



open access in the european research
area through scholarly communication

OPERAS White Paper



Common Standards

July 2018

CONTENTS

EXECUTIVE SUMMARY	4
1. Introduction.....	5
2. Framework and Scope	6
2.1 Open Science	6
2.2 Mandates and Principles to Support Open Science at EU Level	7
2.3 European Open Science Cloud (EOSC).....	8
2.4 Communities Engaged in Scholarly Communication	8
3. State of the Art	11
3.1 Emerging Trends	11
3.2 Initiatives for the Implementation of Standards.....	12
3.3 Operational and Technical Standards for e-Infrastructures	14
3.3.1 Suggested Framework for the Implementation of Standards.....	14
3.3.2 Content Quality and Impact Assessment.....	15
3.3.3 Interoperability	16
3.3.4 Availability	17
3.3.5 Processability	18
4. Where OPERAS Stands	19
4.1 Mapping of Partners' Publishing Workflows, Technical and Content Quality Standards	19
4.1.1 OpenEdition.....	19
4.1.2 EKT ePublishing.....	20
4.1.3 OAPEN	21
4.2 HIRMEOS	22
4.3 Future Orientations for a Common Set of Standards to be Provided by OPERAS	23
LIST OF ACRONYMS AND ABBREVIATIONS	25
REFERENCES	26
ANNEXES	28

This White Paper has been prepared by the **OPERAS Common Standards Working Group**

Coordinator: ***National Documentation Centre (EKT/NHRF)***
Marina Angelaki
Irakleitos Souyioultzoglou

Participants: ***OAPEN***
Eelco Ferwerda
Ronald Snijder

OpenEdition
Arnaud Gingold
Pierre Mounier

University of Milan
Paola Galimberti

All OPERAS White Papers: operas-eu.org/aboutoperas/working-groups

OPERAS Common Standards Working Group contact: irakleitos@ekt.gr

OPERAS website: operas-eu.org

OPERAS contact: pierre.mounier@openedition.org and suzanne.dumouchel@huma-num.fr

CC BY 4.0

EXECUTIVE SUMMARY

The OPERAS Working Group for Common Standards aims at exploring the workflows, mediums and technical standards that have recently emerged as a result of the changes brought about by the transition to Open Science. It places focus on the importance of common standards, and traces the improvements required to ensure content quality and interconnectivity for scholarly output in the SSH and beyond.

The White Paper on Common Standards comprises desk research and identifies key operational and technical aspects to be addressed by digital research infrastructures and service providers. It particularly sketches the landscape of Open Science in Europe, focusing on the policy framework and the institutional initiatives at EU level; it also describes current and emerging research practices and highlights the needs of the stakeholders and communities engaged in scholarly communication.

Reference is specifically made to technical and operational standards for publishing infrastructures, and their importance in providing a digital scholarly communication framework that fosters content reuse and collaboration among researchers, while enabling the implementation of innovative research methods. To this end, the white paper identifies needs yet to be met, introduces 4 complementary areas (content quality and impact assessment, interoperability, availability and processability) for the introduction of common standards, and provides basic recommendations for their future implementation.

The white paper also examines where OPERAS members stand and suggests a roadmap for the community-wide adoption of standards.. As effective implementation of common standards is highly depended upon stakeholders' increased awareness and commitment towards more effective ways of conducting, presenting and communicating research, the white paper underlines the instrumental role of the OPERAS network in specifying new standards and updating existing ones.

1. INTRODUCTION

This White Paper is a deliverable of the OPERAS¹ Working Group on common standards. It explores the new publishing workflows, mediums and technical standards that have recently emerged as a result of the changes brought about by the transition to Open Science so as to provide general guidelines and recommendations towards the development of a unified scholarly communication framework in the Social Sciences and Humanities (SSH), and beyond.

In examining the current trends and challenges deriving from the extensive adoption of digital research and communication flows, the white paper identifies key operational and technical aspects to be addressed by publishing e-infrastructures and service providers. It also highlights the importance of common standards, and the need for them to be collectively implemented by all agents involved in the digital scholarly communication processes.

The term **common standards** refers to features, workflows, tools and practices applied combinedly to upgrade e-infrastructures into a state-of-the-art condition. It indicates the need of these standards to be globally introduced, as an essential step in shaping an integrated digital scholarly communication framework.

Specifically, the white paper will look into the existing landscape and trace the standards required to ensure content quality, availability and discoverability; moreover, it will proceed into examining where OPERAS members stand and the work needed to reach these standards. OPERAS aims at successfully deploying a suite of services and complementary infrastructures, by investing on engaged partners' capacity and expertise to enable the integration of SSH outputs into the European Open Science Cloud (EOSC). To this end, this paper identifies important operational and technical aspects to be addressed towards the realisation of a comprehensive publication management approach; it also provides basic guidelines for the standardization of (a) content management processes and (b) publishing models, along the following lines:

- Editorial and publishing workflows
- Dissemination and preservation of scholarly outputs
- Metadata documentation and organization
- Content discoverability and processability

Therefore, the main scope of the paper is to highlight the necessity for long-term commitments towards the improvement of infrastructures and the enhancement of scholarly publishing processes, in light of the recent developments in the digital scholarly communication landscape.

¹ OPERAS (<https://operas-eu.org/>) is a European research infrastructure for the development of open scholarly communication, particularly in the Social Sciences and Humanities

2. FRAMEWORK AND SCOPE

This section describes the recent developments in research and the complementary roles of researchers, funders, research institutions, infrastructure providers and the EU in realising the Open Science paradigm. In addition, it identifies existing and emerging challenges that stipulate the central role of e-infrastructures and the importance of standards in shaping a global communication framework for all communities engaged in research.

2.1 OPEN SCIENCE

Open Science represents a new approach to the scientific process that seeks to ensure that access to the entire life-cycle of research remains fundamentally open and replicable. This approach shifts the emphasis from the standard practices of publishing research results in scientific publications towards sharing and using knowledge. At a more practical level, this new paradigm entails important and on-going transitions in the way research is performed, researchers collaborate, knowledge is shared and science is organized. A key component of Open Science is open access to publications and research data, yet it is not limited to these two aspects as it also includes aspects like open peer review, open methodologies, open educational resources, and other participatory processes like citizen science.

Within the EU, Open Science forms part of a broader EU strategy and in particular of the three goals for EU research and innovation policy summarized as “Open Innovation, Open Science and Open to the World”². The EU’s interest in supporting Open Science has been confirmed in Council Conclusions on the transition towards an Open Science system adopted on 27 May 2016. The Council acknowledged “that open science has the potential to increase the quality, impact and benefits of science and to accelerate advancement of knowledge” and called on the Commission, the Member States and the stakeholders to “take the necessary actions needed to making open science a reality and to advocate the need for concerted actions.”³

To support further the development of Open Science policy the Directorate General for Research and Innovation (DG RTD) set up an Open Science Policy Platform (OSPP). The platform is intended to provide a forum for a structured discussion with key stakeholders including inter alia research funding and research performing organisations, libraries, and scientific publication associations, and give advice to the Commission on the basis of the European Open Science agenda. The latter is structured around the following themes: 1) fostering and creating incentives for Open Science, 2) removing barriers for Open Science, 3) mainstreaming and further promoting open access policies, 4) developing research infrastructures for Open Science and 5) embedding Open Science in society as a socio-economic driver. These five action lines are in turn translated into eight topics of policy concern, namely: rewards, altmetrics, Open Science Cloud, changing business models for publishing, research integrity, citizen science, open education and skills and FAIR open data. The work of the OSPP is further supported through the Open Science Monitor commissioned also by DG RTD, developed by several partners and led by RAND Europe, an independent non-profit research institute. The monitor is “a pilot project to test the viability and value of assessing Open Science activity in Europe and beyond.”

