
CONTACT	Luddy School of Informatics, Computing, and Engineering Indiana University Bloomington 47408, IN, USA	<i>E-mail:</i> hwkwak@iu.edu http://haewoon.github.io http://soda-labo.github.io/
HIGHLIGHTS	Highly cited research (H-index: 46, i10-index: 95, Lifetime citations > 22k)	
	<ul style="list-style-type: none"> • Twitter as an information network [C8]: 10,441 times • YouTube video popularity characterization [C3]: 2,214 times • Huge online social network (OSN) analysis [C2]: 1,422 times 	
EDUCATION	Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea	
	Ph.D. in Computer Science (4 years) Advisor: Professor Sue Moon Dissertation: Multi-level Analysis on Structures and Dynamics of Online Social Networks	August 2011
	M.S. in Computer Science (1.5 years)	August 2007
	B.S. in Computer Science (3.5 years)	February 2006
PROFESSIONAL EXPERIENCE	Indiana University Bloomington, IN, USA	
	Associate Professor	Jan 2023 - Present
	Singapore Management University, Singapore	
	Associate Professor	Oct 2020 - Dec 2022
	Qatar Computing Research Institute, HBKU, Doha, Qatar	
	Senior Scientist	Jul 2016 - Sep 2020
	Scientist	Apr 2014 - Jun 2016
	Telefonica Research, Barcelona, Spain	
	Researcher	Jun 2012 - Apr 2014
	KAIST, Daejeon, Korea	
	Research Assistant	Aug 2011 - Oct 2011
	Telefonica Research, Barcelona, Spain	
	Intern	Sep 2007 - Dec 2007
AWARDS	2025 IU Trustees Teaching Award	
	2024 Luddy Faculty Fellow (2024-25)	
	2023 Best Metareviewer award, The Web conference, Austin, Texas, US	
	2022 ACM Internet Measurement Conference Test of Time Award, Nice, France	

- 2021 The D.S. Lee Foundation Fellowship (2021-22)
- 2021 [AI 2000 Most Influential Scholars Honorable Mention \(Human-Computer Interaction, 2011-2020\)](#) by AMiner
- 2021 Best Program Committee Member, The Web conference, Virtual conference
- 2020 [AI 2000 Most Influential Scholars \(Information Retrieval and Recommendation, 2010-2019\)](#) by AMiner
- 2019 Best paper award, International Conference on Social Informatics (SocInfo), Doha, Qatar
- 2018 [HBKU Innovation Idea Development Grants](#), Doha, Qatar
- 2016 [Qatar Science & Technology Park \(QSTP\) Accelerator program](#), Doha, Qatar
- 2011 Best Ph.D. Thesis, KAIST Computer Science Department, Daejeon, Korea
- 2007 Best paper award, ACM Internet Measurement Conference, San Diego, CA, USA

FUNDING

Current Awarded Grants

- Global Research Support Program in the Digital Field Program (Task: Development of Reliable, Secure, and Safe Human-AI Alignment Techniques), *Co-PI* (PI: Jinyoung Han, SKKU, Korea), KRW 681,000,000, Ministry of Science and ICT, Korea, 2024–2026
- BRAIN (Belief Resonance and AI Narratives): Understanding Belief Narrative Resonance in the Era of Generative AI, *Key-personnel* (PI: Young-Yeol Ahn, IU), USD 7,500,000, DoD AFOSR MURI, 2024–2029

Past Awarded Grants

- Developing AI systems to monitor diversity and fairness in media, *PI*, SGD 99,998, Ministry of Education (MOE) Tier 1 Academic Research Fund (AcRF), Singapore, 2020–2021

PUBLICATIONS

Journals

- [J20] B. Lee, R. Aiyappa, Y.-Y. Ahn, **H. Kwak**, and J. An. A semantic embedding space based on large language models for modeling human beliefs. *Nature Human Behaviour*, 2025
- [J19] K. Miyazaki, T. Uchiba, **H. Kwak**, J. An, and K. Sasahara. The influence of toxic trolling comments on anti-vaccine youtube videos. *Scientific Reports*, 14(5088), 2024
- [J18] K. Miyazaki, T. Murayama, T. Uchiba, J. An, and **H. Kwak**. Public perception of generative AI on Twitter: An empirical study based on occupation and usage. *EPJ Data Science*, 13, 2024
- [J17] Y. Mejova, J. An, G.D.F. Morales, and **H. Kwak**. Modeling political activism around gun debate via social media. *ACM Transactions on Social Computing*, 2022
- [J16] H. Almerexhi, **H. Kwak**, J. Salminen, and B. J. Jansen. Provoke: Toxicity trigger detection in conversations from the top 100 subreddits. *Data and Information Management*, page 100019, 2022

