Table of Contents

Table of Contents................................................................. 1
1. General Information .................................................................. 1
2. Summary ................................................................................. 2
3. Introduction ............................................................................. 3
4. Project Budget .......................................................................... 3
5. Lessons Learned.......................................................................... 3
   5.1. Project timelines for the 2015 Phase ......................................... 4
   5.2. Steering Committee and Project Team ........................................ 5
   5.3. Community of Interest ............................................................. 6
   5.4. Communication ...................................................................... 6
6. Recommendations for the 2016 phase ........................................... 7
7. Conclusions ................................................................................ 7

1. General Information

<table>
<thead>
<tr>
<th>Project Title:</th>
<th>Ontologies Mapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brief Project Description:</td>
<td>Ontologies Mapping guidelines, tool and service.</td>
</tr>
<tr>
<td>Prepared By:</td>
<td>Ian D Harrow</td>
</tr>
<tr>
<td>Date:</td>
<td>22nd December 2015</td>
</tr>
</tbody>
</table>
2. Summary

Achievements of the 2015 phase

- Established a working project team which is supported by the Steering Committee and Pistoia Operations.
- Engaged a Community of Interest of considerable size and influence in the ontologies field.
- Delivered a set of guidelines for best practise and a checklist to exemplify their value.
- Delivered the detailed requirements for an ontologies mapping tool for the next phase.
- Conducted an initial survey of existing mapping tools which concluded that some of these are contenders to substantially meet our requirements.
- Exceeded the deliverables agreed at the start of the project with a guidelines checklist and an initial survey of existing tools.
- Delivered these achievements on time and under budget (by 39%).

Deliverables for the 2016 phase

- To develop a process (most likely an RFP or challenge) to evaluate and select existing mapping tools to ensure at least one meets our requirements to complete by mid 2016.
- To define our requirements for an ontologies mapping service which builds on the deliverables of the 2015 phase to complete by mid 2016.
- To implement the requirements through evaluation and selection of a provider for an ontologies mapping service using an RFP process to complete by the end of 2016.
- To deliver the proposed plan for the 2016 phase with existing funds allocated already to the project, if funders agree to give their support.

Benefits from the 2016 phase

The 2016 phase will result in the selection of a “state of the art” ontologies mapping tool that meets the requirements defined by the project team during the 2015 phase. This tool can be used in conjunction with the best practise guidelines checklist to perform a “health check” on source ontologies being considered for mapping by a tool applied to the domain “test case” disease, phenotype and experimental investigation.

The project will also develop requirements for an ontologies mapping service, resulting in an agreed “service level agreement (SLA)” standard to satisfy the demands of the project participants, especially the funders. These service requirements will be used to implement an Request For Proposals (RFP) process to evaluate and select providers for an ontologies mapping service which will be available to project participants by the end of 2016.

The proposed plan for the 2016 phase offers a highly cost effective return on investment for delivery of a “state of the art” ontologies mapping tool and service in an efficient and timely manner with shared cost and risk by the funders.
3. Introduction

The aim of this project report is to review the outcome of the Ontologies Mapping project at the end of the first phase against what was planned, identify the key factors which contributed to the success of the project and the areas which could have been improved upon.

The Ontologies Mapping project has been set up to encourage best practices and to create better tools and services for ontology management and mapping in the Life Sciences.

Ontologies can include hierarchical relationships; taxonomies; classifications and vocabularies which are becoming increasingly important for support of research and development. They have numerous applications such as knowledge management, data integration and text mining where researchers need to analyse large quantities of complex data as part of their daily work. The Ontologies Mapping Project will give users access to standardised tools and methodologies to visualise source ontologies, map between ontologies and to exploit ontology structure. The outcome of this project will be to help users to better integrate, understand and analyse their data more effectively.

4. Project Budget

<table>
<thead>
<tr>
<th></th>
<th>Projected Costs</th>
<th>Actual Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Budget</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part-time Project Analyst / Manager</td>
<td>$ 58,500.00</td>
<td>$ 35,449.00</td>
</tr>
<tr>
<td>Publicity</td>
<td>$ 1,000.00</td>
<td>$ 544.55</td>
</tr>
<tr>
<td>Grand Total</td>
<td>$ 59,500.00</td>
<td>$ 36,034.55</td>
</tr>
<tr>
<td>Project underspend</td>
<td></td>
<td>$ 23,465.45</td>
</tr>
<tr>
<td>Available Funding</td>
<td></td>
<td>$ 98,965.45</td>
</tr>
</tbody>
</table>

The project was able to reduce the actual costs compared to projected costs by $23,465 (39%).

The original objectives were achieved in significantly less project management time than planned, resulting in fewer consulting days being required. Publicity was achieved through numerous networking events, monthly community of interest meetings, web site pieces and press release. Most speaking engagements and conference attendance were undertaken by project team members who were based nearby.

The available funding ($98,965) is sufficient to cover the costs of implementing the project plan for the 2016 phase of the project, which is described in the next section of this report. The funders on the steering committee are asked to endorse this proposal so that the available funding can be used to implement the proposed project plan.
5. Lessons Learned

5.1. Project timelines for the 2015 Phase

The idea for Ontologies Mapping project was proposed through the Pistoia Alliance Ideas Portfolio Platform (IP3) which was selected by the Operations Team and Pistoia Board for development of a formal business case as published in IP3. GSK, Merck & Co, Novartis and Roche agreed to fund the initial phase of the project as shown in the project timeline and deliverables. More recently, BIOVIA 3DS have also contributed funds to the project.

