Nandita **Rahman**

Data and AI Strategist · M.S. ·

Washington, DC 20009

(559) 593-9250 | nandita.rahman@nih.gov | rahmann3 | nanditarahman

Summary

Interoperability issues in the U.S. create significant barriers to accessing health data, negatively impacting research, wasting IT investments, and hindering the government's efforts to enhance health outcomes. The lack of technical expertise in the federal workforce further complicates IT procurement strategies.

As an interdisciplinary data scientist with over ten years of experience in the private and public sectors, spanning health data, machine learning, and AI policy domains, I work with teams to develop tailored AI and data governance programs that address their specific challenges and maturity levels. I provide valuable technical expertise during IT procurement, ensuring agencies make informed investment decisions and evaluate investments using data-driven and analytical methods to determine how they serve the American people's interest and well-being.

I take a creative approach to build consensus and requirements with stakeholders across each layer of an enterprise's technical stack. With a strong technical background, strategic writing, and effective interpersonal style, I facilitate collaborative requirement development and simplify complex discussions among key stakeholders. My ultimate goal is to improve collaboration between the private and public sectors, reducing waste and delivering quality results through government procurement in the technical space.

Skills

Policy Focus Areas	Award Portfolio Analysis, Metric-development, Evidence-driven Strategy Design, AI Research, Data Governance
Communication	Contract Management, Strategic Workshop Development, Facilitation, Community Engagement,
Programming Languages	Python, R, Linux, LaTeX, SQL, Tableau
Computational Skills	Confluence, Jira, Git, Jupyter, Cloud (AWS,GCP) and High-performance Computing, Visual Studio
AI Evaluation Focus Areas	Model Interpretablity (global, local), Fairness Metrics, Bias Detection-Mitigation, Calibration Testing
Machine-learning Skills	Natural Language Processing, Topic-modeling, Ensemble-methods, SHAP, LIME, Feature-engineering, LLMs

Work Experience

National Institutes of Health

Bethesda, MD

Program Officer (Data Scientist) (GS-14 Step 10; 40.0 Hours/Week) *Nov. 2023 - PRESENT*

- **(Current Role) National Institutes of Health - Office of Strategic Coordination, Catalytic Data Resource Team**: I support the Common Fund Data Ecosystem (CFDE) which aims to create a federated infrastructure to integrate data, resources, and knowledge across Common Fund Programs. Together, these resources will empower the research community to use Common Fund data sets for novel scientific research that was impossible before. This may include hypothesis generation, discovery, or validation that leads to new insights in health and disease. The CFDE is also supporting pilot projects that engage a broad community of end-users, collect feedback on the utility of Common Fund data resources, and enable novel cross-cutting biological questions to be formulated and addressed.
- **(Previous Role) National Institutes of Health - All of Us Research Program**: Data Governance Director, at the Office of Data and Analytics. Policy and Research lead developing and maturing Program-wide data governance framework, and best practices.
- **Data Operations**: Subject matter expert for drafting language and technical milestones for data ecosystem awards that include having authoritative data sources using schema metadata; and ubiquitous (consistent) language for referencing data, data modeling, data domains, data products, data elements.
- **Data Governance**: Established the program's initial data governance committee and program. This includes getting buy-in and participation around data ownership models as part of a collaborative data management practice focused on improving the efficiency of communication, integration, and automation of data flows between teams across an organization.
- **Strategic Writing and Engagement**: Facilitating workshops to deploy maturity assessment collateral to engage and measure opportunities and gaps across data owners and systems. Using results from such tools to build strategic data objectives and implementation plans to demonstrate measured improvement in data management practices.
- **AI Risk Management**: Experienced in building AI/ML models that employ fairness assessment, bias mitigation, explainability and privacy-preserving methods. In adherence to federal mandates, characterized and developed AI inventory and change management process. Leveraging engagement experience, also developed training collateral, AI-researcher toolkits (e.g., AI audit checklists) and hosted hackathons to promote responsible use of AI and development.

Deloitte

Rosslyn, VA

Advisory Senior Consultant (GS-14 Step 1 Equivalent; 40.0 Hours/Week) *Aug. 2021 - Nov. 2023*

- **U.S. Department of Veteran Affairs- Office of Information and Technology (VA-OIT)**: Independently authored agency-adopted guidance documents to support AI leaders in developing ethical AI frameworks within the field of cybersecurity.
- **U.S. Department of Health and Human Services-Office of the Chief AI Officer (HHS-OCAIO)**: Developed the agency-adopted **HHS Trustworthy AI Playbook**, to support AI leaders in creating ethical AI solutions. This documentation has been widely distributed on public forums as guidance and best practices.
- **Center for Medicaid Services-Office of Minority Health (CMS-OMH)**: Model evaluation and guidance framework development to promote equitable health outcomes. Investigated CMS models used in Medicare and Medicaid for fraud, waste, and abuse, enabling equitable outcomes for beneficiaries.
- **Relevant skills:** AI and data governance policy development, Model evaluation, Data management, Python, R, Cloud-native solutions.

