e-InfraNet

CAPACITIES

e-InfraNet: 'Open' as the Default Modus Operandi for Research and Higher Education

European Network for co-ordination of policies and programmes on e-infrastructures (e-InfraNet)

e-InfraNet: ‘Open’ as the default modus operandi for research and higher education

“...an invention has to make sense in the world in which it is finished, not the world in which it is started”

Ray Kurzweil

Acknowledgements

This report has been developed within the project “European Network for co-ordination of policies and programmes on e-infrastructures” (e-InfraNet.eu). The project is an ERA-NET co-funded by the 7th Framework Programme of the European Commission. The members of the consortium are:

Participant organisation name	Short Name	Country
Higher Education Authority	HEA	IE
Higher Education Funding Council for England	HEFCE (JISC)	UK
SURF	SURF	NL
Foundation for Science and Technology	FCT	PT
Latvian Academy of Sciences	LAS	LV
Ministry of Economy and Competitiveness	MINECO	ES
CSC - IT Center for Science	CSC	FI
National Information Infrastructure Development Institute	NIIFI	HU
The Scientific and Technological Research Council of Turkey	TUBITAK	TR
The Israel-Europe R&D Directorate	ISERD	IL
The General Secretariat for Research and Technology	GSRT	GR
Department of Economy, Science and Innovation, Flemish Government	EWI	BE

Disclaimer: The information in this document is subject to change without notice. Company or product names mentioned in this document may be trademarks or registered trademarks of their respective companies.

Front page quote from “Open Source Paradigm Shift”, Tim O’Reilly, June 2004, http://tim.oreilly.com/lpt/a/4868

Illustrations: The public domain work of numerous illustrators was used to formulate some of the graphics in this publication. Their contribution is hereby acknowledged.

Author(s):	Lilian van der Vaart (SURF)
Co-authors/editors:	Marnix van Berchum (SURF), Rachel Bruce (JISC), Maureen Burgess (HEA), Gabriel Hanganu (JISC), Neil Jacobs (JISC), Damien Lecarpentier (CSC), Stephen Pinfield (Sheffield University), Paul Stokes (JISC)

Table of Contents

Acknowledgements 2

Table of Contents 3

Figures 4

Executive summary 5

1 Introduction 7

1.1 Objectives 7
1.2 Structure 8
1.3 Intended audience 9

2 Context and definitions 10

2.1 Context 10
2.2 The meaning of ‘Open’—Definitions 11
2.3 The meaning of ‘Open’—Motivations 13

3 State of Play—Where has ‘Open’ got us so far? 17

3.1 Open Access to research literature 18
3.2 Open Data 20
3.3 Open Educational Resources 22
3.4 Open Infrastructure and Open Standards 24
3.5 Open Source Software and Open Development 26
3.6 Open Education 28
3.7 Open Peer Review 30
3.8 Open Research 32
3.9 Open Innovation 34

4 Challenges 36

4.1 How open is ‘Open’? 36
4.2 Evaluation and reward or perverse triggers? 41
4.3 The Longer Term 42
4.4 Diversity 45
4.5 Skills 45

5 Ways forward 47

5.1 Research and higher education as an Open Enterprise? 47
5.2 Individual or coordinated development? 48
5.3 Implications for policy and implementation 49

6 Conclusions and recommendations 53

Annexes 56

A Stakeholder map 57

B Reports and projects consulted 62

B.1 Reports 62
B.2 Projects / Initiatives 70
B.3 Other sources 73

C Survey of sustainability models 76

D Data/Services model 77

Figures

Figure 1—one of many open access logos in common use 7
Figure 2—The range of 'Opens' 11
Figure 3 5 star system for open content 37
Figure 4 The 5 stars of linked data 37
Figure 5 The 5 stars of online journal access 37
Figure 6—Stakeholder impact—Research institutions 58
Figure 7—Stakeholder impact—Teachers/Tutors 58
Figure 8—Stakeholder impact—Researchers 58
Figure 9—Stakeholder impact—Learners (students. lifelong) 58
Figure 10—Stakeholder impact—Research funders 59
Figure 11—Stakeholder impact—Infrastructure providers 59
Figure 12—Stakeholder impact—Content providers 59
Figure 13—Stakeholder impact—Data centres 59
Figure 14—Stakeholder impact—Standards bodies 60
Figure 15—Stakeholder impact—Publishers 60
Figure 16—Stakeholder impact—Developer communities 60
Figure 17—Stakeholder impact—Commercial organisations 60
Figure 18—Stakeholder impact—Non-commercial organisations 61
Figure 19—Stakeholder impact—European Commission 61
Figure 20—Data/Services model 77

Executive summary

The document shows that 'Open' is an important 'modus operandi' for an effective and efficient research and higher education system. It is seen as an approach, not as an ideology or an end in itself. It fits this sector naturally since it optimizes the possibilities for the advancement of knowledge that is so necessary to tackle the increasing complexity and scale of the world's questions for Research. The big challenges require cross-disciplinary approaches, and conditions that nourish serendipity, unforeseen collaborations and re-combinations of available research outputs and data into new discoveries Openness in research *and* higher education is crucial to make this possible, to cut across the current barriers that prohibit bringing the necessary minds and computing power to known and new problems. There is by now sufficient evidence for the beneficial impact of 'Open', both within and outside academia.

Openness in both research *and* higher education is crucial to make this possible, to cut across the current barriers that prohibit bringing the necessary minds and computing power to known and new problems

This is increasingly being recognized in the sector, also by policymakers and funding organisations, as witnessed for example by the EC Communication and Recommendation on better access to scientific information[1] and the Royal Society's Report on 'Science as an Open Enterprise'[2]. To realize the full potential of 'Open' e-InfraNet recommends a broad policy framework that covers open access to content and infrastructure as well as open approaches to the further development of 'Open' itself, and to the way research and higher education are conducted. Some practical recommendations for implementation are also included.

To realize the full potential of 'Open' e-InfraNet recommends a broad policy framework that covers open access to content and infrastructure as well as open approaches to the further development of 'Open' itself, and to the way research and higher education are conducted.

For this broad 'Open' policy framework e-InfraNet has opted for a broad definition that describes the optimal situation for 'Open':

"Open means ensuring that there is little or no barrier to access for anyone who can, or wants to, contribute to a particular development or use its output."

The basis for the policy framework is an overview of the current 'Open' landscape outlining contexts, drivers, achievements and effects of the various 'opens', as well as a number of common issues. Because of this commonality, coordinating the vision and approach can benefit all 'opens' individually, *and* contribute to the development of 'Open' as the *default* modus operandi for the research and higher

[1] European Commission Research & Innovation – Science in Society http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1301, July 2012
[2] Science as an Open Enterprise http://royalsociety.org/policy/projects/science-public-enterprise/report/, June 2012

education sectors. A pragmatic approach to the implementation of the vision will ensure the necessary flexibility to adjust for the diversity in the various 'opens' themselves and in their geographic and disciplinary contexts.

In view of its findings, e-InfraNet makes the following recommendations for the adoption and implementation of a generic 'Open' policy framework:

1. e-InfraNet recommends that European and National governments and funding organisations, and research and higher education institutions adopt the policy that for publicly funded activities, 'Open' is the default modus operandi, for the physical infrastructure, the content and the use of both in open processes.
2. e-InfraNet recommends the adoption of the principles of the 'Open Development Methodology' for the further coordinated development of 'Open' in publicly funded projects. The EC and other funding organisations should build into their Funding requirements for projects to comply with these 'open development' principles. Institutions should build the 'open development' principles into their project management methods.
3. e-InfraNet recommends that the involvement and cooperation of researchers in the further development of 'Open' is actively stimulated and that barriers are removed through the adoption of appropriate evaluation and reward systems including qualitative and quantitative criteria that focus on open research communication in aid of the advancement of knowledge; and by support and advice in copyright and licensing issues
4. e-InfraNet recommends that researchers and other staff who work on 'Open' and in an 'open development process' are actively supported with appropriate training and support. Open Education initiatives should be developed as a vital, natural training ground for the 'researcher of the future'.
5. e-InfraNet recommends that interoperability of systems, services and content is continuously supported as a key component in the 'Open' e-infrastructure.
6. e-InfraNet recommends that the EC continues to invest in research and development work in the area of 'Open', with respect to policies, good practices and technical approaches.

With a clear and conscious choice for 'Open' as the default modus operandi in research and higher education, the EC, national governments, funders, infrastructure providers and institutions lay a firm foundation for a more coordinated and sustained development of open approaches. This choice will be a great support to the communities that shape the development. It will contribute towards more intensive, shared use of the valuable resources in the research and higher education system. It will also optimize the conditions for innovative Research and Development, both within and outside academia, that rely upon openness in infrastructure, content and work processes.

1 Introduction

The e-InfraNet project has been set up "to further the uptake and integration of innovative policy at national and EC level, in aid of greater effectiveness and efficiency of e-infrastructure and e-infrastructure programmes. Its approach strives for a good balance between general progress, joint actions and flexibility to adjust to diverse national requirements". e-InfraNet builds on existing policy initiatives and seeks to identify gaps in the uptake and integration of innovative policy at national and EU level.
Its long-term impact is geared towards more efficient use of resources, the adoption of best practice in national e-infrastructure programmes, a convergence of funding programmes and avoiding unnecessary overlap. e-InfraNet's approach will allow for progress to be made, while giving flexibility to adjust to national requirements and leave the route open to all to build on successful joint actions.[3]

e-InfraNet has chosen three focus areas in the wide-ranging field of research and higher education e-Infrastructures: Green ICT, Cloud Computing and 'Open'. This policy document is the output from the project's work on 'Open' which started mid-2011 with an inventory of current (national) programmes and policies in the partner countries, to establish the level of uptake of the 'Open' Agenda. It was found that work on 'Open' up until now has been a patchwork development of multiple open approaches, in response to different drivers in different contexts, that vary in maturity[4, 5]; there is not yet an 'Open' Agenda as such.

Given the present levels of maturity, a more coordinated approach in the further development of the various open approaches may prove more fruitful in contributing to an effective, efficient and sustainable research and higher education system. Therefore, e-InfraNet has undertaken to provide a document scoping a generic policy framework for 'Open' that would stimulate general progress in the development of open approaches, and allow for joint action where applicable, and flexibility to adjust to diverse requirements.

Figure 1—one of many open access logos in common use

1.1 Objectives

- a broad policy framework for 'Open'
- an open system of sharing and (re-)using content and other resources for more effective, efficient, innovative and innovating research and higher education

This document shows that 'Open' has an important role to play in an effective and efficient research and higher education system. The increasing complexity and scale of the world's questions for research, of the way research is conducted, and the impact this has on (the demands on) higher education require the traditional system to evolve its 'modus operandi', in order that it responds to this new complexity, scale and demand. The big challenges ask for cross-disciplinary approaches as well as conditions that nourish serendipity—unforeseen collaborations and re-combinations of available research outputs and data into new discoveries. Openness in research *and* higher education is crucial to make this possible, to cut across

[3] http://e-InfraNet.eu/
[4] We use 'opens' to indicate the various individual open approaches listed in section 2.2. The phrase 'Open' or 'Openness' is used to describe the combined 'opens' and 'Open' as an approach in general.
[5] Proceedings of the Workshop on Open (D.1.24) and the Overview of National Programmes on Open (D.1.1.4)

the current barriers that prohibit bringing the necessary minds and computing power to known and new problems—to paraphrase Donald Rumsfeld[6], unknown knowns.

'Open' fits the research and higher education sector naturally, since it optimizes the possibilities for advancement of knowledge that is so necessary to tackle the increasing complexity and scale of the world's questions for Research. There is sufficient evidence by now that 'Open' leads to better quality and more innovative research and improves re-use of data and other resources, both within and outside academia, so that it contributes to greater innovation and agility in the wider economic context[7]. It has important societal impacts, for instance in terms of improved education beyond institutional and national boundaries and also improved transparency and accountability. Last but not least, it contributes to greater efficiency because it enables shared use of expensive, scarce resources and avoids the duplication of investments and efforts. This is increasingly being recognized in the sector, also by policymakers and funding organisations, as witnessed for example by the European Commission's Communication[8] and Recommendation[9] on better access to scientific information (July 2012) and the Royal Society's Report on 'Science as an Open Enterprise'[10] (June 2012).

In line with e-InfraNet's aims, and to realize the full potential of 'Open', this policy document sets out to provide a broad policy framework for 'Open', which aims for an open system of sharing and (re)using content and other resources for more effective, efficient, innovative and innovating research and higher education. This broad framework covers open access to content and infrastructure as well as open approaches to the further development of 'Open' itself, and to the way research and higher education are conducted. It includes some practical recommendations for implementation.

This policy framework will enable a more coordinated and cross-fertilizing approach to the further development of the various 'opens', thus forming an 'Open' Agenda in research and higher education that better aligns the sector to the role it has to play in fostering growth in knowledge circulation and innovation in society.

1.2 Structure

Section 2 continues with a brief description of the context for this document. It further describes what is understood in this policy document by 'Open' and the various 'opens' constituting the 'Open' Agenda. It also summarizes the key motivations for 'Open'.

Section 3 sketches the present 'Open' landscape—primarily in Europe—in its development, aims, achievements, effects and current issues. This is based on desk research into key projects, activities and reports in Europe (particularly in partner countries) and seen in the wider global context.

Section 4 discusses a number of important issues identified as generic for 'Open'. Based on these issues, Section 5 proposes ways forward for a more coordinated 'Open' Agenda. Section 6 summarizes e-InfraNet's main recommendations. The focus is on necessary shifts in perspectives and on approaches which identify ways out of present debates & dilemmas and leave scope for diversity and phased realization.

[6] http://en.wikipedia.org/wiki/There_are_known_knowns
[7] http://open-access.net/de_en/general_information/pros_and_cons_of_open_access/
[8] Towards better access to scientific information: Boosting the benefits of public investments in research http://ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-towards-better-access-to-scientific-information_en.pdf
[9] COMMISSION RECOMMENDATION of 17.7.2012 on access to and preservation of scientific information http://ec.europa.eu/research/science-society/document_library/pdf_06/recommendation-access-and-preservation-scientific-information_en.pdf
[10] Science as an Open enterprise http://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf

1.3 Intended audience

This document is written primarily for policymakers and policy influencers at EU and national levels, as well as for other funding organisations, e-Infrastructure providers and research & higher education institutions in Europe.

In addition, it is aimed at managers of research, education and e-Infrastructures programmes and projects that may encompass one or more of the 'opens' discussed in the document.

2 Context and definitions

2.1 Context

e-InfraNet's work in the 'Open' Focus area so far has led to two main conclusions:

1. Work on 'Open' up until now has been a variegated development of different open approaches that have grown up in distinct contexts in response to different drivers and that vary in maturity; there is not yet an 'Open' Agenda as such.
2. Given the present levels of maturity, there is a need and scope for a more coordinated approach in the further development of the various open approaches; a generic 'Open' policy framework may prove more fruitful for the further development of all open approaches to contribute to an effective, efficient and sustainable research and higher education system.

The e-InfraNet Action Plan for Open[11] identifies the production of a policy document outlining such a framework as its main deliverable.

The approach recommended here aligns with (and potentially enhances) other EU initiatives. The flagship initiatives 'Innovation Union'[12] and 'Digital Agenda for Europe'[13] form the wider policy context for e-InfraNet's work on 'Open', as does the development of the European Research Area (ERA) as the Europe-wide space for knowledge and technologies. 'Open' as an approach has an important role to play in realising innovation. The realisation of the Digital Agenda and ERA also, to a large extent, build upon the concept of 'Openness'—openness of content, infrastructure and processes.

In the context of this document, 'Open' is therefore seen as **an approach, a modus operandi**[14]**,** not as an end in itself or as an ideology. As is stated in the *OpenAIRE Studies on Subject-Specific Requirements for Open Access Infrastructure*[14], 'infrastructure' is a way of sharing (often scarce and expensive) resources among researchers; and this sharing can lead to synergies and developments that would otherwise not have been possible[15]. 'Open' also implies and facilitates sharing. An 'Open e-Infrastructure' thus maximises "the degree to which the sharing of resources can be exploited"[16]. In this way, 'Open' can contribute to greater efficiency and economic value of the e-Infrastructure, and it provides important and necessary conditions for innovation.

e-InfraNet focuses on 'Open' in the research and higher education sector. Online work, collaboration and education are a growing reality in this sector, and so is 'Open'. Work carried out in this sector is also a key driver for wider innovation, with the sector acting as a supplier of people, ideas, answers and approaches. Thus 'Open' in this sector potentially has a double impact: it contributes to innovation in the research and higher education sector itself, as well as to innovation in other sectors. This potential double impact makes it all the more important to look at ways to strengthen the development of 'Open'.

[11] Action Plan for Open: D.2.13
[12] Europe 2020 Flagship Initiative: Innovation Union COM(2010) 546 final
http://ec.europa.eu/research/innovation-union/pdf/innovation-union-communication_en.pdf
[13] A Digital Agenda for Europe COM(2010) 245 final
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0245:FIN:EN:PDF
[14] Studies on Subject- Specific Requirements for Open Access Infrastructure, Edited by Christian Meier zu Verl and Wolfram Horstmann, http://pub.uni-bielefeld.de/publication/2445229
[15] The Studies mention the Human Genome Project as an example (p. 366). 'Infrastructure' in this context is perceived to include content (see also section 5.2)
[16] Studies on Subject-Specific Requirements, p. 367

2.2 The meaning of 'Open'—Definitions

2.2.1 The range of 'opens'

The phrase 'Open' encompasses a range of open approaches including the following

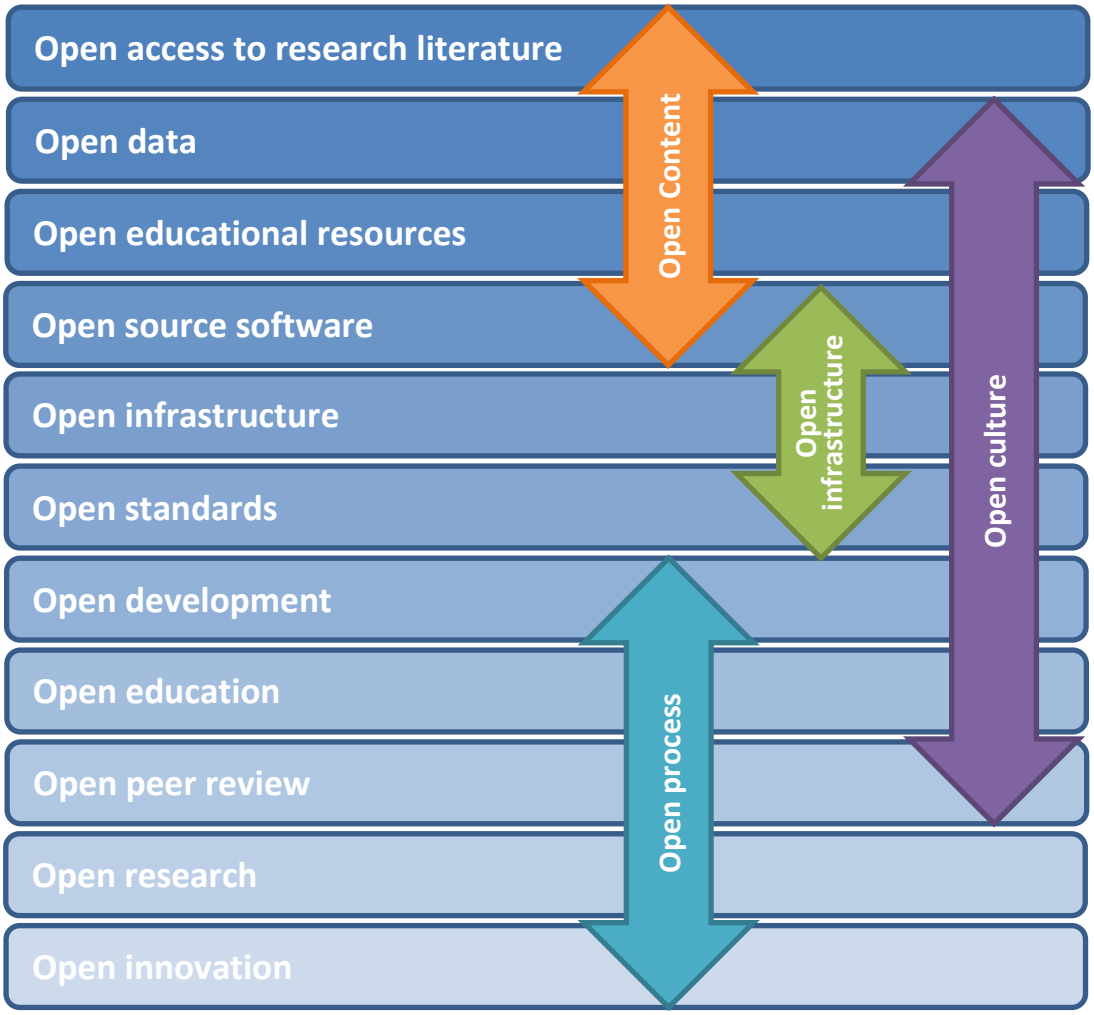

Figure 2—The range of 'Opens'

Together, these open approaches contribute to an open research and higher education system, and more generally, to open borders, open opportunity, and innovation in the wider sense of the word.

