

Statement of Research Interests

Artificial Intelligence (AI) and technology have become increasingly embedded in our daily lives. Every day, we read the news, watch videos and our favorite series, search for nearby restaurants, order food, buy things, and chat with our friends. Online platforms, together with AI and technology, serve all of these actions, thus becoming ‘essentials’ of modern society. Billions of people use them and spend a significant amount of time there. On average, an individual is online for 6 hours and 42 minutes every day [1] and spends 143 minutes on social media [2]. These tremendous interactions on online platforms enable us to understand individual or collective human behavior: what people think, what people care about, how people feel, where people go, what people buy, what people do, and who people associate with. Tools that can extract hidden patterns and insights from large-scale data are essential to address complex social challenges.

The recent advancement of AI has empowered researchers and practitioners to investigate such massive data in an innovative way. For example, Natural Language Processing (NLP) has enabled us to examine the semantic differences of the same words across communities and the nuanced framing of news media, which has been considered challenging and heavily relied on manual efforts. Moreover, deep neural network architecture effectively incorporates various data modalities, such as text and image, or multiple data dimensions, such as network and time-series data, handling the high complexities of data in the wild. For instance, information spreading may be better studied through individual modeling and underlying networks, and urban challenges may be better addressed through environmental conditions and social context.

My main research theme is developing AI-based methods and tools to 1) understand, predict, and nudge online human behavior and 2) tackle a wide range of social problems. I have explored various types of large-scale data and investigated and compared existing AI methods to overcome their limitations. Also, I have been developing new measurements, machine learning models, and NLP methods to understand human behaviors online and solve social problems from media bias and framing, polarization and diversity, online hate, misinformation to user engagement, healthy lifestyle, and urban changes. In this context, I summarize my past research and plans in the following sections.

Novel AI and NLP methods for understanding and predicting online human behavior

Among various formats of online data, my research focuses on the textual content to facilitate the nuanced analyses of human behavior on online platforms. Users’ online posts and conversations can provide insights into emotions, opinions, and actions of individuals, groups, or the public in the past, current, and future. My recent research addresses a range of problems relating to the analysis of structured and unstructured textual information used in, or arising from, content analysis for understanding and predicting online human behavior. A selection of my research in this direction is described below.

Characterizing human behaviors from texts. In online human behavior research, one of the vital tasks is a ‘comparative text analysis’—how two (or more) groups differ in terms of their language use that can be used in many applications from politics to business. Studies considering comparative text analysis have primarily focused on issue salience using topic analysis or tone differences via sentiment analysis. However, language is context-dependent—the same word can mean differently in different communities. For example, the word ‘soft’ can be positively associated with a community of animal lovers while it can have a negative connotation in a sport community [3]. Such subtlety would make crucial differences when studying sensitive issues such as politics or health.

The recent advancement of NLP, particularly the idea of using neural networks to learn ‘good’ representations of words and sentences (i.e., word embeddings), has enabled us to do nuanced text analysis with large corpora.

In a paper published at ACL 2018 (a top-tier NLP conference), we proposed a novel NLP method, ‘SEMAxis,’ that measures semantic changes in words across communities [4]. Using transfer learning of word embeddings, SEMAxis offers a framework to examine and interpret words on diverse semantic axes (732 systematically created semantic axes that capture common antonyms, such as respectful vs. disrespectful). For example, how would Trump supporters perceive ‘gun’ compared to Sanders supporters? One may expect that ‘gun’ would be more positively perceived for Trump supporters than Sanders supporters, but can we go beyond the simple sentiment and characterize the semantic differences more nuancedly? Our approach can examine how ‘gun’ is perceived in two communities in relationship to ‘Arousal’ and ‘Safety’ axes. Using posts and comments on the two subreddits, /r/The_Donald and /r/SandersForPresident, we found that Sanders supporters associate ‘gun’ more with arousal and danger than Trump supporters. Our study can facilitate further investigations on context-dependent text analysis techniques and applications.

More recently, we applied the idea of SEMAxis to detect framing in the text. We proposed FRAMEAxis, an unsupervised method for characterizing documents by identifying the most biased and prominent semantic axes that are overrepresented in the text using word embedding [5]. Our method may accelerate scalable and sophisticated computational analyses of framing across disciplines.

