

# Toolkit for Researchers on Legal issues



JAN. 2019

Advancing Open Scholarship

D3.2 – Toolkit for Researchers on Legal issues

Version 1.0 – Final

PUBLIC

A set of support materials that are targeted to research organisations, researchers, research groups on legal issues for the re-use of research data, in particular privacy, copyright and GDPR issues.



H2020-EINFRA-2017  
Grant Agreement 777541

# Document Description

## D3.2 – Toolkit for researchers on legal issues

WP3 - NOADs Network Operation	
WP participating organizations: UGOE, NKUA, CNR, UNIWARSAW, ATHENA RC, UMINHO, EIFL, UGent, CERES, KNAW, UEDIN, UG, NHRF (EIE), UNIVIE, IMI BAS, RBI, UCY, SDU, UTARTU, UH, COUPERIN, UKON, DE, LUH, TCD, BIU, LU, KTU, ULUX, UOM, TU Delft, UEFISCDI, UB, CVTISR, UL, FECYT, KBSE, UZH, IIT, JISC, MU	
Contractual Delivery Date: 10/2018	Actual Delivery Date: 01/2019
Nature: Report	Version: 1.0 Final
Public Deliverable	

### Preparation Slip

	Name	Organisation	Date
<b>From</b>	Thomas Margoni Prodromos Tsiavos	UG UOA	28/11/2018
<b>Edited by</b>	Najla Rettberg Daniel Bangert	UGOE	07/12/2018
<b>Reviewed by</b>	Pedro Principe Natalia Manola	UMINHO ARC	10/12/2018 30/12/2018
<b>Approved by</b>	Pedro Principe	UMINHO	30/12/2018
<b>For delivery</b>	Mike Chatzopoulos	UoA	03/01/2019

## Table of Contents

<b>1  GUIDES ON COPYRIGHT AND RELATED RIGHTS</b>	<b>6</b>
1.1 GUIDE 1 – USER GUIDE ON COPYRIGHT, OPEN SCIENCE AND DATA	6
1.2 USER GUIDE 2 – APPLYING THE RIGHT LICENSE TO YOUR REPOSITORY	15
3. CONTENT SHOULD ALSO BE LICENSED	21
4. DATA AND DATASETS AND DATABASES SHOULD BE UNDER CCO	24
5. OTHER WORKS OF AUTHORSHIP (ARTICLES, IMAGES ETC.) SHOULD BE LICENSED UNDER A CC-BY 4.0 LICENSE	25
6. REPOSITORIES SHOULD RECOMMEND THE BEST OS LICENSES BUT IT SHOULD BE THE UPLOADER WHO CHOOSES WHICH ONE	26
<b>2  RECOMMENDATIONS REGARDING OPENAIRE SERVICES IN THE CONTEXT OF PERSONAL DATA AND IPR</b>	<b>30</b>
2.1 INTRODUCTION	30
2.2 PERSONAL DATA PROTECTION AND ETHICS	30
2.3 IPR	35
<b>3  REFERENCES</b>	<b>41</b>

## Disclaimer

---

This document contains description of the OpenAIRE-Advance project findings, work and products. Certain parts of it might be under partner Intellectual Property Right (IPR) rules so, prior to using its content please contact the consortium head for approval.

In case you believe that this document harms in any way IPR held by you as a person or as a representative of an entity, please do notify us immediately.

The authors of this document have taken any available measure in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

This publication has been produced with the assistance of the European Union. The content of this publication is the sole responsibility of the OpenAIRE-Advance consortium and can in no way be taken to reflect the views of the European Union.

OpenAIRE-Advance is a project funded by the European Union  
(Grant Agreement No 777541).



# Publishable Summary

---

The OpenAIRE-Advance project runs a series of task forces in three topical areas relevant to open science: Research Data Management, Open Science Policies and Legal Issues for Research Data. The aim is to exploit the expertise in the network (primarily the National Open Access Desks, NOADs) and together to produce a series of materials for public use.

In the context of this task force, legal issues, the team identified a number of pertinent areas to develop materials useful to research communities.

To this end, this deliverable focuses on the emerging field of research data from a legal perspective. It looks at the proper legal and technological classifications and taxonomies for data (data may have very different meaning in hard and social sciences and in legal terms), their status, protection, reusability, licences, interoperability and more in general to any aspect that may make data more or better fit to meet open science goals.

This deliverable is divided into two main sections:

- Section 1 Guides on copyright and copyright related aspects (licenses, reusability, etc.),
- Section 2 Recommendations on data protection and privacy.

## **Section 1**

The first part of this deliverable focuses on the copyright and related rights dimension of open science in the field of data, datasets and databases. The field of data, especially research data, is quite complex and the law does not necessarily call works (copyright) and data (not protected or protected by related rights) following the same categories that hard or social sciences do. Therefore, the field of open access/science to data and research data is particularly complex for users.

The two guides that form this first section, a user FAQs and a companion Open Science (OS) checklist that users can employ to direct repository managers at university and research centres to help them realise their OS potential, are meant to offer a state of the art, legally advanced, but still manageable set of rules, guidelines, and resources to enable the full potential of OS in the EU research field with a view to addressing copyright and related rights issues. The user guide and the companion checklist are intended to be online documents and have been (companion checklist) or will be (user FAQs) implemented on the OpenAIRE website. Given the nature of the documents, they will be regularly updated and revised in the light of new developments in the legal, policy, community and technological fields.

## **Section 2**

The second part of this deliverable provides guidance in the fields of Intellectual Property Rights (IPR) and Personal Data protection in the form of a checklist and a set of recommendations regarding what different players in the open science ecosystem could be doing in order to support and promote open science. The perspective adopted in this paper is one that focuses on the perspective of the researcher but also takes into consideration the types of activities necessary in order to foster Open Science in a broader context. In that sense, it takes forward and expands the

content of section 1: it focuses on suggestions that aim at the institutional level and set the necessary conditions for opening up data of different forms and kinds while remaining compliant with the relevant legislation. Therefore, it uses two tools: (a) a checklist of things to do in the area of Personal Data protection focusing on the compliance with the new General Data Protection Regulation (GDPR), mostly at the institutional level and with a research focus; and (b) a checklist of recommendations regarding the establishment of open IPR policies by different stakeholders in the Open Science ecosystem.

In relation to personal data protection we analyse the key necessary steps for compliance in accordance to the GDPR and with emphasis on the measures that a Research Performing Organisation should take in order to adhere to the provisions of GDPR. Our interest here is to maintain compliance while maximising the opening up of data.

In relation to Intellectual Property Rights, we take a holistic view focusing not just on copyright but on the broader interrelationship between different kinds of rights, such as patents, trade secrets and copyright. Such interrelationship takes particular interest in the establishment of coherent policies and monitors the moment of information disclosure in order to ensure that open science practices do not prejudice opportunities of exploitation. In order to achieve these objectives, and while the researcher always remains the focus, it is necessary to provide recommendations to a wider array of institutions that cover the entirety of Open Science ecosystem.

## 1 | GUIDES ON COPYRIGHT AND RELATED RIGHTS

The first part of this deliverable gives comprehensive support via a series of FAQs on the copyright and related rights dimension of open science in the field of data, datasets and databases. The field of data, especially research data, is quite complex and the law does not necessarily call works (copyright) and data (not protected or protected by related rights) following the same categories that hard or social sciences do. Therefore, the field of open access/science to data and research data is particularly complex for users.

The two guides here contained, a user FAQs and a companion Open Science (OS) checklist that users can employ to direct repository managers at university and research centres to help them realise their OS potential, are meant to offer a state of the art, legally advanced, but still manageable set of rules, guidelines, and resources to enable the full potential of OS in the EU research field with a view to addressing copyright and related rights issues.

The user guide and the companion checklist are intended to be online documents and have been or will be implemented on the OpenAIRE website. Given the nature of the documents, they will be regularly updated and revised in the light of new developments in the legal, policy, community and technological fields.

Future work will focus on the above-mentioned maintenance and updating, as well as on the enrichment of the two guides with new sections that will emerge from the continuous dialogue with NOADs, users, repositories and the community at large. Examples of possible future work may be in the field of ownership (who owns the data), repository liability (are academic repositories information society services that enjoy safe harbour exemptions), and compatibility of licences when derivative works are created.

Guides will be made publicly available via the OpenAIRE portal support information:  
[www.openaire.eu/guides](http://www.openaire.eu/guides).

### 1.1 Guide 1 – User guide on copyright, open science and data

#### 1. How do I know if my research data is protected?

---

##### 1.1 What is Research Data?

In this information guide, the definition of research data is the one used in the Concordat on Open Research Data. Research data are the evidence that underpins the answer to the research question, and can be used to validate findings regardless of its form (e.g. print, digital, or physical). These might be quantitative information or qualitative statements collected by researchers in the course of their work by experimentation, observation, modelling, interview or other methods, or information derived from existing evidence. Data may be raw or primary (e.g. direct from

measurement or collection) or derived from primary data for subsequent analysis or interpretation (e.g. cleaned up or as an extract from a larger data set), or derived from existing sources where the rights may be held by others. Data may be defined as ‘relational’ or ‘functional’ components of research, thus signalling that their identification and value lies in whether and how researchers use them as evidence for claims. They may include, for example, statistics, collections of digital images, sound recordings, transcripts of interviews, survey data and fieldwork observations with appropriate annotations, an interpretation, an artwork, archives, found objects, published texts or a manuscript.