- [J15] H. Almerekhi, **H. Kwak**, and B.J. Jansen. Investigating toxicity changes of cross-community redditors from 2 billion posts and comments. *PeerJ Computer Science*, 8:e1059, 2022
- [J14] Sercan Şengün, Peter Mawhorter, James Bowie-Wilson, Yusef Audeh, **Haewoon Kwak**, and D. Fox Harrell. Contours of virtual enfreakment in fighting game characters. *Technological Forecasting and Social Change*, 180:121707, 2022
- [J13] J. An, **H. Kwak**, H. M. Qureshi, and I. Weber. Precision public health campaign: Delivering persuasive messages to relevant segments through targeted advertisements on social media. *JMIR Formative Research*, 5(9):e22313, 2021
- [J12] **H. Kwak**, J. An, E. Jing, and Y.-Y. Ahn. Frameaxis: characterizing microframe bias and intensity with word embedding. *PeerJ Computer Science*, 7:e644, 2021
- [J11] J. Salminen, J. M. Santos, **H. Kwak**, J. An, S. Jung, and B. J. Jansen. Persona perception scale: Development and exploratory validation of an instrument for evaluating individuals' perceptions of personas. *International Journal of Human-Computer Studies*, 141:102437, 2020
- [J10] A. Zubiaga, B. Heravi, J. An, and **H. Kwak**. Social media mining for journalism. *Online Information Review*, 43(1):2–6, 2019
- [J9] J. Salminen, S. Jung, J. An, **H. Kwak**, L. Nielsen, and B. J. Jansen. Confusion and information triggered by photos in persona profiles. *International Journal of Human-Computer Studies*, 129:1–14, 2019
- [J8] J. An, **H. Kwak**, S. Jung, J. Salminen, M. Admad, and B.J. Jansen. Imaginary people representing real numbers: generating personas from online social media data. *ACM Transactions on the Web (TWEB)*, 12(4):27, 2018
- [J7] J. An, **H. Kwak**, S. Jung, J. Salminen, and B. J. Jansen. Customer segmentation using online platforms: isolating behavioral and demographic segments for persona creation via aggregated user data. *Social Network Analysis and Mining*, 8(1):54, 2018
- [J6] J. Salminen, B. J. Jansen, J. An, **H. Kwak**, and S. Jung. Are personas done?: Evaluating the usefulness of personas in the age of online analytics. *Persona Studies*, 4(2):47–65, 2018
- [J5] P. Mawhorter, S. Şengün, **H. Kwak**, and D.F. Harrell. Identifying regional trends in avatar customization. *IEEE Transactions on Games*, pages 1–11, 2018
- [J4] J. Salminen, S. Şengün, **H. Kwak**, B. J. Jansen, J. An, S. Jung, and S. Vieweg. From 2,772 segments to five personas: Summarizing a diverse online audience by generating culturally adapted personas. *First Monday*, 23(6), 2018
- [J3] S.-Y. Shin, D.-W. Seo, J. An, **H. Kwak**, S.-H. Kim, J. Gwack, and M.-W. Jo. High correlation of Middle East respiratory syndrome spread with Google search and Twitter trends in Korea. *Scientific reports*, 6:32920, 2016
- [J2] **H. Kwak**, S. Moon, Y. Eom, Y. Choi, and H. Jeong. Consistent community identification in complex networks. *Journal of the Korean Physical Society*, 59(5):3128–3132, 2011
- [J1] M. Cha, **H. Kwak**, P. Rodriguez, Y. Ahn, and S. Moon. Analyzing the video popularity characteristics of large-scale user generated content systems. *IEEE/ACM Transactions on Networking*, 17(5):1357–1370, Oct 2009

Conferences

- [C90] Zoher Kachwala, Bao Tran Truong, Rasika Muralidharan, **H. Kwak**, Jisun An, and Filippo Menczer. PLURULE: A challenging benchmark for detecting rule violations of pluralistic communities on social media. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Accepted)*, 2026

- [C89] Jisun An Fan Huang, **H. Kwak**. Vulnerability of llms’ stated belief? llms belief resistance check through strategic persuasive conversation interventions. In *Findings of the Association for Computational Linguistics: ACL 2026 (Accepted)*, 2026
- [C88] M. Kang, J. Kim, H. Jeon, S. Hwang, J. An, Y. Kim, **H. Kwak**, J. An, and J. Han. SynSym: A synthetic data generation framework for psychiatric symptom identification. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD ’26)*, 2026
- [C87] H. Won, **H. Kwak**, J. An, and J. Han. When behavior is not enough: Identity biases in human–ai creative collaboration. In *The ACM CHI conference on Human Factors in Computing Systems (CHI’26) Poster*, 2026
- [C86] C. Kang, Y. Lee, **H. Kwak**, J. An, and J. Han. If i change your eyes, are you still you? identity thresholds in deepfake perception. In *The ACM CHI conference on Human Factors in Computing Systems (CHI’26) Poster*, 2026
- [C85] W. Qi, F. Huang, J. An, and **H. Kwak**. A cross-cultural comparison of LLM-based public opinion simulation: Evaluating Chinese and U.S. models on diverse societies. In *Workshop on AI for Computational Social Science at ICDM*, 2025
- [C84] Z. Kachwala, J. An, **H. Kwak**, and F. Menczer. Rematch: Robust and efficient matching of local knowledge graphs to improve structural and semantic similarity. In *Findings of the Association for Computational Linguistics: NAACL 2024*, pages 1018–1028, 2024
- [C83] P. Nakov, J. An, **H. Kwak**, M. A. Manzoor, Z. Mujahid, and H. Sencar. A survey on predicting the factuality and the bias of news media. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, *Findings of the Association for Computational Linguistics: ACL 2024*, pages 15947–15962, Bangkok, Thailand, August 2024. Association for Computational Linguistics
- [C82] F. Huang, **H. Kwak**, K. Park, and J. An. ChatGPT rates natural language explanation quality like humans: But on which scales? In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 3111–3132, 2024
- [C81] F. Huang, **H. Kwak**, and J. An. Is ChatGPT better than human annotators? potential and limitations of ChatGPT in explaining implicit hate speech. In *Companion Proceedings of the ACM Web Conference 2023*, pages 294–297, 2023
- [C80] F. Huang, **H. Kwak**, and J. An. Chain of explanation: New prompting method to generate quality natural language explanation for implicit hate speech. In *Companion Proceedings of the ACM Web Conference 2023*, pages 90–93, 2023
- [C79] R. Aiyappa, J. An, **H. Kwak**, and Y.-Y. Ahn. Can we trust the evaluation on ChatGPT? In *Proceedings of the 3rd Workshop on Trustworthy Natural Language Processing (TrustNLP 2023)*, pages 47–54, Toronto, Canada, July 2023. Association for Computational Linguistics
- [C78] S. Y. Liaw, F. Huang, F. Benevenuto, **H. Kwak**, and J. An. Younicon: Youtube’s community of conspiracy videos. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 17, pages 1102–1111, 2023
- [C77] T. Uchiba K. Miyazaki, K. Tanaka, J. An, **H. Kwak**, and K. Sasahara. “This is fake news”: Characterizing the spontaneous debunking from Twitter users to COVID-19 false information. In *Proceedings of the 17th International AAAI Conference on Web and Social Media (ICWSM)*, 2023
- [C76] H. Zhang, **H. Kwak**, W. Gao, and J. An. Wearing masks implies refuting Trump?: Towards target-specific user stance prediction across events in COVID-19 and US election 2020. In *Proceedings of the 15th International ACM Conference on Web Science (WebSci)*. ACM, 2023

