RDA EUROPE IRISH NATIONAL WORKSHOP
IRELAND, 8TH SEPT 2017 REPORT

Summary:
On Friday 8th September 2017, The National Library of Ireland, the Digital Repository of Ireland, the Research Data Alliance (RDA) and Open Research Ireland held a workshop with the Irish National Open Research Forum. The workshop was supported by the Research Data Alliance Europe (RDA Europe) project.

The main objective of the workshop was to provide international, European and national perspectives on open research as well as focus on the developments in terms of Open Research in Ireland. In particular, the workshop was designed to provide constructive feedback and contributions on the draft principles on open research data generated by the stakeholders participating in the National Open Research Forum (NORF) Working Group (WG) on the development of open principles for research in Ireland. A further objective was to illustrate initiatives that could provide models and best practices for future Irish initiatives and activities.

The main output of the workshop was the production of a detailed report with feedback and contributions, including concrete actions, on the Principles for Open Research in Ireland. This document will be revised by the NORF WG and circulated to a series of stakeholders for further consultation. The final version is expected to be distributed in late 2017 / early 2018 and will be widely disseminated in Ireland and via RDA.

The programme committee was formed of Sandra Collins, National Library of Ireland, Natalie Harrower, Royal Irish Academy – Digital Repository of Ireland, Brain Galvin – Health Research Board and Hilary Hanahoe, Research Data Alliance.
Table of Contents
Summary: .......................................................................................................................... 1
Table of Contents ............................................................................................................ 2
Overview .......................................................................................................................... 2
Agenda & Presentations ................................................................................................. 3
International, European and National perspectives on Research Data .......................... 3
World Café Overview .................................................................................................... 8
World Café Schedule ..................................................................................................... 8
World Café Breakout Sessions ....................................................................................... 9
Material Distributed ....................................................................................................... 9
Conclusions ..................................................................................................................... 9
Annex 1: Open Research in Ireland Draft Principles ....................................................... 10
Annex 2: Workshop Participants ................................................................................... 19
Annex 3: A National Approach to Open Data in Ireland agenda ..................................... 20

Overview
On Friday 8th September 2017, The National Library of Ireland, the Digital Repository of Ireland, the Research Data Alliance (RDA) and Open Research Ireland held a workshop with the Irish National Open Research Forum. The workshop was supported by the Research Data Alliance Europe (RDA Europe) project and was one of the series of National RDA events.

Entitled “National Approach to Open Research Data”, the purpose of the workshop was to collectively develop proposed national principles for Open Research Data in Ireland, building upon the draft principles created by the National Open Research Forum. Topics addressed included skills & training, infrastructure, incentives & rewards, priorities, policy & practice, communicating & engaging with the research community.

The workshop was divided into two distinct sessions, the morning focusing on setting the scene and informing the invited participants on perspectives on Research Data from a series of international experts present.

The afternoon session was dedicated to the discussion and feedback on the draft principles for a National Approach to Open Research Data in Ireland. The afternoon session followed the World Café hands-on workshop format with participants addressing specific questions in small groups, round-table style to support the definition of a set of concrete action points for implementing the principles in Ireland leveraging on national as well as international activities.

Attendance to the workshop was by invitation only and over 100 data experts in Ireland were invited to attend. Over 60 Irish stakeholders, representing the major funding bodies, academic organisations, libraries and archives as well as research and development agencies, registered to attend the event with 58 people in attendance on the day.
Agenda & Presentations

10.00 – 10.10 Sandra Collins, National Library of Ireland: Welcome and goal of workshop

10.10 – 10.30 Hilary Hanahoe, Research Data Alliance: Overview of Research Data Alliance (presentation)

10.30 – 10.50 Birgit Schmidt, University of Göttingen: RDA for Libraries from an International Perspective (presentation)

10.50 – 11.10 Sarah Jones, Digital Curation Centre UK: National Data Management Approaches (presentation)

11.10 – 11.30 Coffee Break

11.30 – 11.50 Ingrid Dillo, Data Archiving and Networked Services (DANS): Trusted Repositories for Open Research Data (presentation)

11.50 – 12.10 Natalie Harrower, Digital Repository of Ireland: Ireland and RDA (presentation)

12.10 – 12.30 Rebecca Grant, Springer Nature: Publishers and RDM (presentation)

12.30 – 12:50 Kate Kelly, Royal College of Surgeons in Ireland (RCSI): Open Research Ireland (presentation)

12.50 – 13.30 Lunch (provided)