The various 'Opens' differ in context, drivers and in maturity, ranging from 'established' (such as Open Source Software—OSS—and OSS development) to experimental (for example open peer review and open research). This difference in maturity is an inherent characteristic of an 'Open' system; it has to, by definition, remain open to new developments.

2.2.2 'Open' defined

There are individual definitions for each of the 'open' approaches. Indeed there are often more than one. In section 3 one or two definitions are given for each. Thought-provoking work on definitions has been undertaken by the Open Knowledge Definition Project.[17].

In the context of this wide-ranging document a very broad definition of 'Open' has been chosen based on the one developed by JISC CETIS[18]. It describes the optimal situation for 'Open': [19]

> *"Open means ensuring that there is little or no barrier to access for anyone who can, or wants to, contribute to a particular development or use its output."*

This definition has the advantage that it

- is simple
- contains or implies the important attributes 'access', '(re)use' and 'participation'
- is easily and recognizably applicable to the whole range of open approaches
- is non-discriminatory
- leaves space for controlled access when that is essential
- is as applicable within the e-infrastructure as outside it

The danger of simplicity in the definition is that it hides the factors that account for the complexity encountered when implementing 'Open'. The main concepts are as follows:

- ***"little or no barrier to access"*** means there are little or no technological[20], organisational[21], financial[22], legal[23], or even cultural[24] restrictions to access. It also implies that access remains possible over time[25].
- ***"for anyone who can, or wants to"*** means whether (s)he is a regular participant in the research & higher education system or not, and whether (s)he actually contributes/(re)uses or not.
- ***"contribute to a particular development or use its output"*** means that 'little or no barrier to access" extends to "little or no barrier to participate in development and/or use the results of that development". It requires that outputs are available in their entirety (full text, complete data, source code and so on), in formats that allow processing by humans and machines; that this remain the case over time; and that access, participation and (re)use can be immediate. It also requires that full documentation is available to enable understanding of what has been made open, to allow for appropriate (re)use.

2.2.3 'Open' as Content, Process, Infrastructure and Culture

The definition above indicates that 'Open' does not stop at 'open access', whether to content[26] or to e-Infrastructure[27]. 'Open' also implies and encompasses activity: use, reuse, participation in creative and

[17] http://opendefinition.org
[18] Definition by Wilbert Kraan, CETIS Assistant Director, http://jisc.cetis.ac.uk/topic/open
[19] Section 4 goes into the issue of the necessity and legitimacy of restrictions to 'Open'
[20] For example those engendered by standards, formats, authorization & authentication
[21] Access should be time & place independent and there should be seamless access independent of organisation
[22] No upfront payment by end-user
[23] For example rights & permissions, licensing
[24] For example notions about sharing data and what constitutes 'good research'; such cultural aspects are outside the e-Infrastructure as such, but do influence its character
[25] Content is preserved and guaranteed to be available through the use of interoperable e-infrastructure hardware and software

development processes. The availability of and access to e-infrastructures and content are necessary conditions for efficiency and effectiveness in modern research and higher education. For sustained and sustainable development and innovation—both within and outside research and higher education—open participatory and collaborative approaches are also required. As the availability of and access to content and infrastructural resources increases, the need for and use of 'open processes'[28] becomes more evident. Where 'open content' is used and produced in 'open processes' within an open infrastructural setting, a culture of 'openness'[29] gradually emerges.

2.3 The meaning of 'Open'—Motivations

Since the inception of the various 'opens' considerable evidence has been gathered relating to the benefits and challenges involved. Even though the challenges of and perceptions about 'Open' may differ, levels of awareness among communities of stakeholders are considered to be high enough not to have to give a comprehensive account of the benefits of 'Open' in this document. However, a general summary of the main motivations for 'Open' is beneficial in this context; Annex A provides an overview of how various stakeholders are impacted in different ways.

In this document, challenges are seen as the problems, the seeking of solutions to which drives the further development of 'Open'. Section 3 identifies important ones for the different 'opens' and Section 4 discusses the common themes among them.

The main motivations for 'Open' can be summarised in three important, and partly overlapping, categories:

- research & development,
- societal/political, and
- economic

2.3.1 Research & Development

2.3.1.1 Greater visibility and impact from reuse of material

Easier and earlier discovery of and access to content leads to greater impact. It enables recombining material (often in unanticipated ways). Curating it adequately ensures greater longevity—and therefore long-term reuse—as well as large scale data sharing research, data and text mining. It also improves the timeliness of communication. Examples of this were given as early as 1995 by the Committee on Geophysical and Environmental Data of the US National Research Council in their communication "On the full and open exchange of scientific data". They wrote:

> *Experience has shown that increased access to scientific data, information, and related products has often led to significant scientific discoveries and the opportunity for educational enhancement. For example, the declassification of GEOSAT (a U.S. Navy geodetic satellite) data below 30 degrees south latitude led to a breakthrough in the study of global ocean floor topography and ocean sediment thickness (Smith and Sandwell, 1994a). The researchers also produced a global sea floor topography map that is being distributed internationally through*

[26] Such as literature, (primary) data, educational resources, bibliographies, software, images, video, audio etc., normally online—the actual materials available for access and (re)use
[27] Exemplified by software, standards, authorization & authentication—the technology/system enabling open content and open process
[28] As used for example in open source software development, open review, open research, open learning, open innovation—the way in which the materials are dealt with in research & HE—implying a participatory process
[29] The context in which (and the mind-set with which) the infrastructure, processes and content are implemented, carried out and used

> *the World Data Center system (Smith and Sandwell, 1994b). An example where not only the scientific community but the general public was engaged, was the near-real-time monitoring of Shoemaker-Levy comet fragments colliding with Jupiter. Impact phenomena from around the world were shared over the Internet, allowing astronomers an unprecedented opportunity to continuously modify their plans to make the optimal observations (Kerr, 1994). A third example concerns the transmission of real-time weather information into elementary and high schools with the goal of fostering science education. A one-year pilot program by the American Meteorological Society was so successful in engaging teachers and students on problems ranging from science to social studies that it will be expanded to include other types of environmental information.*
>
> *(Geer et al., 1995)*[30]

More recent examples can be found at the University of Manchester where researchers developed a text-mining technique to map Open Access articles to specific gene data, allowing new and unexpected discoveries relevant to genetic disorders. And the example of White Design (Bristol-based architects) shows the effects of successful knowledge transfer across the academic border. They obtained data on carbon content for low carbon buildings using straw construction in an OA paper, which creates the potential for a low carbon offset market that could, eventually, be worth millions of pounds[31].

2.3.1.2 Improved quality

The greater visibility of material enables the input from others via review creating a virtuous circle which, in turn leads, to improving quality of learning, research, software and administration. Furthermore because of concerns around reputation, people are more likely to share better quality work when it is openly available.

A typical example would be that illustrated by the Polymath Project. Tim Gowers, a mathematician, posted a problem on his blog and this 'open science' resulted in a collaboration of 27 contributors and 800 comments which solved the problem better and more quickly and pushed the research forward in a way he could never have achieved alone[32].

Error detection is also quicker and easier, as is shown in the example of the Committee on Geophysical and Environmental Data concerning the Advanced Very High Resolution Radiometer (AVHRR) satellite: "delays in obtaining data allowed a calibration error to go undetected for months… …Misinterpretation of these data led to erroneous estimates of global warming (Reynolds, 1993).[33]

2.3.1.3 Enhancement and protection of reputation and trust in institutions

Use of open courseware has been about demonstrating the quality of education on offer and about strengthening reputation and international connections (as in the case of MIT). Another example is that of research where it is shared openly. The UK House of Commons inquiry into climate data reinforced the 1995 messages of the Committee on Geophysical and Environmental data when it concluded that "…climate scientists need to take steps to make available all the data that support their work and full methodological workings, including their computer codes."[34]

[30] http://www.nap.edu/readingroom/books/exch/exch.html
[31] UK OAIG report: http://open-access.org.uk/2011/10/24/how-can-universities-support-economic-growth-and-innovation-take-the-open-road
[32] http://michaelnielsen.org/blog/the-polymath-project-scope-of-participation/
[33] http://www.nap.edu/readingroom.php?book=exch&page=summary.html
[34] http://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/inquiries/uea/

2.3.2 Societal/political

2.3.2.1 Innovation and agility

The Open Educational Quality Initiative "Beyond OER" report found that the use of Open Educational Resources and the implementation of open educational practices lead to innovations in pedagogical terms and in learning strategies, in addition to the innovations at an institutional level. Furthermore, the MITx example (see Section 3) shows societal impact in that it very effectively breaks down traditional borders in place, time, age as well as the borders between research and education. The White Design and Polymath projects mentioned above clearly show the societal impact as well, in environmental gains and effective use of available knowledge outside the traditional academic boundaries.

2.3.2.2 Transparency and accountability

Transparency is an important motivation in open government and open government data, as a means for improved democratic control of national, regional and local governmental organisations. Transparency is equally important in research and higher education as a precondition for accountability of spent (public) funding. But it is also important for detection of errors, plagiarism or fraud. The AVHRR satellite calibration error mentioned above is an example of error detection. A recent example of a significant measurement error concerns the preliminary findings of a CERN research team about neutrinos travelling faster than the speed of light. Examples of fraud are Hwang's faked stem cell cloning research published in Science[35], and Stapel's social psychology research based on fabricated data[36]; open availability of data would have given the opportunity for checking and testing the validity of the data and the research outcomes based on them.

2.3.2.3 Democracy

Democracy is another motivation stemming from the area of open government (data) that has importance for research and higher education. Open Educational Resources and practices have a democratizing effect on education, as is shown by the MIT example (Section 3); and a better educated population has an important societal impact.

Democratization of research is taking place in open research and citizen science, where members of the public take part in research projects. Galaxy Zoo[37] and Open Source Drug Discovery for Malaria[38] provide recent examples; the project instantiated by the Committee on Geophysical and Environmental Data that involved observing Shoemaker-Levy comet fragments collisions with Jupiter is an older one[39].

Democratization is even more evident in open research projects where members of the public actually influence the aims and scope of the research project, as happens for example in the Citizen Cyberscience programme[40]. Developments like this have the potential to reduce the 'ivory tower'-quality of research by making use of available knowledge outside academia, bringing in research questions based on practical needs, and making the results available publicly without claims of ownership.

[35] http://en.wikipedia.org/wiki/Hwang_Woo-suk , http://www.nytimes.com/2006/05/12/world/asia/12korea.html?_r=0
[36] http://en.wikipedia.org/wiki/Diederik_Stapel, http://universonline.nl/2012/11/28/levelt-report-fraud-detected-in-55-publications/
[37] http://www.galaxyzoo.org/
[38] http://openwetware.org/wiki/Open_Source_Drug_Discovery
[39] http://onlinelibrary.wiley.com/doi/10.1029/94EO00965/pdf
[40] http://citizencyberscience.net/

2.3.3 Economic

2.3.3.1 Business opportunities

Open access to research results enables commercial organisations (especially SMEs) to find new business opportunities and innovate. The White Design example mentioned above shows clearly how research data on low carbon content could be turned into an application of great value to society as well as a business opportunity for the company itself. In the US, open meteorological data has enabled a range of local commercial weather services[41]. AppliSci (a specialist SME supporting big pharmaceutical and healthcare companies with leading edge research services) found directly relevant 'process' evidence in a critical OA article that led to a development thought capable of opening the way for an entirely new medical application for treatment of a rare infectious disease, with a market potential estimated at tens of millions of dollars[42].

2.3.3.2 Cost effectiveness

Open approaches can enable efficient use of expensive resources, shared approaches, reduce duplication of effort and can save time and support collaboration. There are many examples where open research, open access, open source software solutions and open infrastructure can contribute to being more cost effective. Sharing openly on the Web is the most obvious example. Open approaches in e-Infrastructures such as European grids, High Performance Computing and Cloud Computing can provide a much more powerful system in a much more cost effective way than any 'closed' system at a national level could offer. The Open Access model is pursued because of the great inefficiencies in the current scholarly communications system: it is all too common to find researchers at a university where the research is actually produced with no access to it; the contribution of universities to the process is not accounted for in the high subscription costs; access to the results of research is slow. In short, the current system is not cost effective and open approaches can alleviate this.

[41] http://www.noaa.gov/partnershippolicy/

[42] UK OAIG report: http://open-access.org.uk/2011/10/24/how-can-universities-support-economic-growth-and-innovation-take-the-open-road http://open-access.org.uk/wp-content/uploads/2011/10/OAIG_Benefits_OA_PrivateSector.pdf

3 State of Play—Where has ‘Open’ got us so far?

To understand how a generic ‘Open’ policy framework may strengthen the further development of the various ‘opens’ individually and jointly, insight is needed in the main achievements and effects of developments so far, as well as in the important issues to be tackled. The desk research relating to projects and literature undertaken for this document was not exhaustive. The authors chose rather to look for the most recent and representative. The following section covers the main findings and impressions. Annex B outlines the literature and projects consulted, with brief summaries and links to their websites for further reference.

The primary focus in this section concerns the developments in Europe; for some topics, non-European developments are also referenced.

3.1 Open Access to research literature

Open Access (OA) to Research Literature	
Definition	• The practice of providing unrestricted access via the Internet to peer-reviewed scholarly journal articles. OA is also increasingly being implemented for theses, scholarly monographs and book chapters[43]. • Free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited[44].
Roots	• Online archives of preprints or working papers like arXiv[45] in Physics and RePEc[46] in Economics • Dissatisfaction with price models of scientific publishers
Aims/ motivations	• Easy & early exchange of information amongst researchers • Open up research literature to the world, contribute to innovation • Find effective & efficient alternatives to subscription models
Key developments / projects / policy drivers	• OAI-PMH protocol[47] • Budapest Open Access Initiative and the Berlin Declaration[48] • DRIVER / DRIVER II[49] • EU OA Pilot and OpenAIRE(plus)[50] • OA journal Publishing and OA dissemination via repositories

[43] http://en.wikipedia.org/wiki/Open_access
[44] Budapest Open Access Initiative http://www.opensocietyfoundations.org/openaccess/read
[45] http://arxiv.org/
[46] http://repec.org/
[47] http://www.openarchives.org/pmh/
[48] http://oa.mpg.de/lang/en-uk/berlin-prozess/berliner-erklarung/
[49] http://www.driver-repository.eu/
[50] http://www.openaire.eu/en/component/content/article/326-openaireplus-press-release

Achievements/ effects	• OA-supporting (e-)infrastructure being built / in place in most European countries • Policy development at EC-level and in countries is growing and boosts development • Technical support and good practices shared through European projects/initiatives • OA Success stories are increasingly available • OA via institutional repositories widens access to all kinds of research literature/results in addition over and above access via OA journal publishing • Open, distributed network approaches using existing infrastructure contributes to sustainability

Issues	• Uneven repository infrastructure development in the European countries and a range of use cases not yet well supported by repositories due to lack of appropriate interoperability
	• Uneven and sometimes slow policy development, especially at national level; policy initiatives are often at institutional level, and / or at funding bodies level
	• In OA, the many questions concerning the (long-term) preservation of digital research materials and the division of responsibilities regarding preservation may have an even deeper impact than elsewhere
	• Evidence for growth and impact of OA is hard to measure. The required work on metrics / usage statistics in distributed/federated network of systems is in its early stages
	• The continued claim of lack of awareness among researchers about Open Access is heard in many OA projects
	• The three obstacles mentioned most often are:
	  ○ copyright and licensing issues
	  ○ the continued practice of career assessment for researchers based on numbers of publications in high Impact Factor journals
	  ○ the funding of OA publication article process charges

3.2 Open Data

Open Data	
Definition	• A piece of data is open if anyone is free to use, reuse, and redistribute it—subject only, at most, to the requirement to attribute and/or share-like[51].
Roots	• Advocacy by World Meteorological Organization (WMO)[52] and the Committee on Geophysical and Environmental Data of the US National Research Council[53] to open up access to scientific data • Open government data • 'Data deluge'
Aims/ motivations	• To resist commercial interests in changing the practice of "free and unrestricted exchange" of data • Deal with and optimize use of increasing amounts of available data
Key developments / projects / policy drivers	• OECD Principles and Guidelines for Access to Research Data from Public Funding (2007)[54] • Report Riding the Wave: How Europe can gain from the rising tide of scientific data[55] • European Strategy Forum on Research Infrastructures (ESFRI)[56], and numerous Research Infrastructures (e.g. ENVRI, DASISH, CRISP, BiomedBridges) • EUDAT[57] • OpenAIREplus • DataCite[58] • RDA[59]

[51] Open Definition http://opendefinition.org/
[52] http://www.wmo.int/pages/index_en.html
[53] http://www.nationalacademies.org/nrc/index.html
[54] http://www.oecd.org/science/scienceandtechnologypolicy/oecdprinciplesandguidelinesforaccesstoresearchdatafrompublicfunding.htm
[55] http://ec.europa.eu/information_society/newsroom/cf/itemlongdetail.cfm?item_id=6204
[56] http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri
[57] http://www.eudat.eu/
[58] http://datacite.org/
[59] http://www.rda-jsc.org/rda.html

Achievements/ effects	• In various places, work on components of a Data Infrastructure is going on and the Collaborative Data Infrastructure (CDI) called upon by the High Level Expert Group on Scientific Data is emerging. • Scientific communities are increasingly organised, bridging with other disciplines and across national borders. The emergence of pan-European research infrastructures (RIs) in almost all scientific fields reflects this trend. • Identifying and developing common solutions in the domains of environmental sciences, social science and humanities, physical sciences, and biomedical sciences. • Overarching structures at European level cooperate internationally on interoperability and standardization issues in the architecture. • In various regions of the globe there is now a shared understanding that solutions must be global and that the development of an integrated and interoperable data domain can only be achieved through increasing global collaboration. The "Research Data Alliance" (RDA) is a recent international initiative aiming to steer the efforts in this area.
Issues	• Very uneven development in infrastructure, support and policy development across and within countries • Many uncertainties with respect to long-term preservation. • The growth and impact of Open Data are hard to measure. • Greater complications in the data field than in OA to research literature are created for example by the variety of sources, repositories, data sets structures, data formats and metadata, making it hard to combine data (sets) and hampering accessibility. • The desirability of 'Openness' of data is not uncontested; questions and objections concern security and legal issues (IPR, licensing, privacy), competition versus cooperation between researchers and research groups, and implications for career assessment; with the attempts to create a generic collaborative data infrastructure that can, for instance, facilitate interdisciplinary and open research, the complexity of and sensitivity to these issues increases. • Subject specificity is mentioned as an obstacle and a reason not to adopt a generic 'open' policy in the case of data.