- [C75] K. Miyazaki, T. Murayama, A. Matsui, M. Nishikawa, T. Uchiba, **H. Kwak**, and J. An. Political honeymoon effect on social media: Characterizing social media reaction to the changes of prime minister in Japan. In *Proceedings of the 15th International ACM Conference on Web Science (WebSci)*. ACM, 2023
- [C74] K. Miyazaki, T. Uchiba, **H. Kwak**, and J. An. MAANG? MANGA? characterizing spontaneous ideation contest on social media: Case study on the name change of Facebook to Meta. In *2022 IEEE International Conference on Big Data (Big Data)*, pages 776–783, 2022
- [C73] **H. Kwak**. You have earned a trophy: Characterize in-game achievements and their completions. In *Proceedings of the 14th International ACM Conference on Web Science (WebSci)*. ACM, 2022
- [C72] Y. Y. Chong and **H. Kwak**. Understanding toxicity triggers on reddit in the context of singapore. In *Proceedings of the 16th International AAAI Conference on Web and Social Media (ICWSM) (Short paper)*, 2022
- [C71] **H. Kwak**, J. An, and K. Park. Who is missing? characterizing the participation of different demographic groups in a korean nationwide daily conversation corpus. In *Proceedings of the 16th International AAAI Conference on Web and Social Media (ICWSM) (Short paper)*, 2022
- [C70] J. An, **H. Kwak**, C. S. Lee, B. Jun, and Y.-Y. Ahn. Predicting anti-asian hateful users on twitter during covid-19. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 4655–4666, 2021
- [C69] K. Park, **H. Kwak**, J. An, and S. Chawla. How-to present news on social media: A causal analysis of editing news headlines for boosting user engagement. In *Proceedings of the 15th International AAAI Conference on Web and Social Media (ICWSM)*, 2021
- [C68] **H. Kwak**, J. An, and Y.-Y. Ahn. A systematic media frame analysis of 1.5 million new york times articles from 2000 to 2017. In *Proceedings of the 12th International ACM Conference on Web Science (WebSci)*. ACM, 2020
- [C67] H. Almerexhi, S. B. J. Jansen, and **H. Kwak**. Investigating toxicity across multiple reddit communities, users, and moderators. In *Companion proceedings of the web conference 2020*, pages 294–298, 2020
- [C66] R. Baly, J. An G. Karadzhov, **H. Kwak**, Y. Dinkov, A. Ali, J. Glass, and P. 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 (ACL)*. Association for Computational Linguistics, 2020
- [C65] K. Park, **H. Kwak**, H. Song, and M. Cha. “Trust me, I have a Ph.D.”: Exposing a high offline status leads to heightened positive interactions in online communities. In *Proceedings of the 14th International AAAI Conference on Web and Social Media (ICWSM)*, 2020
- [C64] S. Ganguly, J. Kulshrestha, J. An, and **H. Kwak**. Empirical evaluation of three common assumptions in building political media bias datasets. In *Proceedings of the 14th International AAAI Conference on Web and Social Media (ICWSM) (short paper)*, 2020
- [C63] H. Almerexhi, **H. Kwak**, B. Jansen, and J. Salminen. Are these comments triggering? Predicting triggers of toxicity in online discussions. In *Proceedings of the The Web Conference (short paper)*. International World Wide Web Conferences Steering Committee, 2020
- [C62] Y. Zhang, G.D.S. Martino, A. Barrón-Cedeño, S. Romeo, J. An, **H. Kwak**, T. Staykovski, I. Jaradat, G. Karadzhov, R. Baly, K. Darwish, J. Glass, and P. 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