13.30 – 17.00 World Café

International, European and National perspectives on Research Data

Following the warm welcome from the director of the National Library of Ireland, RDA Council member and workshop host – Dr Sandra Collins – the opening presentation covered the international landscape from the Research Data Alliance. Hilary Hanahoe, former Communications Manager for RDA, outlined the alliance, its vision, mission and achievements to date. She provided the participants with an understanding of how RDA could support them with their data challenges, how they could contribute and gain from involvement in RDA as a member and how the solutions produced by RDA to date are being implemented to achieve the RDA vision. “Researchers and innovators openly sharing data across technologies, disciplines and countries to address the grand challenges of society”. Some concrete examples of
implementable RDA solutions\(^1\), also including some recognised as ICT technical specifications (standards) in Europe, were illustrated. The **Scalable Dynamic-data Citation Methodology** supports accurate citation of data subjected to change, for the efficient processing of data and linking from publications. This solution, produced by the RDA Data Citation Working Group, has 10 documented cases of adoption including domain specific examples in environment, forestry, oceanography, biomedicine. Other examples included the **Legal Interoperability of Research Data Principles and Implementation Guidelines** and the **23 Things: Libraries for Research Data**, an overview of practical, free, online resources and tools that users can immediately take advantage of to incorporate research data management into the practice of librarianship. This example transitioned well to the next presenter from the Electronic Publishing unit at Göttingen State and University Library.

Dr. Birgit Schmidt, University of Göttingen, in her presentation on **RDA for Libraries from an International Perspective**, delved deeper into the examples of adoption of RDA solutions in the world of libraries. Her main reasons why libraries engage with research data are that Research data is core to the records of science; data management is not the sole responsibility of researchers and that support by information professionals is essential for securing long-term access and re-use. She illustrated her own organisation’s experience and gain from involvement in the RDA. In her words, "RDA offers members international communities of practice across professions (research, data centres, libraries, publishers, etc.). It provides a problem-oriented approach targeting core data management challenges and a platform for exchange and feedback on approaches and standards, inspiration & opportunities for collaboration." Birgit is one of the co-chairs of the Libraries for Research Data (L4RD) Interest Group in RDA, the same group responsible for the highly successful **23 Things: Libraries for Research Data**\(^2\).

Birgit also gave insights to the European Open Science Agenda and the European Open Science Cloud (EOSC) activities currently being taken forward in Europe. The agenda has eight top level ambitions of which four were highlighted as of interest to the audience and those interested in the use and management of research results and data.

Birgit gave excellent details on the current EOSC pilot project and expected services to emerge from that 24-month project (start date Jan 2017).

Moving from international and European perspectives, Sarah Jones from the Digital Curation Centre UK provided a fascinating overview of a series of different **National Data Management Approaches** as food for thought for the

---

\(^1\) rd-alliance.org/recommendations-and-outputs/all-recommendations-and-outputs

\(^2\) http://dx.doi.org/10.15497/RDA00005
Irish participants. Sarah’s presentation included a huge selection of initiatives and activities that could be leveraged upon as both examples and stepping stones in the creation of a Research Data Management broker or service provider. Some examples include the Tuuli Data Management Plan (DMP)\(^3\) case study an interesting example of a national level initiative supported by the Finnish Ministry of Education with the creation of a common DM template for funders and academic institutes in Finland through a single national service.

As a potential national portal example, Research Data Australia\(^4\) delivers both a portal and a data catalogue to access Australian research data. The portal harvests metadata from 100+ research organisations, government agencies, and cultural institutions. It has also provided funding for projects within institutions to ‘seed the commons’.

One example of training support was the UK based Mantra\(^5\) online portal which, when designed, kept in mind flexibility and other potential uses. It has been used in its original state but has also inspired many variants and been embedded in Virtual Learning Environments (VLEs). The result of seeding other work has led to cost effectiveness.

Sarah also provided examples of (SPAAR EU) data champions programme, National centres of expertise / competence e.g. DCC, ANDS with focused services …, agencies like JISC (UK) that are managing Research Data Funding programmes. She also covered membership (subscription) networks, e.g. Canadian Portage network\(^6\), which while focused on library community have been very careful to collaborate with others. A good example of a formal coordination / strategic board to facilitate national strategy can be found in the Dutch National RDM Coordination Point\(^7\) which is a formal coordination to facilitate national strategy through five working groups that have members from universities of applied sciences, data centres, national services and other institutes. The world of research data management will always be a hybrid environment so collaboration is key. Lessons from these and other initiatives for Irish endeavours are that:

- Many services can be offered at a national level
- It helps to build on good practice emerging from universities or other countries
- Understand community needs and priorities
- Tackle one aspect at a time but keep the broader vision in mind (e.g. Research Data Shared Service Vision) and work in a stepwise fashion.

