BioSharing Survey - Summary

Peter McQuilton[1], Pascale Gaudet[2] and Susanna-Assunta Sansone[1].

1. Oxford e-Research Centre, University of Oxford, UK
2. CALIPHO, Swiss Institute of Bioinformatics, CH

EXECUTIVE SUMMARY

A 10-question survey was conducted from the 18th December 2015 to 22nd February 2016 to gather users' views on which features and content they need to make informed decisions, e.g. on how to best select standards and understand their maturity, or to find the databases that implement them. A link to the questions can be found here: https://bd2kccc.org/2016/01/15/biosharing-standards-registry-survey.

The 533 respondents (operating in the life, environmental and biomedical sciences) ranged from researchers, standard developers, database curators and industry scientists to librarians, funders and journal editors, and were drawn from the ELIXIR, NIH BD2K communities, the BioSharing RDA and Force11 working group and the International Society for Biocuration.

This document provides a brief overview of the survey results, and relates these to current (and planned) functionality and data in BioSharing (https://www.biosharing.org). The results show that BioSharing already fulfills ~80% of user needs, which (i) are not limited to standards but extend to databases and policies, and (ii) require an (curated) informational and educational system, not simply a registry.

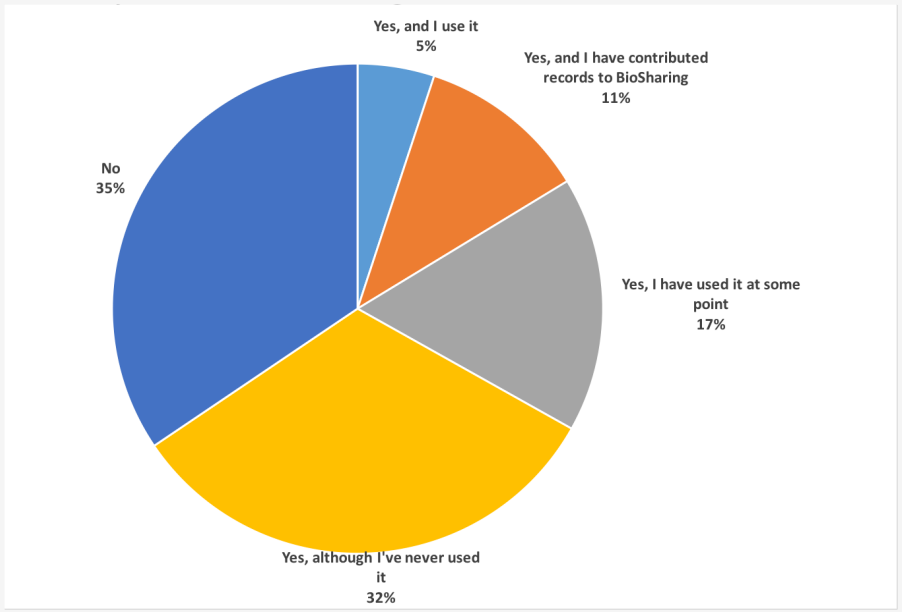
RESULT SUMMARY

The survey received 533 responses from researchers, standard developers, database curators, librarians, funding agencies, industry scientists and journal editors. A breakdown of their profile types is shown in the table below; respondents were allowed to select as many 'positions' as appropriate.

Position	Number
Researcher	323
Tool/database developer	274
Standard developer/maintainer	206
Data curator	151
Data manager	150
Journal publisher/editor	31

