Briefing Memo on Open Science

To: Office of Health Information Programs Development
Office of the Director
National Library of Medicine

From: Kendra Godwin
NLM Associate Fellow

Date: August 25, 2017

Subject: Assessing the Scope of Open Science:
Meaning, Benefits, and Applications for the Biomedical Enterprise

Purpose

This project emerged in response to the strategic planning process of the National Library of Medicine (NLM). A key part of planning for the next ten years concerns how NLM will advance data science, open science, and biomedical informatics. Neither NLM nor the National Institutes of Health (NIH) have open science definitions. This document isn't meant to be comprehensive, but rather to serve as an introductory framework for exploring the meaning, benefits, and applications of open science for the biomedical enterprise. The goal of this project is to develop a set of resources to inform discussion on how NLM could conceptualize and communicate open science, and provide fresh ideas for initiatives that could be undertaken to encourage its advancement.

Methodology

This was an exploration of the open science ecosystem. Methods included 1) conducting a literature review to gather a set of open science definitions and exemplars of open science benefits [see Appendix A]; 2) interviewing open science thought leaders to build summaries of definitions, key points, and recommendations for advancing open science at NLM [see Appendix B]; and 3) compiling resource lists of publications, websites, and visualizations related to open science [see Appendix C, D, and E].

The example set of definitions is representative of seven different governments, organizations, projects, and communities, to show how different institutions that are committed to furthering open science formally define the term. The exemplars are a selection to show how open science advances science and benefits the health of humankind in ways that would not be possible if processes were closed, and includes different areas to illustrate how open science operates for different groups in different disciplines to produce different product solutions. The interviews were with eleven individuals, a mix of genders, ages, and locations, and each were selected for a different perspective on open science. These are thought leaders in the field of open science, many dedicating their recent careers to its exploration and promotion, as well as open policy, data science, data, and librarianship. Roles of those interviewed

include director, executive, policy influencer, computer scientist, researcher, professor, and librarian, with some of these distinctions overlapping. The three resource lists present different ways of understanding and conceptualizing open science, and provide avenues for further learning.

Findings

Rather than being a self-encompassing domain of science, open science is an umbrella term, with many elements branching out from beneath it. Open science is, but is also far more than, the sharing of research data. Much can be missed by equating a lesser term for all of open science itself.

Open science isn't a well-defined term with a singular definition. Some argue that open science is simply science as it was originally practiced at the dawn of the scientific journal, that this term is needed to differentiate science from the closed practice it's become. Others have an entirely different concept of it, as demonstrated in collected definitions. What emerged from my research and conversations was variation, and that the meaning of open science is closely linked to the context in which the term is being used.

The problem with a singular definition for open science is that there isn't one, and to create one is to imply that there's a singular solution to advancing it. Instead of compiling a single definition, key components can be teased out to form a guide for identification via fundamental characteristics. Open science, as seen in the definitions collected, involves the following four characteristics: 1) collaboration with people at different levels and in differing fields; 2) a process rooted in and relying on digital technologies; 3) science conducted in a way that will allow for sharing and reuse; and 4) involvement with any or all parts of the research life cycle.

Conclusion

Open science is a rapidly evolving field, and when exploring its complexities, it's important to set guideposts to enable a shared understanding of its meaning, benefits, and applications. To help advance NLM's understanding of this dynamic space, my project focused on three areas of open science: characteristics, examples, and recommendations. It is my hope that the open science characteristics exposed through this research project can help to frame meaning; that the exemplars identified can help to communicate the tangible benefits of practicing open science; and that the recommendations can help to suggest new application areas for evolving open science. When combined, it becomes a collective narrative for advancing the compelling story of open science and provides a multi-tiered approach that can be embraced by NLM.

To communicate the concept of open science and its role in enabling the biomedical research enterprise, NLM may find it helpful in the future to create a definitional framework for discussing open science. In an attempt at synthesizing this research, here is my definitional framework as a single sentence:

"Open science is a concept to encompass the activities of the research lifecycle, wherein science is purposefully conducted with digital technologies and in collaboration with others, a combination that both allows for and facilitates the intentional sharing and reuse of all generated products."

Kendra Godwin, NLM Associate Fellow

Project Sponsors

- Elizabeth Kittrie, Strategic Advisor for Data and Open Science
- Mike Huerta, Associate Director for Program Development and NLM Coordinator of Data and Open Science Initiatives
- Jerry Sheehan, Deputy Director

Appendices

Appendix A: Open Science Literature Review

Appendix A.1: Open Science Definitions

Selection Criteria: The example set of definitions is representative of seven different governments, organizations, projects, and communities, to show how different institutions that are committed to furthering open science formally define the term.

European Commission, European Union (EU)

Open innovation, open science, open to the world

- Source: http://dx.doi.org/10.2777/061652
- Definition: "Open Science represents a new approach to the scientific process based on cooperative work and new ways of diffusing knowledge by using digital technologies and new collaborative tools."

Organisation for Economic Cooperation and Development (OECD)

Open and Inclusive Collaboration in Science: A Framework

- Source: Global Science Forum. Draft for discussion. OECD Headquarters, 13-14 October 2016.
- Definition: "Open science in its broadest sense refers to efforts to make the scientific process more open and inclusive to all relevant actors, within and beyond the scientific community, as enabled by digitalization."

Making Open Science a Reality

- Source: http://dx.doi.org/10.1787/5jrs2f963zs1-en
- Definition: "Open science commonly refers to efforts to make the output of publicly funded research more widely accessible in digital format to the scientific community, the business sector, or society more generally."
- Definition: "Open science is the encounter between the age-old tradition of openness in science and the tools of information and communications technologies (ICTs) that have reshaped the scientific enterprise and require a critical look from policy makers seeking to promote long-term research as well as innovation."

Facilitate Open Science Training for European Research (FOSTER)

"Open Science Taxonomy"

- Source: https://www.fosteropenscience.eu/foster-taxonomy/open-science
- Definition: "Open science is the movement to make scientific research, data and dissemination accessible to all levels of an inquiring society."
- Source: https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition
- Definition: "Open Science is the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods."

F1000Research

Guide to Open Science Publishing

- Source: https://blog.f1000.com/2015/03/06/guide-to-open-science-publishing/

- Definition: “Open science is the concept of opening up all aspects of scientific research, to allow others to follow the process and collaborate. There is no formal definition of open science, but it usually incorporates some of the aspects we looked at before, such as open access, open peer review, post-publication peer review, and open data. Additionally, it includes other ways to make science more transparent and accessible during the research process, and we will discuss them here: open notebook science, citizen science, and aspects of open source software and crowdfunded research projects.”

Avoin tiede ja tutkimus (ATT) -hanke / Open Science and Research Initiative, Ministry of Education and Culture, Finland

The Open Science and Research Handbook

- Source: http://openscience.fi/handbook
- Definition: “Open science means the promotion of an open operating model in scientific research. The key objective is to publish research results, along with the data and methods used, so they can be examined and used by any interested party. Open science includes practices such as promoting open access publishing, open access publishing itself, harnessing open-source software and open standards, and the public documentation of research processes with ‘memoing.'”

Science & Technology Section, Association of College & Research Libraries

“Open Science and Crowd Science: Selected Sites and Resources”

- Source: http://www.istl.org/12-spring/internet2.html
- Definition: “Open science is a broad concept that includes these closely related areas of open notebook science and open data. Advocates of open science believe that there should be no insider information, and all protocols and results -- even those of failed experiments -- should be made visible and open to reuse as soon as possible in open lab notebooks and data repositories.”

Centre National de la Recherche Scientifique (CNRS) / National Center for Scientific Research

Open Science in a Digital Republic

- Source: http://dx.doi.org/10.4000/books.oep.1635
- Definition: “Open Science is a new horizontal approach to access to [sic] scientific work and objectives, and to sharing of scientific results, as well as a new way of doing science, by opening up its processes, codes and methods.”
- Definition: “Open Science is therefore a change in perspective that can be compared with other earlier major stages, such as the advent of the telescope or the microscope.”
- Definition: “Open Science, a field that is far wider than open access, which is limited solely to publication, refers to all the different ways and means of enhancing scientific work offered by digital technologies.”

Appendix A.2: Exemplars of Open Science Benefits

Selection Criteria: The exemplars are a selection to show how open science advances science and benefits the health of humankind in ways that would not be possible if processes were closed, and includes different areas to illustrate how open science operates for different groups in different disciplines to produce different product solutions.

i. Category: Open Science Tools

Example: Nextstrain

- Source: http://www.nextstrain.org/
- Synopsis from Source: "Nextstrain is an open-source project to harness the scientific and public health potential of pathogen genome data. We provide a continually-updated view of publicly available data with powerful analytics and visualizations showing pathogen evolution and epidemic spread. Our goal is to aid epidemiological understanding and improve outbreak response."

Example: "Ebola, Zika modelers aim to inform policy decisions"

- Source: https://www.fic.nih.gov/News/GlobalHealthMatters/march-april-2016/Pages/disease-modeling-informs-health-policy.aspx
- Synopsis from Source: "Mathematical modelers face numerous challenges as they try to predict the course of epidemics, such as Ebola and Zika. Their forecasts can inform policymakers' decisions on how to most effectively deploy resources to contain and manage the outbreaks."

Example: Insight Segmentation and Registration Toolkit (ITK)

- Source: https://itk.org/
- Synopsis from Source: "ITK is an open-source, cross-platform system that provides developers with an extensive suite of software tools for image analysis. Developed through extreme programming methodologies, ITK employs leading-edge algorithms for registering and segmenting multidimensional data."

Example: "Ion-Abrasion Scanning Electron Microscopy Reveals Surface-Connected Tubular Conduits in HIV-Infected Macrophages"

- Source: http://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1000591
- Synopsis from Source: "We have used a new technique called ion abrasion scanning electron microscopy (IA-SEM) to image entire HIV-infected human macrophages at a resolution high enough to see individual HIV virions and their location within the cell. This approach revealed that HIV is present in a system of nanoscale tubes, barely larger than a virus at some places, which connect internal viral reservoirs to the cell surface. These tubes could allow the macrophage to deliver HIV virions to bystander cells from its continually replenished stores of ammunition, held deep within the cell. Our work provides a glimpse of how the structure of these reservoirs allows macrophages to accomplish viral delivery. Discovery of these virion-channeling tubes provides a potential drug target to address the problem of persistent HIV infection."

Example: NIH 3D Print Exchange

- Source: https://3dprint.nih.gov/
- Synopsis from Source: "3D printing technology is advancing at a rapid pace, but it is difficult to find or create 3D-printable models that are scientifically accurate or medically applicable. The NIH 3D Print Exchange provides models in formats that are readily compatible with 3D printers, and offers a unique set of tools to create and share 3D-printable models related to biomedical science."

Example: "3D-Printable Prosthetic Devices"

- Source: https://3dprint.nih.gov/collections/prosthetics
- Synopsis from Source: "3D-printable prosthetics are changing the face of medicine, as engineers and physicians are able to develop prosthetics that are fully customized to the wearer. Consumer 3D printing is leading to an even bigger revolution: 'DIY' assistive devices that can be printed by virtually anyone, anywhere."

Example: NCBI Hackathons

- Source: https://ncbi-hackathons.github.io/
- Synopsis from Source: "Perfoming [sic] advance bioinformatics analysis of next-generation sequencing data and metadata"

Example: "Closing gaps between open software and public data in a hackathon setting: User-centered software prototyping [version 2; referees: not peer reviewed]"

- Source: https://f1000research.com/articles/5-672/v2
- Synopsis from Source: "In genomics, bioinformatics and other areas of data science, gaps exist between extant public datasets and the open-source software tools built by the community to analyze similar data types. The purpose of biological data science hackathons is to assemble groups of genomics or bioinformatics professionals and software developers to rapidly prototype software to address these gaps."

Example: rOpenSci

- Source: https://ropensci.org/
- Synopsis from Source: "At rOpenSci we are creating packages that allow access to data repositories through the R statistical programming environment that is already a familiar part of the workflow of many scientists. Our tools not only facilitate drawing data into an environment where it can readily be manipulated, but also one in which those analyses and methods can be easily shared, replicated, and extended by other researchers."

Example: "Reproducibility in Science"

- Source: http://ropensci.github.io/reproducibility-guide/
- Synopsis from Source: "In every field of science, scientists are increasingly using electronic tools, from collecting their data to publishing their results. With the increase in computational tools comes the advantage of reproducibility at an extent not previously possible. Unfortunately, reproducibility of results is actually becoming increasingly more difficult, owing to the variety of ways of approaching analysis and incapability of data structures and file types. This is a guide to make scientific research more easily communicated and performed by using tools that promote reproducibility."

Example: CellMiner

- Source: https://discover.nci.nih.gov/cellminer/home.do
- Synopsis from Source: "CellMiner™ is a web application generated by the Genomics & Bioinformatics Group, LMP, CCR, NCI that facilitates systems biology through the retrieval and integration of the molecular and pharmacological data sets for the NCI-60 cell lines."

Example: "CellMiner: A Web-Based Suite of Genomic and Pharmacologic Tools to Explore Transcript and Drug Patterns in the NCI-60 Cell Line Set"

- Source: http://cancerres.aacrjournals.org/content/72/14/3499
- Synopsis from Source: "CellMiner greatly broadens applications of the extensive NCI-60 database for discovery by creating web-based processes that are rapid, flexible, and readily applied by users without bioinformatics expertise."

Example: Basic Local Alignment Search Tool (BLAST)

- Source: https://blast.ncbi.nlm.nih.gov/Blast.cgi
- Synopsis from Source: "The Basic Local Alignment Search Tool (BLAST) finds regions of local similarity between sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches. BLAST can be used to infer functional and evolutionary relationships between sequences as well as help identify members of gene families."

Example: "Basic Local Alignment Search Tool (BLAST)"

- Source: https://www.nature.com/scitable/topicpage/basic-local-alignment-search-tool-blast-29096
- Synopsis from Source: "In fact, the initial paper describing the program, published in the Journal of Molecular Biology and entitled "Basic Local Alignment Search Tool," was the most highly cited publication of the 1990s (Taubs, 2000). In recent years, the parallel development of large-scale sequencing projects and bioinformatic tools like BLAST has enabled scientists to study the genetic blueprint of life across many species, and it has also helped connect biology and computer science in the maturing field of bioinformatics."

ii. Category: Open Science Data Resources

Example: PubMed

- Source: https://www.ncbi.nlm.nih.gov/pubmed/
- Synopsis from Source: "PubMed comprises more than 27 million citations for biomedical literature from MEDLINE, life science journals, and online books."

Example: "Conceptual biology, hypothesis discovery, and text mining: Swanson's legacy."

- Source: https://www.ncbi.nlm.nih.gov/pubmed/16584552
- Synopsis from Source: "Swanson's vision of the hidden value in the literature of science and, by extension, in biomedical digital databases, is still remarkably generative for information scientists, biologists, and physicians."

Example: PubMed Central (PMC)

- Source: https://www.ncbi.nlm.nih.gov/pmc/
- Synopsis from Source: "PubMed Central® (PMC) is a free full-text archive of biomedical and life sciences journal literature at the U.S. National Institutes of Health's National Library of Medicine (NIH/NLM)."

Example: *Making Open Science a Reality*

- Source: http://www.oecd-ilibrary.org/science-and-technology/making-open-science-a-reality_5jrs2f963zs1-en

- Synopsis from Source: “As of 2014, PMC contained more than 3.2 million research articles, and has become a key repository searched regularly by researchers in academia and industry, educators, students, the general public – and major search engines, making deposited papers more visible and accessible. It has been estimated that on a typical weekday, more than 1 million different users download more than 2 million different articles from PMC.”

Example: Human Genome Project (HGP)

- Source: https://www.genome.gov/10001772/all-about-the--human-genome-project-hgp/
- Synopsis from Source: “The Human Genome Project (HGP) was the international, collaborative research program whose goal was the complete mapping and understanding of all the genes of human beings. All our genes together are known as our ‘genome.’”

