

# TRAVIS AARON HOPPE

---

AI Leader, Curriculum Vitae  
travis.hoppe@gmail.com

PhD Physics  
(775) 287-4033

## HIGHLIGHTS

- Federal expert in artificial intelligence (AI), machine learning (ML), and data science. Leader and innovator at the White House, coordinating policy agendas around research and development with partners in the National Science Foundation and the President's Council of Advisors on Science and Technology.
- Led multiple teams across CDC to develop and implement cutting-edge policies, guidance, and strategies for AI/ML adoption, including conversational AI (Chat GPT), and infrastructure needs to support AI/ML implementation. Extensive experience in presenting to federal and industry venues on topics around Trustworthy AI, implementation, and development.
- Respected and well cited author in AI/ML and biomedical research to address questions on multiple topics including: disparities in NIH funding (**Topic Choice**: 600 citations, key evidence used by the **Congressional Black Caucus**), open citations (**NIH Open Citation Collection**: 100 citations), and a fundamental AI dataset for training large language models (**The Pile**: 850 citations).

## PROFESSIONAL EXPERIENCE

### Assistant Director of AI Research and Development

White House Office of Science and Technology Policy (OSTP)

February 2024 - January 2025 (detailed from CDC)

- **Set national vision for AI R&D** for major White House initiatives including the National AI Research Resource (NAIRR), the National Secure Data Service (NSDS), the National Science and Technology Council (NSTC) subcommittee AI and ML, and the OMB Memorandum (**M-24-10**), "Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence"
- **Built coalitions** across the Federal government by connecting State, Tribal, Local, and Territorial governments through many engagements via the White House **Tech Policy Network**, including the National League of Cities, the American Public Human Services Association, and the Council of State and Territorial Epidemiologists
- **Managed, prioritized, and directed resources** for projects within agencies including the Federal AI use-case inventory, the Networking and Information Technology Research and Development (NITRD) AI program repository, and the Chief AI Officer's Council working groups.

**Acting Chief AI Officer (2025),**  
**Associate Director for Data Analytics and Data Science (2023-2025)**  
Centers for Disease Control and Prevention (CDC)  
December 2023 - Present

- **Led organizational change** across CDC by bringing generative AI to all staff, the first Federal agency to do so. Accomplished results by leading 15 pilots programs across the Centers, coordinating financial, human, and IT resources, cleaning the program through cybersecurity, legal, and labor concerns.
- **Modernized** NCHS data systems through cloud adoption through a multi-year plan. Aligned and coordinated the plan with CDC's data modernization initiative (DMI) to help build the foundation for data sharing across all levels of public health. Coordinated policy, governance, and legal issues around the usage of Confidential Information Protection and Statistical Efficiency Act (CIPSEA) protected data. Worked with Office of Management and Budget (OMB) to adopt CIPSEA guidance and lead NCHS as an early cloud adopter for statistical agencies.
- **Piloted** innovative AI, ML, and data science projects to help identify, analyze, and report on emerging public health data. Projects include an item non-response detection model for survey text which resulted in the **first AI model** released by NCHS, privacy enhancing technologies (PETs) like homomorphic encryption and privacy preserving record linkage, text-to-speech transcription (which resulted in a 10-fold improvement from prior methods), and the creation of a bibliometrics dashboard for reporting.
- **Built community** by starting and leading three different inclusive organizations within the agency: NCHS Innovation, NCHS Data Science Community of Practice, and the agency-wide EDAV Best Practices group. Organizations helped build community, foster communication, and spur innovative practices across CDC.

**Senior Service Fellow: Chief Data Scientist**  
Centers for Disease Control and Prevention (CDC)  
2020 - 2023

- **Lead** for NCHS Data Modernization Initiative: *Increase Use, Discoverability, and Access to NCHS Data*. Established pilot projects and worked with stakeholders to create statements of work and business needs.
- **Advised** on CDC's response to the Executive Order on AI and the NCHS Data Science Strategic Plan. Presented to Board of Scientific Council, NCHS All-Hands, and delivered subject matter talks on Bias in AI, advances in Natural Language Processing, bibliometrics, and more.
- **Developed and coordinated** Center-wide metadata standards and built a custom ontology using evidence based sources: publications, web searches, and market research.

- **Researched** new methodology to study free text responses from the Research and Development Survey (RANDS), including non-response detection and zero-shot learning objectives.
- **Implemented** PII detection processes for restricted microlevel data to allow privacy preserving research and development.