² <https://publications.europa.eu/en/publication-detail/-/publication/3213b335-1cbc-11e6-ba9a-01aa75ed71a1>

³ The transition towards an Open Science system - Council conclusions (adopted on 27/05/2016). Available at <http://data.consilium.europa.eu/doc/document/ST-9526-2016-INIT/en/pdf>

2.2 MANDATES AND PRINCIPLES TO SUPPORT OPEN SCIENCE AT EU LEVEL

The European Commission has been an active supporter of open access (to both publications and research data) based on the notion that “there should be no need to pay for information funded from the public purse each time it is accessed or used”. Open access is expected to contribute in generating growth through greater efficiency, faster progress and improved transparency of the scientific process through the involvement of citizens and society. The benefits for researchers are associated with the positive impact on the visibility of research outputs and on the increase in usage and impact.

The support provided by the EU to open access has been further strengthened through the Council of the European Union conclusions of 27 May 2016. The Council recognized that “the exponential growth of data, the increasingly powerful digital technologies, together with the globalization of the scientific community and the increasing demand for addressing the societal challenges contribute to the ongoing transformation and the opening up of science and research which is referred to as “open science”. It called on Member States, the Commission and stakeholders to remove financial and legal barriers and agreed to promote the mainstreaming of open access to publications by continuing to support a transition to immediate open access as the default by 2020.

Open Access is required (mandatory) for all peer-reviewed publications resulting from projects funded under Horizon 2020. This decision follows the pilot action on Open Access, which was implemented in FP7 for part of the funding period. Following also on the **pilot action on Open Access to research data generated in Horizon 2020**, the Commission decided to extend the pilot to all thematic areas as stated in the 2017 Work Programme. Acknowledging that not all data can be open, the possibility of opting out (at any stage before or after signing the Grant) is provided. The Commission’s approach is therefore best described as “as open as possible as closed as necessary”.

The open access mandate is translated into specific requirements in the Model Grant Agreement (articles 29.2 and 29.3) and in the H2020 work programme.

In the context of the European Research Area (ERA) open access is discussed under Priority 5b: “Open access to publications and data in an open science context” (Priority 5 “Optimal circulation, access to and transfer of scientific knowledge”) and headline indicator 5b- “Open Access”. On the basis of the indicator used to track performance and progress for sub-priority 5b- “Open access” (share of papers in open access) approximately 52% of publications in the EU-28 are available in open access. As also highlighted in the same report, the green route makes a more significant contribution to the overall levels of open access compared to the gold route as almost 2/3 of papers are made available through the green route. Depositing in repositories is important as articles are easily discoverable through search engines and retrievable.

In relation to research data, the European Commission has also produced a set of guidelines on FAIR data management in Horizon 2020 to help beneficiaries make their research data findable, accessible, interoperable and reusable (FAIR). The Commission stresses the importance of Data Management Plans (DMPs) as key components of good data management and as such provides guidance to support researchers in developing their DMPs.

A further important initiative relates to the launch of the ‘European Open Science Cloud’ that “aims to create a trusted environment for hosting and processing research data to support EU science in its global leading role.”

2.3 EUROPEAN OPEN SCIENCE CLOUD (EOSC)

The European Open Science Cloud (EOSC) is a vision of the European Commission to provide an infrastructure to support open science and open innovation through the creation of a virtual environment with open and seamless service that will allow researchers to store, manage, analyze and reuse data and results. Through the EOSC Europe wants to ensure that its researchers reap the benefits of data-driven science. Its use will not be limited to researchers though, as it is also expected to serve education and training purposes and to be used by governments and the business sector. Overall, EOSC is expected to leverage other related EU initiatives and actions under the Open Science agenda.

Within this context, the EOSC pilot project (<https://eoscpilot.eu>) looks into the technical, scientific and cultural challenges that need to be addressed in the deployment of EOSC. To achieve this, EOSC pilot will propose and trial a governance framework, develop a number of demonstrators, engage with a broad number of stakeholders to build the trust and skills required.

2.4 COMMUNITIES ENGAGED IN SCHOLARLY COMMUNICATION

The changes brought about by this new approach to conducting and communicating research result in an increased diversity of practices as well as stakeholders' roles and needs, which are briefly presented below:

Researchers: The fundamental research practices of collecting, organizing, processing and disseminating scientific information are highly related to the availability and discoverability of primary resources. Thus, for research to be effective and fruitful, scientific content has to be widely disseminated and effortlessly accessed - a condition that could potentially be met within the digital academic ecosystem, where scholarly communication is performed across a variety of channels and venues: developments such as the advent of Web 2.0 functionalities open new pathways in scholarly communication and significantly increase researchers' capacity to discover and exchange resources and information; moreover, an increasing number of dedicated tools and mediums underpin researchers' capacity to process and enrich a variety of sources available in different formats (e.g. texts, images, datasets).

In this constantly evolving context, emphasis has to be placed on the implications of researchers' enhanced digital skills and the importance of the recently adopted processes and research methods. The advent of Digital Humanities⁴ raises issues related to the sufficient support of prevailing scholarly activities, which now involve a wide spectrum of user-driven innovative practices that entail providers' commitment in designing long-term strategies and tools for managing and preserving resources, enabling collaborative work, and disseminating research outputs. As researchers ask for inclusive publishing venues that can accommodate new types of research outputs (such as media), link research data to publications, and allow users' intervention (commenting, annotating), the quantity and quality of user-generated content becomes a question of crucial importance.

⁴ Digital humanities (DH) is an area of scholarly activity at the intersection of computing or digital technologies and the disciplines of the humanities (source: Wikipedia) https://en.wikipedia.org/wiki/Digital_humanities

Publishers: Academic publishing has evolved into a diverse enterprise, involving small-, medium-, and large-scale independent or commercial publishers, different business models⁵, practices and dissemination venues. As digital publishing becomes a norm, the existing variety of actors and models often results into wide discrepancies in terms of operational and technical standards.

Moreover, recent developments challenge the perceived role of academic publishers, who need to maintain their central place within the scholarly communication landscape, while asked to correspond to an increasing diversification of publishing practices and mediums. Nowadays, delivered value is equally generated by scientific quality and a variety of digital content-related attributes and features, such as availability and process ability. Publishers are asked to develop new tools and services for researchers, and engage in incentives towards the optimization of digital workflows and content.

As the OPERAS survey indicates⁶, publishers have largely conceived this necessity, and share common interest in developing integrated services as well as standardized publishing and dissemination practices. To this end, the establishment of common standards enables more systematic collaboration among existing publishing initiatives and facilitates the deployment of innovative publication models and tools that help researchers to discover resources, communicate effectively, and assess the impact of their work.

Research funders: Public and private research funding bodies have been widely acknowledged as drivers of Open Science. Research funders provide a range of grant schemes to facilitate innovative research, and have adopted policies to make research outputs available in open access.