Example: “The Human Genome Project and its importance in clinical medicine”

- Source: http://www.sciencedirect.com/science/article/pii/S0531513101005702
- Synopsis from Source: “In clinical medicine, the human genome gives important clues in the understanding of human diseases in terms of human biology and pathology. Medicine will be revolutionized in improving diagnosis, prognostic, treatments and prevention.”

Example: GenBank

- Source: https://www.ncbi.nlm.nih.gov/genbank/
- Synopsis from Source: “GenBank ® is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences.”

Example: “Genome data in the public domain: unleashed synergies”

- Source: http://www.columbia.edu/cu/21stC/issue-1.3/dna-genome.html
- Synopsis from Source: “’The database was key to cloning an Alzheimer's gene,’ asserts molecular biologist David Galas, who reported cloning a gene that when mutated caused a predisposition to a variant of early-onset Alzheimer's disease. ‘It would have taken us another year to get there’ without access to such a library, says Galas, chief scientific officer of Darwin Molecular in Bothell, Wash.”

Example: database of Genotypes and Phenotypes (dbGaP)

- Source: https://www.ncbi.nlm.nih.gov/gap
- Synopsis from Source: “The database of Genotypes and Phenotypes (dbGaP) was developed to archive and distribute the data and results from studies that have investigated the interaction of genotype and phenotype in Humans.”

Example: “Mind the dbGAP: The Application of Data Mining to Identify Biological Mechanisms”

- Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3086918/
- Synopsis from Source: “Here we discuss how, in our view, the availability of dbGAP data has changed the traditional scientific approach to the identification of the genetic contributors to human disease and traits. Further, dbGAP has created new opportunities to discover genes important for mammalian development and disease traits through the targeted analysis of coding variants and the application of pathway-based approaches.”

Example: NCI-60 Human Tumor Cell Lines Screen

- Source: https://dtp.cancer.gov/discovery_development/nci-60/

- Synopsis from Source: “The screen was implemented in fully operational form in 1990 and utilizes 60 different human tumor cell lines to identify and characterize novel compounds with growth inhibition or killing of tumor cell lines.”

Example: “NCI-60 Cell Line Screening: A Radical Departure in its Time.”

- Source: https://academic.oup.com/jnci/article-lookup/doi/10.1093/jnci/djv388
- Synopsis from Source: “In conclusion, the NCI 60 cell panel proved to be a most important step forward in cancer drug screening. The concept of incorporating the diversity of human cancers into a cell line screen set the stage for extraordinary advances in targeted drug development in the past two decades.”

Example: The Cancer Genome Atlas

- Source: https://cancergenome.nih.gov/
- Synopsis from Source: “The Cancer Genome Atlas (TCGA) is a collaboration between the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI) that has generated comprehensive, multi-dimensional maps of the key genomic changes in 33 types of cancer. The TCGA dataset, comprising more than two petabytes of genomic data, has been made publically available, and this genomic information helps the cancer research community to improve the prevention, diagnosis, and treatment of cancer.”

Example: “The Impact of the Cancer Genome Atlas on Lung Cancer”

- Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4656061/
- Synopsis from Source: “The consequences of this research has led to an improved understanding of the molecular mechanisms underlying LUAD and LUSC, the elucidation of the clinical implications of these molecular mechanisms, the development of new and improved tools and algorithms, and the discovery of potential new therapies.”

Example: Genomic Data Commons (GDC)

- Source: https://gdc.cancer.gov
- Synopsis from Source: “The GDC obtains validated datasets from NCI programs in which the strategies for tissue collection couples quantity with high quality. The GDC encourages data sharing in support of precision medicine. Tools are provided to guide data submissions by researchers and institutions.”

Example: “Newly launched Genomic Data Commons to facilitate data and clinical information sharing”

- Source: https://www.nih.gov/news-events/news-releases/newly-launched-genomic-data-commons-facilitate-data-clinical-information-sharing
- Synopsis from Source: “Together, TCGA and TARGET represent some of the largest and most comprehensive cancer genomics datasets in the world, comprising more than two petabytes of data (one petabyte is equivalent to 223,000 DVDs filled to capacity with data). In addition, the GDC will accept submissions of cancer genomic and clinical data from researchers around the world who wish to share their data broadly. In so doing, researchers will be able to use the state-of-the-art analytic methods of the GDC, allowing them to compare their findings with other data in the GDC.”

Example: Embryonic Stem Cells Atlas of Pluripotency Evidence (ESCAPE)

- Source: http://www.maayanlab.net/ESCAPE/
- Synopsis from Source: "The database provides web-based interactive search and visualization tools that can be used to build subnetworks and identify known and novel regulatory interactions across various regulatory layers, as well as predict the effects of combinatorial knockdowns."

Example: "ESCAPE: database for integrating high-content published data collected from human and mouse embryonic stem cells"

- Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3689438/
- Synopsis from Source: "Overall, such analyses can be used to link relevant phenotypes to specific regulatory mechanisms in embryonic stem cells, as well as help experimental stem cell biologists who perform high throughput experiments to place their results in context of prior studies."

Example: ClinicalTrials.gov

- Source: https://clinicaltrials.gov/
- Synopsis from Source: "ClinicalTrials.gov is a registry and results database of publicly and privately supported clinical studies of human participants conducted around the world."
- Selected Publications: https://clinicaltrials.gov/ct2/resources/pubs

Example: "Big Data Mining and Adverse Event Pattern Analysis in Clinical Drug Trials"

- Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5175440/
- Synopsis from Source: "Drug adverse events (AEs) are a major health threat to patients seeking medical treatment and a significant barrier in drug discovery and development. AEs are now required to be submitted during clinical trials and can be extracted from ClinicalTrials.gov (https://clinicaltrials.gov/), a database of clinical studies around the world. By extracting drug and AE information from ClinicalTrials.gov and structuring it into a database, drug-AEs could be established for future drug development and repositioning."

Example: Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC)

- Source: https://biolincc.nhlbi.nih.gov/home/
- Synopsis from Source: "The mission of BioLINCC is to facilitate access to, and maximize the scientific value of, the Biorepository and Data Repository and promote the availability and use of other NHLBI funded population-based biospecimen and data resources."

Example: "Providing researchers with online access to NHLBI biospecimen collections: The results of the first six years of the NHLBI BioLINCC program"

- Source: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0178141
- Synopsis from Source: "Program utilization metrics were developed to measure the impact of BioLINCC on Biorepository access by researchers, including visibility, program efficiency, user characteristics, scientific impact, and research types."

Example: International Stroke Trial (IST)

- Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3104487/

- Synopsis from Source: “We aimed to make individual patient data from the International Stroke Trial (IST), one of the largest randomised trials ever conducted in acute stroke, available for public use, to facilitate the planning of future trials and to permit additional secondary analyses.”

Example: “The International Stroke Trial (IST): a randomised trial of aspirin, subcutaneous heparin, both, or neither among 19435 patients with acute ischaemic stroke. International Stroke Trial Collaborative Group.”

- Source: https://www.ncbi.nlm.nih.gov/pubmed/9174558
- Synopsis from Source: “Taking the IST together with the comparably large Chinese Acute Stroke Trial, aspirin produces a small but real reduction of about 10 deaths or recurrent strokes per 1000 during the first few weeks. Both trials suggest that aspirin should be started as soon as possible after the onset of ischaemic stroke; previous trials have already shown that continuation of low-dose aspirin gives protection in the longer term.”

Example: Yale University Open Data Access (YODA) Project

- Source: http://yoda.yale.edu/
- Synopsis from Source: “The Yale University Open Data Access (YODA) Project’s mission is to advocate for the responsible sharing of clinical research data, open science, and research transparency. The Project is committed to supporting research focused on improving the health of patients and informing science and public health.”

Example: “A Historic Moment for Open Science: The Yale University Open Data Access Project and Medtronic”

- Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5459387/
- Synopsis from Source: “The process that produced the reviews in this issue ensured that they were based on the totality of the evidence. That we have established a new standard should be a headline. Let us resolve to have this effort be the first step in the next era of cooperation among industry, academia, clinicians, and the public— one that rewards data sharing, promotes open science, and ultimately makes it untenable to obscure data relevant to the risks and benefits of approved medical products.”

Example: Encyclopedia of DNA Elements (ENCODE) Consortium

- Source: https://www.encodeproject.org/
- Synopsis from Source: “The goal of ENCODE is to build a comprehensive parts list of functional elements in the human genome, including elements that act at the protein and RNA levels, and regulatory elements that control cells and circumstances in which a gene is active.”

Example: “Deciphering ENCODE.”

- Source: https://www.ncbi.nlm.nih.gov/pubmed/26962025
- Synopsis from Source: “The ENCODE project represents a major leap from merely describing and comparing genomic sequences to surveying them for direct indicators of function. The astounding quantity of data produced by the ENCODE consortium can serve as a map to locate specific landmarks, guide hypothesis generation, and lead us to principles and mechanisms underlying genome biology.”

Example: Genotype-Tissue Expression Project (GTEx)

- Source: https://www.genome.gov/27543767/
- Synopsis from Source: "GTEx researchers are studying genes in different tissues obtained from many different people. Thus every donor's generous gift of tissues and medical information to the GTEx project makes possible research that will help improve our understanding of diseases, giving hope that we will find better ways to prevent, diagnose, treat and eventually cure these diseases in the future. In addition, the GTEx project includes a study to explore the effectiveness of the GTEx donor consent process."

Example: "The impact of structural variation on human gene expression"

- Source: http://www.biorxiv.org/content/early/2016/06/09/055962
- Synopsis from Source: "Structural variants (SVs) are an important source of human genetic diversity but their contribution to traits, disease, and gene regulation remains unclear. The Genotype-Tissue Expression (GTEx) project presents an unprecedented opportunity to address this question due to the availability of deep whole genome sequencing (WGS) and multi-tissue RNA-seq data from 147 individuals."

Example: UK Biobank

- Source: http://www.ukbiobank.ac.uk/
- Synopsis from Source: "UK Biobank is a national and international health resource with unparalleled research opportunities, open to all bona fide health researchers. UK Biobank aims to improve the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses – including cancer, heart diseases, stroke, diabetes, arthritis, osteoporosis, eye disorders, depression and forms of dementia."

Example: "UK Biobank: opportunities for cardiovascular research"

- Source: https://academic.oup.com/eurheartj/article/3836952/UK-Biobank-opportunities-for-cardiovascular
- Synopsis from Source: "Observational cohort studies have been essential in informing the prevention and treatment of cardiovascular diseases and identifying the role of cardiovascular risk factors in disease development. However, previous cohorts have either been too small to investigate less common diseases or lacked the depth of data to explore the complex interplay between different factors and cardiovascular disease risk. UK Biobank combines a large sample size of half a million participants with an unprecedented amount of phenotypic and genotypic data as well as ongoing linkage to health records."

Example: Protein Data Bank (PDB)

- Source: https://www.rcsb.org/pdb/home/home.do
- Synopsis from Source: "This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease."

Example: "Drug repurposing to target Ebola virus replication and virulence using structural systems pharmacology"

- Source: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-016-0941-9

- Synopsis from Source: “Here, the Indinavir-HIV complex was downloaded from Protein Data Bank (PDB id 2AVO) [58]. Interestingly, HIV protease, reveals that the same atoms (O2, N4 and O4) of Indinavir form the hydrogen bonds with residues; Ala28, Asp29, Asp25 in Chain A and Asp25 in Chain B of HIV protease. Consequently, the predicted similar binding pattern of VP24 and HIV protease to Indinavir suggest that this HIV protease inhibitor may be repurposed to target Ebola VP24.”

iii. Category: Open Science Drug Discovery

Example: Open Source Malaria

- Source: http://opensourcemalaria.org/
- Synopsis from Source: “The Open Source Malaria project is trying a different approach to curing malaria. Guided by open source principles, everything is open and anyone can contribute.”

Example: “Open Source Drug Discovery: Highly Potent Antimalarial Compounds Derived from the Tres Cantos Arylpyrroles”

- Source: http://pubs.acs.org/doi/full/10.1021/acscentsci.6b00086
- Synopsis from Source: “One of the unique features of this project, the open source research method, ensures that the unexplored lines of inquiry remain open alongside the attendant data posted online that makes it straightforward for others to resume any portion of the research project as fully fledged participants, with access to both positive and negative data, details of all procedures as they were carried out (to aid reproducibility), and anecdotal insight into project loose ends that are easy to explore.”

Example: Medicines for Malaria Venture (MMV)

- Source: https://www.mmv.org/
- Synopsis from Source: “Our mission is to reduce the burden of malaria in disease-endemic countries by discovering, developing and facilitating delivery of new, effective and affordable antimalarial drugs.”

Example: “Our Impact”

- Source: https://www.mmv.org/our-impact
- Synopsis from Source: “‘MMV has transformed the product landscape for malaria and created hope. If MMV didn’t exist, we would have to invent it.’”

Example: OpenZika

- Source: https://www.worldcommunitygrid.org/research/zika/overview.do
- Synopsis from Source: “The Zika virus has been linked to serious neurological conditions, including birth defects in children whose mothers were infected during pregnancy. An international team of researchers is using World Community Grid to search for a critically needed anti-viral drug to combat Zika, and they need your help.”

Example: “OpenZika: An IBM World Community Grid Project to Accelerate Zika Virus Drug Discovery”

- Source: http://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0005023
- Synopsis from Source: “This exemplifies open science, which should help scientists around the world as they address the long and arduous process of discovering and developing new drugs. Screening millions of compounds against many different protein models in this way would take

far more resources and time than any academic researcher could generally obtain or spend. As of August 16, 2016, we have submitted 894 million docking jobs."

Example: ClinicalStudyDataRequest.com (CSDR)

- Source: https://www.clinicalstudydatarequest.com/
- Synopsis from Source: "ClinicalStudyDataRequest.com (CSDR) is a consortium of clinical study data providers. It is a leader in the data sharing community inspired to drive scientific innovation and improve medical care by facilitating access to patient-level data from clinical studies."

Example: "Big Pharma Opens Up Its Big Data"

- Source: https://www.technologyreview.com/s/529046/big-pharma-opens-up-its-big-data/
- Synopsis from Source: "The site allows researchers, for the first time, to obtain data on the same topic from many companies simultaneously. 'We can combine all of the evidence from many studies to get one overall result,' says Sarah Nolan, a research assistant in medical statistics at the University of Liverpool. Nolan is working on a three-year research project designed to determine the clinical utility and cost-effectiveness of various epilepsy treatments."

iv. Category: Open Science Research Initiatives

Example: All of Us

- Source: https://allofus.nih.gov/
- Synopsis from Source: "The All of Us Research Program is a historic effort to gather data from one million or more people living in the United States to accelerate research and improve health. By taking into account individual differences in lifestyle, environment, and biology, researchers will uncover paths toward delivering precision medicine."

Example: "Data Collaboratives: Exchanging Data to Improve People's Lives"

- Source: https://medium.com/@sverhulst/data-collaboratives-exchanging-data-to-improve-people-s-lives-d0fcfc1bdd9a
- Synopsis from Source: "The term data collaborative refers to a new form of collaboration, beyond the public-private partnership model, in which participants from different sectors — including private companies, research institutions, and government agencies — can exchange data to help solve public problems. In the coming months and years, data collaboratives will be essential vehicles for harnessing the vast stores of privately held data toward the public good."

