Proposal

This is the original proposal that was prepared for the Colorado WASH Symposium 2021. It is also available as a Google Doc if you are interested to comment on it.

# Proposal

**Why is the topic relevant to the current trends in the WASH sector?**

WASH professionals are increasingly challenged with the need for (research) data management. The requirements for sharing data openly are increasing, whether that is by funding agencies, scientific journals, or civil society. Open by default. A term for a new era of Open Research Data and Open Government Data. A term which requires publication of non-sensitive data while following personal data protection laws. A term that puts the publication of data first before asking questions of why it should be shared.

A community within the WASH sector that actively works on preparing data for publication has yet to be created. Few have been exposed to FAIR Data Principles for sharing data or understand the benefits of open file formats, machine-readability or machine-actionable metadata. While unprocessed (raw) data is rarely shared, even the processed (analysis-ready) data is most likely stuck in proprietary spreadsheet-based and word processing software, or worse in PDFs.

The benefits for academics to apply open data practices to their work are well-researched and understood (McKiernan et al. 2016), but need to be clarified for professionals working at the policy, implementation or evaluation components within a WASH project. As a result, data remains unpublished (or published not following FAIR principles (Wilkinson et al. 2016)), collaboration is stifled, resources are wasted, and the wheel is constantly reinvented at extensive cost to donors and project beneficiaries. Further, the lack of competencies can make it challenging to establish partnerships between researchers applying ORD practices and partners that are not (GSMA 2021).

**Consider unique or innovative presentation options**

This workshop continues the “Data Science for WASH” efforts, which started at the Colorado WASH Symposium 2021. Once again, participants will be able to gain hands-on practical experience using Posit Cloud for interactive live coding and GitHub for the publication of data. A brief slideshow presentation will be used to introduce participants to openwashdata. This newly formed community addresses the issues outlined above. Participants will get to discuss and share their personal data management experiences in break-out sessions where the outputs are captured using a digital whiteboard solution (e.g. Padlet). The event will be hosted remotely using Zoom, which allows physical participation from the conference and online participation. The only requirements for participation are an internet connection, a browser and an account on GitHub.

**How does the session support the mission of the Colorado [[WASH]] Symposium**

This workshop supports the mission of the Colorado WASH Symposium by offering participants the opportunity to learn about a community effort that may positively impact how we collaborate in the WASH sector. Participants will be able to actively participate and contribute during the session and beyond.

**How does your work confront (or perpetuate) systemic injustice and inequity?**

The workshop and the broader openwashdata community confront injustice and inequity by offering active participation without needing a physical presence, costly software or managed infrastructure. The computing environment is provided and is based on open-source and free software (R and RStudio IDE). The entire process of sharing WASH data openly ensures that those who have contributed to collecting and publishing data also receive appropriate credit. Lastly, the community and this workshop have a Code of Conduct that supports a harassment-free experience for everyone.

GSMA. 2021. “Innovative Data for Urban Planning: The Opportunities and Challenges of Public-Private Data Partnerships.” *Mobile for Development*. <https://www.gsma.com/mobilefordevelopment/resources/innovative-data-for-urban-planning-the-opportunities-and-challenges-of-public-private-data-partnerships/>.

McKiernan, Erin C, Philip E Bourne, C Titus Brown, Stuart Buck, Amye Kenall, Jennifer Lin, Damon McDougall, et al. 2016. “How Open Science Helps Researchers Succeed.” Edited by Peter Rodgers. *eLife* 5 (July): e16800. <https://doi.org/10.7554/eLife.16800>.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. “The FAIR Guiding Principles for Scientific Data Management and Stewardship.” *Scientific Data* 3 (1): 160018. <https://doi.org/10.1038/sdata.2016.18>.