## 1.2 When is research data protected?

Research data are likely to be a ‘bundle’ of different types of information and content, sourced from third parties, or created by the researcher. Research data may be factual and/or creative. Data as such, like facts, principles, mathematical concepts and methods are not protected by copyright. However, there are cases in which data, not as such but part of collections, can be protected. Additionally, whereas data as such are not protected by copyright, that does not mean data are not protected by other laws (e.g. confidential information or personal data). This guide will offer guidance in this area with a specific focus on research data.

This information guide will focus on those elements of research data that, might be automatically protected by intellectual property rights such as copyright protected works and databases protected by copyright and/or the Sui Generis Database Right (SGDR)

## 1.3 How do rules on research data impact on use?

Research data may be:

- Automatically protected by the law;
- Regulated by contract;
- Subject to community norms such as academic best practices.

Multiple types of protection might exist in research data, or there may be elements that have no legal protection. As explained in this information guide, the default position is that the owners of protected rights can restrict the use of those rights. Licensing is a way to stipulate when and how protected rights in research data can be used by others. Before thinking about licensing or using research data, it is important to understand the rights that can arise automatically in research data. The rest of this guide will discuss copyright and the database right. Sui Generis Database Right – SGDR is a right protecting databases where there has been a substantial investment in obtaining, verifying or presenting the data.

Copyright is a property right in certain types of original literary, artistic and scientific works. Copyright does not protect ideas. Confidentiality protects confidential information. This might be imposed by a contract or if the information is marked confidential. Use of confidential information might give rise to a claim for compensation if confidentiality is breached. Data Subject Rights arise in information that identifies individuals and are recognised by data protection laws in the EU. Patents are registered rights in novel inventions of products or processes. Patent protection is not

specifically discussed in this FAQs. Some research data may not benefit from any legal protection, although moral and ethical considerations may apply.

#### **1.4 What is the Sui Generis Database Right (SGDR) and how does it relate to other rights in Databases?**

For legal purposes, a ‘database’ means “a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means.”

Databases are automatically protected by law if:

1. the selection or arrangement of the contents are the author’s own intellectual creation – in which case copyright protection applies to the structure of the database (not to its content); or
2. they qualify for a special IP right (called the sui generis database right (SGDR)) because there has been a substantial investment in obtaining, verifying or presenting the contents of a database (e.g. a database of poetry titles). Importantly, the investment in creating data does not count towards the threshold for SGDR protection. This may be quite confusing (especially because the difference between creating and obtaining data should not be read with the eyes of a scientists, i.e. it is not an epistemological issue, but with the eyes of competition law, i.e. the goal is to limit as much as possible so called single source databases). At present time, the best we know about how to distinguish creation from obtaining is in the field of sporting events: match fixtures are created data, therefore sports organisers do not have a SGDR in that. Live sporting event data (e.g. who scores at what minutes) are obtained data and therefore can be protected by SGDR. The SGDR has been heavily criticised especially from an innovation and scientific point of view, but as for now it is part of EU law and the above example is as accurate as it can be right now.

The content of a database can also be composed of copyright protected works in the first place, such as a database of scholarly articles. However, in this case it is not the copyright in the database that extends to the content, there are simply two independent copyrights in the same database: the copyright in the database structure if it is sufficiently original, and the (independent) copyright in the element constituting the database. In addition, there may also be a SGDR right in the substantial investment in the database. The latter protects the database from substantial extractions, therefore effectively offering a form of protection to data that would otherwise not be protected.

In conclusion, a database may be protected by up to 3 different rights that regulate 3 different uses. They may belong to 3 different authors/owners and follow 3 different rules in terms of exceptions and limitations. Easy isn’t it?

#### **1.5 Who owns SGDR?**

SGDR is owned by the maker of the database, which is usually the person or entity who bears the financial risk. Financial risk in academic research. In academic research the financial risk of projects is usually borne by the institution. It is therefore possible that the SGDR in a database produced during research may be owned by the institution, even where copyright arising in the database may be owned by the researcher. However, contracts among the different parties (funding grants, universities IP policies, researchers employment contracts, etc.) may provide differently.

It is important to check the terms of any funding or collaboration agreement to see whether ownership of the deliverables is specified.

### **1.6 How long does the SGDR last?**

15 years, renewable if a new substantial investment is made.

### **1.7 What can SGDR owners do?**

The SGDR gives the right holder the right to perform or to authorise the extraction (copying) of substantial amount of data, or repeated extraction of insubstantial amount, and their reuse therefore effectively extending protection to the data.

### **1.8 What is Copyright law?**

Original literary, musical, dramatic and artistic works may be protected by copyright law. Certain non-original works such as the typographical arrangement of published editions, sound recordings, broadcasts, non-original photographs, the first fixation of films, etc. are also protected by so called rights related to copyright (or neighbouring rights). Copyright is automatic, there no need to register, deposit, apply a © symbol, etc., although all these activities are permitted and may be even helpful. However, they cannot be a condition for copyright protection (this rule applies almost everywhere in the world, as it is enshrined in the Berne Convention, the most relevant international copyright conventions). Copyright law does not protect ideas that are not expressed in an original form of expression, and does not protect data as such, principles, facts, etc. Copyright law grants the owner the exclusive right to control certain rights such as reproduction, redistribution, communication to the public etc. For example, when you write a research paper you are automatically the author of that paper. If you write it with others you all become co-authors (as long as everyone made an “authorial” contribution: proofreading, commenting, etc. are usually not considered authorial contributions).

### **1.9 Who owns copyright?**

The author is usually the person creating the work, but if the work is created jointly, it may be jointly owned, and if a work is created by someone in the course of a contract of employment it may be owned by the employer (depending on the contract).

### **1.10 How long does copyright last?**

The default rule in the EU is the life of the author plus 70 years. Related rights (e.g. sound recording, film fixations, non-original photographs, etc.) all have different periods, which are more or less harmonised in the EU.

### 1.11 What can copyright owners do?

Copyright owners have rights to restrict certain acts in relation to the protected work such as copying, distributing and communicating the work to the public. The effect is that in most instances use of copyright protected works requires permission from the owner.

Moral rights are additional rights in copyright works. They are not harmonised at the EU level therefore each EU Member State can have different rules on moral rights. However, the Berne Convention stipulates some minimum standards that must be met. These are:

- The right to be attributed as the author
- The right to protect the integrity of the work (i.e. to object to alteration or derogatory treatment of the work that are prejudicial to the honour or reputation).

Moral rights cannot be transferred, but in certain jurisdictions can be waived. They must last at least as long as the economic rights (in the EU 70 + author's life), but can last longer and in certain jurisdictions they are perpetual.

### 1.12 How can copyright protected works be used?

In the EU, there is no concept of 'fair use' of works protected by copyright (or SGDR).

There are however exceptions and limitations that play a similar role, however they do not reach the same flexibility or width as fair use (which is a US concept). Art. 5 of the Copyright in the Information Society Directive of 2001 has a list of about 21 exceptions and limitation to copyright. However, whereas this is a closed list (MS cannot create new exceptions) it is not mandatory (except for the case of temporary acts of reproduction), meaning that each MS can choose which ones to implement. This aspect has attracted strong criticisms as it is not conducive to a level playing field in the EU scientific landscape. For a list of the exceptions and limitations implemented in your country you should refer to your NOAD, which may maintain an updated list.

In absence of an exception or limitation, uses covered by copyright are reserved to their authors. You may however still be able to reuse a work if it is licensed to you or it is generally licensed to the public with a public licence. Creative Commons Licences or FLOSS licences are examples of this type of public (and usually conditional) permissions.

Depending on the type of licence, free use of the research data may be permitted, allowing for use/re-use or re-mixing. Other research data may be subject to certain conditions such as attribution of the original research data owner, or attribution of the owners of other rights in the research data. More restrictive licences may prevent using the research data for commercial purposes or not altering the research data.

If the research data does not specify a licence, it is necessary to consider whether use of any protected content within the research data would infringe the rights of anyone

### **1.13 Are there any automatic legal rights in the dataset?**

Copyright and the sui generis database right (SGDR) arise automatically in qualifying works in the EU. If a dataset contains any protected works, the default position is that these rights are owned by the author or maker.

### **1.14 What if the researcher is an employee of a research institution or a student?**

Depending on the employment contract and the institution's intellectual property policies, the law may recognise an employee's copyright as being owned by the employer.

### **1.15 What if there is a contractual relationship between the researcher and any other party (e.g. funder)?**

A collaboration or funding agreement may specify that the ownership of the deliverable research lies with the funder, or that the funder has an automatic licence in the research.

### **1.16 What if any of the data has been obtained from a third party who imposed terms and conditions?**

For example, datasets collated through social media platforms will be subject to that platform's terms and conditions which may specify how the data can be used or made available.

### **1.17 Who took the financial risk for the research?**

While copyright is usually owned by the author of the work, the SGDR in a database is owned by the person or organisation that bore the financial risk in creating the database. This should be verified on a case by case basis, but it is commonly the employer or the unit that has budgetary autonomy.

## **2. How do I licence my research data?**

---

### **2.1 What licence should be applied to the research data?**

It depends on what rights protect your research data, if at all.

In the light of what is explained above:

if your research data qualifies as a work (literary work such as a journal article or a software), then CC BY 4.0 is usually the best choice. The use of the Share Alike (SA) is also compatible with the Open Access definition and reinforced in Plan S licensing guidance for publications. Non-commercial should be avoided as it is not Open Access compliant. Non-derivative is a tricky issue and should be avoided, especially if you do not know what you are doing. That said, it may not be incompatible with the Open Access definition.

If your research data is a database or a dataset (unstructured data that do not meet the database definition) usually the best option is a CC0, which waives all your rights in the database.

Keep in mind that CC licences only deal with copyright and copyright related matter. Personal data are not included in CC and are analysed separately.

