

Open Research Glossary

Jon Tennant and Ross Mounce

January 2020

This resource is based on the original created by [Jon Tennant](#) and [Ross Mounce](#).

This glossary is designed to help to inform people about the culture of “open research”. It was written by the community, for the community, and depends on the community to stay current.

Core Definitions underlying Open Research

- **Citizen Science** - Citizen science is scientific research conducted, in whole or in part, by amateur (or nonprofessional) scientists. Citizen science is sometimes described as “public participation in scientific research,” participatory monitoring, and participatory action research whose outcomes are often advancements in scientific research, as well as an increase in the public’s understanding of science; [Source](#).
- **Open Access (OA)** - Making peer reviewed scholarly manuscripts freely available via the Internet, permitting any user to read, download, copy, distribute, print, search, or link to the full text of these articles, crawl them for indexing, pass them as data to software, or use them for any lawful purpose, without financial, legal or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited. May also refer to theses, books, book chapters, monographs and other content; [Source](#).
- **Open Data (OD)** - Making data freely available on the public internet permitting any user to download, copy, analyse, re-process, pass them to software or use them for any other purpose without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself; [Source](#).
- **Open Educational Resources (OER)** - Making high quality, openly licensed, online educational materials available for sharing, use, and reuse. They act as a mechanism for instructional innovation as networks of teachers and learners share best practices; [Source](#).
- **Open Lab Notebooks (OLN)** - A concept of blogging about research on a regular basis, such that research notes and data are accumulated and published online as soon as they are obtained; [Source](#).
- **Open Materials (OM)** - Providing access to materials made by and for the wider research community. This can include elements from experiments, tasks, and

questionnaires made by researchers, and helps to increase the transparency, accessibility, and reproducibility of published research.

- **Open Peer Review (OPR)** - An umbrella term for a number of overlapping ways that peer review models can be adapted in line with the aims of Open Science, including making reviewer and author identities open, publishing review reports and enabling greater participation in the peer review process; [Source](#).
- **Open Source Software (OSS)** - Availability of source code for a piece of software, along with an [Open Source license](#) permitting reuse, adaptation, and further distribution; [Source](#).
- **Scholarly Communication** - The creation, transformation, dissemination, and preservation of knowledge related to teaching, research, and scholarly endeavors; the process of academics, scholars and researchers sharing and publishing their research findings so that they are available to the wider academic community; [Source](#).

Scholarly publishing

- **Article Processing Charge (APC)** - A fee charged to the author, creator, or institution to cover the cost of publishing an article, rather than charging the potential reader of the article. APCs may apply to both subscription-based and Open Access publications. APCs are sometimes charged to authors in order to cover the cost of publishing and disseminating an article in an Open Access scholarly journal; [Source](#).
- **Citation** - A reference to a published or unpublished source embedded in content, for the purposes of acknowledging the work and relevance of others to the topic of discussion where the citation appears.
- **Copy editing** - A type of editing designed to improve the formatting, style, and accuracy of text. It usually does not involve changing the content of the original text.
- **Curation** - The selection, preservation, maintenance, collection and archiving of [digital] assets. Curation establishes, maintains, and adds value to repositories of digital data for present and future use; [Source](#).
- **Double-dipping** - In the context of OA, double-dipping occurs when a journal has an APC for publishing an author's work, as well as requiring payment (usually through a subscription fee) by the potential user of the work. This model makes the institution or author pay twice to access the work; [Source](#).
- **Embargo period** - A length of time imposed on a research output for users who have not paid for access, or do not have institutional access, before it is made freely available. Typically refers to peer-reviewed versions of manuscripts.
- **Fee waiver** - If an institution, research funder or author cannot pay for an APC, many publishers or journals will offer partial or total waiving for fees; [Source](#).

