Agents for Social (Media) Change

Blue Sky Ideas Track

Sandip Sen
The University of Tulsa
Tulsa, Oklahoma, USA
sandip@utulsa.edu

Zenefa Rahaman
The University of Tulsa
Tulsa, Oklahoma, USA
zenefa-rahaman@utulsa.edu

Chad Crawford
The University of Tulsa
Tulsa, Oklahoma, USA
chad-crawford@utulsa.edu

Osman Yücel
The University of Tulsa
Tulsa, Oklahoma, USA
osman-yucel@utulsa.edu

ABSTRACT

We are addicted to the Internet and spend a significant portion of our waking hours engaged to that virtual world through the "window" of our electronic devices. A large majority of these interactions occur on online social media. From advertising campaigns to political debates and from trending news topics to communications from family and social circles, social media platforms and services have become invaluable and irreplaceable tools for most of us. Our beliefs and preferences are increasingly shaped and defined by what we see and experience on social media. With this increased reliance also comes the uneasy realization that information and knowledge of value to us is being drowned out in the deluge of forwarded messages and targeted communication from paid advertisers on various social media platforms. This paper seeks to highlight the research challenges underlying the potential for intelligent agents to help stem the tide, to help us deliberate, prioritize and process information of value to us and to our communities, as well as help us reach out, connect to, share and disseminate mutually interesting knowledge with other users. We posit an agent-based ecosystem, where both individual users and organizations see the value and creative possibilities of agent-based solutions to critical problems of connectivity, relevance, varying interest profiles, context, etc.

KEYWORDS

Social Media Analysis; Agent Surrogates; System Enhancers

ACM Reference Format:
Sandip Sen, Zenefa Rahaman, Chad Crawford, and Osman Yücel. 2018. Agents for Social (Media) Change. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018,* IFAAMAS, 5 pages.

1 INTRODUCTION

The idea of an *agent surrogate* encapsulating and representing the interests of a user in a complex information environment has always been an enticing prospect for multiagent systems researchers. Classically, a user's surrogate agent provides a personalized interface between the user and the environment, being knowledgeable both about the environment and of the user's interests and preferences, and enhances the user's participatory experience and operational capabilities in the environment. There are many important enabling, facilitating, and amplifying services such an agent can provide. Some of these, at an abstract level, include the following:

- keeping the user abreast of relevant information and events,
- collecting and fusing information from multiple sources to provide more comprehensive view of items of interest,
- helping identify and connecting to other users or entities that can add value or contribute to the user's activities and projects,
- support and guide the user in making decisions based on context, information sources, competing interests, etc.,
- communicating user decisions and disseminating user-generated information to relevant targets, adapted and presented in appropriate format to the norms and style appropriate for the recipient,
- shielding the user from targeted misinformation or security threats.

While various embodiments of these agent capabilities have been deployed in many domains, in this paper we reimagine the potential of the agent model in changing the asymmetric nature of information flow and user targeting currently pervasive in many social media platforms. The preponderance of social media platforms is hard to overstate, and the deep and personal attachment of most people with access to the internet and handheld smart devices is pervasive in our society. While social scientists are discovering the far-reaching effect such omnipresent and continual access to the virtual world has on our individual psychie and personal and social relationships, we are ill-equipped to comment on those issues. Rather, our goal here is to argue that agents research and technology can (i) help users be more informed consumers and producers of social media content, and (ii) contribute to developing more responsive and adaptive social media platforms that deliver content more efficiently and create more rewarding and meaningful connection between their users.

In this paper, we posit a number of functional applications for both personal agents, those who represent individual users, as well as institutional or system agents, those who work at the system level for social media platforms and services, In the process, we reimagine classical agent capabilities as applicable in social media scenarios. These *agents of change*, in how social media platforms process and deliver content and how users receive information from and upload information to such platforms, will support our citizens to separate the "wheat from the chaff" to become more

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden.

informed, support them in articulating and disseminating their viewpoints, protect them against security threats, support forming and maintaining rewarding virtual relationships and interactions, and help create a vibrant, responsive, and participatory information ecosystem for a more civil, measured, celebratory, well-knit and effective virtual social communities.

