RDM service development checklist

[The text below is presented on as a guide on the OpenAIRE website: https://www.openaire.eu/rdm-service-development-checklist]

Introduction:

Based on the DCC's RISE framework [https://www.dcc.ac.uk/guidance/how-guides/RISE], this checklist gives an overview of the different capabilities required to develop effective research data management (RDM) support across several different levels. The rubric indicates whether the different capabilities are recommended when developing RDM services. It takes the capabilities originally included in RISE and adds two, covering EOSC participation and support for FAIR data and services. The capabilities assessed for each level are:

- **RDM policy** (*covering the development and maintenance of RDM policy and associated documents and processes that enable its implementation*)
- **Business plans and sustainability** (*focusing on the approach to securing the sustainability of RDM services, including staff investment, technological investment, and cost modelling*)
- **Data Management Planning** (*concerning support for researchers to effectively plan the data component of their research and produce associated data management plan (DMP) documentation*)
- **Active data management** (*to do with services that enable data management, including scalability and synchronisation of services, collaboration support, and security*)
- **Access and Publishing** (*covering the support for depositing and publishing open access data*)
- **Appraisal and risk assessment** (*including processes to identify valuable data and research outputs and mitigate any associated risks*)
- **Preservation** (*addressing the need to ensure data integrity and access to data*)
- **Training** (*both developing and delivering of this to researchers and research support staff, in online and in-person formats*)
- **Advisory services** (*concerning the provision of online and in-person advice for researchers and/or support staff who need support with aspects of RDM*)
- **EOSC participation** (*whether as service provider or user of European Open Science Cloud services*)
- **FAIR assessment** (*including assessments of both datasets and FAIR enabling services*).

More detail each of these capabilities is available in the RISE framework.

Note that this resource, following on from RISE, does not explicitly address the maturity of services. It is hoped that the resources and secondary items it points to, and the references throughout to the areas where capabilities are addressed at different levels, can assist users in addressing service maturity and ensure that there is buy-in and input from different actors to work toward a more holistic service. Fundamental to this is the question, 'What do we need to

provide?', which ensures that any discussions around service development centres on the needs of the researchers and service users, and the ability of the group, organisation, or institution to meet these requirements. This also has the effect of giving the group, organisation, or institution a basis on which to formulate any plans for future service provision and developments.

Service Levels

Research Group

- ***RDM policy**:*
    - **Possibly**: When dealing directly with data, either data gathered in experimental settings or data re-used from other sources, a research group or standalone project may have policy in place detailing its approach to research data management (RDM). This could be a fairly informal set of guidelines and local practices rather than an official policy. It will more often be the case that a research group will be required to comply with institutional and/or funder policy requirements. OpenAIRE provides a useful primer for researchers on complying with Horizon 2020 mandates [https://www.openaire.eu/how-to-comply-to-h2020-mandates-for-data], while JISC has produced a guide more specifically focused on compliance with funder requirements [https://rdmtoolkit.jisc.ac.uk/plan-and-design/policy-compliance/].
- ***Business plans and sustainability**:*
    - **Possibly**: Plans around the sustainability of RDM services are often addressed at a higher position in the host organisation, most likely at Institutional or Research infrastructure level. However, a lot of research groups are set up as part of project activities and are dependent on external funding to continue. In these cases, this could be a very important aspect for some groups. The key thing is that the Research Group needs to know their funding source and, where necessary, plan for sustainability.
- ***Data Management Planning:***
    - **Yes**: Data Management Planning in general and the use of a data management plan (DMP) document enables a research group to set out what they will do with their data during a project, and what plans are in place for the data beyond the project's end. A DMP is a living document, meaning that it can be updated throughout the project's lifespan, and typically includes information on data description, data collection methods, metadata, licencing and long-term preservation among other criteria. Research groups can provide a great deal of knowledge on data management planning and workflows for specific disciplines and that this is an invaluable resource that could be tapped into by the host institution central RDM services. Communications between the two levels should be established to exploit this local knowledge and also to ensure that local RDM support aligns with institutional policies and expectations. For more information, a report from Science Europe outlines the core requirements for DMPs [link: https://www.scienceeurope.org/media/4brkxxe5/se_rdm_practical_guide_extended_final.pdf], while this OpenAIRE RDM starter kit contains further practical resources on data management planning for beginners and experts [https://www.openaire.eu/rdm-noads-starter-kit].
- ***Active data management:***