- **Legacy publisher** - A publisher that historically has operated primarily on a paywall-based business model.
- **Library-based publishing** - Many academic libraries are now beginning to act as publishers for scholarly works produced in their institutions and elsewhere. In some cases, the library works with the university scholarly press to publish works. In other cases, the library publishes works independently or separately from the academic press. Library-based publishers are often strongly in favor of Open Access; [Source](#).
- **Loginwall** - A barrier that prevents access by requiring users to register and login to a service, without asking for money to unlock access.
- **Open Access publisher** - A publisher that publishes all research articles as Open Access articles. Most legacy publishers have options to make journals at least partially Open Access (i.e., the 'hybrid' model).
- **Paywall** - Restriction via a financial barrier to access research, often implemented by legacy publishers. Can be removed by personal or institutional subscriptions.
- **Publisher** - A company whose purpose is to make the outputs of research publicly available, via digital or analogue methods.
- **Publishing** - To make a research output available to the public. Commonly refers to the release of works by publishers, irrespective of whether public access is granted or not.
- **References** - Defines a relationship between one object, a designator, and a second object, a source. Usually takes the form of a bibliography of academic papers at the end of a research manuscript.
- **Repository, general** - An online archive to deposit manuscripts. These can be personal websites, or include platforms such as [ResearchGate](#) or [edu](#).
- **Repository, institutional** - An online database designed to collect the intellectual output of a particular institution or university, including digital collections such as electronic theses and dissertations (ETDs), preprints, or faculty scholarship, and presents associated metadata regarding the these items; [Source](#).
- **Repository, software** - An online collection of files managed with version control software (e.g., bzd, hg, git, csv, svn, etc.). Can be hosted by third-party (e.g., [GitHub](#), [Bitbucket](#), [SourceForge](#)), by an institution, or self-hosted locally.
- **Repository, subject-based** - A thematic repository based around research disciplines, geographic regions, or languages, such as [arXiv](#).
- **Self-archiving** - Making a copy of a manuscript available through a personal website, institutional repository, or other repository. This often happens in parallel with publication at a journal.

- **Submission fee** - A fee levied by some publishers for submitting a manuscript to their journals.
- **Subscription** - A form of business model whereby a fee is paid in order to gain access to a product or service - in this case, the outputs of scholarly research in journals.
- **Toll access** - A system whereby a fee is required to pass a paywall to access research.
- **Typesetting** - The composition of text by arranging physical pieces of type or by using software to prepare a version of the text suitable for printing. Stored letters and other symbols are retrieved and ordered according to a language's orthography (conventional spelling system of a language) for visual display.

Journal or publication Types

- **Book** - A medium for recording information in the form of writing or images, typically composed of many pages (made of papyrus, parchment, vellum, or paper) bound together and protected by a cover; [Source](#).
- **Epub** - A free and open e-book standard by the [International Digital publishing Forum](#).
- **Hybrid journal** - Some traditional journals offer an option for authors to make their individual articles freely accessible to anyone worldwide, for an additional fee. All other articles in the journal remain accessible only through subscription; [Source](#).
- **Journal** - An aggregation of published research articles, either digital or printed or both. Historically divided into volumes and issues.
- **Megajournal** - A journal with editorial criteria based on scientific soundness instead of a priori estimated newsworthiness or "impact". They typically cover all research disciplines.
- **Monograph** - A specialist work of writing or exhibition on a single subject or an aspect of a subject, often by a single author or artist, and usually on a scholarly subject; [Source](#).
- **Open Access journal** - A journal that exclusively comprises Open Access articles.
- **Overlay journals** - An OA journal that does not produce its own content, but selects and curates groups of articles that are already freely available online. An example of this is [Discrete Analysis](#); [Source](#).

Types of work often made Open Access

- **Accepted author manuscript (AAM)** - The version of a manuscript that has been formally accepted by a publisher for publication.
- **Eprint** - A digital version of a research document available online within a repository.

- **Journal to Wiki publication (J2W)** - Copying text from a published paper to a wiki (such as Wikipedia or Wikibooks), with attribution: legally possible if the licence of the paper is less restrictive than the licence of the wiki.
- **Postprint** - A manuscript draft after it has been peer reviewed, but before it has undergone further modification by a publisher.
- **Preprint** - A manuscript draft that has not yet been subject to formal peer review, distributed to receive early feedback on research from peers.
- **Preregistration**- The practice of registering a scientific study before it is conducted. Preregistration of studies serves to prevent publication bias and reduce data dredging. It arose as a means to address the replication crisis. Preregistration requires the submission of a registered report, which is then accepted for publication or rejected by a journal based on theoretical justification, experimental design, and the proposed statistical analysis; [Source](#).
- **Registered Reports** - A type of publication in which peer review of the suggested methods and protocol is completed prior to data collection and analysis. Accepted papers then are guaranteed publication in the journal if the authors follow through with the registered methodology; [Source](#).
- **Version of Record (VOR)** - The final published version of a manuscript in a journal, after peer review and processing by a publisher.
- **Wiki to Journal publication (W2J)** - Creating a paper on a wiki, using its features for collaboration and informal review, for submission to a journal for formal peer review. Might involve a public wiki such as Wikipedia or Wikiversity, or a specially-created wiki.

