Steph Buongiorno, PhD

Curriculum Vitae

Department of Video Game Design (Guildhall)   sbuongiorno@smu.edu
Southern Methodist University
Ford Hall 306E
3100 McFarlin Blvd, Dallas, TX 75205

Education

2023	PhD, Applied Science, Lyle School of Engineering, Southern Methodist University

- Subject areas: Computer Science, English, and History.
  - Dissertation: "Grammatical Triples Extraction for the Distant Reading of Textual Corpora." Committee Members: Jo Guldi (Chair), Corey Clark (Co-Chair), David Lin, Mark Fontenot, Miju Ahn, Tim Cassedy.

2017	MA, English, West Virginia University
2014	BA, English & Linguistics, University of Texas at Arlington (Summa Cum Laude)

Professional Employment

2023-Present	Postdoctoral Fellow and Lab Co-Director, Guildhall Video Game Department, Southern Methodist University

- I co-direct a 10 person lab while playing a hands-on role in designing intelligent agents and hierarchical, multi-agent systems that autonomously learn new subjects (by parsing plain-language text into knowledge graphs), retrieve information that encodes domain knowledge, and self-validate (using techniques based on self-reflection/self-prompting).

Publications

(peer reviewed = *)

Book Projects

* Text Mining for Historical Analysis. Steph Buongiorno [**Corresponding Author**] and Jo Guldi. Under contract in June at Cambridge University Press. Expected publication EOY 2025.

* Text Mining for Historical Analysis. eBook Edition. Steph Buongiorno [**Corresponding Author**] and Jo Guldi. Under contract in June at Cambridge University Press. Expected publication EOY 2025.

Articles (Under Review)[1]

* "[PANGeA: Procedural Artificial Narrative using Generative AI for Turn-Based Video Games](#)." Under review at *AAAI Artificial Intelligence and Interactive Digital Entertainment*, a peer-reviewed CS conference. Steph Buongiorno [**Corresponding Author**], Jake Klinkert, Tanishq Chawla, Zixin Zhaung, and Corey Clark. *In this article, I was responsible for all writing, designing the narrative and validation gameplay systems, drawing all diagrams, and designing the experimental tests.*

* "[A Framework for Leveraging Human Computation Gaming to Enhance Knowledge Graphs for Accuracy Critical Generative AI Applications](#)." Under review at *IEEE Conference on Games*, a peer-reviewed CS conference. Steph Buongiorno [**Corresponding Author**] and Corey Clark. *In this article, I was responsible for all writing, conceptualization, and visualizations.*

* "Disambiguating Speakers in the Hansard 19th-Century British Parliamentary Debates." Under review at *Journal of Cultural Analytics*. Steph Buongiorno [**Corresponding Author**], Alexander Cerpa, Jo Guldi. *In this article, I was responsible for all writing, designing the disambiguation pipeline, and analyzing the results.*

* "The Hansard 19th-Century British Parliamentary Debates: Discovering Lost Records and the Creation of an Analysis Ready Data Set." Under review at *Journal of Cultural Analytics*. Steph Buongiorno [**Corresponding Author**], Rob Kalescky, Jo Guldi

Article Projects

Steph Buongiorno [**Corresponding Author**], Ananya Das Manolyl, Corey Clark. "Hierarchies of Thought: A Development Methodology for Explainable Multi-Agent Planning Systems Driven by Generative AI and

---

[1] In my profession it is expected to include co-authors who do not contribute writing. As corresponding author I performed all writing. I directed the student co-authors following my name in conceptualization, design, and development of their respective artifacts (e.g. a video game mechanic, artwork, story narrative) while enabling their own creative agency in our shared project. In this respect, co-authorship is used here to convey transparency and signal their contributions to dimensions of a larger project. When acting as 2nd+ author, I offered an intellectual contribution and produced writing.

Specialized Knowledge Graphs." Target Journal: "ACM Transactions on Intelligent Systems and Technology." Estimated Submission: End of July/August.

Steph Buongiorno [**Corresponding Author**], Aiyou Tan, Ryan Schaefer, Jo Guldi. "Democratizing Text-Based Data Analytics and Data Sharing Across the Humanities and Social Sciences". Target Journal: "International Journal of Digital Humanities." Estimated Submission: End of August.

Jake Klinkert, Steph Buongiorno [**Second Author**], Corey Clark. "Driving Generative Agents with their Personality."Target Journal: "IEEE Transactions on Games." Estimated Submission: End of August.