We recognize that some of the functionalities and applications mentioned below exist in some rudimentary or imperfect forms in current systems. However, we believe that significant improvements to the quality, consistency, robustness and effectiveness of such functionalities are necessary and we are well past the time when they became a necessity and not just a nice application feature. We argue that research and development in agent technology can provide the necessary performance guarantees on these fronts. In addition, we identify other functionalities that agent technology can engender, making the social media experience of the future a much more rewarding and empowering one.

We note that agent researchers are increasingly paying attention to some of the challenges we present. A recent IEEE Expert special issue showcases a handful of those efforts. Of note is the editorial piece [45], which presents a collection of research challenges [1, 11, 27] that do overlap with our proposal. The organization of this paper, separating agent versus system level agents and their collaboration, as well as a number of specific research challenges identified herein are distinct from current state of the art.

2 AGENTS OF CHANGE (IN SOCIAL MEDIA)

2.1 User surrogates

We identify the scope of agents for assisting or representing users to better receive or present information on social media platforms.

2.1.1 Research Assistants for Social Media. Remarkably, people increasingly look towards their social media feeds for news rather than relying on more traditional news outlets. While this trend have its downside, with preponderance of false information as well as slanted and heightened rhetoric, it does have the potential of providing real-time information dissemination from local and distributed sources.

On a related note, social media empowers people to more readily share their viewpoints and positions on any topic of interest to them. Other users also find information or entertainment value, and certainly relevance, in the postings of fellow posters. We posit that together with search engines for the web, there exists a latent demand for "research assistants for social media" which will allow users to access organized, cross-referenced, and easily explorable presentation of social media content related to any issue of interest to a user. For example, if a user searches for a hot-button political issue, such as tax or health care reform, or social issues such as drug overdose epidemic, is it possible to collect, organize and present multiple viewpoints and perspectives as well as their relationships as expressed by different types of users or groups?

Research assistant agents serve the critical role of collating, aggregating, and organizing information to make it consumable for its target user. Some aspects of the research assistant role could be handled by a system agent (Section 2.2.1). System agents generally may have unrestricted access to any information stored on that system. Conversely, personal agents only have access to their target user's information and whatever information the target user is allowed to access. While the system agent is often constrained to information available on that single platform, a personal agent can be aware of social media activities for its target user across multiple platforms. Given the more comprehensive access to how and when a user searches for, stores, shares and uses information, the personal agent has a more complete view of the user's preferences and hence provide targeted personalization that is infeasible for site-specific, system level agents. We discuss below some of the task challenges to be solved for deploying successful personal research assistants.

Ranking refers to the task of selecting social media posts that are relevant and of interest to the user. Personal agents have a distinct advantage over system-level agents when ranking items. For example, a user who prefers to read news articles rather than watch videos might not explicitly express the preference on every social media outlet, or may be new to a certain platform. However, a personal agent may not have the computational power or resources to produce recommendations like a system agent. While personal agents have wider access to a target user's social media activity, they do not have easy access to the activity of other users which could enable the use of collaborative mechanisms.

Challenge: *Personal research agents need to coordinate with system agents to produce effective rankings while protecting user's privacy.*

Aggregation refers to collecting a multitude of social media posts and summarizing them for their target user. There has been research on effectively summarizing social media posts as a general tool for readers [15]. A generic summarization, however is not informative to every user, as people are often interested in different aspects of a story. Summarization needs to be applied often to even dissimilar social media posts on related topics, users are better informed when exposed to the multiple perspectives that individuals have on an issue, as well as arguments for each side. Argumentation mining is a developing field that seeks to produce mechanisms that identify structured arguments within a conversational corpus. However, informal and illogical arguments on social media are also important to mine and understand. Additionally, such techniques need to be largely unsupervised as new topics are continually developing. Some of the correlated issues can be grouped into multiple hierarchies, e.g., sub-issues of liberal vs conservative debate.