‘Flavours’ or ‘types’ of Open Access

- **Black** - Refers to the illegal/illicit sharing of copies of research articles via channels such as ResearchGate or Sci-Hub. “Black as in the classical pirate flag, or in black market!”; [Source](#).
- **Bronze** - Delayed OA journals publish articles initially as subscription-only, then release them as free to read (but not to reuse, adapt and share, so not open access), typically after an embargo period (varying from months to years). In this way subscribers get early access to content and it is not licensed for reuse; [Source](#).
- **Diamond** - A form of “Gold” OA in which there is no author-facing fee (APC). Sometimes also called ‘Platinum’ OA.
- **Gold** - Making the final version of manuscript (VOR) freely available immediately upon publication by the publisher.
- **Green** - Making a version of the manuscript freely available in a repository. Often referred to as ‘self-archiving’, and happens in parallel with publication in a journal.

- **Gratis** - A paper is available to read free-of-charge, though its reuse is still restricted, for example by “All Rights Reserved” or" copyright; [Source](#).
- **Libre** - A paper is made available under an open licence, allowing it to be shared and reused, depending on which licence is used; [Source](#)). Both ‘Libre’ and ‘Gratis’ refer to copyright and licensing restrictions. The difference is often alluded to with the slogan, “Free as in beer, not as in speech.”

Declarations And Principles

- **1964** - [Declaration of Helsinki](#).
- **2001** - [Declaration of Havana Towards Equitable Access to Health Information](#).
- **2002** - [Budapest Open Access Initiative \(BOAI\)](#).
- **2003** - [Bethesda Statement on Open Access Publishing](#).
- **2003** - [Declaration of Berlin on Open Access to Knowledge in the Sciences and Humanities](#).
- **2007** - [Kronberg Declaration on the Future of Knowledge Acquisition and Sharing](#).
- **2007** - [Declaration of Brisbane](#).
- **2009** - [Panton Principles for Open Data](#).
- **2012** - [Cost of Knowledge Manifesto](#).
- **2013** - [The San Francisco Declaration on Research Assessment \(DORA\)](#).
- **2016** - [Vienna Principles](#).
- **2017** - [OCSDNet Open Science Manifesto](#).
- **2018** - [Declaración de Panamá de Open Science](#).

Licensing and copyright

- **Apache License** - A free software license by the Apache Software Foundation; [Source](#).
- **Author addendum** - An author addendum is a supplemental or added agreement to a publishing contract that defines or changes the terms of the contract, often focusing on the transfer of copyright ownership. For authors of scholarly works, an author addendum to a publisher’s standard publication contract may be necessary to help ensure that authors protect important rights, such as the right to post their articles online to a personal website or in a digital repository; the right to use their works within a classroom setting; or the right to use their works as the foundation for future research; [Source](#).

- **Berkeley Software Distribution (BSD)** - A family of UNIX-like operating systems; [Source](#).
- **Copyleft** - A form of licensing that makes a creative work freely available to be modified, and requiring all modified and extended versions of the creative work to be free as well. OA does not require works to be copyleft, nor does it necessarily exclude copyleft works from being OA. The recommended licence (CC BY) for academic publishing is not copyleft.
- **Copyright** - The aspect of Intellectual Property that gives creators the right to permit (or not permit) what happens to their creations, as opposed to trademark rights or moral rights.
- **Creative Commons (CC)** - A popular suite of licences that sets out the rights of authors and users, providing alternatives to the standard copyright. CC licences are widely used, simple to state, machine readable and have been created by legal experts. There are a variety of CC licences, each of which use one or more clauses, examples of which are given below. Some licences are compatible with OA in the 'Budapest sense', and some are not; [Source](#)).
- **CC Zero (CC 0)** - Waiver of copyright, with no rights reserved. Places content as openly as possible in the public domain; [Source](#).
- **CC Attribution (CC BY)** - A licence clause that allows the reuse, sharing, and remixing of materials providing the original author is appropriately attributed. Aside from attribution the CC BY licence has no other restrictions on copying or reuse. Compatible with free cultural works.
- **CC NonCommercial (CC NC)** - A licence clause allowing the reuse, sharing, and remixing of materials providing that it is for non-commercial purposes. Not compatible with free cultural works.
- **CC No Derivatives (CC ND)** - A licence clause requiring that derivatives are not made of the original works. Not compatible with free cultural works.
- **CC ShareAlike (CC SA)** - A licence clause requiring that derivative works have the same licence as the original. Compatible with free cultural works.
- **Creative work** - An original, identifiable piece of content, such as an academic paper, a diagram, a photograph, or a video clip. Owners of creative works have rights, such as copyright, that they might reserve to keep control of the content, or relinquish to allow others to share and reuse that content.
- **Derivative work** - A work based upon one or more pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which a work may be recast, transformed, or adapted; [Source](#).