3.3 Open Educational Resources

Open Educational Resources (OER)	
Definition	• Open Educational Resources (OER) are learning materials that are freely available to use, remix, and redistribute.[60] • Freely accessible, openly formatted and openly licensed documents and media that are useful for teaching, learning, education, assessment and research purposes[61].
Roots	• Developments around digital learning objects • Open and Distance Learning
Aims/ motivations	• Reuse of digital materials in a variety of pedagogical situations • Application of the principles of the open source / free software movements to content • Using technology to improve, and greatly widen access to education • Combining the 'technically possible' and the 'educationally desirable'

[60] Creative commons, http://wiki.creativecommons.org/OER
[61] Wikipedia ,http://en.wikipedia.org/wiki/Open_educational_resources

Key developments / projects / policy drivers	• Several standards efforts (for example with respect to metadata, content exchange, and other standards necessary for discovery and reuse of digital educational content) • Use of Creative Commons Licences in OER • OpenCourseware[62] (OCW, by MIT in 2000) and OpenCourseWare Consortium[63] • Connexions[64] (Rice University) • Wikipedia/Wiktionary • 'MOOC' (Massive open online course) • OER Commons[65] • Paris OER Declaration (UNESCO, 2012)[66] • MITx[67] and edX (2012)[68]
Achievements/ effects	• OER development taking place all over the world, spreading quickly • Attraction of more students in a relatively easy way • Opening up quality education to students who would otherwise not have access to it • Enhancing educational institutions reputation • Contributing to innovation of education and learning as a practice, and as a field of research, with respect to technologies, processes as well as content, by virtue of the Open Development Methodology which is part of OER development.
Issues	• Funding models: how to finance start-ups of OER initiatives and how to make these initiatives sustainable after their initial stage • To a lesser degree, IPR, non-compatible licenses • Discoverability • Recognition • Many initiatives rather 'local' • New educational 'imperialism' (the developed world taking the lead in educating the developing world)

[62] http://ocw.mit.edu/index.htm
[63] http://www.ocwconsortium.org/
[64] http://cnx.org/
[65] http://www.oercommons.org/
[66] http://www.unesco.org/new/en/communication-and-information/resources/news-and-in-focus-articles/all-news/news/unesco_world_oer_congress_releases_2012_paris_oer_declaration/
[67] http://mitx.org/
[68] https://www.edx.org/

3.4 Open Infrastructure and Open Standards

Open Infrastructure and Open Standards	
Definition	• An open standard is a standard that is publicly available and has various rights to use associated with it[69], and may also have various properties of how it was designed (for example, open process). There is no single definition and interpretations vary with usage.[70] • An open infrastructure allows for the aggregation of collective action online without technological, legal and economic hindrances[71]
Roots	• The 'global information infrastructure (GII)' emerging in the 1990s, promoted as an 'open system'. • EC-driven development of pan-European e-infrastructure for European research and education since late 1990s • European telecoms liberalisation legislation (1999)
Aims/ motivations	• Provision of a level playing field and prevention of dominance and lock-in by a few commercial parties • Focus on interoperability and open standards • Support for multiple communities in research and higher education, to face international competition and strive for European excellence

[69] The QR code standard is a good example, it is open ISO standard (ISO/IEC18004), but the patent belongs to Denso wave, http://www.qrcode.com/en/qrstandard.html
[70] Wikipedia, http://en.wikipedia.org/wiki/Open_standard
[71] Paraphrased from FCForum, http://fcforum.net/sustainable-models-for-creativity/how-to-manual#collaboration

Key developments / projects / policy drivers	• European Research Area (ERA)[72] • The GÉANT network[73], provided on behalf of the European consortium of National Research and Education Networks (NRENs) • The European Grid Initiative (EGI)[74], in collaboration with National Grid Initiatives (NGIs). • DEISA[75] and PRACE[76] • Scientific literature and data infrastructures/services such as OpenAIRE(plus), DOAJ[77], BASE and SherpaRoMEO[78]. • D4-Science[79], SCIDIP-ES[80], and EUDAT[81] (data services across disciplines to be operated as part of a generic, open and distributed infrastructure)
Achievements/ effects	• A pan-European 'e-infrastructure system' consisting of several intertwined domains, together providing a variety of functions and services for (in principle) anyone in research and higher education in Europe • Larger, European cooperative structures which facilitate 'openness'—access across disciplinary and national borders • Physical e-infrastructure and e-content are merging into one interconnected research e-infrastructure

Issues	• Technical, legal, organisational and policy 'interoperability', e.g.:
	  ○ Legal complications (data protection legislation); lack of a common legal framework and interpretation in Europe
	  ○ Current Authentication and Authorization Infrastructure (AAI) approaches present a barrier to access for non-traditional participants in the Research and Higher educational processes, such as academically trained professionals in SME's (as participants in innovation processes), and members of the general public (as participants in for example open/citizens research)
	  ○ At the technical level, Single-Sign-On (SSO) for users is not sufficiently widespread across the entire e-Infrastructure. It is acknowledged that "no technology can be universally adopted, but there should be mechanisms in place to allow for integration of different technologies." Some systems, for instance for grid and supercomputing, are complicated and face usability and scalability issues
	  ○ The potential and challenges offered by Cloud infrastructure

[72] http://ec.europa.eu/research/era/index_en.htm
[73] http://www.geant.net/pages/home.aspx
[74] http://www.egi.eu/
[75] http://www.deisa.eu/
[76] http://www.prace-ri.eu/
[77] http://www.doaj.org/
[78] http://www.sherpa.ac.uk/romeo/
[79] http://www.d4science.eu/
[80] http://www.scidip-es.eu/
[81] http://www.eudat.eu/

3.5 Open Source Software and Open Development

Open Source Software and Open Development(OSS/ODM)	
Definition	• Open Source Software is computer software that is available with source code: the source code and certain other rights normally reserved for copyright holders are provided under an open-source license that permits users to study, change, improve and at times also to distribute the software[82] • Open Development is a community-led development methodology used in open source software projects whereby the software is developed collaboratively and openly. Key attributes are: a deep level of user engagement, transparency, collaboration, agility, sustainability, supporting tools[83]
Roots	• Free Software movement • Development of the internet and in particular the world wide web
Aims/ motivations	• Source code viewable to all, no (or relaxed) copyright restrictions to allow for reuse, change and distribution • Maintain and develop the original product, whatever that might be • Collaborate to make something better than any individual could make on his own
Key developments / projects / policy drivers	• Free Software Foundation[84] • Open Source Initiative[85] and Open Source Definition • GNU Licence[86] • Linux • Apache Software Foundation[87] • Development of the world wide web using open standards and open development practices(for example the request for comment process maintained be IETF[88])

[82] Wikipedia, http://en.wikipedia.org/wiki/Open_source_software
[83] OSS Watch, http://www.oss-watch.ac.uk/resources/odm
[84] http://www.fsf.org/
[85] http://opensource.org/
[86] http://www.gnu.org/licenses/gpl.html
[87] http://www.apache.org/
[88] http://www.ietf.org/rfc.html

Achievements/ effects	• OSS/ODM now offers an alternative business model that has proven to be successful; for example:
	  ○ ODM's distributed peer-review method and transparent development process/decisions ensure high standards and continuous development, which contributes to the sustainability of products created in temporary projects
	  ○ OSS products with high market penetration and developer loyalty, which contributes to continuity/sustainability (such as Apache HTTP server, GNU/Linux operating software, Mozilla Firefox browser, Android)
	  ○ ODM offers the potential for a more flexible technology and quicker innovation, because of the typically large number of independent programmers collaborating on the software
	• OSS is now important for reuse of data sets

Issues	• Significant people and management skills are required to deal with the open, distributed and informal nature of the community • Participants may branch off with the product and create a competitor • Continuity of the Open Development community depends to a large extent on the founders being able to constitute a clear governance model as the community itself grows and develops

3.6 Open Education

Open Education	
Definition	• Open education is the collective term that refers to educational organizations/programmes that seek to eliminate barriers to entry. Such institutions, for example, would not have academic admission requirements. Programs are commonly (but not necessarily) distance learning programs like e-learning, MOOC and OpenCourseWare. Whereas many e-learning programs are free to follow, the costs of acquiring a certification may be a barrier, many open education institutes offer free certification schemes.[89] • The development of a vast pool of educational resources on the Internet, open and free for all to use, as well as open technologies that facilitate collaborative, flexible learning and the open sharing of teaching practices that empower educators to benefit from the best ideas of their colleagues. It may also grow to include new approaches to assessment, accreditation and collaborative learning[90].
Roots	• Education of citizens by researchers at public universities through informal education programs • Open universities, open distance learning[91] • OpenCourseWare developments • Open access to knowledge developments/movement
Aims/ motivations	• Learners have control over what, why, how and where they learn • Using technology to make education more accessible and effective • Facilitate collaborative, flexible learning and the open sharing of teaching and learning practices • Produce a threefold educational experience that combines learning/research, communication & collaboration, and the ability to share findings with a specific population

[89] Wikipedia, http://en.wikipedia.org/wiki/Open_education
[90] Cape Town Open Education Declaration, http://www.capetowndeclaration.org/
[91] Open Education does differ from Online (Distance) Education (ODE), which is formal education, including certification, in an online environment.

Key developments / projects / policy drivers	• Open Education Conferences and Open Education Week • Cape Town Open Education Declaration • Non-certificate-granting programs, including edX, MITx, Harvardx[92], Coursera[93], Udacity[94], Open Learning Initiative[95] • OER university[96]
Achievements/ effects	• Inclusion of developing countries; South Africa, Brazil, India and Nigeria are some of the most engaged in the use of Open Education • Contributing to innovation of education and learning as a practice, and as a field of research, with respect to technologies, processes as well as content. • Tying research and education & learning closer together • Nourishing the kind of participatory culture of learning, creating, sharing and cooperation that rapidly changing knowledge societies need
Issues	• Funding models • Certification; issue revolves around the question whether students get the proper recognition of the qualifications gained through open learning • Quality assurance of teachers and OER • Availability of adequate/appropriate technology in developing countries

[92] https://www.edx.org/university_profile/HarvardX
[93] https://www.coursera.org/
[94] http://www.udacity.com/
[95] http://oli.cmu.edu/
[96] http://wikieducator.org/OER_university

3.7 Open Peer Review

Open Peer Review	
Definition	• Scientific literature concept and process, central to which is the various transparency and disclosure of the identities of those reviewing scientific publications[97]
Roots	• Dissatisfaction with the existing anonymous peer review system because of a number of defects; research showed that the open variety performed equally well in addition to having a number of advantages • The availability of internet and online access to research publications • Open Development methodology
Aims/ motivations	• Provide an alternative that gives more transparency and accountability, faster claim of precedence and less possibility for abuse and bias • Prove peer review's added value to the scientific process and thus its place in an electronic world where authors can potentially go straight to readers • In ODM: better quality, faster development
Key developments / projects / policy drivers	• Spread in different fields, from Physics and Computer Science to the Arts & Humanities • Experiments in journals, e.g. British Medical Journal and Shakespeare Quarterly, as well as in book chapters in the A&H • The development methodology in OSS and Open Standards

[97] Wikipedia, http://en.wikipedia.org/wiki/Open_peer_review

Achievements/ effects	• The experiments mentioned above apply the available modern ICT tools and new perspectives on the way research is 'produced' and communicated in order to address challenges in the traditional system. These include: ○ reliance on a small, elite and closed group of experts can limit the range of feedback and institutionalize specific schools of thinking in subject areas ○ monopoly effect of the present system on 'access to publication in established journals' and thus on academic career and tenure They also allow the investigation of the perceived benefits and pitfalls/challenges such as ○ better prevention against fraud and other abuse of peer reviewers' privilege ○ accountability (for reviewers' comments) versus more honesty in anonymous system ○ credit for (often unpaid) expert work ○ faster claim of precedence ○ better quality, faster, when opened up to "the swift collective judgment of a much broader interested audience[98]" versus quality only possible from expert peers and overload in comments
Issues	• As long as Open Review is not common practice, researchers remain uncertain about the effect it may have on their career development and are therefore wary to participate in experiments • Portability of experiments' results to other disciplines • Applicability of Open Review in funding processes (particularly of publicly funded research)

[98] Patricia Cohen: Scholars Test Web Alternative to Peer Review, http://netage.org/2010/08/24/scholars-test-web-alternative/

3.8 Open Research

Open Research	
Definition	• The central theme of open research is to make clear accounts of the methodology freely available via the internet, along with any data or results extracted or derived from them. This permits a massively distributed collaboration, and one in which anyone may participate at any level of the project[99].
Roots	• Open research as a methodology follows from, and combines in it, several other 'opens': Open Access to research literature, Open Data, Open Source Software, Open Development and Open Review

Aims/ motivations	• Make research (in particular that which is publicly funded) more transparent, efficient and accessible to a larger public by, for example:
	  ○ Transparency in experimental methodology, observation, and collection of data
	  ○ Public availability and reusability of research data
	  ○ Public accessibility and transparency of research communication
	  ○ Using web-based tools to facilitate collaboration on research
	  ○ Engaging/active participation of a wider audience, also outside academia

Key developments / projects / policy drivers	• Early adopters in the Life and Computer Sciences, in Mathematics and Astronomy • Important initiatives in the field of medicine/medical treatments, such as Neglected Diseases[100] and the Open Source Drug Discovery Network[101] • 'Citizen Science' projects like Galaxy Zoo[102] and Foldit[103], in which the general public (whether academically trained or not) contribute to research projects • Initiatives promoting 'publicly funded research should be publicly available', as mentioned for other 'opens' • Funders increasingly require collaboration across institutions and countries

[99] Wikipedia, http://en.wikipedia.org/wiki/Open_research
[100] http://www.thesynapticleap.org/
[101] http://www.osdd.net/
[102] http://www.galaxyzoo.org/
[103] http://fold.it/portal/

Achievements/ effects	• Current projects show the potential for research by bringing more minds together on research questions than has ever been possible before, with remarkable results (examples include the Polymath project, the 2011 EHEC-crisis in Germany).
	• These experiments provide testing ground for ideas such as:
	  ○ For the increasingly complex questions facing research, collaboration with others is essential
	  ○ Collaboration on complex research questions will bring individual institutions greater benefits than doing the research on their own
	  ○ Useful knowledge is widely available, both inside and also outside academic institutions

Issues	• Risks and inefficiencies of information overload • Risk of abuse of information for harmful purposes. • Commercial/financial interests versus societal impacts and the public good

3.9 Open Innovation

Open Innovation	
Definition	• Open innovation is a paradigm that assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as the firms look to advance their technology[104] • Innovating with partners by sharing risk and sharing reward[105]
Roots	• New methods for new business development • The term became a standard with the publication of Henry Chesbrough's book "Open Innovation: The new imperative for creating and profiting from technology"
Aims/ motivations	• "Useful knowledge is… …widely distributed and of generally high quality".[106] • To tap into those knowledge resources not available within the organisation (the so called 'outside-in' approach) • To exploit more inventions more successfully by selling or licensing those the organisation cannot and will not exploit itself (the so called 'inside-out' approach) • To be able to deal with more complex questions • To open up the discovery process itself
Key developments/ projects/ policy drivers	• Crowd sourcing • Google Summer of Code project[107] • MITx, edX • Citizen science projects (see also Open Research above)

[104] Chesbrough, Henry, (2003). Open Innovation: The New Imperative for Creating and Profiting from Technology. Boston: Harvard Business School Press
[105] http://www.100open.com/2011/03/open-innovation-defined
[106] Chesbrough
[107] http://code.google.com/soc/

Achievements / effects	• Adoption of the practice has an innovating influence on the way research and higher education is conducted • Stronger innovative effects / more or unexpected spin-off because of the 'out-of-the-box' character of open innovation • Stronger collaboration and ties between organisations/projects and (outside) participants because of the active involvement in the development • Wider societal benefits
Issues	• Exploitation rights, intellectual property (IP) • Calculation of return on investment[108] • Competition • Control

[108] Difficulties include for example the open-ended nature of open innovation, and the effect of societal impacts. In terms of 'sustainable development' the latter is a factor that should be calculated in, as is done by some firms that operate on the basis of sustainability, rather than profit maximization. It is also a relevant factor in publicly funded organisations.

4 Challenges

The survey of the State of Play in Section 3 reveals a number of common themes that can provide challenges in the development of 'Open' as a modus operandi in research and higher education. They are discussed in greater detail in this section to allow the identification of possible ways forward in Section 5.

4.1 How open is 'Open'?

So far in this document, 'Open' has been used as a single, generic term. Yet, closer inspection of the definition already shows there are different levels of 'Open'. The same can be said for a number of the issues encountered in the implementation of the various 'opens'. One of the challenges to be dealt with is the fact that

> *"ideas of openness remain open to interpretation"*[109],

and we probably must be open with respect to these interpretations. However, even within these constraints, there is much that can be said about different levels of openness, and the way they impact discussions about the implementation of 'Open' as a modus operandi.

4.1.1 Practical aspects

4.1.1.1 Free versus Open

In a lot of the discussions about open content and open software (in particular), confusion often arises because the words 'free' and 'open' are used synonymously, and 'free' is interpreted as 'free of charge'. In the Open Source Software community, where this discussion dates back to the 1990s, 'free' has the wider connotation of 'freedom': source code is also open, and one has the right to modify and redistribute it.[110] There is often also a moral connotation to the discussion, with 'freedom' being an end in itself. In the late 1990s, the term 'open source' was created as an alternative to 'free', to escape from the confusion and for the pragmatic reason that open source was a promising business strategy for the corporate world[110].

In the Open Source community, 'open' now *can* mean 'free of charge', but more important is its 'freedom for (re)use'[111]. When *'free of charge' is seen as only one of the possible flavours of 'open', a wider range of options is opened up,* depending for instance on what is available openly, at what level of openness, and at what service level. Open Source and Open Standards have proven that it is possible to have 'open' products, projects and specifications on the basis of which further specific applications are built that are not open—these do not harm or limit the availability and use of the original 'open' version. On the contrary, they can even contribute towards the sustainability of the 'open' version. Examples are online collaboration platforms like Alfresco[112] and Liferay[113], that offer open source platforms in free-of-charge and subscription-based versions. Drupal[114] offers a free-of-charge open source/open development collaboration platform and points to a 'market-place' of providers for paid services, hosting and training.

[109] "Open Source and Open Innovation", OSS Watch, 2010, http://www.oss-watch.ac.uk/resources/openinnov.xml
[110] http://producingoss.com/en/introduction.html#free-vs-open-source, which also quotes the battle between the Netscape and Microsoft IE browsers as an example that zero-cost software is not 'free as in freedom to use'
[111] In the sense of right to modify and re-distribute.
[112] http://www.alfresco.com/
[113] http://www.liferay.com/
[114] http://drupal.org/

4.1.1.2 Levels of Openness

Related to the Free/Open distinction is the question what level of openness is provided. Are resources 'open' if they can be accessed free of charge but only after log-in, or upon receipt of a certificate (as is the case, for example, with many grids)? Is content 'open' if it can be accessed freely, but not easily reused due to technical or legal barriers (barriers such as incompatible formats, copyright claims, and so on)? Is research or innovation 'open' if participants can only join on invitation and/or have no influence on the direction of the project?

These are only a few of the questions which make clear how varied the open landscape can be, and how important it is in discussions to *clarify the possible distinctions and perspectives* between participants; *not to find 'the one and only true definition of Open' but to find the range within which cooperation and progress is feasible*. Examples of how this ranging could be presented can be seen in the 'five-star-system'[115] for open content, embodied in the 5 levels of openness for data (Tim Berners-Lee[116]) and the 'five stars evaluation scheme for Online Journal Articles' as presented by David Shotton (Oxford University[117]). Another example is the distinction between 'open source' and 'community source' in the OSS context. "…in 'community source' projects… …a consortium of institutions or commercial companies sign an agreement whereby they decide to contribute a certain amount of financial and/or human resources, and get, in exchange, exclusivity in influencing the development of the project during an initial closed stage. After this closed period, the code can be made available to everyone, and the development moves towards a more traditional open source model."[118]

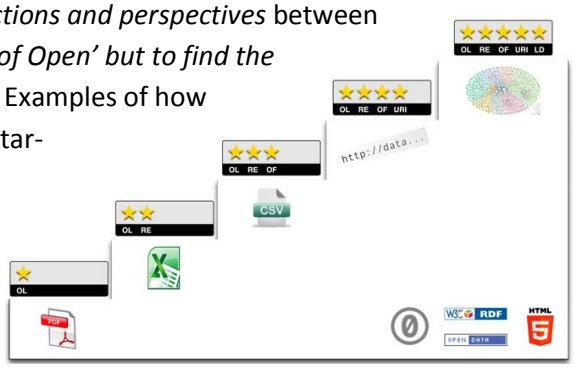

Figure 3 5 star system for open content

Figure 4 The 5 stars of linked data

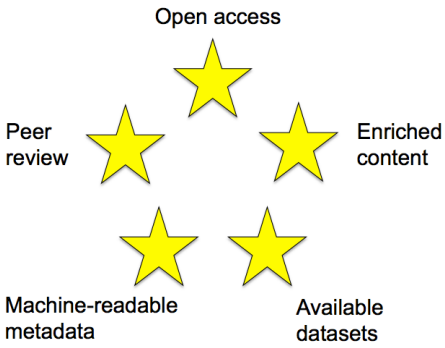

Figure 5 The 5 stars of online journal access

4.1.1.3 Intellectual Property

One of the most often raised issues in the discussions about openness is that relating to Intellectual Property (IP), its protection and its exploitation/commercialization. All parties involved, including authors/creators and their institutions, point to IP as a reason not to provide open research outputs. This may be because IP-rights are habitually transferred exclusively to other parties so that their own control over their outputs is lost. Protection of exploitation interests is another reason.

[115] http://5stardata.info
[116] http://www.w3.org/DesignIssues/LinkedData.html
[117] Shotton D (2012). "The Five Stars of Online Journal Articles — a Framework for Article Evaluation", *D-Lib Magazine* 18 (1/2) (January/February 2012 issue). doi:10.1045/january2012-shotton. (Preprint, giving original text prior to revision and expansion following post-publication peer review, available in Nature Proceedings at http://precedings.nature.com/documents/6542/version/1)
[118] http://www.oss-watch.ac.uk/resources/communityvsopen.xml

It is clear that IP is not an easy issue to deal with, and the legitimacy of the wish for exploitation of rights is seldom debated.

Yet here too, lessons from Open Source Software, Development and Innovation indicate there may be scope for more openness than now presumed. The TexGen example[119] shows the open source benefits were both financial (additional funding) and intangible (increased knowledge about potential applications for example); and the commercial applications developed from the original open source product did not impact the availability and usefulness of the latter. The same phenomenon has been true for other Open Source Software.

Open Innovation literature states that many patents lie unused, because the time, manpower and/or expertise is lacking to exploit them all[120]. Furthermore, Rufus Pollock (co-founder Open Knowledge Foundation) points out that "the best thing to do with your data will be thought of by someone else"[121]— which is characteristic for innovation as a whole. A recent 'Future of Research' report[122] suggests that for UK researchers to stay competitive globally

> *"they need to get used to sharing data and move freely between sectors and countries[123]."*

Holding on to research outputs can mean serious loss of capital (intellectual and financial), to both the originator and society; and it hampers innovation.

4.1.1.4 Competition versus cooperation

Related to the issue of IP but ranging wider is the question of competition versus cooperation. An example is the competition between higher education institutions for more students. In the edX experiment[124] competitive universities join forces in a new offering which can be expected to benefit each of them by drawing in larger number of students and by learning about online education. They could probably each do this independently but no doubt at much higher cost, with no better guarantee of success. Joining forces is likely to increase the chance of achieving the critical mass necessary for this sort of experiment to succeed. edX is an interesting example of a relatively new way of working for organizations: "coopetition"[125].