Challenge: *Develop personalized information summarization schemes to organize diverse, often conflicting, viewpoints on related issues.*

To understand a user, we can learn how they communicate. Current text analysis approaches use statistical techniques like mixed membership topic modeling approaches, e.g., Latent Dirichlet Allocation [8], Probabilistic Latent Semantic Analysis [22], to cluster text. The primary limitation of these is that the small number of words in social media posts leads to sparse word co-occurrences. Several works address this issue by aggregating posts by author [25, 29, 38]. Nonetheless, there have been little to no work on statistical models for conversational text, or *topic flow* in conversations.

Challenge: *Develop topic models for social media conversations.*

2.1.2 Guarding against misinformation. People or organizations can run misinformation campaigns on social media platforms to manipulate users. Such campaigns include false news to direct people's opinions, misleading advertisements about products, etc. Though reputation and trust mechanisms [30, 46] can be used to detect malevolent sources, they typically assume static user preference.

Challenge: *Develop trust and reputation models that track user's changing belief landscape to adapt social media content ratings.*

2.1.3 Social acquaintance finder. Individual's opinions are shaped based on their social media circle and exposure, the people they follow and the people who follow them. Personalized agents can help a user build a social network of acquaintances they enjoy communicating with. Social acquaintance finder is an application of Recommendation systems [9, 10, 12, 17] which rely on gathering, extracting, and inferring information about users and recommend them diverse types of information on social media: From friends to food, from movies to books.

Users write reviews, share images and videos in social media, which provides information about both users and items. It is impossible for the target users to consume all the information available on social media. Hence, it will be useful to create agents which analyzes information on candidate items on behalf of the target user, to learn about candidate items and which sources to trust.

Sun et al. [39] created a recommendation system where the conversations between users on social media are analyzed for increased accuracy. An agent based tagging recommender is presented by An et al. [3]. These agents, however, do not fully leverage the duality of social network topology and content-based similarity between non-neighbor users and fail to find other acquaintances that help broaden the mental and knowledge landscape of the user.

Challenge: *Develop social media focused recommender systems that can connect users to other that are not just similar but provide rewarding mutually complementary experiences.*

2.1.4 Protection from trolls. Though several computational studies have developed automated mechanisms for detection of unwarranted victimization and harassing attacks on social network and microblogging platforms [14, 16, 18, 19, 26, 36, 43, 44], none of them address the issue from the user's perspective. User perception of cyber harassment and tolerance level for cyber harassment can vary widely based on personality traits. For example, a recreational social media user [32] may be more prone to be affected by cyber bullying compared to a professional social media user such as journalist [28]. However, the latter is more likely to be the target of concerted attacks from groups championing certain causes about which a journalist may write a critical commentary. Personal guardian gents can be developed to filter communication based on the target user's threat perception and can also provide morale boosting supportive communication when the user is a victim of aggression. Other agents can also help users from inadvertently posting offensive communication by flagging communication that can be viewed to be insensitive or offensive by other users.

Challenge: *Develop a personalized guardian agent that can learn user's threat perception and tolerance and filter/organize abusive communications directed towards the user and help counter cyber harassment by providing effective support to targeted users.*

2.1.5 Dissemination of user content. Users are interested in sharing personal information and expressing themselves on social media. Beyond the act of posting itself, the user is particularly interested in the feedback (preferably positive) received from others about the post. Agents can learn the characteristics of successful (receiving many positive responses) posts and suggest to the user appropriate enhancements and expressions to elicit positive feedback. Other functionalities of such agents include adapting post content to conform to site-specific requirements, such as word limits on Twitter, suggesting posts be forwarded to groups or users who will value the post, and even suggesting other social media platforms the user might want to join and post the said content.

Challenge: *Develop agents that can transform user content to multiple forms, adapted stylistically and content-wise to different targeted social media platforms to have maximum visibility and impact.*

2.2 System enhancers

In this section, we identify the scope of agents for adding functionality and improving performance of social media platforms at the system, rather than the individual user level.