Mining opinion in social media. Research on online human behaviors is often challenged by the difficult measurement tasks of aggregating large but noisy and non-representative data. Since online users do not tend to be a representative population sample, using raw online data may incur sampling bias in the analyses. In my research program, I have been tackling measurement problems and seeking new methods to resolve them.

With a growing interest in using online data to study and quantify phenomena in the offline real world, much effort has been made to use social media to measure public opinion or for now-casting indices such as levels of flu activity or unemployment. In this context, social media users, who publicly report their health status or job losses, may act as ‘social sensors.’ For example, in our study [6], which was published at *Scientific Report*, we showed high correlations between Google and Twitter search volumes linked to Middle East Respiratory Syndrome (MERS)-related keywords and the number of confirmed MERS cases with a four-day delay. The result provides evidence that digital surveillance using a search engine and Twitter data is helpful in monitoring an emerging infectious disease outbreak.

Yet, to create reliable and auditable social sensing systems, it is critical to understand those people who talk online. Are there particular demographic groups that should be given more or less attention? To this end, we analyzed a 10% sample of all public tweets in 2014, which were in their billions. We investigated how different sampling strategies (first-person [real user, not agencies], gender, age, geography, number of followers, etc.) affect the performance of the now-casting of two common offline indices: flu activity and unemployment [7]. We found that now-casting performs best when the fraction of ‘normal’ people, who have a complete profile, but with not too many followers, and who generally babble about their personal lives, is at its highest. This work provides not only methodological contributions for user attribute inference (e.g., first-person classifier for Twitter users) but also valuable guidelines for designing a social sensing system.

Predicting user engagement. User engagement on social media has been topics of perpetual interest. In the era of attention scarcity, but ironically under information overload, brands, businesses, and organizations strive to make users engage with them because it holds the key to their success. I have focused on the news industry and attempted to predict user engagement levels before publishing news posts on social media. We considered how stylistic features [8], topics [9], and textual features [10] of news content are associated with popularity across multiple platforms using 3.1M news posts of 53 news organizations on Facebook, Twitter, Instagram, and YouTube. We then provided prediction models to determine whether news content becomes popular or not on each platform. As one would expect, our results showed that factors associated with user engagement vary across platforms and news organizations. We have provided a framework to examine news posts’ semantic and syntactic features in users’ engagement through a series of work using cross-platform datasets.

Tackling social problems with AI methods

While it is tempting to think of online platforms as safe and diverse places with lots of ‘good’ opportunities, unfortunately, this may not always be true. Online spaces can be full of hate, bias and prejudice, stereotypes, and fakery; thus, it is crucial to understand the risks and limitations of online platforms and their negative influences on individuals and society. I have been tackling various social problems by leveraging state-of-the-art AI and NLP methods.

Anti-Asian Hate during COVID-19. With the rise of xenophobia and polarization that has accompanied widespread social media usage in many nations, online hate has become a major social issue, attracting many researchers. The novel coronavirus (COVID-19) pandemic provides a unique opportunity to study the development of targeted racial animus at an unprecedented scale. In a recent paper published at EMNLP 2022 Findings (a top-tier NLP conference) [11], we investigated predictors of anti-Asian hate among Twitter users throughout COVID-19. In particular, we applied NLP techniques to characterize social media users who began to post anti-Asian hate messages during COVID-19. We compared two user groups—those who posted anti-Asian slurs and those who did not—with respect to a rich set of features measured with data prior to COVID-19 and showed that it is possible to predict who later publicly posted anti-Asian slurs. Our analysis of predictive features underlines the potential impact of news media and information sources that report on online hate and calls for further investigation into the role of polarized communication networks and news media.

Media Bias and Framing. Cultivation theory holds that long-term exposure to media shapes how the consumers of media perceive the world and conduct themselves [12]. Thus, it is vital to monitor potential biases of media on what they report (‘bias’) and how they report (‘framing’). Given the increasing number of new media and a large amount of news content published every day, automatic ways of detecting media bias and framing would be helpful to understand better the effect of media on our society. I have been exploiting various social media features to detect political bias of news media, such as co-following relationships [13, 14] or published content [15]. Furthermore, we empirically showed that not all news articles follow their source media’s political bias in a recent study published at ICWSM 2020 [16], calling attention to the importance of article-level political bias detection.

