

Active, actionable DMPs

Sarah Jones | Digital Curation Centre | sarah.jones@glasgow.ac.uk

Stephanie Simms | California Digital Library | stephanie.simms@ucop.edu

Daniel Mietchen | National Institutes of Health | daniel.mietchen@nih.gov

Tomasz Miksa | TU Wien | tmiksa@sba-research.org



#ActiveDMPs



Planning &
administration

Create, analyse,
manage data

Publishing & reuse

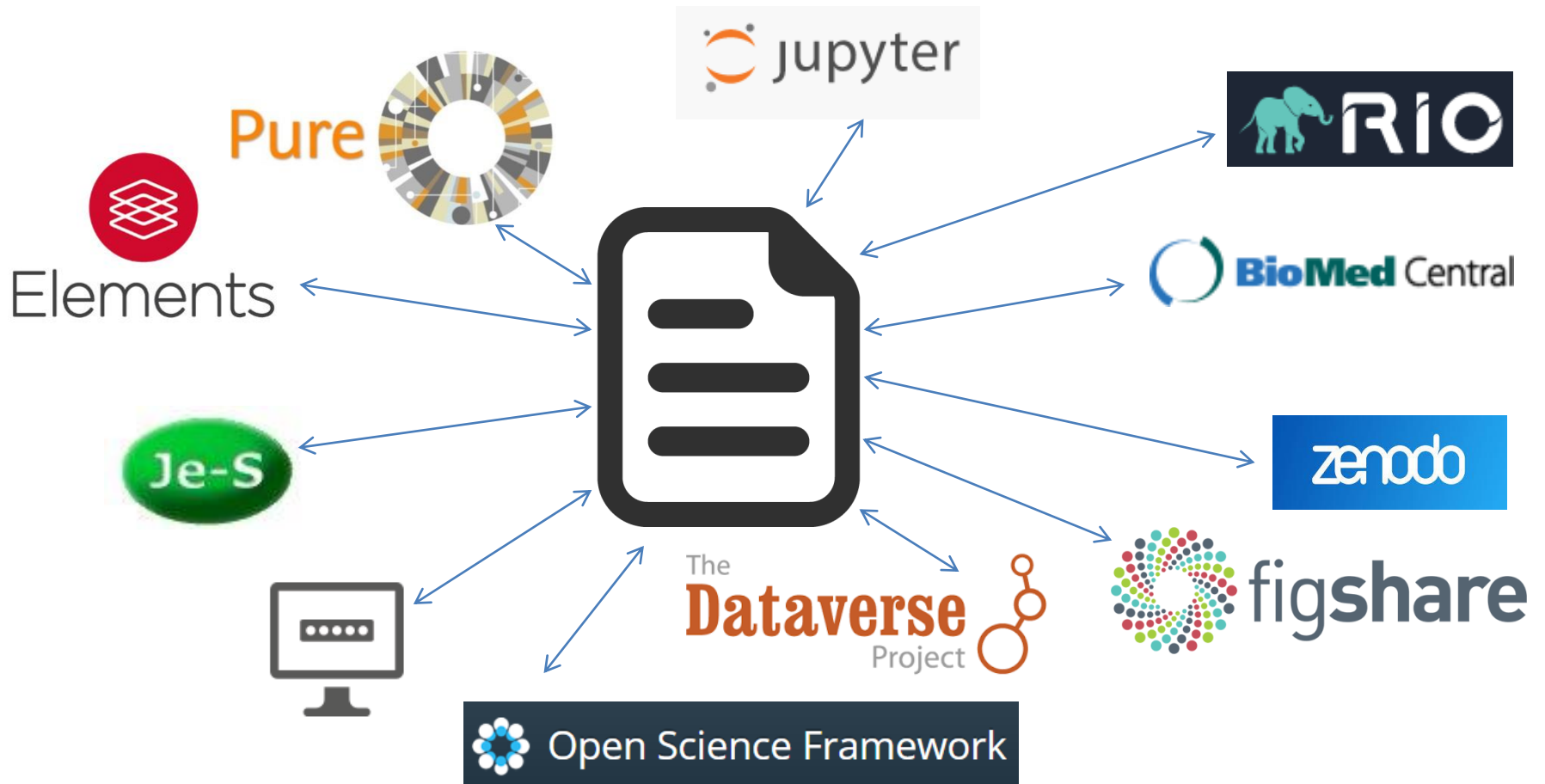


- DMP on periphery
- Often done at grant stage and not looked at again
- Opportunities to (re)use information being missed
- Disconnected & unlinked