2.2.1 Information organizers. While part of the appeal of social media is free flow of information between users, collecting related information in topic-based repositories, e.g., collection of videos of news events of interest to the user base, can be of value to users. User modeling, pattern recognition and text/audio/video/image processing techniques can be leveraged to develop agents that collate, organize and provide structured access to diverse information content on popular and/or trending issues.

2.2.2 Solving the "connection problem": Social media relationships often mirror and reaffirm real-life relationships. Agent applications that help people forge new and beneficial relationships over social media can have measurable social benefits, from protecting children from cyberbullying [23] to mitigating depression [33].

One of the challenges of connecting such people is reliably predicting when users will have positive interactions. There are significant challenges in addressing this issue: (1) Lack of labeled corpora of conversations deemed as "friendship building" or "hostile." (2) A conversational predictive model of friendship development is needed to learn personality and conversational traits for each user and quickly find predicted strong friendship links between users.

2.2.3 Guarding against "information attacks". With increasing reliance of users on social media as news source, we are witnessing the rapid proliferation of fake news and online hoaxes [6, 13, 41, 42]. One particularly significant current case is how fake news [5, 24] has effected the results of the 2016 presidential election in United State [7]. A system level agent can employ probabilistic reasoning and text analysis mechanisms and correlate information from verifiable and trustworthy sources to estimate the likeihood of a news article being fake. Existing tools for microblogging sites such as Hoaxy, Snopes.com, BotOrNot and Sebenarnya.my are built to eradicate the fake news problem. While these tools deal with current embodiments of online fake news, several new techniques are emerging for creating online fake news and spreading those among social media users to shape their opinions. An agent can continue to learn different patterns of emergence of fake news, and can also detect and isolate potential sources of fake news.

2.2.4 Guarding against personal attacks. We argue that social media platforms must protect individual users against personal attacks, such as harassment and abuse, as well as against security threats to steal and misuse sensitive personal information. Manual culling used by social media platforms to eliminate online abuse has proved inadequate. Use of automated agent based systems to curb online harassment is a potential game changer. Despite various computational studies on harassment detection [31, 34–37], research on automated intervention is sparse. Some recent work [20] uses

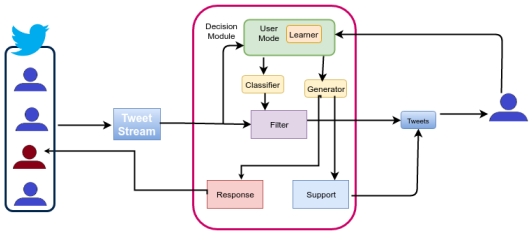

Figure 1: Agent intervention against online harassment.

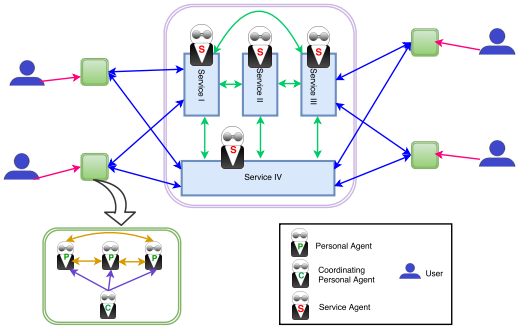

Figure 2: Cooperating/collaborating social media agents.

bots to help maintain block-lists and TrollHunter [2] uses an agent, imitating a person with similar demographic characteristics to the troll, to provide feedback to dissuade hateful comments. Both approaches, however, require manual intervention. A system agent, with access to system-wide information, can mitigate this problem by detecting existing and newly emerging abuse patterns and alerting authorities or users about abusive communication and sources. These agents can also detect changes in intensity, type, and patterns of online aggression by individuals over time as well as analyze how groups of aggressive users mutually reinforce each other's messages, leading to escalation of aggression and conflict. The system level agent can work to protect individual users against personal attacks, such as harassment and abuse, as well as against security threats to steal and misuse sensitive personal information.