- **GNU GPL (General Public License)** - A free copyleft license for software and other kinds of works; [Source](#).
- **Intellectual property (IP)** - A legal term that refers to creations of the mind. Examples of intellectual property include music, literature, and other artistic works; discoveries and inventions; and phrases, symbols, and designs.
- **Intellectual Property Rights (IPR)** - The rights given to the owners of intellectual property. IPR is protected either automatically (e.g., via copyright or design rights) or by registering or applying for it (e.g., trademarks, patents). Protecting your IP makes it easier to take legal action against anyone who steals or copies it. IPR can be legally sold, assigned or licensed by the creator to other parties, or jointly-owned.
- **MIT License** - A popular open and permissive software license; [Source](#).
- **Rehost** - The practice of moving content from one machine or location to another. This can involve deactivating the content at its source and then activating it again at the new new location. Rehosting is a way to preserve investments in hardware and software, while opening paths to future modernisation by moving to an open and more extensible architecture.
- **Redistribute** - Describes the means to copy a piece of software or content and make it available elsewhere.
- **Reuse** - - The practice of using a work, whether for its original purpose (conventional reuse) or to fulfil a different function (creative reuse or repurposing). Reuse – by taking, but not reprocessing, previously used items – helps save time, money, energy and resources. In broader economic terms, it can make quality products available to people and organizations with limited means, while generating jobs and business activity that contribute to the economy; [Source](#).

Peer Review

- **Annotations** - A comment with specific location and context, either inline or in the margin of a text document, or within a region of an image or video, or located within a specific row or cell of data in a data set.
- **Comments** - The enrichment of content, usually through the addition of text towards the end of an article. Can be compiled into comment sections for discussion on a published work.
- **Double-blind peer review** - When the reviewers do not know who the authors are, and vice versa.
- **Open review reports** - When individual reviews are made openly available, typically alongside the article after it is published.

- **Peer review** - A formal process by which a research article is vetted by experts in community before publication; [Source](#).
- **Portable peer review** - Independent peer review that travels with a manuscript that is submitted to subsequent different journals, designed to combat redundancy in the peer review process.
- **Post-publication peer review** - Standard peer review, but after a research article has been formally published. Can also be performed on preprints (e.g., through an Overlay Journal).
- **Signed peer review** - When individual reviews are publicly signed by those who conducted them.
- **Single-blind peer review** - Where reviewers know the identity of authors, but authors do not know the identity of reviewers.
- **Transferable peer review** - Reviews that travel with a paper if it is rejected from a journal.
- **Triple-blind peer review** - Where not only are authors and reviewers blind to each other's identities but where editors are also blind to the identities of both.

Assessment And Metrics

- **Altmetrics** - Altmetrics are "alternative" ways of recording and measuring the use and impact of scholarship. Rather than solely counting the number of times a work is cited in scholarly literature, alternative metrics also measure and analyze social media (e.g., Facebook, Twitter, blogs, wikis, etc.), document downloads, links to publishing and unpublished research, and other uses of research literature, in order to provide a more comprehensive measurement of scholarships reach and impact; [Source](#).
- **Article-level metrics** - All types of article-level metrics including download and usage statistics, citations, and article-level altmetrics; [Source](#).
- **Bibliometrics** - Bibliometrics is the branch of library and information science concerned with the application of mathematical and statistical analysis to bibliography. Bibliometrics involves the statistical analysis of books, articles, or other publications.
- **H-index** - A personal metric that relates the number of citations to the number of published papers for an academic. It measures the number of papers (n) that have been cited n times; [Source](#).
- **Impact factor** - A numerical measure that indicates the average number of citations to articles published over the previous two years in a journal, and frequently used as a proxy for a journal's relative importance. It is frequently mis-used to evaluate articles and individuals.