- **Yes**: Active data management at the Research Group or project level involves the use of services such as cloud computing, file storage, and data back-up. Though these services themselves are often provided at institutional level, the onus is on the Research Group to ensure that active data management procedures are in place at all stages of the research data lifecycle. Similar to the point regarding DMPs, local Research Group knowledge about internal and external infrastructures and RDM workflows could be useful to tap into by central services; in these cases, an effort must also be made to align with institutional policies on data protection and integrity. This OpenAIRE primer gives advice on the handling of raw data, storage, versioning and data back-up [https://www.openaire.eu/RAW-DATA-BACKUP-AND-VERSIONING], while this Data Management Expert Guide from CESSDA is designed to help those working in social sciences to manage their data [www.cessda.eu/dmeg].

- ***Access and Publishing:***
  - **No**: Publishing data and providing long-term access to it is an important aspect to consider at Research Group level, where Research Groups can offer valuable insights into publishing routes and associated costs, also providing insights into issues to be considered for longer-term data access, especially for sensitive data. (e.g., Data access committees). However, the type of infrastructure needed for this means that it is more appropriate that support for data access and publishing is provided in-house at an institutional level, or via an external infrastructure (i.e., repository or subject-specific data centre).
- ***Appraisal and risk assessment:***
  - **Yes**: This involves determining whether the data that the Research Group holds is of potential value to their organisation or to the wider research community, and identifying appropriate preservation strategies based on this. Attention can also be given at this stage to integrating ethics approval processes with data appraisal and risk assessment. The DCC has produced an in-depth guide on how to appraise research data for long-term storage that can assist on this task at a Research Group level [https://www.dcc.ac.uk/guidance/how-guides/appraise-select-data].
- ***Preservation:***
  - **Possibly**: Similar to Access and Publishing, it is important to think about preservation and to include planning for this in a DMP. However the provision of preservation infrastructure will depend on whether the group has their own data repository or databases that are used to support storage and access beyond the life of the project. For the Research Group level, there is an OpenAIRE guide that looks at appropriate data formats for preservation [https://www.openaire.eu/data-formats-preservation-guide].
- ***Training:***
  - **Yes:** Training at Research Group level is often dependent on the current competencies of Group members and how these competencies apply in the context of the work undertaken by the Group. Support from a higher level, such as the host institution, from experienced Research Group members or from other

external bodies can help identify whether RDM training is needed and what the most appropriate types are. It is often the case at this level that researchers can learn from their colleagues, with this peer-to-peer knowledge exchange providing a practical alternative to more formal training courses.

- ***Advisory services:***
    - **No**: Similar to training, Advisory services will usually be available from higher levels, such as the institution at which the research group is based. These advisory services will aim to provide support to researchers on various aspects of RDM, so it is practical for this service to be made available at an institute's professional services level, where it can be offered advice to multiple groups or projects.
- ***EOSC participation:***
    - **Possibly**: The manner of a Research Group's participation in the European Open Science Cloud depends on a number of factors. A Research Group, depending on the work it does, may interact with the EOSC as a data/service provider or a data/service user (or a mixture of both). In any case, the EOSC Rules of Participation Working Group's latest set of guidelines has set out the expectations and obligations of those participating in EOSC in all forms [https://www.eoscsecretariat.eu/sites/default/files/draft_eosc_rop_version_0.5_20-10-2020.pdf].
- ***FAIR assessment:***
    - **Yes**: Making sure that the data produced by your Research Group is Findable, Accessible, Interoperable, and Reusable will result in your data being potentially of more value for other researchers and will make it easier to select an appropriate strategy for long-term storage and preservation. There are a number of resources available to aid researchers and Research Groups in making their data more FAIR; including the FAIR Aware tool , which is designed to help researchers and data managers learn about the requirements for data to be FAIR [https://fairaware.dans.knaw.nl/]