### **Senior Data Scientist / Portfolio Analyst**

National Institutes of Health (NIH)

2016 - 2020

- Scientific team leader for a novel inter-agency government blockchain to detect grant duplication with minimal shared data. Coordinated research, oversaw design, and developed protocols within the NIH and National Science Foundation (NSF) teams.
- Developed new analytic tools to process the text of NIH grants and publications using distributional embeddings (word2vec) and transformers (BERT). Tools were deployed for analysis presented to NIH senior leadership, Congress, and publications in high-ranking journals.
- Architected and productionized machine learning models for classification, regression, outlier detection, and language modeling. Creator and maintainer of several open-source tools used internationally in the scientific community.
- Analyzed grant and publication portfolios, evaluating metrics such as clinical impact, technological impact, and award rates to build quantitative comparisons between various populations.
- Restored historical texts from books and generated new structured data from free text. Expanded NIH grant coverage by thirty years from archival documents. Cross-linked publications to an NIH application's biographical sketch and literature cited. Data used internally with the NIH for analysis on racial disparity, topic analysis, mentorship, and grant efficacy.

### **Postdoctoral Fellowship (IRTA) at National Institutes of Health**

Research Scientist

2014 - 2016

- Researched novel integration schemes for molecular dynamics simulations (MDS). Developed protein models for tertiary structure prediction from primary sequence.
- Designed and managed high-performance computing models on the NIH supercomputer, Biowulf. First to investigate containerized solutions for MDS using a graphics processing card.
- Worked in collaboration with experimentalists to test and validate models.
- Mentored post-baccalaureate researchers, guiding them in their research and professional development.

**Postdoctoral Fellowship at National Institutes of Health**  
Research Scientist  
2011 - 2014

- Developed multi-scale theoretical and computational models to study protein folding, structure, and protein-protein. Derived hard-sphere models to account for crowding in biomolecular simulations and potentials to model anisotropic charge distributions.
- Managed large-scale parallel projects (1000+ cores) to simulate the cellular environment.

**EDUCATION**

- 2011 **Doctor of Philosophy, Physics**  
Drexel University  
*On the Role of Entropy in the Protein Folding Process, Thesis*
- 2008 **Master of Science, Physics**  
Drexel University
- 2005 **Bachelor of Science, Physics**  
University of Nevada
- 2005 **Bachelor of Science, Mathematics**  
University of Nevada

**COMMITTEES  
SERVED**

- 2021-Present **National Science and Technology Council (NSTC) subcommittee on AI/ML, Co-chair (2024-2025), CDC representative (2021-present)**
- 2021-Present **Federal Committee on Statistical Methodology (FCSM), Board Member**
- 2022-2024 **Building Trust and FAIRness into the Process for Finding and Using Government Data (Chief Data Officers Council and FCSM), Co-lead**
- 2023 **Utility and Risks to CDC of Conversational Artificial Intelligence (AI) Technologies like Chat GPT (CDC), Tiger Team Lead**
- 2023 **Supporting the U.S. Public Health Workforce, President's Council of Advisors on Science and Technology (PCAST), External Expert**
- 2022 **Analytics and Machine Learning Implementation within CDC's Cloud Environment (CDC), Tiger Team Lead**
- 2021 **National Science and Technology Council (NSTC): Epidemic Modeling and Forecasting Fast Track Action Committee (FTAC), Contributor: Plan to Advance Data Innovation**
- 2021 **Health and Human Services (HHS): Open Data Task Force, Committee member**

**PUBLICATIONS**

- Policy*
- 2025 **AI-Ready Federal Statistical Data: An Extension of Communicating Data Quality, Travis Hoppe, Jeffrey Gonzalez, Lisa Mirel, Rolf Schmitt, Federal Committee on Statistical Methodology**

