

Next Steps in Modeling Social Media: The Importance of Narrative

Mark A. Finlayson

Florida International University
Knight Foundation School of Computing and Information Sciences
11200 S.W. 8th Street, CASE Building Room 362, Miami, FL 33199
markaf@fiu.edu

Abstract

I argue that automatic detection and understanding of *narratives* is critical to advancing our fundamental theories of, and ability to model, social media. Understanding the narratives in play has always been a key component to understanding human communication, but this has traditionally been a manual task. Now the huge and growing volume, velocity, and variety of social media content means that we must automate narrative understanding as much as possible if we are to have any hope of effectively tracking and modeling social media dynamics. Importantly, automated narrative understanding is useful not only for understanding the content of communications, but also for assessing the psychological impact of those communications, providing insight into the nature of human groups, and predicting future beliefs and behaviors. As a basis for next steps, I identify a critical common mistake that computationalists often make when approaching narrative, and point out two foundational fields where researchers should look for guidance.

Why Automatic Narrative Understanding?

There has been a great deal of mathematical and experimental work on modeling the forms and dynamics of the networks found in social media. While this work has been invaluable, it has mostly focused on network structure (dynamic or not), and has mostly left aside what *content* flows within those structures. I propose that effectively incorporating content into our models is the next big frontier in modeling social media.

Content in the media in general has become of much greater concern of late—and here I use *content* to denote anything that participants in a (social) media network might communicate within that network, including propaganda or mis-/dis-information. Content has come to the forefront because of social media’s driving a dramatic increase in the well-known “3 V’s”—volume, velocity, and variety—where a lot more people are saying a lot more different things a lot more often. This democratization and expansion of mass communication is in stark contrast with previous generations of media forms where the channels themselves contained natural chokepoints, with more responsibility and care being exercised, on average, with regard to content’s veracity

Copyright © 2022, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

or impact. Propaganda and mis-/dis-information were still of course disseminated through older forms of media, but the variety was less, and they were laid against a much smaller background of competing content. It was possible for human analysts and modelers to keep up via manual means.

Today, the huge variety of messages can be easily highlighted by considering only the vast resources that certain countries pour into generating malicious content. For example, the Skripal Poisoning of 2018 saw the Russian government generate several hundred different individual narratives—begetting tens of thousands of news articles and an orders of magnitude more social media activity—to flood the information space in an attempt to deflect blame (Ramsay and Robertshaw 2018). This was only one event fairly localized in time and space; multiply this by any number of events occurring everyday, and any number of well-resourced interested parties, all being continuously amplified and mutated through hundreds of millions of human users and bots. Add to that all the authentic discussion about all topics under the sun, and the problem of just understanding what is being said—not to mention incorporating it in models—becomes quite intimidating indeed. Thus, automated efforts to understand content, including its connotations and denotations, its relationships, and its provenance, are a necessary condition to advancing the science of modeling social media.

Why Automatic Narrative Understanding?

So far I have not said anything really all that new. So perhaps here is some novelty: when working toward modeling content in social media, we should focus explicitly on developing techniques for automatically understanding *narratives* present in the system.

What’s a narrative? Put simply, a narrative is a recounting of things that happen, but with a *point*. Meaning, a narrative is (1) a discourse that presents a sequence of events which are causally related, concern specific people, places, and times, and, crucially, (2) has higher order structure that conveys meaning beyond the commonsense or *prima facie* interpretation of the events in question. As I have discussed in other work (Finlayson and Corman 2013), if you have only (1) you have what one can call an *action discourse*. Only when you have (2) as well do you get a narrative, which gives the narrative its larger point, its “oomph,” if you will.

What kinds of higher order structures are possible? They are myriad, and can include things like “the moral of the story”, culturally relevant “tropes” or “plot patterns”, or “narrative arcs”. These structures imbue a vanilla sequence of events with a greater meaning that drives human behavior in important ways. As such, narratives are all over social media, in either full-fledged or fragmented form, and drive the dynamics of the system out of proportion to their number.