Planning & administration

Create, analyse, manage data

Publishing & reuse



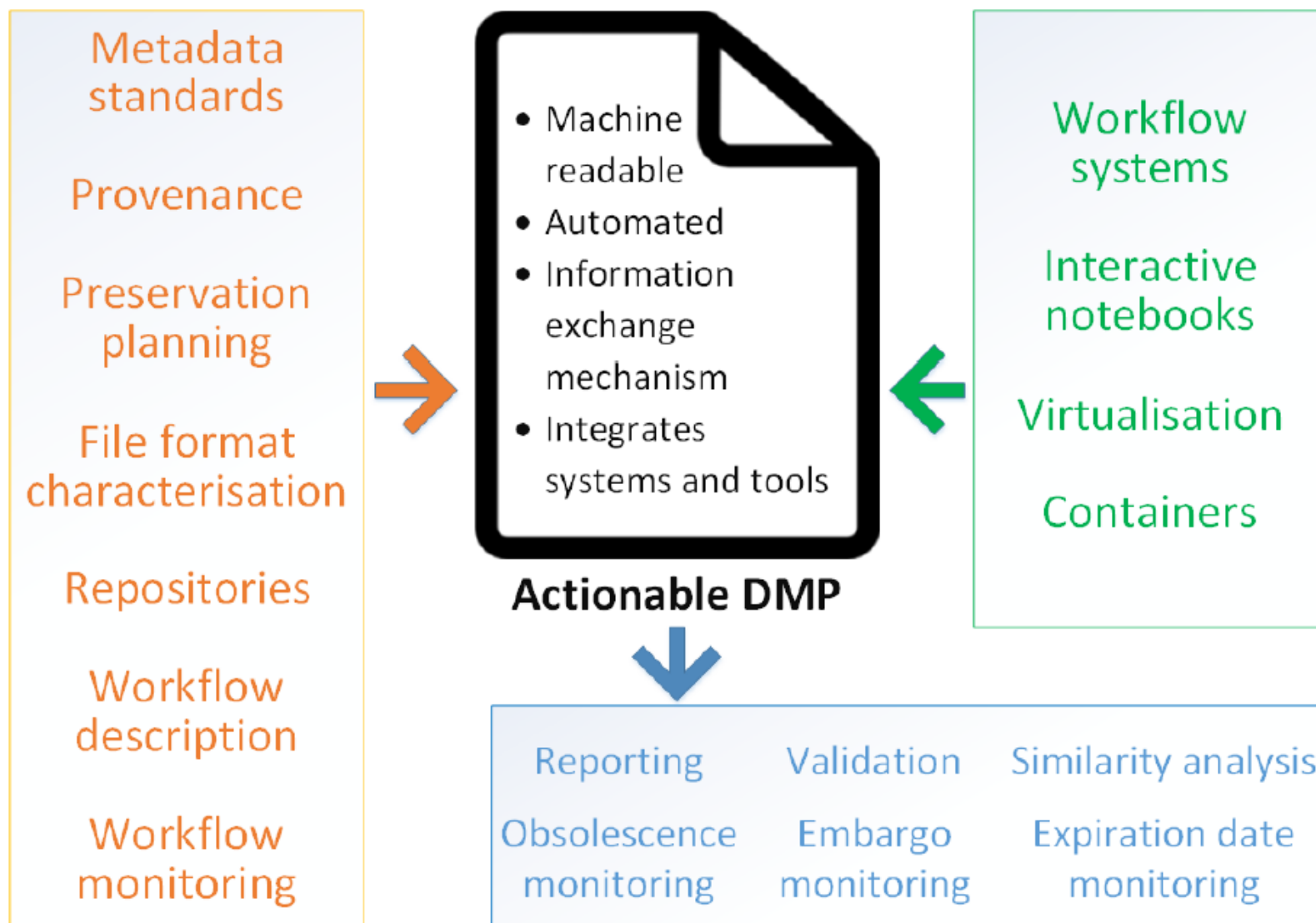
Identifiers:




Resource Identification



What if we had all this information in one place?