- **Journal-level metrics** - Metrics that apply to all papers published within a journal. A common example is the journal Impact Factor.
- **Leiden Manifesto** - A distillation of best practice in metrics-based research assessment so that researchers can hold evaluators to account, and evaluators can hold their indicators to account; [Source](#).
- **Metrics** - Measures of quantitative assessment commonly used for assessing, comparing, and tracking performance or production.
- **San Francisco Declaration on Research Assessment (DORA)** - Recognises the need to improve the ways in which the outputs of scholarly research are evaluated. The declaration was developed in 2012 during the Annual Meeting of the American Society for Cell Biology in San Francisco. It has become a worldwide initiative covering all scholarly disciplines and all key stakeholders including funders, publishers, professional societies, institutions, and researchers; [Source](#).
- **Scientometrics** - The field of study which concerns itself with measuring and analysing scientific literature. Scientometrics is a sub-field of bibliometrics. Major research issues include the measurement of the impact of research papers and academic journals, the understanding of scientific citations, and the use of such measurements in policy and management contexts; [Source](#).
- **Source-Normalised Impact Per-Paper (SNIP)** - Measures contextual citation impact by weighting citations based on the total number of citations in a subject field. The impact of a single citation is given higher value in subject areas where citations are less likely, and vice versa. Unlike the well-known journal impact factor, SNIP corrects for differences in citation practices between scientific fields, thereby allowing for more accurate between-field comparisons of citation impact; [Source](#).

Organisations

Funders and policy-related

- **Gates Foundation** - A funding body co-ordinated by Melinda and Bill Gates, with a strong [Open Access policy](#).
- **Jisc** - An educational charity based in the UK Provides expertise to universities, colleges and cultural institutions on the use of technology to support research, including publication models, repositories, licensing, and infrastructure; [Source](#).
- **Max Planck Society** - A German research organisation with 82 Institutes worldwide and a strong [Open Access policy](#).
- **National Institute of Health (NIH)** - The national medical research agency in the USA, with a [PublicAccess policy](#).

- **National Science Foundation (NSF)** - An independent federal agency in the USA for the funding of research, with a [Public Access policy](#).
- **Plan S** - Plan S is an initiative for Open Access publishing that was launched in September 2018. The plan is supported by cOAlition S, an international consortium of research funders. Plan S requires that, from 2021, scientific publications that result from research funded by public grants must be published in compliant Open Access journals or platforms; [Source](#).
- **Projekt DEAL** - The German Rectors' Conference was tasked by the Alliance of German Science Organizations to institute Projekt DEAL to conclude nationwide "publish and read" agreements with the largest commercial publishers of scholarly journals. DEAL has negotiated on behalf of all German academic institutions including universities, universities of applied sciences, research institutions, state and regional libraries. All academic and research institutions that are entitled to participate in the Alliance and national licenses funded by the German Research Foundation (DFG) are thereby eligible to participate in DEAL agreements; [Source](#).
- **Public Knowledge Project (PKP)** - A multi-university initiative developing free, Open Source software and conducting research to improve the quality and reach of scholarly publishing; [Source](#).
- **Research Excellence Framework (REF)** - An initiative to assess the quality of research undertaken in UK higher education institutions. Coordinated by HEFCE.
- **Wellcome Trust** - A life sciences charity and funding body in the UK ([Open Access policy](#))

Open Research Infrastructure

- **AmeliCA** - A communication infrastructure for scholarly publishing and open science. Sustained cooperatively this initiative focuses on a non-profit publishing model to preserve the scholarly and open nature of scientific communication; [Source](#).
- **Connecting Repositories (CORE)** - A collection of OA repositories; [Source](#).
- **Confederation of Open Access Repositories (COAR)** - An international association that brings together individual repositories and repository networks in order to build capacity, align policies and practices, and act as a global voice for the repository community; [Source](#).
- **DARIAH** - A pan-european infrastructure for arts and humanities scholars working with computational methods. It supports digital research as well as the teaching of digital research methods; [Source](#).
- **Directory of Open Access Repositories (OpenDOAR)** - A directory of academic OA repositories. Also has a search function for repositories and repository contents; [Source](#).