She concluded with a very relevant and useful list of questions, not just for the Irish stakeholders but that could be essential for any national endeavour in its nascent or indeed planning stages:

<table>
<thead>
<tr>
<th>Do you have all the stakeholders at the table?</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ What resources do you have to commit?</td>
</tr>
<tr>
<td>✓ Where are things at now?</td>
</tr>
</tbody>
</table>

\(^3\) https://www.dmptuuli.fi
\(^4\) https://researchdata.ands.org.au
\(^5\) http://mantra.edina.ac.uk
\(^6\) portagenetwork.ca
\(^7\) www.surf.nl/en/lcrdm
Ingrid Dillo from the Data Archiving and Networked Services (DANS), The Netherlands and ad-interim RDA Secretary General provided an extensive overview to the challenge of creating and maintaining Trusted Repositories for (Open) Research Data. Based on years of collaboration and intense activities the CoreTrustSeal® Data Repository certification service was launched in September 2017. It is “the culmination of a cooperative effort between DSA and WDS under the umbrella of the Research Data Alliance to merge their data repositories certifications”. Certification efforts are necessary to:

- build stakeholder confidence in the repository (funders, publishers, etc.);
- raise awareness about digital preservation;
- improve communication within the repository;
- improve repository processes;
- ensure transparency; and
- differentiate the repository from others.

Ingrid illustrated examples of certified repositories to date, including her own organisation DANS. And pointed out to the audience that “ensuring long-term accessibility of open research data through certified repositories” should be included the current draft of the Open Research Data Principles.

Moving to a more national view, Natalie Harrower, Director of Digital Repository of Ireland, illustrated the collaboration between DRI and the RDA and how that has impacted on Ireland and Irish stakeholders to date. Ireland’s involvement in the RDA goes back to its launch in March 2013 and DRI was one of the hosts of the RDA 3rd Plenary meeting in March 2014. As a result, many Irish stakeholders joined as members and DRI and NLI became partners of the RDA Europe initiative, giving them an official mandate to promote and raise awareness of RDA in Ireland and of Ireland in RDA.

Natalie also outlined the RDA outputs that have been adopted to date by the Digital Repository of Ireland, namely:

1. the Data Foundation Terminology (DFT) Core data model (PIIDs on publication, widely accepted metadata standards, certification of TDR, checksums, and FAIR data compliant)\(^8\)
2. DSA/WDS integrated approach to repository certification (2017)\(^9\)

---

\(^8\) [https://www.coretrustseal.org/](https://www.coretrustseal.org/)
\(^9\) [http://dx.doi.org/10.15497/06825049-8CA4-40BD-BCAF-DE9FoEA2FADF](http://dx.doi.org/10.15497/06825049-8CA4-40BD-BCAF-DE9FoEA2FADF)
3. Metadata Standards Directory (eg. Dublin Core)\textsuperscript{11}
4. RDA/WDS Publishing data (in progress)\textsuperscript{12}

As a partner of the current RDA Europe project (funded by the European Commission) and a future partner of the next contract (due to commence on 1 March 2018), DRI is considerably involved in the successful promotion of RDA in Europe and raising awareness on the benefits of collaboration with RDA, particularly with the Social Science stakeholders. The four key aspects of DRI’s mandate that contribute to this are: 1. Trusted digital repository – long term data preservation, access, reuse; 2. Education & Outreach; 3. Research active: projects, technologies, publications; 4. Advocacy, community building, policy input.

Rebecca Grant, Research Data Manager, Springer Nature (and former DRI and NLI employee), provided insight from the publisher point-of-view. She outlined a series of steps that Springer Nature are taking to support open research and particularly researchers in their research data management challenges.

Activities that Springer Nature are working on cover:

- Content types e.g. data articles and journals
- Credit and incentives e.g. data citation and data articles
- Encouraging reuse e.g. open licences
- Data quality e.g. data peer review
- Data discoverability e.g. linking data to publications; supporting repositories
- Raising awareness e.g. editorials, outreach
- Guidance and policy e.g. information for authors, policy harmonisation
- Technology e.g. platform developments, repository integration and other partnerships

Springer Nature has recently proposed a new RDA Interest Group (IG) on Data Standardisation and Implementation\textsuperscript{13} to help define common frameworks for research data policy allowing for different levels of commitment and requirements and disciplinary differences that could be agreed by multiple stakeholders.