As funders need to assess grantees' compliance with their open access requirements, it is essential to introduce standards and services that link funded research and researchers with all relevant published content. On the other hand, provisions need to be made for proper assessment of the impact of funding mechanisms and open access policies: data for funders and authors has to be registered with published content, allowing proper identification and interlinking of all agents involved in a funded research project.

Service and infrastructure providers: As the OPERAS landscape study⁷ indicates, fragmentation (both in terms of the size and nature of publishers and of their business models) is a key characteristic of the academic publishing landscape. In this context, the main challenge in designing sustainable open access publishing models is to identify current needs and limitations that permeate the scholarly communication framework.

A successful publishing service should deploy infrastructures designed to interoperate with a multitude of systems designed for the management and provisioning of digital content. Thus, platform providers cope with a series of administrative and technical issues related to the potential for content reuse, such as the need for effective integration with repositories and/ or search engines; the incorporation of procedures that would ensure the long-term preservation and utilization of the content; and the development of tools to enable identification, authentication, metadata enrichment and discovery.

⁵ The OPERAS white paper in Business Models provides a detailed overview of the different approaches to Open Access publishing. Available at: <https://zenodo.org/record/1323707>

⁶ https://f-origin.hypotheses.org/wp-content/blogs.dir/2465/files/2018/05/operas_online_survey_optimizing_e-infrastructure.pdf

⁷ Landscape Study on Open Access Publishing. DOI 10.5281/zenodo.1009554. See also Tsoukala, V. (2015) University based Open Access Publishing. State of Play, SPARC Europe. http://sparceurope.org/wp-content/uploads/2015/12/SE_UPublishing_Report_0315.pdf

The introduction of common standards tackles the main obstacles towards the full interoperability of publishing infrastructures and paves the way for innovative services at inter-platform level by providing additional data, links and interactions to published material. It also allows a wide adoption of the fast technological developments that occur in the fields of open public data and of open digital content and enables broad reuse and organization of published content.

Libraries: The main role of libraries is to collect, preserve and provide access to scholarly resources. Due to their active participation in the research cycle, libraries face a number of important challenges stemming from the increasing volume of digital content, the predominance of digital dissemination mediums that scholars choose to make their work publicly available, as well as the varying types of material to be curated. Libraries are required to handle digital versions of printed content and, at the same time, make provisions for the preservation of digital resources.

Moreover, academic institutions develop publishing models in the context of which libraries assume various and combined roles in regard to content management. The realization of libraries as publishers and curators implies that their technical infrastructure and operational principles are compatible to the wider context of the digital research ecosystem, and entails challenges related to the introduction of additional workflows and outputs (publications, datasets, multimedia etc.).

The introduction of shared and collectively applicable standards will enhance libraries' capacity to serve researchers in their binary status as producers and consumers of scientific content.

The table below summarizes the importance of common standards for each stakeholder category:

Researchers: inclusive publishing venues to accommodate new types of research outputs; link research data to publications; support of collaborative work

Publishers: new tools and services for researchers; optimization of digital workflows; innovative publication models; content delivered in multiple formats

Funders: identification and interlinking of all agents and outputs of a funded research project

Infrastructure providers: long-term preservation and utilization of the content; tools to enable identification, authentication, metadata enrichment and discovery

Libraries: assume new roles as publishers and curators; handle digital versions of printed content; long-term preservation of digital resources

3. STATE OF THE ART

This section will focus on technical and operational standards for publishing infrastructures, and highlight their importance in providing a sustainable framework for open scholarly communication. Four complementary areas of assessment have been identified:

1. Content quality and impact
2. Interoperability
3. Availability
4. Processability

3.1 EMERGING TRENDS

Research has evolved into a multifaceted activity that encompasses complex methodologies and workflows:⁸ text has ceased to be the exclusive resource for researchers, as the use of new applications enables scholars to discover, process and reproduce a wide range of digital-born or digitized sources (such as image sets, corpora, data sets, visualizations), and introduces techniques that allow collective contributions and metadata production. Unhindered flow of information gradually becomes a precondition for the incorporation of Humanities research into the digital ecosystem, as scholarly communication encompasses innovative practices such as information commenting, data extraction and metadata harvesting.

In this context, research practices in the Humanities increasingly relate to the systematic use of digital resources and tools. Digital Humanities (DH) has recently emerged as an innovative scholarly activity that successfully deploys digital workflows and introduces new methodologies based on collaborative and interdisciplinary work;⁹ thus, it reflects the ways in which research practices progress and science is taught, performed and communicated within the digital ecosystem.

This, in turn, suggests the implementation of new principles and standards that ensure openness, interoperability and processability for all scientific information (cf. Warren 2015). A significantly increasing proportion of published material in the Humanities is available in open access and new venues of communicating research are emerging (cf. Eyman 2015): in addition to publishing in conventional form (e.g. journal articles and monographs in print or digital form), researchers publish pre-prints, compile and deposit datasets and post their work in scientific blogs and other alternative dissemination venues.

Thus, openness emerges as a cultural value that permeates research; in addition to removing access barriers, openness also becomes a norm in the processes of reviewing and assessing research outputs. The recently introduced concept of open peer review (OPR) encompasses a wide spectrum of practices, ranging from revealing authors' and reviewers' names, to collective commenting by "non-experts". OPR is an essential component of Open Science and closely intertwined with digital publishing, as it is performed with the use of specific features (e.g. annotation technologies) and

⁸ Harley, Acord and Earl-Novell's "Assessing the Future Landscape of Scholarly Communication: An Exploration of Faculty Values and Needs in Seven Disciplines" provides a thorough review of the current and emerging scholarly communication practices. Available at: <https://escholarship.org/uc/item/15x7385g>

⁹ Programming Historian (<https://programminghistorian.org/>) is an indicative example of collaborative initiatives by the DH research community.

generates new outputs (e.g conversation threads around published content) that diverge from conventional publication forms.

Within this composite context, copyright issues and proper licensing of publicly available material become a question of critical importance. As the pool of resources that can be accessed, distributed and reused is growing, it is essential for researchers to encourage access in a standardized way that allows others to share and build upon existing content. The need of all subsequent versions of the originally published work to be granted appropriate permissions lead to the emergence of flexible licensing processes, whereupon publishers and funders allow for non-exclusive distribution of the originally published version of the work, or even prior to and during the review process, as this can lead to productive exchanges, as well as earlier and greater citation.

Public access to scientific content (prior to or after publication) results in innovative collaborative workflows, successful scientific ventures, increased impact and widespread dissemination of researchers' work. On the other hand, it requires the global adoption of common operational and technical standards, to facilitate the dissemination of content in an organized manner that regulates copyright and access issues (Hutchens 2013, Pentz & Tananbaum 2014, Browning, Guedon and Kaplan 2013), while stimulating the activities of institutional stakeholders engaged in the scholarly communication cycle.

3.2 INITIATIVES FOR THE IMPLEMENTATION OF STANDARDS

Due to the multiplicity of workflows, object types and content carriers/mediums, the introduction of common standards into the scholarly communication digital landscape becomes a subject of crucial importance. To this end, an increasing number of international organisations have been collaboratively working towards the effective regulation of all issues related to knowledge representation and content dissemination and reuse, by developing protocols for digital content and online communication processes.