- [C61] J. Salminen, M. Nagpal, **H. Kwak**, J. An, S. Jung, and B. J. Jansen. Confusion prediction from eye-tracking data: Experiments with machine learning. In *Proceedings of the 9th International Conference on Information Systems and Technologies*, pages 1–11, 2019
- [C60] J. An and **H. Kwak**. Gender and racial diversity in commercial brands’ advertising images on social media. In *International Conference on Social Informatics (SocInfo)*. Springer, 2019
- [C59] H. Almerexhi, **H. Kwak**, B.J. Jansen, and J. Salminen. Detecting toxicity triggers in online discussions. In *Proceedings of the 30th ACM Conference on Hypertext and Social Media*, pages 291–292. ACM, 2019
- [C58] J. An, **H. Kwak**, O. Posegga, and A. Jungherr. Political discussions in homogeneous and cross-cutting communication spaces. In *Proceedings of the 13th International AAAI Conference on Web and Social Media (ICWSM)*, volume 13, pages 68–79, 2019
- [C57] S. Jung, J. An, **H. Kwak**, J. Salminen, and B.J. Jansen. Assessing the accuracy of four popular face recognition tools for inferring gender, age, and race. In *Proceedings of the 12th International AAAI Conference on Web and Social Media (ICWSM)*, pages 624–627, 2018
- [C56] J. Salminen, H. Almerexhi, M. Milenković, S. Jung, J. An, **H. Kwak**, , and B.J. Jansen. Anatomy of online hate: developing a taxonomy and machine learning models for identifying and classifying hate in online news media. In *Proceedings of the 12th International AAAI Conference on Web and Social Media*, pages 330–339, 2018
- [C55] J. An, **H. Kwak**, and Y.-Y. 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)*, pages 2450–2461, Melbourne, Australia, July 2018. Association for Computational Linguistics
- [C54] S. Zannettou, B. Bradlyn, E. De Cristofaro, **H. Kwak**, M. Sirivianos, G. Stringini, and J. Blackburn. What is Gab: A bastion of free speech or an alt-right echo chamber. In *Companion Proceedings of the The Web Conference 2018*, pages 1007–1014. International World Wide Web Conferences Steering Committee, 2018
- [C53] S. Jung, J. O. Salminen, J. An, **H. Kwak**, and B. J. Jansen. Automatically conceptualizing social media analytics data via personas. In *Twelfth International AAAI Conference on Web and Social Media*, 2018
- [C52] J. Salminen, B.J. Jansen, J. An, S. Jung, L. Nielsen, and **H. Kwak**. Fixation and confusion: Investigating eye-tracking participants’ exposure to information in personas. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval*, pages 110–119. ACM, 2018
- [C51] J. Salminen, **H. Kwak**, J. Santos, S. Jung, J. An, and B.J. Jansen. Persona perception scale: developing and validating an instrument for human-like representations of data. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, page LBW075. ACM, 2018
- [C50] J. Salminen, L. Nielsen, S. Jung, J. An, **H. Kwak**, and B. J. Jansen. ”is more better?” impact of multiple photos on perception of persona profiles. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2018

- [C49] **H. Kwak**, J. An, J. Salminen, S.-G. Jung, and B. Jansen. What we read, what we search: Media attention and public attention among 193 countries. In *Proceedings of the 2018 World Wide Web Conference*, pages 893–902. International World Wide Web Conferences Steering Committee, 2018
- [C48] J. Salminen, S. Jung, J. An, **H. Kwak**, and B. J. Jansen. Findings of a user study of automatically generated personas. In *Extended abstracts of the 2018 CHI conference on human factors in computing systems*, pages 1–6, 2018
- [C47] S. Jung, J. Salminen, **H. Kwak**, J. An, and B. J. Jansen. Automatic persona generation (apg) a rationale and demonstration. In *Proceedings of the 2018 conference on human information interaction and retrieval*, pages 293–296, 2018
- [C46] J. An and **H. Kwak**. Multidimensional analysis of the news consumption of different demographic groups on a nationwide scale. In *International Conference on Social Informatics (SocInfo)*, pages 124–142. Springer, 2017
- [C45] J. An, H. Aldarbesti, and **H. Kwak**. Convergence of media attention across 129 countries. In *International Conference on Social Informatics (SocInfo)*, pages 159–168. Springer, 2017
- [C44] S. Jung, J. An, **H. Kwak**, M. Ahmad, L. Nielsen, and B. J. Jansen. Persona generation from aggregated social media data. In *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems*, pages 2133–2139, 2017
- [C43] **H. Kwak** and J. An. Multiplex media attention and disregard network among 129 countries. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) 2017*, pages 1225–1232. ACM, 2017
- [C42] B. J. Jansen, S. Jung, J. Salminen, J. An, and **H. Kwak**. Viewed by too many or viewed too little: Using information dissemination for audience segmentation. In *Proceedings of the Association for Information Science and Technology*, volume 54, pages 740–742, 2017
- [C41] S. Jung, J. An, **H. Kwak**, J. Salminen, and B. J. Jansen. Inferring social media users’ demographics from profile pictures: A face++ analysis on twitter users. In *Proceedings of The 17th International Conference on Electronic Business*, pages 140–145, 2017
- [C40] J. Salminen, S. Sengun, **H. Kwak**, B. Jansen, J. An, S. Jung, and S. Vieweg. Generating cultural personas from social data: a perspective of middle eastern users. In *2017 5th international conference on future internet of things and cloud (FiCloud)*, pages 255–262. IEEE, 2017
- [C39] J. Reis, **H. Kwak**, J. An, J. Messias, and F. Benevenuto. Demographics of news sharing in the US Twittersphere. In *Proceedings of the 28th ACM Conference on Hypertext and Social Media*, pages 195–204. ACM, 2017
- [C38] D.F. Harrell, S. Vieweg, **H. Kwak**, C.-U. Lim, S. Sengun, A. Jahanian, and P. Ortiz. Culturally-grounded analysis of everyday creativity in social media: A case study in qatari context. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition*, pages 209–221. ACM, 2017
- [C37] J. An, **H. Kwak**, and B. J. Jansen. Personas for content creators via decomposed aggregate audience statistics. In *Proceedings of the 2017 IEEE/ACM international conference on advances in social networks analysis and mining 2017*, pages 1273–1276, 2017
- [C36] J. An, **H. Kwak**, and B. J. Jansen. Automatic generation of personas using youtube social media data. In *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017