Example: Open Science Prize

- Source: https://www.opensciencеprize.org/
- Synopsis from Source: "The Open Science Prize is a new initiative from the Wellcome Trust, US National Institutes of Health and Howard Hughes Medical Institute to encourage and support the prototyping and development of services, tools and/or platforms that enable open content – including publications, datasets, code and other research outputs – to be discovered, accessed and re-used in ways that will advance research, spark innovation and generate new societal benefits."
- Six Finalist Proposals:
    - Open AQ: Real-Time Air Quality Data
        - Source: https://www.opensciencеprize.org/res/p/finalists/#1

        - Synopsis from Source: “Providing real-time information on poor air quality by combining data from across the globe”
    - Real-Time Evolutionary Tracking for Pathogen Surveillance and Epidemiological Investigation
        - Source: https://www.opensciencеprize.org/res/p/finalists/#2
        - Synopsis from Source: “Permitting analysis of emerging epidemics such as Ebola, MERS-CoV and Zika”
    - Open Neuroimaging Laboratory
        - Source: https://www.opensciencеprize.org/res/p/finalists/#3
        - Synopsis from Source: “Advancing brain research by enabling collaborative annotation, discovery and analysis of brain imaging data”
    - OpenTrialsFDA
        - Source: https://www.opensciencеprize.org/res/p/finalists/#4
        - Synopsis from Source: “Enabling better access to drug approval packages submitted to and made available by the Food and Drug Administration”
    - Fruit Fly Brain Observatory
        - Source: https://www.opensciencеprize.org/res/p/finalists/#5
        - Synopsis from Source: “Allowing researchers to better conduct modeling of mental and neurological diseases by connecting data related to the fly brain”
    - MyGene2: Accelerating Gene Discovery with Radically Open Data Sharing
        - Source: https://www.opensciencеprize.org/res/p/finalists/#6
        - Synopsis from Source: “Facilitating the public sharing of health and genetic data through integration with publicly available information”

Example: “International Prize Winners Demonstrate the Future of Open Science”

- Source: https://www.hhs.gov/idealab/2016/05/10/international-prize-winners-demonstrate-future-open-science/
- Synopsis from Source: “The fascinating stories of how these partnerships formed and what they have proposed provide important and inspiring examples of the power of open science. These stories demonstrate how open science can enable interdisciplinary teams from across the globe to work together to creatively advance public health and biomedical research.”

Example: ASAPbio

- Source: http://asapbio.org/
- Synopsis from Source: “ASAPbio is a scientist-driven initiative to promote the productive use of preprints in the life sciences.”

Example: “Preprint Stories”

- Source: http://asapbio.org/preprint-info/preprint-stories
- Synopsis from Source: “Some scientists are concerned about possible negative repercussions of preprinting, like being scooped or being disqualified from submitting to a journal of choice, and others feel that preprinting may have no effect at all. However, many biologists have had very positive experiences with preprinting – including receiving helpful feedback, invitations to submit to journals, jobs, and grants as a result. We’ve collected some of their stories here.”

v. Category: Open Science Community Engagement

Example: Genos

- Source: https://genos.co/
- Synopsis from Source: “Personalized healthcare is in its infancy. Many new insights will be driven by successful research and require community help. Genos provides such opportunities - engage with the Impact Network studies that resonate to you.”

Example: “Should you profit from your genome?”

- Source: http://www.nature.com/nbt/journal/v35/n1/full/nbt.3757.html
- Synopsis from Source: “Finally, the ability to personally profit from genetic data may serve the greater good. These commercial rights could encourage individuals to participate in research or to share their information, which in turn could lead to greater scientific advances. Some authors have described the possibility of “network effects” with respect to the sharing of medical data: the more medical data we have available, the greater the potential benefit of those data9, 12. However, at present we are unable to extract the full value of medical data. Allowing individuals to profit from the use of their genetic and other health data could build trust in systems of data exchange and create incentives to share data across multiple platforms13. Because of network effects, this increased sharing could actually raise the value of the information being shared, resulting in greater social and financial gains.”

Example: InnoCentive

- Source: https://www.innocentive.com/
- Synopsis from Source: “Our Challenge Driven Innovation™ methodology and purpose-built technology result in fresh thinking and cost-effective problem solving. Crowdsource solutions from our diverse network of highly educated problem solvers or internally within your organization.”

Example: “Case Studies”

- Source: https://www.innocentive.com/resources-overview/case-studies/
- Synopsis from Source: “Below you will find case studies from a range of industries, providing snapshots of some of the things that have been achieved through Challenge Driven Innovation. Will you be part of our next success story?”

Example: American Gut Project (AGP)

- Source: http://americangut.org/
- Synopsis from Source: “The American Gut is one of the largest crowd sourced, citizen science projects in the country. We discover new information daily to shed light on the connections between the human microbiome and health.”

Example: “The microbiome and big data”

- Source: http://www.sciencedirect.com/science/article/pii/S2452310017301270
- Synopsis from Source: “Microbiome citizen science initiatives such as the American Gut Project (AGP; americangut.org) have made significant contributions to the field by ‘democratizing’ microbiome research and thus providing large-scale datasets that can be used as comparative frameworks for other studies.”

Example: One Mind

- Source: http://onemind.org/
- Synopsis from Source: "One Mind is an independent, 501(c)(3) non-profit organization. We are dedicated to benefiting all affected by brain illness and injury through fostering fundamental changes that will radically accelerate the development and implementation of improved diagnostics, treatments and cures — while eliminating the stigma that comes with mental illness. The mission of One Mind is to alleviate human suffering from the diseases of the brain by funding scientific research into the causes, prevention, and new treatments leading to cures for brain disease and injury. Our mission is fueled by our belief in open science principles and creating global public-private partnerships among governmental, corporate, scientific and philanthropic communities."

Example: "CDISC and One Mind Announce Availability of Research Data Standard for Traumatic Brain Injury"

- Source: http://www.onemind.org/About-Us/Press/CDISC-TBI-Standard-Now-Available/
- Synopsis from Source: "The Clinical Data Interchange Standards Consortium (CDISC) and One Mind are pleased to announce the availability of version 1.0 of the Traumatic Brain Injury (TBI) research data standard, openly accessible on the CDISC website. The TBI standard was developed by a consortium of medical and technical experts to enable the harmonization of clinical data pertaining to TBI studies, streamlining research processes and improving analysis capabilities to ensure accelerated learning, better science, and to advance treatments for patients with TBI."

vi. Category: Open Science Projects and Tools within Organizations

Example: Sage Bionetworks

- Source: http://sagebase.org/
- Synopsis from Source: "Sage Bionetworks is a non-profit research organization that seeks to develop predictors of disease and accelerate health research through the creation of open systems, incentives, and standards. We create strategies and platforms that empower researchers to share and interpret data on a colossal scale, crowdsource tests for new hypotheses, and contribute to knowledge through community challenges."
- Projects:
    - Neurofibromatosis Therapeutic Acceleration Program (NTAP)
        - Source: http://sagebase.org/research-projects/neurofibromatosis-therapeutic-acceleration-program-ntap/
        - Synopsis from Source: "NTAP is working with Sage on numerous fronts to accelerate the pace of drug discovery in Neurofibromatosis, a genetic pre-cancer syndrome that causes the growth of neurofibromas throughout affected patients."
    - iAtlas
        - Source: http://sagebase.org/research-projects/iatlas/
        - Synopsis from Source: "Sage Bionetworks and the Institute for Systems Biology will partner to build the Cancer Research Institute iAtlas, an online database and web resource designed to help basic and clinical researchers navigate immunological data across multiple tumor types."

- Tumor neoantigEn SeLection Alliance (TESLA)
  - Source: http://sagebase.org/research-projects/tumor-neoantigen-selection-alliance-bioinformatics-project/
  - Synopsis from Source: "The Parker Institute for Cancer Immunotherapy, started by Napster founder Sean Parker, and the Cancer Research Institute have joined together to launch a global research alliance with 30 public and private groups to advance the development of more effective personalized cancer treatments."
- Cancer Systems Biology Consortium (CSBC)
  - Source: http://sagebase.org/research-projects/coordinating-center-for-the-new-cancer-systems-biology-consortium-csbc/
  - Synopsis from Source: "Sage Bionetworks serves as the Coordinating Center for the new Cancer Systems Biology Consortium (CSBC). The interdisciplinary investigators of the CSBC will integrate experimental biology with mathematical and computational modeling to gain insight into processes relevant to cancer initiation, progression, and treatment options. The Coordinating Center lead by Sage Bionetworks will facilitate data and resource sharing, as well as collaborative scientific activities and systems biology-oriented outreach."
- Metastatic Breast Cancer Alliance (MBC)
  - Source: http://sagebase.org/research-projects/metastatic-breast-cancer-alliance/
  - Synopsis from Source: "The Metastatic Breast Cancer Alliance (MBC) is an advocacy group to evaluate and quantify the extent of research in metastatic breast cancer, and assess the overall adequacy of funding in this area."
- The Resilience Project
  - Source: http://sagebase.org/research-projects/the-resilience-project/
  - Synopsis from Source: "The Resilience Project launched in 2014 with a unique vision by Stephen Friend and Eric Schadt that by studying massive numbers of healthy adults, scientists might find rare individuals who are unaffected by genetic variants that should induce disease."
- mPower Researcher Portal
  - Source: http://sagebase.org/research-projects/mpower-researcher-portal/
  - Synopsis from Source: "The mPower Public Researcher Portal is the data access point for the mPower mobile Parkinson Disease study. mPower is an mobile application-based study piloting new approaches to monitoring key indicators of Parkinson Disease progression and diagnosis by supplementing traditional behavioral symptom measurements with novel metrics gleaned from sensor-rich mobile devices."
- Colorectal Cancer Subtyping Consortium (CRCSC)
  - Source: http://sagebase.org/research-projects/colorectal-cancer-subtyping-consortium-crcsc/
  - Synopsis from Source: "The CRCSC is a collection of colorectal cancer researchers and data scientists from 15 academic, pharmaceutical, and non-profit partners which represents the first effort to generate consensus cancer subtypes, taking a major step toward precision medicine in CRC."

- Molecular Mechanisms of the Vascular Etiology of Alzheimer's Disease (M2OVE-AD) Consortium
    - Source: http://sagebase.org/research-projects/m2ove-ad-consortium/
    - Synopsis from Source: "The Molecular Mechanisms of the Vascular Etiology of Alzheimer's Disease (M2OVE-AD) Consortium seeks to better understand how the vascular system may be involved in the onset and progression of AD and related dementias. Developed by the NIA and the NINDS, this program brings together over a dozen research teams working on five complementary projects."
- Accelerating Medicine Partnership in Alzheimer's Disease (AMP-AD)
    - Source: http://sagebase.org/research-projects/accelerating-medicine-partnership-in-alzheimers-disease/
    - Synopsis from Source: "The Accelerating Medicine Partnership for Alzheimer's Disease (AMP-AD) Knowledge Portal is a premier repository and workspace for multi-omic data generated through the AMP-AD consortia – a public private partnership between academic research groups, industry partners, and non-profit partners. The primary goal of the consortia is to generate and test hypotheses to identify new targets for intervention in the treatment of Alzheimer's disease."
- Progenitor Cell Biology Consortium (PCBC)
    - Source: http://sagebase.org/research-projects/tcga-pancan-gi/
    - Synopsis from Source: "The Progenitor Cell Biology Consortium (PCBC) is a NHLBI sponsored effort to identify and characterize progenitor cell lineages, to direct the differentiation of stem and progenitor cells to desired cell fates, and to develop new strategies to address the unique challenges presented by the transplantation of these cells."
- Genomics Evidence Neoplasia Information Exchange (GENIE) Project
    - Source: http://sagebase.org/research-projects/aacr-project-genie/
    - Synopsis from Source: "The Genomics Evidence Neoplasia Information Exchange (GENIE) Project is funded by the American Association for Cancer Research (AACR). The goal of GENIE is to harmonize precision medicine studies across major cancer studies. Sage Bionetworks serves an integral role in this project as an informatics partner."
- PsychENCODE Consortium
    - Source: http://sagebase.org/research-projects/ajcc-biomarker-project-in-colon-cancer/
    - Synopsis from Source: "The PsychENCODE program is funded by the NIMH with the goal of accelerating discovery of non-coding functional genomic elements in human brain, and elucidating their role in the molecular pathophysiology of psychiatric disorders. The Consortium investigators are producing a public resource of multi-dimensional genomic data using unbiased genome-wide approaches on tissue and cell-type specific samples from approximately 1000 phenotypically well-characterized healthy and diseased human post-mortem brains."

- SYNODOS
  - Source: http://sagebase.org/research-projects/childrens-tumor-foundation-ctf/
  - Synopsis from Source: "Synodos is a first-of-its-kind Neurofibromatosis research collaboration dedicated to defeating a variety of forms of neurofibromatosis. To date there are three Synodos efforts funded by Children's Tumor Foundation and supported by Sage. The Synodos model brings together a multidisciplinary team of scientists from world-class labs at academic and medical centers of excellence, who have pledged to work closely together – sharing information, datasets, results and more – at every step in research development, with the goal of speeding up the drug discovery process."
- The Cancer Genome Atlas Pancancer Analysis Working Group
  - Source: http://sagebase.org/research-projects/the-cancer-genome-atlas-pancancer-analysis-working-group/
  - Synopsis from Source: "In 2013, with data from 12 cancers being released from the embargo The Cancer Genome Atlas (TCGA) embarked on an effort to simultaneously analyze six molecular platforms (miRNA-Seq, RNASeq, methylation arrays, Protein arrays, DNA-Seq and copy number arrays) for patterns of similarity and differences across cancer types. Working in Synapse this group not only created a common data resource of normalized data but also made all intermediate and final results available for the community in the form of supplemental compendiums to 18 published papers."
- The CommonMind Consortium
  - Source: http://sagebase.org/research-projects/the-commonmind-consortium/
  - Synopsis from Source: "The CommonMind Consortium Knowledge Portal is the main distribution site for data and analysis results generated by the CommonMind Consortium members. Sage Bionetworks initiated the CommonMind Consortium (CMC) in 2010 as a Public-Private Pre-Competitive partnership that brings together disease area expertise, large scale and well curated brain sample collections, and data management and analysis expertise with a goal to generate and analyze large-scale genomic data from human subjects with neuropsychiatric disease."

Example: "Researchers from Mount Sinai and Sage Bionetworks Report Analysis of Nearly 600,000 Genomes for Resilience Project"

- Source: http://www.mountsinai.org/about-us/newsroom/press-releases/researchers-from-mount-sinai-and-sage-bionetworks-report-analysis-of-genomes-for-resilience-project
- Synopsis from Source: "'This study demonstrates the power of using big data to ask new biological questions,' said Anne Wojcicki, co-founder and CEO of 23andMe, which participated in the project. 'More than 400,000 23andMe customers contributed to this effort, showing that engaged consumers can make a real impact on scientific research.'"

Center for Open Science

- Source: https://cos.io/

- Synopsis from Source: “Our mission is to increase the openness, integrity, and reproducibility of scholarly research. We work on ways to make the process, content, and outcomes of research openly accessible for others to discuss and build on.”
- Projects:
    - SHARE
        - Source: http://www.share-research.org/about/about-share/
        - Synopsis from Source: “SHARE is a higher education initiative whose mission is to maximize research impact by making research widely accessible, discoverable, and reusable. To fulfill this mission SHARE is developing services to gather and freely share information about research and scholarly activities across their life cycle. Making research and scholarship freely and openly available encourages innovation and increases the diversity of innovators.”
    - Preregistration Challenge
        - Source: https://cos.io/prereg/
        - Synopsis from Source: “If you have a project that is entering the planning or data collection phase, we'd like you to try out a preregistration. Through our $1 Million Preregistration Challenge, we're giving away $1,000 to 1,000 researchers who preregister their projects before they publish them. It's straightforward to complete and will really enhance your research output.”
    - Open Science Framework
        - Source: https://cos.io/our-products/open-science-framework/
        - Synopsis from Source: “The Open Science Framework (OSF) provides free and open source project management support for researchers across the entire research lifecycle.”
    - Reproducibility Project: Psychology
        - Source: https://osf.io/ezcuj/
        - Synopsis from Source: “The Reproducibility Project uses an open methodology to test the reproducibility of psychological science. It also models procedures designed to simplify and improve reproducibility.”
    - Reproducibility Project: Cancer Biology
        - Source: https://osf.io/e81xl/
        - Synopsis from Source: “The Reproducibility Project: Cancer Biology is a collaboration between Science Exchange and the Center for Open Science, and is independently replicating a subset of experimental results from a number of high-profile papers in the field of cancer biology published between 2010-2012 using the Science Exchange network of expert scientific labs.”