The **Dublin Core Metadata Initiative**¹⁰ (DCMI) has long experience in the field of metadata standardization, and emerged as one of the main agents involved in monitoring, maintaining, and promoting standards. DCMI has adopted a federated structure, which comprises several communities and specialized Task Groups, committed to maintaining and updating metadata vocabularies. This collective effort leads to specific deliverables, updates of existing guidelines, and the adoption of additional recommendations and suggested terminology refinements for the **Dublin Core Schema**¹¹ (DC), a core set of vocabulary terms used in the identification of digital objects (images, texts, web pages, etc.). DC consists of 15 elements describing the content, carrying medium, licensing and other properties of digital objects, and has been recently supplemented by additional metadata elements as well as a set of controlled vocabularies for the interpretation of element values.

The **World Wide Web Consortium**¹² (W3C) is an extensive network of collaborating communities for the development of Web standards. Among the different working groups operating under the supervision of the W3 consortium, the Publishing Business Group ("Publishing BG") and the Publishing Working Group ("Publishing WG") are dedicated to the development of technologies and workflows

¹⁰ <http://dublincore.org/>

¹¹ <http://dublincore.org/documents/dces/>

¹² <https://www.w3.org/>

that render the Web into a suitable ecosystem for publishing. The joint mission of the two Groups is to enhance publication accessibility, usability, distribution, archiving, as well as achieve reliable cross-referencing.

W3C has developed standards and specifications for a spectrum of web-oriented processes, technologies and tools, including default standards for TCP/IP communication protocols. It also provides recommendations for a variety of web-based languages used for knowledge representation (**OWL** - Web Ontology Language)¹³, text (**XML** - Extensible Markup Language)¹⁴ and hyper text (**HTML** - HyperText Markup Language)¹⁵ markup. A main contribution of W3C comes in the form of the Resource Description Framework (**RDF**)¹⁶, a set of specifications that has evolved into a framework for information modeling.

Another body whose work relates to the implementation of electronic publishing standards is the **International Digital Publishing Forum (IDPF)**¹⁷. Its specific goal is to encourage the adoption of standards by identifying, evaluating and maintaining specifications for publishing workflows and technologies. IDPF also specializes in the development of applications and formats, such as the **EPUB**¹⁸ content publication standard that enables the creation and dissemination of various content types as digital publications.

EDItEUR¹⁹, an international group for the implementation of standards designed to support e-commerce activities in the publishing sector, provides recommendations covering such diverse areas as e-infrastructures, bibliographic information and licensing. EDItEUR has developed a family of machine- and human- readable XML formats for the transmission of publication metadata records. **ONIX for Books Product Information Message**²⁰ and the set of Onix for Subscription Products formats provide a consistent way of incorporating metadata into electronic resource management systems and provide to end users links to these resources, as well as information on their licensing and terms of use.

The **Research Data Alliance**²¹ (RDA) establishes standards to overcome current fragmentation within the research data landscape and facilitates the implementation of the FAIR data principles. As many other organisations involved in the field of standards, RDA Comprises several working groups, dedicated to the establishment of a common framework for data production and reuse in a variety of SSH and STEM disciplines. RDA regularly issues recommendations and guidelines, introduces best practices and updates in issues of data curation, exchange and dissemination. It also assists research communities in understanding and following optimal data publishing workflows and increases researchers' awareness of emerging standards and best practices.

¹³ <https://www.w3.org/OWL/>

¹⁴ <https://www.w3.org/XML/>

¹⁵ <https://www.w3.org/html/>

¹⁶ <https://www.w3.org/RDF/>

¹⁷ <http://idpf.org/>

¹⁸ <http://idpf.org/epub>

¹⁹ <http://www.editeur.org/>

²⁰ <http://www.editeur.org/83/Overview/>

²¹ <https://www.rd-alliance.org/>

The **Text Encoding Initiative**²² (TEI) is a long standing community of practice, composed of institutions and researchers committed to developing and updating standards for the annotation of digital texts, with a special focus on the Humanities and Linguistics. The TEI consortium provides guidelines and other resources (trainings, bibliography and TEI-adopted software) that have been widely used by cultural and academic institutions for the digital representation of texts.

Standards apply not only to content and metadata, but also to information integrity and publishing workflows. The **Committee on Publication Ethics**²³ (COPE) has released a series of core codes of conduct, with an aim to introduce documented practices, publication ethics guidelines and recommendations for editors and publishers. To support editorial teams in their effort to ensure integrity and transparency, COPE releases mandates addressing important aspects of the editorial and publishing processes, such as content reproducibility, licensing and issues of intellectual property, peer review and journal management.

3.3 OPERATIONAL AND TECHNICAL STANDARDS FOR E-INFRASTRUCTURES

This section focuses on technical and operational standards for publishing infrastructures, highlighting their importance in providing a digital scholarly communication framework that fosters content reuse and improved user experience.²⁴ Integrated publishing platforms perform a series of basic functions related to content and user management, metadata indexing, identification and interlinking of resources and contributors. As digital publishing gradually becomes a norm, support of online editorial workflows and interoperability have also emerged as essential features for publishing software.

3.3.1 Suggested Framework for the Implementation of Standards

The current diversity of workflows and operational models underpins the necessity for a global introduction of standards that will serve as a framework for shaping an integrated scholarly communication landscape. This may prove a rather difficult venture, and any sustainable approach should make provisions for these standards' scalable implementation and adjustment to the different infrastructure types and editorial workflows. The suggested framework for the adoption of standards across infrastructures identifies two different levels/layers for the introduction technical and operational improvements:

Platform/system level: the standard functionalities of publishing platforms should be deployed in a manner that allows basic functionalities such as a) content/metadata retrieval and disposal to third-party applications b) online browsing and retrieval of content c) access to metadata related to intellectual property issues and d) meta-search and access to content with persistent identifiers. To increase the potential of content reuse and effectively correspond to the needs of the research and publishing communities, e-infrastructures for scholarly publishing should also support long-term preservation schemes and generate content usage/access statistics.

Inter-platform/semantic level: in the digital scholarly communication landscape, semantic interoperability becomes an element of crucial importance, as it enables the design of advanced

²² <http://www.tei-c.org/>

²³ <https://publicationethics.org/>

²⁴ See also Stathopoulos & Houssos. "Specifications and interoperability features for open digital content" <http://helios-eie.ekt.gr/EIE/handle/10442/8887>

content identification, delivery and processing services. Communication across research infrastructures requires a) the provision of “meaningful”, (i.e. machine-readable) metadata b) the use of standardized ontologies and controlled vocabularies c) the use of widely adopted knowledge representation languages d) the compliance of metadata with a specific encoding and e) the introduction of a common set of principles for data interlinking in the Semantic Web.

At this introductory stage, the main goal of the Working Group is to trace the standards at platform and inter-platform level, and identify key areas for their implementation, as an essential step to ensure content quality, availability and discoverability.²⁵

3.3.2 Content Quality and Impact Assessment

In the context of this section, content refers to a) publicly available information about digital scholarly editions (e.g journals) and b) published scholarly content (e.g. monographs, journal articles) available via digital infrastructures.

