

Defining Success in Open Science

A REPORT FROM THE LEADERS' FORUM, WASHINGTON DC,
OCTOBER 5 - 6, 2017

AUTHORS: SARAH E. ALI-KHAN^{1,2}, ANTOINE JEAN¹, EMILY MACDONALD¹
AND E. RICHARD GOLD^{1,3}

1. CENTRE FOR INTELLECTUAL PROPERTY POLICY (CIPP), MCGILL UNIVERSITY
2. TANENBAUM OPEN SCIENCE INSTITUTE (TOSI), MCGILL UNIVERSITY
3. DEPARTMENT OF HUMAN GENETICS, MCGILL UNIVERSITY



This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Foreword

In October 2017, thought-leaders from around the globe gathered at an Open Science Leadership Forum in the Washington DC offices of the Bill and Melinda Gates Foundation to share their views on what successful Open Science (OS) looks like. Delegates from developed and developing nations, national governments, science agencies and funding bodies, philanthropy, researchers, patient organizations and the biotechnology, pharma and artificial intelligence (AI) industries discussed the outcomes that would rally them to invest in OS, as well as wider issues of policy and implementation.

We aim to capture the breadth of this unique conversation in two reports: this first report summarizes the OS success outcomes identified at the Leadership Forum. A second longer report aimed for early Spring 2018 will address the broader topics that emerged.

Once again, we extend our sincere thanks to everyone who attended the Leadership Forum for their contributions – the scope and enthusiasm of the discussions delighted us and far exceeded our expectations.

Context

The Leadership Forum was the first of a multi-step process to develop a ‘toolbox’ of practical and transparent indicators for assessing where and when OS models of collaboration best advance science, innovation and public benefit. This project was inspired by the recent adoption of a broad institution-wide OS policy at the Montreal Neurological Institute (the Neuro), Canada’s leading centre for neuroscience research and patient care, and one of the largest of its kind in North America. Applying this policy across all its 60-some research labs, the Neuro committed to principles of open access publication, open data and withdrew its support of patenting on any direct research outputs. Spurred by the opportunity presented by this ‘natural experiment’, McGill University created the Tanenbaum Open Science Institute (TOSI) to advance independent study of OS, and to share knowledge about its benefits, costs and mechanisms of action. Given this starting point, the focus of this project is OS in the life sciences, including disciplines and industries such as AI, that may benefit from OS collaborations in these areas. We anticipate that in coming years, our team or others may expand the indicator toolkit to assess OS practice across other scientific fields.

The next step in the process will occur on May 31 - June 1, 2018, when we will bring together experts in innovation measurement, bibliometrics, economics, sociology and other fields to translate the OS success outcomes identified at the Forum into rational and measurable indicators. Based on these conversations, we will draft an indicator toolbox over the remainder of 2018, consisting of a ‘codebook’ of indicators, their definitions, sources, qualitative methods, and associated guidance, and distribute this to stakeholders for comment. In 2019, we anticipate distributing the resulting codebook to global partners and the general public. Working with partners, we will begin to collect, analyze and openly disseminate the resulting data. Throughout this process, we invite those who attended the Forum to provide feedback on the reports and the indicators, and perhaps participate in the OS measurement and assessment activities.

This report – and the downstream development of the indicator toolbox and codebook – is funded and supported by partners with a shared interest in advancing OS: the Bill and Melinda Gates Foundation, the Wellcome Trust, the UK Government Department for Business, Energy and Industrial Strategy, and TOSI.

Introduction

Collaborations offer the potential to not only advance basic research, but to lead to the development of new products and services on the market. In recent years, public research organizations, industry and clients, with the backing, financial support and strategic assistance of governments and philanthropy have experimented with a variety of collaborative structures.

Decreasing innovation rates and the rising costs of research and development (R&D) have increasingly led stakeholders toward open models of collaboration: ‘open science’ partnerships that rely on all or some of the pillars of open access to publications, wide sharing of data and other research outputs, and eschewing intellectual property rights (IPRs).