Uses of Narrative in Modeling Social Media

What would focusing on narrative (as opposed to other aspects of the content) get us in modeling social media? In other words, suppose we were able to detect and dissect narratives reliably in the social media space: what would this buy us? The obvious answer is it allows us to more reliably uncover what participants in social media ultimately *mean* when they communicate. Rather than focus on the surface words used, the raw emotion expressed, or the bare facts conveyed (all of which can now be automatically tracked, at least to some degree), concentrating on the narratives shows the forest for the trees: what is the real *point*? By developing an ability to track narratives, we will be able to find the real message rather than be distracted by the dressing.

Beyond this, uncovering the narratives at play and how social media participants are interacting with them will give us new insight into otherwise partially opaque social network structures, in particular, group memberships and affinities. What narratives one subscribes to is closely associated with one’s associated groups. Mapping narrative flows around the network will allow us to more deeply model and understand how political, social, religious, or cultural groupings (indeed, any grouping, even those of which we may not yet be aware) affect social media dynamics. Importantly, the associations revealed by narrative flows may be explicit, implicit, or somewhere in between, and this information is important aspect of predicting behavior, for example, assessing the risk of violent extremism.

Conversely, understanding narratives will allow us deeper insight into key players in the network: their goals, motivations, strategic thinking, and so forth. What is a participant in the network trying to achieve via their narratives? What kind of influence are they attempting to wield? Is it benign, hostile, neutral, something else? Furthermore, what do those participants think of *other* participants in the network, whom they are trying to influence? The narratives the influencer tells sheds light not only their own internal states but also on what they think of the internal states of the influenced.

Practical and Theoretical Underpinnings

There are many possible specific next steps for achieving automatic narrative understanding in the social media space, mostly found in the technical bowels of natural language processing, computer vision, multimodal processing, or artificial intelligence more generally. Rather than belabor these details, which change quickly in the short term, I would like to stake out two broad positions from which I believe any long-term approach to developing automatic narrative understanding must proceed.

First, future work on automatic narrative understanding in the social media space must move beyond current simplistic and inaccurate models of narrative. Right now, practitioners in the modeling space often compute sets of topics (such as those derived from topic modeling) or keywords, and call those “narratives”. Comparison of topics or keywords with the sketch of what a narrative given above shows the grave paucity of such a representation. Keywords and topics are easy to measure, but they are not narratives. Any attempt, then, to effectively automatically understand narrative must move beyond these simplistic models

Second, computational researchers must acknowledge that narrative is by no means *terra incognita*, and there are two key fields that lay out theoretical approaches which can and should be used as foundations of our automatic approaches to narrative understanding for social media. The first is *Communications Studies* (Littlejohn, Foss, and Oetzel 2021), which lays out general theories of human communication, including person-to-person communication, mass communication, cross-cultural communication, etc., but importantly acknowledges the use of narrative in all of these. Communication theorists have not been ignorant of social media’s transformative effect on modern communication, and the use of narrative therein, and their insights should be incorporated into our modeling approaches. The second key field is *Narratology* (Bal 2017) which concerns itself with study of narrative and narrative structures *per se*. Narratology is a subdivision of literary studies, but rather than focus on interpreting texts or assessing their aesthetics, narratologists seek to understand the narrative form, dissecting how narratives are constructed and how they achieve their effects. Familiarity with the basics of these two fields is key to driving forward research in automatic narrative understanding.

Acknowledgments

This work was partially supported by DARPA Awards HR001121C0186 under the INCAS program and D21AP10117 under the DARPA Young Faculty Award program. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of DARPA.

References

- Bal, M. 2017. *Narratology: Introduction to the Theory of Narrative, Fourth Edition*. Toronto, Canada: University of Toronto Press.
- Finlayson, M. A.; and Corman, S. R. 2013. The Military Interest in Narrative. *Sprache und Datenverarbeitung (International Journal for Language Data Processing)* 17(1–2): 173–191.
- Littlejohn, S. W.; Foss, K. A.; and Oetzel, J. G. 2021. *Theories of Human Communication, Twelfth Edition*. Long Grove, IL: Waveland Press.
- Ramsay, G.; and Robertshaw, S. 2018. Weaponising News: RT, Sputnik and Targeted Disinformation. Report, King’s College London, Centre for the Study of Media, Communication, and Power.