Institution

- ***RDM policy:***
    - **Yes**: Having an RDM policy in place will allow an institution to align its research activities with its overall strategic objectives and direction. An RDM policy is one of the most important documents that an institution will have in place, enabling it to explicitly define the value of its research activities to the overall organisation, align with funder directives, and engage all levels in the organisation in good RDM practices. The DCC has produced a guide for developing RDM services at institutional level [https://www.dcc.ac.uk/sites/default/files/documents/publications/DCC_Howto_Discover_Requirements.pdf], while OpenAIRE has a checklist for those organisations looking to focus on their Open Science policy [https://www.openaire.eu/open-science-policy-checklist-for-research-performing-organisations]
- ***Business plans and sustainability:***
    - **Yes**: Accounting for the sustainability of RDM services at institutional level is essential. Part of this process includes identifying the costs of maintaining RDM services and forecasting where possible which aspects of the service will need to be developed in line with the institution's future priorities (this includes alignment with funder mandates). It is also crucial to link institutional business plans and sustainability to efforts to maximise the value of research group outputs in the long-term. SPARC Europe has designed a tool [https://sparceurope.org/evaluate-your-rdm-offering/] to evaluate an institute's RDM service offering, which can help in focusing on which areas need to be improved in light of the institute's strategic planning.
- ***Data Management Planning:***
    - **Yes**: At the institutional level, data management planning involves actively supporting researchers in documenting the plans for their research outputs and enabling the institution to take advantage of the information it gathers for the purposes of future RDM service provision and long-term strategic planning. This process increasingly involves producing data management plan (DMP) documentation; there are a number of data management plan platforms that allow for institutional accounts to be set up, with institute- and/or funder-specific requirements and guidance to be included in institutional DMP templates. One such is OpenAIRE's Argos platform [https://argos.openaire.eu/splash/], which is part of the OpenAIRE Research Graph, while another is the DCC's DMPonline tool [https://dmponline.dcc.ac.uk/].
- ***Active data management:***
    - **Yes**: Rather than engaging directly with research outputs, an institution's responsibility when it comes to active data management is generally in the form of service provision. These services include things like file storage and synchronisation, data back-up, security, and networked/linked storage. Other aspects to consider from the institution's perspective will be whether researchers at the institution are working with data that they are generating themselves or

with acquired data (which may require differing capabilities in terms of storage capacity and security), and the extent to which the institution's services compete with or complement any third-party services used by researchers. On the topic of data back-up, OpenAIRE has made available this useful 10-point checklist [https://www.openaire.eu/data-backup?highlight=WyJyaXNrIl0=]

- ***Access and Publishing:***
    - **Yes**: From an institutional perspective, this refers to the facilitation of data deposit and making this data as openly available as possible. Depending on the type of institution (whether it is a research-intensive organisation or otherwise) this may mean that setting up an institutional repository is the best approach in ensuring that researchers are supported in publishing their data. In any case, information on data produced by local research projects, often in the form of metadata, should be recorded by the host institution. It is also important to consider here that an institution's approach to the oversight of data publishing and access can add value to its data collections. OpenAIRE Explore provides a platform for searching and linking data and projects funded by Horizon2020 [https://www.openaire.eu/reporting-to-the-ec]
- ***Appraisal and risk assessment:***
    - **Yes**: Having procedures in place to appraise and assess research outputs and identify any potential risks is an important step in maximising the value of the data produced at an institution. This process is linked to Access and Publishing and Preservation steps - having a robust policy in place to appraise an institute's research outputs can inform later decisions on data access, publishing, and the institute's preservation strategies. To this end, the DCC's 'How to Appraise & Select Research Data for Curation' guide sets out the steps to be taken in developing an appraisal policy, as well as the specific roles and responsibilities of those involved [https://www.dcc.ac.uk/sites/default/files/documents/How%20to%20Appraise%20and%20Select%20Research%20Data.pdf]
- ***Preservation:***
    - **Yes**: From an institutional perspective, planning for preservation involves safeguarding not only the institution's research outputs themselves, but also the technologies that ensure their future reuse is possible. This involves a consideration of specific software that is used to read any research data and how feasible it is to maintain this. The institution's appraisal & risk assessment policy is also linked here in deciding which data is of value to organisation (and wider community) and thus worth preserving. The Digital Preservation Coalition Preservation Handbook contains a useful methodology for the development of institutional policies and strategies on preservation [https://www.dpconline.org/handbook/institutional-strategies/institutional-policies-and-strategies].
- ***Training:***
    - **Yes**: Depending on its available resources, and the degree to which good RDM practices are embedded in the institution, policy makers will decide on providing

online or face-to-face training (or a mixture), from in-house or external sources (or a mixture), to researchers and/or research support staff. In the long term, the institution's role descriptions for professional and research support staff can be adapted include the competencies required to deliver RDM training to the institution's research community (and potentially to those outside the institution). For institutions looking to provide external training there are a wide range of resources available, including OpenAIRE's RDM 'Train the Trainer' resources [https://www.openaire.eu/rdm-train-trainer-resources/view-document], a course designed by DCC and Research Data Netherlands on 'Delivering Research Data Management Services' [https://www.futurelearn.com/courses/delivering-research-data-management-services], while the DMT Clearing House provides a registry for online learning resources about RDM [https://dmtclearinghouse.esipfed.org].