Theory, anecdote and early data predict that OS will accelerate discovery and innovation, maximize the value of scientific investment and bring expanding social and economic benefits. Indeed, OS has already gained significant momentum through the support of some governments, politicians and philanthropies. Yet these early supporters are working against the status quo, including entrenched business models, research culture and academic research incentives. Optimizing the outcomes of OS crucially depends on broad community adoption, which in turn depends on policy that supports and provides incentives for open practice. While some policy-makers are actively engaged in actualizing OS, the majority hesitate to enact the needed fundamental structural and cultural change in the absence of evidence. To address this legitimate concern, we are developing a toolkit upon which to build an evidence base of the benefits (and the costs) of OS, so that decision-makers in the public, private and social sectors can systematically create the conditions for success, maximize social value and spur a global transformation in how public and private partners conduct science and innovation.

In the following section, we present the hoped-for outcomes of successfully implemented OS proposed by Leadership Forum participants. These are organized by theme, each comprising a brief summary of the relevant discussion. We list the corresponding success outcomes, including scientific, clinical, social and economic factors in Appendix 1, organized by timeline – short to medium or long-term – in which we expect them to manifest.

In this document, we list the success outcomes that delegates highlighted in the Leadership Forum. Not all delegates agreed to every outcome; our goal was, instead, to capture the variety of outcomes sought rather than to reach consensus. We recognize that many, if not most, of the outcomes result from the complex interaction of factors, including OS, and therefore success or failure cannot be attributed solely to OS. However, statistical analysis and case studies will help to reveal the role that OS plays in attaining these outcomes. Nevertheless, we expect that some of them will be only aspirational and will not be measurable within a reasonable timeframe, while others are too complex or entangled

with other phenomena to be measured separately. Despite this, we believe it is useful to list all outcomes, regardless of their tractability, to mark current thinking about OS and in the hopes that others will find ways to assess them in the future.

As indicated above, in the next phase of the project, we will ‘translate’ these success factors into indicators, survey scripts, case study guides and other assessment tools to identify the role of OS in contributing to each outcome. To do this, we will differentiate between success factors that act as controls (e.g., attitudes toward OS, implementation of OS, etc.), independent variables (e.g., investments in open science, access to data sharing infrastructure, etc.), and dependent variables (e.g., diversity of publications, reaching milestones along the route to introducing new products and services). While some success factors can be quantitatively measured (e.g., number of publications, number of students trained, survey results of OS attitudes), others can only be assessed qualitatively (e.g., how a partnership was created, difficulties overcome, lessons learned, etc.).

Taken together, the success outcomes listed in this document cast a wide net: at the May-June 2018 meeting, we will begin the work of translating them into indicators, and integrating these into a conceptual framework that rationalizes inputs, actors, activities, outputs, impacts and the links between them.

Success outcomes identified at the Leadership Forum

1. Increased quality and efficiency of scientific outputs

Many participants believe that OS will curb the considerable amount of waste within biomedical research and development (R&D), lowering otherwise rising costs and providing a better return on investment than presently exists. In particular, funders and philanthropy expect that OS will lead to increased reuse of data and fewer ‘throwaway’ datasets that, once used by the data generator are afterwards virtually inaccessible. Biotech and Pharma cited the ‘reproducibility crisis’, noting an urgent need for more reliable academic outputs that can be used without lengthy in-house validation. They expect that OS will build open detailed knowledge of the basic biology and biochemistry of drug targets and pathways – outcomes that point to important downstream success factors, including the ability to rapidly select the most promising drug targets, to identify failure earlier in the innovation process and to reduce costly late stage failures – all contributing to lower attrition rates within the R&D pipeline. Additionally, Biotech expects that OS will lead innovation actors to concentrate their efforts where they can excel, reducing the redundancy of roles and activities: this should further contribute to the efficiency gains that many stakeholders anticipate from OS. Most participants across sectors expect OS to generate more, and more diverse, high quality datasets, together with the meta-data necessary to use them (including descriptions of methods, reagents, protocols and workflows, the instruments or platforms used in their generation, how and why data were collected).

