

Digital Science Publisher Day
Vocabulary List



For more information
publishing@digital-science.com

AI - Artificial intelligence (AI) is intelligence exhibited by machines. In computer science, the field of AI research defines itself as the study of "intelligent agents": any device that perceives its environment and takes actions that maximize its chance of success at some goal.

Altmetrics - In scholarly and scientific publishing, altmetrics are non-traditional metrics proposed as an alternative to more traditional citation impact metrics, such as impact factor and h-index. The term altmetrics was proposed in 2010, as a generalization of article level metrics, and has its roots in the #altmetrics hashtag

Altmetric Score - a weighted count of the online attention a single item has received. Calculated by altmetrics provider Altmetric, the score takes into account the volume, type and author of each mention, and acts as an indicator of engagement (positive or potentially negative).

API - Application program interface (API) is a set of routines, protocols, and tools for building software applications. An API specifies how software components should interact.

CC BY/Creative Commons license - A Creative Commons (CC) license is one of several public

copyright licenses that enable the free distribution of an otherwise copyrighted work. A CC license is used when an author wants to give people the right to share, use, and build upon a work that they have created.

CHORUS - a suite of services and best practices that provides a sustainable solution for agencies and publishers to deliver public access to published articles reporting on funded research in the United States.

CiteScore - Elsevier's new metrics that are part of the Scopus basket of journal metrics that includes SNIP (Source Normalized Impact per Paper), SJR (SCImago Journal Rank), citation- and document-counts and percentage cited. The integration of these metrics into Scopus provides insights into the citation impact of more than 22,220 titles.

CLOCKSS -Controlled LOCKSS is a not-for-profit joint venture between the world's leading academic publishers and research libraries whose mission is to build a sustainable, geographically distributed dark archive with which to ensure the long-term survival of Web-based scholarly publications for the benefit of the greater global research community.

COPE - The Committee on Publication Ethics (COPE) provides advice to editors and publishers on all aspects of publication ethics and, in particular, how to handle cases of research and publication misconduct. It also provides a forum for its members to discuss individual cases. COPE does not investigate individual cases but encourages editors to ensure cases are investigated by the appropriate authorities (usually a research institution or employer).

COUNTER - provides the Code of Practice that enables publishers and vendors to report usage of their electronic resources in a consistent way. This enables libraries to compare data received from different publishers and vendors. The COUNTER Code of Practice helps librarians demonstrate the value of electronic resources by facilitating the recording and reporting of online resource usage stats in a consistent and credible way

CrossRef - a not-for-profit membership organization for scholarly publishing working to make content easy to find, cite, link, and assess.

DOI - A digital object identifier (DOI) is a unique alphanumeric string assigned by a registration

agency (the International DOI Foundation) to identify content and provide a persistent link to its location on the Internet. The publisher assigns a DOI when your article is published and made available electronically.

Drupal - content management software used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles.

DTD - A document type definition (DTD) is a set of markup declarations that define a document type for an SGML-family markup language (SGML, XML, HTML). A Document Type Definition (DTD) defines the legal building blocks of an XML document. It defines the document structure with a list of legal elements and attributes.

EPUB - an e-book file format with the extension .epub that can be downloaded and read on devices like smartphones, tablets, computers, or e-readers. It is a technical standard published by the International Digital Publishing Forum (IDPF). The term is short for electronic publication and is sometimes styled ePub.

FTP - an acronym for File Transfer Protocol. As the name suggests, FTP is used to transfer files between computers on a network. You can use FTP to exchange files between computer accounts, transfer files between an account and a desktop computer, or access online software archives.

FundRef - Facilitated by CrossRef, FundRef is a collaborative project of scholarly publishers and funding agencies that supports a standard way of reporting funding sources for published scholarly research.

FundRef - Facilitated by CrossRef, FundRef is a collaborative project of scholarly publishers and funding agencies that supports a standard way of reporting funding sources for published scholarly research.

Google Analytics - a freemium web analytics service offered by Google that tracks and reports website traffic. Google launched the service in November 2005 after acquiring Urchin. Google Analytics is now the most widely used web analytics service on the Internet.