In a recently published report²⁶, the Committee on Publication Ethics (COPE)²⁷, the Directory of Open Access Journals (DOAJ)²⁸, the Open Access Scholarly Publishers Association (OASPA)²⁹, and the World Association of Medical Editors (WAME)³⁰ defined a number of core principles of transparency and provided recommendations for a range of managerial and editorial practices for scientific publications. According to these recommendations, journal websites should provide adequate information about the journal’s identity, focus and scope, and avoid statements that might mislead authors and/or readers. Moreover, the names and affiliations of the journal’s editorial committee and scientific board should be provided in a manner that indicates their expertise in the relevant scientific field. The relevant report also suggests the inclusion of information on the journal’s peer review and editorial processes, with explicit statements on the terms of acceptance and the estimated time span for accepted articles’ publication. These recommendations could be adopted as business standards, and be applied during the design and development of web-based interfaces for academic editions.

Regarding the quality of published scientific content, a key element in ensuring academic transparency and research integrity is peer review. As peer review methods maintain quality standards and provide credibility to scientific editions, editorial teams should encourage the engagement of reviewers and assist them in conducting and communicating their review. Current and emerging peer review practices entail certain challenges for scholarly communication e-infrastructure, which should support all different types of peer review, keep track of and compile records of exchanged communication between engaged parts, store and provide access upon demand to all different versions of submitted manuscripts. With the emergence of open peer review as common practice, provisions should also be made for future introduction of web 2.0 functionalities in publishing platforms.

²⁵ A detailed list of tools and services for e-publishing infrastructures may be found in the relevant OPERAS white paper, available at: <https://zenodo.org/deposit/1324058>

²⁶ “Principles of Transparency and Best Practice in Scholarly Publishing”. Available at: https://publicationethics.org/files/Principles_of_Transparency_and_Best_Practice_in_Scholarly_Publishingv3.pdf

²⁷ <https://publicationethics.org/>

²⁸ <https://doaj.org/>

²⁹ <https://oaspa.org/>

³⁰ <http://www.wame.org/>

As for impact assessment, web-based publication and interlinking of scholarly resources opens new pathways in measuring the outreach of research outputs. While citations remain the most widely acknowledged medium for evaluating research, there is a growing trend to assess impact by documenting users' engagement with published content and scientific data. *Altmetrics* refers to a family of relevant indicators, such as the number of actions and user responses to published content (views, discussions, downloads), references and citations in external resources, even shares in social media platforms³¹. Thus, a comprehensive approach to digital publishing platforms should include measures to define altmetrics standards, and increase their technical capacity to provide usage-related statistics.

In sum, from an operational point of view, high content quality implies – at minimum – the implementation of complementary quality standards and editorial workflows that underpin the potential of digital infrastructures to serve as a venue of scholarly communication, support researchers' enhanced digital skills and encourage users' increased involvement during the peer review procedures. This could be accomplished by designing and implementing publishing models based on the “software as a service” (SaaS) concept, introducing validation criteria for content conformance with adopted quality standards, designing user friendly interfaces, enriching software functionalities with detailed guidelines (e.g. knowledgebase) to proactively support users, and introducing best practices for producing, reviewing and publishing scientific content.

3.3.3 Interoperability

In general terms, interoperability refers to the capacity of digital infrastructures and software to communicate in an automated manner and exchange reciprocally stored information and files. In technical terms, interoperability is achieved with the implementation of common metadata standards across systems, and supported by the introduction of open APIs and Web Services that enable data transfers under a globally applied communication protocol.

Limited interoperability has important implications in research and disturbs scholarly communication processes, as it significantly confines researchers' ability to exploit the full potential of web applications. These possible limitations in discovery, retrieval and dissemination of scientific resources and information need to be taken into account in the designing of e-infrastructures, and be addressed in combination with the prospective developments in research methods and information technologies.

As Almeida, Oliveira and Cruz (2011) suggest, Open Source and Open Standards play a crucial role in interoperability issues. A basic set of recommendations for the implementation of interoperability standards and e-infrastructure functionalities could be as follows:

Harvesting and aggregating features: e-infrastructures must be able to provide data to third party applications, through APIs (Application Programming Interfaces) that conform to appropriate protocols. Data should also be deliverable partially in clusters and/or metadata form, thus allowing combined harvesting with the implementation of certain criteria.

³¹ <https://www.altmetric.com/>

Data exchange: the main challenges that have to be addressed relate to the designing of a common communication framework that allows systems not only to exchange, but also identify data. This implies the use of common (and appropriate for each information type) metadata schemes, the availability of individual metadata records in a structured way (XML), and their compliance with several predefined formats to enable their incorporation into collective schemata.

Semantic interoperability: to ensure semantic interoperability, it is important for infrastructure providers to adopt appropriate knowledge representation languages, and established ontologies for the documentation of their digital resources, in compliance with the principles of Linked Open Data (LOD -Yu 2011, Alexiou et al. 2016). Moreover, semantic interoperability assumes the interlinking of each metadata element with a suitable equivalent in a predefined list of values (e.g. vocabulary, list of standard terms, thesaurus).

3.3.4 Availability

In addition to technical standards, interoperability encompasses several organizational aspects that affect the process of file management and information exchange. A sustainable approach for e-infrastructures should be based on a framework that enhances metadata availability and interlinking (Day 2005), while respecting restrictions deriving from scholarly communication regulations; it should, furthermore, comply with researchers' needs of discovering and accessing a) files and metadata b) resources, identifiers for resources and contributors and c) information on dissemination and reuse rights.

A combination of technical and operational specifications should allow for research-oriented added value services:

Resource and metadata: in addition to online browsing and content downloading, electronic infrastructures should also support advanced search options and combined content retrieval features.

Identification: the use of persistent identifiers for content, contributors, funding agents or institutions is essential, as it facilitates a series of meta-services based on proper element interlinking. In addition to providing and/or displaying as relevant metadata unique persistent identifiers for persons, organisations and digital objects, identification also involves long-term commitment to resolving digital resources.

Licensing: proper licensing is a key element for scholarly communication. Combined with the appropriate technical workflows, it enables optimal data flow across interlinked infrastructures and wide dissemination of research outputs and primary data. The use of CreativeCommons licenses for open access content prevents copyright infringements and allows authors to define the terms of reuse and distribution of their work. Licensing information should be clearly indicated in all published formats

Preservation: content preservation is an essential part of sustainable planning for research infrastructures. A feasible preservation mechanism should be based on provisions for at least one remote copy of digital objects and relevant metadata entries, as well as automated processes for remote backup of digital content. It should also be designed upon commonly applied preservation schemes (OAIS) and eventually incorporate future changes in technologies and data formats.

3.3.5 Processability

With the advent of digital methods and tools, research in the humanities shifts towards large-scale projects, often undertaken by multi-disciplinary, multi-institutional networks. This distributed production of knowledge drives digital workflows away from the basic functionalities of content uploading/downloading, and promotes online collaborative work in a cloud-based environment. It also introduces innovative research methods based on content mining and aggregation, text annotation and markup etc.

Digital research often produces deliverables in multiple formats, which should be supported by enriched workflows applied across publishing platforms. To effectively address issues stemming from the emergence of augmented and dynamic texts as communication medium, publication e-infrastructures should develop tools and workflows to support online/native authoring (e.g. Hyde, 2015), submission and peer review, as well as the conversion of semantic-based inputs into other file types available to end users.