2. Accelerated innovation and impact

Participants across sectors underlined innovation and public health-related factors as key outcomes of success. They expect OS to lead not only to faster innovation, but to deliver truly novel products and services that address unmet needs and bring measurable benefits to communities:

“discovering tomorrow’s medicines, today”. Thus, government, philanthropy and national funding agencies expect improved health outcomes across their populations. Biotech and Pharma anticipate that OS will give rise to a greater diversity of research, penetrating research ‘white space’ and seeding novel research domains, including new interdisciplinary fields. Taken together, many participants expect OS to generate faster development of knowledge and its translation into products and services with marked social value.

3. Increased trust in and accountability of the research enterprise

Many participants agreed that increased trust is a key success outcome, and one that is more likely to be achieved through greater openness than by other means. Many expect that OS will augment transparency, and consequently instill greater accountability across the entire research process, including at the level of reporting on the use of public research funds and the resulting public benefits. Governments expect this information to foster public trust in the research enterprise, and greater appreciation, understanding and support of science. They anticipate that this will result in increased research participation, public funding of science and private donations. In parallel with heightened transparency, many participants said that a successful implementation of OS would require new and improved mechanisms to explain research to communities and transparent governance and technical mechanisms to ensure the security of sensitive or confidential personal data while facilitating legitimate and beneficial uses.

On their side, Biotech and Pharma expect OS to augment trust between innovation actors, leading to streamlined partnering and collaboration. Again, they cited the poor reproducibility of academic outputs and a tendency for universities to compete rather than cooperate, for example by over-valuing their IP or other research outputs, which can slow partnership or knowledge transfer negotiations and fuel industry skepticism. Here again, these players believed that OS could lead to a paradigm shift toward cross-sectoral complementarity and collaboration.

4. Increased equity in research

Participants highlighted increased equity as another key success factor that is most likely to be achievable by way of greater openness. They discussed equity and inclusiveness at the level of i) participation and individual agency in the research process; and ii) access to research outputs and benefits.

First, many expect that OS will foster democratization of the research enterprise, resulting in a greater diversity of people meaningfully involved and gleaning value from the process. Nevertheless, many delegates, including those from developing countries, noted that, to achieve these results, countries with the assistance of funders will first need to develop a sharing infrastructure that includes high-speed internet, as well as local research infrastructure and a critical mass of trained researchers. Bearing this in mind, many delegates believe that OS will lead to increased diversity of research leaders, collaborators and participants, including across communities of colour, gender, ethnicity and socio-economic group. Patient organizations expect OS to result in greater involvement of end-users and communities in the research process – for example, leading studies, framing research questions, making funding decisions

and determining the outputs of value – and that there will be more funding available to ensure that these outputs are accessible to communities that participated in the studies. Many researchers expect OS partnerships to draw more clinical trials to OS research centres, augmenting local patient access to innovative therapeutics. Taken together, many delegates believe that OS will deliver more, and more immediate benefits, to communities from the research process.

Patient organizations also believe that greater openness will shift access and sharing decision-making to the individual donors, rather than researchers. Thus, OS will lead to new mechanisms to keep participants better informed about use of their materials and establish the individual as the ‘unit of openness’. Many participants expect OS will encourage other de-centralizations of power, including increasing collaborations in which developing country actors are equal partners or drivers of the research. In this regard, delegates underlined the need to avoid repeating historical power inequities whereby the benefits of some joint projects were coopted by the researchers from the more powerful or developed settings.

Second, by increasing access to knowledge, many participants expect OS to pave the way to increased scientific capacity in lower income, marginalized and developing communities. Again, they believe that OS will result in broader distribution of benefits, stimulating the development of research and sharing infrastructure, training, jobs and funding opportunities in lower income settings, and increasing retention of highly trained individuals in their local communities. Taken together, many participants agreed that heightened equity and inclusiveness through OS will bolster solidarity and justice, leading to greater empowerment of individuals and communities globally, and more opportunities to participate in the science innovation system, to create impact and to improve local health and well-being.