## 2.2 What is a Creative Commons licence?

Creative Commons, a global not-for-profit organisation which provides legal tools to promote the sharing and reuse of works of authorship, has produced a number of licences some of which meet the criteria for Open Access. These offer different levels of permission.

Creative Commons offers licences readable at three different levels: legal, machine (the metadata) and human (non-legal descriptions). Creative Commons has a useful tool to help you determine the licence best for you. More restrictive CC licences are unlikely to meet Open Access requirements (e.g. because they impose restrictions on commercial use).

## 2.3 How are licences applied to research data?

Licences are not automatic. The owner of rights protected data set must make it clear that a licence is applied. Repositories may help you to select the licence applied to data deposited in their repository. Applying a licence can happen by:

- Choosing a license when uploading your data in a repository
- Referring to the licence on the landing page or host site for a digital research data;
- Attaching the licence to the metadata of the research data;
- Setting up a Read Me file for the data.

If a standard licence from Creative Commons, they will have tools to help attach the licence effectively. See for more info the accompanying OS repository checklist for an explanation on how to use those tools.

## 2.4 I'm really concerned with attribution. How can I make sure others cite me as the source for my research?

Attribution is a genuine concern. To help others cite your research, include a citation in your research that users can copy and paste to give you credit for your hard work. If you licence your data under a CC BY you are legally requiring attribution, but we recommend that you do this only if you are authoring a work such as a journal article or a photograph or a song. If you are producing protected databases (as explained above) probably your best choice is to use CC0. You can still ask for attribution, not as a legal requirement but as “please attribute my data” in line with scientific norms.

## 2.5 But I would like attribution when others use my dataset. In that case, shouldn't I use a CC BY licence?

We recommend that you avoid using a CC BY licence for data. While attribution is a genuine, recognisable concern, not only might using a CC BY licence be legally unenforceable when no underlying copyright or SGDR protects the work, but it may also communicate the wrong message

to the world, as you are requiring attribution for something that the law says there is no attribution (e.g. SGDR does not require moral rights).

A better solution is to use CC0 and simply ask for credit (rather than require attribution), and provide a citation for the dataset that others can copy and paste with ease. Such requests are consistent with scholarly norms for citing source materials.

### **2.6 I'm uncomfortable with others using my research for commercial purposes. Should I use a non-commercial licence for my dataset?**

We recommend you avoid using a non-commercial licence. For legal purposes, drawing a line between what is and is not 'commercial' can be tricky; it's not as black and white as you might think. For example, if you release a dataset under a non-commercial licence, it would clearly prohibit an organisation from selling your dataset to others for a profit. However, it might also prohibit someone using the dataset in their research if they intend to eventually publish that research. This is because most academic journals are commercial businesses that charge some sort of fee for access to their content, hence, such use could qualify as 'commercial'. Consequently, using a non-commercial licence may prevent researchers from using your data in work destined for publication. This can subsequently affect the dissemination, recognition, and impact of your dataset. And it is definitively NOT open access. (see the Berlin Declaration, Bethesda Statement on Open Access Publishing, and Budapest Open Access Initiative).

### **2.7 I'm uncomfortable permitting use of my research for any and all purposes. Should I use a 'No Derivatives' (ND) licence for my dataset?**

We recommend you avoid using a 'No Derivatives' licence. Similar to how a non-commercial licence might restrict meaningful reuse of your dataset, a ND licence can have the same effect: it may prevent someone from recombining and reusing your data for new research. For data to be truly Open Access, it must permit these important types of reuse. It is less clear whether ND is OA compliant or not. The best

view is that it depends on what kind of modifications it prohibits, therefore, there are probably cases where ND is incompatible with OA, and thus you should not use it.

### **2.8 Is there any part of the research data that cannot be made available?**

Consider redacting research data to remove personal data, confidential information or third party intellectual property.

### **2.9 I want to CC licence my work, but I'm concerned because it contains copyright protected material made available by others that I cited or quoted. Will this affect their copyright?**

Your CC licence applies only to your original contributions and does not supersede any rights retained by authors whose works you have cited or have permission to use.

### **2.10 How should I licence my data for the purposes of Open Science?**

We recommend you use the CC0 Public Domain Dedication, which is first and foremost a waiver, but can act as a licence when a waiver is not possible. By applying CC0 to your data you enable everyone to freely reuse your data as they see fit by waiving (giving up) your copyright and related rights in that data.

### **2.11 How should I licence my work for the purposes of Open Access?**

CC BY 4.0.

If you work for an educational institution, it is good practice to first check with your research director and library. Your institution may already have an Open Access publishing policy for you to consult, and your library will be able to help you decide how to best proceed.

### **2.12 Is data always subject to copyright?**

You should keep in mind that there are many situations in which data is not protected as a matter of copyright and related laws. Such data can include facts, names, numbers – things that are considered ‘non-original’ and part of the public domain thus not subject to copyright protection. Similarly, your database (which is a structured collection of data) might be considered ‘non-original’ and thus ineligible for copyright, and it might additionally be excluded from other forms of protection (like the EU sui generis database right, also known as the ‘SGDR’, for non-original databases).

In these cases, using a Creative Commons licence such as a CC BY could signal to users that you claim a copyright in the non-original data despite the law, and perhaps despite your real intention. Finally, if your data is in the public domain worldwide, you might state simply and obviously on the material that no restrictions attach to the reuse of your data and apply a Public Domain Mark.

### **2.13 Does the researcher owe any obligations of confidentiality or ethics in respect of the data?**

Obligations of confidentiality may be imposed by contract or implication. Most researchers are expected to abide by ethical codes of conduct.

## **3 Can I reuse someone else’s research data?**

---

### **3.1 How can a protected dataset be used?**

The most common means for the owners of rights in datasets to grant permission is through a licence. A licence is a legal contract that specifies standard restrictions and permissions for a work, which can then be used by anyone the licence is granted to, although the ownership rights are not transferred. Again, keep in mind that licences usually regulate only the copyright and copyright related aspects. Personal data protection cannot be addressed through licences.

### **3.2 Where are licences found?**

Licences might have been applied to data in a number of ways: Embedded in the metadata for the data; Communicated by watermarks or notices within the data; Specified on the landing page for the dataset; Specified on a repository website; Detailed in a ReadMe file released with the dataset.

### 3.3 Interoperability and stacking

Using data under a licence can become more complex when that data may be combined with data from other sources. The result may be a dataset with different licensing restrictions. The ability of licences to interact with other licences is called interoperability. Not all licences can accommodate different layers of protected work. Creative Commons licences are better at this than some other licences, and bespoke licences can present particular interoperability issues.

When choosing a licence for protected research data, it is necessary to consider how this might interact across the whole research data (including non-protected elements) and how it might interact with future derivative works produced from re-use or re-purposing.

OpenMinTeD has a useful [matrix](#) presenting the compatibility of different licences. Using two or more licences may require stacking of attribution of rights in the licensed work.

### 3.4 What happens if I use ‘Share Alike’ (SA) licensed material in my work? Does that mean I have to make my work available under the same SA licence?

Not necessarily, but it depends on how you use the SA licensed content. A ‘Share Alike’ CC licence applies only to the content licensed as SA that you have used. It does not require you to also make your work available under a SA licence, so long as you have not combined the independent works into one new work (known as a ‘derivative’ work).

### 3.5 Can a dataset be used if there is no licence?

If the dataset is not protected by copyright or related rights (see above) yes. If they are protected, then in the absence of a licence the owner(s) of rights must be approached directly for permission. There are a limited number of exceptions that permit use without a licence or specific permission. Examples of exceptions for copyright protected works include: Research & Private study; Quotation, Criticism and review; Disability access; Public administration; Text and data mining; News reporting; Parody & Pastiche. As said, exceptions and limitations to copyright are not harmonised at the EU level, so you should verify which ones are available to you. Also, keep in mind that these exemptions are granted in limited circumstances and subject to certain conditions. As usual, specific advice is recommended before relying on an exception.

### 3.6 What are the risks of using a dataset without a licence?

Where a dataset has protected elements, intellectual property laws may permit the owner of the rights to take steps to preserve their exclusive rights to use, copy and make derivative works from the data. Ultimately, the rightsholder could use the courts to enforce these rights, by asking a judge to make an order to stop use of the dataset, to destroy copies or derivative works, or to pay compensation to the rightsholder. The time and cost of any legal action, and the risk of an adverse judgment, may outweigh the benefits of using the dataset.

## 1.2 User Guide 2 – Applying the right license to your repository

### **1.1. One of the best licences you can use for your repository is a CC BY 4.0 licence, specifying that “unless otherwise noted, this repository is under a CC BY 4.0 licence”.**

We recommend using a CC BY 4.0 licence as a repository licence for the following reasons:

- Creative Commons licences are internationally recognised, well-established, and both human-readable and machine-readable;
- CC BY 4.0 licenses meet the definition of “open access” as defined in the Budapest, Bethesda, and Berlin declarations on open access;
- CC BY 4.0 is one of the most compatible licences for interoperability purposes.

#### **LEGAL SOURCES**

The following declarations and statements provide definitions of Open Access:

- [Budapest Open Access Initiative](#) (2002). (last accessed: 3 July 2018)
- [Bethesda Statement on Open Access Publishing](#) (2003). (last accessed: 3 July 2018)
- [Berlin Declaration on Open Access](#) (2003). (last accessed: 3 July 2018)

The following sources from Creative Commons and OpenMinTeD provide details on compatibility between licences (both generally and for the purposes of text and data mining):

- Creative Commons (date unknown) Compatible Licenses. (last accessed: 3 July 2018)
- OpenMinTeD (date unknown) Licence Compatibility Matrix. (last accessed: 3 July 2018)

See the following source by Creative Commons UK for details on both how and why you should be interested in making your work open access:

- Creative Commons UK (2017) Frequently Asked Questions on Creative Commons & Open Access. Zenodo. (last accessed: 3 July 2018), see “How should I licence my work for the purposes of Open Access?”