Example: “Estimating the reproducibility of psychological science”

- Source: http://science.sciencemag.org/content/349/6251/aac4716
- Synopsis from Source: “Innovation points out paths that are possible; replication points out paths that are likely; progress relies on both. Replication can increase certainty when findings are reproduced and promote innovation when they are not.”

vii. Category: General Examples of Benefit

Why Open Science is Necessary

"We need Open Science for Sustainability"

- Source: https://blog.frontiersin.org/2016/09/30/open-science-for-sustainability/
- Synopsis from Source: "Faced with exponentially growing problems, we need exponential growth in solutions – the sort of growth we have seen in IT and in biotech. To date, this kind of growth has not been forthcoming. We suggest that the answer lies in the infrastructure, the knowledge networks, the plumbing that makes the scientific enterprise tick, the machinery that allows science to evolve."

"Open Science and Access to Medical Research"

- Source: https://blogs.scientificamerican.com/guest-blog/open-science-and-access-to-medical-research/
- Synopsis from Source: "The solution to these potential issues that may arise when we transition to an open science format is not to limit the access of the data. Instead, it is imperative that concomitant with the creation of an open science environment we also build independent institutions or organizations that help interpret the available the data in a manner that non-scientists are able to receive accurate and solid information about the nature and significance of the results."

"Is Open Science the Future of Drug Development?"

- Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5369032/
- Synopsis from Source: "This perspective argues that while open science poses some risks—which include the management of collaboration and the protection of proprietary data—these strategies are, in many cases, the more efficient and ethical way to conduct biomedical research."
- Examples:
    - Yale University Open Data Access (YODA) Project
        - Source: http://yoda.yale.edu/
    - Human Genome Project
        - Source: https://www.genome.gov/10001772/all-about-the--human-genome-project-hgp/
    - ClinicalTrials.gov
        - Source: https://clinicaltrials.gov/
    - InnoCentive
        - Source: https://www.innocentive.com/
    - Medicines for Malaria Venture (MMV)
        - Source: https://www.mmv.org/
    - ClinicalStudyDataRequest.com (CSDR)
        - Source: https://www.clinicalstudydatarequest.com/

"Can open science help patients and save pharma?"

- Source: https://opensource.com/health/14/6/can-open-science-help-patients-and-save-pharma

- Synopsis from Source: “These study results support that transparency and collaboration such as that envisioned by open science would be positive for: 1) R&D efficiency and costs, 2) science, 3) patients as individuals, and 4) population health as a whole.”

“Open science is a research accelerator”

- Source: http://www.nature.com/nchem/journal/v3/n10/full/nchem.1149.html
- Synopsis from Source: “An open-source approach to the problem of producing an off-patent drug in enantiopure form serves as an example of how academic and industrial researchers can join forces to make new scientific discoveries that could have a huge impact on human health.”

“How open science helps researchers succeed”

- Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4973366/
- Synopsis from Source: “We review literature demonstrating that open research is associated with increases in citations, media attention, potential collaborators, job opportunities and funding opportunities. These findings are evidence that open research practices bring significant benefits to researchers relative to more traditional closed practices.”

“Why open science matters”

- Source: http://onsnetwork.org/chartgerink/2013/06/29/why-open-science-matters/
- Synopsis from Source: “Several of the most important consequences for the population include (1) more public scrutiny of research, (2) better research in general and (3) a better foundation for businesses to build their company on. Much research is funded by the public via tax money.”

Why Science Should be Open, More Collaborative

“Why is publication of negative clinical trial data important?”

- Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3514755/
- Synopsis from Source: “Because most compounds do fail at phase IIa (Arrowsmith, 2011b), spending large sums of money on compounds doomed to fail is to the commercial disadvantage of everyone; therefore, we strongly believe that all clinical trial data should be made publicly available in a timely manner. This allows assessment as to whether a particular molecular target has been invalidated or not, which, in the long run, is to the benefit of everyone working in the industry. Indeed, Arrowsmith's (2011b) analysis of phase II failures concluded that an increase in collaborative efforts between companies up to the point of PoC for novel targets or mechanisms might be more cost-effective and time-efficient.”

“Stop the privatization of health data”

- Source: http://www.nature.com/news/stop-the-privatization-of-health-data-1.20268
- Synopsis from Source: “Yet there is a major downside to monoliths such as Google or smaller companies such as consumer-genetics firm 23andMe owning health data — or indeed, controlling the tools and methods used to match people's digital health profiles to specific services.”
- Examples:
    - Enlite
        - Source: https://www.medtronicdiabetes.com/products/enlite-sensor
    - 23andMe
        - Source: https://www.23andme.com/

- Sage Bionetworks
  - Source: http://sagebase.org/

"When Evidence Says No, but Doctors Say Yes"

- Source: https://www.theatlantic.com/health/archive/2017/02/when-evidence-says-no-but-doctors-say-yes/517368/
- Synopsis from Source: "For all the truly wondrous developments of modern medicine—imaging technologies that enable precision surgery, routine organ transplants, care that transforms premature infants into perfectly healthy kids, and remarkable chemotherapy treatments, to name a few—it is distressingly ordinary for patients to get treatments that research has shown are ineffective or even dangerous."

"The costs of inequality: Money = quality health care = longer life"

- Source: http://news.harvard.edu/gazette/story/2016/02/money-quality-health-care-longer-life/
- Synopsis from Source: "Addressing those social and behavioral factors would require government officials and community leaders to think innovatively and cooperatively about the everyday realities that affect health, even down to the designs of neighborhoods and transportation systems. There would have to be more flexibility for health care spending to help, for example, an asthmatic child whose medication might be covered but whose need for an air conditioner is not. Physicians would need to be aware of nonmedical pressures that patients face after leaving their offices that might, for instance, leave them without transportation to follow-up appointments or to pharmacies for medicine. Partner organizations would need to help meet routine needs by such things as grocery shopping and cooking for postoperative patients who are well enough to go home, but can't yet push carts down store aisles."

"How Flawed Science Is Undermining Good Medicine"

- Source: http://www.npr.org/sections/health-shots/2017/04/06/522262881/how-flawed-science-is-undermining-good-medicine
- Synopsis from Source: "The impact of weak biomedical research can be especially devastating, Harris learned, as he talked to doctors and patients. And some prominent scientists he interviewed told him they agree that it's time to recognize the dysfunction in the system and fix it."

"Open Science? Try Good Science."

- Source: https://hub.wiley.com/community/exchanges/discover/blog/2014/04/01/open-science-try-good-science
- Synopsis from Source: "Beyond my anecdotal example above, we have documented examples where errors in the literature have significant effects on grants awarded or the ability to publish papers that are in disagreement (e.g., Miller, 2006). All of these have a very real human cost to science and scientists."

"Open Science and Verifiability"

- Source: http://web.stanford.edu/~vcs/Nov21/dg-OpenScienceandVerifiability.pdf
- Synopsis from Source: "Our position on open source and open data in science was arrived at when an increasing number of papers began crossing our desks for review that could not be

subjected to verifiability tests in any meaningful way. Paper A might have used a commercial package that comes with a license that forbids people at university X from viewing the code! Paper 2 might use a code which requires parameter sets that are 'trade secrets' and have never been published in the scientific literature. Our view is that it is not healthy for scientific papers to be supported by computations that cannot be verified except by a few employees at a commercial software developer. Should this kind of work even be considered Science?"

"The Human Cost of a Misleading Drug-Safety Study"

- Source: https://www.theatlantic.com/health/archive/2015/09/paxil-safety-bmj-depression-suicide/406105/
- Synopsis from Source: "The study is now again in the news, as a new reanalysis of the its original data—including about 77,000 pages of formerly inaccessible patient records—shows that Paxil was neither effective nor safe. The reanalysis, published in the scientific journal BMJ, found that the study, underwritten by the drug's maker, GlaxoSmithKline (or GSK), created a false picture of safety partly by misclassifying suicidal acts (such as taking 80 Tylenol) as less-alarming behavior or side effects. Other researchers who have looked at Study 329's data have concluded likewise."

"Registered clinical trials make positive findings vanish"

- Source: http://www.nature.com/news/registered-clinical-trials-make-positive-findings-vanish-1.18181
- Synopsis from Source: "However, he says, this means that at least half of older, published clinical trials could be false positives. 'Loose scientific methods are leading to a massive false positive bias in the literature,' he writes."

"Perspective: Science is still too closed"

- Source: http://www.nature.com/nature/journal/v533/n7602_supp/full/533S70a.html
- Synopsis from Source: "Both monoclonal antibody and phage-display technologies are used to identify precursors to antibody drugs. Both technologies were invented more than 25 years ago, and both have been used to discover successful medicines. But whereas the phage-display technology has been fiercely protected, monoclonal antibody technology was placed in the public domain. As of 2014, there were 47 approved monoclonal antibody drugs and only 7 derived from phage-display technology."

Open Research

"Open research and collaborations"

- Source: https://emckiernan.wordpress.com/2015/04/09/open-research-and-collaborations/
- Synopsis from Source: "People have sent me links to all kinds of interesting projects: open collaborative transcription and editing of humanities texts; biomedical research made possible by publicly available data; projects creating open source tools that can be used by all and continue to be improved by online communities; and personal stories about how they got PhDs, jobs, and authorships through open science."

Open Data

"How open data could save millions of lives"

- Source: https://www.weforum.org/agenda/2016/07/how-open-data-could-save-millions-of-lives/
- Synopsis from Source: "In May, a global group of public health experts convened by Chatham House took a bold step forward that stands to change that, issuing a call for *all* public health surveillance data to be shared as necessary to improve and protect public health."
- Examples of real-time surveillance data:
    - Mekong Basin Disease Surveillance (MBDS)
        - Source: http://www.mbdsnet.org/
        - Synopsis from Source: "Mission: strengthen national and sub-regional capabilities in infectious disease surveillance and outbreak response, especially for priority diseases, to rapidly and effectively control them."
    - Child Health and Mortality Prevention Surveillance (CHAMPS)
        - Source: https://champshealth.org/
        - Synopsis from Source: "CHAMPS Network's goal is to provide accurate, timely and reliable data on the causes of death and sickness for children under five to improve health outcomes and quality of life."

Data Sharing

"Three Ways Communities Are Sharing Data to Improve Health"

- Source: http://dashconnect.org/2016/09/01/three-ways-communities-are-sharing-data-to-improve-health/
- Synopsis from Source: "As communities form multi-sector collaborations to share and use data that supports their ability to better understand and address the social determinants of health, there is a growing recognition of the need to expand partnerships and build capacity beyond the health care sector. A primary aim of Data Across Sectors for Health (DASH) is to better understand local initiatives that are pioneering new ways to integrate data systems from multiple sectors with the ultimate aim of improving community health."
- Examples of multi-sector collaborations:
    - Altair Accountable Care for People with Disabilities
        - Source: http://dashconnect.org/2016/09/01/integrated-care-for-people-with-disabilities-lessons-from-an-accountable-care-organization/
        - Synopsis from Source: "One out of every three people with intellectual and developmental disabilities may also have behavioral health issues. However, very little has been done to standardize the process for assessing this population for depression, anxiety, isolation, and other behavioral health conditions."
    - Center for Innovation through Data Intelligence (CIDI): Community Health Improvement Information System
        - Source: http://www1.nyc.gov/site/cidi/projects/community-health-improvement-information-system.page
        - Synopsis from Source: "Administrative data generally have detailed geographic information, but limited socioeconomic information. Triangulating

administrative data with other data sources such as Census data at the neighborhood level can provide insight into the social determinants of health and community well-being."

- Allegheny County Data Sharing Alliance for Health
  - Source: http://dashconnect.org/2016/06/02/connecting-data-across-sectors-to-address-cardiovascular-disease/
  - Synopsis from Source: "When Hacker learned about Data Across Sectors for Health (DASH), she was excited about the opportunity to leverage existing partnerships in order to combine multi-sector data and address questions about the causes and drivers of cardiovascular disease in the county."

"The Importance of Clinical Trial Data Sharing: Toward More Open Science"

- Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3318983/
- Synopsis from Source: "The clinical and research community often only becomes aware of its shortcomings when safety concerns are raised about a drug, device or other treatment strategy. Past experiences with rofecoxib (Vioxx), rosiglitazone (Avandia), and oseltamivir (Tamiflu) have illustrated that it is in the public's interest to have access to comprehensive clinical trial data to ensure a complete understanding of drug or device safety and effectiveness."
- Examples of clinical trial data sharing:
  - Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC)
    - Source: https://biolincc.nhlbi.nih.gov/home/
  - International Stroke Trial (IST)
    - Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3104487/
  - Yale University Open Data Access (YODA) Project
    - Source: http://yoda.yale.edu/

Open Access

"The academic, economic and societal impacts of Open Access: an evidence-based review [version 3; referees: 4 approved, 1 approved with reservations]"

- Source: https://f1000research.com/articles/5-632/v3
- Synopsis from Source: "This review aims to be a resource for current knowledge on the impacts of Open Access by synthesizing important research in three major areas: academic, economic and societal. While there is clearly much scope for additional research, several key trends are identified, including a broad citation advantage for researchers who publish openly, as well as additional benefits to the non-academic dissemination of their work."

Open Access Publishing

"Let's speed up science by embracing open access publishing"

- Source: https://www.statnews.com/2016/12/19/open-access-publishing/
- Synopsis from Source: "If published research and data were freely accessible and reusable by researchers of diverse interests, urgently needed solutions could be greatly accelerated. Scientists could quickly cross-check important studies, catching potentially consequential mistakes. Medical providers could access the latest technical guidance, improving patient care. And students around the world could build on each other's work. With openness, good ideas could truly come from anyone, anywhere."

- Examples:
    - Harvard scientist Jay Bradner and JQ1 molecule
        - Source: https://www.ted.com/talks/jay_bradner_open_source_cancer_research
    - Human Genome Project
        - Source: https://www.genome.gov/10001772/all-about-the--human-genome-project-hgp/

Preprints

"The findings of medical research are disseminated too slowly"

- Source: https://www.economist.com/news/science-and-technology/21719438-about-change-findings-medical-research-are-disseminated-too
- Synopsis from Source: "The wider use of preprints might also help reduce the problem of pre-publication data-hoarding. Once a preprint is published, its authors need not fear that others will take credit for their work. And it is becoming easier to make data available in a way that lets the originator retain control and garner credit."
- Examples:
    - bioRxiv
        - Source: http://www.biorxiv.org/
    - ASAPbio
        - Source: http://asapbio.org/
    - Chan Zuckerberg Biohub
        - Source: https://www.czbiohub.org/
    - F1000Research
        - Source: https://f1000research.com/
    - Figshare
        - Source: https://figshare.com/

Open Scholarship

"Imagining the 'open' university: Sharing science to improve research and education"

- Source: https://peerj.com/preprints/2711/
- Synopsis from Source: "Open scholarship, such as the sharing of articles, code, data, and educational resources, has the potential to improve university research and education, as well as increase the impact universities can have beyond their own walls. To support this perspective, I present evidence from case studies, published literature, and personal experiences as a practicing open scholar."

Open Writing

"Open Writing is the Necessary Precursor to Open Science"

- Source: https://www.socialsciencespace.com/2016/11/open-writing-necessary-precursor-open-science/
- Synopsis from Source: "As process, our Open Writing ambition thus encourages various types of collaboration. As product, Open Writing is communicative, accessible and sensitive to its potential impact (which can be economic, social, ethical or political). Yet Open Writing does not end here; how we as researchers -through our writing – stimulate and react to innovation and

impact is crucial to enhancing the Open Science agenda. If there is no real connection between the researcher (writer) and the recipient (reader), we risk our writing circulating only within a closed system of peers. Open Writing seeks to open up this closed circuit."

Appendix B: Open Science Interviews

Selection Criteria: The interviews were with eleven individuals, a mix of genders, ages, and locations, and each were selected for a different perspective on open science. These are thought leaders in the field of open science, many dedicating their recent careers to its exploration and promotion, as well as open policy, data science, data, and librarianship. Roles of those interviewed include director, executive, policy influencer, computer scientist, researcher, professor, and librarian, with some of these distinctions overlapping.

Appendix B.1: Open Science Definitions and Interview Summaries

Patricia Flatley Brennan, RN, PhD

Role

Director of the National Library of Medicine (NLM)

Open Science Definition

Open science is a philosophy that drives knowledge, and it identifies commitment through transparency, engagement, and awareness.