Framing is known to be challenging to operationalize and detect. Recently, we have proposed ways to detect frames automatically: 1) we developed a general media frame classifier that achieves state-of-the-art performance [17] and 2) we proposed new NLP methods to identify frames in a large text corpus [5]. As traditional media framing studies relied on manual annotation, it has been challenging to scale up to massive news datasets. We believe that our scalable, computational approaches open up new pathways for systematic media framing studies.

Misinformation. The authenticity of information has become a significant issue affecting businesses and society. To tackle this challenge, we have built *Tanbih*. *Tanbih*, an Arabic term meaning ‘alert,’ is a news aggregator aimed at combating misinformation and promoting news literacy by presenting various meta-information regarding news articles such as whether it is propagandistic or not or a frame expressed as well as media profiles including factuality and political bias [18]. We recently presented our demo at EMNLP 2019 [19]. *Tanbih* has been featured in reports by more than 50 public and university media outlets, including Wired, MIT News, Fast Company, and Forbes.

Furthermore, in a recent paper published at ACL 2020 [15], we proposed a machine learning model that detects source-level factuality. We studied the impact of three dimensions, namely 1) what was written, 2) who reads it, and 3) what was written about the target medium (in Wikipedia). The evaluation results show that what was written matters most, and we further show that putting all information sources together yields huge improvements over the current state-of-the-art. As the proliferation of fake news on social media becomes a primary societal concern, our study provides a practical solution to reduce the time for a fact-check.

Polarization. Polarization has been a threat to our society by damaging all institutions essential to democracy. For any nation formed by multi-racial, multi-cultural, multi-religious heritage, polarization is an inevitable social issue that has to be resolved. Unfortunately, digital and mobile technologies and social media have inflamed those divisions by enabling echo chambers and filter bubbles. In this context, I have empirically revealed the polarized behavior on online platforms. For example, we observed ‘partisan sharing’ on both Facebook and Twitter; people selectively share political news based on their political leanings [20], contributing to the creation of a social media echo chamber. More recently, by comparing the interaction and linguistic patterns of Trump and Clinton supporters using 2.5M posts and 39.8M comments on Reddit [21], we found that Trump supporters are far less likely to engage with Clinton supporters. Moreover, Trump and Clinton’s supporters change the meaning of words significantly when talking to ingroup vs. outgroup persons. Mitigating polarization is known to be extremely hard, and thus, early detection of polarization would be beneficial to society. Our findings suggest that not only social connections but also linguistic behaviors can be indications of polarization.

Underrepresentation of Minorities. Minority groups are continuously under- or misrepresented in both traditional and new media. Since media have the power to influence people, the portrayal of different demographic groups in their content can play a crucial role in shaping the formation of the identities of those groups. The visual analyses have been required manual efforts. However, the advancement of deep learning has have reverberated computer vision, enabling us to conduct large-scale image analysis to investigate the representation of various demographic groups on digital content. As one example, we looked into news media. By analyzing more than two million news photos published in January 2016, we observed a significant gender gap in news photos: 1) there are more women than men in living and health sections; and 2) women smiles more and look younger than men in the news photos [22]. Another area we find essential to monitor is online advertising, as big brands, such as Starbucks or Jeep, have millions of followers on social media. By analyzing 350K online advertising images relating to top US brands on Instagram and Facebook, we characterized the gender and racial diversity among ads [23]. Our study revealed that racial diversity differs across brands, and inter-racial couples are generally underrepresented in online ads. This work provided a framework for building a real-time system that monitors diversity and stereotyping in advertisements at a low cost.