GRID - Global Research Identifier Database (GRID) is an openly accessible dataset of information on world-wide research organizations.

H-index - an author-level metric that attempts to measure both the productivity and citation impact of the publications of a scientist or scholar. The index is based on the set of the scientist's most cited papers and the number of citations that they have received in other publications.

HSS - Humanities and Social Sciences

HTML - Hypertext Markup Language, a standardized system for tagging text files to achieve font, color, graphic, and hyperlink effects on World Wide Web pages.

Impact factor - The impact factor (IF) is a measure of the frequency with which the average article in a journal has been cited in a particular year. It is used to measure the importance or rank of a journal by calculating the times its articles are cited

ISBN - the acronym for International Standard Book Number. This 10 or 13-digit number identifies a specific book, an edition of a book, or a book-like product (such as an audiobook). Since 1970 each published book has a unique ISBN. In 2007, assigned ISBNs changed from 10 digits to 13.

ISSN - An International Standard Serial Number (ISSN) is an eight-digit serial number used to uniquely identify a serial publication. The ISSN is especially helpful in distinguishing between serials with the same title.

JATS - The Journal Publishing Tag Set is a moderately prescriptive set, optimized for the archives who wish to regularize and control their content, not to accept the sequence and arrangement presented to them by any particular publisher. Publishing is also intended for use by publishers for the initial XML tagging of journal material, usually as converted from an authoring form like Microsoft Word.

LaTeX - a document preparation system for high-quality typesetting. It is most often used for medium-to-large technical or scientific documents but it can be used for almost any form of publishing.

MARC records - MACHine-Readable Cataloging record. And what is a machine-readable cataloging record? Machine-readable: "Machine-readable" means that one particular type of machine, a computer, can read and interpret the data in the cataloging record.

MeSH - Medical Subject Headings (MeSH) is a comprehensive controlled vocabulary for the purpose of indexing journal articles and books in the life sciences; it serves as a thesaurus that facilitates searching. Created and updated by the United States National Library of Medicine (NLM), it is used by the MEDLINE/PubMed article database and by NLM's catalog of book holdings. MeSH is also used by ClinicalTrials.gov registry to classify which diseases are studied by trials registered in ClinicalTrials.gov.

Metadata - Metadata is data (information) about data. The <meta> tag provides metadata about the HTML document. ... The metadata can be used by browsers (how to display content or reload page), search engines (keywords), or other web services.

NISO - The National Information Standards Organization is a United States non-profit standards organization that develops, maintains and publishes technical standards related to publishing, bibliographic and library applications.

NLP - Natural language processing (NLP) is a field of computer science, artificial intelligence, and computational linguistics concerned with the interactions between

computers and human (natural) languages and, in particular, concerned with programming computers to fruitfully process large natural language corpora.

ONIX - ONIX for Books is an XML format for sharing bibliographic data pertaining to both traditional books and eBooks. It is the oldest of the three ONIX standards, and is widely implemented in the book trade in North America, Europe and increasingly in the Asia-Pacific region. It allows book and ebook publishers to create and manage a corpus of rich metadata about their products, and to exchange it with their customers (distributors and retailers) in a coherent, unambiguous, and largely automated manner.

Open Access (OA) - Open Access is the free, immediate, online availability of research articles coupled with the rights to use these articles fully in the digital environment. Open Access ensures that anyone can access and use these results—to turn ideas into industries and breakthroughs into better lives.

Green OA is self-archiving of the pre or post-print in repositories
Gold OA refers to articles in fully accessible open access journals - authors pay a fee to be published

Hybrid OA – subscription journals with open access to some individual articles - authors pay a fee to be published

Platinum OA - content that is fully open access and do not charge authors an APC (article-processing charge)

Open Access Mandate - a policy adopted by a research institution, research funder, or government which requires researchers—usually university faculty or research staff and/or research grant recipients—to make their published, peer-reviewed journal articles and conference papers open access (1) by self-archiving their final, peer-reviewed drafts in a freely accessible institutional repository or disciplinary repository ("Green OA") or (2) by publishing them in an open-access journal ("Gold OA") or both.