Interview Summary

Open science is a philosophy that drives knowledge, and it identifies commitment through transparency, engagement, and awareness. It is a way of creating knowledge that is an alternative to industrial science. The open science model is driven by discovery, not solely by profit. Open science is how we frame things, and it's about who gets to make decisions regarding literature and investments. It isn't simply citizen science, and it pulls together many roles. It requires a dialogue. Open science isn't a claim, it's a demonstration. By contrast, data science is about discovery and the use of methods to extract knowledge from data. It's a more inductive approach. At the intersection of open science and data science is an access to data. They are both part of the conversation about what kind of data you collect, and the engagement around original data collection. Indexing and cataloging are involved in both, and they add levels of detail for discovery but also uncertainty to data with how it is labeled. The open science process means that interpretation is attached at many points within the process, versus industrial science where it is closed the entire time. A data driven discovery requires an open science mindset/framework, with rigor and reproducibility as key; need to get good value out of it. With data science investigations, it's more important to show comprehensiveness. As with data science, you're trying to show repeatability of the phenomenon. Open science is a philosophy of discovery, not repudiation. Open science is about adding and subtracting different elements. It has characteristics of completeness, dynamism, the experimental space of what phenomenon brings with it. Doesn't have to be like industrial science and comply with how the study was originally designed. There are degrees of openness, and there are our resources of open. One is not better than the other. NLM's open science role is as an early adopter. Open science is how a library is structured, with an embrace of open science through promoted access to free information and public engagement. NIH ultimately underwrites open science.

Russ Altman, MD, PhD

Role

Professor of Bioengineering, Genetics, Medicine, Biomedical Data Science, and Computer Science, and Director of the Biomedical Informatics Training Program at Stanford University

Open Science Definition

Open science is a series of best practices and expectations for scientists, that science should be shareable, replicable, and accessible, with the overall goal of accelerating science while still respecting scientists and preserving careers.

Interview Summary

Open science is a series of best practices and expectations for scientists, that science should be shareable, replicable, and accessible, with the overall goal of accelerating science while still respecting scientists and preserving careers. Open science is operating procedures and a way of conducting science, not a thing, a discipline, or a science. Difficult term when it comes to reporting, auditing, and marketing. Data science is just part of informatics, and too much data is an informatics problem. Data science is the collection, management, and analysis of data, and it makes the sharing of the products of itself, like data and algorithm sharing, easier. Data science is an important first area for open science, because if you can't do open science with data science, it can't be done. Data sharing only accelerates the sciences by improving the environment, a secondary ripple effect. In the biomedical sciences, open science is needed by the informatics field, experimental scientists, and the clinical realm. DNA sequencing data has set our expectations of data. It's digital, hard to misinterpret, understandable. There's the aspirational goal of the complete capture, and new technologies, perhaps from other fields, could enable the unfeasible thought. There's a treasure trove of data to mine and reuse. Everyone wants clinical data, and if people would share data, we could avoid trials because we'd know the results of what has already been done. The problem is people relinquishing data they worked hard for, and there's always the issue of money. Can measure open science progress by looking at drugs and how tax dollars treat or prevent disease, our ability to measure the health of the American people. Secondary measures include ease and speed for people to test new hypotheses, biomarkers like the availability of drugs, lifespan, infant mortality, etc, and case studies that show its value.

Elizabeth Marincola

Role

Senior Advisor for Science Communications and Advocacy at the African Academy of Sciences

Open Science Definition

Open science means no barriers to leveraging the products of science.

Interview Summary

Data science is any tool that enables open science, the remixing of data, and scientists to do their work. Has the perspective of Africa, where open science allows them to sequence genes and more easily do things at the clinical edge. Envisions world where 95% of science is communicated through a publishing platform, not a journal. Science communicated with data science allows for open science, and we need to develop processes that allow for this like preprints, open evaluation, etc. Data science is critical because the full power of open science won't come about without it. Open science means no barriers to leveraging the products of science. This applies to education, data mining, public health, etc. There should be nothing artificial as a barrier to sharing output. People conflate open access with open science, but it is a subset, and only one aspect. Open access doesn't mean you have the appropriate tools, that you can share or teach, or that you have all the data. The open science notion is that daylight is the best disinfectant. Open science leads to better scrutiny, quality rises to the top, it still involves peer review, and can be more effective and credible. Open science helps to cure or combine redundancy

in funding, and people need to know what's being done. Open science operating at its best is it operating in the interest of the community. People get away with too much in a closed environment, and open will promote good behavior, not the opposite.

Brian Nosek, PhD

Role

Professor of Psychology at the University of Virginia, and Co-founder and Executive Director of The Center for Open Science

Open Science Definition

Open science is science, and that's it.

Interview Summary

Open science is science, and that's it. Openness is supposed to be there, as it is the definition of science. The term of open science and this redundancy is needed because of the reality of science today. Science claims become valid with reproducibility, and for that we need transparency, openness, and accessibility. Much of the academic science world is closed because the primary incentive drivers are tenure, grants, and publications. All of this provides status to researchers, but openness may not be required or relevant. Environment constraints, versus ideals, and societal influences, as in what do other researchers in field think is important, are key considerations. Reproducibility crisis is well known, but needs to be translated into policies, standards, etc. Good infrastructure is needed. Look at ways to simultaneously train and provide tools, as in learn and change behaviors as they go, and don't disrupt workflows. Important to meet researchers where they are, like with the Open Science Framework. It is usable as an entirely private tool for a lab, but from there it is easy to go open. Focus on services, interfaces, and moving to scale.

John Wilbanks

Role

Chief Commons Officer at Sage Bionetworks

Open Science Definition

Open science means multiple ways to follow a scientific process, and it means an idea of thinking about science, not something to point at.

Interview Summary

Can't be a singular definition of open science, much like there is no such thing as a singular science, but sciences. Hard and dangerous to force open science to be one thing, but it is useful to have and use the phrase. Open science is not just the object itself but the object(s) and the process together. Can know what open source software is because it's the objects that relate to the process. If you can distinguish an object, as in it can carry property rights, then you can also decide if it can be open. Sharing should be trivial. Open science means multiple ways to follow a scientific process, and it means an idea of thinking about science, not something to point at. Open science can be closed at first to intentionally create a diversity of thought, but it needs to be open at some point too, viral, so that others adopt their workflows and methodologies. Humans are bad as seeing the negative effects of the typical way of doing things. There's the failure of the existing system, and the only people who can get out are the ones who don't need it. Mandates and the societal pressure of seeing who's complying. Open science is intersectional and self-reinforcing, and it will form its own ecosystem eventually.

Heather Joseph

Role

Executive Director of Scholarly Publishing and Academic Resources Coalition (SPARC)

Open Science Definition

Open science is about thinking about the research process end to end with an open model, as in designing, conducting, and communicating openly, with as few technical, legal, and financial barriers as possible.

Interview Summary

A wider definition of open science beyond open access is needed, as the terms are often used interchangeably. Scholarly Publishing and Academic Resources Coalition (SPARC) doesn't have official open science or data science definitions, and instead uses the terms open access and open data. Senses a shift from open access, as researchers seem to like open science more. Open science is about thinking about the research process end to end with an open model, as in designing, conducting, and communicating openly, with as few technical, legal, and financial barriers as possible. Open access and open data are part of the open science environment/ecosystem. Prefers the term ecosystem for open science as it describes a defined process, whereas environment is evocative of more open scholarship. Open is about empowering science to be more efficient and faster. Open science is an aspirational goal, and people need definitions to aspire to and align themselves with. It is important to consider what compels people.

Christine Borgman, PhD

Role

Distinguished Professor and Presidential Chair in Information Studies, and Director of the Center for Knowledge Infrastructures at University of California, Los Angeles (UCLA)

Open Science Definition

Instead of looking at open, a term that publishers have co-opted and no one comprehends, look at data reuse for framing, and see if there's a way to measure it.

Interview Summary

Instead of looking at open, a term that publishers have co-opted and no one comprehends, look at data reuse for framing, and see if there's a way to measure it. Software, documentation, and calibration are needed to make data useful. Librarians are the ones to think about stewardship issues, not scientists or politicians. Important to think about the new kinds of education and engagement that are needed to make data usable in 20 years, and valuable now. Until it's someone's job, no one will be invested in it, and training programs have the greatest potential. The rewards system is flawed, as no one is promoted for being a good data scientist. There is a need for interoperable infrastructure to cure the assumptions of self-curated data. Look to see what's working in other fields, and what's working in other parts of the world. How can we get people to cite data if we can't even agree on what a dataset is? We need infrastructure for cross referencing data across publications and databases.

Erin McKiernan, PhD

Role

Assistant Professor of Biomedical Physics at Universidad Nacional Autónoma de México / National Autonomous University of Mexico

Open Science Definition

Open science is an umbrella term for sharing research and education.

Interview Summary

Open science is an umbrella term for sharing research and education. Key features of open are availability and access, reuse and redistribution, and universal participation. This can apply to most things, not just as a definition but to products within the research life cycle. Open science is on a spectrum, with sharing the final version of a paper on one end, and sharing e-notebooks in real time on the other. Open science adoption requires incremental steps. Open science is an overall way of looking at how researchers share products of research, not just articles, and involving the concept within the educational process. For effective data science, data must be in organized in a structured and accessible format that allows for manipulation. Data science and open science overlap, but they're not the same thing. Data science can be closed, even when it's better for data to be open.

Lisa Federer

Role

Research Data Informationist at the NIH Library

Open Science Definition

Open science is science along a continuum with its open data, it encompasses the research process, and it is circular.

Interview Summary

It is important to include librarians in this conversation. Open science is science along a continuum with its open data, it encompasses the research process, and it is circular. While open data is a part of that, open science is much larger. What comes to mind is citizen science. This is crowd sourcing data collection where people are engaged, scientific literacy increases, and the larger community benefits. Open data can't just be collected and available, it has to be good and appropriate. These are the resources that enable discovery, and it exists in an ecosystem of models of data that allows for reuses and a symbiotic relationship. There are also a lot of changes that need to take place. Training is important, and researchers need the skills of how to find data, reanalyze, and use open source tools, etc. Incentive structures need to change as well, such as the publication process and scientific sharing. There are other means to share as well, like open access journals and hackathons, and GitHub in general is a great form of both documentation and sharing, and involves immediate fixing and shaping with collaborators. The current scholarly communications model isn't the only way.

Terry Yoo, PhD

Role

Computer scientist in the Office of High Performance Computing and Communications (OHPCC) at the National Library of Medicine (NLM)

Open Science Definition

Open science is open access, open data, and open source.

Interview Summary

Open science is a massive concept that is difficult to define. All science is data science; it is the science of observation. There is data science buzz now because of the Internet and its massive amounts of curated, annotated data to learn from, and the watershed moment of having enough computational power. There are two main motivators: hope and fear. Some NIH Institutes and Centers are driven by fear, as in disease. NLM is the 'hope' institute, with data science the science of hope, because it promises that we can wrangle and make use of data to detect and cure. Science is closed because practices such as competitive grants make it closed. Open science is democratizing. Science needs to be open because of massive storage quantities. National initiatives, such as the 'All of Us' initiative at NIH, are required for open, epidemiological, and longitudinal needs. Open science with its open data is the only solution to this massive task. Open source as important key, as the computer is now the primary method/tool/instrument for discovery. Data has a lifecycle, data is a spectrum. An important aspect of open science is saving data. Value is measured in downloads and citations, and if it's foundational, what would happen if it disappeared? Importance of centralized place like NLM, accomplishing what no singular entity could do, and then giving it away for free. Open science is open access, open data, and open source. This is what is required to be a reproducible science. If not, it's not progress.

Ben Busby, PhD

Role

Lead of Bioinformatics Training and Genomics Outreach Coordinator at the National Center for Biotechnology Information (NCBI) at the National Library of Medicine (NLM)

Open Science Definition

Open science means streaming everything to the world in real time.

Interview Summary

Open science is like big data, in that it is very nebulous. Private companies can do whatever they want, but for public institutions like NIH it's better to share, especially with data that would contribute to data infrastructure. Open science means streaming everything to the world in real time. Open source is a rule of thumb within the bioinformatics community. One tool used by NCBI to encourage open science is hackathons, which are community-based gatherings at which participants build tools and applications using data and code. With hackathons, open code helps to standardize aggregators and skeletons with a common workflow language. NCBI hackathons are important as it's a critical mass of people who understand the problem and have time to work on pre-scoped projects, meaning 80% of projects are built. This helps NCBI share information and moves the biogenomics community forward. Marketing is key after hackathons, because developers need to convince people to use their tools to make data interoperable.

Appendix B.2: Consolidated Open Science Recommendations for NLM

Training

- These recommendations involve training and the funds that support it, critical for the future.
    - Create training programs for data scientist librarians
    - Fund continuing education for the sciences

    - Continue to encourage challenges and competitions
    - Support more collaboration amongst researchers

Tools

- These recommendations relate to how tool development can advance open science efforts.
    - Produce data curation stylebooks
    - Build a library of analytical models to foster specific scientific skillsets
    - Build a CRISPR database
    - Develop a catalog of stem cells

Policy

- These recommendations are opportunities for policy and leading by example.
    - Incentivize data sharing
    - Develop way to measure reuse
    - Make open science an evaluation metric
    - Leverage power at the grant level with a reproducibility section

Infrastructure

- These recommendations relate to infrastructure and expansion. Many of the interviewees saw this area as key to solving open science issues.
    - Continue to support centralized and stable systems
    - Focus on the cloud, collaboration, and coordinating centers
    - Secure funding to become dissemination portal
    - Build infrastructure for cross referencing data across publications and databases
    - Become the intramural data connection between ICs

Appendix C: Open Science Resource List – Publications

Selection Criteria: This resource list presents different perspectives of groups assessing, defining, and advancing open science in written form.

Opening Science: The Evolving Guide on How the Web is Changing Research, Collaboration and Scholarly Publishing.

- Source: http://book.openingscience.org/
- Synopsis from Source: "This book will give researchers, scientists, decision makers, politicians, and stakeholders an overview on the *basics*, the *tools*, and the *vision* behind the current changes we see in the field of knowledge creation. It is meant as a starting point for readers to become an active part in the future of research and to become an informed party during the transition phase. This is pivotal, since research, as a sensitive, complex process with many facets and millions of participants, hierarchies, personal networks, and structures, needs informed participants."

Expanding Perspectives on Open Science: Communities, Cultures and Diversity in Concepts and Practices

- Source: http://www.iospress.nl/book/expanding-perspectives-on-open-science-communities-cultures-and-diversity-in-concepts-and-practices/
- Synopsis from Source: "Continuing the tradition of bringing together academics, publishers, lecturers, librarians, developers, entrepreneurs, users and all other stakeholders interested in the issues surrounding electronic publishing, this edition of the conference focuses on Open Science, and the 27 research and practitioner papers and 1 poster included here reflect the results and ideas of researchers and practitioners with diverse backgrounds from all around the world with regard to this important subject. Intended to generate discussion and debate on the potential and limitations of openness, the book addresses the current challenges and opportunities in the ecosystem of Open Science, and explores how to move forward in developing an inclusive system that will work for a much broader range of participants. It will be of interest to all those concerned with electronic publishing, and Open Science in particular."

Open Science: Sharing Knowledge in the Global Century

- Source: http://www.publish.csiro.au/book/6195/
- Synopsis from Source: "*Open Science* offers practical ways to communicate science in a highly networked world where billions of people still have little or no access to advanced knowledge or technologies. The authors describe low-cost, effective means to transfer knowledge to target audiences in industry, government, the community and to the public at large. The book features sections on good science writing, practical advice on how to develop communication and media strategies, ways to measure communication performance, how to handle institutional 'crises', how to deal with politicians and much more."

European Commission, European Union (EU)

Open innovation, open science, open to the world

- Source: http://dx.doi.org/10.2777/061652

- Synopsis from Source: "This book brings together some of the key conceptual insights behind Open Innovation, Open Science and Open to the World and highlights actions that are already taking place or are being prepared."