Similarly, research groups and institutions can be faced with the choice between competition, cooperation or coopetition. Which choice is made is in any case influenced by the assumptions about expected benefits to the group, the institution and/or the project, and by funding requirements (is a consortium required for instance). An important factor in these assessments is whether societal benefits are included or not. Although

> *"It is difficult to predict the potential pay outs of technology or to assess the costs of withholding it, there is general agreement that the benefit to any single institution of*

[119] A software product originally made in the context of a research project and subsequently provided Open Source in an Open Development community; it became successful in much wider circles, procuring continued funding from new sources and finding new, unexpected applications. TexGen: a case study, 2009, http://www.oss-watch.ac.uk/resources/cs-texgen.xml

[120] "Open Innovation: Researching a New Paradigm", Chesbrough et al., http://www.openinnovation.net/Book/NewParadigm/Chapters/01.pdf p. 5, last update March 2006

[121] "Open Innovation in software", http://www.oss-watch.ac.uk/resources/openinnovsoftware.xml, OSS Watch, 2011

[122] Future of Research report http://www.universitiesuk.ac.uk/Publications/Documents/TheFutureOfResearch.pdf

[123] Quoted from "Open Innovation in software", http://www.oss-watch.ac.uk/resources/openinnovsoftware.xml, OSS Watch, 2011

[124] https://www.edx.org

[125] "Coopetition occurs when companies interact with partial congruence of interests. They cooperate with each other to reach a higher value creation if compared to the value created without interaction, and struggle to achieve competitive advantage", http://en.wikipedia.org/wiki/Coopetition

> *holding technology is not as great as the cost of withholding it from all other research institutions[126].”*

This sentiment—expressed here in relation to technology—applies equally to information and other research outputs. From the perspective of a single institution this may not seem attractive; yet, there is growing recognition that with the increasingly complex questions facing research, collaboration with others is essential and will bring the individual institution greater benefits than ‘going it alone’.

Also of influence on the choice is the available research and management expertise. *Cooperation and coopetition require complex management skills*, and if these are lacking, the best choice may well be to stick to the competition mode. The best choice is dependent on the specific situations and circumstances; none of the working modes is inherently the best.

4.1.1.5 Security

Another frequently mentioned argument against ‘Open’ is security risk: personal safety risks, health hazards, breach of privacy, financial risks, etc. Risks are real and not unique to the online world, though ‘online’ does add a new assortment. No doubt the ‘new and unknown’ of open increases sensitivity to its risks. In itself this is positive, since it triggers important questions for which new answers need to be found. It also creates an alertness to the risks that is necessary to combat them—as is seen in the current discussions about cybercrime, spyware, and the like. The ‘Microsoft Internet Explorer leak’[127] case poses interesting questions not only about the safety of the browser, but also about the decisions where to have equipment produced and assembled. At the same time it is important to remain aware that the interests of organisations specializing in cybercrime and spyware detection may influence the discussions, as happened in the debate on Open Standards and Open Source Software in the 1990s.

The recent H5N1 influenza debate[128] is another significant example of a security dilemma. Publications from two different research groups aimed at easier detection and prevention of the spread of the H5N1 avian influenza virus also opened up the possibility of biological warfare. The debate on this ‘dual use’ risk in the US National Science Advisory Board for Biosecurity (NSABB) eventually resulted in a reversal of its original advice to withhold key details from publication, to a large extent because:

> *“New evidence has emerged that underscores the fact that understanding specific mutations may improve international surveillance and public health and safety. Global cooperation, critical for pandemic influenza preparedness efforts, is predicated upon the free sharing of information and was a fundamental principle in evaluating these manuscripts.”*

An informed debate tipped the scales in favour of the benefits of publication, rather than the risks. This may look counter-intuitive, and that is exactly why the informed debate is so important in issues like this. The NSABB argument echoes the experience in the OSS environment that ‘open source’ is neither more nor less secure than ‘closed source’; the apparent lack of organized control is mitigated by the open source principle of ‘many eyes make every problem shallow’[129]. Our world is becoming increasingly open, and we cannot afford to be naive about its effects; but *neither can we afford to keep relying on rules, notions and intuitions that have been formed in a paper-based information society*.

[126] http://en.wikipedia.org/wiki/Open_science
[127] http://technet.microsoft.com/en-us/security/bulletin/ms12-063
[128] http://news.sciencemag.org/scienceinsider/2012/03/breaking-news-nsabb-reverses-pos.html
[129] http://www.oss-watch.ac.uk/resources/securityintro.xml

4.1.2 Cultural aspects

4.1.2.1 Mindset

'Habit is one's second nature' and it often unconsciously influences our behaviour in adoption of new technologies and other innovations[130]. Habits are formed within a frame of reference created by education and social and organisational reward systems. They can lead to innovations, but also hinder them.[131]

'Open' is an innovation at the global cultural level—notwithstanding the fact that it impacts different cultures in different ways—with deep impacts at all 'lower' levels: national, political, organisational, and personal. The impact of open developments in research and higher education is mixed. On the one hand, there is happy adoption of tools supporting relatively open communication and collaboration—Google Docs, Mendeley, Blogs, Twitter, etc. On the other hand, the more organised introduction of 'Open' as in Open Access, Open Data, Open Review and Open Research triggers hesitation (at least), if not resistance. This may, in part, have to do with lack of freedom of choice; but the responses in surveys among researchers on the open developments indicate there is more at stake. The open web-tools mentioned above facilitate or build upon the present process of research communication and do not by their nature bring into question ingrained research publication habits. These habits underpin the current frame of reference for the entire research and higher education system: for research, funding, career assessment, review and publication, education. The introduction of 'Open' most definitely does affect these habits.

Where critical examination of new theories and experiments is *the* method used in one's everyday work—a scenario that is particularly prevalent in the world of research—valued habits are likely to be retained until an alternative is proven to be efficacious. As Chesbrough says in the context of innovation research:

> *"scholars withhold their support of these novelties, unless and until they can demonstrate a more enduring contribution to the advancement of knowledge"*[132].

In addition to the current research publication habits, this scholarly mind set may be an important factor at play in the lack of adoption of 'Open'. Firstly, 'Open' is not mature enough yet to demonstrate such enduring contribution sufficiently, but perhaps more importantly, *it is hard to simultaneously apply this scholarly perspective of 'demonstrate first' and be participant in the experiment*.

4.1.2.2 Distinctions in disciplines and generations

Perspectives on and adoption of 'Open' differ among disciplines and generations. The 'early adoption' disciplines and generations provide the experiments needed to demonstrate the 'enduring contribution to the advancement of knowledge' (and/or to society at large). And in a growing number of examples this demonstration is successful, or on-going. OSS is probably the easiest example to use of 'proof' because it is more established than Open Access or Open Data. Yet early examples of Open Access publishing and Open Data also prove to have benefits hitherto unknown or unimagined (see section 2.3)

The main issue with disciplinary distinctions is that success in one discipline is not perceived as necessarily applicable in another and so is not accepted as 'proof'. This isn't entirely unjustified. Every discipline must find its own applications in its own specific context. The Arts and Humanities clearly show that given

[130] "Intention does not always matter: The contingent role of habit on it usage behaviour", M. Limayem et al., *Global Co-Operation in the New Millennium, The 9th European Conference on Information Systems*, Slovenia, June 2001
[131] Innovation and the role of habits: a conceptual analysis, D. Cavagnoli, Discussion Paper, August 2008
[132] "Open Innovation: Researching a New Paradigm", Chesbrough et al., http://www.openinnovation.net/Book/NewParadigm/Chapters/01.pdf, p. 4, last update March 2006

time and interest, scholars will find new applications that enrich their research and publications. Annotation tools, text mining techniques and the OAPEN[133] publication platform are clear cases in point.

Distinctions in generations are not easy to interpret. Younger researchers and students—the so called digital natives—have grown up in, and are believed to be used to, online environments and tools; they would seem to be the likely candidates to adopt the changes. However, when they move into the world of academic education and research, many show fairly conservative behaviour; *the influence of the habits of their senior mentors and the current reward systems are strong and engender a 'must fit in' frame of mind*, particularly when it comes to safeguarding career prospects[134].

Established researchers are steeped in, and owe their position to, the current traditions and therefore appear to be less likely to change their habits. Yet, they could be powerful innovators precisely because of their established position, in their roles as tutors for young researchers, as researchers and as participants in funding decisions. Some actually do take that innovator role, as do some young researchers, but it is not enough yet to create the necessary critical mass.

4.2 Evaluation and reward or perverse triggers?

> *"**Elsevier boycott: Academic spring in Germany slightly rainy***
> *A ZBW survey shows that German economists sympathize with the Elsevier boycott and Open Access, but still support the established reputation and ranking systems*
> *…*
> *Young researchers indicated that they are reluctant to join the boycott for fear of the negative impact on their further career. One of the participants stated: 'Boycott is only an option for tenured professors. A young scholar would criminally endanger his personal future if he did not publish in the magazines of Elsevier, some of which enjoy a very high reputation.'"*
>
> *Leibniz Information Centre for Economics, Kiel/Hamburg, 16 May 2012*

> *"The most daunting obstacle to opening up the process is that peer-review publishing is the path to a job and tenure, and no would-be professor wants to be the academic canary in the coal mine.*
> *The first question that Alan Galey, a junior faculty member at the University of Toronto, asked when deciding to participate in* The Shakespeare Quarterly's *experiment was whether his essay would ultimately count toward tenure. 'I went straight to the dean with it," Mr. Galey said. (It would.)' "*
>
> *Scholars Test Web Alternative to Peer Review, NY Times, 24 August 2010*

> *"The European Commission will help if it gives value to the publications with open access policies and if it convinces governments to create valuable indicators related to open access policies'. Moreover, as highlighted by another, 'open access policies will have an impact if its use is somehow prized and merited according to standards for promotion."*

[133] http://www.oapen.org/home
[134] JISC reports on generation X and Y; http://www.jisc.ac.uk/publications/reports/2012/researchers-of-tomorrow.aspx and http://www.jisc.ac.uk/media/documents/programmes/reppres/ggworkpackageii.pdf

These quotes from researchers in the context of Open Access to research literature and Open Peer Review are equally applicable to other 'Opens', as was shown in section 3. If one's career is perceived to be dependent primarily on the number of research articles published in a very particular set of journals, all requests for participation in other activities will be considered in that light and therefore be judged as at least ineffective and at worst damaging. These other activities could include for example provision of Open Data, cooperation in the creation of research-informed OER, cooperating in the development of fit-for-purpose disciplinary data infrastructures and in the innovation of education.

One could argue that research and higher education should be about the advancement of knowledge, not about career points, and no doubt this view is endorsed widely. Yet if an evaluation and reward system relies heavily on a single quantitative measure that has proven to be flawed even in what it measures[135], it is arguably lacking in effectiveness. It is a legitimate question whether it is not more a perverse trigger than a reward system.

> *"Innovation reflects a process of learning. Learning is affected by the system of rewards within each organisation; therefore, workplace practices, management policies and government policies affect the system of rewards, and so, learning; thereby contributing to, or harnessing the, creation and diffusion of innovation…*
> *…Habits of behaviour and of thoughts affect the priority of what needs to be learned. Habits help to adapt. They are maintained and reinforced via a system of rewards. However, rewards influence the process of learning, as they perpetuate values and beliefs of what to learn; and so, habits can also hinder adaptation and innovation."*[136]

4.3 The Longer Term

4.3.1 Preservation

One of the issues mentioned in section 3 is that there are many questions and uncertainties with respect to long(er) term preservation of open access research materials. Particular questions of note include:

- what to preserve—particularly in relation to the different impacts of preservation in different disciplines
- how that can be done over decades—problems include ever-changing hardware and software configurations
- who is responsible—what level of aggregation do we look at (individual, institution, discipline, and so on) and how do we take into consideration discipline-related differences in needs
- who pays the substantial bills involved (not just now, but also in the future)

The analysis of the EC's questionnaire on National Open Access and Preservation Policies in Europe states that:

> *"While many of the responding countries have put in place notable initiatives or strategies regarding the digital preservation of cultural heritage in general, specific*

[135] Experience has shown that there can be a marked difference between a journal's reputation amongst peers and its impact factor, not just because of time delays in the IF, but also structurally – e.g. in the case of interdisciplinary journals. The manipulation of the IF e.g. by selective citing and regular publishing of Review articles/issues is also a known phenomenon.

[136] *Innovation and the role of habits: a conceptual analysis*, D. Cavagnoli, Working Papers 2008.03, School of Economics, La Trobe University, http://ideas.repec.org/p/trb/wpaper/2008.03.html.

> *attention to the preservation of scientific information needs to be further developed within most existing national policies and legislative frameworks. Moreover, researchers do not seem to always be aware of preservation of scientific information, articles and data as a key issue..."*[137]

Considering the complexity and scope of the issues involved, although a nationally-based policy and legislative framework is necessary, long-term preservation seems to be an area of research and development that would benefit from a collaborative approach across nations to prevent costly duplication of efforts. It also seems to be an area that needs to look for 'approaches' rather than 'solutions'. A 'solution' implies a finality that is illusionary in an activity that is in principle never completed. 'Approaches' can evolve over time as circumstances change, and help find the 'best solutions for the time being'. The Blue Ribbon Task Force[138] has set an example of how this could be done, by treating *long-term preservation as a process, rather than a project* and setting time-horizons for further or renewed decision-making rather than trying to cater for eternity. Also, it is important to realize that open approaches themselves have a role to play in long-term preservation, as is evidenced by the LOCKSS system, which is an open-source, library-led digital preservation system predicated on the principle that "lots of copies keep stuff safe"[139].

Long-term preservation is a much bigger agenda than 'Open', but specific attention to it in the context of 'Open' is important also because it *may help build trust among researchers*. They need to know that this development is not a 'hype' that may disappear in four years' time when the project funding is over. The issue of the 'trusted party' is highly important here: the responsibility should not be left to commercial parties with a stake in research publishing. The fact that they do have a stake could make them eligible funding partners.

4.3.1.1 Governance

Equally important in terms of continuity and trust-building is the necessity to find governance models for open developments that start out as projects. Even if projects initially manage to continue in the form of a new project, at some point a model needs to be found to guarantee its continuity. A project that is considered to be successful and worthwhile should be able to find such a model; that it can be done is shown by numerous examples in the worlds of Open Source Software, Open Standards and also by way of example the European Research and Higher Education network. The European NREN's have found a modality by setting up a jointly owned not-for-profit, limited liability company, responsible for operational management and project management for new developments. Another example can be seen in the Apache Software Foundation (ASF) which is a foundation, functioning as a meritocracy, built around the principles of open development. Both examples show that governance models may use traditional legal entities and operate in traditional or completely new and experimental ways. Both have developed from existing cooperation as projects that felt it was necessary and worthwhile to collaborate in a more permanent form. Funding sources vary, but may (and do) include commercial parties. *The ASF is a good example of how the interests of the open source community and those of commercial partners can be combined without detrimental effect to open source principles and development*.[140]

[137] National open access and preservation policies in Europe, Analysis of a questionnaire to the European Research Area Committee, EC-DG for Research and Innovation, p 38
[138] http://brtf.sdsc.edu/about.html
[139] http://www.lockss.org/about/principles
[140] OSS watch provides two governance templates for academic open source projects based on two well- known governance models in open source communities: 'meritocratic' and 'benevolent dictator' http://www.oss-watch.ac.uk/resources/meritocraticGovernanceModel.xml and http://www.oss-watch.ac.uk/resources/benevolentdictatorgovernancemodel.xml

4.3.1.2 Sustainability

"How can this project be made sustainable?" is an issue in practically all 'opens'. Several questions come to mind in this context. Firstly, do all projects need to be sustained? Secondly, why set up a new activity or service as a project in the first place rather than as a regular new activity or service? Thirdly, in addition to asking how a project can become sustainable *the question that could (and perhaps should) be asked is how the project can contribute to the sustainability of an effective, efficient and innovative research and higher education system*. This shift in perspective means that the narrow project focus—'how can we keep our project alive'—is widened to include the project's place in the entire research and higher education ecosystem. This wider and more outward looking approach may make it easier to identify, or open up to, possible new applications and funding sources; or integrate it into regular activities or services. In the latter case, asking the sustainability question also means asking how those new activities and services impact the existing organisation(s) and how that can free up resources. It also means that a wider group of stakeholders may get involved in dealing with the question, rather than just the original project group. As the experience from the Open Development area shows this is an important success factor.

Sustainability is not a popular topic in not-for-profit sectors, nor a natural one in sectors where activities are often organised in project-form. The Open Source Software and Standards Development shows it can be done and provides a variety of examples for approaches. One size does not fit all. Each project will have to find an approach most fitting to its purpose and needs. Annex C provides a survey of available sustainability approaches.

OSS Watch also has guidance on this topic, guidance which is especially useful to researchers who start planning for the sustainability of their projects at bid writing stage[141]. To keep it actively in the project focus from day one and treat finding the answers as a process throughout the project increases the chance that sustainability is achieved as long as it is required. This calls to mind the approach advocated by the Blue Ribbon Task Force for long-term preservation: *'sustainability as a process'*, rather than as a fixed solution. The recently issued Knowledge Exchange Report on Sustainability of Open Access Services reinforces this message[142].

Another lesson to be learned from the OSS/ODM field is that the involvement of for-profit organizations can contribute to sustainability. Maintaining the open principles and approaches as a requirement in the collaboration will 'keep things honest', as has been shown by Open Standards development and the ASF. This 'principled pragmatism' does, however, require firm negotiation skills.

It is important to realize as well that sustainability is only partly about money; it is equally about things such as tools, infrastructure, work processes, people, time, skills. Obviously, money is a facilitator, but no amount of money will suffice if the people with the necessary skills are lacking. And skilled people often appear capable of achieving a lot with relatively little money, given time. Money and time are always scarce; creativity and skill will compensate for much if recognition of achievements is clear and appropriate.

The often-used concept 'business model' may actually contribute to the confusion and resistance in the sustainability discussion. For though it is a wide ranging concept it may easily be misunderstood, or interpreted narrowly as a 'financial model'. Furthermore, its commercial origins may cause unnecessary resistance in not-for-profit environments where the emphasis is on delivering services for the public good

[141] http://www.oss-watch.ac.uk/resources/planningsustainability.xml
[142] Sustainability of Open Access Services, Report Phase 1 and 2: Scoping the challenge and consulting the stakeholders, http://www.knowledge-exchange.info/Default.aspx?ID=536

instead of on doing business. *There is more involved than just semantics when we would replace the phrase 'business model' by 'sustainability approach'*.

4.4 Diversity

In the implementation of infrastructure projects or concepts across a wide range of countries, regions, disciplines, sectors, organisations and people, diversity is always remarked upon as an issue. We also see this happening with the various 'opens', especially in the area of open data where disciplines regularly raise the issue of their subject-specific requirements, the implication being that they are "different" to those of other disciplines.

As diversity is the rule, not the exception, it could perhaps be treated as a given rather than an issue. One of the reasons that diversity becomes an issue is that the 'one size fits all' approach is widely considered to bring about efficiency (particularly in process such as, for example, the maintenance of systems). This notion is understandable, and up to a certain point is no doubt true. But the efficiency that may exist at the technical implementation level gets seriously compromised—if not wiped out—by the inefficiencies that appear at the wider organisational, disciplinary or other aggregation levels.

Dealing with diversity is also often considered to be difficult. From the 'one size fits all' perspective this is often the case. It is also true in the sense that dealing with diversity means trying to understand and assess the views, needs and wishes of people and groups from completely different backgrounds than one's own. This takes up valuable time.

Taking diversity as a given means adopting a wider perspective on 'efficiency'. One way of doing this is by asking 'whose efficiency?' It also means taking a wider perspective on implementation than 'efficiency'. Effectiveness, contribution to communication, innovation, advancement of knowledge, other societal impacts, may also come into play. This is true for geographical and disciplinary diversity, as well as for diversity in levels of development.

In open source communities diversity is actually encouraged as the one of the most important factors for community growth and for the improvement of the product being developed. For Mårten Mickos, former CEO of MySQL, diversity is a quintessence for innovation, which is:

> *"…most likely to happen when people encounter each other in social spaces where different views and solutions can be explored. Diversity brings long-term sustainability precisely because diverse people pursuing their own self-interests within a community are likely to end up working together for the benefit of everyone."*[143]

4.5 Skills

Section 3, and preceding paragraphs in this section, have shown that implementation of the various opens requires the development of new skills sets for almost everyone involved. It concerns both specialist and management skills. Examples include:

- Library staff involved in Open Access publishing and dissemination via institutional repositories have to learn to deal in a much more pro-active way with researchers and in different phases of the research cycle than they have been used to, to inform them about and assist them in OA publishing; including, but not limited to, dealing with copyright in the appropriate way.
- Journal editors, authors and reviewers may want to learn to work with open review.

[143] http://www.oss-watch.ac.uk/resources/openinnovsoftware.xml

- Management of 'Open Data' is a whole new field, with a call for new types of staff and expertise (for example selecting appropriate storage, metadata and classification, legal issues); in the broader definition of Research Data, the 'Data librarian' is a relatively new role.
- In Open Source projects, basic licensing knowledge is needed to avoid building or bundling together bits of code released under conflicting licenses. Also needed is specific knowledge and experience of building communities of users and developers around the software being produced in order to create sustainable ecosystems.
- Project leaders involved in Open Research projects need to acquire new management and leadership skills to maximize the benefits of involvement of external experts and/or the general public.
- Project managers need to learn how to build in sustainability approaches into their projects effectively. Strong negotiation skills to deal with commercial partners in collaborations on 'open' projects need to be developed. As do the skills required to deal effectively with diversity.