Mapping

 Checklist for a Data Management Plan, v4.0

Please cite as: DCC. (2013). Checklist for a Data Management Plan. v.4.0. Edinburgh: Digital Curation Centre. Available online: <http://www.dcc.ac.uk/resources/data-management-plans>

| DCC Checklist | DCC Guidance and questions to consider |
|--|--|
| Administrative Data | |
| ID | A pertinent ID as determined by the funder and/or institution. |
| Funder | State research funder if relevant |
| Grant Reference Number | Enter grant reference number if applicable [POST-AWARD DMPs ONLY] |
| Project Name | If applying for funding, state the name exactly as in the grant proposal. |
| Project Description | <p>Questions to consider:</p> <ul style="list-style-type: none"> - What is the nature of your research project? - What research questions are you addressing? - For what purpose are the data being collected or created? <p>Guidance:</p> <p>Briefly summarise the type of study (or studies) to help others understand the purpose for which the data are being collected or created.</p> |
| PI / Researcher | Name of Principal Investigator(s) or main researcher(s) on the project. |
| PI / Researcher ID | E.g ORCID http://orcid.org/ |
| Project Data Contact | Name (if different to above), telephone and email contact details |
| Date of First Version | Date the first version of the DMP was completed |
| Date of Last Update | Date the DMP was last changed |
| Related Policies | <p>Questions to consider:</p> <ul style="list-style-type: none"> - Are there any existing procedures that you will base your approach on? - Does your department/group have data management guidelines? - Does your institution have a data protection or security policy that you will follow? - Does your institution have a Research Data Management (RDM) policy? - Does your funder have a Research Data Management policy? - Are there any formal standards that you will adopt? <p>Guidance:</p> <p>List any other relevant funder, institutional, departmental or group policies on data management, data sharing and data security. Some of the information you give in the remainder of the DMP will be determined by the content of other policies. If so, point/link to them here.</p> |
| Data Collection | |
| What data will you collect or create? | <p>Questions to consider:</p> <ul style="list-style-type: none"> - What type, format and volume of data? - Do your chosen formats and software enable sharing and long-term access to the data? - Are there any existing data that you can reuse? <p>Guidance:</p> <p>Give a brief description of the data, including any existing data or third party sources that will be used, in each case noting its content, type and coverage. Outline and justify your choice of format and consider the implications of data format and data volumes in terms of storage, backup and access.</p> |
| How will the data be collected or created? | <p>Questions to Consider:</p> <ul style="list-style-type: none"> - What standards or methodologies will you use? - How will you structure and name your folders and files? - How will you handle versioning? - What quality assurance processes will you adopt? <p>Guidance:</p> <p>Outline how the data will be collected/created and which community data standards (if any) will be used. Consider how the data will be organised during the project, mentioning</p> |



Requirements & Limitations

Dublin Core,
PREMIS,...

W3C PROV,
Janus,...

Plato, C3PO,...

FITS, Droid,...

CKAN, Research
Compendia,
Zenodo...

Research Objects,
Context Model, ...

PMF, CDE,
noWorkflow,...

OSF

GitHub

Taverna,
VisTrails,
Pegasus,...

Jupyter,
Zeppelin, ...

Reprozip,
Vagrant,...

Docker





Utopia workshop

47 participants from 16 countries

- Funders
- Developers
- Librarians
- Service providers
- Researchers

Understand research workflows

Develop use cases for maDMPs

Set priorities for future work

www.dcc.ac.uk/events/workshops/postcard-future-tools-and-services-perfect-dmp-world

Uses cases and prioritisation

- Interoperability with research systems
- Institutional perspective
- Repository use cases
- Evaluation & monitoring
- Utilising PIDs

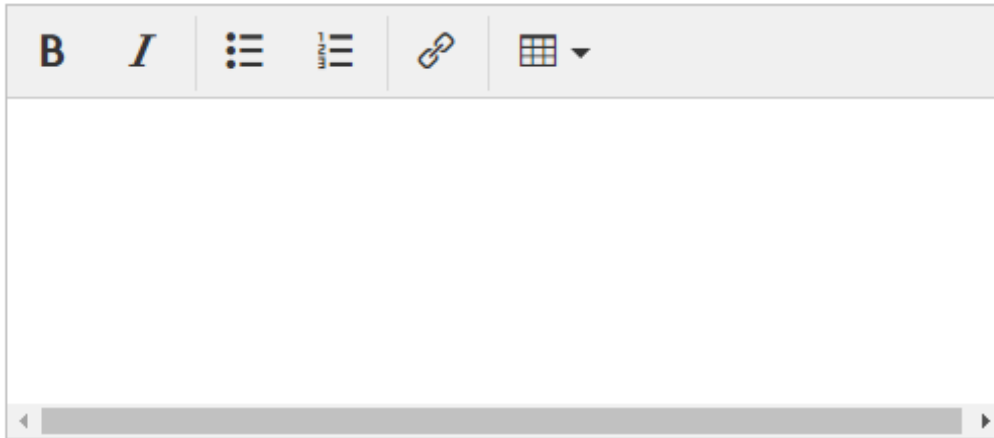


maDMP priority areas

- Common standards and protocols
- Funder integration
- Share/publish/deposit DMPs
- Utilise PIDs for automatic reporting
- Capacity planning (institutional & data centre)
- Automated compliance checks