Halıcıoğlu Data Science Institute, University of California - San Diego

San Diego, CA

Lecturer (15.0 Hours/Week) *Aug. 2022 - August 2024*

- **Course Link:** DSC180-AB: Data Science Senior Capstone: Responsible AI
- The capstone program comprises a two-quarter sequence in which students are mentored by a faculty or industry expert in the domain of responsible AI.
- Quarter 1 (DSC 180A), students gain background information in their mentor's domain by replicating a known result related to responsible AI development. By the end of Q1, students complete a replication project and propose a more independent project focused on ethical AI.
- Quarter 2 (DSC 180B), students execute the Q2 Project proposed at the end of Q1, addressing responsible AI development challenges.
- Core competencies: Python, Anaconda, Machine-learning Model Development, AI Risk Management, Fairness Metrics, Debiasing methods, IBM Fairness 360 Toolkit, Medical Expenditure Data, AI Policy Development

University of Baltimore County

Baltimore, MD

Adjunct-Lecturer (15.0 Hours/Week) *Dec. 2021 - Sep 2022*

- Course: BTEC 330: Software Applications in Biotechnology
- This course will provide an overview and basic practical skills in software tools that are used widely in biological research and development in the areas of general productivity, basic data analysis and databasing, statistical analysis and programming, analysis and comprehension of high-throughput genomic data, biological sequence analysis and bioinformatics, image analysis and morphometry.
- Core competencies: Python, R, Linux, Statistical analysis, Anaconda, Git, Research methods, Machine-learning, and data science methods

Noblis

Reston, VA

Bioinformaticist (40.0 Hours/Week) *Sep. 2019 - Aug. 2021*

- Developed and communicated machine learning findings on large biological datasets using Python, R, MATLAB, and Tableau, to wide range of audiences.
- Wrote Bioinformatics and machine-learning scripts using R, Python, SQL and 3D-printing languages. Wrote documentation for client products using JIRA and Atlassian tools.
- Scientific research awardee. Managed seedling research funds and led team of 3-5 data scientists related to applied AI applications for modernizing government processes.
- Bioinformatics mentor for Noblis intern cohort for Summer/Winter 2020, and currently Summer 2021.
- Primary author for publication in mSystems Biology (2021) for biological 3D-printing research, and contributing author in bioinformatics publication in BioMed Central (2022)

Military HIV Research Program, Walter Reed

Silver Spring, MD

Bioinformatics Research Associate (40.0 Hours/Week) *Sep. 2018 - Sep. 2019*

- Analysis of transcriptomics through RNA-sequencing (single-cell and bulk), to understand potential genetic signatures involved with immunogenic host response to vaccination and infection.
- Developing bioinformatic pipelines on high-throughput NGS data (multiple platforms) from human clinical vaccine cohorts.
- Use of UNIX/LINUX, Job submission and resource allocation on high-performance computing grids (SLURM).
- Use of JIRA, Atlassian tools to develop R, python and pipeline scripts and documentation.

Canon Biomedical

Rockville, MD

Scientist (40.0 Hours/Week) *Aug. 2014 - Sep. 2018*

- FDA regulatory work: experimental design for device-development, preparing reporting for FDA 510k submission packets for Class II medical prototype-device development team, built language and gained consensus with stakeholders for design-history file documents.
- Molecular and laboratory researcher, developing genomic pharmacogenetic assays such as KRAS, BRAF, MTHFR, CYP2C19 and other genetic-screening panels
- Biological Safety Officer, developed framework for BSL-2 laboratory work, including documentation system, training suite, compliance/reporting and building controls.
- Primary author for academic publication in BioTechniques related to novel polymerase chain reaction (PCR) pharmacogenetic-assay methods (2018)

Education

George Mason University

Fairfax, VA

Master in Computational Biology and Bioinformatics *Jun. 2017 - May. 2020 (Graduated)*

Minnesota State University - Mankato *Mankato, MN*

Bachelors in Microbiology, Chemistry *Jun. 2009 - Jul. 2014 (Graduated)*

Certifications *Certification Program*

AWS, GCP Certified Cloud Practitioner *Issued Jul. 2022, 2024*

Publications & Projects

Knowledge-Driven Data Sharing Governance: Developing a data provenance model for constructing context-aware data sharing frameworks for linkage data within the All of Us Research Program *IJPDS*

DOI:10.23889/IJPDS.V9I5.2883 *Sep. 2024*

Advancing cancer disparities research through the integration of new data: Opportunities from the National Institutes of Health All of Us Research Program *AACR*

DOI:10.1158/1538-7755.DISP24-B170 *Sep. 2024*

SARS-CoV-2 Delta variant isolates from vaccinated individuals *BMC Genomics*

DOI:10.1186/S12864-022-08652-Z *Jun. 2022*

Biomimetic Gut Model Systems for Development of Targeted Microbial Solutions for Enhancing Warfighter Health and Performance *mSystems Biology*

DOI:10.1128/MSYSTEMS.00487-20 *Oct. 2020*

A vaccine-induced gene expression signature correlates with protection against SIV and HIV in multiple trials *Science Translational Medicine*

DOI:10.1126/SCITRANSLMED.AAW4236 *Aug. 2019*

T-blocker: a simple and robust probe-free quantitative PCR assay to detect somatic mutations down to 0.1% frequency *BioTechniques*

DOI:10.2144/BTN-2018-0111 *Oct. 2018*

Extracurricular Activity

American Association for the Advancement of Science (AAAS) *Washington, D.C.*

Scientist in the Classroom (1.0 Hour/Week) *Sep. 2020 - Present*

An organization of science, technology, engineering and mathematics (STEM) professionals who volunteer in K-12 classrooms in 7 school districts in the Washington DC metro area.

COVID Work-study Group *Washington, D.C.*

Back End Developer (10.0 Hours/Week) *Mar. 2020 - Sep. 2020*

https://www.covidworkstudy.com/team | Our goal is to identify elements of the working environment and novel coronavirus-related safety protocols that are associated with higher or lower prevalence of COVID-19. This information will be made publicly available so that professionals and business owners can make informed decisions regarding workplace safety. All individual data and identifying information will be kept strictly confidential. We are collecting survey responses from each participant on an ongoing, weekly basis so that we can provide data visualizations and analyses that are timely and sensitive to changing conditions.