Data Visualisation & Open Source Technology

Working Group Scope
Data Visualisation and Open Source Technology aims to support, address, and answer pertinent questions around Data Visualisation and Open Source Technology. The combination of these two subjects is natural in today's environment given the powerful Data Visualisation tools within the Open Source languages available today. Some of the questions, amongst others, that we intend to address are: • How do you safely use Open Source languages for analytics and submissions within a Regulatory environment? • What are the potential uses of Open Source software within a company outside of data analysis for a submission? • How can interactive visualisations be leveraged appropriately within a clinical environment? • What are the best practices for creating powerful interactive visualisations?

Current Projects
Best Practices for Interactive Analysis for Decision Making Submissions
Clinical Visual Analytics for Review and Submission (CVARS)
Comparing Analysis Method Implementations in Software (CAMIS)
Demonstrating Real-World Impact of Modernization of Statistical Analytics (MSA) Framework
PharmaDevOps: Development Operations Guidance for Clinical Reporting Codebases

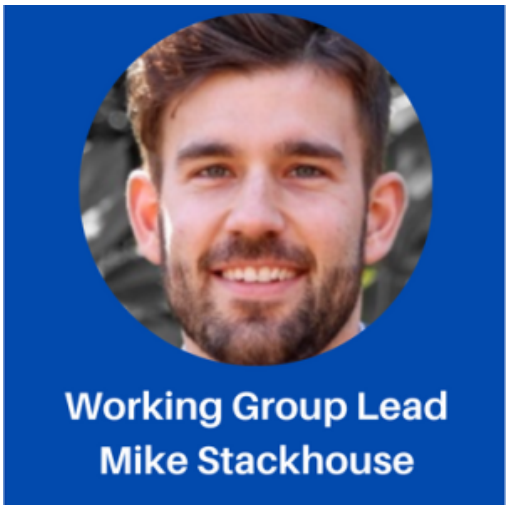
hanming.tu@gmail.com

Hanming Tu is a highly experienced professional with over 25 years of expertise in information technology, information security, database administration, data integrity, data analytics and project management, with more than 20 years of experience in the pharmaceutical industry and CDISC standards. He is currently the Co-Founder and Chief Information Officer at Ashanda, having previously held positions such as Security Officer, VP of IT and Database Administration (DBA) at Frontage Laboratories, Director of Clinical IT at Octagon Research and DBA Manager at Accenture.

Hanming has participated in several CDISC standard development teams, where he has contributed to the initial development of CDISC standards such as the Study Data Tabulation Model (SDTM), the Operational Data Model (ODM) and the Protocol Representation Model (OPM). He has also been involved in e-submission and data conversion projects and architected the Automated Data Conversion Development (AutoDCD) program for a large FDA data conversion project at Octagon Research. Moreover, Hanming has developed a Define-XML generator and contributed to the CheckPoint application for SDTM data validation using Oracle PL/SQL language at Accenture.

In addition, Hanming has been a Co-Lead for the Standard Analyses and Coding Sharing Working Group at PHUSE since 2012. This Working Group merged with others to become the current Data Visualization and Open Source Technology Working Group in 2020. Through his participation in these Working Groups, Hanming has contributed to the discussion and presented numerous times on emerging technologies, data standardisation, data visualisation, data transformation and automation intelligence at industry conferences such as DIA, CDISC, PHUSE and PharmaSUG. He is also an active contributor to the open-source community, having published several R and R shiny packages in Comprehensive R Archive Network (CRAN), Perl modules in the Comprehensive Perl Archive Network (CPAN), and developed several Python packages.

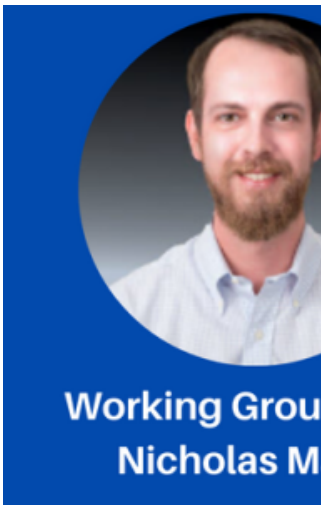
Apart from his work in information technology and the pharmaceutical industry, Hanming holds a master's degree in city and regional planning from Ohio State University and a master's degree in physical geography from Central China Normal University.



mike.stackhouse@atorusresearch.com

Michael Stackhouse is the Chief Innovation Officer at Atorus Research. He has extensive CDISC experience, working with both Study Data Tabulation Model (SDTM) and Analysis Data Model (ADaM) standards, and serving as a subject matter expert for Define.xml. He holds a bachelor's degree from Arcadia University, where he studied business administration, economics, and statistics. He is a 2020 UC Berkeley School of Information Master of Information and Data Science (MIDS) program graduate, where he worked on projects involving computer vision, natural language processing, cluster computing, and deep learning. His special interests include automation, machine learning, big data technology, and mentoring rising programmers.

Previously, Michael was a senior manager of statistical programming at Covance, where he led U.S. innovation activities for the FSP department. Under his guidance, projects achieved data standardization according to SDTM standards on upwards of 75 studies, including database integration and data warehousing. He also managed programming activities through a multiagency submission for multiple studies across a single compound. In addition, he took on multiple automation projects, including the development of a tool capable of dynamically locating programming independence violations and automatically detecting protocol deviations, as well as the creation of data pipelines around tracking systems for programming deliverables. Michael and his team at Atorus have been actively developing and releasing open source R packages, such as pharmaRTF and Tplyr.



nmasel@its.jnj.co

Nicholas Masel is the Associate D
Team Lead within Clinical & Statisti
Janssen R&D. In this role, he is res
the design and implementation of te
global department use, as well as e
tools and technologies for potential
in support of portfolio needs. Some
to the Data Visualization and
Technologies scope of work include
transition of Janssen R&D's Statis
group from an SAS-based env
environment that supports SAS
Nicholas's responsibilities includ
Implementation Team, developing i
R packages, validating R package
and deploying Shiny applications
ecosystem. He is also an activ
contributor to several external orga
PHUSE. He is the Janssen repre
pharmaverse Adoption Working Gro
and Templates Working Group,
Modernization of Statistical Analy
member of R Consortium's Infras
Committee. He also contributes a
developer to logrx, envsetup

Prior to transitioning to the Associate
Innovation Team Lead role, Nicholas
Programming Lead responsible for t
oversight of statistical programming
of clinical projects. He holds a maste
economics from East Carolina Unive
Raleigh, NC with his wife and dog. I
enjoys organising the RTP R User G
to the local community through the F
Hills, helping in the local community
recently joined CrossFit (where he fi
everything).