- ***Advisory services:***
    - **Yes**: Providing advice on RDM at an institutional level is an effective way to guide researchers through the array of RDM tools and services that are widely available, and aid researchers in selecting the most relevant service for their work. This service can also help to direct researchers toward any appropriate training resources that might benefit their data management practices. Institutions will need to position any advisory services such that they complement the overall strategic aims of the institution; ideally, this service will be able to provide researchers with support throughout the lifecycle of their projects, from grant identification and submission, through to the storage and preservation of research outputs.
- ***EOSC participation:***
    - **Yes**: Institutions themselves will typically engage with the EOSC from the perspective of being a research performing organisation holding important data collections and perhaps as a service provider. These services can be in the form of technical resources (such as repositories or cloud storage) or 'human' services (such as training and consultancy). To provide a service to EOSC, an institution must adhere to the EOSC Rules of Participation [https://www.eoscsecretariat.eu/sites/default/files/draft_eosc_rop_version_0.5_20-10-2020.pdf] along with the more specific requirements to be met by service providers [https://eosc-portal.eu/for-providers]. Depending on an institution's capacity, engagement with EOSC services should be encouraged where possible, especially where research receives European Commission funding; this can be done by advocating for EOSC engagement in the institution's RDM strategic policy.
- ***FAIR assessment:***
    - **Yes**: Assessment of FAIRness at institutional level involves both the services that the institution provides and the levels of awareness of FAIR amongst students, researchers, and policy makers. In terms of services, an article [https://doi.org/10.1016/j.patter.2020.100058] resulting from a collaboration between the OpenAIRE, FAIRsFAIR, RDA Europe, EOSC-hub, and FREYA

projects sets out a series of recommendations which provide a framework for aligning policies with the FAIR principles. Amongst the recommendations for institutions is the establishment of data stewardship programmes and defining roles for institutional data stewards, who take responsibility for providing training and advice on FAIR. Linked to this is the idea of increasing FAIR awareness at all levels of the institution; the FAIRsFAIR-developed FAIR-Aware tool [https://fairaware.dans.knaw.nl/] can be of use to those looking to gauge the level of engagement with FAIR at their institution.

Repository

- ***RDM policy:***
    - **Yes**: An RDM policy document for a repository will focus on the types of service the repository provides to potential users and the procedures and protocols for maintaining the services. For example, if a repository only accepts datasets under a certain size, then its policy will outline the reasons for this and if there are any exception cases. The policy will include information on acceptable formats, ownership rights, licensing, metadata formats, PIDs, versioning, and access procedures. This policy may also be supplemented by a 'terms of use' document or something similar, which outlines the requirements for users of any services. The NI4OS Europe project has developed a Repository Policy Generator tool that allows for the creation of customised policies based on provided information [https://repol.ni4os.eu/].
- ***Business plans and sustainability:***
    - **Yes**: Repositories often hold large amounts of data and other research related outputs which are potentially of commercial and scientific value to both contributing researchers and the wider research community. In addition, funding for researchers and research projects often requires that data and outputs produced be deposited in a relevant repository to ensure that their value is maximised. As such, plans for the long term sustainability and the business case for services of a repository should be in place to ensure that there is confidence in the service and that resources for the maintenance and development of services are accounted for. The Curation Costs Exchange 'Digital Curation Sustainability Model' (DCSM) [https://www.curationexchange.org/understand-your-costs/76-using-dcsm] can assist in this regard (its 'Example questions for organisations' section on this page is especially relevant for repositories).
- ***Data Management Planning:***
    - **No**: As the data and research outputs held in repositories is submitted by users themselves, it can be said that a large part of the responsibility for planning rests with them. However, on behalf of the repository, long term planning in terms of the capacity of the repository to continue to provide its services is necessary; this can indirectly involve data management planning (for example, planning for adequate storage over the long term, or how changes to any policies and impact on data currently held).
- ***Active data management:***
    - **No**: Similar to the Data Management Planning aspect above, responsibility for Active data management also rests largely with the researchers and principal investigators who generate the data and research outputs. However, review of data is carried out by some repositories to ensure that submissions meet the standards required. The extent to which this is carried out depends on the repositories' selectiveness and the amount of resource it has to carry out detailed checks on submissions.
- ***Access and Publishing:***