Open Peer Review - also called "public peer review", "transparent peer review") denotes several, closely related forms of scholarly peer review. ... Open-invitation or post-publication peer review, where anyone interested can contribute to the peer review process.

ORCID - Open Researcher and Contributor ID (ORCID) is a nonproprietary alphanumeric code to uniquely identify scientific and

other academic authors and contributors.

PDF - Portable Document Format (PDF) is a file format used to present and exchange documents reliably, independent of software, hardware, or operating system. Invented by Adobe, PDF is now an open standard maintained by the International Organization for Standardization (ISO). PDFs can contain links and buttons, form fields, audio, video, and business logic.

Peer review - Scholarly peer review (also known as refereeing) is the process of subjecting an author's scholarly work, research, or ideas to the scrutiny of others who are experts in the same field, before a paper describing this work is published in a journal or as a book.

Portoco - The Portico archive is a centralized repository of tens of thousands of e-journals, e-books, and other electronic content, replicated to ensure security. Content comes into the archive under formal preservation agreements with publishers.

PubMed - PubMed is a service of the US National Library of Medicine® that: Provides free access to MEDLINE®, the NLM® database of indexed citations and

abstracts to medical, nursing, dental, veterinary, health care, and preclinical sciences journal articles. Includes additional selected life sciences journals not in MEDLINE.

RCR - Relative Citation Ratio (RCR) is a field-normalized metric that shows the citation impact of one or more articles relative to the average NIH-funded paper

Reproducibility - the ability of an entire analysis of an experiment or study to be duplicated, either by the same researcher or by someone else working independently, whereas reproducing an experiment is called replicating it.

Science Citation Index - The Science Citation Index (SCI) is a citation index originally produced by the Institute for Scientific Information (ISI) and created by Eugene Garfield. The larger version (Science Citation Index Expanded) covers more than 8,500 notable and significant journals, across 150 disciplines, from 1900 to the present. These are alternatively described as the world's leading journals of science and technology, because of a rigorous selection process.

Sci-Hub - a controversial pirate online search engine with over 58,000,000 academic papers and

articles available for direct download, bypassing publisher paywalls. New papers are uploaded daily when accessed through educational institution proxies, and papers that have been accessed through Sci-Hub are stored in the LibGen repository.

SCImago - SCImago Journal Rank (SJR indicator) is a measure of scientific influence of scholarly journals that accounts for both the number of citations received by a journal and the importance or prestige of the journals where such citations come from.

Semantic HTML - The use of HTML markup to reinforce the semantics, or meaning, of the information in webpages and web applications rather than merely to define its presentation or look. Semantic HTML is processed by traditional web browsers as well as by many other user agents.

SEO - Stands for “search engine optimization.” It is the process of getting traffic from the “free,” “organic,” “editorial” or “natural” search results on search engines.

SGML - Standard Generalized Markup Language (SGML) is a standard for how to specify a document markup language or tag set. Such a specification is itself a document type definition (DTD). SGML is not in itself a document

language, but a description of how to specify one.

STM - Scientific, Technical and Medical Publishing
Submission systems - A submission management system is a software system, also known as submission processing, that streamlines and eases out the collection, tracking and management of electronic submissions. ... A submission management system may be regarded as an application-specific content management system.

Supplemental data - data supplemental to a journal article, including additional figures, materials & methods, tables, data sets, etc.

Tagging - To identify content for delivery to multiple publishing channels, it is necessary to label each piece of content with a "tag." Standard tagging schemes have been developed to make the content "machine-readable," and these universal tagging sets are referred to as markup languages. Examples of markup languages include XML (eXtensible Markup Language), SGML (Standard Generalized Markup Language), and HTML (HyperText Markup Language). Tagging content with different markup languages allows it to be delivered and properly translated to print products (XML),

Web sites (HTML), wireless products (WML), and other media channels.

XML - a metalanguage which allows users to define their own customized markup languages, especially in order to display documents on the Internet.