [Thinking big with small stories in narrative and identity analysis](#). *Narrative Inquiry*, 16(1):122–130.
- Jesse Graham, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean P. Wojcik, and Peter H. Ditto. 2013. [Moral Foundations Theory](#). In *Advances in Experimental Social Psychology*, volume 47, pages 55–130. Elsevier.
- Algirdas Julien Greimas and Ronald Schleifer. 1983. *Structural semantics: an attempt at a method*. University of Nebraska Press, Lincoln London.
- Paula Helm and Selin Gerlek. 2026. [Empirical AI Ethics: Grounded, relational, transformative](#). *Cambridge Forum on AI: Culture and Society*, 2:e13.
- David G Hobson, Haiqi Zhou, Derek Ruths, and Andrew Piper. 2024. [Story Morals: Surfacing value-driven narrative schemas using large language models](#). In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 12998–13032, Miami, Florida, USA. Association for Computational Linguistics.
- Dirk Hovy and Shrimai Prabhumoye. 2021. [Five sources of bias in natural language processing](#). *Language and Linguistics Compass*, 15(8):e12432.
- Nassim JafariNaimi, Lisa Nathan, and Ian Hargraves. 2015. [Values as Hypotheses: Design, Inquiry, and the Service of Values](#). *Design Issues*, 31(4):91–104.
- Richard Kearney and James Williams. 1996. [Narrative and Ethics](#). *Proceedings of the Aristotelian Society, Supplementary Volumes*, 70:29–61.
- Johannes Kiesel, Milad Alshomary, Nicolas Handke, Xiaoni Cai, Henning Wachsmuth, and Benno Stein. 2022. [Identifying the Human Values behind Arguments](#). In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4459–4471, Dublin, Ireland. Association for Computational Linguistics.
- Liesbeth Korthals Altes and Sjoerd-Jeroen Moenandar. 2024. [By Way of Introduction – Reflections on Narrative and Values, and the Value of Narratives](#). In Sjoerd-Jeroen Moenandar and Barend Van Heusden, editors, *Narrative Values, the Value of Narratives*, volume 91 of *Narratologia*, pages 1–21. De Gruyter.
- Christopher A. Le Dantec, Erika Shehan Poole, and Susan P. Wyche. 2009. [Values as lived experience: evolving value sensitive design in support of value discovery](#). In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1141–1150, Boston MA USA. ACM.
- Hanna Meretoja. 2018. *The ethics of storytelling: Narrative hermeneutics, history, and the possible*. Oxford University Press.
- Hanna Meretoja. 2023. [Implicit narratives and narrative agency: Evaluating pandemic storytelling](#). *Narrative Inquiry*, 33(2):288–316.
- Maria Mäkelä, Samuli Björninen, Laura Karttunen, Matias Nurminen, Juha Raipola, and Tytti Rantanen. 2021. [Dangers of Narrative: A Critical Approach to Narratives of Personal Experience in Contemporary Story Economy](#). *Narrative*, 29(2):139–159.
- Maria Mäkelä and Hanna Meretoja. 2022. [Critical Approaches to the Storytelling Boom](#). *Poetics Today*, 43(2):191–218.
- Dong Nguyen, Dolf Trieschnigg, and Mariët Theune. 2014. Using crowdsourcing to investigate perception of narrative similarity. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pages 321–330.

- Andrew Piper. 2023. [Computational Narrative Understanding: A Big Picture Analysis](#). In *Proceedings of the Big Picture Workshop*, pages 28–39, Singapore, Singapore. Association for Computational Linguistics.
- Andrew Piper. 2025. [Towards a Perspectival Moral History of the Novel Using LLMs](#). *Journal of Computational Literary Studies*, 4(1).
- Marie-Laure Ryan. 2001. *Narrative as virtual reality: Immersion and interactivity in literature and electronic media*. Parallax.
- Paul Röttger, Valentin Hofmann, Valentina Pyatkin, Musashi Hinck, Hannah Kirk, Hinrich Schuetze, and Dirk Hovy. 2024. [Political Compass or Spinning Arrow? Towards More Meaningful Evaluations for Values and Opinions in Large Language Models](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15295–15311, Bangkok, Thailand. Association for Computational Linguistics.
- Shalom H. Schwartz. 2012. [An Overview of the Schwartz Theory of Basic Values](#). *Online Readings in Psychology and Culture*, 2(1).
- Zhivar Sourati, Filip Ilievski, Pia Sommerauer, and Yifan Jiang. 2024. [ARN: Analogical Reasoning on Narratives](#). *Transactions of the Association for Computational Linguistics*, 12:1063–1086.
- Jackson Trager, Alireza S. Ziabari, Aida Mostafazadeh Davani, Preni Golazizian, Farzan Karimi-Malekabadi, Ali Omrani, Zhihe Li, Brendan Kennedy, Nils Karl Reimer, Melissa Reyes, Kelsey Cheng, Mellow Wei, Christina Merrifield, Arta Khosravi, Evans Alvarez, and Morteza Dehghani. 2022. [The Moral Foundations Reddit Corpus](#). *arXiv preprint*.
- Karina Vida, Judith Simon, and Anne Lauscher. 2023. [Values, Ethics, Morals? On the Use of Moral Concepts in NLP Research](#). In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 5534–5554, Singapore. Association for Computational Linguistics.
- Dustin Wright, Arnav Arora, Nadav Borenstein, Srishti Yadav, Serge Belongie, and Isabelle Augenstein. 2024. [LLM Tropes: Revealing Fine-Grained Values and Opinions in Large Language Models](#). In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 17085–17112, Miami, Florida, USA. Association for Computational Linguistics.
- Haoran Ye, Yuhang Xie, Yuanyi Ren, Hanjun Fang, Xin Zhang, and Guojie Song. 2025. [Measuring Human and AI Values Based on Generative Psychometrics with Large Language Models](#). *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(25):26400–26408.