#### **EXAMPLES**

Creative Commons provide guidance on how to select the most appropriate licence for your work (depending on your sharing preferences), as well as how to mark your work once your appropriate licence has been identified:

- Creative Commons (date unknown) [Licence Chooser](#). (last accessed: 12 September 2018)
- Creative Commons (date unknown) [Marking your work with a CC license](#). See section “Example: Website”

#### **1.1.1. If you follow point 1.1 it means that the licence applies to all “works” or other subject matter in the repository.**

This includes:

- The repository as a copyright protected database (in case it qualifies);

- The repository as a sui generis database right protected database (in case it qualifies);
- The elements composing the database which can be:
  - Not protected, such as a database of temperature measurements. In this case, as these data are not protected in themselves you don't need a licence. CC licences are written in a way that you only have to accept them if you need permission to use something.
  - Protected (e.g. a database of journal articles)

For University repositories, it is likely that several of these elements co-exist, but it could also well be that the repository is not a protected database. In either case CC licences are a good choice because (avoiding technicalities) they only regulate the use if that use requires a permission. Under this point of view, it could be said that CC licences are self-contained to when permission is necessary.

### LEGAL SOURCES

See the following source for further details on the EU legislation regulating the legal protection of databases:

- European Commission (date unknown) Protection of Databases. (last accessed: 3 September 2018)

For confirmation of same see:

- Margoni, T., Caso, R., Ducato, R., Guarda, P. and Moscon, V. (2016) Open Access, Open Science, Open Society. In Loizides, F. and Schmidt, B., (eds) (2016) Positioning and Power in Academic Publishing: Players, Agents and Agendas. IOS. (last accessed: 1 October 2018) at p.82

**1.1.2. However, this could become problematic, when, as in the case of University repositories, the owner of the repository (the University) and the owner of the journal article (the author unless they transferred the copyright) are different people.**

Therefore, by using the recommended “unless otherwise noted” wording, you clarify that the elements that belong to third parties (e.g. journal articles) are distributed under their own licence terms (which as you will see later, ideally is a CC0 or a CC BY).

It is important to licence the repository as a database under an open access compliant licence. This is because when a user uses aggregated data (such as in data analytics, text and data mining, etc.) in order to crawl, scrape or analyse the database, authorisation (e.g. a licence or an exception if it exists) is often necessary. But if you have applied a CC BY to your repository this is already taken care of!

### LEGAL SOURCES

In order to meet the definition of open access as provided in the Budapest Declaration, users must be able to crawl the database:

- Budapest Open Access Initiative (2002). (last accessed: 3 July 2018)

See also recommendations made to data and e-infrastructure providers in the source below which confirm CC BY 4.0 is the most appropriate licence for data access:

- Guibault, L., Margoni, T. and Spindler, G. (2013) Conclusions and Recommendations in Guibault, L. and Wiebe, A. (eds) (2013) Safe to be open: Study on the protection of research data and recommendations for access and usage. OpenAIRE. (last accessed: 1 October 2018) at p.163/section 5.3

#### EXAMPLES

OpenAIRE provides a tool which tests online repository compliance with Open Science guidelines: [Validator service](#).

### **1.2. The CC BY 4.0 licence should be incorporated into the terms of service of the repository.**

#### LEGAL SOURCE

Terms of service are general rules about how a service, such as a website, can be used. These may include a multitude of conditions, such as privacy policies, limitations of liability, and codes of conduct. All users of the service have to agree to the terms of service.

#### EXAMPLES

Creative Commons provide guidance on how to integrate their licences within your terms of service:

- Creative Commons (2018) [Creative Commons Integration, from A to Z](#). (last accessed: 3 July 2018), see “Checklist for an ideal TOS”

#### **1.2.1. The CC BY 4.0 licence exists as a separate legal document from the terms of service.**

As such, it must be incorporated by reference into the contractual, and broader terms of service which govern all uses of the repository. Creative Commons provide guidance on how to incorporate the CC BY 4.0 licence into the repository terms of use.

#### EXAMPLES

Creative Commons provide guidance on how to integrate their licences within your terms of service:

- Creative Commons (2018) [Creative Commons Integration, from A to Z](#). (last accessed: 3 July 2018), see “Checklist for an ideal TOS”

## **2 Don't forget the metadata**

---

## 2.1. You should provide metadata in order to enhance discoverability of your resources.

### EXAMPLES

See the following sources for a discussion on the merits of using metadata in your repository:

- Kemp, J., Dean, C., and Chodacki, J. (2018) [Can Richer Metadata Rescue Research?](#) *The Serials Librarian*, 74: 1-4, 207-2011
- Knoth, Petr (2013). [From open access metadata to open access content: two principles for increased visibility of open access content.](#) *Open Repositories 2013*, 8-12 Jul 2013, Charlottetown, Prince Edward Island, Canada.

Creative Commons provide technical guidance on how to implement metadata via HTML, as well as providing a generation tool for embedding metadata within files:

- Creative Commons (2014) *Marking Works Technical*. (last accessed: 3 July 2018)
- Creative Commons (2015) *XMP*. (last accessed: 4 July 2018)
- Creative Commons (2018) *Creative Commons Integration, from A to Z*. (last accessed: 3 July 2018), see “License metadata”

Application of a CC0 licence for metadata is increasingly recognised as a community standard in the following institutions:

- DPLA (date unknown) *Elements of the DPLA*. (last accessed: 3 July 2018), see “Metadata”
- Europeana Collections (date unknown) *Usage Guidelines for Metadata*. (last accessed: 4 July 2018)
- World Bank Open Knowledge Repository (date unknown) *About The World Bank Open Knowledge Repository*. (last accessed: 12 September 2018)

### 2.1.1. Providing machine-readable bibliographic metadata is a requirement for projects which are funded under Horizon 2020.

#### LEGAL SOURCE

See H2020 Framework Programme Regulation (EU) No 1291/2013 of the European Parliament and of the Council of 11 December 2013 establishing Horizon 2020 - The Framework Programme for Research and Innovation (2014 - 2020) (OJ 347, 20.12.2013, p. 104) for legal basis, and more specifically article 29.2 of the H2020 Programme: AGA - Annotated Model Grant Agreement (2018). (last accessed 1 October 2018)

#### EXAMPLES

The requirement for machine-readable bibliographic metadata is detailed by the European Commission in the following sources:

- European Commission (date unknown) [“Open Access”](#). (last accessed 13 July 2018), see “Step 2 - Providing open access to publications”

Any projects funded by Horizon 2020 must have machine-readable bibliographic metadata, as detailed in:

- European Commission (2017) [H2020 Programme: Guidelines to the Rules of Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020](#).

## **2.2. Metadata often are not protected as such because they are factual information, thus not original or not substantial.**

However, in certain cases, complex and elaborate metadata could perhaps be protected. For the avoidance of doubt, apply a CC0 to your metadata. In this way, in those cases when a right exists, you are waiving it and allowing other people to reuse the metadata information.

CC BY should be avoided unless you know exactly what it entails.

### **LEGAL SOURCES**

Applying a CC BY licence may result in “copyfraud” in countries where metadata is not eligible for copyright protection (as in this case the application of a CC BY licence imposes more restrictive conditions than what the metadata is actually entitled to). Confirmed in:

- Kreutzer, T. (date unknown) [Validity of the Creative Commons Zero 1.0 Universal Public Domain Dedication and its usability for bibliographic metadata from the perspective of German Copyright Law](#). (last accessed: 5 July 2018) at p10

### **EXAMPLES**

Currently, 61% of the open access academic repositories listed on OpenDOAR have no clear metadata policy, as detailed in:

- Guadamuz, A. and Cabell, D. (2014) [Data Mining in UK Higher Education Institutions: Law and Policy](#). Queen Mary Intellectual Property Review 4:1 pp. 3-29

### **2.2.1. In those few cases when metadata can be considered original works, thus protected by copyright, they will enjoy both economic and moral rights.**

Moral rights are recognised in most countries (but with important exceptions, such as the US), and may be unwaivable.

This should not represent an issue in the case of CC0, as the waiver clarifies that it only waives the rights as long as this is permitted under applicable law. So if you enjoy unwaivable moral rights, CC0 will not affect your moral rights.

### **LEGAL SOURCES**

For an example of a jurisdiction with unwaivable moral rights, see Kreutzer’s discussion of the German position, and why the application of a CC0 licence in this situation is still valid:

- Kreutzer, T. (date unknown) [Validity of the Creative Commons Zero 1.0 Universal Public Domain Dedication and its usability for bibliographic metadata from the perspective of German Copyright Law](#). (last accessed: 5 July 2018)
- Cox, K.L. (2017) [Metadata and Copyright: Should Institutions License Their Data about Scholarship?](#) Copyright, Fair Use, Scholarly Communication, etc. 59. See section “Recommendation”

**2.3. For metadata to be used meaningfully, it must be standardised to optimise machine-reading (a requirement of H2020 projects). A commonly used format in libraries and cultural heritage institutions is Dublin Core.**

#### LEGAL SOURCE

The following source details issues which arise from using inconsistent metadata practices:

- Knoth, Petr (2013). From open access metadata to open access content: two principles for increased visibility of open access content. Open Repositories 2013, 8-12 Jul 2013, Charlottetown, Prince Edward Island, Canada. (online resource available here: <http://oro.open.ac.uk/37824/>, last accessed: 4 July 2018)

#### EXAMPLES

RIOXX provide a tool which tests repositories metadata compliance with open access standards:

- RIOXX (date unknown) The RIOXX Metadata Profile and Guidelines <http://rioxx.net/> (last accessed: 4 July 2018)

For details on the formatting and implementation of Dublin Core bibliographic metadata, see:

- Dublin Core (date unknown) Dublin Core Metadata Initiative <http://dublincore.org/> (last accessed: 3 September 2018)

### 3. Content should also be licensed

---

**3.1. In point 1 you have applied a licence to your repository, and to its content “unless otherwise noted”. Now let’s take a look at the “unless otherwise noted” part.**

*As a repository manager, you (or the University) usually don’t own the copyright in the articles uploaded (unless you have written them).*

Therefore, the repository has to implement a licence selection procedure that allows the uploader (author or rightsholder) to choose the proper licence. As further detailed in [point 6](#), this process should offer a number of choices to the author, but it should be the author who ultimately decides what licence to use.