Not only the media but online platforms also need to consider fair representation when designing features. For example, ‘trending topics’ on social media influence our perceptions of globally significant issues. But, are those topics genuinely trend for all? By analyzing tweets and profile images of 350K Twitter users from New York, we showed that the trending topics are driven mainly by a dominant, majority demographic group as the influence of the data for minorities is too small to register [24]. Another example of such features is ‘most read news,’ typical in news portals or online news sites. These lists often give perceptions on which news topics are salient. However, who are these topics important to? Using the daily top-30 news items for each gender and age group in Daum News, the second most popular news portal in South Korea, for the whole of 2015, we found that there is a striking difference between the popular news items across groups [25]. Our work highlights that these platforms need to go beyond analyzing population-level behavior, as majority groups will always dominate this.

Future directions

My research aims to understand, predict, and nudge online human behavior and tackle social problems by developing theoretical models and AI-based methods and tools using large-scale real-world data. My interests span a broad area of computational social science, including communication and journalism, politics, and health; I believe that the rich interdisciplinary environment of the school will enable me to pursue my research objectives successfully.

Computational framework to understand changes of human opinion and behavior. My primary plan is to 1) develop a new holistic representation learning framework to computationally model individuals by combining Beliefs, Emotions, Attitudes, and their Relationships (BEAR) and 2) apply the framework to uncover fundamental mechanisms of opinion and behavioral changes, particularly regarding polarization.

I believe that individuals' BEAR determine their opinions and behaviors. For example, those who believe that scientists are lying (**beliefs**) and those who hate the government (**emotion**) may possess negative attitudes toward the government's science-related activities (**attitude**). They may be susceptible to believing false information, such as "the COVID-19 vaccine changes humans' DNA," causing an aversion to vaccination (**behavior**). Therefore, understanding BEAR may reveal the cognitive mechanisms underlying changes in opinions and behaviors.

Yet, one's intrinsic characteristics may not fully explain such changes as social relationships and interactions also influence us. Together with social context, the BEAR framework can be extended to tackle various problems, including misinformation spreading, development of racial animus, social polarization, algorithmic and platform bias, and their impact. For this goal, I am collaborating with the psychologist at A*STAR (Agency for Science, Technology and Research) to develop methods to infer the attitude strength from linguistic features. The school and I will benefit from each other because many faculty members have expertise in social media, information dissemination, and online trust (e.g., Professors Menczer, Flammini, Bollen, and Ahn), and my experience in developing and applying AI and NLP methods can help to extract meaningful signals from the text data for research in those area.

Developing computational methods for computational social science. As textual data is a primary source of online data, I apply and develop NLP methods in my work. I will further leverage and pursue rich linguistic analysis methods to answer questions in both academia and industry. My short-term goal includes developing context-dependent text analysis techniques for non-English languages, particularly resource-scarce languages, by extending SemAxis [4] and FrameAxis [5] to foster research in other non-English cultures.

Images and videos are presently abundant on social media. These visual data can be crucial cues to understanding the psycho-sociological status of individuals. However, due to the difficulty in processing large-scale visual data, the relationship between visual data and human behavior on social media is understudied. Based on my previous experience with visual analysis [22, 23], I aim to use visual data to understand individuals' interests and behaviors and develop tools that make it more comfortable to analyze visual data for other research communities.

Media analysis and its impact on society. Our ongoing project, 'Discursive power of contemporary media system,' funded by the Volkswagen Foundation, aims to analyze interdependencies and power relationships among media, groups, and individuals contributing information to national and international news flows. Under the current media system, anyone can contribute to communication spaces to introduce, amplify, and maintain topics, frames, and speakers, thus shaping public discourses and controversies; this concept is called discursive power [26]. Working with a political scientist and an informatics scholar at the University of Bamberg, I will reveal the power relationships in contemporary media systems and provide theoretical, methodological, and practical implications. We are currently building software that collates news articles, political talk, and social media interactions in real-time, which other scholars will be able to access in the future.