\textsuperscript{10} https://www.rd-alliance.org/group/repository-audit-and-certification-dsa%E2%80%93wds-partnership-wg/outcomes/dsa-wds-partnership
\textsuperscript{11} http://rd-alliance.github.io/metadata-directory/
\textsuperscript{12} http://dx.doi.org/10.15497/RDA00004
\textsuperscript{13} https://www.rd-alliance.org/groups/data-policy-standardisation-and-implementation
The final presentation of the morning session from Kate Kelly, Royal College of Surgeons in Ireland (RCSI) focused on Open Research Ireland (ORI). Kate outlined that ORI was established to seek more defined role in implementation of Open Science in Ireland and gave insight and details to the work that is being executed by the different stakeholders and working groups involved to achieve their aims. The specific NORF working group on drafting Open Data principles for Ireland, which was the main focus of the afternoon World Café session, is only one of the six that are focusing on different areas to achieve the ORI aims.

In specific reference to the Librarians and libraries as stakeholders, she affirmed that they are very much part of the Open Science infrastructure in Ireland and in a strong position to play the “enabling” role within research institutions and wider community. Achieving the aims will require consistency from funders as well as a national approach to data management plans. This presentation wrapped up the morning session and paved the way for the afternoon discussions in World Café style.

World Café Overview
The purpose of the World Café is to discuss and provide feedback on the proposed national principles for Open Research Data in Ireland, building upon the work to date by the National Open Research Forum. The World Café will support the definition of a set of concrete action points for implementing the principles in Ireland leveraging on national as well as international activities.

World Café Schedule
13:30 – 13:45: Introduction to World Café objectives & guidelines – Hilary Hanahoe, Moderator
13:45 – 15:00: World Café Session 1 breakout (60 minutes) & report back (15 minutes) – Table Hosts & Guests
15:00 – 15:30: Networking Coffee / Tea break
15:30 – 16:45: World Café Session 2 breakout (60 minutes) & report back (15 minutes) – Table Hosts & Guests
16:45 – 17:00: Conclusions and next steps – Hilary Hanahoe
World Café Breakout Sessions

Based on the “National Open Research Forum: Principles for Open Research in Ireland” that were shared with all participants prior to the event, the aim of the afternoon “World Café” session was to get:

1. **Eventual feedback and contributions for the principles, to improve them (rather than beginning the conversation from a ‘blank page’).**

2. **A set of actions / steps for implementing the principles in Ireland**

The World Café was structured to allow the participants to discuss the main questions in small groups, facilitated by hosts. Table hosts were volunteers from the NORF working group that drafted the principles document. There were 7 tables for breakout discussions and each table discussed and reported back on the same set of questions.

The first session focused on answering: **Do the “Principles for Open Research in Ireland” reflect your (organisation, institute, community, etc.) view of what is important?** Providing practical and implementable feedback to the NORF WG on the draft principles.

The second session focus was on practical next steps with participants answering the following questions:

- **What does Ireland need to do next?**
  - What needs to be done in Ireland to achieve the implementation of these principles?
  - What can you (your organisation, community, institution, etc.) do to support the implementation of these principles?
  - How can RDA or other international initiatives help?

Material Distributed

All participants received host organisation, RDA and RDA Europe material highlighting the outputs and recommendations to date, the workshop agenda, draft open principles and sources document upon registration.

Conclusions

The participants actively engaged throughout the day and particularly in the afternoon session where they all provided contributions to the discussions. This wealth of feedback and contributions from the afternoon discussion was the subject of a detailed report that was delivered to the NORF WG at the beginning of October 2017 and has been used to generate the pre-final version of the principles.

The workshop model was positively received and is an excellent example of how to combine raising awareness of RDA and related activities with a pragmatic and successful approach to collecting input and contributions from experts. Having all stakeholders represented at the workshop meant that the feedback from libraries, funders, research and academic organisations was part of a collaborative and collective effort with concrete outputs to ensure a balanced and representative set of principles.
Annex 1: Open Research in Ireland Draft Principles
As circulated to all participants on 5th September 2017

National Open Research Forum: Principles for Open Research in Ireland

1. Open research data improves the quality of research and innovation

2. Publicly-funded knowledge is used to benefit the whole of society and create public value

3. Research data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control

4. Open Research must be supported by extensive capacity development of researchers, librarians and data specialists

5. Good data management is fundamental to high quality research and should be established early on in the research process primarily by use of a Data Management Plan.