Nevertheless, in order to help researchers to make the right choice, you can offer and implement some guidance that will help researchers to make the right choice and to adhere to Open Science principles.

In points [4](#) and [5](#) you will see what choices are recommended for a) data and databases; and b) articles.

#### [LEGAL SOURCE](#)

See guidelines on open access as provided by the European Commission for details on self-archiving and open access publishing:

- European Commission (date unknown) "[Open Access](#)". (last accessed 13 July 2018), see "Step 2 - Providing open access to publications"

**3.2. We recommend a CC BY 4.0 licence in respect of the content of the repository. This is detailed further in Point 5. This may not be appropriate for data or datasets as detailed in Point 4.**

#### [LEGAL SOURCE](#)

For discussions on the merits of CC BY see the following resources and details in [point 5](#):

- Amiel, T. and Soares, T.C. (2016) [Identifying Tensions in the Use of Open Licenses in OER Repositories](#). International Review of Research in Open and Distributed Learning, 17(3) (last accessed: 5 July 2018), see "Licensing"
- Creative Commons UK (2017) [Frequently Asked Questions on Creative Commons & Open Access](#). Zenodo. (last accessed: 3 July 2018), see "How should I licence my work for the purposes of Open Access?"
- Mewhort, K. (2012) [Creative Commons Licenses: Options for Canadian Open Data Providers](#). (last accessed: 5 July 2018)

#### [EXAMPLES](#)

Some funders and institutions may require that any outputs are made available under a CC BY licence, some of which are detailed in the following source:

- Springer Nature (date unknown) [Funders and institutions requiring a CC BY licence for OA articles](#). (last accessed: 4 July 2018)

**3.3. Creative Commons licences are not appropriate for software.**

*Instead, we would recommend that a GNU GPL v3.0, BSD/Apache style licence is applied. These are some of the most well-established free libre open source (FLOSS) public licences for software. The GNU GPL licence is known as a "copyleft" licence (with provisions similar to Share Alike in CC licences), whereas the BSD/Apache style licences are known as permissive free software licences. All licences detailed here are also highly interoperable with other licences.*

#### [LEGAL SOURCES](#)

Creative Commons confirm that their licences should not be applied to software:

- Creative Commons (2018) [Frequently Asked Questions](#). (last accessed: 4 July 2018), see “Can I apply a Creative Commons license to software?”

#### EXAMPLES

The following open access software licences are appropriate substitutes for a CC BY or CC0 licence:

- The Apache Software Foundation (2018) [Licensing of Distributions](#). (last accessed: 19 August 2018)
- GNU Operating System (2018) [How to choose a license for your own work](#). (last accessed: 4 July 2018), see “Software”
- Open Source Initiative (date unknown) [The 2-Clause BSD License](#). (last accessed: 19 August 2018)

### **3.4. Whilst Creative Commons licences may apply to both digital and non-digital content, this guide currently provides advice only in respect of fully digital repositories.**

*Libraries and cultural heritage institutions may need to audit their non-digitised resources, and check for complex or multilayered content (e.g. multiple authors, orphan works etc.)*

#### LEGAL SOURCE

For guidance on digitising works for the purpose of creating a digital repository, see the following sources:

- Hamilton, G. and Saunderson, F. (2017) Open Licensing for Cultural Heritage. London, Facet Publishing, chapter 12 (p167-p193)
- Jordan, M. (2006) Putting Content Online: A Practical Guide for Libraries. Oxford, Chandos Publishing.

#### EXAMPLES

This may be a particularly relevant consideration for any projects funded by the European Research Council, which require that all project materials be machine-readable. In this case, scans are not acceptable, which may impact non-digital repositories:

- European Research Council (2017) “[Guidelines on Implementation of Open Access to Scientific Publications and Research Data in projects supported by the European Research Council under Horizon 2020](#)”. (last accessed 13 July 2018)

### **3.4.1. If you need to digitise materials for use on an online repository, note that there may be special considerations for traditional knowledge works.**

#### EXAMPLES

Both the Alaska Native Knowledge Network and Charles Darwin University provide examples of special considerations and guidance when digitising or collecting traditional knowledge works:

- Alaska Native Knowledge Network (date unknown) [Guidelines for Respecting Cultural Knowledge](#). (last accessed 3 September 2018)
- Charles Darwin University (date unknown) [Aboriginal and Torres Strait Islander protocols for libraries, archives and information services](#). (last accessed 3 September 2018)

## 4. Data and datasets and databases should be under CC0

---

### 4.1. Data and datasets and databases should be offered without restrictions on use, meaning under a CC0.

#### EXPLANATION

- **Data:** as such not protected by copyright.
- **Dataset:** not defined by law, can include database (as defined by law) and other structured and unstructured data.
- **Database:** defined by law as “a collection of independent works, \*data\* or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means.”

Remember that a database can be protected by copyright (database structure) and/or SGDR (substantial investment in obtaining verifying and presenting data), without prejudice to any copyright or other rights in the underlying material.

Applying a CC0 to a database means that if any rights exist they are waived, if they don't exist CC0 does not create any obligation. If waiver is not possible then CC0 operates as a waiver or as a licence to the same effect within the limits of applicable law.

#### LEGAL SOURCES

- Creative Commons (2018) Open Data Guide. (last accessed: 3 July 2018).
- Creative Commons (2018) Open Science. (last accessed: 3 July 2018).

#### 4.1.1. The advantages of making data available without restrictions include:

*Greater availability and accessibility of publicly funded scientific research outputs;*

- Possibility for rigorous peer-review processes;
- Greater reproducibility and transparency of scientific works;
- Greater impact of scientific research.

#### LEGAL SOURCE

- United Nations Educational, Scientific and Cultural Organisation (2018) Global Open Access Portal. (last accessed: 20 August 2018).

**4.2. Where the uploader is concerned with regards attribution they can ‘kindly request’ to be attributed, rather than using the more legally restrictive CC-BY licence.**

*This is not legally binding although follows standard scholarly practice in crediting researchers for their work.*

**LEGAL SOURCE**

- Creative Commons UK. (2017). Frequently Asked Questions on Creative Commons & Open Access. Zenodo. (last accessed: 3 July 2018).

**4.2.1. When ‘kindly requesting’ attribution of a work, the uploader should be advised to offer a citation which can be easily copy and pasted by subsequent users.**

**LEGAL SOURCE**

- Creative Commons UK. (2017). Fact Sheet on Creative Commons & Open Science. Zenodo. (last accessed: 20 August 2018).

**4.2.2. Tools such as the Creative Commons ‘Open Attribute’ tool are available to assist with ensuring adequate attribution.**

**LEGAL SOURCE**

- Creative Commons (2018) Open Attribute Tool. (last accessed: 3 July 2018).

**4.3. Where there are concerns with regards privacy issues or data protection, these should be dealt with under the relevant legislation or ethics policies.**

**LEGAL SOURCE**

- European Commission (2018) 2018 reform of EU data protection rules. (last accessed 23 July 2018).

**4.3.1. Both the BioMed Central consultation and the PLOS data policy address these issues.**

**LEGAL SOURCES**

- Bloom T, Ganley E, Winker M (2014) Data Access for the Open Access. (last accessed 18 July 2018).
- Hrynaszkiewicz et al (2013) Licensing the future: report on BioMed Central’s public consultation on open data in peer-reviewed journals. (last accessed 18 July 2018).

**5. Other works of authorship (articles, images etc.) should be licensed under a CC-BY 4.0 license**

---

*Suggest which licence should be chosen to meet OS requirements, but let the uploader choose.*

### 5.1. Give the uploader the possibility to choose the licence.

*You can indicate which licences are better for OA/OS, but don't choose for them.*

To avoid ambiguity, uploaders should be expected to apply a licence at the point of upload. Failure to apply a licence at upload results in 'All Rights Reserved', which generally means people are unable to use, re-use, modify or data-mine the unlicensed content, without authorisation.

#### LEGAL SOURCES

- Creative Commons UK. (2017) Frequently Asked Questions on Creative Commons & Open Access. Zenodo. (last accessed: 3 July 2018).
- Creative Commons (2018) Creative Commons Integration, from A to Z: 'Incorporating CC licences into a user interface'. (last accessed: 3 July 2018).

### 5.2. Repositories may play an important role in educating uploaders with regards open licensing.

#### LEGAL SOURCE

- Creative Commons (2018) Creative Commons Integration from A-Z: Communication and Education. (last accessed 23 July 2018).

#### 5.2.1. The importance of making their work open access, should be explained to uploaders prior to upload.

The benefits of open access work include: Researchers and their institutions benefit from having a wider audience Open access allows use of text and data mining tools, without legal barriers. Funders receive a greater return on their investment when results of research can be utilised by more people and at an earlier date.