- 2023 **A Framework for Data Quality: Case Studies**, Lisa Mirel, Darius Singpurwalla, Travis Hoppe, Rolf Schmitt, Julie Weber, Erika Liliedahl, [Federal Committee on Statistical Methodology](#)
- 2023 **Dark citations to Federal resources and their contribution public health**, Jessica Keralis, Juan Albertorio-Díaz, & Travis Hoppe, [Frontiers in Research Metrics and Analytics](#)
- 2023 **Application of a Novel Machine Learning Technique in a Bibliometric Analysis of Health Disparities Articles**, Pascaline Ezouah, Bao-Ping Zhu, & Travis Hoppe, Manuscript in preparation
- 2019 **Topic Choice Contributes to Lower Rate of NIH Awards to African-American/Black Scientists**, Travis Hoppe, Aviva Litovitz, Kristine Willis, Rebecca Meseroll, Matthew Perkins, B. Ian Hutchins, Alison Davis, Michael Lauer, Hannah Valantine, James Anderson, & George Santangelo, [Science Advances](#)
- Data Science*
- 2024 **Semi-Automated Nonresponse Detection for Open-text Survey Data**, Kristen Cibelli Hibben, Zachary Smith, Ben Rogers, Valerie Ryan, Paul Scanlon, Travis Hoppe, [Social Science Computer Review](#)
- 2023 **Model Release: Semi-Automated Nonresponse Detection for Surveys (SANDS)**, Kristen Cibelli Hibben, Zachary Smith, Ben Rogers, Valerie Ryan, Travis Hoppe, [Hugging Face](#)
- 2023 **Predicting causal citations without full text**, Travis Hoppe, Salsabil Arabi, Ian Hutchins, [Proceedings of the National Academy of Sciences of the United States of America](#)
- 2020 **The Pile: An 800GB Dataset of Diverse Text for Language Modeling**, Leo Gao, Stella Biderman, Travis Hoppe, et al., [arXiv](#)
- 2019 **The NIH Open Citation Collection: A public access, broad coverage resource**, Ian Hutchins, Kirk Baker, Matthew Davis, Mario Diwersy, Ehsanul Haque, Robert Harriman, Travis Hoppe, Stephen Leicht, Payam Meyer, George Santangelo, [PLoS Biology](#)
- 2017 **Additional support for RCR: A validated article-level measure of scientific influence**, Ian Hutchins, Travis Hoppe, Rebecca Meseroll, James Anderson, & George Santangelo, [PLoS Biology](#)
- Protein Topology and Interactions*
- 2019 **Non-specific Interactions Between Macromolecular Solutes in Concentrated Solution: Physico-Chemical Manifestations and Biochemical Consequences**, Travis Hoppe & Allen Minton, [Frontiers in Molecular Biosciences](#)
- 2016 **Incorporation of Hard and Soft Protein-Protein Interactions into Models for Crowding Effects in Binary and Ternary Protein Mixtures**, Travis Hoppe & Allen Minton, [Journal of the Physical Chemistry B](#)

- 2015 **Dependence of Internal Friction on Folding Mechanism**, Wenwei Zheng, David De Sancho, Travis Hoppe & Robert B. Best, *Journal of the American Chemical Society*
- 2015 **An equilibrium model for the combined effect of macromolecular crowding and surface adsorption on the formation of linear protein fibrils**, Travis Hoppe, Allen Minton, *Biophysical Journal*
- 2013 **A simplified representation of anisotropic charge distributions in proteins**, Travis Hoppe, *Journal of Chemical Physics*
- 2013 **Singular Value Decomposition of the Radial Distribution Function for Hard Sphere and Square Well Potentials**, Travis Hoppe, *PLoS ONE*
- 2010 **Protein Folding with Implicit Crowders: A Study of Conformational States Using the Wang-Landau Method**, Travis Hoppe, Jian-Min Yuan, *Journal of Physical Chemistry B*
- 2009 **Entropic flows, crowding effects, and stability of asymmetric proteins**, Travis Hoppe, Jian-Min Yuan, *Physical Review E*
- Graph theory*
- 2014 **Integer sequence discovery from small graphs**, Travis Hoppe, Anna Petrone, *Discrete Applied Mathematics*
- Experimental Modeling*
- 2014 **Programmable Nanoscaffolds that Control Ligand Display to a G-Protein Coupled-Receptor in Membranes allow Dissection of Multivalent Effects**, Andrew Dix, Daniel Appella, Travis Hoppe, et al., *Journal of the American Chemical Society*
- 2014 **Quantification of plasma HIV RNA using chemically engineered peptide nucleic acids**, Chao Zhao, Daniel Appella, Travis Hoppe, et al., *Nature Communications*
- 2008 **The importance of EBIT data for Z-pinch plasma diagnostics**, A S Safronova, Travis Hoppe, et al., *Canadian Journal of Physics*
- 2006 **Spectroscopic and Imaging Study of Combined W and Mo-pinches at 1 MA-pinch Generators**, Alla Safronova, Travis Hoppe, et al., *IEEE Transactions on Plasma Science*