6. Research Data should be as open as possible, as closed as necessary

7. Every effort should be taken to increase the interoperability of research data within and across domain boundaries and, to the greatest extent reasonable, to enable direct re-use by intelligent machine agents.
1. Open research data improves the quality of research and innovation

Open research data allows access for all to view, check and make use of research data which might otherwise be difficult to access, unclear in nature, or completely unknown to other researchers around the world. There are clear benefits from this transparency to the quality, integrity and reach of research and so to the broader research environment.

Open research data should act as a further self-check mechanism, improving the quality and efficiency of research:

- It enables others to check the accuracy of the data
- It allows others to analyse and assess the data for themselves, perhaps verifying the conclusions or perhaps coming to different conclusions from the same data.
- It prevents others from needing to gather identical data, allowing greater efficiency in research spending and the global research environment
- If shared in an easily interoperable way, complying with FAIR principles, it allows others to be able to combine open data with other similar data to improve the strength of evidence in any conclusions.
- It encourages the reuse of data in unforeseen ways to further research and understanding in other fields of study.

Research Data being published through open repositories brings a greater transparency to the whole research process:

- It enables data to be published, then cited and acknowledged in its own right, whether linked to a published paper or not
- It encourages and enables an open research data culture with more scope for the sharing of unsurprising and negative results which might otherwise remain unknown

Though there may be understandable reasons for withholding some details from a research data set or requiring a period of exclusive access, open research data ultimately brings benefits to the wider research environment, stimulating innovation and economic growth.
2. Publicly-funded knowledge is used to benefit the whole of society and create public value

Open Research is built on the ideals of openness, the free exchange of information and equality of opportunity in the creation and use of knowledge, ideals that are essential to the notion of public value. A society that supports, encourages and requires researchers to make their research output, and the data on which it based, publicly available, can be said to have created greater public value than one that does not.

With the possibility of making research resources available to all, we look for greater social and economic return on public investment in research by providing access to data that can be used to create new knowledge. Much of the momentum towards greater openness of research data is driven by the potential economic benefits from data reuse. Government understanding of this potential impacts on funding policy and Europe is moving towards a funding model that will require applicants to demonstrate their capacity to manage research data in a way compatible with open research. Another driver is the gains to the research community through increased dissemination of work, career advancement and wider and more rigorous scrutiny of the data underlying published outputs.

While the 'carrot and stick' approach will help to convince researchers of the value of open data, funders and policymakers must also emphasise the generous contribution to public value that open research entails. A policy that emphasises civic responsibility will complement more orthodox economic and career-oriented incentives. While openness and fair exchange of knowledge are inherently public goods, there will also be more tangible benefits to individual members of the public through greater participation in citizen and civic research. This will in turn increase awareness and appreciation of science, itself a further public value.

Public policy experts have developed refined instruments to measure the contribution of public managers to public value and to demonstrate to the general public the social gains associated with particular actions. Public awareness of the enormous potential of open research will be needed in the difficult work of moving towards a policy supportive of the open research agenda. While knowledge of the economic benefits will be an important part of this awareness it will be necessary to emphasise the social and political gains resulting from greater access to research data. If these benefits are expressed using the language of public value then the contribution of open research to society will understood better and increase the likelihood of public support.
3. Research data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control

Protocols should be developed on whether, when and how data that may be commercially sensitive should be made openly accessible taking account of the weight and nature of contributions to the funding of collaborative research projects and providing an appropriate balance between openness and commercial incentives. Any restrictions must be sound and justifiable and should not deter companies from collaborating with universities and other research organisations.

Creators
- Data access arrangements should consider the applicability of copyright, trade secret, patent, database rights or of other intellectual property laws that may be relevant.
- To promote data reuse the owners of rights should mark the data clearly with associated permissions.
- Researchers who generate original data must have reasonable right of exclusive first use for an appropriate and well-defined period which may vary by subject and disciplinary area. Any period of exclusive first use should be considered at project planning and included in the Data Management Plan (Concordat 4)
- Data supporting publications should be accessible by the publication date and in a citeable form.

Funders
- Research funders should provide guidance on sharing restrictions based around intellectual property as part of the application process.

Users
- Use of others’ data should always conform to legal, ethical and regulatory frameworks including appropriate acknowledgement.
- In order to recognise the intellectual contributions of researchers who generate, preserve and share key research datasets, all users of research data should acknowledge the sources of their data and abide by the terms and conditions under which they are accessed.