- [C35] J. An and **H. Kwak**. Data-driven approach to measuring the level of press freedom using media attention diversity from unfiltered news. In *The Workshops of News and Public Opinion (NECO) co-located with Eleventh International AAAI Conference on Web and Social Media (AAAI Technical Report WS-17-17)*, 2017
- [C34] J. An and **H. Kwak**. What gets media attention and how media attention evolves over time: large-scale empirical evidence from 196 countries. In *Proceedings of the 11th International AAAI Conference on Web and Social Media (ICWSM)*, pages 464–467, 2017
- [C33] L. Nielsen, S. Jung, J. An, J. Salminen, **H. Kwak**, and B. J. Jansen. Who are your users? comparing media professionals’ preconception of users to data-driven personas. In *Proceedings of the 29th Australian conference on computer-human interaction*, pages 602–606, 2017
- [C32] A.R. Kang, J. Blackburn, **H. Kwak**, and H.K. Kim. I would not plant apple trees if the world will be wiped: Analyzing hundreds of millions of behavioral records of players during an MMORPG beta test. In *Proceedings of the 26th International Conference on World Wide Web Companion*, pages 435–444. International World Wide Web Conferences Steering Committee, 2017
- [C31] K. Park, M. Cha, **H. Kwak**, and K.-T. Chen. Achievement and friends: Key factors of player retention vary across player levels in online multiplayer games. In *Proceedings of the 26th International Conference on World Wide Web Companion*, pages 445–453. International World Wide Web Conferences Steering Committee, 2017
- [C30] **H. Kwak** and J. An. Revealing the hidden patterns of news photos: Analysis of millions of news photos through GDELT and deep learning-based vision APIs. In *The workshop on News and Public Opinion co-located with Tenth International AAAI Conference on Web and Social Media (AAAI Technical Report WS-16-18)*, 2016
- [C29] **H. Kwak** and J. An. Two tales of the world: Comparison of widely used world news datasets GDELT and EventRegistry. In *Tenth International AAAI Conference on Web and Social Media*, pages 619–622, 2016
- [C28] J. An, **H. Kwak**, and B. J. Jansen. Validating social media data for automatic persona generation. In *2016 IEEE/ACS 13th international conference of computer systems and applications (AICCSA)*, pages 1–6. IEEE, 2016
- [C27] J. An, H. Cho, **H. Kwak**, M. Z. Hassen, and B. J. Jansen. Towards automatic persona generation using social media. In *2016 IEEE 4th international conference on future internet of things and cloud (FiCloud)*, pages 216–221. IEEE, 2016
- [C26] J. An, **H. Kwak**, Y. Mejova, S.A.S. De Ogar, and B.G. Fortes. Are you Charlie or Ahmed? cultural pluralism in Charlie Hebdo response on Twitter. In *Proceedings of the 10th International AAAI Conference on Web and Social Media (ICWSM)*, pages 2–11, 2016
- [C25] S. Abbar, A. Jisun, **H. Kwak**, M. Yacine, and B.-H. Javier. Consumers and suppliers: attention asymmetries. a case study of Al Jazeera’s news coverage and comments. In *Computational Journalism Symposium*, volume 2, 2015
- [C24] D. Reis, F. Benevenuto, P. Olmo, P. Raquel, **H. Kwak**, and J. An. Breaking the news: First impressions matter on online news. In *Ninth International AAAI Conference on Web and Social Media (ICWSM)*, 2015
- [C23] **H. Kwak**, J. Blackburn, and S. Han. Exploring cyberbullying and other toxic behavior in team competition online games. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI)*, pages 3739–3748. ACM, 2015