5. Better opportunities and recognition of early career researchers and youth

Many participants believe that greater openness will lead to development of new high-value jobs, and better and more diverse opportunities for students, post-docs and the next generation to launch their careers. For example, some of the new roles they foresee include novel positions and pathways in academia around data management, including data scientists, curators and stewards. They also expect OS to decrease barriers to students moving between academia and industry, by increasing collaboration and knowledge flow between the two settings. Many participants strongly underlined that at the very least, OS would not disadvantage early career researchers and youth who are considering entering the sciences. However, several noted that fears about the consequences for post-docs establishing their labs are a key reason that stakeholders may hesitate to embrace OS.

6. Positive economic impact

Many participants expect that OS will lead to equitable and positive economic impact. First, governments expect OS to prompt the private sector and venture capital (VC) to invest in research, where otherwise they would not. By augmenting OS public-private collaboration, OS is expected to increase the resources available to universities both through access to industry infrastructure and knowledge, and through additional funding. In particular, many participants believe that OS will lead to economic development in the communities housing OS research centres: in order for firms to take full advantage of

the expertise, know-how and relationships embodied in local researchers and infrastructure, they will be obliged to set up in the environs, bringing jobs and investment. Thus, participants anticipate that OS will catalyze the development of vibrant local ecosystems, make launching start-up firms easier, and create more skilled jobs, and more jobs overall, at all levels. Industry and philanthropy further expect that OS will lead to the creation of new business models, including for VC and investment.

7. Implementation Success

The Leadership Forum discussions often turned toward what will be needed for a successful implementation of OS. Many times, participants across the spectrum of sectors present stated that OS will require a paradigm shift in scientific research culture in order to realize its full potential. At the same time, they said that such a transformation would, to a large degree, show that execution of OS was well underway.

As noted in previous sections, Biotech, Pharma and governments said that successful OS implementation will result in better definition of the activities and roles of the various actors within the innovation system, including their specific responsibilities in the integration of OS.

Most researchers and industry believe that OS implementation will lead to an attitudinal shift amongst researchers in favour of sharing data and collaboration: individual researchers will come to view their outputs as part of a broader initiative to build a discovery platform for the benefit for all, rather than as belonging to them. Thus, successful implementation of OS will be characterized by researchers freely sharing data, publishing by default in open access journals and avoiding the use of restrictive IPRs: open practice will be fully integrated into every-day workflow by research institutions, governments and philanthropy. To aid in this, there will be many new resources available to practitioners, including training in how to conduct open practice and manage data, and tools, such as model workflows, sharing protocols and templates.

Governments and philanthropy noted that success will be reflected in the availability of long-term and sustainable funding to support OS infrastructure and more trusted open repositories for housing research outputs.

Many participants expect there to be an increasing number of data professionals, including scientists, curators and stewards, to ensure that data are managed and put to their best use. Likewise, many said that tracking of scientific outputs by DOI or other means will become standard. In parallel, participants agreed that a critical indicator of successful OS implementation will be the recognition of a broader range of outputs as publishable material by journals, funders and institutions – including reproducibility studies, datasets, policy publications, clinical guidelines, etc. – and the assigning of value to these in promotion, tenure and funding processes.

Next Steps

This Report serves as the basis for discussions on May 31 - June 1, 2018 in the Wellcome Trust's London offices to start work on translating success outcomes into indicators. We anticipate that these

indicators will include both quantitative and qualitative measures. We anticipate that this work will lead to the development of a toolkit consisting of indicators, a code-book of how to assess them, survey templates, and qualitative methods that we anticipate disseminating in 2019.

Appendix 1

We list here the success outcomes extracted from discussions among delegates on October 5-6, 2017 in Washington, DC. As noted in the Report, we opted to list a broad selection of outcomes without regard to whether they can, in practice, be measured or separately assessed. Our goal is to capture the breadth of hoped-for outcomes of OS rather than to limit discussion to only those outcomes that we know, in advance, we will be able to measure. We also note that not all delegates at the Leadership Forum agreed to all of these outcomes. Rather than being a consensus list, this represents the variety of outcomes we heard.