4. WHERE OPERAS STANDS

As all partners have developed custom services and workflows, the identification of common operational and technical standards across the OPERAS network is a complex exercise. Nevertheless, standardization doesn't mean uniformization and, in fact, the existing diversity should not be considered as an obstacle to standardization, but rather as an opportunity to effectively address the emerging challenges.

4.1 MAPPING OF PARTNERS' PUBLISHING WORKFLOWS, TECHNICAL AND CONTENT QUALITY STANDARDS

Part of the work undertaken in the context of the OPERAS-D project related to the technical mapping of the OPERAS partners' infrastructures. The technical mapping report³² identified commonly applied practices or standards, and provided general recommendations towards the accomplishment of full interoperability between the consortium partners.

Regarding technical standards, there is a general and wide use of interoperable metadata schemata. As a common minimum standard within the Consortium, the report identifies the Dublin Core schema and a corresponding OAI repository. Due to the technical implementations proposed by the HIRMEOS project (presented below in detail), the display of unique identifiers (DOI for documents, ISSN/ISBN for digital editions, ORCID for authors) will soon become a standard in five important publishing and indexing platforms.

Nonetheless, not every consortium partner uses the same set of standards, especially in the case of content indexing standards (like BIC or LCSH), or dissemination standards (e.g. ONIX). There is also a low uptake of technologies like XML or RDF. The semantic interoperability through OWL and SKOS is also still to be built within the Consortium; this will be one of the main prospects to develop further standardization in OPERAS.

The three examples listed below (OpenEdition, EKT, OAPEN) provide a general framework of the services and workflows delivered by the OPERAS partners and indicate the challenges (and opportunities) related to the implementation of consortium-wide standards:

4.1.1 OpenEdition

Ingestion of formatted content to produce different publishing formats through a structured language.

Service and workflow description: OpenEdition (OE) publishes 4 types of content and manages a corresponding number of platforms for journals, books, academic blogs, scientific events. For books and journals, the general workflow consists of XML-TEI generation from .doc and .odt files, which are then imported into the OE CMS Lodel. Lodel is also used by the users to create new scientific events. The blogs are published with the use of templates in Wordpress-based websites.

The peer review process is not managed by the OE staff. The requirements in this field are that the journals should be peer-reviewed by publishers and the books should be examined by reading committees. The blogs and the events are free and published in open access. Journals and books are disseminated through a Freemium model: the HTML version is generally full OA, PDF and Epub files can be full OA or accessible by subscription.

³² "Technical Mapping of OPERAS Consortium - Annex to OPERAS Design Study". <https://doi.org/10.5281/zenodo.1247926>

Technical standards

- Metadata curation and indexing:
 - Subject indexing: partly by publishers / partly by OE staff
 - Indexing standards and tools:
 - Internal OE subject index
 - OST/ISI index (partially used)
 - Bisac index (only for Amazon and simil.)
 - Bic index (only for OAPEN DOAB - automated generation from Bisac)
 - Interoperable Standards:
 - DublinCore for OAI
 - Qualified DublinCore for OAI
 - METS for OAI
 - XML-TEI (chapter/article level)
 - Mets/Mods (book/issue level)
 - Full-text automated indexing (Solr)
- Unique identifiers: ISBN, ISSN, DOI, ORCID, Funding registry
- Search: faceted search on OE platforms
- Discoverability:
 - Distribution on OE platforms
 - Commercial distribution by third-party (<http://www.immateriel.fr/>)
 - Referencing mainly on Isidore (<http://recherche-isidore.fr/>), Google Scholar
- Output formats: HTML, PDF, EPUB

4.1.2 EKT ePublishing

Integrated process with peer-reviewing, enrichment and publishing using OJS software.

Service and workflow description: EKT ePublishing provides access to content and services from a single point on the web. It hosts three distinct platforms for journals <https://ejournals.epublishing.ekt.gr/>, monographs <http://ebooks.epublishing.ekt.gr/>, and proceedings series <https://eproceedings.epublishing.ekt.gr/>. Services include, most significantly, the organization, documentation and organized dissemination of content and metadata, training and consulting services on issues such as the standardization of editorial processes, intellectual property, the inclusion of content and metadata in content indexers and harvesters via interoperable systems, digitization and ingestion of digital content into the platform, as well as production of metadata for past issues.

Article and book submissions are accepted online. During the submission process, end users upload files in PDF, HTML or Word format and add related metadata. The peer review process is conducted online, under publishers' supervision. Article galleys are externally processed and uploaded on the platform for publication. Published content is available in open access and appropriately licensed.

Technical standards

- Metadata curation and indexing: manual metadata enrichment and optimization
- Metadata standards: DC for OAI, MARC/MARC21
- Unique identifiers: ORCID, DOI, Fundref, ISBN/ISSN
- Search and content discoverability: full-text search is provided at platform level. Content referenced in DOAJ, PKP Index Service, DOAB, Zenodo
- Output formats: PDF, HTML ePUB (books only)
- Preservation: PKP LOCKSS network

4.1.3 OAPEN

Integration of files through FTP and further metadata enrichment and dissemination.

OAPEN manages both a library and, together with OpenEdition, the Directory of Open Access Books (DOAB). The OAPEN library publishes books, which the publishers upload with their metadata files on a FTP server. For its selected partners, OAPEN retrieves the metadata through OAI-PMH. Direct upload of metadata is possible with .csv or ONIX files.

The DOAB disseminates the metadata of books. These books are always in full OA and have a peer-reviewing process validated. Metadata can be registered in various ways, using .csv files, ONIX files or through a manual action.

Technical standards

- Identifiers: DOI, ISBN, ORCID, Funding registry
- Content indexing:
 - OAPEN library: BIC subject headings
 - DOAB: LCC subject headings
- Discoverability:
 - Metadata feeds:
 - ONIX (3.0) – XML
 - MARC - MACHine-Readable Cataloging file
 - MARCXML – based on MARC 21 XML Schema
 - CSV – comma delimited text file
 - TSV - tab delimited text file
 - XML - optimised for import in Excel
 - Search engine optimization: The OAPEN Library website uses the schema.org model for books, a special mark-up used on the book's landing pages telling search engines that the OAPEN site contains books, and points to the title, author/s, etc. This data is used by Google Scholar, to index the contents of the OAPEN Library
- Output format: HTML, PDF

4.2 HIRMEOS

The project HIRMEOS, (the acronym stands for “High Integration of Research Monographs in the European Open Science Infrastructure”) is an EU funded project dedicated to the integration of high quality scientific content in the European Open Science ecosystem, with a special focus on the Social Sciences and the Humanities.

The project is undertaken by 9 members of the OPERAS network (research centres, university presses, university libraries and public foundations for the promotion of research), with common orientations towards the enhancement of Open Access through the development of European-wide infrastructures for scholarly communication. HIRMEOS comprises 7 WPs, of which 5 are exclusively technical, with a general scope to improve five important OA monograph publishing platforms, by designing and implementing common operational as well as technical standards, in light of their future incorporation into the European Open Science Cloud.

During this ongoing process, there are certain challenges to be met, mainly related to the different technologies, functionalities and features of the participating infrastructures. To this end, HIRMEOS will enhance their interoperability, by designing common services for the identification and validation of content and its metadata, as well as tools that enrich information and entity extraction. In the future, end users will annotate, extract and share content, while content providers and infrastructure administrators will be gathering usage data and metrics. HIRMEOS will also enhance the technical capacities of DOAB to import and ingest enriched and structured metadata, and also design peer reviewing validation processes.