- [C22] I. Leontiadis, A. Lima, **H. Kwak**, R. Stanojevic, D. Wetherall, and D. Papagiannaki. From cells to streets: Estimating mobile paths with cellular-side data. In *Proceedings of the 10th ACM International on Conference on emerging Networking Experiments and Technologies (CoNEXT)*, pages 121–132. ACM, 2014
- [C21] **H. Kwak** and J. An. Understanding news geography and major determinants of global news coverage of disasters. In *Computation+ Journalism Symposium*, 2014
- [C20] **H. Kwak** and J. An. A first look at global news coverage of disasters by using the GDELT dataset. In *International Conference on Social Informatics (SocInfo)*, pages 300–308. Springer, 2014
- [C19] **H. Kwak** and J. Blackburn. Linguistic analysis of toxic behavior in an online video game. In *Workshop on Exploration on Games and Gamers co-located with international Conference on Social Informatics*, pages 209–217. Springer, 2014
- [C18] L. Gyarmati, **H. Kwak**, and P. Rodriguez. Searching for a unique style in soccer. In *2014 KDD Workshop on Large-Scale Sports Analytics*, 2014
- [C17] M. Martin, R.D. Oliveira, **H. Kwak**, , and N. Oliver. Didn’t you see my message?: predicting attentiveness to mobile instant messages. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3319–3328. ACM, 2014
- [C16] J. Blackburn and **H. Kwak**. STFU NOOB!: predicting crowdsourced decisions on toxic behavior in online games. In *Proceedings of the 23rd international conference on World wide web*, pages 877–888. ACM, 2014
- [C15] Y. Hong, **H. Kwak**, Y. Baek, and M. Sue. Tower of babel: A crowdsourcing game building sentiment lexicons for resource-scarce languages. In *Proceedings of the 22nd International Conference on World Wide Web, WWW ’13 Companion*, pages 549–556, New York, NY, USA, 2013. ACM
- [C14] B. Xu, Y. Huang, **H. Kwak**, and N. Contractor. Structures of broken ties: Exploring unfollow behavior on Twitter. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work, CSCW ’13*, pages 871–876, New York, NY, USA, 2013. ACM
- [C13] **H. Kwak**, S. Moon, and W. Lee. More of a receiver than a giver: why do people unfollow in Twitter? In *Sixth International AAAI Conference on Weblogs and Social Media*, pages 499–502, 2012
- [C12] **H. Kwak**, Y. Hong, J. You, and S. Moon. Visualizing a personal timeline by adding multiple social contexts. In *The workshop on Social Media Visualization co-located with International AAAI Conference on Web and Social Media (AAAI Technical Report WS-12-03)*, 2012
- [C11] **H. Kwak**, H. Chun, and S. Moon. Fragile online relationship: a first look at unfollow dynamics in Twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1091–1100. ACM, 2011
- [C10] C. Lee, **H. Kwak**, H. Park, and S. Moon. Finding influentials based on the temporal order of information adoption in Twitter. In *Proceedings of the 19th international conference on World wide web*, pages 1137–1138. ACM, 2010
- [C9] **H. Kwak**, C. Lee, H. Park, and S. Moon. What is Twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, pages 591–600. ACM, 2010
- [C8] S. Moon, J. You, **H. Kwak**, D. Kim, and H. Jeong. Understanding topological mesoscale features in community mining. In *2010 Second International Conference on COMMunication Systems and NETWORKS (COMSNETS 2010)*, pages 1–10. IEEE, 2010

- [C7] **H. Kwak**, Y. Choi, Y. Eom, H. Jeong, and S. Moon. Mining communities in networks: a solution for consistency and its evaluation. In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement*, pages 301–314. ACM, 2009
- [C6] **H. Kwak**, H. Y. Shin, J. I. Yoon, and S. Moon. Connecting users with similar interests across multiple web services. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 3, pages 254–257, 2009
- [C5] X. Amatriain, N. Lathia, J. Puyol, **H. Kwak**, and N. Oliver. The wisdom of the few: a collaborative filtering approach based on expert opinions from the web. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 532–539. ACM, 2009
- [C4] H. Chun, **H. Kwak**, Y. Eom, Y.-Y. Ahn, S. Moon, and H. Jeong. Comparison of online social relations in volume vs interaction: A case study of cyworld. In *Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement*, IMC '08, pages 57–70, New York, NY, USA, 2008. ACM
- [C3] M. Cha, **H. Kwak**, P. Rodriguez, Y.-Y. Ahn, and S. Moon. I tube, you tube, everybody tubes: Analyzing the world’s largest user generated content video system. In *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*, IMC '07, pages 1–14, New York, NY, USA, 2007. ACM
- [C2] Y.-Y. Ahn, S. Han, **H. Kwak**, S. Moon, and H. Jeong. Analysis of topological characteristics of huge online social networking services. In *Proceedings of the 16th International Conference on World Wide Web*, WWW '07, pages 835–844, New York, NY, USA, 2007. ACM
- [C1] **H. Kwak**, S. Han, Y.-Y. Ahn, S. Moon, and H. Jeong. Impact of snowball sampling ratios on network characteristics estimation: A case study of cyworld. In *Proceedings of the Korean Information Science Society Conference*, 2006

Book chapters

- [B1] J. Salminen, B. J. Jansen, J. An, **H. Kwak**, and S. Jung. Automatic persona generation for online content creators: Conceptual rationale and a research agenda. In *Personas-User focused design*, pages 135–160. Springer, 2019

Patents

- [P1] S. Moon, **H. Kwak**, H. Shin, and J. Yoon. System and method for extracting users of similar interests between various types of web servers, April 16 2013. US Patent 8,423,542