For successful further implementation of the various 'opens', individually and collectively, awareness of the necessity to develop such skills is important, as is the development or provision of training programmes. It is also important not to discourage people who have already acquired such skills in using them as tends to happen now to young people entering into research.

More important is the necessity that *due recognition be given to the new jobs and skills required to create a fully functioning open ecosystem*. The required career paths need to be made attractive if people are going to want to adopt them. Researchers with an interest in the application of ICT in their field of expertise and the possibilities for data management may well turn to these careers if they get the appropriate credit for it, and are not treated as 'second-rate researchers' (see also section 4.2 above).

5 Ways forward

The previous sections have shown that the 'Open' Agenda consists of a patchwork of developments in a series of 'individual opens' and have explored, to some extent, the state of play of those various 'opens'. In Section 2 it was stated that there is need for and possible benefit in undertaking a more coordinated approach to the development of those various opens in order to truly realize the full potential of the 'Open' Agenda; that is to say a more effective, efficient and innovative research and higher education system. This section looks at this claim more closely and explores how it could be realised.

5.1 Research and higher education as an Open Enterprise?

The claim that the research and higher education system will benefit from a greater level of openness could itself still be subject to criticism (and the various opens do have their critics). Yet, in section 2.3 we have shown the substantial benefits that are being realised. And, despite the criticism, the notion of free exchange of knowledge in research is a very basic one that is shared widely because everyone involved is aware that that is a necessary ingredient for the advancement of knowledge. 'Free' in this context has the same meaning as in our definition of 'Open' in section 2.2: free from barriers to access or use. *For optimal 'free exchange of knowledge', openness is a necessity*. This has always been the case, but the traditional ways of research communication—conferences, one to one communication, and publishing—by their nature limited the achievable level of openness. The modern ICT facilities enable levels of openness that have never been possible before; and consequently they can facilitate unprecedented step changes in the advancement of knowledge.

It is clear that openness has its issues and that it changes existing behaviours and power structures in research and higher education. The issue at question is whether that should be a reason *not* to go ahead with it or to slow down its development. The EC has made its choice with its recent *Communication and Recommendation on better access to scientific information*[144]. The UK, too, has a clear stance with the Royal Society report "Science as an Open Enterprise"[145]. It states that **"**openness is central to scientific enquiry" in that it supports better research, reviewed research and the development of new knowledge. New technologies open up new improved methods of science and whilst changed behaviour is witnessed this is piecemeal and there are no systematic policies and fit-for-purpose infrastructures in place; on the whole, the report states, many behaviours in research still duplicate that of the paper world and the paper journal and that this is holding back advances in science.

It would seem that Openness has reached the 'point of no return'. Even the question whether it should be either 'a possible' or 'the preferred' modus operandi may already have been answered in favour of the latter. To realize the beneficial effects for the entire research and higher education system, 'Open' needs to be adopted across the board as the preferred modus operandi, notwithstanding the fact that implementation may be gradual and phased, and situations may remain that require exceptions.

[144] http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1301
[145] published in June 2012, http://royalsociety.org/policy/projects/science-public-enterprise/report

5.2 Individual or coordinated development?

5.2.1 Interconnectedness of open infrastructure, content and process

The technical deployment of a complex e-infrastructure is one step towards fostering new ways of conducting research. The developments in this area are such that, by now, people are starting to view 'information' as a part of the research and higher education e-infrastructure: the *'physical' infrastructure is "invisible", the information is the infrastructure* and is "perceived as a matter of course"[146]. Thus, the hardware, software and content used and produced in the e-infrastructure are all seen as a set of necessary, interlinked, resources. This perceived interconnectedness is strengthened by the 'data deluge': the capture of masses of data that has been facilitated by a massively improved physical infrastructure; the handling of the mass of data requires ever more powerful infrastructure. The increased complexity and diversity across varying data collections necessitates increasingly complex e-infrastructures.

The interconnectedness shows why, in a world of increasingly open content, the enabling physical infrastructure needs to be open as well: if content is open, the means with which to access and process it—manually and/or through machine processing—needs to be open as well. The international nature of research—and increasingly of higher education—reinforces this since the research and educational processes do not stop at institutional or national borders. Although this has long been the case, it becomes a more pressing matter in the virtual world. Hence also the need for open participatory and collaborative processes: with increased complexity and specialisation, the colleague with the necessary knowledge may be located on the other side of the globe or may not even be identifiable beforehand.

An equally important step to the technical deployment is encouraging researchers to use this technical framework to its full potential. Addressing this issue, a number of reports mention the need to embed the technical research framework in a *'human infrastructure'*. In this context, human infrastructure refers to the social and organisational arrangements enabling technologies to be used effectively. The AVROSS report, for example, states that uptake of e-infrastructure is as often hindered by human and organisational issues as it is by technical ones. Focusing on the UK, it recommends continued technical innovation in e-research. At the same time, it suggests, the social framework that would allow research communities to better exploit these technical assets should be improved.[147]

5.2.2 Interconnectedness in work on open infrastructure, content and process

The perceived interconnectedness between the open categories described above suggests that the work to be done in and on them, and the parties involved in that work, should also be interconnected. e-InfraNet's work on 'Open' to date demonstrates this is only partially the case. As with many innovations, the development of individual opens has started 'bottom-up', each community having their own reason for embracing 'Open'.

Although it is clear that each of these opens has its own specific characteristics, audiences, and issues, there are important generic concerns. Solutions or approaches found and developed in one area might well be of benefit to other areas. Furthermore, at some point in time bottom-up development needs top-down policy and funding support to become stronger and to spread. This is evidenced by the developments in Open Access and Open Data, where the coordinated policy support of the EC has provided serious impetus for developments at the national level. Policy support makes it clear that the new development is taken seriously at the highest levels.

[146] http://pub.uni-bielefeld.de/publication/2445229
[147] http://www.oss-watch.ac.uk/resources/researchinfrastructure.xml

Leaving the individual opens to continue their separate development is an option in theory, and may seem superficially less complex. Developments that have been around and incrementally developing for 10 to 20 years have, by definition, proven to be viable and no doubt further progress could be made in a similar way. However, the present fragmentation also poses a number of important risks and disadvantages:

- Fragmentation in development easily leads to fragmentation in understanding of the importance of the development, especially in relation to the entire system of research and higher education. Each community has its own rationale for the development; each open will impact different stakeholders in different ways (see Annex A). Within a fragmented perspective, it is easy to make decisions on the development of one open that may adversely affect the development of others and thus of the entire system.
- Fragmentation, and the resulting lack of clarity, can slow down or even undermine efforts in advocacy, debate, policy development and implementation; and it can hide from view that each open contributes to, and forms part of, an overall important cultural change.
- Fragmentation leads to sub-optimization in the sense that each open is judged separately on issues such as sustainability.
- Fragmentation between groups hinders learning from each other's experiences and development methods.

In short, fragmentation leads to sub-optimal development decisions and progress, and hampers the innovation of the research and higher education system as a whole by trying to optimize the development of its constituent parts.

No doubt, for practical reasons, the actual work at operational level will primarily remain divided among the present groups, with probably more conscious attempts at cross-fertilization. Yet, they will *need to be able to look at the European, national and institutional strategic/policy level for guidance and support which stem from a more unified and forward-looking picture of 'Open'*. This will reduce sub-optimization and foster progress and innovation.

5.3 Implications for policy and implementation

5.3.1 The choice of 'Open' as a modus operandi

Taking 'Open' as the preferred default modus operandi across the research and higher education system, and adopting a coordinated approach to its further development, implies that at the national, funders and institutional decision making and policy level a *conscious choice* is made to follow this route and support its development. It requires

- all stakeholders to be aware that this has implications for all constituent parts of the system (including the stakeholders themselves) and that these stakeholders will have to work together, or in any case refrain from working against each other;
- that the issues encountered are seen and treated as implementation issues for which approaches or solutions need to be found; not as arguments for or against 'Open';
- removing existing barriers to implementation, so that natural bottom-up developments and cooperation across borders (disciplinary, institutional, departmental and so on) can be fostered.

5.3.2 An open approach to the implementation of 'Open'

Making the choice for 'Open' requires an open approach to its implementation, and this may be a bigger paradigm shift than the choice for 'Open' itself. The 'Open' choice can be rationalized, but an open approach to implementation impacts behaviour.

The choice for 'Open' does not necessitate a one size fits all approach, nor a centralized approach. In fact, *'Open' inherently requires a distributed network approach*, *not only at the technical level, but also at the organisational and financial level*. This allows for diversity in development speed which can ease implementation pressures on countries, disciplines and institutions. It also allows for diversity and phasing in approaches and in levels of openness (as exemplified by the 5 star-system mentioned in section 4.1.1.2) which can alleviate privacy, security and competition issues. It is a practical way for people to get used to the new approach, and to gradually upgrade or replace necessary systems and software.

A key concept in openness is interoperability (as opposed to homogeneity); interoperability in hardware, software, formats and between people. *The focus is not on how to change them, but on how they can work together*. This requires use of open standards[148], for hardware, software and formats, as well as for cooperation between people (in the use of open licences, transparent governance rules, and so on).

There is no end point or fixed timeline for the development of 'Open'. It would therefore benefit from being approached as *a continuing process with transparent governance principles among dedicated stakeholders*, rather than as a final project (as advocated by the Blue Ribbon Taskforce for long-term preservation and sustainability and the principles of the 'Open Development Method'). Adopting this approach will ease implementation issues and contribute to its sustainable development, as demonstrated by Open Development of Open Software and Standards. It will also help to build confidence among the various stakeholders (and in particular researchers).

The sustainability question needs to be considered at least at research and higher education system level (or perhaps even higher in the hierarchy) rather than just at the level of each different 'Open' and each different open project, to reduce the sub-optimization effect. Just how each different open project contributes to a sustainable open research and higher education system requires different stakeholders to cooperate on the question; this methodology also fosters cross-fertilization and reduces the unnecessary duplication of efforts. It simultaneously builds the dedicated communities necessary for sustainable development[149]. It shifts the focus from the individual interests at project level to the larger interests at systems level, which opens up new development and funding possibilities.

Furthermore, if the sustainability question is included in such discussions as part of the open development of 'Open' its integration is ensured from the outset. This increases the likelihood of actually achieving a sustainable 'Open', especially if specialists from, for example, Human Resources and Finance are involved in the discussions. It is also important that the scope of the sustainability question is widened from purely economic costs/benefits to include aspects such as people, skills, timeframes and societal impacts.

An important lesson from the Open Development Method, also espoused by Open Access and Research advocates like Cameron Neylon[150], is *that the development of 'Open' relies on the inclusion of diverse parties*, including people and organisations that are traditionally not part of the research and higher education system; in other words the general public and the commercial sector. Their inclusion or exclusion has in the past been framed as part of an ideological debate. Practically speaking, however,

[148] Or in any case standards of which the specifications are open.
[149] See the OSS Watch literature on (the relation between) sustainability and community building, http://www.oss-watch.ac.uk/resources/buildingcommunities.xml
[150] http://cameronneylon.net/about/

they are necessary participants in the process of open innovation. The Open Development model shows that, as long as the principles of transparent governance, community building and sustainable development are preserved, everyone should be able to use, adapt or re-distribute the product, including those who wish to use it for commercial purposes. A similar principle in the Open Content area can be found in the so called Data/Services model (see Annex D): the basic data/information is published under an open licence; on top of that commercial services are possible as long as these do not pose any restrictions on the availability and use of the basic open content originally provided. This Data/Services model underpins the respective roles of repositories, OA journals, peer review and the like in the research communication cycle and reduces the importance of the 'Green' or 'Gold' debate. It can also provide guidance in the sustainability discussions: at what level of data and/or service provision does an institution, funder or nation want or need to operate to safeguard 'Open' and what can be left to other parties.

The development of 'Open' involves a variety of stakeholder groups, who each have a vital role to play. The involvement and cooperation of researchers, in their various roles: author, reader, editor, reviewer, conductor of research, project leader, decision maker etc is essential. Without them, 'Open' is just an empty container.

In the 10 to 20 years of development of the various opens, we have seen researchers get involved out of personal interest and/or an interest in the public cause. But these researchers are still primarily a minority group of early adopters. The larger group of researchers is still not on board. And the most quoted reason for this is that it doesn't help them in their career. Whether this is correct or not, it is in any case the overriding perception and it is perhaps the most significant factor in researcher's behaviour and willingness to cooperate in anything else but research.

Efforts so far to change this seem to focus on different types and means of usage statistics and citation analysis. Although this can help, it is not sufficient to solve the problem, for it still focuses on the final, peer-reviewed publication as the measure of 'good' research and for career assessment.

If a change is really wanted, a change in perspective is needed. Research and researchers assessment needs to be based on a broader set of criteria that focus on the contribution to the advancement of knowledge. Such a contribution can be made in many different ways: by publishing an article, but also by educating students, by communicating about research questions in forums and blogs, by making datasets available, by cooperating in 'open' projects to name but a few examples.

The call is often made to offer rewards or incentives for researchers who cooperate in open projects or to force them to cooperate. The only thing that really needs to be done is to take away the barrier imposed by the 'publish or perish' imperative. This can be done by *introducing multi-criteria evaluation for research and researchers based on research communication in aid of advancement of knowledge, rather than on research publication in aid of career advancement.*

Such a measure will means that researchers will stand to gain from cooperation in open developments. Young researchers do not need to be re-socialized into 'traditional' research and publishing behaviour anymore and their knowledge of the use of new tools and ways of working can be put to good use in, and speed up, the further development of 'Open'.

The measure can have additional important effects as well. It can for example reduce the importance of the Green or Gold route debate and ease the copyright worries, because the role of the traditional journal and their publishers is no longer all important. It can also boost the integration of research into teaching. Most importantly, it can result in better research.

5.3.3 The importance of open education as training ground for future researchers

For many, working on 'Open', in an open development setting, is new, so it requires support, training and educating. It can, and will in the shorter term often be training on the job, by participation in projects. However, the new developments in higher education with respect to learning and teaching have an important role to play too. The development of OER contributes to innovation of education and learning as a practice, and as a field of research, with respect to technologies, processes as well as content. As the MITx and edX projects show, in an open educational environment which includes open development methodologies, students from all levels and backgrounds (including professionals and retirees) and tutors both teach and learn from each other about the effect of online tools and means on their learning and teaching processes. *Open Education (as a form of ODM) nourishes a participatory culture of learning, creating, sharing and cooperation and it therefore is a vital and natural training ground for current and future researchers and educators*, turning them into confident users and designers of open approaches in research and higher education. Open Education therefore has an important contribution to make to the sustainable development of an effective and innovative Open research and higher education system.

6 Conclusions and recommendations

This policy document has looked at the development of a range of 'Open' approaches to ascertain whether a more coordinated Open policy framework would benefit the growth of each of them and their joint contribution to a more effective, efficient and innovative research and higher education system. This fits the e-InfraNet project's purpose to further the integration of national, European and global e-infrastructure programmes and to work towards a more efficient use of resources, the adoption of best practice in national e-infrastructure programmes, a convergence of funding programmes and avoiding unnecessary overlap.

There are clear differences in context, drivers, communities and maturity between the various 'opens'. To a certain extent this diversity is a necessity for learning and growth (section 4.4). Yet it is also necessary to strike a careful balance between diversity and commonalities. To realize the benefits from diversity as well as the balance a coordinated approach is needed to create opportunities for the different communities to get to know each other; and to stimulate them to work together on common issues. The cooperation and cross-fertilization achieved in this way will bring solutions that all can benefit from, but that no single community can achieve on its own.

A number of the open approaches are sufficiently mature to enable a move to a more coordinated approach for 'Open' in general. Evidence for this can be found in a new collaborative initiative like the TERENA/LIBER report on Authentication and Authorization[151], and projects like EUDAT (section 3). This development can and should be strengthened and stimulated by the adoption and implementation of a generic 'Open' policy framework. e-InfraNet 's recommendations for this are in line with its focus on the generic level.

1. e-InfraNet recommends that European and National governments and funding organisations, and research and higher education institutions adopt the policy that for publicly funded activities, 'Open' is the default modus operandi.
   'Open' applies to the entire e-Infrastructure for research and higher education including content, as well as to the use of the e-infrastructure in open processes.
   Within this policy, restrictions to access, (re)use and participation can be legitimate but must be justified.
   For the actual implementation of the policy a practical, phased approach is recommended that allows for different speeds for the different 'open' strands and the different national and disciplinary contexts and cultures.
   Acceleration could be made on those strands which are more mature (such as open content and open infrastructure). In the less mature strands more exploratory work could be facilitated.
2. e-InfraNet recommends the adoption of the principles of the 'Open Development Methodology' for the further coordinated development of 'Open' in publicly funded projects. These principles include transparent governance and development processes, building of stakeholder communities and sustainable development with relevant partners.
   - The EC and other funding organisations should build into their Funding Calls (and the presentations of these Calls) requirements for projects to comply with these open development principles and information on how to do this throughout the project life

[151] https://confluence.terena.org/display/aaastudy/AAA+Study+Home+Page;jsessionid=48255D13910F9F0AABD5AED4FE01896B

cycle (from bid stage to final evaluation of project results). Compliance should also be included in the project monitoring and evaluation systems.
    - Institutions should build the open development principles into their project management methods.

3. e-InfraNet recommends that the involvement and cooperation of researchers in the further development of 'Open' is actively stimulated and that barriers are removed by:
    a) the adoption of appropriate evaluation and reward systems
    b) provision of support and advice in copyright and licensing issues
    - European and national funding organisations and institutions should expand present research evaluation and reward systems from those based primarily on quantitative publication criteria into multi-criteria-evaluation systems including quantitative and qualitative criteria. These criteria should focus on open research communication in the broadest sense of aiding in the advancement of knowledge.
    - European and national funding organisations should require that all research outputs resulting from publicly funded projects are published under non-exclusive licenses (Creative Commons or similar).
    - Institutions should offer structural support and advice with respect to licensing to researchers and other staff involved in publishing and dissemination of research and educational materials.
4. e-InfraNet recommends that researchers and other staff who work on 'Open' and in an 'open development process' are actively supported with appropriate training and support.
    In the short term, dedicated training and support should be provided in the context of open projects.
    For the long term, curricula should be developed for the training of specialized research support staff.
    Furthermore, Open Education initiatives should be developed as a vital, natural training ground for the 'researcher of the future' (and other professionals).
    - European and national funding organisations should allow for a portion of open project budgets to be allocated for dedicated training and support of researchers and other staff.
    - Institutions should invest in the appointment and dedicated training of support staff for 'Open'.
    - Institutions are recommended to seek cooperation in the design and development of curricula for training of professional research support staff specialized in 'open' approaches.
    - European and national funding organisations, and institutions, should invest in the (collaborative) development of Open Education initiatives as a training ground for the researcher of the future, to develop the non-technical skills and mindset required for working in an open environment.
5. e-InfraNet recommends that interoperability of systems, services and content is continuously supported as a key component in the open e-infrastructure.
    - European and national funding organisations and e-infrastructure providers keep investing in and supporting the development and application of open standards and other open technical approaches, to improve the interoperability across systems, services and content.
6. e-InfraNet recommends that the EC continues to invest in research and development work in the area of 'Open', with respect to policies, good practices and technical approaches.

With a clear and conscious choice for 'Open' as the default modus operandi in research and higher education, the EC, national governments, funders, infrastructure providers and institutions lay a firm basis for a more coordinated and sustained development of open approaches. This choice wil be a great support to the communities that shape the development. It will contribute towards more intensive, shared use of the valuable resources in the research and higher education system. It will also optimize the conditions for innovative Research and Development, both within and outside academia, which rely upon openness in infrastructure, content and work processes.