Nudging human behavior. Large-scale observational online data and the advancement of AI and NLP enable us to work on 'understanding' and 'predicting,' but not 'nudging' online human behavior. I aim to pursue working in the area of 'nudging' human behavior. I have worked on a fascinating project, Precision Public Health Campaigns, to determine how we could (i) use data analysis and machine learning to create the right message for the right person; and (ii) evaluate the impact on offline behavior beyond simple impressions and click-based metrics. Digital channels facilitate the tailoring and distribution of messages to particular target audiences and

decide on an effective approach and content. It enables us to do A/B testing and evaluate which message works best for which sub-population in creating physical participation, and not just a digital ‘Like.’ We address the opportunities and challenges of using targeted advertising tools for personalized health messaging for public health campaigns [27]. We show that tailored messages by gender, culture, and psychological attributes are more effective for user engagement. I further aim to investigate its effectiveness on real-world activities—how online advertising increases offline activities, such as benefiting from vaccinations or cancer screenings.

Urban computing. The ‘social’ context is imperative for tackling urban computing, and a framework that could effectively exploit social context would be helpful. Urban challenges require considering multiple factors—for example, if wanting to predict traffic congestion, not only the road conditions, but also other contexts such as spatial characteristics such as point of interest and popular activities, nearby social events, or even weather. The deep neural architecture has been more effective in integrating all those various types of information, resulting in much better performance in tackling such urban problems. I recently started collaborating with researchers at KAIST (the Korea Advanced Institute of Science and Technology) to develop a deep learning framework for multi-modal and multi-dimensional data, that enables us to tackle various urban problems, such as traffic congestion, urban sprawl, gentrification, and real estate prices.

Concluding remarks. AI and data science have opened a new door to developing methods that can leverage large-scale online data to understand individuals and society, provide decision-makers with insights into the impacts of online data, and build systems to support organizations. I have remained fascinated by the opportunities at the intersection of AI and social science. Fortunately, I have worked not only with computer science researchers but also with communication scholars, political scientists, psychologists, and journalists. This diversity of academic training has brought unique perspectives to the table. While collaborations are always tricky, they are also the most effective and engaging way to perform multidisciplinary research; I plan to continue conducting research in this way. Lastly, AI provides excellent convenience and enables us to tackle various social problems. However, even without wrongful intentions, it can bring unexpected negative consequences. My previous work on trending topics already showed how widely-used algorithms could easily marginalize minority groups on social media. AI’s negative impact should be carefully considered as important as its promises. I believe that I can study such problems with the eyes of a social scientist and the hands of a computer scientist.

References

- [1] Saima Salim at Digital Information World. More than six hours of our day is spent online, 2019. <https://www.digitalinformationworld.com/2019/02/internet-users-spend-more-than-a-quarter-of-their-lives-online.html>.
- [2] Famemass. How much time do people spend on social media in 2019?, 2019. <https://famemass.com/time-spent-on-social-media>.
- [3] William L. Hamilton, Jure Leskovec, and Dan Jurafsky. Diachronic word embeddings reveal statistical laws of semantic change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1489–1501. Association for Computational Linguistics, 2016.
- [4] Jisun An, Haewoon Kwak, and Yong-Yeol Ahn. SemAxis: A lightweight framework to characterize domain-specific word semantics beyond sentiment. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, July 2018.
- [5] Haewoon Kwak, Jisun An, Elise Jing, and Yong-Yeol Ahn. Frameaxis: characterizing microframe bias and intensity with word embedding. *PeerJ Computer Science*, 7:e644, 2021.
- [6] Soo-Yong Shin, Dong-Woo Seo, Jisun An, Haewoon Kwak, Sung-Han Kim, Jin Gwack, and Min-Woo Jo. High correlation of middle east respiratory syndrome spread with google search and twitter trends in korea. *Scientific reports*, 6:32920, 2016.
- [7] Jisun An and Ingmar Weber. Whom should we sense in “social sensing” - analyzing which users work best for social media now-casting. *EPJ Data Science*, 4(1):22, 2015.
- [8] Kholoud Khalil Aldous, Jisun An, and Bernard J. Jansen. Stylistic features usage: Similarities and differences using multiple social networks. In *International Conference on Social Informatics*, pages 309–318. Springer, 2019.
- [9] Kholoud Khalil Aldous, Jisun An, and Bernard J Jansen. View, like, comment, post: Analyzing user engagement by topic at 4 levels across 5 social media platforms for 53 news organizations. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 13, pages 47–57, 2019.