As for the project’s impact, HIRMEOS will enhance the platforms’ capacity to serve as venues for the discovery of high quality scientific content as well as mediums for the provision of metadata to third party aggregators, such as the OPENAIRE infrastructure. In addition, HIRMEOS will establish criteria and standards for the validation of e-publishing platforms, along with procedures for the certification of trusted partners. Finally, the project stands next to existing initiatives, and aspires to have a catalyst effect in including more disciplines into the Open Science paradigm, widening its boundaries towards the SSH.

On technical level, if HIRMEOS aims at implementing existing standards (namely identifiers and annotations) on publishing platforms, it will also contribute to the emergence of new standards in two sectors, namely peer-reviewing and metrics:

Peer review: the work package dedicated to the enhancement of DOAB technical capacity will enable the Directory to assign standardized peer review types certificates to academic books that will lead to a clarification of that crucial domain for scientific quality.

Metrics: the work package will eventually provide a service that combines standardized measurements of usage of open access resources, whether in downloads/views and citations in the open web environment. Aiming at providing a, richer, more balanced, more transparent and definitely more open alternative to the de facto standard Impact Factor, HIRMEOS work package on metrics will entail the creation of a metrics service open to the whole community and supported by OPERAS infrastructure.

4.3 FUTURE ORIENTATIONS FOR A COMMON SET OF STANDARDS TO BE PROVIDED BY OPERAS

The current multiplicity of formats, publication versions and content types raises questions of accessibility and usability of digital scholarly output and entails new roles for infrastructure providers, who are actively participating in the design of policies and procedures with an expressed aim to ensure content availability. This is a complex exercise, as it implies the implementation of long-term strategies and business models for sustainable resource management, as well as the utilization of standards to prevent content from becoming obsolete, and ensure its wide dissemination and interoperability. One of the main challenges to be addressed is the current fragmentation and prevailing imbalances, in terms of delivered content quality and e-infrastructure capacities.

Change could only be brought about by acknowledging the importance of a common operational framework for digital publishing, and adopting a collaboration-oriented approach that brings together all parts engaged in scholarly communication. Thus, effective implementation of common standards is highly depended upon stakeholders' increased awareness and commitment towards more effective ways of conducting, presenting and communicating research.

The OPERAS network could play an instrumental role in this collaborative effort, as it comprises a significant number of European institutions with publishing experience and combined expertise in all aspects of digital scholarly communication. To these, one should add the partners' research/scientific profile and participation in EU-wide projects and networks of researchers, publishers, and e-infrastructure providers. OPERAS has the potential to coordinate a future initiative for the introduction of publishing standards, in the SSH and beyond. Through its extensive partnership structure, OPERAS could allocate standards requirements to different research or IT communities, contribute in specifying new standards and updating existing ones, and assist other organisations in practically implementing operational and technical recommendations.

This role implies a variety of challenges that need to be met, such as: the active involvement and coordination of different stakeholders; the identification of discipline-specific needs and dissemination processes; the effective promotion and uniform implementation of recently introduced standards. Thus, the realization of a centralized standardization mechanism supported by OPERAS requires extensive preparatory work towards the identification of current needs and existing standards, the inauguration of transdisciplinary communication networks and the consolidation of monitoring and validation processes.

This introductory phase could be followed by a multi-stage process leading to a community-wide adoption of standards. A first step would be aiming at establishing a common framework of minimum standards and best practices within the OPERAS consortium. To this end, the Standards Working Group could serve as a central node for incorporating existing quality requirements and workflows into a unified set of recommendations to be respected by all OPERAS members. In the long term, a sustainable approach for the introduction of operational and technical standards for electronic publishing should further a) take into account discipline-specific standards, needs and workflows, b) focus on extensive membership and build upon members' expertise and specialization. As described in section 3.2 of the present document, all institutions involved in standards modeling have adopted management practices and a federated structure to maintain their recommendations' credibility and sustainability. Given the foreseen expansion of its network, OPERAS could proceed in assigning standards-oriented tasks into specialized Working Groups (e.g. Platforms and Services WG, Best

Practices WG), under the overall supervision of the Standards WG and a commonly acknowledged framework for the validation and communication of recently introduced or updated standards.

As most of the OPERAS partners have established bonds with the research and publishing communities, these recommendations and suggested standards could be widely promoted to enhance researchers' awareness of the importance of a common and regulated scholarly communication framework. A final milestone in the standards implementation roadmap would be the establishment of continuous communication flows with these communities, so as to promptly identify and effectively address emerging needs or current trends. This, in turn, would further encourage the collaboration of the different OPERAS working groups (as well as between OPERAS and other bodies) in the field of standards development and lead to more concrete deliverables, such as software toolboxes for content enrichment and validation, and the documentation of best practices for publishing workflows.

OPERAS is on track to becoming a major infrastructure to support open scholarly communication. The consortium members invest on common principles of governance and sustainability, for the future development of a comprehensive set of services for researchers and publishers. At this introductory point, the OPERAS Working Groups work together towards drafting a roadmap for the implementation of best practices within the OPERAS network and beyond. In this context, the white paper on Common Standards provided general orientations for the adoption -in a structured and organised manner- of minimum standards across the composite landscape of scholarly publishing.