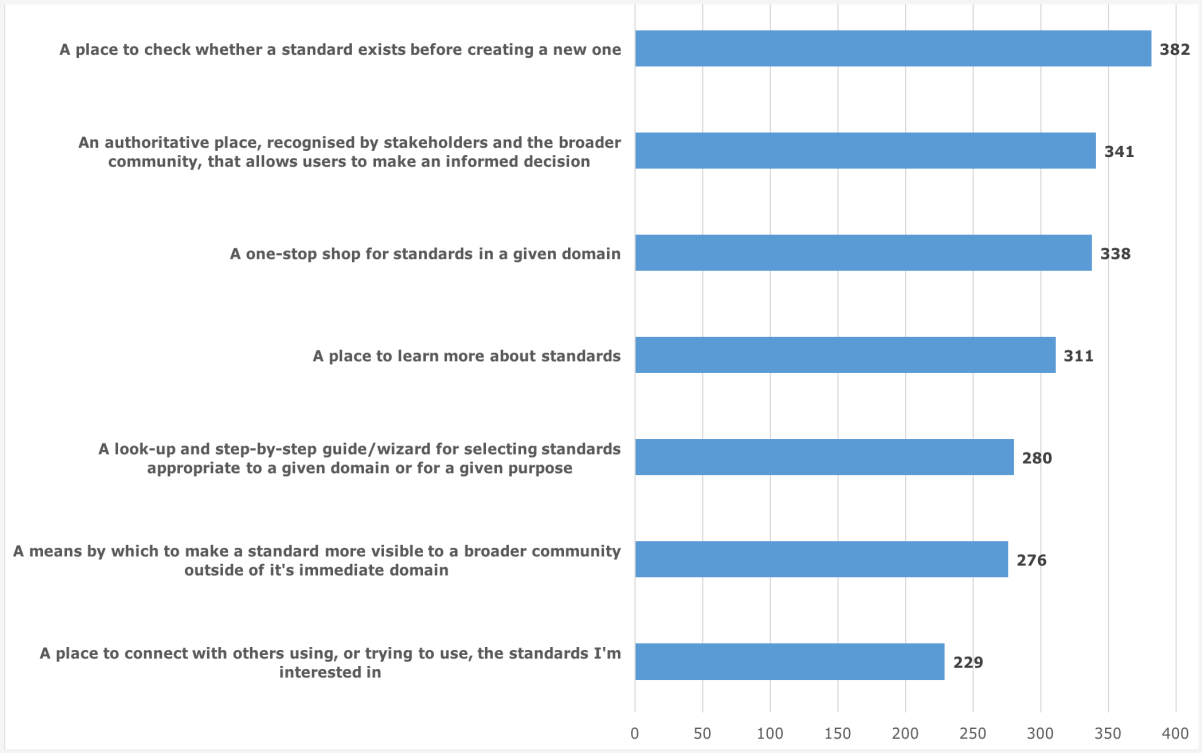
Librarian	20
Funding agency	20

Approximately 25% of respondents are associated with ELIXIR and 21% with NIH BD2K. Replies came from all over the world, with a predictable concentration in Europe and the USA. Approximately 65% of the respondents know of BioSharing, of which 33% have used it.



The following sections are generated with data from all respondents, not only the ELIXIR member subset.

1. **What do you need from a standards registry?** Respondents were allowed to select as many as appropriate.

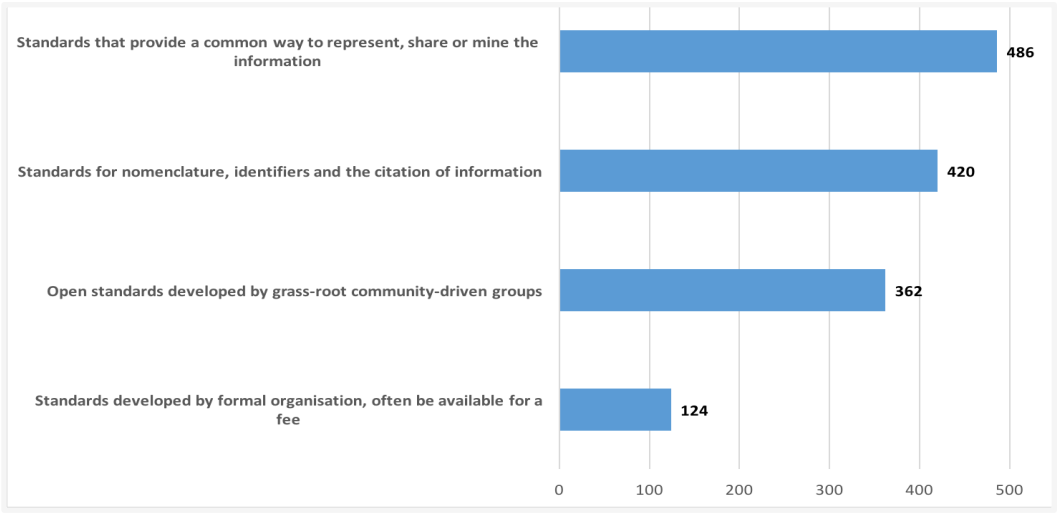


BioSharing already provides many of the requirements selected by the respondents, although new features and improvements are required to completely fulfill the diverse needs of the different stakeholders (non-technical vs. technical).

Despite the high number of respondents with a technical background, only a few (<10) have requested an API or some kind of programmatic access to the data (e.g. to have a record available for download in JSON format), or the creation of validation tools to check whether a dataset conforms to a particular standard. Although all these features are desirable, these results will help prioritise the work to be done.