TEACHING
EXPERIENCE

Indiana University Bloomington, US

- [I590: Topics in Informatics - Data Visualization](#)
FA23, FA24, SP25, FA25, SP26 (Online)
- I210: Intro to Programming
SP23

Singapore Management University, Singapore

- [IS457: Fairness in Socio-technical Systems](#)
FA21, FA22
- IS111: Intro to Programming
FA21, FA22

- IS112: Data Management
SP21
- SMT483: Project Experience
SP21

PROFESSIONAL
SERVICE

Editorial services

- Editorial board member, [Humanities and Social Sciences Communications](#) (published by Springer Nature, 2023 Impact Factor: 3.7): 2025-present
- Editorial board member, [Online Social Networks and Media](#) (Science Citation Index Expanded journal): 2019-present
- Editorial board member, [Digital Journalism](#) (2018 Impact Factor 2.679): 2020-2021

Organization

- Co-organizer, RespMultimodal 2026: Responsible Multimodal Foundation Models for Knowledge Discovery workshop co-located with KDD'26, Jeju, Korea (2026)
- PhD Symposium chair, The ACM Web Conference, Singapore (2024)
- Virtual experience chair, The fifth annual ACM FAccT conference, Seoul, Korea (2022)
- Co-organizer, 4th International Workshop on NEws and publiC Opinion (NECO 2020) co-located with ICWSM'20, Atlanta, GA, USA (2020)
- Co-organizer, Workshop on Bias, Disinformation, Misinformation, and Propaganda in Online News and in Social Media co-located with SocInfo'19, Doha, Qatar (2019)
- Workshop chair, 29th ACM Conference on Hypertext and Social Media (HT), Baltimore, Maryland, USA (2018)
- Co-organizer, 3rd International Workshop on NEws and publiC Opinion (NECO 2018) co-located with ICWSM'18, Stanford, CA, USA (2018)
- Co-organizer, 4th workshop on Social News on the Web (SNOW 2017) Co-located with WWW'17, Perth, Australia (2017)
- Co-organizer, 2nd International Workshop on NEws and publiC Opinion (NECO 2017) co-located with ICWSM'17, Montreal, Canada (2017)
- Co-organizer, 3rd workshop on Social News on the Web (SNOW 2016) Co-located with WWW'16, Montreal, Canada (2016)
- Co-organizer, 1st International Workshop on NEws and publiC Opinion (NECO 2016) co-located with ICWSM'16, Cologne, Germany (2016)
- Track chair (with Alessandra Sala and Aristides Gionis), Social Networks and Graph Analysis Track in 2016 International World Wide Web Conference, Montreal, Canada (2016)
- Co-organizer, 1st International Workshop on Exploration Games and Gamers (EGG) co-located with SocInfo'14, Barcelona, Spain (2014)
- Registration chair, ACM Internet Measurement Conference (IMC), Barcelona, Spain (2013)

Program Committee Member

- ACM The Web (WWW) Industry Track Area Chair: 2026
- ACM The Web (WWW) Social Networks and Social Media Track Senior PC: 2025, 2026

- ACM The Web (WWW) Responsible Web Track Senior PC: 2025
- AAAI ICWSM Senior PC: 2014, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025
- ACM WebSci Senior PC: 2023, 2024, 2025, 2026
- ACM WSDM Senior PC: 2025
- ACM CIKM Senior PC: 2025, 2026
- COLM: 2025, 2026
- ACL ARR AC/Reviewer: 2023, 2024, 2025, 2026
- ACM The Web (WWW) Web and Society Track Senior PC: 2023
- ACM The Web (WWW) Web and Society Track PC: 2018, 2019, 2020, 2021, 2022
- ACM The Web (WWW) Fairness, Accountability, Transparency, and Ethics on the Web Track PC: 2023
- ACL Computational Social Science and Cultural Analytics Track: 2023
- AAAI/ACM Conference on AI, Ethics, and Society (AIES) PC: 2022
- EACL Ethical and Sustainable NLP track: 2023
- EMNLP: 2023
- AACL-IJCNLP: 2022
- International Conference on Computational Social Science (IC2S2): 2023, 2024
- Conference on Complex Systems: 2024
- SocInfo PC: 2014, 2015, 2016, 2020
- ACL NLP and Application Track: 2020
- The Web (WWW) Social Network and Graph Analysis Track: 2014, 2015, 2016, 2018, 2019, 2021
- Complex Networks: 2017, 2018, 2020, 2021, 2022, 2024
- IEEE ICDCS Social Networks and Crowdsourcing: 2018
- European Symposium Societal Challenges in Computational Social Science: 2017
- LC2S2 (At the crossroads: lessons and challenges in Computational Social Science): 2016
- ACM WebSci PC: 2014, 2022
- ACM Hypertext: 2014, 2023, 2024
- The Web (WWW) Poster: 2014
- ICWSM PC: 2013
- European Conference on Complex Systems (ECCS): 2013

Program Committee Member - Workshop

- NLP+CSS: 2024, 2026
- ICWSM PhoMemes Workshop: 2023
- ACL WOAHS Workshop: 2023
- ACL NLP COVID-19 Workshop: 2020
- CIKM MAISO Workshop: 2020