- **Yes**: One of the key functions of a repository is to maintain access to data and other research outputs deposited for re-use by others. In order to facilitate greater access, repositories can consider enlisting as part of a registry such as re3data.org, which allows users to locate appropriate repositories for accessing relevant data and sharing data with the most interested communities. A report from the APARSEN project [http://www.alliancepermanentaccess.org/wp-content/uploads/sites/7/downloads/2014/06/APARSEN-REP-D31_1-01-1_4_incURN.pdf] outlines the responsibilities for continuing to upload digital rights within the context of data access.

- ***Appraisal and risk assessment:***
  - **Yes**: Appraisal and risk assessment at repository level will focus on two levels: on the services the repository provides and on the research outputs deposited by users. For the former, important aspects to consider are whether the services provided are those that are most effective at maximising the value of the deposited research outputs and whether the changing needs of users will continue to be met by the services; for the latter, appraisal and risk assessment will tend to centre on the legal and ethical guidelines which the producer adhered to in creating any research outputs, and whether these criteria impact access and re-use across different disciplinary and regional boundaries.
- ***Preservation:***
  - **Yes**: The preservation of research outputs is at the core of the services that repositories offer. This preservation refers to the integrity of the deposited data objects, research outputs and attendant metadata, as well as the continued access to these. Another report from the APARSEN project [http://www.alliancepermanentaccess.org/wp-content/uploads/sites/7/downloads/2014/06/APARSEN-REP-D21_1-01-2_1_incURN.pdf] outlines the high-level services that can enable repositories and other similar organisations in reinforcing their preservation processes and ensuring the sustainability of the repositories' services.
- ***Training:***
  - **Possibly**: Repositories may provide training materials or support to users; for example, DANS [https://dans.knaw.nl/en/about/services/training-consultancy/training], the Dutch national centre of expertise and repository for research data, provides training on Open and FAIR data, open science, RDM, and long-term preservation with a view to improving the data that will be deposited in its archive. Where it is not the case that a repository can provide extensive training and support materials to users, most repositories will generally provide training materials or support in some form, usually supplying guides for users on submitting and accessing the data and other research outputs that it holds.
- ***Advisory services:***
  - **Yes**: Some repositories offer the chance to contact staff about the suitability of a repository to deposit their data prior to submission. They can offer advice on things like standards that the repository recommends, appropriate formats and

ontologies, and queries to do with costs, etc., making it more straightforward for depositors to ensure their data is in the right state for submission. Similar to training, advice for users who are not depositing data (i.e. those looking to access and re-use data) can be provided in the form of guide documents or multimedia, or in a Frequently Asked Questions webpage.

- ***EOSC participation/readiness:***
    - **Yes**: A repository can join the EOSC as a provider, where its services can potentially be accessed by users beyond its original community. The criteria that the EOSC requires of providers can also be of benefit to the repository itself in encouraging it to focus on its FAIR enabling services and making available user statistics and feedback. These requirements can also feed into the strategic planning around the sustainability of the organisation and its services.
- ***FAIR assessment:***
    - **Yes**: From the perspective of repositories, incorporating the FAIR principles with mean focusing on FAIR enabling services. There are many resources available to assist repositories in this respect, including two papers from the FAIRsFAIR project, one on a framework for FAIR services [https://doi.org/10.5281/zenodo.4292599] and another on FAIRsFAIR's support to help organisations meet the CoreTrustSeal Requirements with an assessment of repositories' ability to enable FAIR data [https://doi.org/10.5281/zenodo.3835698], while the FAIRsFAIR project has also developed the F-UJI tool to assess the FAIRness of datasets [https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool]. A repository may also consider implementing a FAIR Data Point [https://www.fairdatapoint.org/], which allows metadata to be stored, searched, and accessed in a FAIR manner by users.