- **Directory of Open Access Journals (DOAJ)** - A directory indexing OA peer-reviewed journals; [Source](#).
- **Directory of Open Access Books (DOAB)** - A directory indexing open access peer-reviewed books; [Source](#).
- **Europe PubMed Central (EuroPMC)** - Based on PubMed Central, and part of a network of repositories supported by funders of life sciences and biomedical research; [Source](#).
- **FOSTER** - 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; [Source](#).
- **OpenAIRE** - A pan-European infrastructure that supports the European Commission's OA Mandate in Horizon2020. All publications funded by the EC should be made available in Open Access and OpenAIRE harvests from a range of data sources namely repositories, OA publishers. ([Source](#))
- **OpenUP** - An open, dynamic and collaborative knowledge environment that systematically captures, organizes and categorizes research outcomes, best practices, tools and guidelines; [Source](#).
- **Open Scholarly Communication in the European Research Area for Social Science and Humanities (OPERAS)** - A project to coordinate and pool university-led scholarly communication activities in Europe in the Social Sciences and Humanities, in view of enabling Open Science as the standard practice.
- **PubMed** - A repository comprising more than 24 million citations for the biomedical literature; [Source](#).
- **PubMed Central (PMC)** - A free full-text archive of biomedical and life sciences journal literature at the US National Institutes of Health's Library of Medicine; [Source](#).
- **Repository 66** - A mash-up of data from ROAR and OpenDOAR overlaid onto Google maps; [Source](#).
- **Red de Revistas Científicas de América Latina y El Caribe, España y Portugal (REDALYC)** - A bibliographic database and a digital library of Open Access journals, supported by the Universidad Autónoma del Estado de México with the help of numerous other higher education institutions and information systems; [Source](#).
- **Registry of Open Access Repositories (ROAR)** - A registry for OA repositories, hosted by the University of Southampton, UK; [Source](#).
- **Scientific Electronic Library Online (SciELO)** - A programme started in Brazil in 1998 which has now expanded to 15 other countries, developed by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) and BIREME (Centro Latino-americano

e do Caribe em Informação em Ciências da Saúde). The objectives are to develop a common methodology for the preparation, storage and dissemination of scientific literature, including standardized evaluation and quality control processes. This comprises a model for cooperative electronic publication of scientific periodicals on the internet using organised bibliographic databases with full text access; Source, in [Portuguese](#) and in [English](#).

- **Securing a Hybrid Environment for Research Preservation and Access - Rights of METadata for Open archiving (SHERPA-RoMEO)** - A tool to check what the self-archiving policies for individual journals are; [Source](#).

Advocacy organisations

- **Open Access Scholarly Publishers Association (OASPA)** - Represents the interests of OA journal and book publishers in all scientific, technical, and scholarly disciplines; [Source](#).
- **Right to Research Coalition (R2RC)** - Founded by students in the summer of 2009 to promote an open scholarly publishing system based on the belief that no student should be denied access to the articles they need because their institution cannot afford the often high cost of access. Since its launch, the Coalition has grown to represent nearly 7 million students internationally and counts among its members the largest student organizations in both the United States and Canada. While the Coalition currently has a strong base in North America, it is by no means solely a North American organization and is expanding to incorporate student organizations from around the world; [Source](#).
- **Scholarly Publishing and Academic Resources Coalition (SPARC)** - An international alliance of academic and research libraries working to create a more open system of scholarly communication. Primarily based in the [USA](#), but now with branches in Europe, South Africa, and Japan.

Tools and platforms

Formats, identifiers and standards

- **Comma-Separated Values, or Character-Separated Values (CSV)** - A plain-text (non-binary) format for tabular data.
- **Digital Object Identifier (DOI)** - A unique text string that is used to identify digital objects such as journal articles or OSS releases; [Source](#).
- **Extensible Markup Language (XML)** - A language that defines a set of rules for encoding documents in a format that is readable by both machines and humans; [Source](#).
- **Hypertext Markup Language (HTML)** - The set of markup symbols or codes inserted in a file intended for display on a browser page; [Source](#).

- **Journal Article Tag Suite (JATS)** - A common XML format in which publishers and archives can exchange journal content; [Source](#).
- **LaTeX** - A markup language for typesetting documents, particularly common in mathematics and the sciences. Many academic journals accept submissions in LaTeX; [Source](#).
- **Machine readable** - Data or metadata in a format that can be understood by a computer.
- **Machine Readable Cataloguing (MARC)** - A set of digital formats for the description of items catalogued by libraries; [Source](#).
- **Markdown** - A syntax for adding formatting to documents allowing correctly formatted articles to be written in plain text; [Source](#).
- **Open Archives Initiative (OAI)** - Develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content; [Source](#).
- **ORCID** - A persistent digital identifier that distinguishes individual researchers. Also supports integration in research workflows; [Source](#).
- **ResearcherID** - Assigns a unique identifier for researchers to manager publication lists, track citations, and avoid author mis-identification; [Source](#).
- **Scalable Vector Graphics (SVG)** - A format for images that is open rather than tied to particular software, resolution-independent (unlike GIF, PNG and JPG), and structured so that with appropriate software it is relatively easy, for example, to translate labels into different languages.
- **Uniform Resource Identifier (URI)** - A string of characters used to identify a name of a resource to enable its digital and networked representation and interaction; [Source](#).