**Project Timeline and Deliverables**

The project started in April 2015 and was expected to deliver the agreed initial objectives for the 2015 phase by the end of 2015. These aimed to 1) select an Ontologies domain as a “test case”; 2) provide guidelines for minimal standards and best practices for the application and mapping of ontologies and 3) define requirements for an Ontologies Mapping tool. In addition, the project team has developed a checklist process to encourage adoption of the guidelines. Furthermore, the project team has used the mapping tool requirements to determined whether any existing mapping tools substantially meet them.

Delivery of the objectives as planned for 2015 has been due to the commitment of the project team which has completed tasks in an efficient manner. Plans for the 2016 phase of the project are described in the next section.
5.2. Steering Committee and Project Team

The Project Steering committee, which includes the funders Roche, GSK, Novartis, Merck & Co and BIOVIA 3DS, is responsible for making decisions, informed by recommendations from the Project Team. They comprise of interested Pistoia Alliance members (including the funders) and meet biweekly. The Project Team consults with the Community of Interest each month. The Community of Interest is open to any organisation or individual with relevant skills and experience. The relationship between the three groups is illustrated in the Project Structure and Communication diagram below.

The project team consulted with members of the Pistoia Alliance and Community of Interest to inform the selection of Disease, Phenotype and Experimental Investigation as “test cases” to serve the project. Guidelines for best practise to support application and mapping of ontologies have delivered by the project team as planned. These have served as a useful checklist for selection of source ontologies for mapping to support a particular domain or application.

The project team have delivered the detailed requirements for an Ontologies Mapping tool as planned. Comparison of existing ontologies mapping tools with our requirements through an initial survey has identified that they are likely to be met substantially by a few commercial and academic providers. The project team recommends evaluation and selection from existing mapping tools through a process such as Request For Proposal (RFP) or challenge. In tandem with this, the project will define the requirements for an ontologies mapping service. This will be used to evaluate and select at least one service provider, through an
RFP process, to complete by the end of 2016. These deliverables and timings are shown in the figure on page 4 of this report.

There are eighteen members of the project team including the project manager and Pistoia Operations and most have attended more than 50% of the biweekly meetings since the start of the project in April 2015. Ten different organisations (Roche – project champion, Instem – project leader, Merck & Co, Novartis, GSK, Eagle Genomics, Osthaus, Jansen J&J, Bayer and Takeda) are represented by these team members. A new funding member has joined the project team recently from BIOVIA 3DS.

It has been important to maintain the enthusiasm and engagement of the project team by close interaction with the decision-making of the steering committee and a sense of making progress with the project. However, this is not sufficient without their company having an interest in consumption and exploitation of the project deliverables.

The time availability of team members sometimes appeared to be problematic for some, which seemed to depend on other on-going commitments for their company. However, this challenge was alleviated to some extent by the use of the project wiki on Confluence for the collaborative development of project documents.

5.3. Community of Interest

The Community of Interest (CoI) consists of over 70 individuals with expertise and experience with ontologies representing close to 50 different organisations. This includes more Pistoia members and some key academics in the ontology field such as Prof Barry Smith, one of the founders of the OBO foundry and Helen Parkinson of EMBL-EBI and the GA4GH consortium. The monthly CoI meetings have been attended well throughout and are valuable consultative body for the project. They are already proving to be a powerful mechanism for influence and adoption which is important to maximise the impact of this project.

5.4. Communication

Internal communication was primarily through email and a project wiki on Confluence for the collaborative development of project documents. Biweekly project team meetings and monthly CoI meetings have made use of the GoToMeetings videoconference tool which has worked well as the preferred conversational method, along with the wiki and emails.

Communication is not just about the mechanism used, but also the way in which it is done. The project worked hard to ensure that there was an open and supportive atmosphere where everyone’s opinion was valued and taken seriously. This is vital in keeping a volunteer workforce engaged as a project like this must be an enjoyable experience for those in it.

External communication consisted primarily of presentations and receptions at key conferences, workshops and webinars. Face to face events led to extremely useful conversations and enabled the team to make contacts and learn more about the way the project was perceived.
6. Recommendations for the 2016 phase

- Use the available funding to implement the proposed project plan for the 2016 phase:
  - To select existing mapping tools from commercial and academic sources by evaluation of their ability to meet our requirements through an RFP or challenge process by mid 2016.
  - To define an ontologies mapping service by mid 2016 and select service provider(s) through a second RFP process to complete by end of 2016.
- To maintain flexibility in the project planning to allow the project team to respond to implementation of RFPs for the mapping tool and service.
- To focus publicity on personal attendance at conferences, workshops and the CoI to gather feedback on the project and maintain contact with key influencers and interested organisations.
- To maintain enthusiasm in the project team by celebrating success and keeping a positive, open and inclusive style of leadership.
- To advertise RFP or challenge processes extensively and be aware of barriers deterring respondents from participating.
- To ensure that RFP requirements are not too onerous up front.

Note: The last two recommendations are derived from the successful HELM project

7. Conclusions

The Ontologies Mapping project has completed successfully the 2015 phase. It has delivered beyond the agreed objectives on time with a 39% underspend of the projected budget. The cost of the proposed plan for the 2016 phase can be met through carryover of existing available budget of $98,965, if agreed by the funders.