

Ozan Aygun, Ph.D.

E-mail: ozan.aygun5@gmail.com | Phone: (240) 599-6002 | Portfolio: <https://ayguno.github.io/curious>

Senior Health Data Scientist with over 10 years of experience driving transformative data analytics, solutions in complex medical research, regulatory and technology environments. Expertise in all stages of data analysis circle to implement data-driven strategies that improve stakeholder outcomes, user experience and operational efficiency. Skilled in using advanced analytics, AI/ML, and data science to enhance decision-making and drive innovation. Proven success in developing comprehensive data strategies, optimizing data governance frameworks, and ensuring data quality, privacy, and security across systems. Extensive experience working with medical research organizations such as the FDA, NIH, and MIT, with a focus on healthcare data initiatives. Demonstrated leadership in advancing data literacy, mentoring teams, and building successful cross-functional partnerships to accelerate data-driven innovation. Currently teaching at the Data Analytics program at Georgia Institute of Technology.

KEY DATA SCIENCE, ANALYTICS, AND COMPUTATIONAL SKILLS

- **Machine Learning and Statistical Modeling using R and Python:** built and deployed production-quality machine learning and statistical models including; regularized regression (Ridge and Lasso), Mixed Effect Models, Longitudinal Analysis, Time-Series analysis, forecasting, supervised learning and unsupervised learning. Simulation, Optimization and Prescriptive Analytics.
- **Natural Language Processing (NLP) and Text Analytics:** performed production scale large-scale document analysis, embedding analysis, feature engineering from unstructured text, Latent Semantic Analysis (LSA), and Topic Analysis (LDA). Hugging Face, GenAI/LLM-based solution development (Ollama, LangChain, AWS Bedrock).
- **Data Visualization Solutions by R, Python and BI Tools:** built and deployed numerous web-based and stand-alone desktop applications using R-shiny. Deployed Python machine learning models using R-shiny as well as REST APIs. Tableau and Power BI.
- **Cloud Computing:** AWS Cloud ecosystem (S3, SageMaker, Athena, Glue, State Machine), Databricks.
- **Front-end web programming:** HTML, JavaScript and D3, CSS, version control, Git, GitHub.

PROFESSIONAL EXPERIENCE

Supervisory Data Scientist (November 2022 - Present) (Public Trust) **Silver Spring, MD**
U.S. Food and Drug Administration (FDA), Center for Devices and Radiological Health

- Delivered two AI solutions to automate clinical evidence extraction from millions of unstructured documents, cutting retrieval costs by 70%, and earned the FDA's AI application award in 2024.
- Modernized data governance by eliminating redundant cloud storage, and reducing IT program costs by \$500K. Led open-source adoption, cutting text processing costs by 80%.
- Developed a data governance framework for medical device safety surveillance, addressing data gaps and integrating internal/external data with AWS cloud solutions to serve over 1,500 users.
- Created roadmaps for implementing AI algorithms, including Gen AI and LLM solutions. Presented to senior leaders to align AI strategies with agency goals, earning an FDA incentive award in 2024.

- Led strategic and cost-effective multi-year cloud migration for the AI program, collaborating with IT officials and contractors to deliver impactful business solutions.

Senior Health Informatics Specialist (June 2022 – November 2022)

Bethesda, MD

National Institutes of Health (NIH), Office of the NIH Director

- Advised executive leadership on health data analytics solutions and architecture for \$1.5 billion budget public health data initiatives, providing expertise in Big Data technologies, Natural language Processing (NLP), AI strategies, EHR integration, data governance, and statistical analysis planning.
- Led the characterization of 100,000 precision medicine study participants by integrating EHRs, demographic, and genomic data, and identifying opportunities for auxiliary public health studies.
- Accelerated external public health studies and data governance frameworks, guiding the integration of new data assets into the existing governance ecosystem.
- Counseled senior leadership on evaluating complex, multi-center public health program research, providing strategic input on work products, delivery streams, and communications.

Senior Operations Research Analyst (August 2019 – June 2022)

Silver Spring, MD

U.S. Food and Drug Administration (FDA), Center for Drug Evaluation and Research

- Led AL/ML solutions for the FDA's \$1.1 billion user fee program, driving financial effectiveness and supporting the agency's public health mission.
- Built and scaled 12 distinct AI/ML frameworks to forecast workload outcomes and enable data-driven decision-making, while collaborating with cross-functional teams to streamline data platform production.
- Modernized the FDA's machine learning initiative for forecasting workload, improving financial planning and staffing by over 100% compared to legacy methods, earning the FDA Honor Award (2021) and FDA Incentive Awards (2020 & 2021).
- Oversaw data-driven optimization for strategic hiring and financial planning in the FDA's largest clinical drug review office, developing an algorithm that influenced an annual budget allocation of \$4.5 million.
- Guided a team of 6 data scientists, evaluating machine learning pipelines and improving AI/ML model performance, saving the agency thousands of dollars annually.