Success factors

1. Increased quality and efficiency of scientific outputs

Short to medium term success outcomes:

- a. Greater diversity and number of accessible, useable, and inter-operable datasets with detailed meta-data
- b. Standard and consistent data management and curation of datasets
- c. Greater access to negative and null results
- d. Greater use, re-use and re-combination of datasets
- e. Access to non-curated data (but with meta-data) to support AI approaches

Long-term success outcomes:

- f. Reduced quantity of research on the same targets or drugs
- g. Increased reliability and reproducibility of scientific outputs
- h. Reduced rates of scientific misconduct and retractions
- i. Reduced attrition rates in the R&D pipeline
- j. Reduced redundancy of roles and activities amongst players along the innovation pathway

2. Accelerated innovation and impact

Short to medium term success outcomes:

- a. Faster generation and greater translation of valuable knowledge to policy and practice
- b. Increased access to meaningful, understandable and useable health information
- c. Increasingly diverse research questions, penetration of research 'white space'
- d. More local clinical trials

Long-term success outcomes:

- e. Faster innovation to treatments, products and services of measurable public health impact
- f. Faster introduction of new clinical interventions
- g. Increased diversity of clinical intervention options
- h. Emergence of new inter-disciplinary fields

3. Increased trust in and accountability of the research enterprise

Short to medium term success outcomes:

- a. Greater transparency across the research process
- b. Improved data management practices, including marketing and stewardship
- c. The establishment of effective and transparent mechanisms to govern data access
- d. More effective oversight of research by government and ethics committees
- e. Greater public trust, appreciation and understanding of science and the research process
- f. Greater mutual trust among innovation actors

Long-term success outcomes:

- g. More and faster establishment of partnerships and collaborations with better outcomes
- h. Greater diversity of partners within partnerships and collaborations
- i. Greater public support of science

4. Increased equity in research

Short to medium term success outcomes:

- a. Greater diversity – including across colour, gender, ethnicity, socio-economic group etc. – participating in research including as participants, collaborators and leaders
- b. More international collaborations to develop infrastructure to support data-sharing and re-use
- c. Increased data use outside of existing networks, including in developing countries
- d. More funding and personnel to assist communities in accessing and making best use of research outputs
- e. More developing country researchers, governments and NGOs involved in research and as equal partners in research collaborations
- f. The development of mechanisms to enable donors to make access and sharing decisions, and keep them better informed of how their materials may be used

Long-term success outcomes:

- g. Increased development of knowledge and innovations of real value to developing country, low income and marginalized populations
- h. Greater scientific capacity in developing countries, including improved research infrastructures, training, jobs and funding opportunities
- i. Greater retention of highly trained individuals in their home countries or communities

5. Better opportunities and recognition of early career researchers and youth

Short to medium term success outcomes:

- a. Neutral or better and more diverse opportunities for students, post-docs and young researchers to launch their careers
- b. Reduced barriers to graduate students moving between academia and industry

6. Positive economic impact

Short to medium term success outcomes:

- a. Increased resources available to universities through access and collaborations with industry
- b. Augmented private and VC R&D investment locally
- c. Increased attraction and embedding of firms in local communities
- d. Greater ease in launching start-ups

Long-term success outcomes:

- e. Development of new business models to support OS and maximize public benefit
- f. Creation of more skilled jobs and of new types of high value jobs at all levels
- g. Increased local/regional private sector investment in R&D

7. Implementation Success

Short to medium term success outcomes:

- a. More researchers engage in sharing activities such as publishing open access, freely sharing their data, and avoiding restrictive IP
- b. Journals, funders and public research institutions recognize a wider range of scientific outputs as publishable material, and credit these in promotion, tenure and funding decisions
- c. Greater levels of education and sharing of best practices about OS
- d. More effective and comprehensive tracking of the range of scientific outputs by DOI or other means
- e. Greater availability of OS resources and training for researchers, including how to conduct open practice and to manage data, model workflows, templates, protocols, etc.
- f. Incorporation of OS into standard researcher workflows by research institutions, governments and philanthropy
- g. Increased use of open processes and tools across the entire research workflows
- h. More trusted repositories
- i. Increased numbers of data professionals, including scientists, curators and stewards