Science Ecosystem 2.0

- Source: http://dx.doi.org/10.2777/67279
- Synopsis from Source: "The report analyses the potential impact of a transition towards Open Science on the stakeholders of the research ecosystem."

Organisation for Economic Cooperation and Development (OECD)

Open and Inclusive Collaboration in Science: A Framework

- Source: Global Science Forum. Draft for discussion. OECD Headquarters, 13-14 October 2016.
- Synopsis from Source: "The challenge then is to develop a holistic framework for open science based on its 3 basic elements: scientific process, key actors and digitalization." (Page 3)

Making Open Science a Reality

- Source: http://dx.doi.org/10.1787/5jrs2f963zs1-en
- Synopsis from Source: "This report, *Making open science a reality* reviews the progress in OECD countries in making the results of publicly funded research, namely scientific publications and research data openly accessible to researchers and innovators alike. The report i) reviews the policy rationale behind open science and open data; ii) discusses and presents evidence on the impacts of policies to promote open science and open data; iii) explores the legal barriers and solutions to greater access to research data; iv) provides a description of the key actors involved in open science and their roles; and finally v) assesses progress in OECD and selected non-member countries based on a survey of recent policy trends."

F1000Research

Guide to Open Science Publishing

- Source: https://blog.f1000.com/2015/03/06/guide-to-open-science-publishing/
- Synopsis from Source: "It contains information about open access, open peer review, post-publication peer review, open data, and other aspects of open science, and it was based on a series of blog posts we published last year."

Avoin tiede ja tutkimus (ATT) -hanke / Open Science and Research Initiative, Ministry of Education and Culture, Finland

The Open Science and Research Handbook

- Source: http://openscience.fi/handbook
- Synopsis from Source: "The Open Science and Research Handbook aims to help researchers, research organisations, decision-makers, financiers, and the general public promoting the adoption and use of open science and research. This handbook directs all these communities to further develop practical means to facilitate this process."

Open science and research leads to surprising discoveries and creative insights: Open science and research roadmap 2014–2017

- Source: http://julkaisut.valtioneuvosto.fi/handle/10024/75210

- Synopsis from Source: “This roadmap is based on the work of the Open Science and Research Initiative (ATT), a cross-administrative initiative established by the Ministry of Education and Culture, whose goal is to promote open science and the availability of information.”

Centre National de la Recherche Scientifique (CNRS) / National Center for Scientific Research

Open Science in a Digital Republic

- Source: http://dx.doi.org/10.4000/books.oep.1635
- Synopsis from Source: “This White Paper gives an account of these reflections on the practices of researchers with regard to the use of scientific and technical information and digital tools. The package of proposals for the creation of Open Science is the result of combined efforts and powerful testimonies from the world of research.”

Committee on Open Science, Science Council of Japan (SCJ)

Recommendations Concerning an Approach to Open Science That Will Contribute to Open Innovation

- Source: http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-23-t230-en.pdf
- Synopsis from Source: “In 2010 the Science Council of Japan recommended the establishment of a ‘comprehensive scientific journal consortium’ for the purposes of resolving issues relating to scientific journals, while in 2015 it issued ‘Recommendations for the 5th Science and Technology Basic Plan.’ In response to subsequent developments, such as studies conducted by Cabinet Office’s Expert Panel on Open Science based on Global Perspectives, this Committee was launched with a focus on crafting a vision for ‘openness in research data’ and ‘data sharing.’”

Expert Panel on Open Science, Global Perspectives Cabinet Office, Government of Japan

“Promoting Open Science in Japan: Opening up a new era for the advancement of science”

- Source: http://www8.cao.go.jp/cstp/sonota/openscience/150330_openscience_summary_en.pdf
- Synopsis from Source: “It is vital for Japan to participate in international discussions and to demonstrate a proactive approach to the promotion of open science. The Expert Panel on Open Science based on Global Perspectives has discussed various relevant issues of immediate importance for Japan. Based on these discussions, the Panel presented the guiding principles for promotion of open science in Japan.”

SPARC Europe + Digital Curation Center (DCC)

“A Snapshot of Open Data and Open Science Policies”

- Source: http://sparceurope.org/open-data-open-science-policy-europe/
- Synopsis from Source: “This report is the first in a series that will seek to shed light on similarities and differences in Open Data and Open Science policies between European nations, and to assess their effectiveness in opening research data.”

An Analysis of Open Data and Open Science Policies in Europe

- Source: http://sparceurope.org/new-sparc-europe-report-analyses-open-data-open-science-policies-europe/

- Synopsis from Source: “Providing an analysis of Open Data and Open Science policies across Europe, today we are releasing a new report. Produced in collaboration with the Digital Curation Center (DCC), it follows on the heels of a previous work that listed national research data policies. This latest companion piece takes that work several steps further, analysing the types of policies in place, their processes of creation, and some of their details. Included in the study are the 28 EU member states, as well as countries from the European Research Area, including Iceland, Norway, and Switzerland.

Ministry of Education, Culture and Science, Netherlands

National Plan Open Science

- Source: https://www.openscience.nl/binaries/content/assets/subsites-evenementen/open-science/national_plan_open_science_the_netherlands_february_2017_en_.pdf
- Synopsis from Source: “Many national initiatives have already been launched. A major boost is required if these initiatives are to be coordinated and the great ambition realised. That is why this Plan lists the ambitions and provides details of the parties intending to take action, as well as the timeframes within which they believe they can realise their objectives.”

European Grid Infrastructure (EGI)

Open Science Commons

- Source: https://documents.egi.eu/public/ShowDocument?docid=2410
- Synopsis from Source: “With this paper, the European Grid Infrastructure (EGI) proposes the Open Science Commons as a new approach to digital research, tackling policy challenges and embracing open science as a new paradigm for knowledge creation and collaboration. EGI invites organisations from the research landscape to join it in this journey to develop these concepts, and through them to advance the implementation of the European Research Area.”

The Royal Society

Science as an open enterprise

- Source: https://royalsociety.org/topics-policy/projects/science-public-enterprise/report/
- Synopsis from Source: “The Science as an open enterprise report highlights the need to grapple with the huge deluge of data created by modern technologies in order to preserve the principle of openness and to exploit data in ways that have the potential to create a second open science revolution. Exploring massive amounts of data using modern digital technologies has enormous potential for science and its application in public policy and business. The report maps out the changes that are required by scientists, their institutions and those that fund and support science if this potential is to be realised.”

Interagency Working Group on Open Science (IWGOS), Committee on Science (CoS), National Science and Technology Council (NSTC), The White House

Charter of the Interagency Working Group on Open Science

- Source: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/cos-iwgos-charter-1016-signed.pdf

- Synopsis from Source: “The purpose of the IWGOS is to advance Federal efforts to support open science by making the results of Federally funded scientific research more accessible and useful to the public, industry, and the scientific community.”

“Alchemy & algorithms: perspectives on the philosophy and history of open science”

- Source: https://riojournal.com/article/13593/
- Synopsis from Source: “Our aim was a critical assessment of the defining features, underlying narratives, and overall objectives of the contemporary open science movement. The event brought together contemporary open science scholars, publishers, and advocates to discuss the philosophical foundations and historical roots of openness in academic research.”

“History, ontology, science studies: How to study open science and scientific data”

- Source: https://easst.net/article/history-ontology-science-studies-how-to-study-open-science-and-scientific-data/
- Synopsis from Source: “It might be fruitful, if we want to reflect on our own methods as scientists, to look at our ways how we define data, or open science (or anything else as a matter of fact) and at the causes of selection of things that seem worthy of analysing. Also, I propose to analyse to greater extent the ontology of data or open science: what is regarded as data or open/closed science, which scientific, methodological or other traditions influence how these notions came into being in a specific scientific discipline at a specific time in history, at a specific place on Earth.”

“Opening science: towards an agenda of open science in academia and industry”

- Source: https://link.springer.com/article/10.1007/s10961-014-9375-6
- Synopsis from Source: “So far, the impact and implications of the general tendency towards more openness in academic and industrial science at the very front-end of the innovation process have been mostly neglected. Our paper presents a conceptualization of this open science as a new research paradigm. Based on empirical data and current literature, we analyze the phenomenon and propose four perspectives of open science. Furthermore, we outline current trends and propose directions for future developments.”

“From Science 2.0 to Open Science - Turning rhetoric into action?”

- Source: http://stcsn.ieee.net/e-letter/stcsn-e-letter-vol-3-no-1/from-science-2-0-to-open-science
- Synopsis from Source: “Policy rhetoric is furthermore closely linking Open Science to a specific concept of innovation and the potential of economic growth and targeting all markets including the job market. While reading the respective work programme one gains the impression Open Science might also just serve as empty container for still to be defined concepts, as it appears only twice in the document, without being clearly defined.”

Appendix D: Open Science Resource List – Websites

Selection Criteria: This resource list presents different platforms for groups understanding and advancing open science.

European Commission, European Union (EU)

"Open Science"

- Source: https://ec.europa.eu/research/openscience/index.cfm
- Platform for latest EU Open Science news and events.

Facilitate Open Science Training for European Research (FOSTER)

"The Foster Portal"

- Source: https://www.fosteropenscience.eu/
- Synopsis from Source: "The FOSTER portal is an e-learning platform that brings together the best training resources addressed to those who need to know more about Open Science, or need to develop strategies and skills for implementing Open Science practices in their daily workflows."

Science & Technology Section, Association of College & Research Libraries

"Open Science and Crowd Science: Selected Sites and Resources"

- Source: http://dx.doi.org/10.5062/F48913SM
- Science and Technology Resources on the Internet. Issues in Science and Technology Librarianship Spring 2012.

opensource.com

"What is open science?"

- Source: https://opensource.com/resources/open-science
- Open science history and context.

Year of Open 2017

"May Open Perspective: What is Open Science?"

- Source: https://www.yearofopen.org/may-open-perspective-what-is-open-science/
- Synopsis from Source: "In addition, we argued that it is important to consider the term "open" in relation to values like accessibility and responsibility. Up till now, focus has been on the accessibility aspect of open science. The responsibility aspect can and should be strengthened."

The OpenScience Project

- Source: http://openscience.org/
- Personal blog of Dan Gezelter

openscienceASAP

- Source: http://openscienceasap.org/
- Synopsis from Source: "openscienceASAP is a [community] platform for specific activities on Open Science."

Scholarly Publishing and Academic Resources Coalition (SPARC)

- Source: https://sparcopen.org/
- Synopsis from Source: "SPARC is a global coalition committed to making Open the default for research and education. SPARC empowers people to solve big problems and make new discoveries through the adoption of policies and practices that advance Open Access, Open Data, and Open Education."

Center for Open Science

- Source: https://cos.io/
- Synopsis from Source: "Our mission is to increase openness, integrity, and reproducibility of research. These are core values of scholarship and practicing them is presumed to increase the efficiency of acquiring knowledge. For COS to achieve our mission, we must drive change in the culture and incentives that drive researchers' behavior, the infrastructure that supports their research, and the business models that dominate scholarly communication."

Open Knowledge International

- Source: https://okfn.org/
- Synopsis from Source: "Open Knowledge International is a global non-profit organisation focused on realising open data's value to society by helping civil society groups access and use data to take action on social problems. Open Knowledge International does this in three ways: 1.) We show the value of open data for the work of civil society organizations; 2.) We provide organisations with the tools and skills to effectively use open data; 3.) We make government information systems responsive to civil society. Open Knowledge International is a worldwide network of people passionate about openness, using advocacy, technology and training to unlock information and enable people to work with it to create and share knowledge.

Open Science Federation

- Source: http://opensciencefederation.com/
- Synopsis from Source: "The Open Science Federation is a nonprofit alliance working to improve the conduct and communication of science. We are scientists and citizen scientists, writers, journalists, and educators, and makers of and advocates for Open Data, Open Access, and Open Source and Standards."

Global Young Academy

"Open Science"

- Source: https://globalyoungacademy.net/activities/open-science/
- Synopsis from Source: "The Working Group on Open Science aims to inform current transformations in publication systems, institutions and technologies by (1) garnering and voicing young researchers' views on which scientific outputs should be disseminated, how, to whom and with which expectations; (2) investigating the challenges and opportunities involved in implementing Open Science mandates across highly diverse research environments; (3) promoting Open Science mandates across the GYA membership and collaborators."

Center for Open Data Enterprise

- Source: http://www.opendataenterprise.org/

- Synopsis from Source: “The Center for Open Data Enterprise works to improve the open data ecosystem in three ways. In the U.S. and internationally, we host interactive roundtables, conduct user-centered research, and provide tools and trainings for greater application of open data.”

Open Science Prize

- Source: https://www.openscienceprize.org/
- Synopsis from Source: “The Open Science Prize is a new initiative from the Wellcome Trust, US National Institutes of Health and Howard Hughes Medical Institute to encourage and support the prototyping and development of services, tools and/or platforms that enable open content – including publications, datasets, code and other research outputs – to be discovered, accessed and re-used in ways that will advance research, spark innovation and generate new societal benefits.”

Board on Research Data and Information (BRDI), Policy and Global Affairs Division, The National Academies of Sciences, Engineering, and Medicine

“Toward an Open Science Enterprise”

- Source: http://sites.nationalacademies.org/pga/brdi/open_science_enterprise/index.htm
- Synopsis from Source: “An ad hoc committee under the Board on Research Data and Information (BRDI) will conduct a study on the challenges of broadening access to the results of scientific research, described as ‘open science.’ […] The committee will produce a consensus report with findings and recommendations that address these issues, with the majority of the focus on solutions that move the research enterprise toward open science.”

Global Open Access Portal (GOAP), United Nations Educational, Scientific and Cultural Organization (UNESCO)

“Open Science Movement”

- Source: http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/open-science-movement/
- Explanation, historical principles, and the current situation of the movement in Africa, Asia, Europe, North America, Latin America, and the Caribbean.

The European Open Science Cloud for Research Pilot Project

- Source: https://eoscpilot.eu/
- Synopsis from Source: “The European Open Science Cloud will offer 1.7 million European researchers and 70 million professionals in science and technology a virtual environment with open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines by federating existing scientific data infrastructures, today scattered across disciplines and Member States. The EOSCpilot project has been funded to support the first phase in the development of the European Open Science Cloud (EOSC).”

Ministry of Science, Technology and Higher Education, Portugal

“Ciência Aberta” / “Open Science”

- Source: http://www.ciencia-aberta.pt/home

- Synopsis from Source: “The Open Science webpage gathers information, initiatives and training content. Its target audience is everyone involved in the national scientific system and society in general. Following the principles of Open Science this is a collaborative project, in development, done for the community and with the community.”

Committee on Data (CODATA), International Council for Science (ICSU)

“Open Science in Africa and in South Africa”

- Source: http://www.codata.org/index.php?mact=News,cntnt01,detail,0&cntnt01articleid=151&cntnt01returnid=62
- Synopsis from Source: “The Science International Accord on Open Data in a Big Data World was launched at the Science Forum South Africa (SFSA) in December 2015. It was fitting that the next edition of SFSA should see the launch of an African Open Science Platform intiatiative [sic] that seeks to put into practice the principles and enabling practices for Open Data and Open Science that are laid out in the Accord.”

Open Science Fair

- Source: http://www.opensciencefair.eu/
- Synopsis from Source: “Open Science Fair will critically showcase the elements required for the transition to Open Science: e-infrastructures and services, policies as guidance for good practices, research flows and new types of activities (disseminate, mine, review, assess, etc.), the roles of the respective actors and their networks. OSFair2017 is organized as an emblematic initiative of four EU projects in the area of Open Science: OpenAIRE, OpenUP, FOSTER and OpenMinTeD. It is locally curated by the National Kapodestrian University of Athens and the ‘Athena’ Research and Innovation Center and co-sponsored by the National Library of Greece.”

Open Science Data Cloud

- Source: https://www.opensciencedatacloud.org/
- Synopsis from Source: “The Open Science Data Cloud provides the scientific community with resources for storing, sharing, and analyzing terabyte and petabyte-scale scientific datasets. The OSDC is a data science ecosystem in which researchers can house and share their own scientific data, access complementary public datasets, build and share customized virtual machines with whatever tools necessary to analyze their data, and perform the analysis to answer their research questions. It is a one-stop shop for making scientific research faster and easier.”

