

**EGI-Engage**

Deliverable Review Form

|  |
| --- |
| **Details of the document being reviewed** |
| *Title:* | **D2.4 Data Management Plan** | *Document identifier:* | **EGI-doc-2556-v1**  |
| *Project:* | **EGI-Engage** | *Document url:* | [https://documents.egi.eu/document/2556](https://documents.egi.eu/document/2556%22%20%5Ct%20%22_new)  |
| *Author(s):* | **S. Andreozzi** | *Date: 21/8/2015* |  |

|  |
| --- |
| **Identification of the reviewer** |
| *Reviewer:* | **Matthew Dovey** | *Activity:* | **PMB** |

**General comments on the content**

|  |
| --- |
| **Comments from Reviewer:** |
| The current plan only covers data emerging from the partner projects within the EGI-Engage work-programme. What is missing is a data management plan for data which is created by the EGI-Engage project itself, such as datasets emerging from any user surveys or other community engagement activities, usage statistics for EGI services or infrastructure etc.In a number of cases, no agreed standards exist for the data – in such cases, is there any consideration of documenting the data formats (e.g. using DFDL)? |
| **Response from Author:**  |
| **My understanding of the Open Research Data pilot is that it mainly refers to research data that may underpins scientific publications. Data that are used mainly to extract requirements and develop new services are not under the scope of the DMP (and the open research data pilot more in general).****About providing more details on standards, this can be improved in the future iterations of the document. The guidelines state that this is a living document that needs updates when important events happen or before each project review (i.e., M12/M24/M30 for EGI-Engage).** |

**Additional comments**

*(not affecting the document content e.g. recommendations for the future)*

|  |
| --- |
| **From reviewer:** |
|  |

**Detailed comments on the content**

| **N°** | **Page** | **§** | **Observations** | **Reply from author(correction / reject,  …)** |
| --- | --- | --- | --- | --- |
|  | **8** | **2.2.3.3** | **states that datasets released are limited in scope (1/10 of total data) – why is there this additional limitation beyond the embargo period? What are the criteria applied to determine what data is in “scope”** | **This was clarified in the document; only data linked to scientific publications will be made publicly available (usually 1/10th)** |
|  | **9** | **2.3.4** | **Why no more than three copies? Is this a typo for no less than?** | **The contributor agreed to write “no less than”**  |
|  | **10** | **2.4.4** | **I’m not sure I understand precisely what “large short term” and “reduced long term” mean – what are the criteria for decidedly which? is there tiering between these? Is there a weeding policy?** | **Sent clarification request to DMP contact** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**English and other corrections:**

Note: English and typo corrections can be made directly in the document as comments.