**Nikolaus Nova Parulian**
School of Information Sciences
University of Illinois at Urbana-Champaign
501 E. Daniel Street, Room 112
Champaign, IL 61820-6211
nnp2@illinois.edu

Expected Graduation Date: June 2023

| | |
|---|---|
| ***Professional Interests*** | Research in the areas of data quality and provenance on data science/machine learning application, machine learning, text mining. |
| ***Education*** | Ph.D. Candidate, Information Sciences Fall 2018 - Now, University of Illinois at Urbana Champaign (Expected Graduation: June 2023)<br>M.S. Information Management Fall 2016 - Fall 2017, University of Illinois at Urbana Champaign<br>B.E. Informatics Engineering 2006, Telkom University, Indonesia |
| ***Employment History*** | |
| ***2016 – Now*** | **Graduate Research Assistant at School of Information Sciences UIUC** |
| | • Center of Informatics Research in Science and Scholarship (CIRSS)<br>Supervised by: Dr. Bertram Ludaescher<br>• Hathi Trust Research Center (HTRC): https://wiki.htrc.illinois.edu<br>Supervised by: Dr. J. Stephen Downie<br>• Social Computing Lab: http://jdiesnerlab.ischool.illinois.edu<br>Supervised by: Dr. Jana Diesner<br>• Blender Lab (Natural Language Processing): http://blender.cs.illinois.edu<br>Supervised by: Dr. Heng Ji |
| ***2019 – 2020*** | **John Deere, Champaign, Illinois**<br>Robotics Student Intern at John Deere Technology and Innovation Center |
| | Main Responsibilities:<br>• Data warehouse administration and development for data analytics workflow.<br>• Research and Development for sensor fusion systems for Autonomous Vehicles.<br>• Working with computer vision (image processing), for object/obstacle detection systems.<br>• Research and Design for the Internet of Things (IoT) technology, sensors, messaging protocol, and AWS cloud infrastructure |
| ***Publications*** | Parulian, N. N., Dubnicek, R., Worthey, G., Downie, J.S. (2022). Uncovering Black Fantastic: Piloting A Word Feature Analysis and Machine Learning Approach for Genre Classification. In *Proceedings of the Association for Information Science and Technology*.<br><br>Parulian, N. N., & Ludäscher, B. (2022). Conceptual Model and Framework for Collaborative Data Cleaning. In *International Digital Curation Conference 2022*.<br><br>Parulian, N. N., & Ludäscher, B. (2022, June). DCM explorer: a tool to support transparent data cleaning through provenance exploration. In *Proceedings of the 14th International Workshop on the Theory and Practice of Provenance* (pp. 1-6).<br><br>Wang, Q., Li, M., Wang, X., Parulian, N., Han, G., Ma, J., ... & Onyshkevych, B. (2021, June). COVID-19 Literature Knowledge Graph Construction and Drug Repurposing Report Generation. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Demonstrations* (pp. 66-77). |

| | |
|---|---|
| | Huang, L., Cao, S., Parulian, N., Ji, H., & Wang, L. (2021, June). Efficient Attentions for Long Document Summarization. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 1419-1436).<br><br>Parulian, N. N., Li, L., & Ludäscher, B. (2021). or2yw: Modeling and Visualizing OpenRefine Histories as YesWorkflow Diagrams.<br><br>Parulian, N. N., & Worthey, G. (2021). Identifying Creative Content at the Page Level in the HathiTrust Digital Library Using Machine Learning Methods on Text and Image Features. In *Diversity, Divergence, Dialogue: 16th International Conference, iConference 2021, Beijing, China, March 17–31, 2021, Proceedings, Part I 16* (pp. 478-489). Springer International Publishing.<br><br>Parulian, N. N., McPhillips, T. M., & Ludäscher, B. (2020). A Model and System for Querying Provenance from Data Cleaning Workflows. In *Provenance and Annotation of Data and Processes* (pp. 183-197). Springer, Cham.<br><br>Dinh, L., & Parulian, N. (2020). COVID-19 pandemic and information diffusion analysis on Twitter. *Proceedings of the Association for Information Science and Technology*, *57*(1), e252.<br><br>Parulian, N. N., Lu, T., Mishra, S., Avram, M., & Diesner, J. (2020). Effectiveness of the Execution and Prevention of Metric-Based Adversarial Attacks on Social Network Data. *Information*, 11(6), 306.<br><br>Parulian, N., Dubnicek, R., Eden, K., Hu, Y., & Downie, S. (2020), Evaluating a machine learning approach to identify expressive content at page level in HathiTrust. *Advance Issue of Digital Scholarship in the Humanities*.<br><br>Avram, M. V., Mishra, S., Parulian, N. N., & Diesner, J. (2019, August). Adversarial perturbations to manipulate the perception of power and influence in networks. In *2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 986-994). IEEE.<br><br>Dinh, L., Sarol, J., Cheng, Y. Y., Hsiao, T. K., Parulian, N., & Schneider, J. (2019). Systematic examination of pre-and post-retraction citations. *Proceedings of the Association for Information Science and Technology*, *56*(1), 390-394.<br><br>Cheng, Y. Y., Parulian, N., Hsiao, T. K., Dinh, L., Sarol, J., & Schneider, J. (2019). ReTracker: actively and automatically matching retraction metadata in Zotero. *Proceedings of the Association for Information Science and Technology*, *56*(1), 372-376.<br><br>McPhillips, T., Li, L., Parulian, N., & Ludäscher, B. (2019). Modeling provenance and understanding reproducibility for openrefine data cleaning workflows. In *11th International Workshop on Theory and Practice of Provenance (TaPP 2019)*.<br><br>Dinh, L., Cheng, Y.-Y., & Parulian, N. (2019). ReTracker: an Open-Source Plugin for Automated and Standardized Tracking of Retracted Scholarly Publications. *2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, 406–407.<br><br>Arlitsch, K., Wheeler, J., Pham, M. T. N., & Parulian, N. N. (2020). An analysis of use and performance data aggregated from 35 institutional repositories. *Online Information Review.* |
| ***Presentation and Poster*** | • Digital Humanities (DH) 2022 (July 2022): Uncovering the Black Fantastic: Piloting Text Similarity Methods for Finding "Lost" Genre Fiction in HathiTrust<br>• Digital Humanities (DH) 2020 (July 2020): Evaluating a Machine Learning Approach to Identifying Expressive Content at Page Level in HathiTrust |

| | |
|---|---|
| | • Coalition for Networked Information (CNI) Fall 2019 Meeting (December 2019): Visualizing Use and Performance Data from a Global Cross-platform Set of Institutional Repositories |
| ***Teaching Assistant*** | • Fall 2018: Data Cleaning, Theory and Practice for Information Sciences (IS537)<br>• Summer 2019, 2021, 2022: Data Cleaning, Theory and Practice course for Computer Science (CS513) Coursera online course. |
| ***Synergetic Activities*** | • 2020, COVID-19 Literature Surveillance Team, Technology Coordinator, https://www.covid19lst.org. Providing crowdsourcing mechanism for literature survey, critics, and evidence mining for COVID-19 literature-related papers/publications/articles.<br>• 2020, Association of Computing Machinery for Woman (ACM-W), Technology Administrator, https://acmw.illinois.edu.<br>• 2019, Awardee of LIS Education and Data Science for the National Digital Platform (LEADS-4-NDP) Program by Drexel University, Philadelphia, USA.<br>• Fall 2018, Summer 2019, Summer 2021, Teaching Assistant for Data Cleaning, Theory and Practice course (IS537 and CS513). |