Annexes

A Stakeholder map

	Enables reuse (easier discovery and access, clearer more permissive rights position, easier curation)	**Supports innovation and agility** (promotes change, enables organisations to respond well to change, innovation in technologies, business models)	**Increases cost effectiveness** (enables collective and shared approaches, reduces duplication of effort, saves time, supports collaboration)	**Improves quality** (by ensuring visibility of material, uses peoples' concern with their reputation to see better quality work shared, supports review of work)	**Enables better risk management** (including easier compliance with legal/regulatory requirements)
Research institutions	high	high	high	medium	medium
Teachers/ Tutors	medium	medium	medium	high	medium
Researchers	high	medium	medium	high	medium
Learners (students, lifelong)	medium	low	low	medium	low
Research funders	medium	medium	high	high	medium
Infrastructure providers	medium	medium	high	medium	high
Content providers	high	high	high	high	high
Data centres	low	medium	high	medium	high
Standards bodies	high	medium	medium	high	high
Publishers	medium	medium	high	high	high
Developer communities	medium	medium	medium	medium	medium
Commercial organisations	medium	medium	high	high	high
Non-commercial organisations	medium	high	medium	high	high
European Commission	high	high	high	high	high

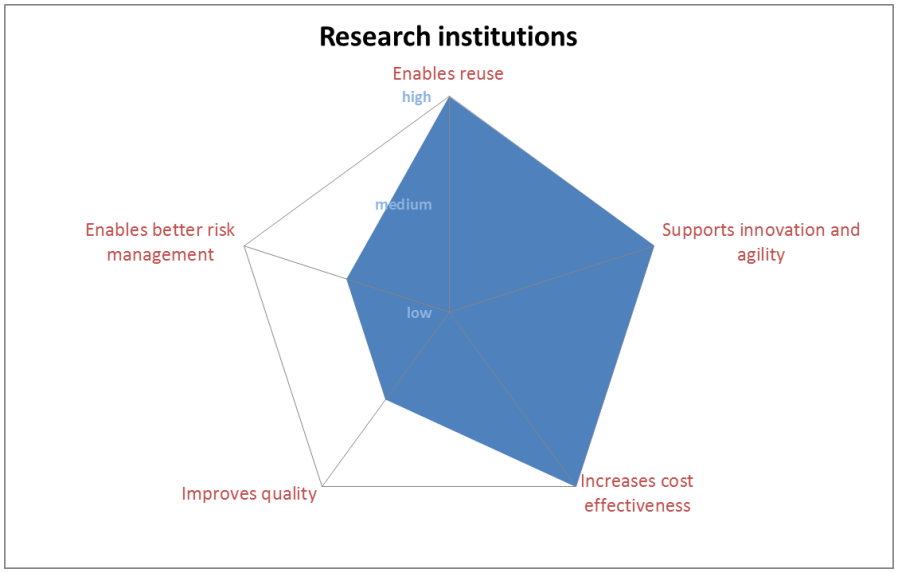

Figure 6—Stakeholder impact—Research institutions

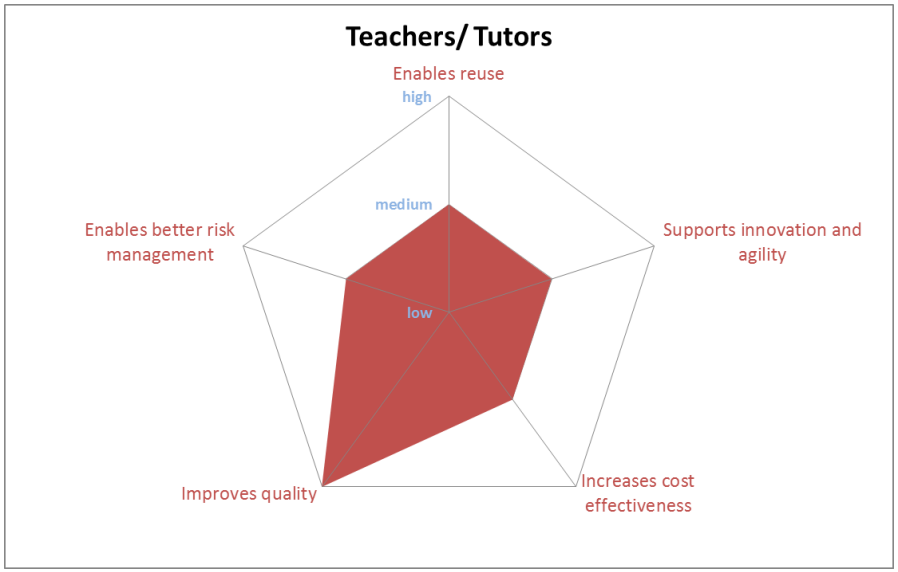

Figure 7—Stakeholder impact—Teachers/Tutors

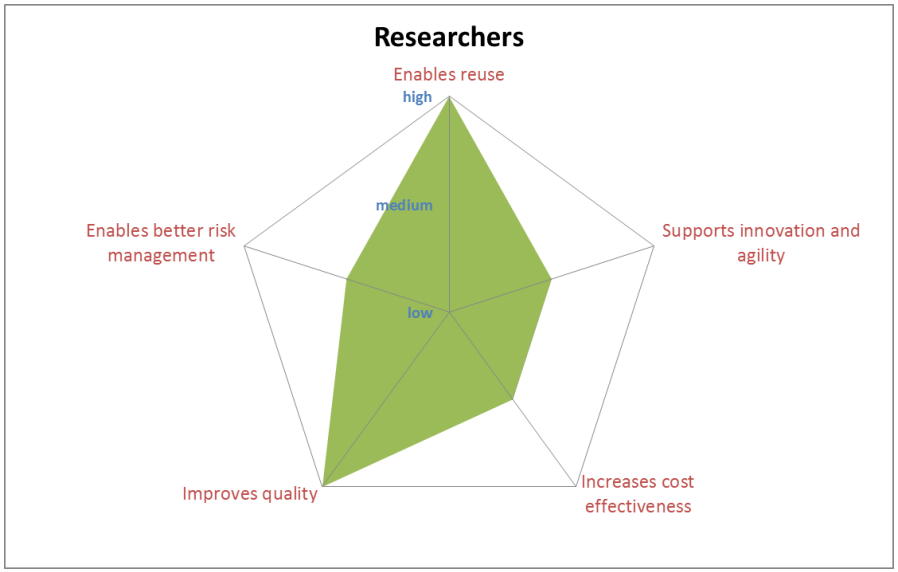

Figure 8—Stakeholder impact—Researchers

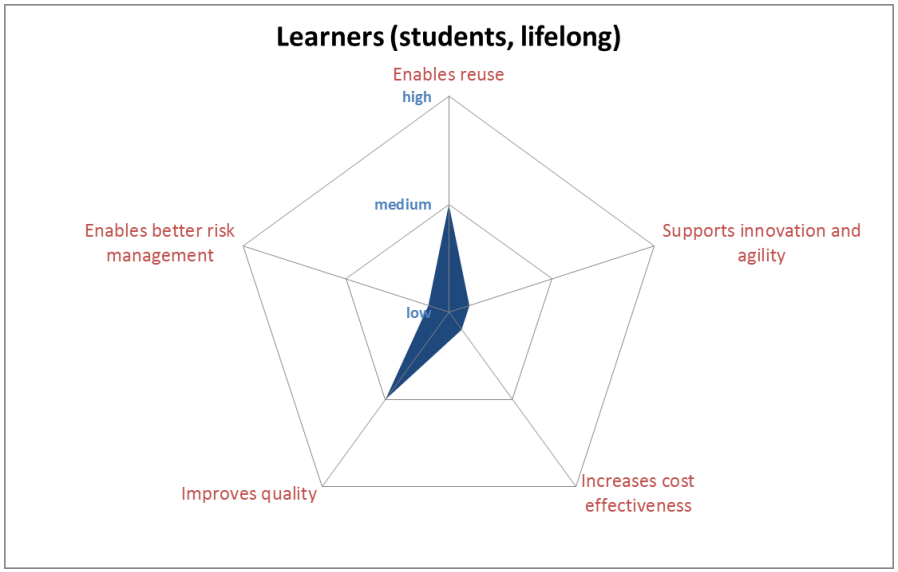

Figure 9—Stakeholder impact—Learners (students. lifelong)

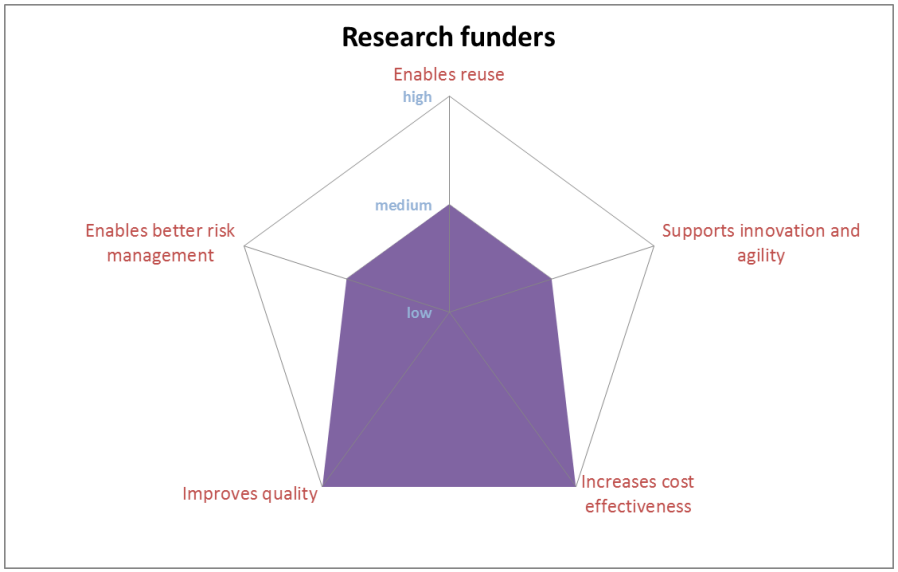

Figure 10—Stakeholder impact—Research funders

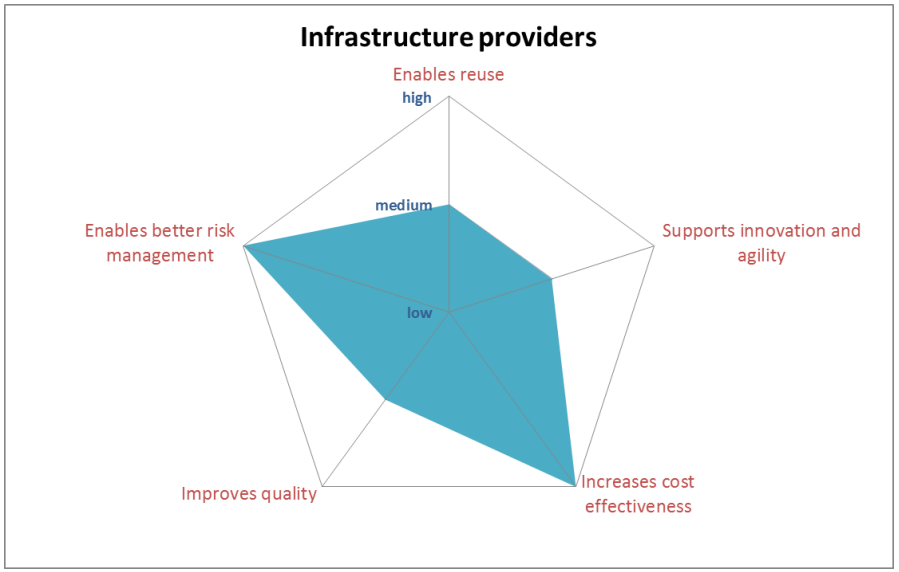

Figure 11—Stakeholder impact—Infrastructure providers

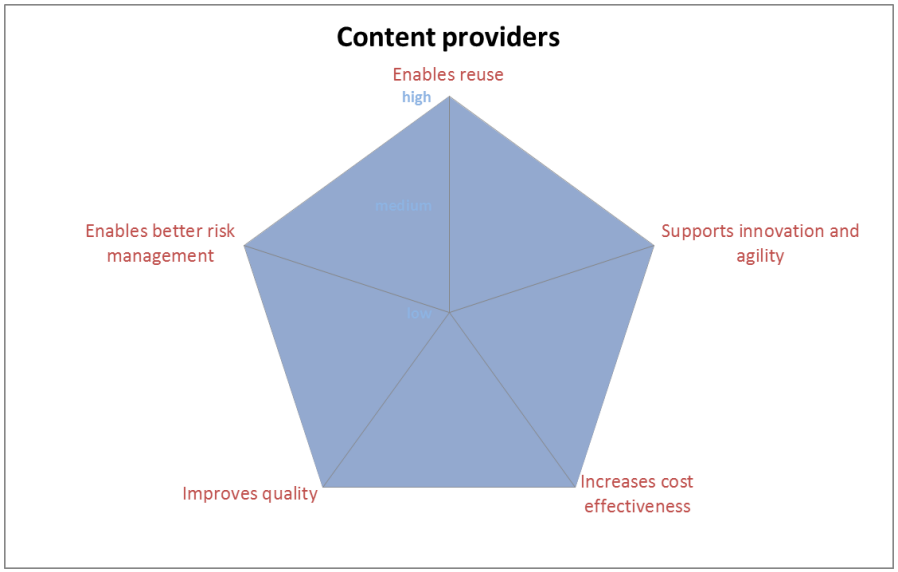

Figure 12—Stakeholder impact—Content providers

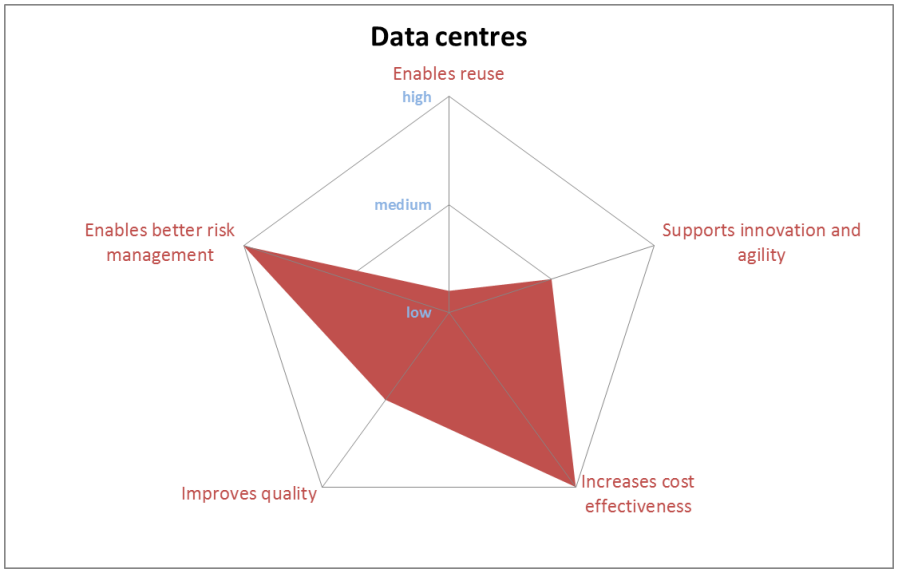

Figure 13—Stakeholder impact—Data centres

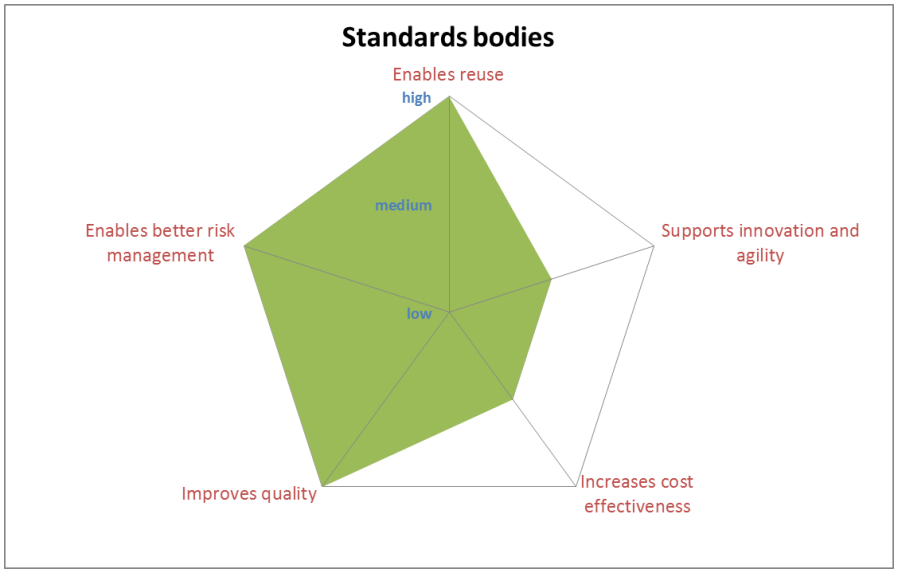

Figure 14—Stakeholder impact—Standards bodies

Figure 15—Stakeholder impact—Publishers

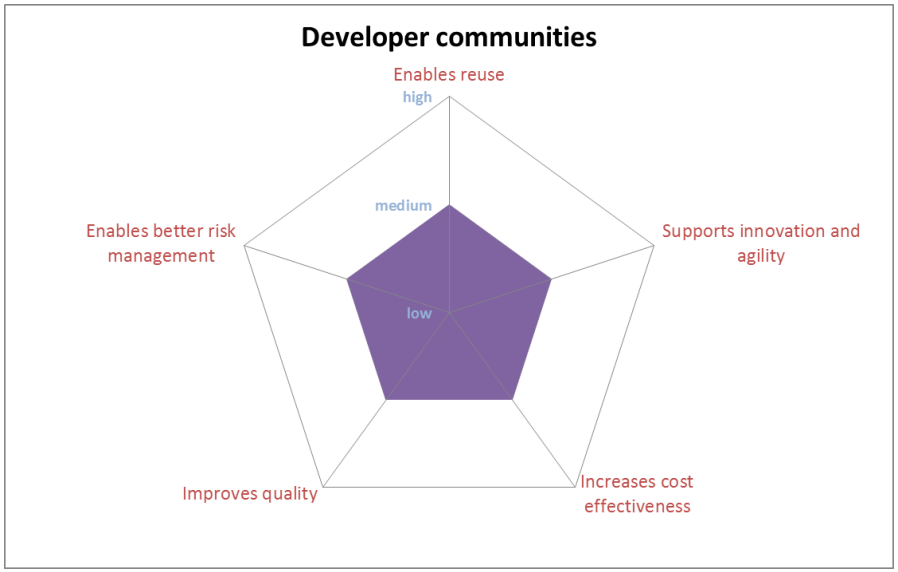

Figure 16—Stakeholder impact—Developer communities

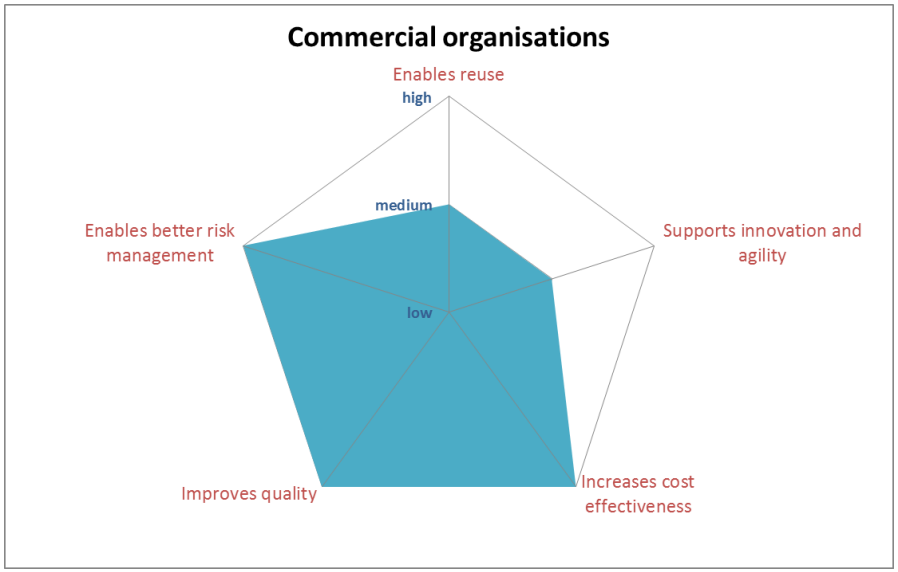

Figure 17—Stakeholder impact—Commercial organisations

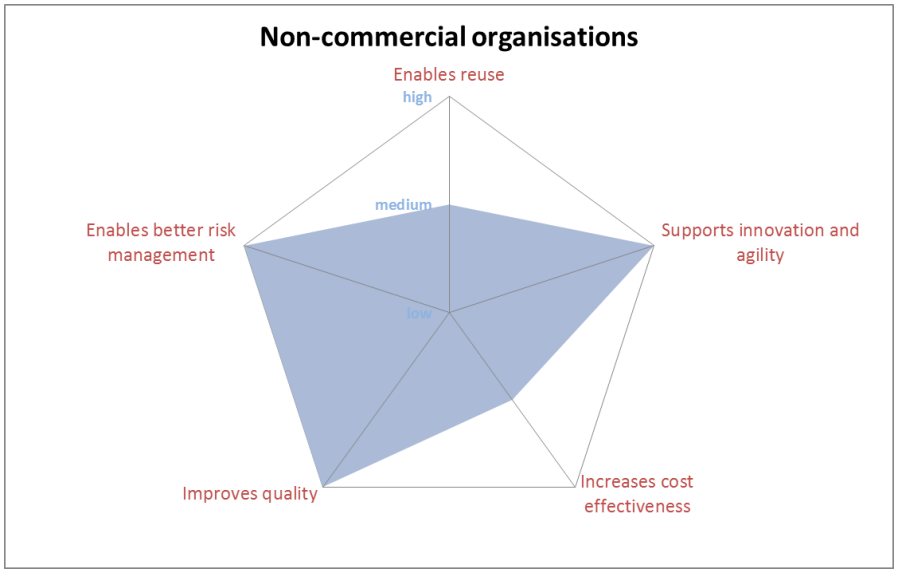

Figure 18—Stakeholder impact—Non-commercial organisations

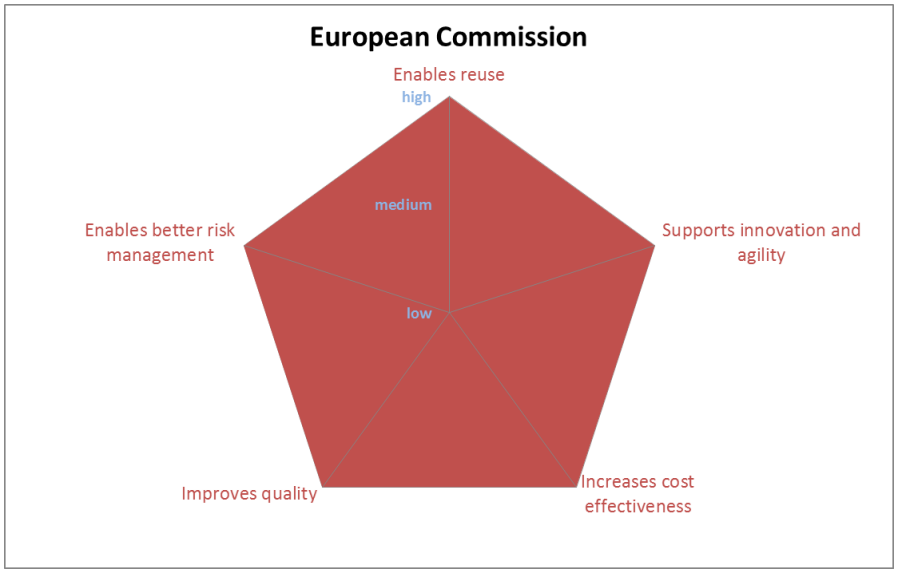

Figure 19—Stakeholder impact—European Commission

B Reports and projects consulted

B.1 Reports

i. Open Access to research literature

Accessibility, sustainability, excellence: how to expand access to research publications, Report of the Working Group on Expanding Access to Published Research Findings—the Finch Group 2012, http://www.researchinfonet.org/publish/finch/

In 2011, following the generation of a strong evidence base in support of OA, the UK Government initiated an independent working group under Professor Dame Janet Finch whose remit was to examine and make recommendations on both expanding access to UK research outputs, and expanding access for UK researchers to global research outputs. The report was released in June 2012. Its main recommendations are that the UK should transition to Gold OA, and that various extensions should be made to licensing arrangements to ensure good access for UK researchers. The report estimated that an extra £50m-£60m per year would be needed during the transition to pay for Gold OA article processing charges and other costs. Responses to the report have been varied, but were mainly positive. The UK Government has endorsed the report's conclusions and recommendations, though suggested a longer embargo period for Green OA than many hoped. Research intensive universities have expressed some disappointment that the Finch Report and Government response downplayed Green OA. The UK Research Councils have released a combined policy on OA, to come into force in April 2013, which requires grant-holders to make their published papers OA via either Green or Gold OA, and will make block grants available to universities for article processing charges. The Higher Education Funding Council for England (HEFCE)[152] has announced its intention to consult with the sector about any implications for the Research Excellence Framework 2020.