LIST OF ACRONYMS AND ABBREVIATIONS

COPE: Committee on Publication Ethics

DC: Dublin Core

DCMI: Dublin Core Metadata Initiative

DG RTD: Directorate General for Research and Innovation

DH: Digital Humanities

DMP: Data Management Plan

DOAB: Directory of Open Access Books

DOAJ: Directory of Open Access Journals

DOI: Digital Object Identifier

ERA: European Research Area

EOSC: European Open Science Cloud

FAIR (data): Findable, Accessible, Interoperable and Reusable

H2020: Horizon 2020 Work Programme

HTML: HyperText Markup Language

IDPF: International Digital Publishing Forum

LOCKSS (preservation system): Lots of Copies Keep Stuff Safe

LOD: Linked Open Data

MARC: Machine Readable Cataloguing

OA: Open Access

OAI: Open Archives Initiative

OAI-PMH: Open Archives Initiative – Protocol for Metadata Harvesting

OASPA: Open Access Scholarly Publishers Association

OJS: Open Journal Systems

ORCID: Open Researcher and Contributor ID

OS: Open Science

OSPP: Open Science Policy Platform

OWL: Web Ontology Language

PKP: Public Knowledge Project

RDA: Research Data Alliance

RDF: Resource Description Framework

SKOS: Simple Knowledge Organization System

SSH: Social Sciences and Humanities

STEM: Science, Technology, Engineering, Mathematics

TEI: Text Encoding Initiative

WAME: World Association of Medical Editors

W3C: World Wide Web Consortium

XML: Extensible Markup Language

REFERENCES

- Alexiou, G., Vahdati, S., Lange, C., Papastefanatos, G., & Lohmann, S. (2016). "OpenAIRE LOD services: scholarly communication data as linked data". In International Workshop on Semantic, Analytics, Visualization (pp. 45-50). Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-319-53637-8_6
- Almeida, F., Oliveira J., Cruz J. (2011). "Open Standards And Open Source: Enabling Interoperability". International Journal of Software Engineering & Applications (IJSEA) 2(1). <http://airccse.org/journal/ijsea/papers/0111ijsea01.pdf>
- Ball, C. E., Eyman, D. (2015). "Editorial Workflows for Multimedia-Rich Scholarship". Journal of Electronic Publishing 18(4). DOI: <http://dx.doi.org/10.3998/3336451.0018.406>
- Browning, S., Guedon, J., Kaplan, L. (2013). " Metadata and Open Access: Reliably Finding Content and Finding Reliable Content". Proceedings of the Charleston Library Conference. <http://dx.doi.org/10.5703/1288284315314>
- COPE, DOAJ, OASPA, WAME. "Principles of Transparency and Best Practice in Scholarly Publishing". <https://publicationethics.org/resources/guidelines-new/principles-transparency-and-best-practice-scholarly-publishing> (accessed 29/6/2018)
- Day, M,. (2005). "Metadata". In S.Ross, M.Day (eds) DCC Digital Curation Manual. <http://www.dcc.ac.uk/resource/curation-manual/chapters/metadata/> (accessed 29/6/2018)
- European Commission, Directorate General for Research and Innovation. (2016). "Open innovation, Open Science, Open to the World". <https://publications.europa.eu/en/publication-detail/-/publication/3213b335-1cbc-11e6-ba9a-01aa75ed71a1> (accessed 16/4/2018)
- Ferwerda, E., de Vries, S., Dumouchel, S., Gatti, R., Mounier, P., Radovic, D., Smaniotto, A., Stojanovski, J., & Waaijers, L. (2018). OPERAS Platforms and Services White Paper. Zenodo. <https://zenodo.org/deposit/1324058>
- Harley, D., Acord S., Earl-Novell, S., Lawrence, S., King, J. (2010) "Assessing the Future Landscape of Scholarly Communication: An Exploration of Faculty Values and Needs in Seven Disciplines". Berkeley: The Center for Studies in Higher Education. <https://escholarship.org/uc/item/15x7385g> (accessed 29/6/2018).
- Hutchens, C. (2013). "Open access metadata: current practices and proposed solutions". Learned Publishing, 26 (3) pp: 159-165. DOI: <http://dx.doi.org/10.1087/20130302>
- Hyde, A. (2015). "Open Access and Standards". Journal of Electronic Publishing 18(1). DOI: <http://dx.doi.org/10.3998/3336451.0018.117>
- Neylon, C., Pentz, E., Tananbaum, G. (2014). "Standardized Metadata Elements to Identify Access and License Information". Information Standards Quarterly,(2) pp: 35-37. <https://espace.curtin.edu.au/handle/20.500.11937/24522> (accessed 29/6/2018).

OPERAS Consortium (2017a). “Landscape Study on Open Access Publishing”. DOI: <https://doi.org/10.5281/zenodo.1299151>

OPERAS Consortium (2017b). “Technical Mapping of OPERAS Consortium”. DOI: <https://doi.org/10.5281/zenodo.1247926>

OPERAS Consortium. (2018). “Report on online survey on optimizing e-infrastructure”. https://f-origin.hypotheses.org/wp-content/blogs.dir/2465/files/2018/05/operas_online_survey_optimizing_e-infrastructure.pdf (accessed 30/6/2018).

Speicher, L., Armando, L., Bargheer, M., Eve, M., Fund, S., Leão, D., Mosterd, M., Pinter, F., & Souyioultzoglou, I. (2018). OPERAS Open Access Business Models White Paper. Zenodo. <https://zenodo.org/record/1323707>

Stathopoulos, P., Houssos, N.. (2013). “Specifications and interoperability features for open digital content” (trans. Georgiadis, H.). <http://helios-eie.ekt.gr/EIE/handle/10442/8887>

Tsoukala, V. (2015). “University based Open Access Publishing. State of Play”. SPARC Europe. http://sparceurope.org/wp-content/uploads/2015/12/SE_UPublishing_Report_0315.pdf (accessed 29/6/2018).

Warren, J. (2015). “Zen and the Art of Metadata Maintenance”. Journal of Electronic Publishing 18(3). <http://dx.doi.org/10.3998/3336451.0018.305>

Yu, L. (2011). “Linked Open Data”. A Developer’s Guide to the Semantic Web. Springer, Berlin, Heidelberg

List of Websites

Altmetric. <https://www.altmetric.com/>

Committee on Publication Ethics. <https://publicationethics.org/>

Directory of Open Access Journals (DOAJ). <https://doaj.org/>

Dublin Core metadata Initiative. <http://dublincore.org/>

EDiTEUR. <http://www.editeur.org/>

European Open Science Cloud Pilot Project. <https://eoscpilot.eu/node>

International Digital Publishing Forum. <http://idpf.org/>

Open Access Scholarly Publishers Association (OASPA). <https://oaspa.org/>

Programming Historian. <https://programminghistorian.org/>

Research Data Alliance (RDA). <https://www.rd-alliance.org/>

Text Encoding Initiative (TEI). <http://www.tei-c.org/>

World Association of Medical Editors. <http://www.wame.org/>

World Wide Web Consortium (W3C). <https://www.w3.org/>

ANNEXES

Annex 1: Poster of the Common Standards Working Group presented at the OPERAS Conference “Open Scholarly Communication in Europe. Addressing the Coordination Challenge”, 31 May - 1 June 2018, Athens

http://operas-eu.org/files/2018/05/operas_standards_poster.pdf



Common Standards

OPERAS Working Group

Focus and scope:

The current multiplicity of research practices, dissemination mediums and content types raises questions of accessibility and usability of digital scholarly output and entails new roles for publishers and infrastructure providers.

The Working Group on common standards

- Explores recently emerged workflows, mediums and technical standards related to academic publishing
- identifies key operational and technical aspects to be addressed by digital publishing infrastructures and service providers
- highlights the importance of common standards, and trace the standards required to ensure content quality, availability and discoverability

Key areas for the implementation of standards

1 | Content quality and impact assessment

- ▶ Validation criteria
- ▶ Users' increased engagement
- ▶ Transparent publishing workflows
- ▶ Metrics / altmetrics

2 | Interoperability

Content harvest / aggregation: data available to third party applications

Capacity for data exchange and identification across systems

Semantic interoperability

3 | Availability

Resources and metadata: support online browsing, content downloading, advanced search options and combined content retrieval features

Identification: persistent identifiers for content (DOI, handle), contributors (ORCID), funding agents (Fundref)

Licensing: proper licensing to prevent copyright infringements and define the terms of content reuse and distribution

Preservation: provision for remote copies and relevant metadata entries, automated processes for remote backup of digital content

4 | Processability

Enriched workflows applied across publishing platforms

- native authoring
- Online submission and peer review

Multiple output formats

Towards an integrated publishing ecosystem

- Mapping of partners' publishing workflows, technical and content quality standards
- HIRMEOS – technical improvements and alignment of publishing and indexing platforms
- Recommendations for a common set of standards to be applied and provided by OPERAS



**www.
operas-eu.
org**

WG members:

National Documentation Centre - GR (contact point),
OAPEN – NL, OpenEdition – FR, University of Milan - IT

WG Contact info:

Irakleitos Souyioultzoglou (EKT) Irakleitos@ekt.gr



Horizon 2020
European Union Funding
for Research & Innovation