Publishers
- Publishers should enable the formal citation of data in articles to support attribution and other terms and conditions that may be applied to the access and use of data by data creators.
- In line with the practical limits for the field of research and as a condition of publication, scientific journals should require data on which the argument of the article depends to be accessible, assessable, usable and traceable through information in the article.
- The article should indicate when and under what conditions the data will be available for others to access.
4. Open Research must be supported by extensive capacity development of researchers, librarians and data specialists

The successful implementation of Open Research policies in Ireland will depend on the capacity of those involved at each stage of the research process to undertake their responsibilities with efficiency, skill and an awareness of the needs of other spheres of the research environment. The transition towards an Open Research system will require increasingly sophisticated data stewardship. Education and training will form key parts of the the infrastructure supporting researchers and data managers through the stages of idea generation, funding application, project completion, publication and impact measurement and evaluation.

Researchers will need to use new protocols and templates to manage data in a way that is compatible with Open Research. Stewarding research data will need knowledge of, and technical skill in, the areas of archiving, curating and other information management capacities. Librarians and other data managers apply these skills in their routine work and will play an important role in helping to develop these skills in researchers. Librarians will need more advanced training in these areas and there should be learning opportunities through collaborative training and project management with researchers.

Facilitating reuse of data will need increased awareness of licensing, intellectual property, data protection and ethical aspects of data publishing. As funders increase emphasis on open data, researchers will need to understand how this affects the progress of their work and other aspects of research administration.

For open research to become a reality researchers, the creators of data, must themselves become more data aware. Increasing data awareness will encourage researchers to make their research workflows open, allowing the sharing and reuse of research outputs. Educating the research community, policy makers and consumers of research outputs is an essential part of creating an Open Research environment. Raising awareness of the benefits of open research and open data is work that needs to be done alongside the development of services, training of data stewards and the establishment of incentives and recognition mechanisms. A national data management plan must include provision to train and support advocates and educators for Open Research.

An open research environment will require more open sharing of research outputs and librarians’ knowledge of publications and publishing practices will be valuable in ensuring that this is done in a sustainable manner and in the researcher’s best interests. As they are active at each stage of the research cycle, and will become equally engaged in the data cycle, librarians are ideally placed to work as advocates and educators for Open Science. Early career librarians will need to have the opportunity to move into data-related professions including: data analysts, data managers, data curators and data librarians.

Data skills will be developed through self direct learning but most learning will happen through specialised tuition and formal instruction. A national data management plan will assign responsibility for providing researchers with the learning opportunities they need to contribute to Open Research. Funders and research institutions will have the primary responsibility to provide these opportunities and will need to develop an educational and training plan to ensure data management needs are met and career paths developed.
5. Good data management is fundamental to high quality research and should be established early on in the research process primarily by use of a Data Management Plan.

Research data is the data collated by researchers in response to their research questions. It can be in many formats and may be primary data or data derived from primary sources. Open research data is research data that is freely accessible and used, modified and shared with appropriate acknowledgment. Open access to research data is an enabler of high quality research, encourages innovation and upholds good research practice while promoting and developing high levels of integrity in the research process. Data Management encourages a culture of transparency and sharing.

Good data management is essential and preparation of data management plan should be informed by the following objectives:

- Open Research data should be prepared in such a way that it is as widely usable as possible. Data should be in non-proprietary formats. If this is not possible, researchers should indicate what proprietary software is required.
- Researchers, where possible, should make their research data open and usable within a short period of time. Data supporting publications should become available in conjunction with the publication and should be in a citable form.
- Data should be made findable and accessible by lodging in an appropriate repository/web interface with a sufficient level of metadata/description. Supporting metadata should, at a minimum, provide details of how the data was collected/generated and information on the methodology used.
- Data should be correctly cited noting the intellectual contribution of the original creators. Original creators may have a limited period of privileged use of their research data in order to publish but this period of time should be specified in accordance with the norms of the discipline.
- Production of open research data should be acknowledged formally as a legitimate output of the research process and should by recognised as such by employers, funders and others in relation to career development, research assessment and research funding decisions. Data citations should be ranked as of equal important to publication citations.
6. Research Data should be as open as possible, as closed as necessary

Access to and reuse of research data should be open and unrestricted as a default rule, or otherwise be granted to users with the fewest limitations possible. There are sound reasons why the openness of research data may need to be restricted but any restrictions must be justified and justifiable.

