

TEXT AND DATA MINING FOR NON-COMMERCIAL SCIENTIFIC RESEARCH

A STATEMENT OF COMMITMENT BY STM PUBLISHERS TO A ROADMAP TO ENABLE TEXT AND DATA MINING (TDM) FOR NON COMMERCIAL SCIENTIFIC RESEARCH IN THE EUROPEAN UNION

Recalling that the International Association of STM Publishers (STM) and member publishers have actively and constructively supported the discussions in the Licences for Europe Working Group Four,

Recognising that a high level of copyright and database protection, together with interoperability standards, technology innovations and sustainable business models, are vital for the viability of the creative industries.

Reaffirming that licensing is the smart and speedy route to providing access and the rights needed for text and data mining, and for related technologies such as text-to-speech and automated translation services,

Emphasising that the International Association of STM Publishers (STM) and its member publishers are committed to on-going discussions with research communities to help them achieve the most when using the research published in science journals, whether subscription based or open access.

Recalling that offering publications and new information services, such as mine-able high- quality content in useable formats and high-quality content delivery systems, to enable successful text and data mining to the commercial and non-commercial communities is central to what publishers do,

STM PUBLISHERS DECLARE THEIR COMMITMENT TO THE FACILITATION OF TEXT AND DATA MINING FOR NON-COMMERCIAL SCIENTIFIC RESEARCH IN THE EUROPEAN UNION:

Publishers' Commitments and Toolbox of Solutions

Signatories commit to granting the necessary copyright licenses to permit the text and data mining of copyright-protected content and other subject-matter on reasonable terms for non-commercial scientific research purposes in the European Union. Whilst enabling managed use of their subscription platforms and electronic resources, either online or offline, signatories will retain the ability to apply and develop safeguards, such as technological solutions to minimise online system performance degradation.

The signatories recognise that there should not be an undue burden on bona fide researchers at non-commercial institutions intending to engage in text and data mining of their subscribed or open access journal content for non-commercial scientific research that does not compete with and does not substitute for the content and services of STM publishers.

Therefore, signatory publishers will work with rights clearance agents as necessary and implement licensing systems to facilitate easy, “one-to-many” rights clearance, such as PLS Clear, and will continue to work on relevant technological solutions, such as CrossRef’s Prospect service to facilitate the access to content in an efficient way. Any licensing solutions offered by publishers are without prejudice to the author’s moral rights, including the integrity and ethical aspects of academic research.

Commitment 1: The Sample Clause for non-commercial Text and Data Mining in scientific research

As part of authorising text and data mining, signatories commit either to implementing the STM sample clause for non-commercial text and data mining into existing agreements (on request or as part of subscription renewal), or to offering researchers and institutions to which researchers are affiliated comparable and equivalent access rights for the purpose of non-commercial text and data mining of subscribed journal content for non-commercial scientific research, at no additional cost to researchers/subscribing institutions.

Commitment 2: Developing the Mine-ability of Content

Signatories commit to finding the best methods to use existing proprietary electronic infrastructure and to enable the content located on or via that infrastructure to be used to facilitate text and data mining. Signatories also undertake to examine and improve the mine-ability of content, taking into account available resources, technical advances and current technology, including for content migration and formatting. Signatory publishers will also be sensitive to market demand in the different academic disciplines.

Commitment 3: Developing the Platforms

Signatories undertake to find solutions whereby researchers resident at European non-commercial institutions holding relevant subscriptions are able to integrate subscribed journal content and open access licensed content for text and data mining through suitable technical applications and infrastructure, such as CrossRef Prospect.

A Roadmap for Successful Text and Data Mining for Non-commercial Scientific Research Purposes

The signatories will monitor and evaluate the evolution of the dynamic and fast-moving field of text and data mining, and adapt practices and services as required.

Continuing to work collaboratively with all stakeholders to offer new publications and information services, such as mine-able normalized content and powerful content delivery systems to enable successful text and data mining by non-commercial communities and commercial entities, is part and parcel of what digital publishers do. The interests of research and knowledge creation are best served by encouraging and incentivizing the creation of new copyright works, new publications and new information services that contribute to the development of these communities and to society at large.

Public-private partnerships between publishers, libraries, governments, research funders and data curation centres and repositories are eminently possible and hold the promise of ever greater integration of Big Data with Big Science.

SIGNATORY PUBLISHERS:

American Chemical Society
American Institute of Physics
American Physical Society
BMJ
Brill
Cambridge University Press
EDP Sciences
Elsevier BV
George Thieme Verlag KG
Hogrefe Verlag GmbH & Co. KG
IEEE
Institute of Physics Publishing
John Wiley & Sons
Karger Publishers
Springer Nature
New England Journal of Medicine (Massachusetts Medical Society)
Oxford University Press
Royal Society of Chemistry
Sage Publishing
Taylor & Francis
Wolters Kluwer Health (Medical Research) Ltd

Version 1.1 November 2013

Version 1.2 September 2015 (additional signatory)

Version 1.3 October 2015 (additional signatory)

Version 1.4 November 2015 (additional signatories)

Version 1.5 May 2017 (additional signatories)