

**EGI-Engage**

Deliverable Review Form

|  |  |  |  |
| --- | --- | --- | --- |
| **Details of the document being reviewed** | | | |
| *Title:* | **D6.1 Integration of assisted pattern recognition tools** | *Document identifier:* | EGI-doc-2643 |
| *Project:* | **EGI-Engage** | *Document url:* | <https://documents.egi.eu/document/2647> |
| *Author(s):* | Eduardo Lostal, Francisco Sanz | *Date:* | 13 November 2015 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Identification of the reviewer** | | | |
| *Reviewer:* | **Eric Yen** | *Activity:* | **PMB** |

**General comments on the content**

|  |
| --- |
| **Comments from Reviewer:** |
| 1. Machine learning is a good direction for EGI. Existing or successful use case are the most persuasive if there is any. If any EGI project or user community has included similar machine learning tools or framework in their computing models, it will be much of reference and be inspiring to introduce their deployment experiences, especially to use what kind of tools to resolve what problems in what way. 2. There are many production open source machine learning tools being used for various application fields. In choosing Caffe in this deliverable, it is better to explain the criteria or measurement for this selection in addition to performance. 3. Some applications might have only few data sets in the early stage. When deploying the supervised learning framework, limitation of data set volume would lead to low quality result because of insufficient learning data for example. So, the guideline of deploying the recommended machine learning framework or tool should be included in the future. 4. It’s good that this deliverable already included benchmarking for the suggested tool. However, Stress test by raw data which is largely unlabeled and un-categorized should be also covered in the future. 5. Integration with EGI infrastructure and Federated Cloud is also an important plan. Whenever it is possible, it is better to include the strategy how to realize this. 6. Verification by user community is also an important step for the machine learning development of EGI. The strategy of how to support and collaboration with early adopter or targeted/volunteer projects should be described in the future plan. |
| **Response from Author:** |
|  |

**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,  …)** |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**English and other corrections:**

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

1. P5, the last sentence of first paragraph of the Introduction section. “Besides being able to learn, that is, the framework must include a neural network able to be trained, it must be possible to run in on the cloud taking advantage of its benefits.”
   1. Should be “In the cloud”
   2. “taking advantage of its benefits” is semantic repetition and is not concrete enough. Better to revise the whole sentence to be affirmative and specific.
2. P11, the last sentence of the paragraph below Figure 2, “Keeping in mind those considerations, given the provided framework, building a consistent and goo data set will affect significantly on the final accuracy of the framework.” 🡪 “goo” should be “good”.