System level agent interventions can include the following:

Filtering Offensive Communication: Agents can employ machine learning approaches [4, 21, 40] for classifying communications by harassment types and intensity level. The agent can then personalize the delivery of communication to the user by selectively filtering or tagging offensive Communications.

Supportive Communication to Target Users: Targets of aggressive communication can be heartened by receiving supportive communication, which is one of the purported goals of the website TrollBuster (http://www.troll-busters.com/). An agent can send supportive messages to dilute the chain of aggressive communication, divert attention from these negative inputs, and bolster the confidence and self-esteem of the target of aggression.

Response to Online Aggressors: System agents can identify and classify sources of aggressive communications based on the frequency, intensity, and number of users attacked by such a source. The agent can take various counter-measures such as shaming, aggressive counter-attacks, threat of legal recourse, etc. as well as evaluate the relative efficacy of these different intervention. Figure 1 shows system level agent interventions on a social media platform.

Developing an agent based system that protects users from personal attacks on social media gives raise to following research challenges: (i) What monitoring and filtering mechanisms can be employed to detect harassment? (ii) How to use the personal information of users to build personalized filtering? (iii) What intervention mechanisms to apply to reduce the effect of aggressive behavior on target? (iv) What intervention strategies will help stop online harasser from sending abusive communications?

2.3 Case for agent cooperation/collaboration

We assume that each social media platform has a system level agent. Users can also have personal agents for each of the services the user engages with. We argue that these agents at the user and the system level should collaborate (see Figure 2).

Cross-platform agents: Agents can provide seamless integration and presentation of information received from multiple platforms. For example, a core user agent can correlate information available on multiple social media platform to piece together more coherent and comprehensive views of the distributed presence of individual entities, such as friends or organizations, across multiple platforms. Similar information integration and organization will also be useful for topics or issues of importance to the target user.

Helping disseminate user content across platforms: Cross-platform agents can also take more elaborate input from users on certain topics and issues and help reformulate them for various social media platforms, adapting the content to meet the requirements and constraints of individual platforms, such as word and size limits, as well as accepted norms and prevalent posting practices of users on the plaform, e.g., the use of terminology, emoticons, etc.

Collaboration between user and system level agents: We argue that given a service, its user agents should coordinate with its system level agent for that service to provide its users a more comprehensive as well as personally adapted interaction experience with the social media platform. Such collaboration between personal and system level agents have been alluded to in previous sections and is crucial to both improving user satisfaction as well as providing for better utilization of system level resources.

3 CONCLUSIONS

We presented a gamut of possibilities for agents to both inform and empower individual users, by providing a more comprehensive view of issues being discussed across social media platforms and connecting like-minded users, and for developing more agile and responsive platforms that can better organize information, connect its user base and mitigate attacks by leveraging available resources. While the scope and diversity of agent applications for social media presented herein are neither exhaustive nor detailed in specification, we have identified a broad set of key challenges that define fruitful research avenues and will lead to transformative social media applications of agent technology of significant value to citizens. We believe such applications can both drive novel agent research agendas and can showcase agent technology to the populace.

REFERENCES

[1] Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. 2015. Privacy and human behavior in the age of information. *Science* 347, 6221 (2015), 509–514.

[2] Sally Adee. 2016. Troll Hunters: The Twitterbots That Fight Against Online Abuse. *New Scientist* (2016).

[3] Shihu An, Zhikun Zhao, and Hong Zhou. 2017. Research on an Agent-Based Intelligent Social Tagging Recommendation System. In *Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2017 9th International Conference on*, Vol. 1. IEEE, 43–46.

[4] Horace B Barlow. 1989. Unsupervised learning. *Neural computation* 1, 3 (1989), 295–311.

[5] Hal Berghel. 2017. Alt-News and Post-Truths in the" Fake News" Era. *Computer* 50, 4 (2017), 110–114.