Long-term success outcomes:

- j. Greater clarification by actors of their roles and responsibilities in implementing OS
- k. An attitudinal shift among researchers in favour of sharing all research outputs
- l. Long-term and sustainable funding available for infrastructure to support OS

Appendix 2

Leadership Forum Participant List

Dr. Sarah Ali-Khan, Research Associate, Faculty of Law, McGill University

Dr. Lluís Ballell-Pages, Director, External Opportunities, GlaxoSmithKline, Tres Cantos Open Lab for Diseases of the Developing World

Dr. Patricia Brennan, Director, National Library of Medicine, Interim Associate Director for Data Science, National Institutes of Health, US Department of Health and Human Services

Dr. Katja Brose, Science Program Officer, Chan Zuckerberg Science Initiative

Ms. Rachel Bruce, Head of Open Science, Department for Business, Energy & Industrial Strategy, UK Government

Mr. David Carr, Program Manager - Open Research, Wellcome Trust

Dr. Simon Chaplin, Director of Culture and Society, Wellcome Trust

Me Mylène Deschênes, Ethics and Legal Advisor to Chief Scientist of Quebec, Fonds de recherche du Québec

Dr. Aled Edwards, Chief Executive Officer, Structural Genomics Consortium

Ms. Ashley Farley, Associate Officer of Knowledge & Research Services, Bill & Melinda Gates Foundation

Dr. Richard Gold, James McGill Professor, Faculty of Law, McGill University

Ms. Jennifer Hansen, Senior Officer, Knowledge & Research, Bill & Melinda Gates Foundation

Dr. Jason Karamchandani, Associate Professor, Department of Pathology McGill University

Dr. Michael Hawrylycz, Investigator, Allen Institute for Brain Science

Dr. Nadia Khelef, Senior Advisor for Global Affairs, Institut Pasteur

Mr. Robert J Kiley, Head of Open Research Development, Wellcome Trust

Ms. Elizabeth Kittrie, Strategic Advisor for Data and Open Science, US National Institutes of Health

Mr. Manoj Kumar, Head of Entrepreneurship and Innovations, Tata Trusts

Dr. Alexandre Le Bouthillier, Founder and COO, Imagia

Dr. Matthew Lucas, Executive Director, Corporate Strategy and Performance, Social Sciences and Humanities Research Council of Canada

Dr. Thomas Maina Kariuki, Director of the Alliance for Accelerating Excellence in Science in Africa

Dr. Lara Mangravite, President, Sage Bionetworks

Ms. Jessica Mankowski, Manager, Knowledge Translation Strategies, Canadian Institutes of Health Research

Dr. Sanjay Mehendale, Additional Director General, Indian Council of Medical Research

Dr. Mark Namchuk, SVP Research, Non-Clinical and Pharmaceutical Development, Alkermes

Ms. Thea Norman, Senior Program Officer, Bill & Melinda Gates Foundation

Mr. James O'Leary, Chief Innovation Officer, Genetic Alliance

Dr. Sébastien Paquet, Lead Applied Research Scientist & Culture Hacker, Element AI

Mr. Ben Pierson, Senior Program Officer, Bill & Melinda Gates Foundation

Dr. Claude Pirmez, Senior Researcher, Oswald Cruz Institute

Ms. Casey Selwyn, Fellow, Global Health Program, Bill & Melinda Gates Foundation

Ms. Annabel Seyller, Chief Operating Officer of the Open Science Experiment, The Montreal Neurological Institute

Dr. Carthage Smith, Senior Policy Analyst, Organization for Economic Co-Operation and Development (OECD)

Dr. Jeff Spies, Chief Technology Officer and Co-founder, Centre for Open Science

Dr. David Sweeney, Executive Chair Designate of Research England and Director, Research and Knowledge Exchange Higher Education Council for England

Dr. Michiel van Den Hauten, Head/Deputy Director of Research and Science Policy at the Ministry of Education, Culture and Science

Dr. Kate Williams, Scientific Director, Krembil Foundation