FDA Senior Staff Fellow (September 2017 – August 2019)

Silver Spring, MD

U.S. Food and Drug Administration (FDA), Center for Devices and Radiological Health

- Developed the 510(k) TIMELINE PREDICTOR, an ML-based solution that improved predictive accuracy of regulatory project timelines by 45%.
- Created the NETWORK JUMPSTART, a web-based dashboard that streamlined regulatory review, reducing precedent mapping from hours to seconds.
- Spearheaded the development of SIMPLIFY, a web-based reporting app that integrated real-time

enterprise data and automated stakeholder reports, cutting reporting time by 90%.

- Led the regulatory review of medical devices and molecular diagnostics for oncology, guiding interdisciplinary teams and overseeing the entire review process, from submission to executive decision recommendations.
- Chaired stakeholder meetings, led technical discussions, and authored regulatory memorandums and decision summaries.
- Oversaw the regulatory review of innovative oncology diagnostic devices and AI/ML algorithms, authoring over 100 technical review reports that supported senior leadership's decision-making.

Senior Data Scientist (January 2016 – September 2017)

Cambridge, MA

Broad Institute of MIT and Harvard

- Managed 10+ multi-institutional collaborations with 2 pharmaceutical companies and 8 academic teams, including Harvard, MIT, and MGH, and authored 3 peer-reviewed publications.
- Architected JURKAT SCRAPER, an online platform for real-time monitoring of Mass Spectrometer performance, saving thousands in maintenance costs by enabling early issue detection.
- Led a large-scale oncology data project with a pharmaceutical company, driving the discovery of mechanisms of action for 5 clinically relevant anti-cancer drugs.
- Directed multi-institutional oncology research collaborations, leading the design and implementation of data analytics solutions to advance biomarker discovery in molecular oncology.
- Built strong partnerships with academic and industry leaders, promoting best practices in data sharing and analytics.

Charles King Trust Fellow (September 2013-December 2015)

Cambridge, MA

Massachusetts Institute of Technology (MIT)

- Accelerated statistical analysis of large data sets, developing algorithms and a streamlined pipeline that reduced analysis time by 30%, supporting novel scientific methodologies.
- Modernized an inventory database to optimize chemical and biological specimen use, enabling fast retrieval of over 1,000 critical assets and streamlining lab operations.
- Designed and led two research projects, securing a \$100,000 grant for a two-year period by collecting experimental results and writing successful proposals.

RELEVANT PROFESSIONAL TEACHING EXPERIENCE

Senior teaching faculty member (May 2021 - Present)

Georgia Institute of Technology, Program of Master of Science in Analytics

Remote work, Part-time

Coached and mentored graduate data analytics students, managing weekly workload of a team of 25 teaching assistants and providing guidance on machine learning, statistics, and analytics modeling. Conducted office hours and offered thought leadership to students from diverse professional backgrounds.

- Mentored graduate students to enhance their proficiency in data analytics concepts and their applications across various business domains.
- Provided coaching to bridge the gap between academic knowledge and real-world business applications of AI/ML technologies, emphasizing the responsible and ethical use of AI and data analytics.

EDUCATION

Georgia Institute of Technology (Degree obtained: May 2021)

Master of Science (MS) in Analytics

University College London (UCL) (Degree obtained: February 2010)

Doctor of Philosophy (Ph.D.) in Biochemistry and Molecular Biology

Middle East Technical University (METU) (Degree obtained: June 2006)

Bachelor of Science (BSc) in Molecular Biology and Genetics

AWARDS

- AI for Regulatory Review Award, FDA (2024)
- FDA Incentive Award (2024)
- FDA Honor Award (2021)
- FDA Incentive Award (2021)
- FDA Incentive Award (2020)
- Insight Health Data Science Fellowship (2017)
- The Medical Foundation, Charles King Trust Fellowship (2014)
- European Molecular Biology Organization (EMBO), Long-term Fellowship (2010)
- London Research Institute (LRI), PhD Student Seminar Prize (2008)
- Cancer Research UK, Overseas Graduate Research Fellowship (2006)
- University of Glasgow, Graduate Research Fellowship (2006)
- German Cancer Research Centre (DKFZ), Graduate Research Fellowship (2006)
- German Cancer Research Centre (DKFZ), Summer Research Fellowship (2005)
- German Academic Exchange Service (DAAD), Summer Research Fellowship (2004)