National open access and preservation policies in Europe: Analysis of a questionnaire to the European Research Area Committee—EC-DG for Research and Innovation 2011, http://ec.europa.eu/research/science-society/document_library/pdf_06/open-access-report-2011_en.pdf

This report gives an overview of how open access is developing in the European Research Area. It's based on a survey conducted via the European Research Area Committee and shows that open access is backed by a growing number of universities, research centres and funding agencies across Europe. In addition it highlights the dynamic growth of open access. It also stresses, however, that national initiatives and practices are still fragmented, thus preventing the European Union from realising its full research and innovation potential.

Survey on open access in FP7—EC-DG for Research and Innovation 2012, http://ec.europa.eu/research/science-society/document_library/pdf_06/survey-on-open-access-in-fp7_en.pdf

In August 2008, the European Commission launched the open access pilot: in seven FP7 research areas grant beneficiaries are expected to deposit peer-reviewed research articles or final manuscripts resulting from their projects into an online repository and make their best efforts to ensure open access to those articles within a set period of time after publication. All projects also have open access fees eligible for

[152] http://www.hefce.ac.uk/

reimbursement during the time of the grant agreement. In May 2011, the Commission identified the 811 projects designated at the time and sent a questionnaire to all project coordinators in order to collect feedback on their experiences of both the implementation of the pilot and the reimbursement of open access publishing costs. A total of 194 answers were received by the end of August 2011. They provide important input for the future of the open access policy and practices in Horizon 2020, and for the preparation of a communication from the Commission and a recommendation to Member States on scientific publications in the digital age.

Online survey on scientific information in the digital age—EC-DG for Research and Innovation 2012, http://ec.europa.eu/research/science-society/document_library/pdf_06/survey-on-scientific-information-digital-age_en.pdf

In preparation of the European Commission's intended communication and recommendation on access to and preservation of scientific information in the digital age, the 'Online survey on scientific information in the digital age' (open from 15.07.2011 to 09.09.2011) provided stakeholders with the opportunity to comment on the state of play, barriers and potential policy actions in the area of access to and preservation of scientific results. All citizens and organisations concerned were welcomed to contribute to this consultation. Contributions were particularly sought from governments, research institutions and universities, libraries, scientific publishers, research funding organisations, businesses, individual researchers and other interested parties. The consultation spurred great interest among those stakeholders, with 1140 responses received from 42 countries.

The purpose of the consultation was to gather information from as many sources as possible and receive important input for the future development of policy options in the area of scientific information in the digital age.

Open Access success stories—Knowledge Exchange, http://www.oastories.org/

Open Access success stories complements the existing, more scientific, reports on the advantages of Open Access. The value of the success stories lies in the fact they offer a more personal account of the positive experiences scientists, editors or publishers have with Open Access and can be used as motivators for people for OA. The disadvantage is they might be too anecdotal to be used as input for policy development.

ii. Open Data

Riding the Wave: How Europe can gain from the rising tide of scientific data—High level Expert Group on Scientific Data 2010, http://www.grdi2020.eu/Repository/FileScaricati/c2194260-3ddf-47bd-93e4-68f8912a3564.pdf

A fundamental characteristic of our age is the rising tide of data—global, diverse, valuable and complex. In the realm of science, this is both an opportunity and a challenge. This report, prepared for the European Commission's Directorate-General for Information Society and Media, identifies the benefits and costs of accelerating the development of a fully functional e-infrastructure for scientific data—a system already emerging piecemeal and spontaneously across the globe, but now in need of a far-seeing, global framework. The outcome will be a vital scientific asset: flexible, reliable, efficient, cross-disciplinary and cross-border.

A Surfboard for Riding the Wave: Towards a four country action programme on research data—Knowledge Exchange 2011, http://www.knowledge-exchange.info/default.aspx?id=469

The *Riding the Wave* report by the high level expert group on research data (see above) called for a collaborative data infrastructure that will enable researchers and other stakeholders from education, society and business to use, reuse and exploit research data to the maximum benefit of science and society. The Knowledge Exchange partners embraced this vision and commissioned a report that translates Riding the Wave into actions for the four partner countries and beyond.

This paper builds on the Riding the Wave report and presents an overview of the present situation with regard to research data in Denmark, Germany, the Netherlands and the United Kingdom. It offers broad outlines for a possible action programme for the four countries in realising the envisaged collaborative data infrastructure. An action programme at the level of four countries will require the involvement of all stakeholders from the scientific community.

The report formulates three long-term strategic goals, and describes an action programme to achieve them:

- Data sharing will be part of the academic culture
- Data logistics will be an integral component of academic professional life
- Data infrastructure will be sound, both operationally and financially.

Science as an Open Enterprise, final report—UK Royal Society 2012, http://royalsociety.org/policy/projects/science-public-enterprise/report

The Royal Society report *Science as an Open Enterprise* was published in June 2012, alongside a plethora of reports in the UK with regard to opening up access to research and data; it is closely aligned to the UK Governments *Open Data White Paper* (see below), although the UK Government White paper is focused largely on government data and the Royal Society focuses on data produced as part of research. The report aims to address how science should change in response to technological advances. It states that openness is central to scientific enquiry in that it supports better research, reviewed research and the development of new knowledge. However with current technology there are now new ways to manipulate, store, share and communicate data as part of science. These new technologies open up new improved methods of science and, whilst we are witnessing changed behaviour, this is piecemeal and there are no systematic policies and fit for purpose infrastructures in place. The report states that much behaviour in research duplicates that of the paper world and the paper journal and that this is holding back advances in science.

The report recognises that openness is central, but that there are key issues that need to be addressed and recognised. Thus the report proposes "intelligently open data" as the answer. Intelligently open means data must be accessible and easy to locate and it must be intelligible. The report does not shy away from some of the dry technicalities. For example it states that metadata is essential, that citation is a must, especially to reward open science practice, and the report fully recognises that there will be costs involved. However, it is acknowledged that the benefits can far outweigh the costs, and that a short term, limited view must not be taken; this would lead to a "tragedy of the commons". An example of the benefits of intelligently open data is illustrated by the following excerpt:

> *"[…] the events following an outbreak of a severe gastro-intestinal infection in Hamburg in Germany in May 2011. This spread through several European countries and the US, affecting about 4000 people and resulting in over 50 deaths. All tested positive for an unusual and little-known Shiga-toxin–producing E. coli bacterium. The*

> *strain was initially analysed by scientists at BGI-Shenzhen in China, working together with those in Hamburg, and three days later a draft genome was released under an open data licence. This generated interest from bio-informaticians on four continents. 24 hours after the release of the genome it had been assembled. Within a week two dozen reports had been filed on an open-source site dedicated to the analysis of the strain. These analyses provided crucial information about the strain's virulence and resistance genes – how it spreads and which antibiotics are effective against it. They produced results in time to help contain the outbreak. By July 2011, scientists published papers based on this work. By opening up their early sequencing results to international collaboration, researchers in Hamburg produced results that were quickly tested by a wide range of experts, used to produce new knowledge and ultimately to control a public health emergency." (p. 15)*

The report covers the concerns of privacy and is robust about the fact that personal data is required in research highlighting that provisions need to be in place with regard to governance and management in order to provide individuals with protection. Consent is not the way to deal with this issue. It does not adequately defend interests that go beyond the individual. Of course the report does call for responsible data sharing that respects privacy and security. Another thorny issue that the report addresses is that of competition—for example if the UK opens up its research will it be giving away its knowledge assets? Besides the point that research is a global endeavour the report outlines that opening up research is essential and that with

> *"rising scientific powers such as China, India and Brazil and the growth of scientific efforts in the Middle East, South-East Asia and North Africa we find many have signed up to the principles of open data through membership of the International Council of Science (ICSU). In addition, international collaboration that depends on the open data principle is increasingly supported by inter-governmental funding or funding from international agencies. Such collaboration focuses on matters of global concern such as climate change, energy, sustainability, trade, migration and pandemics. The OECD Global Science Forum Expert Group on Data and Research infrastructure for the Social Sciences will produce a report in Autumn 2012." (p. 18)*

In summary the report recommends:

- Openness as standard—the default position is openness and responsible sharing, with only justifiable and reasonable restrictions on access and use.
- Clear policies—all custodians should have transparent policies for custodianship, data quality and access.
- Clear routes to access—positive action must be taken to facilitate openness through the creation of clearly signposted data registers.
- Pre-specification of data release—when and how newly acquired data will be released for reuse must be specified in advance.
- Respect for values that bound openness—governance mechanisms need to be in place for protection of privacy and commercial interests.
- Rules of sharing must be made explicit in the terms and conditions for data sharing.

Open Data White Paper, Unleashing the potential—UK Government 2012, http://data.gov.uk/library/open-data-white-paper

The UK Government undertook a consultation around Open Data in 2011 and in June 2012 released its White Paper on the matter. The premise is that open data supports transparent government, improved policy and understanding, innovation and collaborations across public, private and the academic sectors and citizen participation. The UK Government has undertaken an initiative on open data like many governments over the past five years, however this new white paper clarifies the position on the benefits, the standards and infrastructure required and some of the areas where open needs to be managed carefully. The paper recognises that pushing data out is only part of the answer and that there is a requirement for up-to-date legislation and common standards for the data to support what the paper calls "an effective Open Data ecosystem".

Whilst promoting openness to government data to the full, the paper stresses the importance of anonymity and privacy of the individual and attempts to strike the right balance; calling for the maximum openness but implementing new structures such as 'data safe havens' that will allow the support of modern personalised services but not undermine privacy. The Government remains convinced that transparency and privacy are compatible if the right governance structures are in place. It quotes a case study from the Ministry of Justice where they worked in partnership with a group of academics, students and professional data security consultants, and delivered a project that ensured that the anonymity of individual re-offending and sentencing data would be as secure as possible. This project showed openness and privacy can work side by side.

In terms of standards and infrastructure for openness the paper promotes open licences for re-use, open linked data (based on Tim Berners-Lee five stars —see section 3 above) and the sharing of APIs for data sources; alongside this they will enhance the http://data.gov.uk site and ensure data is both available and useable. Usability is an issue that is believed to be important to openness and transparency, after all there is little point in being open if the result is more confusion and frustration. The example of the Tabulation Tool[153] developed by the Department of Work and Pensions is an example of how data providers can improve the quality and accessibility of their data; in this case the tool allows users to download national statistics formulated to their own requirements.

However, recognising that standards and usability are not issues that can be addressed immediately, the UK Government says it will ensure training of staff to comply with necessary standards and also ensure the public has easy to use channels to address queries. Furthermore, the Freedom of Information legislation has been altered so that people can ask for a data set in a standard that is useable for them.

Many benefits are referenced, such as the possibility of new discoveries and the use of government data by the researcher or the citizen to enhance understanding. An example of how the paper sees open data driving better services and innovations is the Nation Archives work to open up legislation data at www.legislation.gov.uk:

> *"This has enabled the National Archives to develop a new, transferable operating model for updating government databases; and delivers an easy to use interface that makes it easy for anyone to access legislation data by adding /data.xml or /data.rdf to any web page containing legislation etc. All the data can be re-used free of charge under the Open Government Licence. This has led to the development of third-party*

[153] http://statistics.dwp.gov.uk/asd/index.php?page=tabtool

> *applications, including smartphone apps and a service for law lecturers to create and self-publish relevant extracts of legislation for their courses." (p. 25)*

The white paper also promotes the open sharing of research outputs including data and, like the Royal Society report (Considering science as a public enterprise), wants this to be shared in ways that those outside of the research community can benefit from it; this is also in line with the UK Governments Research and Innovation strategy. However, the white paper does recognise some of the special conditions that surround this and that need to be addressed; these will be examined via a Research Transparency Sector Board, protecting the integrity of the research and associated intellectual property, while ensuring access to research for those SME entrepreneurs vital for driving growth. In support of the vision for Open Data driving innovation the Government has established The Open Data Institute[154] to demonstrate the commercial value of Open Data by working with the public and private sectors as well as academia in developing its exploitation.

It is interesting to note that the White Paper recognises the value of key information sets and how opening these up can help to link data and drive innovation and efficiency. An example of this might be geographies and the support of the INSPIRE directive so data is geospatially linked. The paper supports the argument that opening up data should mean that data can be improved by receiving feedback and that it can also avoid re-keying of the same data, thereby driving efficiencies in the system. An aspiration of the Government is to look for savings in processes, for example the paper states that the

> *"decennial Census is an example where data is collected for the purpose of counting the population; this cost nearly £500 million for the 2011 Census. However several government data collections already have partial coverage of the population and, if these were brought together and the data re-used, could provide an alternative count to that from a traditional enumeration at a reduced cost." (p. 38)*

Note: The European Commission has recently launched a wide ranging open data initiative which it expects will generate €140 billion a year of income. The Commission will open its own stores of data through a new portal.

iii. Open Educational Resources

Trend Report: Open Educational Resources 2012—Ria Jacobi and Nicolai van der Woert (edd.), Special Interest Group Open Educational Resources 2012, https://www.surfspace.nl/media/bijlagen/artikel-697-ee18ac0f1441bb158e6122818f5f589e.pdf

This report by the Dutch Special Interest Group Open Educational Resources gives an overview on the status quo of OER in the Netherlands. Twelve articles describe trends in the Netherlands and worldwide, taking four perspectives: the educational perspective, the content-related perspective, the technological perspective and the organisational perspective. The report is comprehensive covering both historical developments as well as inspirational for new initiatives.

Guidelines for Open Educational Resources (OER) in Higher Education—COL, UNESCO 2011, http://unesdoc.unesco.org/images/0021/002136/213605E.pdf

These guidelines outline key issues and make suggestions for integrating OER into higher education. Their purpose is to encourage decision makers in governments and institutions to invest in the systematic

[154] www.theodi.org

production, adaptation and use of OER and to bring them into the mainstream of higher education in order to improve the quality of curricula and teaching and to reduce costs.

Open Educational Resources, analysis of responses to the OECD country questionnaire, Jan Hylén, Dirk V. Damme, Fred Mulder, Susan D'Antoni—http://www.oecd-ilibrary.org/education/open-educational-resources_5k990rjhvtlv-en

OECD's Centre for Educational Research and Innovation (CERI) has worked on Open Educational Resources (OER) in the past, which led to the publication *Giving Knowledge for Free—the Emergence of Open Educational Resources* (2007). This questionnaire for member countries was set up to collect information regarding the policy context related to OER. The questionnaire was sent to the 34 OECD member countries in August 2011. It outlined a short informative note about the benefits and challenges of OER. The analysis of the responses to the questionnaire concludes that almost all OECD member countries indicate that they are in one way or another active in the area of OER, mostly by involvement with specific projects or programmes or through the initiative of institutions or engaged individuals. Several countries, especially those with federal systems, indicate that they have insufficient knowledge about the OER activities in their educational institutions. The most frequently cited policy reason for OER activity is the desire to increase access to high-quality learning materials.

In contrast to the conventional understanding that situates OER mainly on the post-secondary educational level, OER activity seems to be spread across the educational spectrum in the view of officials responding to the Questionnaire.

The majority of countries surveyed indicated that governments take great responsibility over the production, financing and distribution of educational resources although most countries have little information about whether these resources are also available in digital formats. Countries indicate that the use of licensing (CC licenses in particular) is well distributed. Nevertheless, precise information in this regard is generally lacking.

Countries attach a high importance to the cited benefits of OER. OER offers open and flexible learning opportunities. Almost equally appreciated is the cost efficiency of OER as well as the possibility of increased quality in learning resources and the increased flexibility that OER can offer.

Policy-relevant research on OER exists in some countries, but here again precise information is often lacking. At present, the two areas best covered by research are raising the quality of learning outcomes and the widening of access to educational opportunities.

iv. Open Infrastructure and standards

e-Infrastructures in support of the Digital Agenda—e-IRG , http://www.e-irg.eu/images/stories/digital_agenda_a5_final.pdf

This survey describes the essential role of the e-Infrastructure community for the realisation of the Digital Agenda and the importance of a proper e-Infrastructure governance. Recent e-IRG observations, recommendations and actions issued in White Papers and other e-IRG policy documents are summarized and linked to the seven problem areas recognised in the Digital Agenda. The aim is to provide guidance on the developments and actions required to achieve the goals of the Digital Agenda.

Advancing technologies and Federating communities" : Study on Authentication, Authorization and Accounting (AAA) Platforms For Scientific data / information Resources in Europe—TERENA/LIBER/University of Amsterdam/University and Library of Debrecen 2012, https://confluence.terena.org/download/attachments/30474266/AAA-Study-Report-0907.pdf?version=1&modificationDate=1341850616400

A study by TERENA, LIBER, University of Amsterdam and DEENK *Advancing technologies and Federating communities* provides an inventory of the main (federated) access systems currently available and in use for networks, web applications, 'the grid', supercomputing and cloud. At the technical level, all studied infrastructures individually provide seamless access/Single-Sign-On (SSO) for their users, although the technology used varies per sub-system. It is acknowledged that "no technology can be universally adopted, but there should be mechanisms in place to allow for integration of different technologies." Apart from the technical dimension, legal, policy and practical aspects also impact the actual level of 'openness' of the systems. The study points out that for example at network level, "the level of deployment, the participation of institutions and the amount of services available … …vary significantly from country to country." Authentication and Authorization (AA) systems for grid and supercomputing are considered to be fairly complicated and face usability and scalability issues. Cloud infrastructure offers yet other challenges such as identity management, but possibly also solutions with wider potential in the entire e-infrastructure.

Legal complications (in particular data protection legislation) play a role here, since there is no common legal framework and interpretation in Europe. Work is going on to address this at EU level with the proposed draft Data Protection Regulation.

The current Authentication and. Authorization Infrastructure (AAI) approaches also present a barrier to access for non-traditional participants in the research and higher educational processes, such as academically trained professionals in SME's (as participants in innovation processes), and members of the general public (as participants in, for example, open/citizens research). The future development of AAI needs to take these groups of users into account as well.

v. Open Research

Open science at web-scale: Optimising participation and predictive potential—JISC 2009, http://www.jisc.ac.uk/publications/reports/2009/opensciencerpt.aspx

This report has attempted to draw together and synthesise evidence and opinion associated with data-intensive open science from a wide range of sources. The potential impact of data-intensive open science on research practice and research outcomes, is both substantive and far-reaching. There are implications for funding organisations, for research and information communities and for higher education institutions.

B.2 Projects / Initiatives

i. Open Access to research literature

OpenAIRE—www.openaire.eu

OpenAIRE is essentially a support infrastructure for the OA policies of the European Commission. The original scope of the project included the OA Pilot in FP7[155] and the ERC Guidelines on OA[156], focussing on peer reviewed articles. With the start of OpenAIREplus in 2011 this scope was expanded to research data.