#### LEGAL SOURCES

- Budapest Open Access Initiative (2002). (last accessed: 3 July 2018).
- Bethesda Statement on Open Access Publishing (2003). (last accessed: 3 July 2018).
- Berlin Declaration on Open Access (2003). (last accessed: 3 July 2018).

#### 5.2.2. SPARC Europe offers a useful summary of the benefits of making work open access.

#### LEGAL SOURCE

- SPARC Europe (2018) Setting the Default to Open: Open Access. (last accessed: 20 August 2018).

## 6. Repositories should recommend the best OS licenses but it should be the uploader who chooses which one

---

**6.1. Uploaders should be offered all possible guidance and explanation with regards the various licences open to them, and the degree to which these are compatible with open access principles.**

#### LEGAL SOURCE

- Creative Commons (2018) Creative Commons Integration, from A to Z: Licence Chooser. (last accessed: 3 July 2018).

**6.1.1. This can be done by incorporating some form of ‘Licence Selector’ tool into the upload process. The tools featured here offer examples of how this can be achieved.**

#### LEGAL SOURCES

- Creative Commons (2018) Creative Commons Integration, from A to Z. (last accessed: 4 July 2018), see “Model Platform”.
- Hamilton, G. and Saunderson, F. (2017) Open Licensing for Cultural Heritage. London, Facet Publishing, chapter 12 (p167-p193).
- Van Piggelen, H. (2018) Eudat: License Selector guide. (last accessed: 18 July 2018).
- ‘Wikimedia’ (2018) Open Content - A Practical Guide to Using Creative Commons Licences/The Creative Commons licencing scheme. (last accessed: 3 July 2018).
- ‘Smartcopying’ (2018) Open Education. (last accessed: 3 July 2018).

**6.2. CC-BY 4.0 may be considered as a default standard licence, except in the case of data and datasets.**

*However, any default licence provided should always be accompanied by a selection of alternative licences and comprehensive explanations about the function of each.*

The CC-BY 4.0 licence is often considered the ‘gold standard’ open access licence, since it is the least restrictive and allows people to use the licensed content as they choose, provided attribution is provided, and is fully OA compliant. As a note of caution, however, it should always be the uploader who makes the final licence selection.

#### LEGAL SOURCE

- Creative Commons UK. (2017). Frequently Asked Questions on Creative Commons & Open Access. Zenodo. (last accessed: 3 July 2018).

**6.3. Where uploaders select a licence which is less compatible with open access/science requirements, this should be made clear to them.**

This is particularly relevant where uploaders choose Creative Commons licences with NC (non-commercial) or ND (no derivatives) conditions. These licences have been described by Creative Commons as failing to promote ‘free culture’.

#### LEGAL SOURCE

- Creative Commons UK. (2017). Frequently Asked Questions on Creative Commons & Open Access. Zenodo. (last accessed: 3 July 2018).

#### **6.4. In the case of software, application of a GNU GPL or BSD/ Apache style licence is recommended.**

These licences are:

- The most well-established public licences for free software, and
- The most interoperable licence both in terms of general use and for TDM purposes.

##### LEGAL SOURCES

- Creative Commons (2018) [Frequently Asked Questions](#). (last accessed: 4 July 2018), see “Can I apply a Creative Commons license to software?”.
- The Apache Software Foundation (2018) [Licensing of Distributions](#). (last accessed: 19 August 2018)
- GNU Operating System (2018) [How to choose a license for your own work](#). (last accessed: 4 July 2018), see “Software”.
- Open Source Initiative (2018) [Licences and Standards](#). (last accessed 20 August 2018).

#### **6.5. In the case of public sector information, application of an Open Government Licence is mandated by the UK Government Licensing Framework (UKGLF) for all public sector information.**

##### LEGAL SOURCE

- The National Archives (not dated) Open Government Licence. (last accessed 18 July 2018).

#### **6.6. Ultimately, however, the final decision with regards which licence is applied should rest with the uploader.**

##### LEGAL SOURCES

- Creative Commons UK. (2017). Fact Sheet on Creative Commons & Open Science. Zenodo. (last accessed: 3 July 2018).
- Creative Commons (2018) Creative Commons Integration, from A to Z: Licence Chooser. (last accessed: 3 July 2018).
- Skills Commons (2018) Open Access Support Sheets: Creative Commons Requirement. (last accessed: 3 July 2018).

#### **6.7. The resources featured here offer comprehensive discussion regarding the benefits of open access principles, and provide an example of how these might be expressed to uploaders.**

##### LEGAL SOURCES

- Goodman D. (2004) The Criteria for Open Access, *Serials Review*, 30:4, 258-270. (last accessed 17 July 2018).
- Sparc Europe (2018) Setting the Default to Open: Open Access. (last accessed: 3 July 2018).
- Wellcome Trust (2018) Creative Commons Attribution licence (CC-BY). (last accessed: 3 July 2018).
- 'Wikimedia' (2018) Open Content - A Practical Guide to Using Creative Commons Licences/The Creative Commons licencing scheme. (last accessed: 3 July 2018).

### **6.8. Account must be taken of any external limitations on the uploader's choice of licence.**

This may be as a result of funding body stipulations or publishers' requirements.

#### **LEGAL SOURCE**

- Creative Commons UK. (2017). Frequently Asked Questions on Creative Commons & Open Access. Zenodo. (last accessed: 3 July 2018).

### **6.8.1. The resources featured here, including the European Commission H2020 guidance, provide an example of possible funding body stipulations, with regards making work open access and how this should be done.**

#### **LEGAL SOURCES**

- European Commission: Directorate General for Research and Innovation (2017) H2020 Programme Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020 (Version 3.2). (last accessed 18 July 2018).
- Hamilton, G. and Saunderson, F. (2017) Open Licensing for Cultural Heritage. London, Facet Publishing, chapter 12 (p167-p193).
- Springer Nature (date unknown) Funders and institutions requiring a CC BY licence for OA articles. (last accessed: 4 July 2018).

## 2 | RECOMMENDATIONS REGARDING OPENAIRE SERVICES IN THE CONTEXT OF PERSONAL DATA AND IPR

### 2.1 Introduction

The Objective of this guide is to identify key issues with regards to Personal Data, Ethics and IPR issues related to stakeholders from different communities and provide a list of items that need to be taken into consideration in order to ensure the seamless provision of OpenAIRE services. In order to achieve this goal, we present different recommendations in relation to the different types of stakeholders appearing in the OpenAIRE ecosystem, mainly, Research Performing Organisations, Funders, E-infrastructures and Research Infrastructures and Policy Makers. These categories include also international or intergovernmental organisations that may assume any of the aforementioned roles. In addition, for each of these categories we focus on how we may lower transaction costs and, thus, increase the utility of the services offered by OpenAIRE. All recommendations take into consideration the broader EOSC context and the need to locate OpenAIRE services within that context.

### 2.2 Personal Data Protection and Ethics

The introduction of the General Data Protection Regulation (GDPR) has marked the emergence of a more thorough European Data Protection Regime with a global outlook and implications. GDPR entails a series of specific obligations to data controllers and processors and, thus, directly influences the research process. By mandating data protection by design and data protection by default, it brings a series of obligations for all involved stakeholders, while at the same time it requires a more holistic approach in relation to the development of Data Management Plans.

The main recommendations regarding all types of stakeholders and communities may be summarized as follows:

#### **1. Create Data Protection compliance structures within the organisation**

Before starting any data protection/ GDPR compliance activities it is necessary that an RPO has set a number of relevant structures in order to facilitate the taking of technical and organisational measures of compliance. These should include *de minimis*:

- Appointment of a Data Protection Officer (DPO)
- Setting up of a Data Protection Governance Forum/ Working Group
- Establishing an independent Ethics Committee

Such a structure will facilitate the implementation of any compliance measure both at the organisational and project level and is the stepping stone for a GDPR compliant sustainable data management approach.

#### **2. Create a GDPR compliance road map**

It is essential to create a concrete roadmap for the implementation of GDPR with particular focus on stages of maturity both at the organisational and project/ research programme level. The roadmap should at the outset of the compliance process to have a duration of 12 to 24 months with three-months progress milestones. In each milestone an assessment of technical, organisational and legal measures should be conducted.

### **3. Focus on processes rather than data as the starting point for the gap analysis**

In order to perform a comprehensive gap analysis it is necessary to focus on different processes rather than on data. As a result it is necessary to perform a diligent recording of the processes and identify on these (a) the range of personal data used (b) the various forms of processing (c) risk areas in terms of probability and impact (d) measures taken to mitigate the risk (e) recurring compliance measures.

### **4. Focus on data rather than documents**

While GDPR compliance entails proper data management, a great part of such data is still collected through documents rather than forms. There needs to be a clear process of data set creation and ensuring that the documents or forms where the data were collected from are thoroughly and properly managed.

### **5. Use existing processes, technologies and internal rules**

While GDPR compliance is not a given in most RPOs, a baseline of data protection practices, processes, organisational and technical measures exists. There is also an abundance of documents, notices and contractual terms that ensure some level of compliance. It is always necessary to collect, assess and maintain those elements that may be used in the GDPR context. In addition, the use of already existing elements may allow faster and more sustainable compliance as it may better relate to existing organisational culture and practices.

### **6. Align GDPR compliance with the Data Management Plan**

GDPR compliance is mostly an issue of processes, organisational and technical measures and always follows the data flows of a research project or administrative process. As a result, it is always advisable not to reinvent the wheel but rather try to match the DMP for a particular project or research protocol with the GDPR compliance process in order to lead to standard and cross-organisational DMPs.