Sharing research data poses ethical and legal questions, particularly around consent and data protection, maintaining confidentiality, guarding against unreasonable costs, managing security, commercial constraints on access, protection of endangered species, legal processes, and around the exploitation and ownership of intellectual property.

Governance arrangements must be in place to facilitate the lawful access to and reuse of research data derived from individuals, while safeguarding privacy and confidentiality. Such arrangements should draw upon well-established models and good practices for managed access to data, and should always be proportionate to the level of risk involved. They need to take full account of legal, regulatory and ethical requirements – including applicable data protection laws and relevant codes on research ethics and research integrity.

Collaborative research involving companies and with public and voluntary bodies should not be deterred by open data practices. Clear protocols on whether, when and how data that may be commercially sensitive should be made openly accessible are needed, taking account of the weight and nature of funding contributions and providing an appropriate balance between openness and commercial incentives. Policymakers should balance the legal interest and consider public interests in developing rules for access to and use of publicly generated research data.

Researchers who generate original data must have reasonable right of exclusive first use for an appropriate and well-defined period which may vary by subject and disciplinary area. Any period of exclusive first use should be considered at project planning and included in the Data Management Plan.

Attribution of research data used in any research output should be a formal acknowledgement process established by good research policy and practice, and production of open research data should be formally recognised in research evaluation.

Clarity on legal, ethical and regulatory issues and necessary exemptions to open data is essential before data is deposited in a repository. An educational process for researchers regarding rights and responsibilities in research data should be developed and adopted by relevant institutions.

Research should be conducted in such a way that others can contribute, collaborate and add to the research effort. The European Union copyright reform should modernise copyright legislation for the digital age and provide legal certainty around cross-border research activities and the use of new technologies for research and innovation. This would allow for the use text and data mining (TDM) methods to analyse and extract new insights and knowledge from vast amounts of digitally-available content. We support the sentiment that "The Right To Read Is The Right To Mine".
7. Every effort should be taken to increase the interoperability of research data within and across domain boundaries and, to the greatest extent reasonable, to enable direct re-use of intelligent machine agents

The commonly endorsed FAIR Data Principles provide high-level guidelines on how to make research data Findable, Accessible, Interoperable and Reusable. The guidelines are not a technical specification and allow for many different approaches to rendering data and services Findable, Accessible, Interoperable, to serve the ultimate goal: the reuse of valuable research objects. The FAIR Data Principles also place particular emphasis on the aspiration that, to the greatest extent reasonable, steps should be taken to enable direct re-use of research data by automated agents within a data environment.

A useful definition of interoperability is provided by the IEEE: "The ability of two or more systems or components to exchange information and to use the information that has been exchanged". Furthermore interoperability is often broken down further:

**Foundational/Functional Interoperability**
This is the most basic level of interoperability and focuses on "data exchange from one information technology system to be received by another but does not require the ability for the receiving information technology system to interpret the data". This degree of data exchange can be achieved by a myriad of standard protocols. In most cases, individual researchers do not need to concern themselves too much with this level of interoperability, although the choice of communication protocol may be practically important for extremely large datasets or computationally intensive workflows.

**Syntactic/Structural interoperability**
This level of interoperability deals with the packaging of the data via message format standards, encoding standards and file-formats. In order for a receiving system to parse and interpret the incoming information at the data-field level, it must be able to understand how that data is structured and arranged. This type of interoperability is often achieved within a particular discipline where systems and products share common interchange formats, either as the result of formal standardisation processes, informal convention, or through de-facto standards derived from commonly used software. XML, JSON, CSV are examples of common data serialisations/formats which can be used to encode information from a wide array of disciplines in such a way that they are parsable (but not necessarily semantically understood) by a wide array of software systems.

This level of interoperability will be familiar to many researchers working with data from different sources. In terms of FAIR Data, many practical steps can be taken to increase re-use of data by, for example, converting data to cross-domain, open formats (e.g. CSV rather than XLS), or ensuring an open source software package is available which is capable of parsing the particular data format in use. Even within common formats such as CSV/XLS, there are considerations about how to structure the data to make it more easily and readily re-usable by colleagues within and across-domains; for example, the simple notion arrange tabular information
as “tidy data”\textsuperscript{1} is gaining popularity as a convention particularly within the data science and statistical domains.