Open Science Grid (OSG)

- Source: https://www.opensciencegrid.org/
- Synopsis from Source: “The OSG facilitates access to distributed high throughput computing for research in the US. The resources accessible through the OSG are contributed by the community, organized by the OSG, and governed by the OSG consortium.”

Appendix E: Open Science Resource List – Visualizations

Selection Criteria: This resource list presents different ways that open science has been communicated visually by different groups.

European Commission, European Union (EU)

Open innovation, open science, open to the world

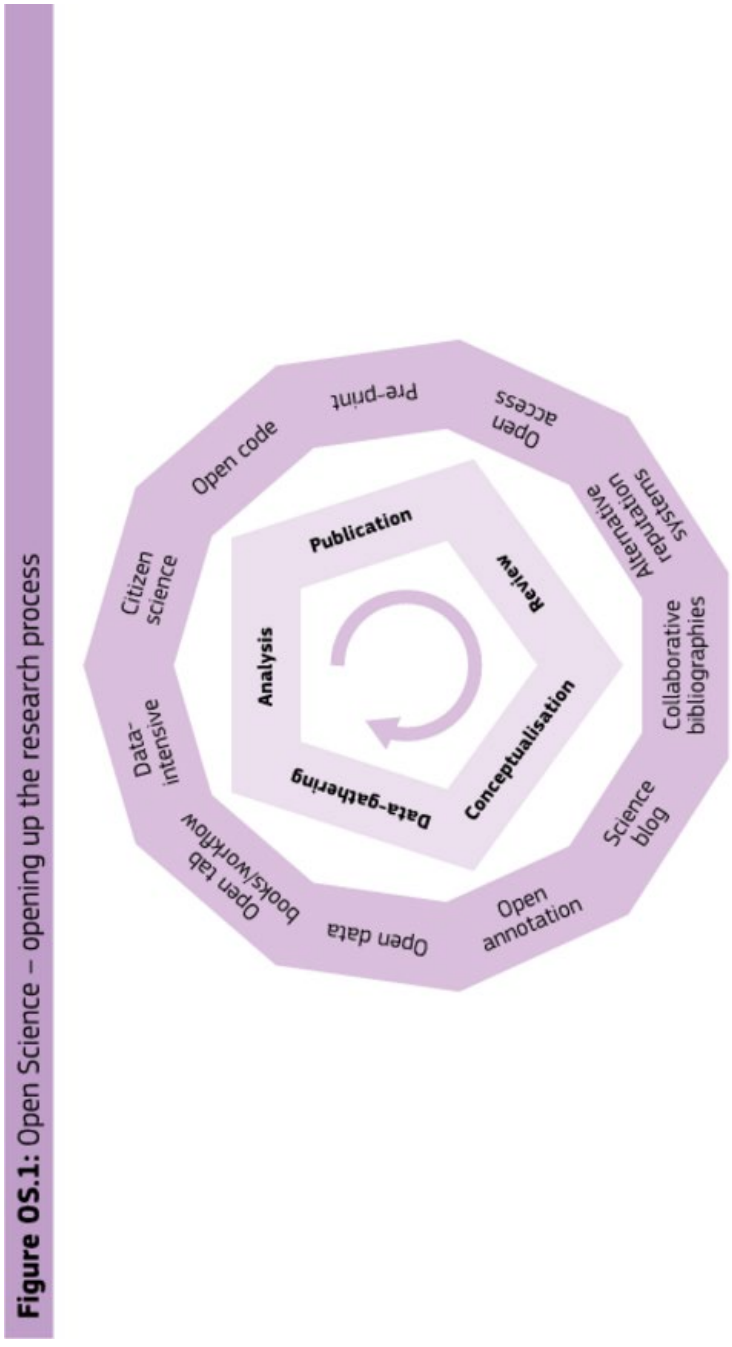

Figure 1: Figure OS.1: Open Science – opening up the research process

- Source: http://dx.doi.org/10.2777/061652
- Page 36

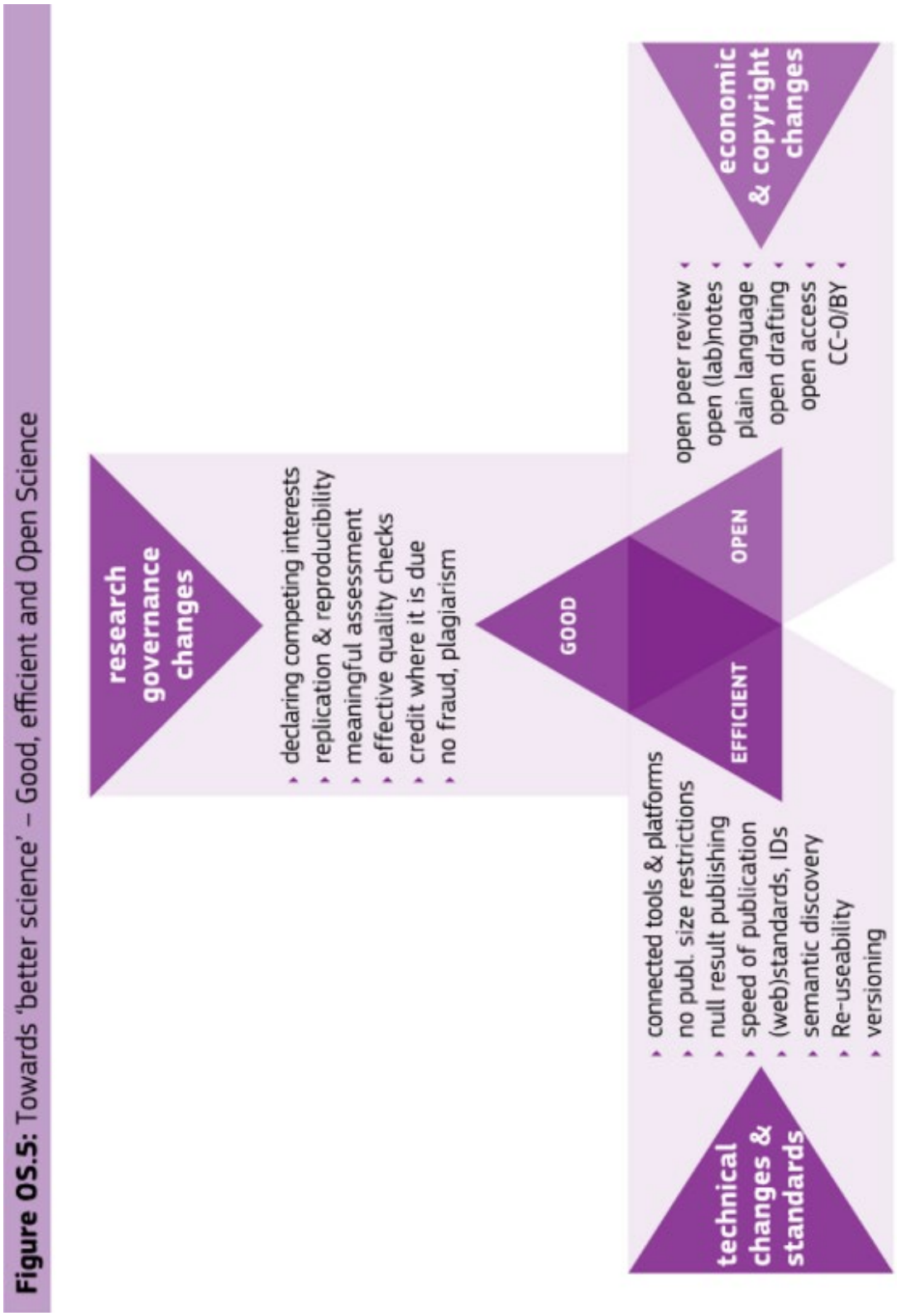

Figure 2: Figure OS.5: Towards 'better science' – Good, efficient and Open Science

- Source: http://dx.doi.org/10.2777/061652
- Page 55

Science Monitor

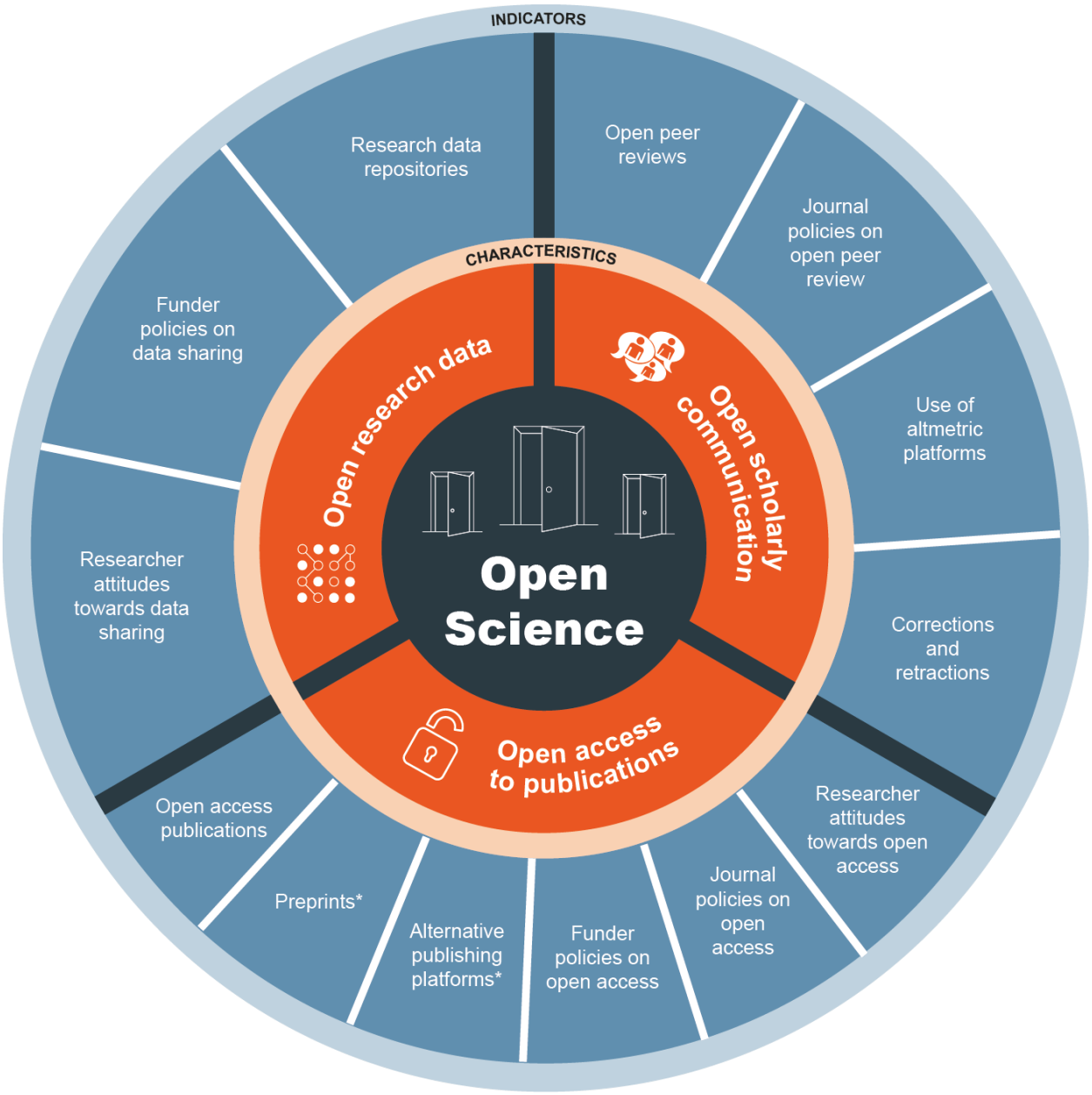

Figure 3: Open Science Wheel

- Source: http://ec.europa.eu/research/openscience/index.cfm?pg=home&section=monitor

Facilitate Open Science Training for European Research (FOSTER)

The FOSTER Portal

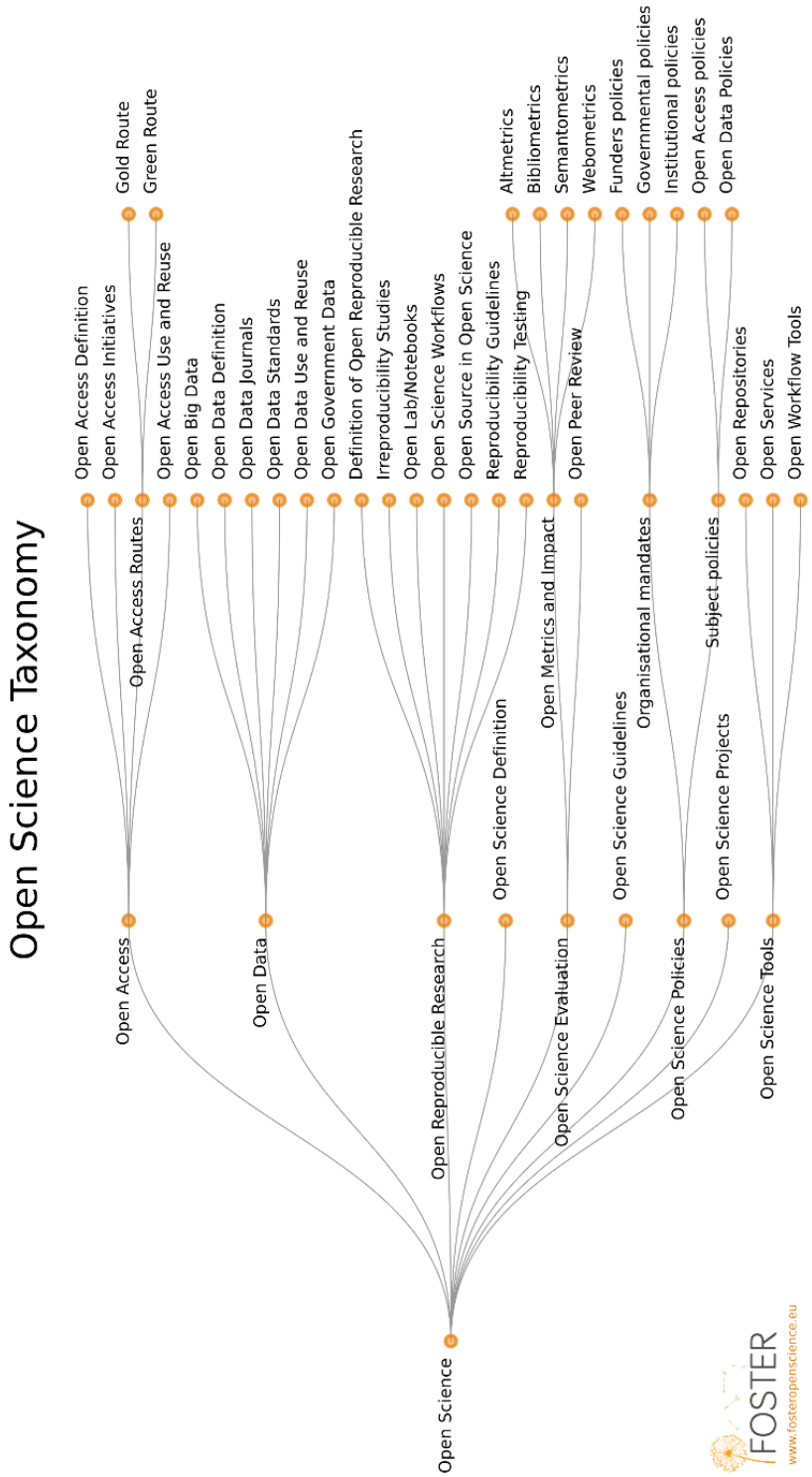

Figure 4: Open Science Taxonomy Tree

- Source: https://www.fosteropenscience.eu/foster-taxonomy/open-science

Figure 5: Open Science facets as a beehive

- Source: https://www.fosteropenscience.eu/content/what-open-science-introduction

Figure 1: Benefits to different parties.

Figure 6: Figure 1: Benefits to different parties.