Research Infrastructure

- ***RDM policy:***
    - **Possibly**: An overarching policy is not necessarily needed at this level since some services may not be suitable for all researchers and a level of user discretion should be allowed. For example, those researchers that are producing or dealing with sensitive data cannot and should not use the same active storage areas that are used in other fields of research. In these cases, where possible, services such as data safe havens are preferred but may not be available in all institutions.
- ***Business plans and sustainability:***
    - **Yes**: Creating and maintaining services can be very expensive. Indeed, a first step in the decision whether to build bespoke services should be to determine the level and amount of research being conducted in a host institution and the income from this through grants and other channels. This should be used to weigh up the value of the data being created and used in that institution and this is connected to the appraisal and risk assessment. Research intensive institutions, where financially feasible, should build their own services for their researchers and students. This will require careful planning and a business case needs to be drawn up as well as outlining the sustainability and scalability of any proposed services. For example see the EC context [https://ec.europa.eu/info/research-and-innovation/strategy/european-research-infrastructures/long-term-sustainability_en].
- ***Data Management Planning:***
    - **Possibly**: Although not always necessary, DMPs are important for service development. They allow institutions and funders to monitor the usage of their research infrastructures and act accordingly where necessary. For example, if there are several projects running concurrently that are using large amounts of storage space of compute power, then there may be bottlenecks in the future which should be accounted for. DMPs, by their very nature outlining what should happen in the future, should allow the host institutions and funders to foresee such hurdles when taken together at scale.
- ***Active data management:***
    - **Yes**: Safeguarding data through the active phase of the curation lifecycle is extremely important. Data loss and breaches of security are risk factors that should be mitigated and the ability to backup data and retrieve it where necessary should be factored into the design of any infrastructure. This includes networked storage and cloud services, and, for sensitive data, secure spaces such as data safe havens. See for example EUDAT services such as B2DROP [https://marketplace.eosc-portal.eu/services/b2drop/information] for syncing and sharing data.
- ***Access and Publishing:***
    - **Yes**: The ability to access data for third parties as well as data owners can pose difficult questions, especially for sensitive data. However, in all cases maintaining data in formats that are easily read, i.e. open file formats, should be given

priority. This is also true for eventual publishing of data. For help on this, see the UK Data Services' recommendations for file formats [https://www.ukdataservice.ac.uk/manage-data/format/recommended-formats].

- ***Appraisal and risk assessment:***
  - **Yes**: When channeling any given data through the various tools and spaces that a research infrastructure provides, it will be necessary to determine the value and sensitivity of that data and act accordingly with respect to safeguarding it. Appraisal is also necessary from the viewpoint of determining which data should be kept for the long-term and which can or should be destroyed. Many times this may be a subjective call, but there should be policies and/or guides in place that will allow researchers to determine this themselves. For example, the DCC has produced this guide on how to appraise and select research data. [https://www.dcc.ac.uk/guidance/how-guides/appraise-select-data].
- ***Preservation***:
  - **Yes**: Essentially, this relates to repositories, whether institutional, domain specific or generalised. The option of which to use lies with the researcher but it is highly recommended that the researcher chooses a domain specific repository where possible, which will increase the value of their data. See for example [https://marketplace.eosc-portal.eu/services/b2share/information].
- ***Training***:
  - **Yes**: For those services that are new or more complex in nature, there should be training available. Many research infrastructures offer training, such as ESFRI clusters [https://www.eosc-hub.eu/eosc-hub-and-esfri-cluster-projects] and the NERC Data Tree course [https://datatree.org.uk/].
- ***Advisory services***:
  - **Yes**: Choosing the appropriate service, knowing what services are available, and ultimately knowing how to use any given service (training) will require advice. For example, at the University of Edinburgh, there is a large and wide-ranging catalogue of services available, some of which are not exclusive to the University. This can be overwhelming to researchers, especially those that are new to the host institution, and guidance must be provided to create workflows for their research based on the tools available to them, as is the case with the University of Edinburgh's Digital Research Services [https://www.digitalresearchservices.ed.ac.uk/].
- ***EOSC participation***:
  - **Yes**: The EOSC has several services that have been built or that are in development or being planned, and that are available through a federated system to any users. Integration with these services and using them where possible, especially for EC funded projects, should be mandated. A catalogue of services can be found here [https://eosc-portal.eu/services-resources].
- ***FAIR assessment***:
  - **Yes**: To amplify FAIR principles compliance, mechanisms should be set in place to provide FAIR metrics. These will allow evaluation of how well any given workflow built of various services offers FAIR compliance and can be a useful

indicator for researchers of how best to deal with their data. See the FAIRsFAIR project's F-UJI tool [https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool].