Steph Buongiorno. "North and South American Cave Diving Fatalities and Comorbid Factors (1970-2021)." Collected from the archives of the National Speleological Society accident reports and public news reports. Target Journal: FIND THE NAME OF IT.

Digital Projects

Steph Buongiorno, Ryan Schaefer, Aiyou Tan, Wes Anderson, Chris Miller, and Matt Swigart. **Democracy Viewer,** Emory University (forthcoming Summer 2024) *I lead the development of a [public-facing web app for exploring, text mining, and visualizing humanities and social sciences data sets in English, German, Spanish, and French. To be deployed on Amazon Web Service (AWS) in Summer 2024.*

Steph Buongiorno. The Hansard Viewer. Southern Methodist University, 2022. A Shiny app for text mining and visualizing the 19th-century British parliamentary debate using data science metrics.

Steph Buongiorno. The Congress Viewer. Southern Methodist University, 2022. Description: A Shiny app for text mining and visualizing the U.S. Congressional Records using data science metrics.

* Steph Buongiorno. usdoj. ROpenGov. Description: An R package for creating a structured version of the U.S. Department of Justice press releases, blogs, and records.

* Steph Buongiorno. oldbailey. ROpenGov. Description: An R package for creating a structured version of the Old Bailey criminal trials. Handles broken tags and messy data and returns an analysis-ready dataset.

* Steph Buongiorno. hansardr. GitHub. Description: An R package for querying a clean version of the 19th-century Hansard Corpus.

* Steph Buongiorno and Omar Alexander Cerpa. hansard-speakers. GitHub. Description: Code for disambiguating speakers in the 19th-century Hansard Corpus using Levenshtein distances and parallel computing.

 Steph Buongiorno. noaa. GitHub. Description: An R package for querying a clean version of NOAA climate and weather data.

 Steph Buongiorno and Omar Alexander Cerpa. posextract. GitHub. Description: A Python package for extracting grammatical subject-predicate triples from data. Tailored for the analysis of agency in text.

Steph Buongiorno. posextractr. GitHub. Description: An R package for extracting grammatical subject-predicate triples from data. Tailored for the analysis of agency in text.

Steph Buongiorno and Jo Guldi. democracy-lab. GitHub. Description: A code repository for text mining techniques for the Digital Humanities.

Steph Buongiorno and Ryan Schaefer. dhmeasures. GitHub. Description: Optimized, "white-box" statistical functions for textual analysis.

Instructional and Interactive Materials

Steph Buongiorno. "Foundations and Applications of Humanities Analytics 2023." Santa Fe Institute. GitHub. *I was the lead instructor for the SFI's "Humanities Analytics" summer camp two years running. My code – shared here as a Github page — formed the basis for all instruction and activity.*

Steph Buongiorno. "Foundations and Applications of Humanities Analytics 2022." Santa Fe Institute. GitHub.

Steph Buongiorno. "Digital History." Southern Methodist University. GitHub.

Video Games

Dark Shadows

Steph Buongiorno, Jake Klinkert, Tanishq Chawla, Zixin Zhaung, and Corey Clark. *Dark Shadows*, Southern Methodist University, Guildhall. Topic: Address the real-world problem of human trafficking while playing a video game.

Dark Shadows, Scene I

I conceived of Dark Shadows as a film noir-style "document thriller" (inspired by *Papers, Please!* and *Night Call*). I designed and developed underlying "human-in-the-loop" mechanics to collect player feedback and train a machine learning model, which can disambiguate speakers, locations, and events in real human trafficking data. I guided graduate students on the design of the gameplay mechanics and narrative. https://www.cbsnews.com/texas/news/i-team-video-game-helps-researchers-fight-sex-trafficking/

Dark Shadows, Scene II

I designed NPCs that leverage large language models (LLMs) and a self-reflection-based validation system to enable dynamic, free-form interactions with the player aligned with a procedural game narrative (re: PANGeA). NPCs express traits from the Big 5 Personality Model in their responses, while the validation system ensures responses stay on narrative track. https://www.youtube.com/watch?v=H-3-9zxhUlg

Dark Shadows, Scene III

[Forthcoming] Dark Shadows will include a social engineering scene, where the player must investigate the NPCs' personalities and use rhetorical devices to gain clues on the mystery.