### **7. Differentiate between administrative, service oriented and research processes**

In order to accelerate the compliance process, while ensuring a low transaction cost for all involved parties, it is necessary to differentiate between three types of processes:

- Administrative processes: these mostly relate to procurement processes, labour and employment issues and public administration obligations.
- Service processes: these refer to e-/research infrastructure type of services both to members of the RPO and third parties that need to have a minimum Service Level and as

such need to follow standard processes of compliance. These tend to be highly standardised and need to fulfil the customer needs.

- Research processes: these are contingent upon the type of research, they are not always standardised and need to follow the ethics rules as well as research protocols.

These differentiations allow a clear separation of types of processes and a faster implementation of the compliance processes. It also takes into consideration the different cultures and processes within the same organisation, while maintaining the required high compliance level.

## **8. Establish a Data Protection Working Group**

In order to ensure that compliance will be achieved at all levels of the organisation and across the range of research projects undertaken by different groups and research teams it is necessary to set up a Data Protection Working Group. These will involve persons from different parts of the organisation, offering e-services or research groups undertaking different forms of research in order to ensure their needs and views are expressed during the compliance process. In the case of broader research communities, the Data Protection Working Group should also take into consideration the needs of different members (organisations) of the research group and the different facets of the conducted research.

## **9. Ensure proper documentation of GDPR compliance**

A core element of compliance is the proper documentation of all stages of the research process in a standard and easily traceable fashion. The use of information systems rather than simple spreadsheets allows the lowering of costs in the long run and the more effective and efficient monitoring of the compliance maturity.

## **10. Establish an Ethics Committee**

The establishment of Ethics Committee at the level of the organisation providing the services, a research consortium or an RPO is essential in order to ensure there is guidance and proper monitoring of both data protection and ethics rules. The Ethics Committee should also liaise and coordinate with the relevant DPO(s) in order to join forces and achieve the desirable compliance results.

## **11. Produce standard data protection compliance protocols**

In order to avoid duplication of effort and repletion of processes, it is necessary to create standard processes, documents, notices and model contracts at the level of the RPO, the cross-organisational research group or the individual research team. These may mature through their daily use and testing and should ideally also be combined with the necessary processes and technical measures.

## **12. Maintain a comprehensive and consistent consent archive**

While consent is obtained in most of the cases where personal data (including those of special categories (sensitive data)) are obtained, the way in which consent is collected is not recorded.

Similarly, the range of data processes and their transformation in the course of time also needs to be carefully monitored and ensure it remains within the scope of consent. In that sense a comprehensive archive of consent documents, accompanied by proper meta-data schemes is necessary for a sustainable data protection scheme.

### **13. Create standard, machine readable and comprehensive cross organizational data protection policies**

There is need for standard or at least modular and machine readable policies, so that it is possible to share them between RPOs or e-services in a seamless and interoperable way. Custom policies in relation to data processing may lead to effectively unusable data sets.

### **14. Tackle high risk areas in data processing**

The most important data processing related risks appear at the points of data collection, access rights and sharing with third parties. In addition, there are always issues in relation to the data detention period and the process of deletion. In that sense, it is necessary that the focus of the compliance process always deals with these cases. In addition, special care should be taken to monitor and ensure proper documentation of the legal bases in the course of the data life cycle or ensure an easy mechanism for obtaining the necessary legal bases.

### **15. Establish regular reviews of processes, documentation and model documents**

While there is a substantial initial investment in time, people and resources in ensuring compliance, the lack of a regular review of processes, documentation and model documents may lead to a depreciation of these assets. Hence, their review and improvement should be part both of any organizational or interorganizational compliance process.

### **16. Train researchers, business units and administrative staff**

Training of researchers, business units offering services and administrative staff is essential for ensuring GDPR compliance at all levels. This needs to happen both at the outset of the compliance process and in regular intervals, particularly when major amendments of the standard processes, documents or technologies take place.

## **Review all data sharing agreements**

Data sharing agreements between research groups, organisations or other third parties is necessary in order to ensure that GDPR compliance is there. Such review should always take into account the different data protection policies in cases of inter-organisational collaborations and, hence, needs to lead to workable results/ recommendations while avoiding the problem of the lowest common denominator in data protection.

## **17. Clearly differentiate IPR from data protection issues**

While both data protection and IPR issues may deal with the same data set and may have as a common objective their maximum sharing, they approach them from an entirely different legal stand point. In that sense, they require a different documentation and monitoring process, which may nevertheless exist on the same system and be part of the same Data Management Plan.

## **18. Ensure legal basis when data are released under an open licence**

The moment of releasing a data set under an open licence has as a precondition a proper legal basis. This is necessary to be established at the moment of data collection rather than at the moment of data release, since the costs increase or may render the establishment of a sound legal basis impossible when they are sought at the end rather than the beginning of the research process.

## **19. Monitor GDPR compliance maturity (both at the level of the service and the organization)**

In order to ensure that there is a progress and improvement in the compliance process it is necessary to develop Key Performance Indicators both at the level of organization and cross organization project and monitor their improvement in the course of time. It is also important to make such monitoring both with regards to specific services (e.g. ways in which consent is obtained or ways in which the data are disseminated) and an organization as such (e.g. in the internal processes it has or the technical and organizational measure it has implemented).

## **20. Set GDPR compliance (data protection by design and by default) as a funding requirement**

Funders should set as a funding requirement the existence of a minimum of maturity regarding data protection compliance, either in the research process (if it is individual researcher) or the organization/ consortium (if it is institutional funding). Close collaboration between RPOs, e-Infrastructures, Policy Makers and Data Protection Authorities is necessary in order to produce workable requirements.

## 2.3 IPR

This section provides an outline of the main policy themes and recommendations falling under the IPR umbrella. It aims at a comprehensive, yet concise, approach addressing the main challenges related to IPR that are identified in previous sections.

The following subsections make recommendations regarding different types of IPR policies, that follow the life-cycle of an intangible asset (identification, clearance, sharing, management, dissemination). In the core of our approach is the reduction of any unnecessary transaction costs throughout the life cycle of the research process in a way that supports the vision and practice of open science.

The main recommendations related to IPR may be summarized under the following titles:

### **1. Coordinate Open Access and IPR reutilisation in a comprehensive and coherent IPR framework**

The issue of IPR and open science, particularly open access, are often presented as antithetical or incompatible. As this section demonstrates, there is limited merit in such an approach. Different forms of IPR exploitation relate to open licensing in a limited fashion only and, mostly, have to do with the choice of time when the protected material is to be released, especially in relation to patents. In addition, IPR exploitation policies are closely related to questions of rights registration and enforcement.

### **2. Have a proper IPR documentation when releasing or accessing a research resource**

Rights documentation is a crucial part of any IPR policy, as it allows all involved stakeholders to have an accurate understanding of the rights status of different assets in different stages of their life-cycle (registration, sharing, licensing). Documentation should cover at least the following:

- (a) type of IPR
- (b) ownership of rights
- (c) licensing of resources

Such documentation should normally exist in three areas:

- (a) on the resource itself (e.g. ownership and copyright notices on a document)
- (b) in the meta-data of the resource file (e.g. in the meta-data of .doc file)
- (c) in the repository meta-data

In the case the resources are offered through a web service (e.g. an API), the API documentation should also include the terms and conditions (TCs) or Terms of Service (ToS) under which they are offered.

Both licences and ToS/ TCs have to be stored in a permanent URI. They also have to follow a clear versioning system and contain a versioning history (versions/ date). To the extent possible, a change of version shall be communicated to the recipients of the service (e.g. registered users) or made visible through a public website.

### **3. Clear IPRs before sharing them over e-Infras/ Research Infrastructures**

Rights clearance is a precondition for sharing any research output or resource and ensuring this happens before the introduction of the resource in a shared environment will substantially reduce research transaction costs and risks.

#### **4. Provide coherent and consistent IPR ownership policies**

One of the greatest challenges in comprehensive IPR policies for all types of organisations is the introduction of clear ownership and rights registration policies. Such policies allow all levels and types of participants in a research process to have a clear understanding of their rights regarding their contribution in a specific creative process.

#### **5. Have a clear access and rights management regime**

Rights management within a research environment, by and large, relates to the access rights that different levels and types of staff have on research results and services. This needs to be in accordance with all the aforementioned points and provide a coherent framework both for reducing potential risks and for ensuring no unnecessary exclusion of persons or institutes requiring access to resources is in place.

#### **6. Ensure that licensing policies accommodate different types of value production**

Licensing schemes are necessary both in relation to the establishment of any type of collaboration related to resources and services and in relation to exploitation of resources in a broader value chain including such resources. The relevant stakeholders should make provision in order to have in place policies both in relation to collaboration and the exploitation/ dissemination of resources.

#### **7. Introduce Open Access enforcement policies and mechanisms**

Enforcement policies should address three issues: First, how the organisation is to monitor the implementation/ application of the licence agreements it grants in relation to its own assets; second, how it is to respond to infringements of its licences and/ or IPR in general; and third, how is it going to respond to infringements that take place through the services/ assets it provides to third parties.

