

New Forms of Data and Scientific Research

David De Roure*

University of Oxford, Oxford, UK david.deroure@oerc.ox.ac.uk

Data is now captured digitally in all scientific disciplines, from large experiments to sensor networks. In some cases this enables an acceleration of research, shortening “time to outcome”. However the new data sources also enable entirely new research questions to be answered. New forms of data require new techniques of analysis and interpretation, and there is a further methodological change whereby scientific practice is becoming increasingly data-centric where previously the collection of data was hypothesis-led. Dealing with the increasing scale and variety of data, and with its new qualities, also raises challenges in computational methods as well as data management, movement and curation. In particular increasing automation is essential for systematic processing of the bulk of data available, and this is leading to new techniques such as computational workflows. The greater participation of the public in the digital world leads to not only new data about and from citizens but also the new techniques of citizen science. All these changing practices bring new requirements to the way that science is recorded for scholarly communication and for reproducibility.

Key Words: Data-Intensive Research, Web Observatory, Computational Workflow.