Open Science: A New Paradigm for the Research Lifecycle

Yannis Ioannidis
University of Athens and Athena Research Center
Athens, Greece
yannis@{di.uoa.gr,athenarc.gr}

ABSTRACT

Open Science is a fundamentally new philosophy about the goals, methods, and business models of scientific research and scholarly communication. The motivation for research remains the same, as Open Science continues to honor scientific curiosity and societal or industrial challenges as its primary drivers. Open Science methodologies, however, continue to spotlight individual researchers for their achievements, but also promote common good. They are motivated by the following needs: increase researchers' collaboration, accountability, and transparency; democratize researchers' and the general public's access to knowledge around the world; reconceive of what constitutes a publication to include data, software, and other artifacts beyond traditional papers; elevate the importance of research results' reproducibility; and increase society's trust to the scientific community. Open Science affects every single stage of the research lifecycle significantly. Naturally, the scientific community and the extended ecosystem around it (university administrators, research funders, journal publishers, conference organizers, …) need ample time to move towards the new mentality. In preparation for this change, this panel exposes the key dimensions of Open Science to the database community, highlights the various choices offered in each dimension, and debates their advantages and disadvantages and the way forward.

1 PANEL MOTIVATION AND CONTENT

Next to the dramatic AI-centered changes in the technological foundations of all sciences (including computing science and, therefore, data management), for the past two decades, we have witnessed dramatic Openness-centered changes in the policy foundations of the operation, management, and communication of science. Bottom-up demands by the scientific community across all disciplines for democratization of access to scientific output have also led to top-down mandates by governments and funding agencies around the world for scientific transparency within the scientific community and accountability to society at large.

Open Access to scientific publications and FAIRness (Findable, Accessible, Interoperable, Reusable) of other research artifacts (datasets,

software, etc.) are critical but capture only one aspect of a much broader openness vision. Open Science is a whole new paradigm of doing research, promoting collaboration, innovation, and reproducibility, aiming at greater social impact and increased public trust in science. It requires new scholarly communication forms, different research assessment criteria, and significant technological support.

In this panel, we examine the basic principles of Open Science and how it may impact the researchers' everyday life. We highlight some of the main steps in the typical research lifecycle and give details on the changes that must occur to serve the open science philosophy, including open access, open reviews, open infrastructures, open research preregistration, and others. Inspired by Open Science, we also explore the foundations of our current publication methodologies in computing (including databases) and investigate the motivation behind publishing in journals vs. conferences and other aspects. Finally, we address newly arising existential and ethical challenges for researchers and acceptability of research results.

All aspects of Open Science mentioned above are general and affect all scientific disciplines. In addition to those, however, in this panel, we also try to address some specialized questions on how Open Science impacts or is impacted by Data Science and, unavoidably, Artificial Intelligence. For example, does Openness in the research lifecycle facilitate the development of AI? Does Openness in AI models facilitate the advancement of science? Does Openness pose significant new challenges to data processing?

2 PANEL COMPOSITION

The panel comprises distinguished professionals, each bringing to the table different characteristics and backgrounds regarding Open Science and our field: different geographies, different experiences with Open Science, involvement (or not) with the data management community, and other characteristics. This diversity promises a balanced discussion on the key issues. The panelists are the following:

- **Yanlei Diao**, Professor of Computer Science, École Polytechnique, France, and Professor of Computer Science, University of Massachusetts – Amherst, USA
- **Dame Wendy Hall**, Regius Professor of Computer Science and Director of the Web Science Institute, University of Southampton, UK
- **Wolfgang Lehner**, Professor of Computer Science, Chair of Databases, and Director of the Institute of Systems Architecture, Technische Universität Dresden, Germany
- **Natalia Manola**, CEO of OpenAIRE AMKE, Greece
- **Julia Stoyanovich**, Institute Professor of Computer Science & Engineering, Professor of Data Science, and Director of the Center for Responsible AI, New York University, USA
- **Matei Zaharia**, Professor of Computer Science, University of California - Berkeley, and CTO, Databricks, USA