Services and service providers

- **AnnotatorJS / Hypothes.is** - A framework and application for annotating resources online according to an [emerging W3C standard for web annotations](#). Focus is on scholarly applications; [Source Annotator](#) and [Source Hypothes.is](#).
- **Athens** - A sign-in system that provides access to library resources; [Source](#).
- **Bitbucket** - Free source code hosting site; [Source](#).
- **Center for Open Science** - A non-profit technology organization with a mission to increase the openness, integrity, and reproducibility of scientific research. Brian Nosek and Jeffrey Spies founded the organization in January 2013, funded mainly by the Laura and John Arnold Foundation and others, after implementation and use of the Open Science Framework (OSF); [Source](#).

- **CrossRef** - An association of scholarly publishers that develops shared infrastructure to support more effective scholarly communication; [Source](#).
- **Dissem.in** - An online tool that detects papers behind paywalls and invites their authors to upload them in one click to an open repository; [Source](#).
- **DSpace** - Software for digital open repositories launched by The Massachusetts Institute of Technology (MIT) in 2002; [Source](#).
- **Eprints** - Software for open digital repositories to self-archiving launched by Southampton University in 2000; [Source](#).
- **Etherpad** - An online, OSS collaborative writing/editing tool operating in real time; [Source](#).
- **Flexible Extensible Digital Object and Repository Architecture (FEDORA)** - Software for digital repositories launched by The Cornell and Virginia Universities in 2003; [Source](#).
- **Git** - An OSS distributed revision control system; [Source](#).
- **GitHub** - A web-based service that provides a source code repository that works exclusively with the Git command-line tool; [Source](#).
- **Google Scholar** - A freely accessible search engine for indexing the scholarly literature across an array of publishing formats and disciplines; [Source](#).
- **IPython notebook** - A web-based interactive computational environment where you can combine code execution, text, mathematics, plots and rich media into a single document; [Source](#).
- **OAI Media Importer Bot** - A computer program, run by [Daniel Mitchen](#), that takes figures and video clips from OA articles in PubMed, and copies them to [Wikimedia Commons](#) with full attribution of the original paper. This facilitates the reuse of those files in educational materials or Wikipedia articles.
- **Open Access Button** - Tracks global encounters with paywalls, and helps provide access to papers through a 'wishlist'; [Source](#).
- **Open Archives Initiative** - Supplies a common framework to web communities that allows them to gain access to content in a standard manner by means of metadata harvesting; [Source](#).
- **Open Conference Systems (OCS)** - A free Web publishing tool that will create a complete Web presence for scholarly conferences; [Source](#).
- **Open Harvester Systems** - A free metadata indexing system; [Source](#).
- **Open Journal Systems (OJS)** - A journal management and publishing system; [Source](#).

- **Open Monograph Press** - An OSS platform for managing the editorial workflow required to see monographs, edited volumes, and scholarly editions through internal and external review, editing, cataloguing, production and publication; [Source](#).
- **Our Research** - A company that specialises in building and maintaining open-source, open-data tools to help power the Open Science revolution; [Source](#).
- **Paperity** - A multidisciplinary aggregator of OA journals and papers, Gold and Hybrid. Aims to include ultimately 100% of OA literature; [Source](#).
- **Protocols.IO** - Up-to-date crowd-sourced protocol repository; [Source](#).
- **Publish or Perish** - Software for retrieving and analyzing academic citations; [Source](#).
- **ScienceOpen** - A discovery platform with interactive features for scholars to enhance their research in the open, make an impact, and receive credit for it; [Source](#).
- **Shibboleth** - A single sign-in system for computer networks and services on the open Internet; [Source](#).
- **Stack Overflow** - A Question and Answer site for programming issues; [Source](#).
- **Symplectic** - A world-leading products and services company specialized in research information management. Their flagship system, [Elements](#), is used by a number of the world's research institutions; [Source](#).
- **Zotero** - A free and open-source reference management software to manage bibliographic data and related research materials (such as PDF files). Notable features include web browser integration, online syncing, generation of in-text citations, footnotes, and bibliographies, as well as integration with the word processors Microsoft Word, LibreOffice Writer, and Google Docs. It is produced by the Center for History and New Media at George Mason University; [Source](#).

Data Repositories

- **Databib** - A searchable registry of research data repositories; [Source](#). Note that the Databib and re3data.org registries merged at the end of 2015.
- **DataONE** - A framework and infrastructure for Earth observational data; [Source](#).
- **Dryad** - A curated resource that makes the data underlying scientific publications discoverable, freely reusable, and citable; [Source](#).
- **Figshare** - A repository where users can make all of their research outputs available in a citable, shareable, and discoverable manner; [Source](#).
- **Genbank** - The NIH sequence database comprising an annotated collection of all publicly available DNA sequences. Part of the [International Nucleotide Sequence Database Collaboration](#); [Source](#).