The problem of freetext

Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how



A screenshot of a rich text editor interface. The toolbar at the top includes icons for bold (B), italic (I), bulleted list, numbered list, link, and table. Below the toolbar is a large, empty text area with a horizontal scrollbar at the bottom.

Guidance

Share note

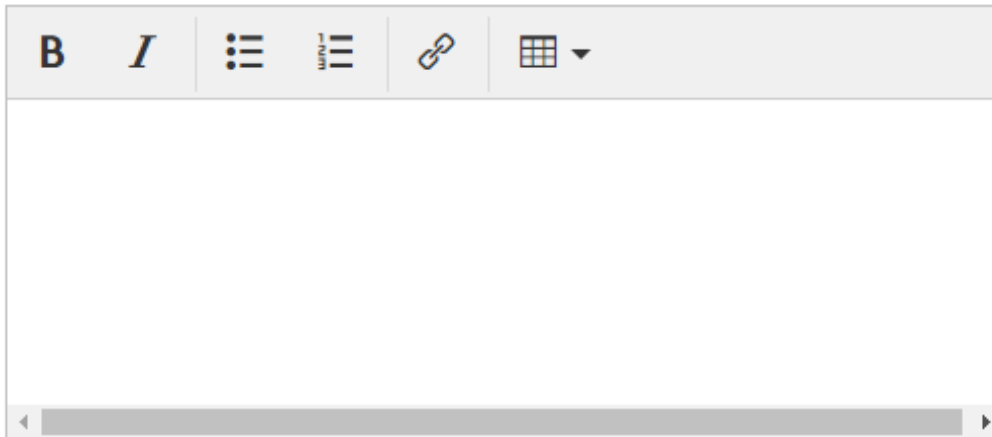
EC Guidance

The Research Data Alliance provides a [Metadata Standards Directory](#) that can be searched for discipline-specific standards and associated tools.

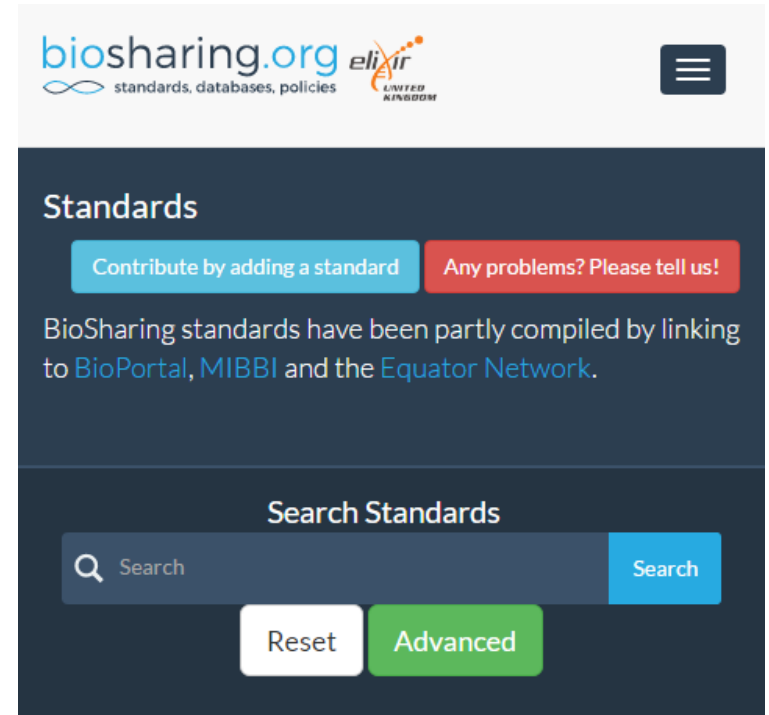
Templates typically ask very broad questions, even when dropdown options are feasible (e.g. metadata standards, file formats, data volumes, repositories, licences...)

Plugins to give structured response

Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how



A rich text editor toolbar with icons for bold (B), italic (I), bulleted list, numbered list, link, and table, followed by a large empty text area.