Data Sets

Steph Buongiorno; Robert Kalescky; Omar Alexander Cerpa; Jo Guldi, 2022, "The Hansard 19th-Century British Parliamentary Debates with Improved Speaker Names: Parsed Debates, N-Gram Counts, Special Vocabulary, Collocates, and Topics", https://doi.org/10.7910/DVN/ZCYJH8, Harvard Dataverse

Steph Buongiorno; Omar Alexander Cerpa; Jo Guldi, 2022, "The Hansard 19th-Century British Parliamentary Debates with Improved Speaker Names: Speaker Metadata", https://doi.org/10.7910/DVN/Z3LTVV, Harvard Dataverse, FORTHCOMING

Other Contributions

30 visualizations in Jo Guldi's *The Dangerous Art of Text Mining*.

Grant Proposals

"Integrating Human Computer Interaction, Machine Learning, Game Design, and Educational Assessment in a STEM+C Curriculum." Topic: To support the development of personalized, educational agents in Minecraft. We propose the integration of an advanced set of agent abilities across three design types: A) Personalized Educational Agents that autonomously monitor student progression and dynamically generate customized curriculum for addressing individuals' needs based on their existing knowledge, personalities, and interests; B) Bridging Agents that demonstrate the "bigger picture" behind computational thinking by creating connections between STEM topics and interdisciplinary topics, such as the Language Arts; and C) Analyst Agents that translate student progression and learning outcomes to teachers for their easy assessment of student progress.

"Gen-AI Distributed Autonomous Agents and Deployment Infrastructure to Accelerate Research, Discovery and Development." Topic: To support finishing the development and design of intelligent agents and hierarchical, multi-agent systems for deployment in a distributed environment. To be applied to research problems with collaborators in psychology, chemistry, education, and economics.

"SmartCADD Machine Learning Drug Discovery." Topic: To support the development of AI for discovering new medical drug compounds.

Invited Lectures and Presentations

2024   [upcoming] "Navigating Privacy, Storage, and Analysis in a Generative AI Enabled World," SMU Human Trafficking Data Conference, Sponsored by Thomson Reuters.

2024   [upcoming] "Generative AI and Agents Integration into Human Trafficking Data and Analysis," SMU Human Trafficking Data Conference, Sponsored by Thomson Reuters.

2024   "Rhetorical Tools for Storytelling," Guildhall, Southern Methodist University

2022   "Text Mining Parliamentary Data," Department of History, Academy of Finland

2022   "Data Storage and Data Processing," Public Health, UT Southwestern

2022   "Text Mining and Data Analytic Techniques," Public Health, UT Southwestern

2018   "Analyzing Gender and Agency using Grammatical Triples Extraction." Think-Play-Hack, Southern Methodist University

Teaching

2024   HGME 6377: Thesis III Software Development. Graduate Course. "gaia: A GEN-AI Gaming Framework for creating Dynamic Gameplay Content via Large Language Models."Guildhall, Southern Methodist University

2024   HGME 6377: Thesis III Level Design. Graduate Course. "Unveiling New Realms: Enhancing Procedural Narrative Generation and NPC Personalization using AI." Guildhall, Southern Methodist University

2024   HGME 6377: Thesis III Production. Graduate Course. "Chinese PC Gamers' Motivation, Game Preferences, and Publishing Advice for Non-Chinese Game Companies." Guildhall, Southern Methodist University

2024   CS 5352: Senior Design II. Undergraduate Course. Computer Science, Southern Methodist University. Topic: Taught a year-long digital humanities project with 3 enrolled undergraduates

2023   Instructor, Guildhall Academy Video Game Design (for high school students), Guildhall, Southern Methodist University

2023   CS 5351: Senior Design I. Undergraduate Course. Computer Science, Southern Methodist University. Topic: Topic: Taught a year-long digital humanities project with 3 enrolled undergraduates

2023   HGME 6276: Thesis II Software Development. Graduate Course. "gaia: A GEN-AI Gaming Framework for creating Dynamic Gameplay Content via Large Language Models." Guildhall, Southern Methodist University

2023   HGME 6276: Thesis II Level Design. Graduate Course. "Unveiling New Realms: Enhancing Procedural Narrative Generation and NPC Personalization using AI." Guildhall, Southern Methodist University

2023   HGME 6592: Team Game Production I. Graduate Course. Topic: Replicability and Version Control. Guildhall, Southern Methodist University

2023   Teaching Fellow, Foundations and Applications of Humanities Analytics, Santa Fe Institute.

2023   HGME 6268: Directed Studies in Creative Narrative Writing for Video Games. Graduate Course. Guildhall, Southern Methodist University