[6] Hal Berghel. 2017. Lies, Damn Lies, and Fake News. *Computer* 50, 2 (2017), 80–85.

[7] Hal Berghel. 2017. Oh, What a Tangled Web: Russian Hacking, Fake News, and the 2016 US Presidential Election. *Computer* 50, 9 (2017), 87–91.

[8] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3, Jan (2003), 993–1022.

[9] Philip Bonhard, Clare Harries, John McCarthy, and M Angela Sasse. 2006. Accounting for taste: using profile similarity to improve recommender systems. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 1057–1066.

[10] Philip Bonhard and Martina Angela Sasse. 2006. Knowing me, knowing you-Using profiles and social networking to improve recommender systems. *BT Technology Journal* 24, 3 (2006), 84–98.

[11] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2011. Information credibility on twitter. In *Proceedings of the 20th international conference on World wide web*. ACM, 675–684.

[12] Li Chen and Feng Wang. 2013. Preference-based clustering reviews for augmenting e-commerce recommendation. *Knowledge-Based Systems* 50 (2013), 44–59.

[13] Yimin Chen, Niall J Conroy, and Victoria L Rubin. 2015. Misleading online content: Recognizing clickbait as false news. In *Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection*. ACM, 15–19.

[14] Ying Chen, Yilu Zhou, Sencun Zhu, and Heng Xu. 2012. Detecting offensive language in social media to protect adolescent online safety. In *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*. IEEE, 71–80.

[15] Freddy Chong Tat Chua and Sitaram Asur. 2013. Automatic Summarization of Events From Social Media. (2013), 81–90.

[16] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting Depression via Social Media.. In *ICWSM*. 2.

[17] Souvik Debnath, Niloy Ganguly, and Pabitra Mitra. 2008. Feature weighting in content based recommendation system using social network analysis. In *Proceedings of the 17th international conference on World Wide Web*. ACM, 1041–1042.

[18] Nemanja Djuric, Jing Zhou, Robin Morris, Mihajlo Grbovic, Vladan Radosavljevic, and Narayan Bhamidipati. 2015. Hate speech detection with comment embeddings. In *Proceedings of the 24th International Conference on World Wide Web*. ACM, 29–30.

[19] Patxi Galán-GarcÍa, José Gaviria De La Puerta, Carlos Laorden Gómez, Igor Santos, and Pablo García Bringas. 2015. Supervised machine learning for the detection of troll profiles in twitter social network: Application to a real case of cyberbullying. *Logic Journal of IGPL* (2015), jzv048.

[20] R Stuart Geiger. 2016. Bot-based collective blocklists in Twitter: the counterpublic moderation of harassment in a networked public space. *Information, Communication & Society* 19, 6 (2016), 787–803.

[21] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. 2009. Overview of supervised learning. In *The Elements of Statistical Learning*. Springer, 9–41.

[22] Thomas Hofmann. 1999. Probabilistic Latent Semantic Indexing. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 50–57.

[23] Kristin Kendrick, Göran Jutengren, and Håkan Stattin. 2012. The protective role of supportive friends against bullying perpetration and victimization. *Journal of adolescence* 35, 4 (2012), 1069–1080.

[24] Geraldine Lee. 2017. The importance of facts in this 'fake news' era. *International Emergency Nursing* 31 (2017), 1.

[25] Yan Liu, Alexandru Niculescu-Mizil, and Wojciech Gryc. 2009. Topic-link LDA: joint models of topic and author community. In *Proceedings of the 26th Annual International Conference on Machine Learning*. ACM, 665–672.

[26] Chikashi Nobata, Joel Tetreault, Achint Thomas, Yashar Mehdad, and Yi Chang. 2016. Abusive Language Detection in Online User Content. In *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 145–153.

[27] Bo Pang, Lillian Lee, et al. 2008. Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval* 2, 1–2 (2008), 1–135.

[28] Sherry Ricchiardi. 2017. Female journalists fight online harassment. *Global Investigative Journalism Network* (2017).