The screenshot shows the biosharing.org website. The header includes the biosharing.org logo, the elixir logo, and the text 'standards, databases, policies'. Below the header, the page is titled 'Standards'. There are two buttons: 'Contribute by adding a standard' and 'Any problems? Please tell us!'. The main content area states: 'BioSharing standards have been partly compiled by linking to BioPortal, MIBBI and the Equator Network.' Below this is a 'Search Standards' section with a search input field, a 'Search' button, a 'Reset' button, and an 'Advanced' button.

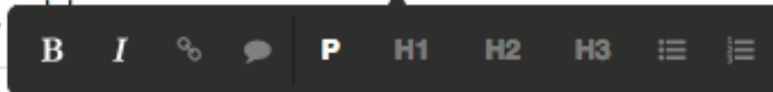
<http://rd-alliance.github.io/metadata-directory>



New text editor

Substance Forms: <http://substance.io>

Formatting tools appear when you **highlight text**. Additional features coming soon: e.g.,



Move the cursor over the **marker** to see comments. Create lists like so:

- bullet 1
 - bullet 2
1. number 1
 2. and so on...

I am a **comment**.

Me too.

[Add comment](#)

Repository use cases

- I. Repository recommender service via re3data.org
- II. Text mine to ping repositories when mentioned in a DMP
- III. Use DMP as metadata to facilitate deposit process
- IV. Deposit DMPs with data



Institutional use cases

- I. Connect researchers to relevant services & support
- II. Gather information to forecast demand and do capacity planning
- III. Embed DMP in research process (domain workflows, ethics, admin systems)



Persistent identifiers (PIDs)


- Assign DOIs to releases of DMP versions
- Leverage other PIDs to populate DMP over time:
 - Researcher IDs ([ORCID](#)s)
 - Funder IDs ([FundRef](#))
 - Grant IDs
 - Research Resource IDs ([RRID](#)s)
 - antibodies, organisms, cell lines



- Also enables compliance monitoring

<http://pidapalooza.org>

ORCID integration

  University of California Office of the President Signed in as super_admin@example.com ▾



[View plans](#) [Create plan](#) [About](#) [Future plans](#) [Help](#) [Public DMPs](#) [Change language ▾](#)

Your account has been connected to orcid

Edit profile


Please note that your email address is used as your username. If you change this, remember to use your new email address on sign in.

You can edit any of the details below.

| | |
|--------------|--|
| Email * | <input type="text" value="super_admin@example.com"/> |
| First name | <input type="text" value="Super"/> |
| Last name | <input type="text" value="Admin"/> |
| Organisation | <input type="text" value="University of California Offi... x ▾"/> |
| Language | <input type="text" value="English (UK) x ▾"/> |
| orcid |  https://orcid.org/0000-0001-9129-3790  |

If you would like to change your password please complete the following fields.

| | |
|------------------|--------------------------|
| Current password | <input type="password"/> |
| New password | <input type="password"/> |



Utilising EC grant IDs in plans

- Harvest grant IDs from OpenAIRE API
- Provide look up when entering project details
- Enables join up of DMP with other outputs

SELECT FUNDING AGENCY

1. Select funding agencies

European Commission (EC) 

2. Select Project(s)

foster|

FosterREG - Fostering public capacity to plan, finance and manage integrated urban REGeneration for sustainable energy uptake - 649846

FOSTER ITS - First Operational, Secured and Trusted galileo Receiver for ITS - 641492

FOSTER - FACILITATE OPEN SCIENCE TRAINING FOR EUROPEAN RESEARCH - 612425

FOSTERNAV - Flash Optical Sensor for Terrain Relative Robotic Navigation - 262996

| | |
|--------------|---|
| Plan name | <input type="text" value="My plan (Horizon 2020 DMP)"/> |
| ID | <input type="text"/> |
| Grant number | <input type="text"/> |

Evaluation & monitoring

- Automated compliance checks
 - did researchers do what they said they would?
- Quality or validation checks
 - closed questions / range of defined options
 - training and evaluation rubrics
 - evaluate FAIRness of data and repository...

Next steps: maDMP pilots



Summary

Think of DMPs as key elements of a networked data management ecosystem:

- connected via a shared vocabulary
- actionable by humans and software
- versioned
- public

[From Flickr by highwaysengland, CC BY 2.0](#)



Join us for more!

Thurs 6th April, 9:30-11:00, @RDA Plenary in BCN, Active DMP IG session