2022   HGME 6266: Directed Studies in Creative Narrative Writing for Video Games. Graduate Course. Guildhall, Southern Methodist University

2022	HIST 3380: Topics in Digital History, Department of History, Southern Methodist University.
2022	Instructor, Guildhall Academy, Southern Methodist University
2022	Teaching Fellow, Foundations and Applications of Humanities Analytics, Santa Fe Institute
2021	HIST 3380: Topics in Digital History, Department of History, Southern Methodist University.
2018	WRTR 1312: Introduction to Academic Writing, Department of English, Southern Methodist University (1 section)
2017	WRTR 1311: Foundations of Writing, Department of English, Southern Methodist University (1 section)
2017	PRW 2144: SCUBA, Department of Physical Fitness, Southern Methodist University (1 section)
2016-17	ENGL 102: Composition and Rhetoric. Department of English, West Virginia University (4 sections)
2015-16	ENGL 101: Introduction to Composition and Rhetoric. Department of English (4 sections)
2014	EXSA 1249: Advanced Scuba Diving. Department of Exercise and Sport Activity, The University of Texas at Arlington (2 sections)
2014	Scuba Diving. Texas Woman's University. Department of Fitness and Recreation. (1 section) (need to find course number)
2012-14	EXSA 1249: Scuba Diving. Department of Exercise and Sport Activity, The University of Texas at Arlington (6 sections)

Service

2024	Track Co-Chair, SciPy, Social Science, Humanities, Economics, and Human Systems
2024	Peer Reviewer, IEEE Conference on Games, Long Papers
2024	Peer Reviewer, IEEE Conference on Games, Short Papers
2023	Track Co-Chair, SciPy, Computational Social Science and Digital Humanities
2023	Mentor, PhD student Joel Nyman in history, Academy of Finland

Previous Employment

Other Research Positions

2022-23	Technical Lead, "The Human Trafficking Project," Department of Computer Science, National Institute of Justice Grant (H.R. 2471), PI Corey Clark, Southern Methodist University
2021-22	Research Assistant, "Toward a History of the Associative-Developmental State," Department of History, Southern Methodist University
2021-22	Research Assistant, "The History of Modern China," Department of History, Southern Methodist University
2018-21	Technical Lead, "Global Urbanization and Housing Affordability: Poverty, Property, and the City," National Science Foundation Grant (no. 1520103), PI Jo Guldi, Department of History, Southern Methodist University

Employment: Teaching Positions

2023-24	Democracy Lab Co-Director, Emory University. Topic: Development of an advanced text mining app, data sharing, and data visualization public web app for the digital humanities.
2023-24	Level Design Thesis Advisor, Guildhall, Southern Methodist University. Topic: creative narrative writing for a film-noir style detective thriller augmented by generative AI.
2023	Support Instructor, Team Game Production I, Guildhall, Southern Methodist University. Topics: Reproducibility and version control.
2022-23	Teaching Fellow, Foundations and Applications of Humanities Analytics, National Endowment of the Humanities Grant (no. HT-272418-20), Santa Fe Institute. Topics: Digital humanities and cultural analytics.
2021-22	Support Instructor, Topics in Digital History, Southern Methodist University. Topics: Code, bias, and the representation and obfuscation of gender in both coding libraries and textual corpora (e.g. U.S. Congress and Reddit).
2017-18	Graduate Teaching Assistant, Academic Writing, Southern Methodist University
2015-2017	Graduate Teaching Assistant, Composition and Rhetoric, West Virginia University

Employment: Related Positions

2023-present	Instructor, Guildhall Academy (high school students), Guildhall, Southern Methodist University
2022-23	Computer Science Senior Design, Southern Methodist University

Employment: Other

2017	Assistant Instructor, Introduction to SCUBA, Southern Methodist University
2012-2014	Assistant Instructor, Advanced SCUBA, University of Texas at Arlington
2014	Assistant Instructor, Introduction to SCUBA, Texas Woman's University
2012-2014	Assistant Instructor, Introduction to SCUBA, University of Texas at Arlington

Awards

2023	Dean's Award for Best Poster Presentation in Computer Science, Southern Methodist University

Other Skills and Certificates

2020	Adv. Side Mount Cave Diver, International Association of Nitrox and Technical Divers
2018	Full Cave Diver, International Association of Nitrox and Technical Divers