[29] Michal Rosen-Zvi, Chaitanya Chemudugunta, Thomas Griffiths, Padhraic Smyth, and Mark Steyvers. 2010. Learning author-topic models from text corpora. *ACM Transactions on Information Systems* 28, 1 (2010), 1–38. https://doi.org/10.1145/1658377.1658381

[30] Fernando P Santos. 2017. Social Norms of Cooperation in Multiagent Systems. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 1859–1860.

[31] Anna Schmidt and Michael Wiegand. 2017. A Survey on Hate Speech Detection using Natural Language Processing. *SocialNLP 2017* (2017), 1.

[32] H Andrew Schwartz, Johannes Eichstaedt, Margaret L Kern, Gregory Park, Maarten Sap, David Stillwell, Michal Kosinski, and Lyle Ungar. 2014. Towards assessing changes in degree of depression through facebook. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*. Citeseer, 118–125.

[33] Maarten HW Selfhout, Susan JT Branje, M Delsing, Tom FM ter Bogt, and Wim HJ Meeus. 2009. Different types of Internet use, depression, and social anxiety: The role of perceived friendship quality. *Journal of adolescence* 32, 4 (2009), 819–833.

[34] Leandro Silva, Mainack Mondal, Denzil Correa, Fabrício Benevenuto, and Ingmar Weber. 2016. Analyzing the targets of hate in online social media. *arXiv preprint arXiv:1603.07709* (2016).

[35] Sara Sood, Judd Antin, and Elizabeth Churchill. 2012. Profanity use in online communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1481–1490.

[36] Sara Owsley Sood, Judd Antin, and Elizabeth F Churchill. 2012. Using Crowdsourcing to Improve Profanity Detection.. In *AAAI Spring Symposium: Wisdom of the Crowd*.

[37] Sara Owsley Sood, Elizabeth F Churchill, and Judd Antin. 2012. Automatic identification of personal insults on social news sites. *Journal of the American Society for Information Science and Technology* 63, 2 (2012), 270–285.

[38] Mark Steyvers, Padhraic Smyth, Michal Rosen-Zvi, and Thomas Griffiths. 2004. Probabilistic author-topic models for information discovery. *Proceedings of the 2004 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1990 (2004), 306. https://doi.org/10.1145/1014052.1014087

[39] Yueming Sun, Yi Zhang, Yunfei Chen, and Roger Jin. 2016. Conversational Recommendation System with Unsupervised Learning. In *Proceedings of the 10th ACM Conference on Recommender Systems*. ACM, 397–398.

[40] Richard S Sutton and Andrew G Barto. 1998. *Reinforcement learning: An introduction*. Vol. 1. MIT Press Cambridge.

[41] Glenn S Tillotson. 2017. Keeping the faith–reporting on antimicrobial resistance in an era of fake news. *The Lancet Infectious Diseases* 17, 5 (2017), 473–474.

[42] Nitin Verma, Kenneth R Fleischmann, and Kolina S Koltai. 2017. Human values and trust in scientific journals, the mainstream media and fake news. *Proceedings of the Association for Information Science and Technology* 54, 1 (2017), 426–435.

[43] William Warner and Julia Hirschberg. 2012. Detecting hate speech on the world wide web. In *Proceedings of the Second Workshop on Language in Social Media*. Association for Computational Linguistics, 19–26.

[44] Dawei Yin, Zhenzhen Xue, Liangjie Hong, Brian D Davison, April Kontostathis, and Lynne Edwards. 2009. Detection of harassment on web 2.0. *Proceedings of the Content Analysis in the WEB* 2 (2009), 1–7.

[45] Pınar Yolum and Michael N Huhns. 2017. Agents for Social Media. *IEEE Internet Computing* 21, 6 (2017), 5–7.

[46] Yanjun Zuo and Jigang Liu. 2017. A reputation-based model for mobile agent migration for information search and retrieval. *International Journal of Information Management* 37, 5 (2017), 357–366.