- [10] Kholoud Khalil Aldous, Jisun An, and Bernard J. Jansen. Predicting audience engagement across social media platforms in the news domain. In *International Conference on Social Informatics*, pages 173–187. Springer, 2019.
- [11] Jisun An, Haewoon Kwak, Claire Seungeun Lee, Bogang Jun, and Yong-Yeol Ahn. Predicting anti-Asian hateful users on Twitter during COVID-19. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 4655–4666, Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.
- [12] George Gerbner. Cultivation analysis: An overview. *Mass communication and society*, 1(3-4):175–194, 1998.
- [13] Jisun An, Meeyoung Cha, Krishna Gummadi, and Jon Crowcroft. Media landscape in twitter: A world of new conventions and political diversity. In *Fifth International AAAI Conference on Weblogs and Social Media*, 2011.
- [14] Jisun An, Meeyoung Cha, Krishna Gummadi, Jon Crowcroft, and Daniele Quercia. Visualizing media bias through twitter. In *Sixth International AAAI Conference on Weblogs and Social Media*, 2012.
- [15] Ramy Baly, Georgi Karadzhov, Jisun An, Haewoon Kwak, Yoan Dinkov, Ahmed Ali, James Glass, and Preslav Nakov. What was written vs. who read it: News media profiling using text analysis and social media context. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 3364–3374, Online, July 2020. Association for Computational Linguistics.
- [16] Soumen Ganguly, Juhi Kulshrestha, Jisun An, and Haewoon Kwak. Empirical evaluation of three common assumptions in building political media bias datasets. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, pages 939–943, 2020.
- [17] Haewoon Kwak, Jisun An, and Yong-Yeol Ahn. A systematic media frame analysis of 1.5 million new york times articles from 2000 to 2017. In *12th ACM Conference on Web Science*, pages 305–314, 2020.
- [18] QCRI. Tanbih, 2019. <https://tanbih.org>.
- [19] Yifan Zhang, Giovanni Da San Martino, Alberto Barrón-Cedeño, Salvatore Romeo, Jisun An, Haewoon Kwak, Todor Staykovski, Israa Jaradat, Georgi Karadzhov, Ramy Baly, Kareem Darwish, James Glass, and Preslav Nakov. Tanbih: Get to know what you are reading. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP): System Demonstrations*, pages 223–228, Hong Kong, China, November 2019. Association for Computational Linguistics.
- [20] Jisun An, Daniele Quercia, and Jon Crowcroft. Partisan sharing: Facebook evidence and societal consequences. In *Proceedings of the second ACM conference on Online social networks*, pages 13–24. ACM, 2014.
- [21] Jisun An, Haewoon Kwak, Oliver Posegga, and Andreas Jungherr. Political discussions in homogeneous and cross-cutting communication spaces. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 13, pages 68–79, 2019.
- [22] Haewoon Kwak and Jisun An. Revealing the hidden patterns of news photos: Analysis of millions of news photos through gdelt and deep learning-based vision apis. *Proceedings of the International AAAI Conference on Web and Social Media*, 10(2):99–107, Aug. 2016.
- [23] Jisun An and Haewoon Kwak. Gender and racial diversity in commercial brands’ advertising images on social media. In *International Conference on Social Informatics*, pages 79–94. Springer, 2019.
- [24] Jisun An and Ingmar Weber. # greysanatomy vs.# yankees: Demographics and hashtag use on twitter. In *Tenth International AAAI Conference on Web and Social Media*, 2016.
- [25] Jisun An and Haewoon Kwak. Multidimensional analysis of the news consumption of different demographic groups on a nationwide scale. In *International Conference on Social Informatics*, pages 124–142. Springer, 2017.
- [26] Andreas Jungherr, Oliver Posegga, and Jisun An. Discursive power in contemporary media systems: A comparative framework. *The International Journal of Press/Politics*, 24(4):404–425, 2019.
- [27] Jisun An, Haewoon Kwak, Hanya M Qureshi, Ingmar Weber, et al. Precision public health campaign: Delivering persuasive messages to relevant segments through targeted advertisements on social media. *JMIR Formative Research*, 5(9):e22313, 2021.