- WWW Cybersafety Workshop: 2018, 2019, 2020
- ACL Abusive Language Workshop (ALW): 2019
- ACL Workshop on NLP and Social Dynamics: 2014
- CIKM Rumours and Deception in Social Media Workshop: 2018
- WWW Connecting online and offline life (COOL) Workshop: 2014
- WWW Simplifying Complex Networks for Practitioners (SIMPLEX) Workshop: 2014
- SocInfo Interaction and Exchange in Social Media Workshop: 2014
- CIKM Search and Mining User-generated Contents Workshop: 2013

Best Paper Committee

- ACM WebSci'22

IN THE PRESS

Selected list of media coverages

- “A semantic embedding space based on large language models for modelling human beliefs” (Nature Human Behaviour’25)
 - [Phys.org Featured Article](#)
- “Public perception of generative AI on Twitter: an empirical study based on occupation and usage” (EPJ Data Science’24)
 - [Blockchain News](#)
- “Investigating toxicity changes of cross-community redditors from 2 billion posts and comments” (PeerJ CS’22)
 - [New Scientist](#)
 - [NewsWise](#)
 - [Independent](#)
 - [DailyMail](#)
 - [PeerJ Spotlight](#)
- “Understanding Toxicity Triggers on Reddit in the Context of Singapore” (ICWSM’22)
 - [AI Ethics Brief Newsletter #111 by Montreal AI Ethics Institute](#)
- “Tanbih: Get to Know What You Are Reading” (EMNLP Demo’19)
 - [Forbes](#)
 - [Boston Globe](#)
 - [Aljazeera](#)
 - [MIT Technology Review](#)
 - [Science Daily](#)
 - [Popular Science](#)
 - [Fast Company](#)
 - [The Register](#)
 - [WIRED](#)
 - [Engadget](#)
- “What is Gab? A Bastion of Free Speech or an Alt-Right Echo Chamber?” (WWW Cybersafety workshop’18)

- New Scientist
- Vice
- “Data-driven Approach to Measuring the Level of Press Freedom Using Media Attention Diversity from Unfiltered News” (ICWSM’17)
 - The Best of the Physics arXiv (week ending April 15, 2017) in MIT Technology Review
- “I Would Not Plant Apple Trees If the World Will Be Wiped: Analyzing Hundreds of Millions of Behavioral Records of Players During an MMORPG Beta Test” (WWW Companion’17)
 - Nifty news
 - New Scientist
 - Daily Mail (2025)
 - Daily Mail (2017)
 - IFL Science (2017, 18.1K shared)
 - IFL Science (2023)
 - PC Gamer (2017)
 - PC Gamer (2023)
 - Massively Overpowered
 - 4Gamers
 - El Confidencial
 - Joongang Ilbo
 - How Stuff Works
 - PC Gamer
 - Digital Trends
 - Engadget
- “Revealing the Hidden Patterns of News Photos: Analysis of Millions of News Photos Using GDELT and Deep Learning-based Vision APIs” (ICWSM Workshop on News and Public Opinion’16)
 - The Best of the Physics arXiv (week ending March 26, 2016) in MIT Technology Review
- “Breaking the News: First Impressions Matter on Online News” (ICWSM’15)
 - Other Interesting arXiv Papers (Week ending April 11, 2015) in MIT Technology Review
 - O Globo
 - Engadget
 - Arkansas Democrat-Gazette
- “Understanding News Geography and Major Determinants of Global News Coverage of Disasters” (Computation+Journalism Symposium’14)
 - MIT Technology Review
 - ACM Tech News
- “Searching for a Unique Style in Soccer” (KDD Workshop on Large-scale Sports Analytics’14)
 - BBC
 - MIT Technology Review
 - The Sunday Times

- The Economist
- Slate
- Pacific Standard
- “Linguistic Analysis of Toxic Behavior in an Online Video Game” (SocInfo Workshop on Exploration Games and Gamers’14)
 - Other Interesting arXiv Papers (Week ending November 1, 2014) in MIT Technology Review
- “STFU NOOB! Predicting Crowdsourced Decisions on Toxic Behavior in Online Games” (WWW’14)
 - Nature News
 - Scientific American
 - Hybrid
 - Kincir
 - Chosun Ilbo
- “Fragile Online Relationship: a First Look at Unfollow Dynamics in Twitter” (CHI’11)
 - Kyunghyang Shinmun
- “What is Twitter, a Social Network or a News Media?” (WWW’10)
 - Mashable
 - Readwrite
 - Giagom
 - The Guardian
 - PC World
 - MIT Technology Review
 - Chosun Ilbo
 - Dong-A Ilbo
- “I Tube, You Tube, Everybody Tubes: Analyzing the World’s Largest User Generated Content Video System” (IMC’07)
 - Work Today
 - Financial News

REFERENCES

Prof. Sue Moon

Professor, School of Computing, KAIST, Korea
sbmoon@kaist.edu

Prof. Jaideep Srivastava

Professor, Department of Computer Science and Engineering, University of Minnesota
Twin Cities
srivasta@cs.umn.edu

Prof. Meeyoung Cha

Associate Professor, School of Computing, KAIST, Korea
Chief Investigator, Institute for Basic Science (IBS), Korea
meeyoungcha@kaist.ac.kr