Funder

- ***RDM policy:***
  - **Yes**: Since they are providing financial backing, usually through public money, funders need to make their grantees aware of their obligations in relation to the work they conduct and how the data generated should be handled. Since data volume is growing at a pace that is ever increasing, data needs to be managed appropriately to allow research integrity to be upheld. Examples of funder mandated policy requirements can be found at the DCC [https://www.dcc.ac.uk/guidance/policy/overview-funders-data-policies] which mainly shows those in the UK but also for Horizon 2020. OpenAIRE provides a useful checklist for research funding organisations to assess their readiness in adopting the Horizon2020 Open Science requirements as part of their RDM policy [10.5281/zenodo.2578036].
- ***Business plans and sustainability:***
  - **Yes**: A common strategy employed by institutions is to assess the gaps in their service provision. Filling in any gaps will require substantial financial investment in many cases and therefore requires proper planning and must exhibit awareness of future challenges that will also be encountered and that need to be addressed. Funders can identify key gaps in provision and prioritise effort there e.g. via open calls (see NWO [https://www.nwo.nl/en/calls?f%5B0%5D=calls_bw_call_status%3A1866&f%5B1%5D=calls_bw_call_status%3A1868&sort_bef_combine=bw_call_start_date_DESC] or Wellcome Open Science grants [https://wellcome.org/what-we-do/our-work/open-research]) or supporting data centres like UKDA or NERC.
- ***Data Management Planning:***
  - **Yes**: As well as a tool that aids researchers themselves, DMPs are also very useful to funders who can monitor how their money is being spent and thus potentially identify problems in the future. This is related to the business plans and sustainability and is again something that funders have the power to implement in terms of building infrastructure. See for example the Horizon 2020 DMP template [https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm].
- ***Active data management:***
  - **No**: Unlike data at the final version stage of the curation lifecycle, which will typically be deposited in a repository, and which requires funder mandates, the active phase is not and should not be subject to funder requirements. This is in large part due to the heterogenous nature of service provision at institutions. See *Institution*.
- ***Access and Publishing:***
  - **Yes**: Publishing of data has become a very prominent topic and one that has been made more so due to increasing awareness by lay persons. The public nature of research in most cases warrants public access to these data and, to facilitate this, most funders will provide guidance on data sharing and data

access statements. Many of the major publishers have aligned with funder policies by putting data sharing policies in place with some commitment to mandatory sharing of data, where applicable. See Springer Nature [https://www.springernature.com/gp/authors/research-data-policy/research-data-policy-types] and PLOS [https://journals.plos.org/plosone/s/data-availability] for two publisher examples.

- ***Appraisal and risk assessment:***
  - **No**: These considerations are likely specific to individual institutions and therefore should not be overseen by funders themselves. Instead, the RDM policy should provide the necessary failsafe in these circumstances. See *Institution*.
- ***Preservation:***
  - **Yes**: The preservation of research data has clear benefits for researchers themselves, but also for funders; if the preservation of research data is carried out correctly, this can reduce the need for studies to be repeated and allow support and resources to be directed toward more innovative and original research. Funders have a key role in incentivising the preservation of research data, and can do so through their policy requirements. For example, the Dutch funder NWO stipulates [https://www.nwo.nl/en/research-data-management] that it expects researchers to aim to preserve their data for ten years beyond the project end (taking into account discipline-specific and legal limitations of this).
- ***Training:***
  - **Possibly**: How to use the services that are developed and built can seem daunting to many researchers and this must be addressed by providing adequate training that will provide the basics at the very least. Training content should not be directly dictated by funders since it is likely that there will be unique situations per institution but financial backing should be provided.
- ***Advisory services:***
  - **Possibly**: Similar to training, this should also be financially backed but not necessarily driven by funder requirements.
- ***EOSC participation:***
  - **Yes**: As part of pan-European efforts to integrate services and data, funder involvement should be encouraged in order to coordinate these efforts [https://eoscpilot.eu/pilots/research-funding-bodies], while many are already members of the EOSC association [https://www.eosc.eu/members].
- ***FAIR assessment:***
  - **Possibly**: FAIR metrics are becoming increasingly important and prevalent as more researchers embrace the principles. It may be a requirement of funders, and in their own best interests, to adequately appraise the research they fund for FAIR compliance.