OpenAIRE combines a human support infrastructure with a technical infrastructure. The latter is a continuation of the DRIVER infrastructure[157]. Characteristic of OpenAIRE is the 'distributed approach' of these infrastructures: on a local—in this context national and institutional—level work is done to achieve the desired results at the central level. The portal of OpenAIRE provides extensive information on the status quo of OA and Open data in the member states.

OpenAIRE is strongly supported by the EC, and is closely involved in the plans of the EC to open all content arising from the Horizon 2020 programme[158].

ii. Open Data

EUDAT—www.eudat.eu

The EUDAT project, which started in October 2012, is exploring ways to build generic services that can support multiple communities and is working with communities from linguistics, earth sciences, climate sciences, environmental sciences, and biological and medical sciences to create a sustainable, open, cross-disciplinary and cross-national data infrastructure that provides a set of shared services for accessing and preserving research data.

The EUDAT vision is to support a Collaborative Data Infrastructure as envisioned by the High level Expert Group on Scientific Data in its report 'Riding the wave'(see above), which will allow researchers to share data within and between communities and enable them to carry out their research effectively. EUDAT aims to provide a solution that will be affordable, trustworthy, robust, persistent and easy to use.

EUDAT aims to:

- Help fulfil the vision of a European Data e-infrastructure by providing a sustainable platform of technologies, tools and services driven by user needs.
- Engage users (including individual researchers along with representatives from universities, research labs, and libraries) in defining and shaping a platform for shared services that makes it possible for data-intensive research to span all the scientific disciplines.
- Produce the common low-level services that are required to provide the level of interoperation and trust of data that is necessary to support both widespread access to data, and the long-term preservation of data for use and re-use.
- Ensure that the data infrastructure is sufficiently robust to keep pace with the expected acceleration of the scale and complexity of scientific data being generated within the ERA and beyond

[155] http://ec.europa.eu/research/science-society/document_library/pdf_06/open-access-pilot_en.pdf
[156] http://erc.europa.eu/sites/default/files/document/file/erc_scc_guidelines_open_access.pdf
[157] http://www.driver-repository.eu
[158] http://europa.eu/rapid/pressReleasesAction.do?reference=IP/12/790&format=HTML&aged=0&language=EN&guiLanguage=en

Research Data Alliance (RDA)—http://www.rd-alliance.org

In various regions of the globe, there is now a shared understanding that solutions must be global and that the development of an integrated and interoperable data domain can only be achieved through increasing global collaboration. The "Research Data Alliance" (RDA) is a recent international initiative aiming to steer the efforts in this area. The vision of RDA is to make it possible for researchers around the world to share and use research data without barriers. They aim to accelerate international data-driven innovation and discovery by facilitating research data sharing and exchange, use and reuse, standards harmonization, and discoverability. This will be achieved through the development and adoption of infrastructures, policies, practises, standards, and other deliverables. The RDA will be based on working groups aiming to accelerate progress in concrete ways for specific communities, by increasing innovation, coordinating the global data community across disciplinary boundaries, and promoting data sharing and exchange, interoperability, use and re-use, discoverability analysis, stewardship and preservation, and best practice.

DataCite—http://www.datacite.org/

DataCite is a not-for-profit organisation formed in London on 1 December 2009 in order to:

- establish easier access to research data on the Internet
- increase acceptance of research data as legitimate, citable contributions to the scholarly record
- support data archiving that will permit results to be verified and re-purposed for future study

It brings together the datasets community to collaboratively address the challenges of making research data visible and accessible. Members of DataCite meet in person every six months at summer and winter conferences, and collaborate in established working groups. The focus is on:

- supporting researchers by helping them to find, identify, and cite research datasets with confidence
- supporting data centres by providing persistent identifiers for datasets, workflows and standards for data publication
- supporting journal publishers by enabling research articles to be linked to the underlying data

iii. Open Educational Resources

CETIS—http://jisc.cetis.ac.uk/

The Centre for Educational Technology and Interoperability Standards (CETIS) provides advice to the UK Higher and Post-16 Education sectors on educational technology and standards. The website brings together educational technology news, comment and analysis, as well as information on community events. The aim of CETIS is to contribute to current debates and future thinking in this rapidly growing and changing field.

OER programme of UNESCO—http://www.unesco.org/new/en/communication-and-information/access-to-knowledge/open-educational-resources

UNESCO is in favour of OER and collaborates on issues of the production and use of OER. Over the last 10 years, UNESCO has helped spur an international movement in support of OERs. The term Open Educational Resources (OER) was coined at a 2002 UNESCO Forum on the Impact of Open Courseware for Higher Education. In 2012 UNESCO hosted the World Open Educational Resources Congress, which

adopted the 2012 Paris OER Declaration[159]. Furthermore UNESCO hosts the Open Training Platform[160], a comprehensive database of OER.

iv. Open Infrastructure and standards

EGI-InSPIRE—http://www.egi.eu/about/egi-inspire/

The EGI-InSPIRE project (Integrated Sustainable Pan-European Infrastructure for Researchers in Europe) co-funded by the European Commission for four years, started on 1 May 2010 as a collaborative effort involving more than 50 institutions in over 40 countries. Its mission is to establish a sustainable European Grid Infrastructure (EGI). EGI-InSPIRE is ideally placed to join together the new Distributed Computing Infrastructures (DCIs) such as clouds, supercomputing networks and desktop grids, for the benefit of user communities within the European Research Area.

v. Open Source Software and Open Development

OSS-Watch—www.oss-watch.ac.uk

OSS Watch is the non-advocacy open source software service for UK higher and further education. The service is funded by Jisc to help academic institutions find the best software solutions for their projects (either open source or proprietary), and to offer them licence advice and expertise on building sustainable communities around the developed software. The service, based at the University of Oxford, publishes reports, briefing documents, blog posts and news about all aspects concerning the procurement, licensing, developing, exploitation and sustainability of open source software in higher and further education. Telephone and email support, as well as one-to-one consultations, are freely provided to the sector, and events are organised regularly on open source topics of interest to the projects. Support with writing the sections of funding applications concerning resource allocation and management of open source development is also available.

[159] http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/Events/English_Paris_OER_Declaration.pdf
[160] http://otp.unesco-ci.org/

B.3 Other sources

i. Open Access to research literature

- Budapest Open Access Initiative—http://www.soros.org/openaccess
- The Berlin Declaration—http://oa.mpg.de/files/2010/04/berlin_declaration.pdf
- OAI-PMH—Open Archives Initiative Protocol for Metadata Harvesting,
  http://www.openarchives.org/OAI/openarchivesprotocol.html#Introduction
  http://en.wikipedia.org/wiki/Open_Archives_Initiative_Protocol_for_Metadata_Harvesting
- DRIVER/DRIVER II—http://www.driver-repository.eu
  Both DRIVER projects were primarily aimed at setting up a network of institutional repositories.

ii. Open Data

- OECD Principles and Guidelines for Access to Research Data from Public Funding—
  http://www.oecd.org/dataoecd/9/61/38500813.pdf
  The Guidelines stress the importance of guaranteeing open access to and use of data from publicly funded research, as they are an essential and valuable resource to share, to maximise their benefits for better, more robust research and for innovative solutions to complex and large-scale problems.
- ENVRI—http://envri.eu/home
  Implementation of common solutions for a cluster of ESFRI infrastructures in the field of "Environmental Sciences" (ENVRI)
- DASISH—http://dasish.eu
  This project brings together all 5 ESFRI research infrastructure initiatives in the social sciences and humanities (SSH)
- CRISP—http://www.crisp-fp7.eu/about-crisp/
  Partnership which builds collaborations and creates long-term synergies between research infrastructures on the ESFRI (European Strategy Forum on Research Infrastructure) Roadmap in the field of physics, astronomy and analytical facilities.
- Biomedbridges—http://www.biomedbridges.eu/
  Joint effort of ten biomedical sciences research infrastructures on the ESFRI roadmap. Together, the project partners will develop the shared e-infrastructure—the technical bridges—to allow interoperability between data and services in the biological, medical, translational and clinical domains.

iii. Open Educational Resources

- Open CourseWare Consortium—http://www.ocwconsortium.org
- OER Commons—http://www.oercommons.org/
- Brief history of OER—http://www.oecd.org/edu/ceri/36224377.pdf
- On funding models for OER:
  - http://www.ijklo.org/Volume3/IJKLOv3p029-044Downes.pdf , pp. 34-36
  - https://www.documentcloud.org/documents/400864-coursera-fully-executed-agreement.html#document/p40
  - http://chronicle.com/article/How-an-Upstart-Company-Might/133065/

iv. Open Infrastructure and standards

- "Competing Definitions of "Openness" on the GII", Jonathan Band, in *The Unpredictable Certainty: White Papers (1997)*, The National Academies Press, http://www.nap.edu/openbook.php?record_id=6062&page=31
- GEANT—http://www.geant.net/About_GEANT/pages/home.aspx
- ESFRI (Roadmap)—http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri
- EGI—http://www.egi.eu/about/
  The European Grid Initiative (EGI) provides the largest capacity computing grid for publicly funded research in collaboration with National Grid Initiatives (NGIs)
- DEISA—http://www.deisa.eu
  DEISA played a key role in setting up a pan-European high performance computing (HPC) infrastructure for supercomputing applications and a production-quality environment for EU scientists to exploit.
- PRACE—http://www.prace-ri.eu/About-PRACE-RI?lang=en
  The Partnership for Advanced Computing in Europe (PRACE) currently represents the top of the European HPC ecosystem, and provides Europe with cutting-edge world-class HPC systems
- D4Science—http://www.d4science.eu/about
  D4Science constitutes a continuation of the DILIGENT and D4Science projects. It will bring together several scientific e-Infrastructures established in areas such as biodiversity, fishery resources management and high energy physics.
- Science Data Infrastructure for Preservation—Earth Science—http://www.scidip-es.eu/
  Science Data Infrastructure for Preservation—Earth Science aims to deliver generic infrastructure services for science data preservation and to build on the experience of the ESA Earth Observation Long Term Data Preservation (LTDP) programme to favour the set-up of a European Framework for the long term preservation of Earth Science (ES) data through the definition of common preservation policies, the harmonization of metadata and semantics and the deployment of the generic infrastructure services in ES domain.

v. Open Source Software and Open Development

- Open Source Initiative—http://opensource.org/
- Open Source Definition—http://opensource.org/docs/definition.php
- Free Software Foundation—http://www.fsf.org/
- "Open Source Paradigm Shift", Tim O'Reilly, June 2004, http://tim.oreilly.com/lpt/a/4868
- Meritocrats, cluebats and the open development method: an interview with Justin Erenkrantz http://www.oss-watch.ac.uk/resources/erenkrantz.xml
- The Apache Software Foundation (ASF)—http://www.apache.org/
  ASF is a not-for-profit organization supporting open development projects. ASF-funders include for-profit companies, code contributors include people working for for-profit companies. ASF has built its organization around the open development method. Their distributed peer-review method and transparent development process ensure high standards. Having not just the source code open, but also the development method and decisions, has made it possible for newcomers to contribute to continuous development. This is an important aspect in open development, because it contributes to the sustainability of products created in temporary projects. Examples of this are Apache Wookie and TexGen, both software products that were originally made in the context of a research project. They were subsequently provided Open Source in an Open

Development community and have become successful in much wider circles, procuring continued funding from new sources and finding new, unexpected applications.

- Wookie: a case in sustainability, 2011- http://www.oss-watch.ac.uk/resources/cs-wookie.xml
- TexGen: a case study, 2009—http://www.oss-watch.ac.uk/resources/cs-texgen.xml

vi. Open Education

- MITx—http://web.mit.edu/newsoffice/2011/mitx-faq-1219.html
- edX—https://www.edx.org/
- Coursera—https://www.coursera.org/#about
- Udacity—http://www.udacity.com/udacity
- Open Learning Initiative—http://oli.cmu.edu/
- The Capetown Open Education Declaration—http://www.capetowndeclaration.org/
  The Cape Town Open Education Declaration arises from a meeting convened in Cape Town in September 2007 by the Open Society Institute and the Shuttleworth Foundation. The aim of this meeting was to accelerate efforts to promote open resources, technology and teaching practices in education. The first concrete outcome of this meeting is the Cape Town Open Education Declaration: a statement of principle, strategy and commitment. It is meant to spark dialogue, to inspire action and to help the open education movement grow. The Declaration has already been signed by hundreds of learners, educators, trainers, authors, schools, colleges, universities, publishers, unions, professional societies, policymakers, governments, foundations and other kindred open education initiatives around the world.
- OER university—http://wikieducator.org/OER_university/Home

vii. Open Peer Review

- Scholars Test Web Alternative to Peer Review", The New York Times, August 23, 2010, http://www.nytimes.com/2010/08/24/arts/24peer.html?pagewanted=all
- "Opening up BMJ peer review; a beginning that should lead to complete transparency", Richard Smith, *BMJ* 1999 January 2; 318(7175): 4–5, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1114535/

viii. Open Research

- What, exactly, is Open Science, Dan Gezelter—http://www.openscience.org/blog/?p=269, 2009 and http://www.openscience.org/blog/?page_id=44
- Open Source Drug Discovery Network—http://www.osdd.net/
- Galaxy Zoo—http://www.galaxyzoo.org/
- Foldit—http://fold.it/portal/
- Polymath project—http://en.wikipedia.org/wiki/Polymath_Project

ix. Open Innovation

- Open Innovation: Researching a New Paradigm, Chesbrough et al., http://www.openinnovation.net/Book/NewParadigm/Chapters/01.pdf,
- Open Source and Open Innovation, http://www.oss-watch.ac.uk/resources/openinnov.xml, 2010
- Google Summer of Code—http://code.google.com/soc/

C Survey of sustainability models

Model	Used for	Organisation
Firm contributions	Linux	Linux Foundation[161]
Commercial contributions (300+)	Eclipse	Eclipse Foundation[162]
Contributions of capital and effort and firms	Kuali Financials and Student	Kuali Foundation[163]
Supporting firms and contributions of effort	uPortal	JA-SIG[164]
Income - percentage of partner revenue; donations	Moodle	Moodle Foundation[165]
Membership fees and contributed effort (and capital)	Sakai CLE PeerJ/PeerJ Preprints	Sakai Foundation[166] PeerJ[167]
Contributions of effort and donations	Apache projects	Apache Software Foundation[168]
Capital and income from service	GEANTnetwork	Dante[169]
(Venture) capital, subscriptions and contributions of effort	Alfresco	Alfresco[170]
Memberships, supporting partners (firms), sponsorships, donations and contributions of effort	Drupal	Drupal Association[171]

Most examples mentioned are Open Source Software products. PeerJ/PeerJ Preprints is a new academic Open Access journal initiative. GEANT is the collaborative network.

In many of the OSS cases, there is a free version in addition to paid-for versions, dependent on the level of sophistication of the product or services. This model ties in with the 'data-services-model' presented in Annex D.

In all cases, sustainability also depends on the efforts of the community of users/contributors around the product or service (contributions of code; contributions of preprints and peer review services in the case of PeerJ).

[161] http://www.linuxfoundation.org/about
[162] http://www.eclipse.org/org/
[163] http://kuali.org/
[164] http://www.jasig.org/
[165] http://moodle.com/partners/about/
[166] http://www.sakaiproject.org/sakai-foundation
[167] https://peerj.com/
[168] http://www.apache.org/foundation/
[169] http://www.dante.net/
[170] http://www.alfresco.com/tour
[171] https://association.drupal.org/

D Data/Services model

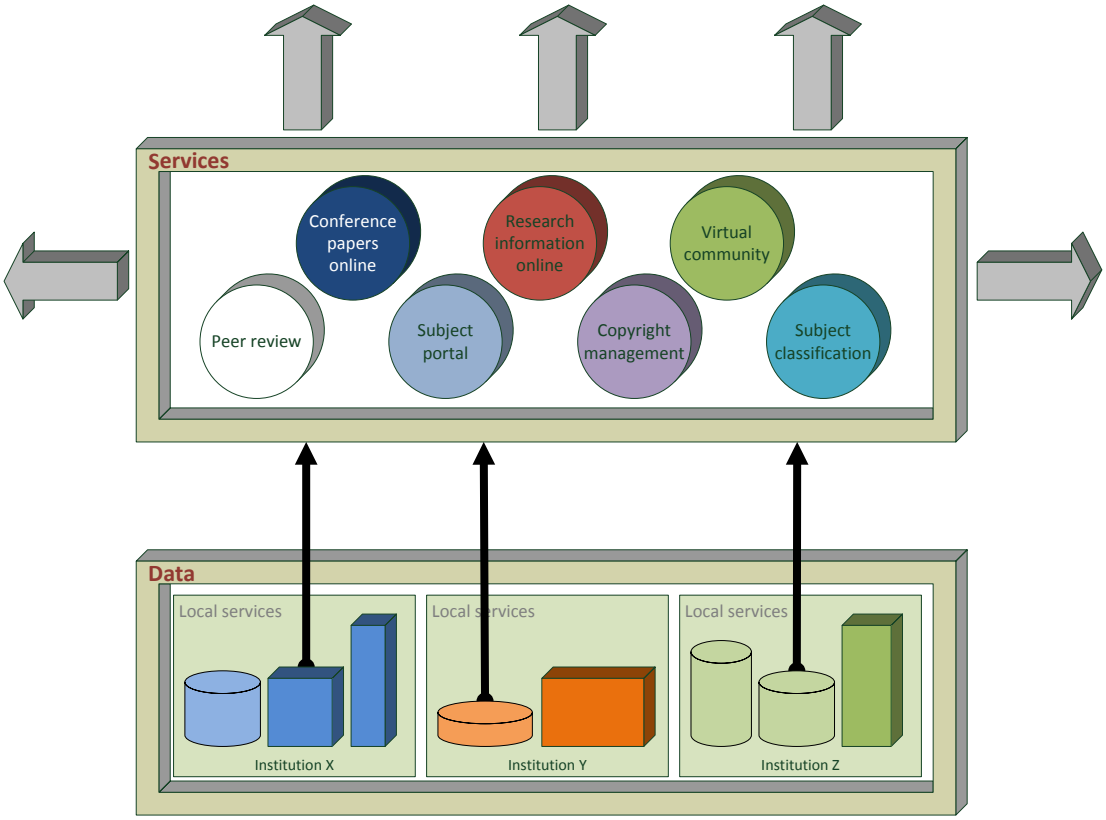

Figure 20—Data/Services model

Shown above is the data and services model[172] that underpins repository infrastructure. It comprises: a) a basic facility (data level) and b) 'services' (services level). These services can be developed using resources from the basic facility and provide certain added value for specific groups of users.

- The data level is where the infrastructure is set up and maintained. Research institutions establish repositories that store research information from that institution and keep it available for use (or reuse) according to uniform international standards: working papers/preprints, dissertations, research reports, datasets, conference reports, multimedia material etc., including the corresponding metadata.
- Services level: basic material from the data level can be used to develop services providing added value for scientists, students, universities, funders and other interested parties. The possibilities are numerous, including current or new services[173] (management information or updating of resumes for example); services the institutions themselves wish to offer, either individually or in cooperation with others (subject portals); or services provided by third parties such as publishers (e-journals). Supply and development can take place at the local level (digital display) as well as at national level (national research information service) or internationally (virtual communities). The

[172] This model was developed in the context of the DARE-programme, coordinated by SURF Foundation (www.surf.nl/en/DARE), but is useable as a more generic model as well; the original text as published by SURF has been slightly modified to reflect this generic applicability, with grateful acknowledgement to the original source.
[173] The services referred to in the diagram and between brackets in the text are intended as examples of the possibilities

most appropriate level of cooperation can be decided upon depending on the situation in question: most important is that institutions retain control of their own information.

The advantage offered by the data-services model is that it provides a guide to sound consideration and decision-making about the costs and expenses involved in the level and extent of services that an institution wishes to offer. The data level offers every institution a basic facility for the reliable, structured digital storage of its own research outputs.

If the data level is set up as simply as possible, the basic facility can be offered free of unnecessary extras and can be guaranteed long-term.

All additions to functionality (and thus work and manpower) aimed at activities above and beyond the basic facility belong at the services level. This makes it possible to ascertain what the corresponding costs are for each service, for whom such services are intended, whether they are worth the effort, and how best to finance them.

note

note

e-InfraNet

Consortium

IRELAND	Higher Education Authority (HEA)
UNITED KINGDOM	Higher Education Funding Council for England HEFCE (JISC)
NETHERLANDS	SURFfoundation (SURF)
PORTUGAL	Foundation for Science and Technology (FCT)
LATVIA	Latvian Academy of Sciences (LAS)
SPAIN	Ministry of Economy and Competitiveness (MINECO)
FINLAND	CSC - IT Center for Science (CSC)
HUNGARY	National Information Infrastructure Development Institute (NIIFI)
TURKEY	The Scientific and Technological Research Council of Turkey (TUBITAK)
ISRAEL	The Israel-Europe R&D Directorate (ISERD)
GREECE	The General Secretariat for Research and Technology (GSRT)
BELGIUM	Department of Economy, Science and Innovation, Flemish Government (EWI)

Visit the e-InfraNet website:

www.e-infranet.eu

Enhanced Participation:

The current partnership is keen to attract further participation.

Those interested in participating in the project's activities should contact the co-ordinator: Pat O'Connor

e-mail: poconnor@hea.ie