These recommendations need to be positioned in a broader context comprising of the following elements that describe actions that need to be taken by different stakeholders in the broader open science ecosystem:

## I. RESEARCH PERFORMING ORGANISATIONS SHOULD

- Adopt a holistic IPR policy that covers all types of IPR, i.e. Copyright, Patents, Trademarks and Design Rights.
- Collaborate with National IPR Offices to create custom IPR awareness campaigns with an emphasis on the interaction between IPR and open access
- Increase the number and quality of IPR courses for non-lawyers focusing on interaction between open access and IPR utilisation
- Adopt minimum IPR documentation policies as a condition for the inclusion of resources in their institutional repositories.
- Ensure IPR documentation is standard and machine readable
- have specific IPR ownership rules for the following instances:
  - regular research activities of the staff
  - research collaboration in the framework of projects funded by third parties
  - research collaborations with commercial parties
  - research conducted in collaboration with RPOs spin offs
  - research collaborations with the government
- specify in clear terms the division of ownership between the RPOs and the individual researcher
- Establish clear access procedures in accordance to their IPR policies and ensure that such policies do not preclude neither open access publication of results nor the utilisation and exploitation of research output.
- Provide clear decision paths for making choices in relation to releasing research results under open licences and the exploitation of research results.
- Provide training and support in relation to the different value production models and set open licensing as the default choice for the publication of research output
- introduce model licensing agreements for open innovation networks
- Establish IPR policies in relation to different forms of exploitation. Such policies should contain at least the following elements:
  - have a patent or other industrial property assessment of research results
  - identify value in monetary and non-monetary terms -at least- in relation to core assets
  - identify possible embargos and specify how the scholarly communication of research results affects the exploitation possibilities of research results
  - specify a life-cycle or asset management plan for different assets contained in the research results
  - introduce model dual/multiple-licence agreements
- Establish standard operational procedures (SOPs) for responding to infringement, reporting to affected owners and limiting damage, including notice and take down procedures.
- Establish risk mitigation strategies, particularly through comprehensive rights clearance at the source of the information entry.
- focus on:
  - violation of attribution terms

- violation of copyleft terms
- violation of the non-commercial clauses
- introduce warning and mediation strategies before escalating legal action in case of infringement.

## II. FUNDERS SHOULD

- Require the existence of comprehensive IPR policies as a precondition for institutional funding
- Have IPR documentation of all research outcomes as a condition for funding a research project.
- have clearance of rights as an eligible cost in their funding programmes
- not accept as deliverables any content/ research resources that remains uncleared
- have clear allocation of IPR ownership as a funding condition
- Establish clear procedures with regards to allowed embargo periods and access limitations in order to maximise open access publications
- Condition funding upon the release of a research output, at a certain stage or certain degree, as open and FAIR content.
- Request a justification on the basis of a comprehensive IPR exploitation plan of any decision not to openly release research output.
- require that individual researchers and RPOs have a clear exploitation plan along with an open scholarly communication plan. In case they fund consortia, they should provide model consortia agreements and in all cases make suggestions in relation to both open licences to be used (mostly those characterised as Free Cultural Work licences<sup>12</sup>), as well as model licensing frameworks.<sup>3</sup>
- Require the existence of SOPs for the enforcement of open licences
- Undertake the funding of the whole or part of the litigation process, as well as encourage collaborations with civil society orgs (e.g. FSF).<sup>4</sup>

## III. E-INFRASTRUCTURES/ RESEARCH INFRASTRUCTURES SHOULD

- condition RPOs participation to e-Infrastructures upon the existence of comprehensive IPR policies for the resources shared on the infrastructures
- only host research content that contains IPR documentation
- Provide tools and guidelines for clearing content
- Ensure that IPR clearance takes place before any resource is shared through the infrastructure and only host IPR cleared material
- ensure there is clear ownership of all resources entering an e-infrastructure
- Provide specific and clear rules for accessing research process and results

---

<sup>1</sup> <https://www.coalition-s.org/feedback/>

<sup>2</sup> For the relevant definition, see <https://creativecommons.org/share-your-work/public-domain/freeworks/>

<sup>3</sup> See e.g. the UK OpenGov Licensing Framework <http://www.nationalarchives.gov.uk/information-management/re-using-public-sector-information/uk-government-licensing-framework/>

<sup>4</sup> See e.g. <https://www.gnu.org/licenses/why-assign.html>

- have very clear rules as to the kind of content they host and how they support scholarly communication and commercial exploitation accordingly.
- Follow a coherent licence policy encouraging Free Cultural Work Licences<sup>5</sup>
- Follow a license compatibility framework, i.e. suggest a limited range of licences and ensure there are licence calculators in place to allow user to re-use and re-combine material.<sup>6</sup>
- introduce Standard Operational Procedures (SOPs) for all kinds of infringements taking place over their network.

#### IV. POLICY MAKERS SHOULD

- Encourage collaboration between National IPR offices and RPOs
- provide incentives for clear rules of ownership and the documentation of the ownership of research resources.
- provide rules specifying minimum embargo periods after which research results should become open
- Ensure that different access levels are based on predefined, rational and transparent parameters that may be monitored and enforced.
- clearly relate public funding to open access and commercial exploitation to state aid or financing from private sources. By identifying a series of different types of value (e.g. monetary and non-monetary), policy makers should opt for open scholarly communication that could be complemented with other types of protection, e.g. patent protection, especially if the disclosure obligations of a patent are fulfilled through the open access publication of the underpinning research.
- take all measures possible in order to reduce licence pollution by encouraging the use of standard and existing licences and also by linking funding and career development with the opening up of research results.
- introduce policies linking the assessment of an RPO with the maturity of its enforcement mechanisms, particularly in relation to the violation of Open licences
- provide guidance and training in relation to the types of liability related to different types of releasing research resources.

#### V. OPENAIRE GOVERNANCE AND POPS SHOULD

- Create an IPR registry containing all IPR policies of participating organisations
- Express IPR policies in a standard and -ideally- machine readable format
- Introduce obligatory IPR documentation as a ground rule for PoP. This includes at least ownership and licensing information.
- Require rights clearance and documentation of the clearance process before any resource is uploaded on OpenAIRE infrastructure.
- Create model collaboration agreements
- Record rights allocation rules in collaboration projects
- Record rights ownership in collaboration projects
- Provide model access policies (modular, standard and machine readable)

<sup>5</sup> See <https://creativecommons.org/share-your-work/public-domain/freeworks/>

<sup>6</sup> See e.g. <http://janelia-flyem.github.io/licenses.html>

- Ensure that all resource providers have an access policy in place
- Produce an OpenAIRE wide modular and standardise model policy for scholarly communications and IPR
- Produce decision trees for the choice of open access policies in accordance to IPR policies
- Provide Licence compatibility charts, wizards and training
- Use standard and documented licences
- Create machine readable licensing policies
- Have an OpenAIRE-wide enforcement policy
- Create SOPs for handling infringement of open licences and communicate it to the users

### 3 | REFERENCES

- [1] [Budapest Open Access Initiative](#) (2002). (last accessed: 3 July 2018)
- [2] [Bethesda Statement on Open Access Publishing](#) (2003). (last accessed: 3 July 2018)
- [3] [Berlin Declaration on Open Access](#) (2003). (last accessed: 3 July 2018)
- [4] Creative Commons UK (2017) [Frequently Asked Questions on Creative Commons & Open Access. Zenodo](#). (last accessed: 3 July 2018)
- [5] Creative Commons (2018) [Creative Commons Integration, from A to Z](#). (last accessed: 3 July 2018)
- [6] Creative Commons (2018) [Open Data Guide](#). (last accessed: 3 July 2018).
- [7] Creative Commons (2018) [Open Science](#). (last accessed: 3 July 2018).
- [8] European Commission (date unknown) "[Open Access](#)". (last accessed 13 July 2018)
- [9] European Commission (date unknown) [Protection of Databases](#). (last accessed: 3 September 2018)
- [10] European Commission: Directorate General for Research and Innovation (2017) [H2020 Programme Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020](#) (Version 3.2). (last accessed 18 July 2018).
- [11] Goodman D. (2004) [The Criteria for Open Access](#), *Serials Review*, 30:4, 258-270. (last accessed 17 July 2018).
- [12] Guibault, L., Margoni, T. and Spindler, G. (2013) Conclusions and Recommendations in Guibault, L. and Wiebe, A. (eds) (2013) [Safe to be open: Study on the protection of research data and recommendations for access and usage](#). OpenAIRE. (last accessed: 1 October 2018)
- [13] [H2020 Framework Programme Regulation \(EU\)](#) No 1291/2013 of the European Parliament and of the Council of 11 December 2013 establishing Horizon 2020 - The Framework Programme for Research and Innovation (2014 - 2020) (OJ 347, 20.12.2013, p. 104) for legal basis, and more specifically article 29.2 of the H2020 Programme: AGA - Annotated Model Grant Agreement (2018). (last accessed 1 October 2018)
- [14] Jordan, M. (2006) *Putting Content Online: A Practical Guide for Libraries*. Oxford, Chandos Publishing.
- [15] Knoth, P. (2013). From open access metadata to open access content: two principles for increased visibility of open access content. *Open Repositories 2013*, 8-12 Jul 2013, Charlottetown, Prince Edward Island, Canada. (online resource available here: <http://oro.open.ac.uk/37824/>, last accessed: 4 July 2018)
- [16] Kreutzer, T. (date unknown) [Validity of the Creative Commons Zero 1.0 Universal Public Domain Dedication and its usability for bibliographic metadata from the perspective of German Copyright Law](#). (last accessed: 5 July 2018)
- [17] Margoni, T., Caso, R., Ducato, R., Guarda, P. and Moscon, V. (2016) Open Access, Open Science, Open Society. In Loizides, F. and Schmidt, B., (eds) (2016) [Positioning and Power in Academic Publishing: Players, Agents and Agendas](#). IOS. (last accessed: 1 October 2018)
- [18] OpenMinTeD (date unknown) [Licence Compatibility Matrix](#). (last accessed: 3 July 2018)
- [19] Sparc Europe (2018) [Setting the Default to Open: Open Access](#). (last accessed: 20 August 2018).
- [20] United Nations Educational, Scientific and Cultural Organisation (2018) [Global Open Access Portal](#). (last accessed: 20 August 2018).