**Semantic interoperability**

This is the layer of interoperability describes a situation where systems are capable of interpreting not only the structure of data but also the formal meaning of that data within the particular context and are able to take appropriate actions based on that interpretation. It is the difference, for example, between correctly interpreting the string “cork” as referring to a particular city, county, geographic location, tree, bottle-top. While this degree of interoperability can be achieved by default within a closed system, making research data generally semantically interoperable, by humans or machines, within or across domains, is a far greater challenge. In common practice, human-to-human semantic interoperability can be facilitated through metadata and other data documentation; a common example of this are the semi-structured code-books or data-dictionaries associated with micro-data or survey datasets which allow the precise variables and allowed values to be interpreted. As emphasised by the F.A.I.R. guidelines this level of interoperability is optimally achieved by the use of “formal, accessible, shared, and broadly applicable language for knowledge representation”; and ideally such shared ontologies, terminologies should include qualified references to associated ontologies from other domains, allowing for semantically accurate combination and inference between disperse datasets. While this level of structured and machine-actionable interoperability is actively in use in some data-intensive disciplines, and forms part of the the core W3C standards and development plans for the world wide web, it is still an area of active research. For individual researchers, semantic interoperability should be seen as a spectrum where practical steps are taken to contextualise and document research data at a granular level to facilitate accurate re-use by colleagues in the same field, colleagues in other disciplines and, to the greatest extent possible, automated processes which seek to discover and act across large, heterogeneous datasets.

\textsuperscript{1} http://dx.doi.org/10.18637/jss.v059.i10
Annex 2: Workshop Participants

Barrett Julia
Brennan Jennifer
Brown Peter
Caffrey Maire
Callaghan Fran
Clarke Anne-Marie
Clarke Patricia
Cleere Liam
Collins Sandra
Cox John
Dillo Ingrid
Donovan John
Duggan-Walls Kay
Farragher Louise
Fox Joseph
Frost Dermot
Furlong Robbie
Galvin Brian
Grant Dolores
Grant Rebecca
Griffith Lisa
Hanahoe Hilary
Harrower Natalie
Healy Arlene
Irvine Gemma
Jennings Brendan
Jones Sarah
Kinsman Oonagh

Lawton Aoife
Long Kevin
Lynn Therese
McDonough John
McMahon Garret
Montesanti Annalisa
Mulligan William
Murphy Lisa
Murphy Orla
Nangle Sarah
Ó Carragáin Eoghan
O Farrell Sean
O'Connell David
O'Connor Stephanie
O'Dowd Niamh
O'Neill Jenny
O'Riordan Gobnait
Parsons Aengus
Quinn Ciaran
Ryan Michael
Ryan Ann
Shenton Helen
Simpson Andrew
Smith Fiona
Stack Pádraic
Swan Rónán
Tray Elizabeth
Wilson Ger
Woods Kevin
Annex 3: A National Approach to Open Data in Ireland agenda

Workshop with the Irish National Open Research Forum and a National Research Data Alliance (RDA) event.

National Library of Ireland | Seminar Room | 8th September 2017

The purpose of this workshop is to collectively develop proposed national principles for Open Research Data in Ireland, building upon, and feeding into, work by the National Open Research Forum. Topics to be addressed include: skills & training, infrastructure, incentives & rewards, priorities, policy & practice, communicating & engaging with the research community.

All presentations from the morning session are available on-line14.

Morning Session: International Experts present perspectives on Research Data

10.00 – 10.10 Sandra Collins, National Library of Ireland: Welcome and goal of workshop
10.10 – 10.30 Hilary Hanahoe, Research Data Alliance: Overview of Research Data Alliance
10.30 – 10.50 Birgit Schmidt, University of Göttingen: RDA for Libraries from an International Perspective
10.50 – 11.10 Sarah Jones, Digital Curation Centre UK: National Data Management Approaches
11.10 – 11.30 Coffee Break
11.30 – 11.50 Ingrid Dillo, Data Archiving and Networked Services (DANS), and RDA Acting Secretary: Trustworthy Repositories for Open Research Data
11.50 – 12.10 Natalie Harrower, Digital Repository of Ireland: Ireland and the RDA
12.10 – 12.30 Rebecca Grant, Springer Nature: Publishers and RDM
12.30 – 12:50 Kate Kelly, Open Research Ireland
12.50 – 13.30 Lunch (provided)

Afternoon Session: 13.30 – 17.00 World Café (facilitated by Hilary Hanahoe)

The World Café is a hands-on workshop to draft principles for a National Approach to Open Research Data in Ireland. Participants will address specific questions in small groups, round-table style.

14 http://dri.ie/presentations-national-approach-research-data-workshop-norf