Developing and Using Data Visualization Requirements Gathering Processes for Clients

Sarah Glidden, Westat

Creating data visualizations and dashboards for clients' decision-making needs requires an iterative requirement gathering process before actual visualization development can begin. Developing consistent requirement gathering steps that are used across all data visualization projects ensures consistency and accuracy as a data visualization team. An interview with the client is the first step to assessing their needs and the properties of the data that will be used, as well as understanding their level of experience communicating about visualizations. Following this interview, formal write-ups and static draft visualizations can be iterated on between the data visualization analyst and the client. This talk will share processes, strategies, and challenges from data visualization requirements gathering that may benefit data visualization analysts working with a diverse array of clients.

Mapping Global Open-Source Collaboration: A Data Science Approach to Measuring International Software Development

Rashi Saluja, Westat

Open-source software (OSS) has emerged as a critical indicator of a nation's science and technology capacity. Understanding who contributes to OSS offers valuable insight into the structure of global innovation. This study presents a data-driven framework for measuring international collaboration in open-source software development, developed as part of the 2026 Science and Engineering Indicators report, Translation to Impact: U.S. and Global Science, Technology, and Innovation Output, published by the National Science Board, at the U.S. National Science Foundation. OSS repositories serve as rich, observable records of developer activity, capturing not just code but contributor identity, institutional affiliation, and geographic origin. We constructed a large-scale data pipeline to collect and process repository- and contributor-level records spanning 2008–2025. This enabled us to identify repositories where developers from multiple countries co-contributed.
The resulting tabulations and network-based visualizations reveal collaboration patterns across the top contributing nations globally. Beyond measurement, this work surfaces questions of representation: whose contributions are visible in global OSS, which sectors and geographies lead, and where the data goes silent. These questions sit at the heart of what it means to create truly global connections - and why diverse perspectives in data science matter how we build and interpret such indicators.

Multimodal AI Fusion for Better Alzheimer Detection

Chipo Zidana, ZIMYAS

Background: Diagnosing Alzheimer's disease (AD) and its precursor, Mild Cognitive Impairment (MCI), is hindered by hidden correlations between genetic factors and neurostructural changes. Isolated modality evaluation misses critical systemic insights.

Methods: We propose a dual-branch multimodal AI fusion framework integrating functional fMRI and genomic data from the ADNI database. Significant differentially expressed genes (DEGs) identified via LIMMA in R were mapped to structural brain regions of interest (ROIs) utilizing the UK Biobank. Volumetric spatial and high-dimensional time-series features were extracted using 3D convolutions and tsfresh. Classification was evaluated across Random Forest, SVM, and deep Neural Network architectures.

Results: Our framework identified 12,214 significant probe sets mapping to 10 unique genes regulatory of structural variations in key ROIs like the lateral occipital cortex. Targeted biological ROI feature selection drastically minimized computational noise, outperforming all-feature models. The Neural Network branch achieved a hold-out accuracy and AUC of 1.00, while SVM and Random Forest reached cross-validation accuracies of 97% and 93% respectively.

Conclusion: Fusing genomic variables with neuroimaging eliminates noise and accelerates MCI-to-AD classification accuracy. This interpretable pipeline establishes a powerful paradigm for target-driven precision diagnostics.

End-of-Life Healthcare Utilization among Medicare Beneficiaries with Parkinson's Disease Dementia at USA Academic Movement Disorders Centers: A 2019 Pre-Intervention Assessment

Adriana Perez, The University of Texas Health Science Center at Houston

Evaluating end-of-life (EOL) care for hospitalized Medicare decedents with Parkinson's disease dementia (PDD), defined here as the last 90 days of life in 2019, poses statistical challenges when analyzing cohorts with PDD beneficiaries clustered within hospitals. The four primary outcomes were: non-elective inpatient hospitalization, intensive care unit (ICU) admission, in-hospital death, or receiving palliative care consultation. PDD beneficiaries first visited either a comparison Academic Movement Disorders Center (AMDC) or an AMDC intervention-scheduled center. The adjusted odds ratios (AOR) for the simultaneous effect of visiting AMDCs, controlling for sex, race, age, Social Vulnerability Index (SVI), and the cluster AMDC effect, are reported. The sample sizes were 675, 628, 507, and 486 for each outcome, respectively. There were no statistically significant differences in any health outcomes between PDD beneficiaries who first visited an AMDC intervention-scheduled center and those who first visited an AMDC comparison center. After adjustment, PDD beneficiaries of races other than non-Hispanic white had higher odds of non-elective inpatient hospitalization (AOR=1.92; 95% CI 1.05–3.52), as well as ICU admission (AOR=2.43; 95% CI 1.29–4.58). How AMDC centers and socio-demographic characteristics influence PDD outcomes 90 days before death helps to inform future program design and policy recommendations aimed at optimizing care for PDD Medicare beneficiaries.