National

- ***RDM policy:***
  - **Yes**: As of the end of 2020, a significant minority of European countries have national level policies in place focussing on open science and research, and this is projected to grow. Those that have devised policies aim to provide researchers in their countries with unambiguous guidelines on improving open science and research. The Netherlands provides a good example of national level ambitions where they have developed a roadmap [https://doi.org/10.5281/zenodo.4486423], while a landscape analysis of policies in Europe can also be referred to. [https://zenodo.org/record/4005612#.YCuianmnzVghttps://zenodo.org/record/4005612#.YCuianmnzVg].
- ***Business plans and sustainability:***
  - **Yes**: As with funder backing, national level investment in services is also important, especially those that are publicly funded. They will likely provide more trustworthy services and ones that will be sustainable and long lasting. See for example the eInfrastructures Austria project, the German National Research Data Initiative (NFDI: [https://www.nfdi.de/en-gb]) and the Swiss Data Lifecycle Management project (DLCM: [https://www.dlcm.ch/about-us/dlcm-project]).
- ***Data Management Planning:***
  - **Yes**: In a similar way to that of funders, national level DMPs will allow a better understanding of resource requirements of researchers and will be the basis of service development planning. These DMPs will also be a useful resource cataloguing the research that has been conducted within these countries. See the Swedish example where the Swedish Research Council have introduced a DMP requirement [https://www.vr.se/english/applying-for-funding/requirements-terms-and-conditions/producing-a-data-management-plan.html] and in the Netherlands the Dutch Research Council has implemented a DMP requirement [https://www.nwo.nl/en/research-data-management].
- ***Active data management:***
  - **No**: These services should be provided ideally at an institutional or research group level.
- ***Access and Publishing:***
  - **Possibly**: Where possible, and as required, national repositories should be established that provide a last resort for data publication for the long-term. It is in the best interests of countries, especially those that provide substantial public funds, to safeguard their research data if institutional and domain repositories do not exist. Indeed, it may also be an option to deposit the data in more than one repository as a backup.
- ***Appraisal and risk assessment:***
  - **No**: Though this should be addressed at National policy level, the responsibility for this should rest with the research group, institution, or infrastructure, who will be better-placed to carry out the necessary checks on data appraisal and risk assessment.

- ***Preservation:***
    - **Yes**: Many countries have decided to implement a national level approach to supporting the preservation of data generated by their researchers. This has ensured that there is always a safety net for collecting all data and that researchers do not need to rely on third parties. Pooling of resources has allowed costs to be reduced and therefore the economic case for national repositories has been very attractive. See for example the Ductch national repository, DANS. [https://dans.knaw.nl/en/front-page?set_language=en]
- ***Training:***
    - **Yes**: Especially when considering local rules and regulations, it is essential that there is training available at a national level to provide the basic foundations for researchers to conduct their work. This is even more important when dealing with data of a sensitive nature and which has to abide within the laws at national and supranational levels. See for example training from EUDAT [https://eudat.eu/training].
- ***Advisory services:***
    - **Yes**: Similarly to training, there should be guidance available to help navigate local and national mandates and other regulations. Exchange of data and services and between countries and the ability to be open, can be impacted if these issues are not addressed.
- ***EOSC participation:***
    - **Yes**: The European Research Area (ERA) is an initiative to allow free exchange of data and services across the continent and the EOSC is central to achieve those goals. Participation by nations is encouraged and uptake and implementation of services developed on a federated model is leading to better integration. See this short overview document from the European Commission on the new European Research Area [https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/aae418f1-06b3-11eb-a511-01aa75ed71a1].
- ***FAIR assessment:***
    - **Yes**: Related to the EOSC, FAIR assessment and metrics will help in better integration of services and data. As a set of guiding principles, FAIR does not set out to dictate the technology but more how data should be managed, and this will have direct implications on how the technology should be used. At a national level, there should be an ability to gauge this, thereby increasing the value of the data produced in their jurisdiction. See F-UJI [https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool].