



European Data
Grant agreement number: RI-283304

Report from the Workshop on Data Access and Re-Use Policies (DARUP), Rome 11-12 November 2014

Author(s)	Richard Blake (STFC), Rob Baxter (EPCC), Damien Lecarpentier (CSC)
Status	Final
Version	1.1
Date	02/03/2015

Abstract: This report reviews the different contributions and summarises the overall discussions from the EUDAT Data Access and Re-Use Policies workshop held in Rome in November 2014.

Document identifier: EUDAT-WP2-DARUP-1	
Deliverable lead	EPCC
Related Work package	WP2
Author(s)	Richard Blake (STFC), Rob Baxter (EPCC), Damien Lecarpentier (CSC)
Contributor(s)	See Annex A
Due date of deliverable	-
Actual submission date	-
Reviewed by	-
Approved by	
Dissemination level	PUBLIC
Website	www.eudat.eu
Call	FP7-INFRA-2011-1.2.2
Project Number	283304
Instrument	CP-CSA
Start date of Project	01/10/2011
Duration	42 months

Disclaimer: The content of the document herein is the sole responsibility of the publishers and it does not necessarily represent the views expressed by the European Commission or its services.

While the information contained in the document is believed to be accurate, the author(s) or any other participant in the EUDAT Consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose.

Neither the EUDAT Consortium nor any of its members, their officers, employees or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

Without derogating from the generality of the foregoing neither the EUDAT Consortium nor any of its members, their officers, employees or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

TABLE OF CONTENTS

EXECUTIVE SUMMARY	4
1. SESSION 1: RESEARCH DATA POLICY: ARE WE WINNING?.....	5
1.1. Over-complicating data management obligations will kill data sharing and re-use	5
1.2. JRC Data Policy: an update on progress	6
2. SESSION 2: THE LEGAL MINEFIELD: HOW TO AVOID SERIOUS INJURY	8
2.1. Report on open data survey from EUDAT	8
2.2. Mapping the right legal issues.....	9
2.3. Lightning Talks and War Stories	11
2.3.1. Experiences – and solutions – from Creative-B.....	11
2.3.2. Data problems for agri-food researchers.....	12
2.3.3. Copyright Reform and WIPO	13
2.3.4. Points of contact between data access/reuse policies and data preservation policies	13
2.4. Further Ideas and Open Discussion.....	14
3. SESSION 3: SHARING DATA: TOO MANY STICKS, NOT ENOUGH CARROTS?.....	15
3.1. Funders’ requests on data sharing: an overview of current practices	15
3.2. Incentives: what works, what doesn’t.....	17
3.3. Funding RD infrastructures: initial results	19
4. DARUP ACTIVITIES FOR 2015.....	19
4.1. The EUDAT Licensing Wizard – a demo	19
4.2. Towards data Policy Harmonisation.....	20
4.3. How can we make best use of DARUP?	21
ANNEX A. LIST OF PARTICIPANTS	23

EXECUTIVE SUMMARY

As part of its preparations for the next phase of its common data services, EUDAT organised a series of workshops in Barcelona in September 2013 to consider establishing the need for Working Groups in a number of topics. One of these addressed policies regarding **Data Access and Reuse Policies (DARUP)** at community and service provider levels.

DARUP is an important area for a number of reasons. Policies must be made clear to potential users of data infrastructures, both the “data providers” – who need to know how their data will be made accessible and who will be able to access it – and the “data consumers” – who need to know about access or reuse restrictions before starting work with any of the data. Data access and reuse policies within communities must also be understood by e-Infrastructure providers so that they can provide adequate resources and tools to handle the data. It is vital for data infrastructures like EUDAT, which deal with data coming from multiple communities, to determine a common approach on this issue and to specify principles regarding data access and reuse within their own infrastructures.

The **1st DARUP Workshop** brought together a broad set of participants from funding agencies, research communities and e-Infrastructure providers. The Workshop considered funders and policy-makers’ views on data access and re-use, data access and re-use policies in pan-European research communities – mapping the current landscape and practices and, managing data and its restrictions: policy and technology solutions. The workshop concluded that there was a strong need for a DARUP Working Group to advise EUDAT on what it should be doing in this area and identified four themes that it should explore in further detail: look at existing EUDAT data policies, licensing, credit and citation and tools and support. The Working Group subsequently established an action plan.

The **2nd DARUP Workshop** organised in Rome in November 2014 brought together a similar set of participants to those involved in the first workshop and considered progress against the action plan. This had been significant in several areas in particular in reviewing EUDAT partners’ approaches to data policy, different licensing options and the development of a tool to support licence selection. The agenda considered the issues as to whether we are winning in terms of the development of data policy, the legal issues around protecting data and the carrots and sticks related to the sharing of data and future DARUP activities.

The Workshop concluded that DARUP was doing a very good job in identifying key areas of importance to users and progressing relevant actions. Priorities for the immediate future included building trust and driving interoperability not only between services but between licences as well; effort would be well spent on developing a manual for data management capturing the best practise from the last twenty years.

It was with great sadness that we learned that our friend and colleague Ville Oksanen passed away shortly after the Workshop. His contribution to our work — a rare combination of expert knowledge and social engagement — cannot be overrated. His input was always very well respected, and on the day of the Workshop he challenged us to take up the open data argument with the World Intellectual Property Organisation. EUDAT would like to extend their deepest condolences to his family and friends and dedicate this report to his memory.

1. SESSION 1: RESEARCH DATA POLICY: ARE WE WINNING?

1.1. Over-complicating data management obligations will kill data sharing and re-use

Max Wilkinson, UCL Research Data Services

Max presented a number of life-cycle models for Research Data Management ranging from the initial research concept, creating/receiving data, appraising and selecting, ingesting, preservation actions, storing, access, data use and re-use and transforming. The lifecycle builds on preservation and curation actions which need to be informed by community watch and participation activities which involve preservation planning and the exploitation of information description and data access technologies. Many researchers are involved in many of these steps, to some degree driven by what is regarded as good practice within their community, using tools that are readily to hand rather than those that may be needed.

The key issue is that researchers are not interested in data per-se. They are driven by research. Data management has always occurred. It is true that there are challenges as the volumes and complexity grow but these considerations are not an explicit part of the process of research. The transition that we are seeking is not one in technology but one in research output and outcomes.

Max went on to explode six myths that are in common circulation in the data management industry.

Myth 1 - We are changing the world: Technology is no longer a passive facilitator of research, it is part of research. Actually, the process of research/science is not being changed; it may be being supplemented or made more vital but it is not being changed in any fundamental way.

Myth 2 - Technology is free(ish): The true costs are rarely made available. No one really knows except those that procure and commission services. This practice needs to change and costs need to be discussed.

Myth 3 - Data re-use is easy. It's not and likely never will be. The research environment changes constantly, techniques change, and infrastructure cannot solve this. Effort will always be required to re-use data and an important lesson is not to over-complicate metadata.

Myth 4 - You should enforce compliance. The only way you can enforce anything in the university domain is with money and history is littered with examples of money drying up and researchers walking away.

Myth 5 - We can standardise this. Standard practices do not mix well with innovation. Standards help interoperability, but so does freedom from standards. Unfortunately each discipline has its own norm and what works well for one may not work well for another. Forcing a particular standard can stifle research.

Myth 6 - You need to licence the public domain. This is perverse as it is rather difficult to enforce liberty. Many universities have a policy of open data, public domain with all rights waived.

In summary:

- there is little benefit in over-complicating a process that exists;
- we need to reveal the actual economic costs;
- be realistic about the benefits as users do not have to use it, reuse existing conventions;
- enabling take-up rather than expecting it;
- enforcing liberty never works;

- compliance – mustn't over-stress this.

There was a general discussion about how to overcome some of the hurdles to data re-use that the talk had highlighted. Some disciplines such as theology and veterinary science do not have a convention of sharing data. Small steps such as including the authors, a title and a unique reference or identifier could enable others to cite the data. Licences can provide a mechanism for supporting the acknowledgement of data, but licences smack of enforcement which does not go down well in academia.

The scale of data re-use cannot be increased significantly over-night; communities need to develop norms for data citation. Conventions do not need to be enforced – they are just accepted. There needs to be clear incentives for researchers to make their data re-usable. The societal benefits are clear, there needs to be much clearer benefits to the individual such as data citations feeding through into career development opportunities.

1.2. JRC Data Policy: an update on progress

Catherine Doldirina, EC JRC

The Joint Research Centre (JRC) is the European Commission's in-house science service with a mission that includes serving society, stimulating innovation and supporting legislation.

The JRC's Data Policy has evolved over a number of years:

- Open Access put forward by the Commission Decision on the reuse of Commission documents (2011/833/EU).
- Open Access put forward by the Commission for scientific publications/data within Horizon 2020.
- Overall open data trend (INSPIRE, G8, ..).
- Open Access to scientific data promoted by Horizon 2020 and by other relevant initiatives (e.g. Research Data Alliance).

The JRC strategy for open access of its scientific publications has been defined and implementation started in 2014.

The JRC Data Policy contributes to the objectives of the JRC Open Data Project through the consistent management of its data, enabling datasets to be cited (for example, by publications), making data cited in publications available through a few data-stores with agreed principles on metadata, storage and accessibility and clear user guidelines and tools. The JRC data policy has been endorsed and work on preparing implementation guidelines has started. Discussions within the JRC for assessing best practices in data management and challenges in the policy implementation are underway.

The JRC data policy has a number of components. Open data principles such as the free (without charge), full (for commercial and non-commercial use), open (without the need to make an individual application) and timely (released as quickly as produced or acquired, or as soon as can be made available or accessible) access to data is a default rule. There are no restrictions on re-use subject to an attribution obligation and a requirement of registration in specific cases. Exceptions are possible if a regulatory basis exists. A harmonised data acquisition strategy aimed at negotiating restriction-free access and use of licensed data is being developed.

The JRC has identified a number of data categories including: that produced exclusively by the JRC (minority), that produced by the JRC with third party data or partners (e.g. as part of EU projects), that produced by the JRC with other data. Data of any type should be shared by the JRC, conditions permitting, and sharing requires the use of licences.

Exceptions regarding data access and use create various scenarios for all of the JRC data categories. These exceptions include: protection of the public interest in security, defence and military matters; international relations; financial, monetary or economic policy of the European Union and commercial interests of natural or legal persons; intellectual property; legal advice related to court proceedings; and for the purpose of inspections, investigations and audit. JRC has some exceptions for making part or all of a dataset available.

A negotiation strategy is required for acquiring third party data. Favourable acquisition conditions cannot be guaranteed. Different licensing conditions may reduce the legal interoperability and availability of the data produced by the JRC with/from third party data. The licensing conditions for already acquired data are not always clear and may render use of open data compatible licences impossible. The JRC's intent is to accept a bare minimum of restrictions so that data can be reused throughout the Commission at best, or at worst across all of the centres even if the data is not available to the public at large. The complexity of the situation could well result in different licences for different data and derivative products, with their re-use restricted by the original licences. What is being sought is a common approach across the JRC with the aim of minimising the number of licences.

The JRC is developing data policy implementation guidelines covering both the governance aspects of implementation (covering stakeholders, Interrelationships, relevant workflows and time frame and costs) and data management aspects of implementation (decisions on making data available, application of the registration requirement, handling various JRC data types and the generation and approval of data management plans). Current decisions include how to provide clear(er) instructions as to making JRC data available, how to support researchers/data stakeholders in this decision-making process, what timeline to introduce and what implementation and data management activities to prioritise.

The announcement of the open data principles, was an important signal to the researchers at the six centres to move towards a data sharing culture as this is not the case at the moment. The discussion explored the tension between the open data principles and the restrictions. The attribution obligation is important to the Commission but is difficult to enforce. Registration procedures can be cumbersome and act as a barrier to data re-use and a good case needs to be made to have them.

The JRC's open data principles are broadly consistent with other initiatives. The question of how much collaboration there had been with ESFRI projects in developing the JRC's principles was raised. The ESFRI principles have no restrictions. It is clear within the exceptions that some data will not be open – what the policy is trying to do is to promote data-sharing leaving the option for restrictions which can be reviewed periodically.

JRC does not have a single policy for the preservation of data. Third party data can be preserved elsewhere but one needs to keep an eye on how long third party data will be made available. JRC does preserve quite a lot of third party data depending on the extent of influence with the suppliers.

The question of what constitutes *sensitive data* within the JRC was raised. An example from the environment area detecting oil spills was given. Satellite data could identify the extent of spillages and the responsible ship. Proceedings could be taken with commercial and or judicial impacts. Other areas of sensitive data include personal data such as health records and security. JRC complies with EU regulations on data privacy in research which are captured in workflows and procedures.

2. SESSION 2: THE LEGAL MINEFIELD: HOW TO AVOID SERIOUS INJURY

2.1. Report on open data survey from EUDAT

Damien Lecarpentier CSC/EUDAT

Damien presented the talk as Marie Sandberg unfortunately could not make the meeting.

Facilitating open access to research data is a principle endorsed by an increasing number of countries and international organizations, and one of the priorities in the European Commission's Horizon 2020 funding framework. But what do researchers themselves think about it? How do they perceive the increasing demand for open access and what are they doing about it? What problems do they face, and what sort of help are they looking for?

As a pan-European research data infrastructure, these are questions that are of fundamental interest to EUDAT. To understand better what researchers think, EUDAT has conducted a programme of interviews with fourteen major research communities from the fields of life sciences, Earth and atmospheric science, astrophysics, climate science, biodiversity, agricultural science, social science and humanities – a broad cross-section of European research interests. While one cannot, of course, interpret the views of any given individual as the official position of a whole research community, they nevertheless provide useful information on the general attitude, requirements and challenges researchers face with regard to opening up their research data. In this article we report on our initial conclusions from this survey.

Open access to research data is increasingly seen as a compelling principle in many research communities. There is a growing awareness of the global move towards open access, the potential benefits of it, and the necessity to implement open access policies within their disciplines. According to preliminary figures on the first wave of open data pilot projects in Horizon 2020, the opt-out rate among proposals submitted to the “open by default” categories was below 30%, and the opt-in rate among other proposals was around about the same. This underlines our findings in EUDAT– researchers are pretty happy about sharing their data.

In practice, though, there are many unsolved challenges still to be addressed, and those most often quoted by researchers were the ethical and legal complications, and the issue of credit.

Not all data can be made open access. Personal data, and especially sensitive personal data, cause particular challenges. In these days of large-scale combination and data mining, can such data truly be anonymised for research purposes? And what about the re-purposing of data for ends very far away from the original research agenda – for military or even criminal purposes? There are no easy answers to these questions, and the culture of ethics surrounding good research is making some communities tread warily.

Our survey highlights a lack of knowledge on the legal aspects of data sharing and data reuse, in particular around intellectual property rights, copyright and licensing, which can act as a barrier not only for opening data but also for re-using someone else's data. Choosing the right licence, for instance, can be a daunting task for some researchers who don't necessarily understand the implications of their actions.

While researchers are naturally keen to see their research published as widely as possible, in an interesting contrast to the open access scholarly paper movement, open data is viewed differently. Often research groups invest significant time and effort in collecting “hard to get data” which can then be used to build careers, offering what can only be termed a competitive advantage over those who do

not have access to the same data. This issue of credit and consequent career progression is a real concern in many communities.

While aware of, and supportive of, the open access data agenda, many research communities are looking for guidance about the practicalities of doing it; training on managing the legal issues, for instance. They also feel that these issues should be addressed at cross-disciplinary level, perhaps rendering the tasks even more challenging. And while much of the open access focus is on coordination efforts, training needs and policies, researchers also stress the importance of developing the right tools and services to enable these policies and, ultimately, the sharing and reuse of data; this is seen as particularly crucial for handling sensitive data.

Compared to scholarly publications, open access to research data is both less developed and more difficult to implement. Although open access to research data has only just begun, the broad spectrum of expectations on EUDAT and other initiatives shows that research communities have the notion that open access to research data cannot be solved through isolated activities or actions; instead it needs to underpin the whole system, reaching from strategic planning and overall policies to the mindset and everyday practice of the individual researcher.

2.2. Mapping the right legal issues

Pawel Kamocki, IDS Mannheim

The talk described the pyramid in which data feed into information which feeds into knowledge and then hopefully into wisdom. The concept of data is not well defined for research although *information* is defined as ‘data plus meaning’. There is also a Diaphoric Definition of Data: ‘A datum is a putative fact regarding some difference or lack of uniformity within some context within the real world, between two physical states (science) or between two symbols (humanities – linguistics). Data are neutral – they have no meaning without context. Data are relational entities, information can consist of different types of data, there can be no information without data representation and data can have meaning independently of whoever reports it.

In information science, data can be symbols that represent a property of an object or be the products of observation, discrete, objective facts or observations, uncategorised and unprocessed, therefore having no meaning because of the lack of context or sensory stimuli that we perceive through our senses.

The OECD Principles and Guidelines for Access to Research Data from Public Funding defines research data as factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings. A research data set constitutes a systematic, partial representation of the subject being investigated. Research data is closest to the OECD definition; however, this has issues in that it doesn’t really help disciplines where data are not the primary sources for “scientific” exploration.

In the spectrum of ‘hard’ to ‘soft’ data in different sciences there are a range of different mechanisms used to protect data as illustrated below.

Hard	Soft
geoscience physics chemistry biology medicine economy sociology linguistics	literary science
Copyright in data	
Copyright in compilations	
Scientific editions	
<i>Sui generis</i> database right	
	Personal data
Patent law	
Trade secret	

Copyright protects original works which are the author’s own intellectual creation, but not ideas, pure facts or discoveries. The copyright grants exclusive rights to copy, to share and to make derivative works. The rights last for the life of the author plus seventy years, and also protect original compilations.

Scientific and critical editions rights apply only in some jurisdictions (DE, IT, PL, ES, PT, UK) and provide quasi-copyright for new editions of public domain works with the scope varying from results of scientific analysis to typographical arrangement in different countries.

Sui generis database rights for databases apply to a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means. There needs to have been a qualitatively and/or quantitatively substantial investment in either the obtaining, verification or presentation of the contents, not their creation. Exclusive rights are granted for the extraction and/or re-utilisation of a substantial part and systematic extraction and/ or re-utilisation of insubstantial parts. The right holder is the maker of the database and the rights last for fifteen years after each substantial investment – potentially unlimited time.

Personal data is defined as any information relating to an identified or identifiable natural person (‘data subject’). There are special categories of personal data – racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, health, sex life – often referred to as sensitive data, although that term has no legal definition. Processing of the data is only allowed in limited cases, e.g.: informed consent (specific consent for special categories of data) and statutory exception (article 83 of the new Regulation). The data subject has the rights of access and rectification, information and objection (if processing for direct marketing purposes). Anonymised data are no longer personal data.

Patents protect patentable inventions that are novel (i.e. not previously known to the public), involve an ‘inventive step’ and are susceptible to industrial application, but which are not discoveries, presentations of information or scientific theories. Exclusive rights are granted for commercial use of the invention but not for private use or research. The term of the protection is for 20 years after filing and the impact on research data is that the patent process both forces disclosure and discourages, but does not prohibit sharing.

A *Trade Secret Directive* was proposed in 2013 to provide protection against the unlawful acquisition, use and disclosure of information of commercial value. The information must be secret (i.e. not generally known among or readily accessible to circles that normally deal with the kind of information in question). It must have commercial value because it is a secret. It must have been subject to reasonable

steps by the rightful holder of the information to keep it secret (e.g. through confidentiality agreements).

It is worth noting that there are statutory exceptions across all of these processes and there are a variety of different licences that can be used, ranging from creative commons through informed consent to individually negotiated agreements.

The workshop discussion started off with the definition of research data: does it, for example, extend to blood samples. The definition is not limited to digital stored data, it could include symbols for example. Samples such as brains etc. can be classed as research materials. Different domains have different levels of abstraction; for example, a scientific paper is not regarded as data by a chemist but linguists could use papers as material. It was felt that the classification of different scientific domains into hard and soft data was not always helpful as areas such as chemistry collect hard data from instruments but needed to interpret the data often using soft hypotheses.

It was noted that different countries have different conventions for defining (so-called) sensitive data which can range from the protection of endangered species to personal data to state secrets. In Europe sensitive data is not defined; there are just special legal categories of data – for example – personal. Different states have different legal authorisation processes to access different documents. The RDA is seeking to harmonise the terms on legal interoperability and had developed twenty principles for research data sharing illustrated through case studies. The DARUP working group should seek to contribute to the development of the principles which are in an early draft form. The principles are relevant to all research, although the incentives in publicly funded versus non-publicly funded research are different.

The issue of transferring personal data across borders was raised. This is fine between member states in the EU but beyond that countries are classified into 1st world states which can provide adequate levels of protection as certified by the Commission. The USA is an exception – here there is a safe harbour agreement which is a private sector agreement that governs the transfer of private data between businesses. In terms of the implications for EUDAT services – we need to distinguish use of services within Europe and outside and appropriate authentication processes have to be built in.

2.3. Lightning Talks and War Stories

2.3.1. Experiences – and solutions – from Creative-B

Wouter Los, University of Amsterdam

Creative-B addresses the Coordination of Research e-infrastructure Activities Toward an International Virtual Environment for Biodiversity: www.creative-b.eu. The project involves centres in the USA, South America, Europe, South Africa, Australia, China as well as a number of global collaborations.

The project has developed a roadmap for interoperability at three levels: community engagement, technology and legal and governance. They plan a service marketplace where users can compare what is offered and select their preferred services and they can consider any imposed conditions. User activity within the marketplace will demonstrate the level of demand to the providers and hopefully this will lead to new developments and support.

An example of a project was presented which uses a spatial decomposition approach to develop habitat models based on suitable rainfall and altitude combined with current observations to describe the potential distribution of an amphibian species in northern South America. A semi-automated workflow for the problem highlighted the need for licences or property rights for algorithms, software,

environmental data layers, permission to use AuthN/AuthZ, moving data from one service to another, third party software and all issues associated with publications. This makes machine/automatic processing of workflows almost impossible. Identical machine readable licence texts might be a solution. The Research Data Alliance has adopted the Creative-B legal analysis as a case study.

2.3.2. Data problems for agri-food researchers

Babis Thanopoulos and Nikos Manouselis, Agro-Know

Agro-Know is a company located in Athens that captures, organises and adds value to the information available in agricultural and biodiversity sciences. They develop and put place in useful services that transform data into meaningful knowledge that are universally accessible. They are leading the data interoperability work in the Agricultural Interoperability Interest Group (IG) at the Research Data Alliance (RDA) and the Database Subgroup, Knowledge & Learning Systems Working Group of the Global Food Safety Partnership (GFSP). The goal of the latter is to support the rapid scaling of food safety practices around the world.

Open models are being developed which leverage other investments, impact on a diverse set of stakeholders and support the development of a range of business models. Technology needs to provide a federated infrastructure as the data sharing backbone. The aim is to develop a modular and scalable architecture that will embrace and facilitate the creation of “global” networks for food safety capacity building.

Requirements for the GFSP project are currently being gathered. An example of an agri-food project in epidemiology was presented. This demonstrated the range of information and data that needed to be collected, including: population demographics and statistical data; research articles in journals, publications, newsletters and conference presentations; papers in published proceedings; chapters in books; corporate outputs; secondary or processed data; food risk assessments; project reports; and general papers and documents. This presents a number of challenges to the current workflow, in particular: data collection and sharing; limited primary data discovery and validation; time-consuming and laborious data identification and documentation; limitation in accessing relevant publications; difficulty in discovering peers with whom to collaborate and to exchange knowledge and experience; multiple databases for processed data storage; and curation and dissemination of own research outcomes.

Possible GFSP platform features are aimed at improving the curation of data with a focus on making data documentation, storage, management and sharing easier. Activities are looking to migrate existing databases into a single central data repository, improve data organisation and classification schemes and improve data curation and filtering workflows (document and store data once, feed multiple sites/access points). The focus is on zoonotic diseases reports, extending the coverage of data types and extending the coverage of data sources (include more sites with zoonoses and infectious agents). Connections with other research groups are important for exchanging knowledge and experience

The next steps include deploying and validating the demonstrator in a Global Food Safety Knowledge Hub) and piloting the setting up local GFSP demonstrators (Local Food Safety Knowledge Hubs) in capacity building projects. There will then follow the need to adapt and provide federated knowledge discovery services to existing or new projects.

The project is ambitious with international scope, bearing in mind the difficulties in just connecting data sources in Europe. It would be interesting to understand how much time went into negotiating access to

the various data sources and how difficult it was. Were institutes happy to open up primary data sources and how, within an open collaboration context, did one select which data to harvest?

The discussion focussed on the various benefits and costs of using commercial versus home developed solutions. Could one not just stick everything on Amazon and use Google? Does research leverage off these capabilities enough? Apparently the astronomy community uses Google all the time. However many of the commercial services do not have sophisticated capabilities for dealing with metadata or lack the structure required to analyse research data. There is also the issue of who preserves the data. If services are developed in-house then they must be as easy to use as current offerings and must be as stable. It is difficult to develop a trust network built on third party companies that can just walk away and many companies assert rights over the data.

2.3.3. Copyright Reform and WIPO

Ville Oksanen, Aalto University

The World Intellectual Property Organisation (WIPO) is headquartered in Geneva. It is part of the United Nations with 183 member countries. It manages copyright, patents and trademark treaties. There is a requirement for unanimity.

Members of the ‘civil society’ may participate as observers – traditionally these are organisations like IFPI, RIAA, IBU and library associations. It is easy to apply to become an observer, but only once a year¹.

For an international or national non-governmental organisation to join it needs to provide details of:

- The text of its constituent instrument (articles of incorporation, bylaws, etc);
- An indication of the date and place where it was established;
- A list of officers (showing their nationality in the case of an international NGO);
- A complete list of its national groups or members (showing their country of origin in the case of an international NGO);
- A description of the composition of the members of its governing body or bodies (including geographical distribution in the case of an international NGO);
- A statement of its objectives;
- An indication of the field or fields of intellectual property (e.g. copyright and related rights) of interest to it.

It was noteworthy that The European Commission and the CLARIN project are both observers. The WIPO has two meetings over ten days within a year. It would be useful for more observers from research organisations such as ERICs to join to push for exceptions in a new treaty on research and education that is currently being discussed.

2.3.4. Points of contact between data access/reuse policies and data preservation policies

Simon Lambert, Science and Technology Facilities Council

The talk discussed the roles played in policies which include who requires and outlines the policy, who devises and imposes the policy, who has rights or obligations under the policy and who implements the policy. ISO 16363 defines a hierarchy of requirements starting with a Repository Mission Statement which feeds into a Preservation Strategic Plan which in turn feeds into a Preservation Policy resulting in a Preservation Implementation Plan.

¹ See <http://www.wipo.int/members/en/admission/observers.html>

APARSEN Deliverable 35.1² discussed exemplar good governance structures and data policies. It defined the various actors: memory institutions, universities and research institutions and research funding agencies and policy makers. It considered a number of aspects including: governance and funding, data policy and data policy adherence, data requirements, preservation and reuse and ownership and responsibilities.

The following key points for policies needed to be considered: the links between preservation and (open) access, update mechanisms, risk assessment and contingency planning, how long to preserve and of course don't forget the Designated Community!

The discussion focussed on how to deal with data that is not part of a controlled experiment that can be repeated. For example, climate data is time dependent and we can't go back to yesterday to re-measure gas concentrations. Archaeological data is precious in that if sites are not preserved indefinitely there is no way to go back and get the data again. So there are dangers in throwing data away and mistakes will be made. Equally it is clear that projections of data volumes cannot be met by current media roadmaps so some value judgements will need to be made.

Is curation different to preservation? Curation actions on digital material need to ensure that the data is preserved for future access and re-use. Curation is the responsibility of the user / owner of the data. Preservation usually means maintaining the integrity of the bit-stream. It is difficult to link preservation to re-use as this is typically not measured. If one could track usage one could work with the different communities to identify appropriate lifetimes.

2.4. Further Ideas and Open Discussion

The question of how easy it is to label data was raised. Common labels include: "research", "personal", "sensitive", "open". The talks had identified the wide range of definitions that existed in different countries and in different initiatives. It is very difficult for an infrastructure to deal with this level of complexity. There are clearly some kinds of data that an e-infrastructure cannot/should not touch, in particular in the areas of personal data (especially where it has not been anonymised), commercial data (as this often violates terms of use of publicly-funded services) and data that impacts on state security.

Should data management be encouraged through guidelines or enforced policies? Enforcement seldom works in the university area so data is usually managed through guidelines supported by enabling activities. But even universities need to recognise that there are legal constraints such as data protection and these need to be respected.

Education and training are key roles for e-infrastructures. Researchers often say that they need training but unfortunately there is a general lack of enthusiasm to attend training workshops. Researchers, in general, are not interested in data management because its value is not tangible to them at moment. Data management has value in the long term but there needs to be more recognition that data is a valid output. The current situation seems to be that data is appreciated retrospectively – hopefully in the future it will be appreciated prospectively.

In general, institutions do not have the people or resources to support data sharing activities. Part of the problem in universities might be that the knowledge management team and the library team are in different parts of the organisation. In the US they tend to have teams that can address the technical and legal aspects of data sharing as a one-stop-shop. There is some way to go in Europe in terms educating

² Available at http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2014/06/APARSEN-REP-D35_1-01-1_0_incURN.pdf (Jan 2015).

researchers to be consumers of specialist services. At an undergraduate/graduate level it is difficult to find ways to integrate data management into the curriculum. Data management is more of a skills gap than an education gap.

Some effort needs to be put into promoting the benefits of good data management. It needs to be recognised as a requirement in research support jobs that are building the infrastructures, the supply side, and a key skill in new generations of researchers and project managers, the demand side.

3. SESSION 3: SHARING DATA: TOO MANY STICKS, NOT ENOUGH CARROTS?

Funders tend to apply sticks but they could also supply carrots. How does EUDAT as an e-infrastructure provider help to demonstrate benefits to researchers' careers rather than others benefitting from ones hard work in the future?

3.1. Funders' requests on data sharing: an overview of current practices

Angus Whyte, Data Curation Centre

The Digital Curation Centre (DCC) recognises that good research needs good data and its mission is to help to build capacity, capability and skills in data management and curation across the UK's higher education research community. DCC provides guidance and support in particular through the development of guides such as *How to Develop Research Data Management Services - a guide for HEIs*³.

The DCC addresses all components of the Research Data Management lifecycle, from the development of policy and strategy through to the development of business plans and sustainability. DCC advice covers data management planning, managing active data, data selection and handover, data repositories, data catalogues all of which require guidance training and support.

Making research data available to users is a core part of the UK Research Councils' remit and is undertaken in a variety of ways. The Research Councils are committed to transparency and to a coherent approach across the research base. These RCUK common principles on data policy provide an overarching framework for individual Research Council policies on data. The principles cover open data benefits, data management policies and plans, useful metadata, legal, commercial and ethical constraints, embargo periods, recognition and management of research data.

There is very good coverage of policy including published outputs and data, policy stipulations covering time limits, data plans, access and sharing and long-term coverage across the Research Councils. Long-term curation, monitoring and guidance have partial coverage whereas the support for repositories, data centres and costs has partial to no support from the various Councils.

The Horizon 2020 programme has developed its own Guidelines on Data Management⁴ and these build on current activities such as the Collaboration to Clarify the Costs of Curation⁵. The FOSTER: Facilitates Open Science Training for European Research⁶ – it is a pilot project in the H2020 program which seeks to develop (and update) a data management plan (DMP), deposit data in a research repository and make it possible for third parties to access, mine, exploit, reproduce and disseminate data – free of charge for

³ http://www.dcc.ac.uk/sites/default/files/documents/publications/How-to-develop-RDM-services_finalMay2013rev.pdf

⁴ http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

⁵ <http://4cproject.eu/>

⁶ <https://www.fosteropenscience.eu/content/horizon-2020-open-research-data-pilot>

any user, and provide information on the tools and instruments needed to validate the results (or provide the tools). The DMP is supported by an online tool.

What kinds of ‘sticks’ can we deploy to encourage data access and reuse? Well we can get national funders to state their policy on Open Access to research data, expect a DMP, mandate sharing in a repository, mandate that a DMP be submitted and monitor that a DMP has been submitted and data has been deposited.

What kind of ‘carrots’ can we deploy to encourage data access and reuse? Well we can get national funders to offer guidance, tools and support staff, support a data repository, state costs will be met and offer rewards through funding or assessment. The following table from the reference in footnote 3 on the previous page summarises what some EU states are doing,

- **Int’l Federation of Data Organizations for Social Science (IFDO)** Report ‘Policies for Sharing Research Data in Social Sciences and Humanities’, <http://ifdo.org/wordpress/?p=532>
- **SIM4RDM** Landscape Report
- **APARSEN D35.1**. Exemplar Good Governance Structures and Data Policies
- **January ’15 RECODE D5 Policy guidelines for open access and data dissemination and preservation** “...review relevant open access and data dissemination and preservation policies at the European and Member State level and in third countries. It will identify policy gaps...”
 - Common ○ Occasional ◊ Rare or never

	Policy	DMP	Deposit	Monitor	Guides	Repos	Costs	Reward
AT	◊	○	◊	○	○	◊	○	○
CZ	○	○	○	○	◊	◊	○	○
FI	●	●	○	●	●	●	◊	○
FR	○	○	○	○	○	◊	○	○
DE	◊	◊	◊	○	◊	●	●	○
IT	◊	○	●	○	○	●	○	○
ND	◊	◊	◊	◊	●	●	◊	○
NO	●	●	●	●	●	●	●	○
PL	●	○	○	○	○	●	○	○
ES	●	○	○	○	○	●	●	○
SE	○	◊	○	○	●	●	○	○
CH	○	○	○	○	○	○	○	○
UK	●	●	◊	◊	●	●	●	○

You can see that repositories are supported across nearly all countries and DMPs are expected in 6 of 13 countries whilst monitoring is in place in only 4 countries. Rewards are rare. In the UK it has been assessed that at a DMP was submitted to six out of seven Research Councils (RCs) and that the plan met

the quality criteria in all of these cases. However, sharing of the data from the project as planned occurred in only two of the seven RCs. In terms of specific funding streams for data sharing or reuse there is one good very good exemplar in the UK, the Biotechnology and Biology Sciences RC (BBSRC) which has a Tools and Resources Development Fund and a Bioinformatics and Biological Resources Fund.

In terms of credit the importance of data is now reflected in the UK's Research Excellence Framework 2014; for example, case studies on impact beyond academia and research environment information are now included.

As a final issue Angus considered trust and what policies build trust in a research organisation's support service. Trust certification clearly promotes trust in repositories but there is still an issue on how research organisations help local groups scale up their research data management activities. Issues here focus on the interactions between data generators and users, community support services and common data services.

The discussion explored the degree of monitoring of DMPs with only two RCs checking that the data was shared as planned. So the issue is what are they checking – availability or use. They are checking that an application was made to deposit the data in the archive – although the data may not be accepted. They are checking that the plan was implemented not that data was reused.

Continuous support of repositories is a problem, notable examples include the Chemical Data Service in the UK which is put out to periodic tender with changing requirements, and archaeology database provision. Data services can (and do!) disappear with very limited notice.

The credit for data publication is limited. More important in most countries is the credit given to the publication of papers. Holland has developed a protocol based on the strategic profile of data including relevant indicators which reward data publication and impact. In astronomy, journals encourage the publication of data and code used in its analysis. The Harvard astronomy data system queries all journals at once and uses the ADS query mechanism to explore citations of source codes for simulations, data reduction, links to publications, data-identifiers and data management. Certain disciplines are pushing this area forward – rather than national funding agencies. Perhaps we should leave it to them. But, on the other hand, we do need cross-cutting activities to meet the needs of the growing number of interdisciplinary applications such as those involved in the Grand Challenges.

Universities should regard the issue of data provenance and curation as an institutional responsibility. They should develop a Policy on Research Data Management (RDM) rather than leaving it to specific disciplines. Good RDM is a vital enabler for good research leading to high impact publications. Failure of data services or archives usually results from a lack of investment. Universities should provide a depository of last resort – not duplicate data which may be stored in project specific repositories. The problem with funding is a mismatch in the timescales between an investment and a subsequent return on investment. The funders have short time-scales, focussed on high-impact projects, whereas institutions and societies have a long-term interest in sustainability.

3.2. Incentives: what works, what doesn't

Libby Bishop, UKDA/Knowledge Exchange

The talk explored the incentives for data sharing. The barriers to data sharing are well known and there is a wide variation in data sharing policies across Europe. Where policies are weak or not present, most countries rely on norms and incentives. The overall benefits of data sharing are clear, but the benefits

for individual researchers can be weak or mixed. Incentives provide a better basis for data / research collaboration.

The Knowledge Exchange undertook a study of incentives in the period March-June 2014⁷. They explored five case studies that involved active data sharing in five European countries, in five disciplines. Some twenty-two researchers were interviewed ranging from PhD students to senior academics. Areas explored included: research, data, sharing practices, motivations, optimal times, barriers and future incentives.

The interviews described diverse modes of data sharing which included: **private management** (sharing within research group), **collaborative sharing** (within a consortium), **peer exchange** (sharing in informal networks), **transparent governance** (sharing with external parties for accountability, research assessment, scrutiny and inspection), **community sharing** (with research community members) and **public sharing** (with any member of the public).

The data sharing practices in case studies identified that data sharing is part of the scientific process, supporting collaborative research, exchange of data between peers and providing supplementary data for publications. There is evidence of sharing of raw data in early research, sharing of processed data at the time of publication and well-established data sharing practices in some disciplines, for example crystallography and genetics. A number of community-topical databases have been enabled and there is some sharing through public repositories, in particular in the areas of chemistry, ethnography and biology.

If we invest in incentives there are direct benefits for research itself, for example in the collaborative analysis of complex data, methods learning, research that depends on data/information and data mining, supplementary data as evidence for publications and the role of research in creating data resources. Researchers may benefit from the increased visibility of the researchers/research group, reciprocity in terms of sharing data, reassurance (e.g. being invited to share, thus supporting the development of the discipline) and the generation of better science

In terms of incentives – we must understand the discipline norms. These days these do include sharing by the default in the research domain, research group or within institutions. Wider openness benefits research, but individual researchers are reluctant to take a lead.

There are also a number of external drivers from external funders, journals expecting data citation and deposition, learned societies developing the appropriate infrastructure, the growth of data support services, and the evolving publisher and funder policies and expectations. The latter may not push data sharing as much as they could do, e.g. supplementary data in journals can be of poor quality; mandated repository deposits tend to be minimal or exclude valuable data; and they have slowly changing general attitudes, practices and norms.

Looking to the future the various stakeholders could provide incentives for researchers that encompass a level playing field for sharing, direct funding for RDM support, student training in data sharing, infrastructure and standards, sharing failed experiments, micro-publishing/micro-citation and broader norms.

The study proposed a number of recommendations with a principal focus on changing norms and encouraging direct benefits. Leadership should be encouraged from funders, institutions, learned societies and publishers to develop a “mixed economy” of incentives that consider the phase in the research data life cycle, the career stage of the researcher and the context of the discipline / research

⁷ See <http://www.data-archive.ac.uk/about/projects/incentive>

environment. At the European level it was recommended that there be investment in ‘rich’ data resources: the ‘data + context’ model.

The discussion focussed on how we get researchers to agree to share data. The usual approach was for them to agree to cooperate on sharing then when the project was funded they could then disagree. Informal peer sharing within a Consortium is probably the most productive model; otherwise, if there are competing claims, then copyrighting or licensing may be appropriate. Within Norway it was noted that embargo periods can provide an incentive to data sharing. An emerging theme is that researchers want more control on access to their data – especially early researchers.

Another issue that was raised was that of a retention policy – after all there is no point spending money on data that is not used. This is a difficult issue: we can’t keep everything for ever even if this avoids difficult value judgements as to what collections will be valuable in the future. This can only be informed by the users who can express value judgements that must be based on access and actual or perceived impact. One high impact access could well be better, in terms of sustaining the research activity, than twenty undergraduate downloads. So the emphasis should be on establishing both the quality and quantity of accesses.

3.3. Funding RD infrastructures: initial results

Irina Kupianen, CSC/KE/Science Europe

Irina presented the current status of a report on the funding of research data infrastructures. The findings were presented in-confidence as a final version of the report has yet to be agreed. The report from Knowledge Exchange and Science Europe is due for publication in early 2015.

4. DARUP ACTIVITIES FOR 2015

The workshop concluded with an update on some of the activities within EUDAT that were prompted by the first DARUP workshop in 2013, and a look ahead to where DARUP and EUDAT might go next.

4.1. The EUDAT Licensing Wizard – a demo

Pawel Kamocki, IDS Mannheim

The first DARUP workshop highlighted the issue of licensing and suggested that a tool might be developed to make it easier for researchers to select an appropriate one. This is particularly important to the EUDAT B2SHARE service where researchers want to deposit and share data. They want to understand which licence to choose and which repositories support which licences.

Pawel demonstrated the features of a licensing wizard that is currently being developed and which includes the selection of a specific licence for both data and software, taking into account their current restrictions such as copyright or third party rights. The wizard checks that the licence selected is consistent with the current restrictions. The current source code is available through Github.

The attendees very much welcomed the development of the wizard and looked forward to it becoming available as part of the EUDAT collaborative data infrastructure. The wizard provided guidance on the differences between licences. It was suggested that by providing information to researchers about all possibilities this may encourage them to be too protective. The guidance could be complemented perhaps by a statement about what EUDAT encourages, for example, if you want to share then use this licence. Most users want a process that allows them to deposit data easily and quickly. This could be

implemented through a tick-box. Of course it depends on who the depositor is – is it an institute a data developer or a user? Users and their institutes may have different views on licences.

The EU JRC is working on something similar to include into a DMP before a project is approved. The first question that the user has to answer is – can you apply an EU copyright notice? Researchers have to assess whether this is applicable or not. If the copyright notice is not applicable then they proceed to other options. It is important to keep explanatory text to a minimum otherwise researchers will not complete the process. The emphasis is on trying to reduce the amount of analysis and decision making required but it is difficult to distil the analysis to simple yes/no questions which can be automated.

It is difficult to change attitudes with an automatic tool, and there are traps, in particular, litigation pitfalls. Some people will just flick through any wizard following the path of least resistance. Perhaps the text explanations should be coupled to access to a helpdesk answering queries, but this is very resource intensive. It was suggested that licence selection should be led by the institutions as, at the end of the day, they are usually the data owners and they invest either directly or indirectly in the infrastructure to preserve it. Institutions are meant to provide basic information as part of the research process and provide tools to enable data management. Researchers want a dishwasher rather than instructions on how to wash dishes. It would be useful if we could automate the whole data upload service so that there was no need for human intervention between collecting data and it being uploaded into a repository.

4.2. Towards data Policy Harmonisation

Rob Baxter, Edinburgh Parallel Computing Centre

The first DARUP workshop addressed the question: What policies, across the national divides, should EUDAT adopt for data deposit & sharing? An approach was suggested that involved reviewing existing EUDAT partners' approaches and identifying a common approach as "unifying principles" for all EUDAT participants. The approach was modified somewhat in the interim, to undertake an internal survey with reference to open data – a traditional "bottom-up" approach and to review external data policies – a more "top-down". The former has been completed whereas the latter has not progressed as quickly.

Four sources have been used to compare external data policies.

- Sheldon/APARSEN
 - Sheldon 2013: *Analysis of Current Digital Preservation Policies*
 - APARSEN 2013: *Exemplar Good Governance Structures and Data Policies*
- DSA Guidelines
 - <http://datasealofapproval.org/en/information/guidelines/>
- GEOSS DS principles, GEO DM principles (draft)
 - <https://www.earthobservations.org/index.php>
- The University of Edinburgh "10 Commandments"
 - <http://www.ed.ac.uk/schools-departments/information-services/about/policies-and-regulations/research-data-policy>

The criteria against which these sources were compared included: Access and Use, Accessioning and Ingest, Audit, Bibliography, Collaboration, Content Scope, Glossary/Terminology, Mandates, Metadata, Documentation, Policy/Strategy Review, Preservation Model/Strategy, Preservation Planning, Rights and Restriction Management, Roles and Responsibilities, Security Management, Selection/Appraisal, Staff Training/Education, Storage, Duplication and Backup and Sustainability Planning.

The suggestion was made that the comparison could well be strengthened by considering the RDA Legal Interoperability guidelines on sharing, the JRC principles and possibly other activities.

The final results of this exercise, including conclusions and a distilled set of potential policy guidelines for EUDAT to follow, will be included in the final EUDAT work package 2 deliverable, *D2.1.3 Sustainability Plan (final)*, due for release April 2015.

4.3. How can we make best use of DARUP?

This question has to be answered in the context of what is EUDAT's mission.

EUDAT is an umbrella organisation that operates data storage and manages trusted services and promotes best practice in the field. So the following questions arise. Which policy areas should EUDAT stick its nose into? Should it guide or prescribe? Should it capture policy in SLAs? What about policy monitoring? Enforcement? So is EUDAT just a network of centres with common technology and services or is it more?

To what extent should traditional research institutions be seen as pillars of the data policy development and delivery along the lines of the Australian NDS which was built from university foundations, and how does this fit into the broader European landscape? Universities across Europe have a clear mandate to support research and data management is an important service. Holland has a data management policy that operates at a number of levels identifying various roles and responsibilities from boards at universities that secure funding to researchers that need to look after data.

It was commented that repositories are good at sharing current data but may not be so good at preserving it given current business models. Both functions require a significant investment in services. The issue of trust is paramount. For example, if I use B2-SHARE – can I throw away my original data file or is it just a mechanism for sharing data. Most researchers are not comfortable with throwing away their only remaining copy. The ability to share and ensure data safety underpins everything. Are more copies better? EUDAT should take a lead in defining/guiding policy in this area rather than abdicating responsibility and leaving it to others to develop best practice because it is just too difficult.

We have to recognise the environment within which EUDAT partners exist. There is a mixture of public/private organisations which have limited opportunities to influence regulatory or policy frameworks. The EUDAT partners can identify certain trends and foster them and can seek to promote common themes and reduce discrepancies across different institutions and countries. EUDAT is well placed to spot differences and similarities and spot things that hinder development of an infrastructure. EUDAT is an umbrella organisation.

EUDAT has a key role in driving interoperability in terms of the services that it provides and, in particular, in terms of licences. EUDAT with its open access mission needs to minimise the complexity of choice in terms of licences and minimise the barriers to the selection of licences through automation, to enable easy data deposition.

The question was raised as to whether we are worrying too much about data management. It will sort itself out – is data management just a glorified notebook? Do researchers need much help? As discussed earlier data management is a skill in doing research but research dissemination through publications are credited more. So EUDAT needs to provide more value than that of a glorified notebook – it needs to provide different ways of doing research facilitated by the infrastructure. The key research drivers are multidisciplinary research, with an explosion of data in individual disciplines needing to be integrated to meet the needs of the Grand Challenges.

There was a call for action, EUDAT has a responsibility here to do whatever it can to support open data sharing. If we don't then the world may muddle through but we have the skills that can make a difference to support new scientific endeavours. Perhaps we should work on a manual for data management that consolidates the best practice from the past twenty years and informs the future agenda. At the end of the day we want to make things simpler for researchers. We do not want to make things even more complicated. We want to support the advancement of science.

The workshop concluded that yes, EUDAT should be involved in policy. But its key role is to enable the next generation of research whilst it is being overwhelmed by data, technology and cultural changes.

ANNEX A. LIST OF PARTICIPANTS

1. **Mark Allen**, University of Strasbourg (France).
2. **Rob Baxter**, Sustainability Lead, EUDAT; EPCC, University of Edinburgh (UK).
3. **Magchiel Bijsterbosch**; SURFsara (The Netherlands).
4. **Libby Bishop**; UK Data Archive (UK).
5. **Richard Blake**, EUDAT; Science and Technology Facilities Council (UK).
6. **Catherine Doldirina**, Researcher, GEOSS; JRC-EC ISPRA (Italy).
7. **Daniela Duca**; JISC (UK).
8. **Giuseppe Fiameni**, CINECA (Italy).
9. **Carmela Freda**; INGV (Italy).
10. **Carl Johan Håkansson**; KTH Stockholm (Sweden).
11. **Margareta Hellstrom**; Scientific Coordinator, ICOS; Lund University (Sweden).
12. **Pawel Kamocki**; Legal Expert, IDS Mannheim (Germany).
13. **Pirjo Kontkanen**; University of Helsinki (Finland).
14. **Irina Kupiainen**, CSC – IT Centre for Science (Finland).
15. **Simon Lambert**; STFC (UK).
16. **Damien Lecarpentier**; Project Director, EUDAT; CSC – IT Centre for Science (Finland).
17. **Wouter Los**; University of Amsterdam (The Netherlands).
18. **Nikos Manouselis**; Agro-Know (Greece).
19. **Jens Nieschulze**; University of Göttingen (Germany).
20. **Ville Oksanen**; Law and Technology Specialist, CLARIN-ERIC; Aalto University (Finland).
21. **Maria Reh binder**; Aalto University (Finland).
22. **Ton Smeele**; Utrecht University (The Netherlands).
23. **Herman Stehouwer**; Max-Planck-Gesellschaft (Germany).
24. **Charalampos (Babis) Thanopoulos**; Agro-Know (Greece).
25. **Angus Whyte**; Digital Curation Centre (UK).
26. **Max Wilkinson**, VPH; Head of Research Data Services, University College London (UK).