- Source: http://openscience.fi/handbook
- Page 4

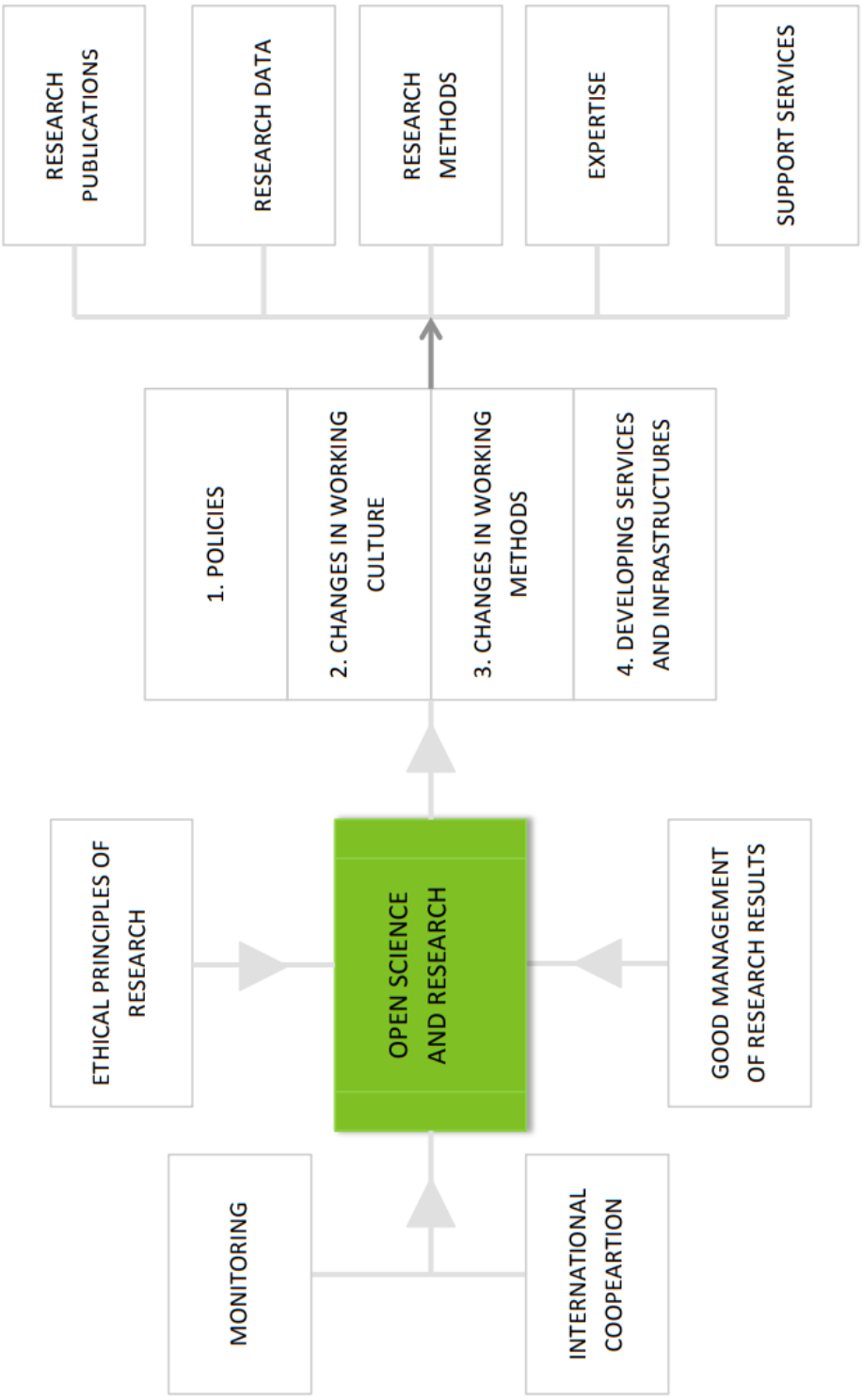

Figure 3: The entire process of the OSR promotion.

Figure 7: Figure 3: The entire process of the OSR promotion.

- Source: http://openscience.fi/handbook
- Page 6

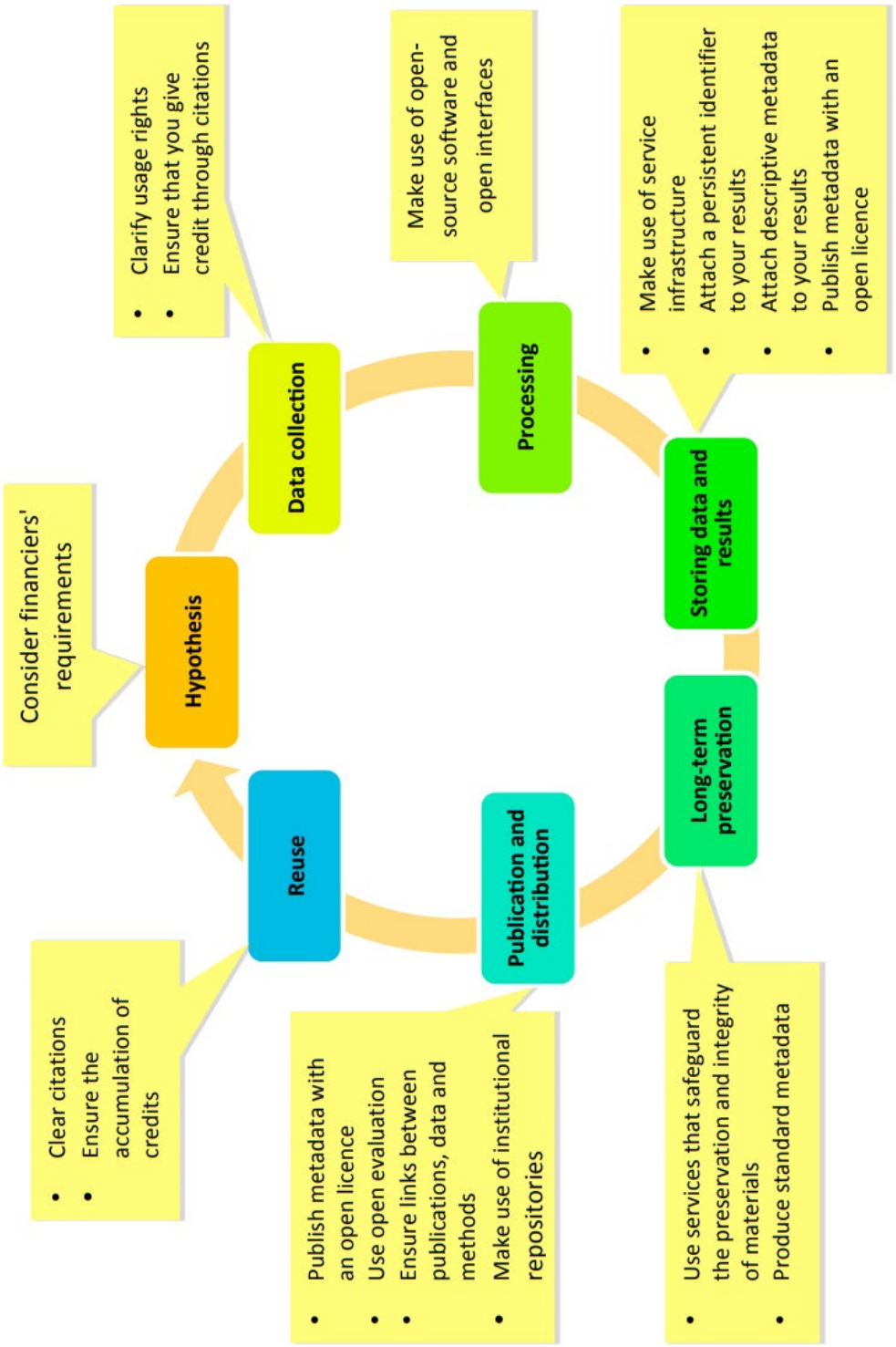

Figure 4: Promoting openness at different stages of the research process.

Figure 8: Figure 4: Promoting openness at different stages of the research process.

- Source: http://openscience.fi/handbook
- Page 8

Opening Science: The Evolving Guide on How the Web is Changing Research, Collaboration and Scholarly Publishing.

"Open Science: One Term, Five Schools of Thought"

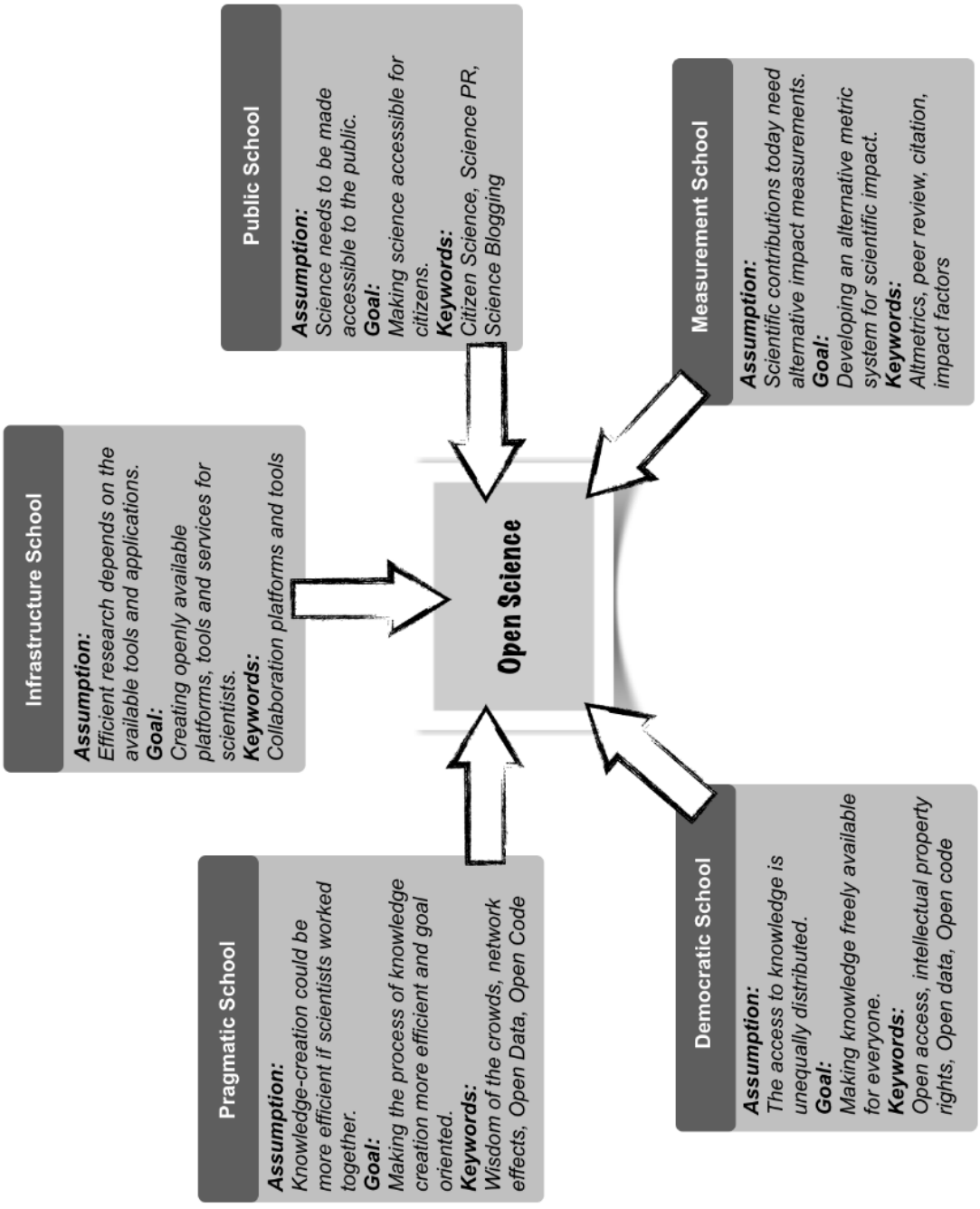

Figure 9: Table 1. Five Open Science schools of thought.

- Source: http://book.openingscience.org/basics_background/open_science_one_term_five_schools_of_thought.html

Ministry of Science, Technology and Higher Education, Portgual

Ciência Aberta / Open Science

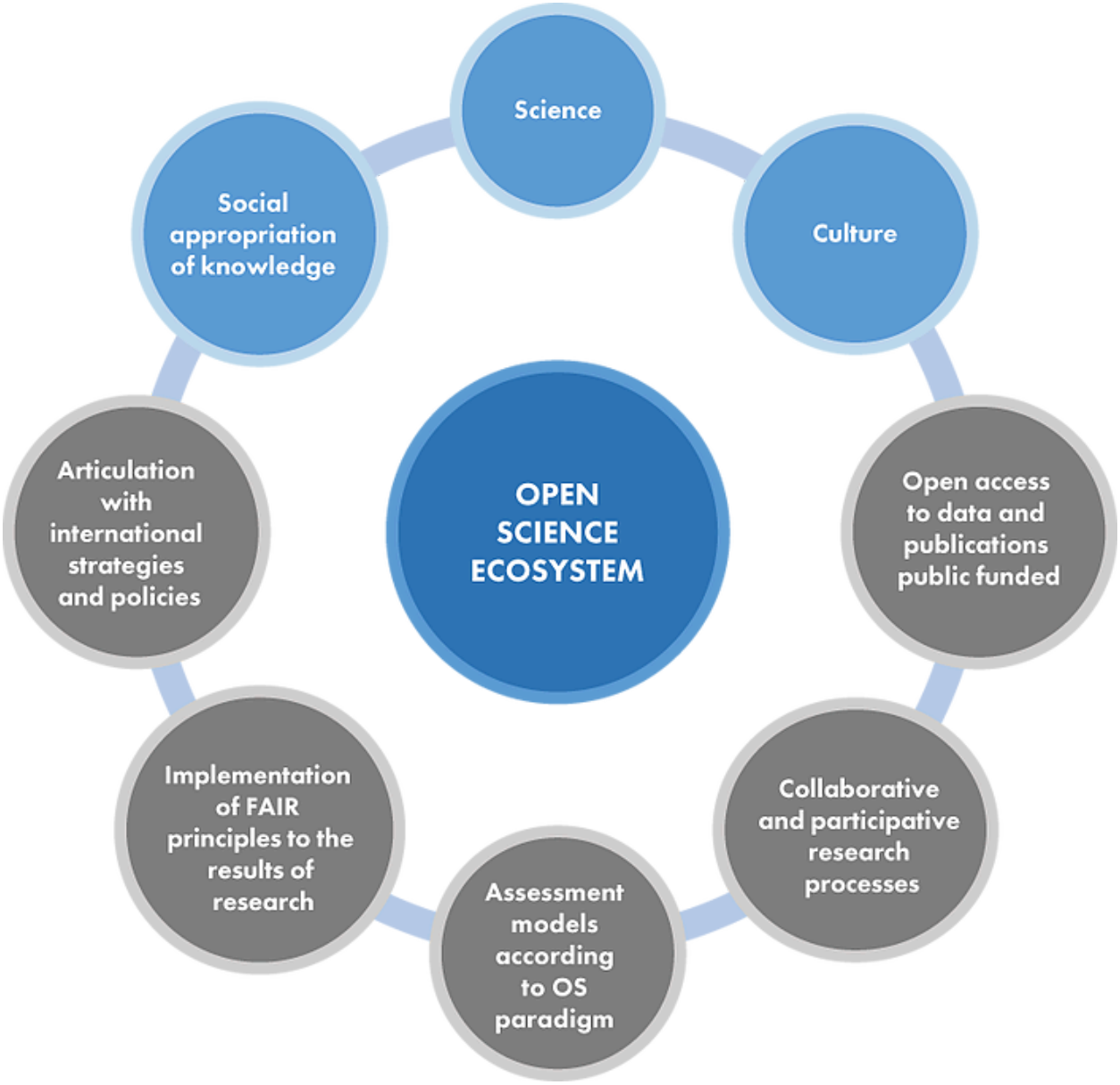

Figure 10: Open Science Ecosystem

- Source: http://www.ciencia-aberta.pt/about-open-science