- **Global Biodiversity Information Facility (GBIF)** - Contains data about all types of life on Earth, published according to common data standards; [Source](#).
- **Knowledge Network for Biocomplexity (KNB)** - A network for the discoverability, access, and interpretation of complex ecological data; [Source](#).
- **Morphbank** - An image database documenting a range of specimen-based research, including comparative anatomy and taxonomy. Funded by the National Science Foundation; [Source](#).
- **Morphobank** - A Web application for collaborative evolutionary research, specifically phylogenetic systematics or cladistics, involving morphology; [Source](#).
- **Open Access Directory Data Repositories** - A list of repositories or databases relating to Open Data. Organised by field subject; [Source](#).
- **Open Science Framework** - A tool created by the Center for Open Science for researchers. It is both a research and workflow management tool and open repository. Their goal is to link up the entire research ecosystem, from conception through publication. They give the user full control over the openness of their work and allow for the creation of registrations, which can be used when submitting registered reports; [Source](#).
- **re3data.org** - A global registry of research data repositories from different academic disciplines; [Source](#).
- **Registry of Research Data Repositories** - An Open Science tool that serves as a global registry of research data repositories; [Source](#).
- **UniProt** - Central repository of protein sequence and annotation data; [Source](#).
- **Worldwide Protein Data Bank (wwPDB)** - Publicly available repository of macromolecular structural data; [Source](#).
- **Zenodo** - An all-purpose free to use repository for all research outputs, with DOIs and flexible licensing; [Source](#).

Miscellaneous

- **Accessibility** - Refers to the degree of access. Defined by an end-user basis, depending on their ability to understand or reuse content.
- **Content mining** - Large-scale extraction of information from content (e.g., photographs, videos, audio, metadata), usually involving thousands of items; for example using [ContentMine](#).
- **Data archiving** - The process of moving data to a storage device for long-term preservation; [Source](#).

- **Data mining** - An analytic process designed to explore data in search of consistent patterns and/or systematic relationships between variables, and then to transform this information into content for future use.
- **Funder** - An institute, corporation or government body that provides financial assistance for research.
- **Impact** - The scale of use of research outputs both inside and outside of academia.
- **Mandate** - An authority to carry out a policy. In this context, largely to conform to OA policies.
- **Open Access Movement (OAM)** - A global movement started in the late 1990s and early 2000s fueled by the widespread public access to the World Wide Web. Its prime objective is the free and unrestricted access and reuse of the world's knowledge.
- **Open Access Directory Events** - Conferences and workshop archive related to OA; [Source](#).
- **Openwashing** - Having an appearance of OSS and open-licensing for marketing purposes, while continuing proprietary practices. Coined by [Audrey Watters](#).
- **Publicly funded research** - Refers to research which is, at least in part, funded by Governments, often through Research Councils.
- **Repeatability** - The similarity between results of a study or experiment and independent results obtained with the same methods and under identical conditions (i.e., pertains to methods and analysis).
- **Reproducibility** - The similarity between results of a study or experiment and independent results obtained with the same methods but under different conditions (i.e., pertains to results).
- **Reproducibility, computational** - When publishing computational findings, include details and access to the underlying code, data, and implementation.
- **Reproducibility, empirical** - Reproduction of results to obtain “verifiable facts”, through improving existing communication standards and reporting.
- **Reproducibility, statistical** - Validating the statistical results, errors, and confidence measures in research. Also the statistical assessment of repeated results for validation purposes; [Source](#).
- **Scraping** - A computing technique to extract information from websites; [Source](#).
- **Sharing** - The joint use of a resource or space. A fundamental aspect of collaborative research. As most research is digitally-authored and digitally-published, the resulting digital content is non-rivalrous and can be shared without any loss to the original creator.

Additional Resources

- The [Open Science MOOC](#) resource pages.
- The [IRIDIUM](#) international research data management glossary.
- List of [data repositories](#), from the Open Access Directory.
- The [Registry of Research data Repositories](#).
- List of [declarations](#) in support of OA, from the Open Access Directory.
- A [list of data repositories](#) approved by F1000Research.
- A [list of data repositories](#) approved by Scientific Data.
- A [map](#) of Open Educational Resource Repositories, containing locations for items in the directory of Open Educational Resource Repositories.