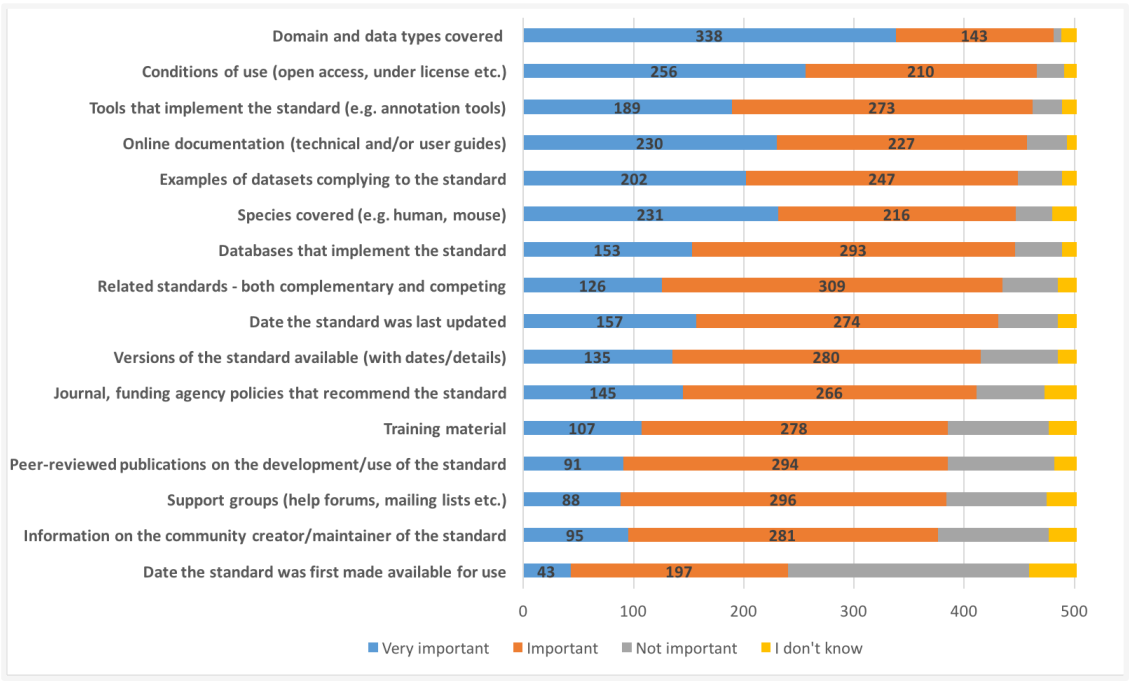
To address the top requirements - the AskBioSharing page has just been launched as a beta feature to guides users through standards, databases and data policies, allowing them to select resources relevant to their domain and species.

2. **What type of standards would respondents be interested in seeing in a standards registry?** Respondents were allowed to select as many as appropriate.



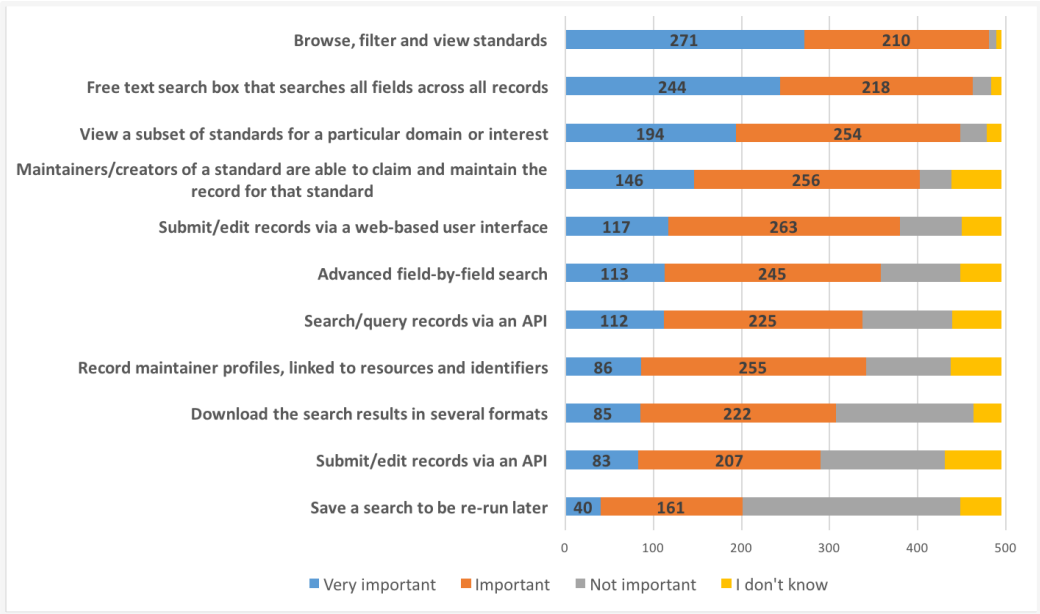
Clearly, respondents are looking for a comprehensive registry grouping all types of digital standards that facilitate representing, sharing and mining of information; also naming, identification and categorization of data that aids unambiguous citations. Currently BioSharing focuses on metadata standards, but other types can be easily added.

3. **What information should a standards registry capture about a standard?** Respondents were allowed to choose among: 'Very important', 'Important', 'Not important', and 'I don't know'.



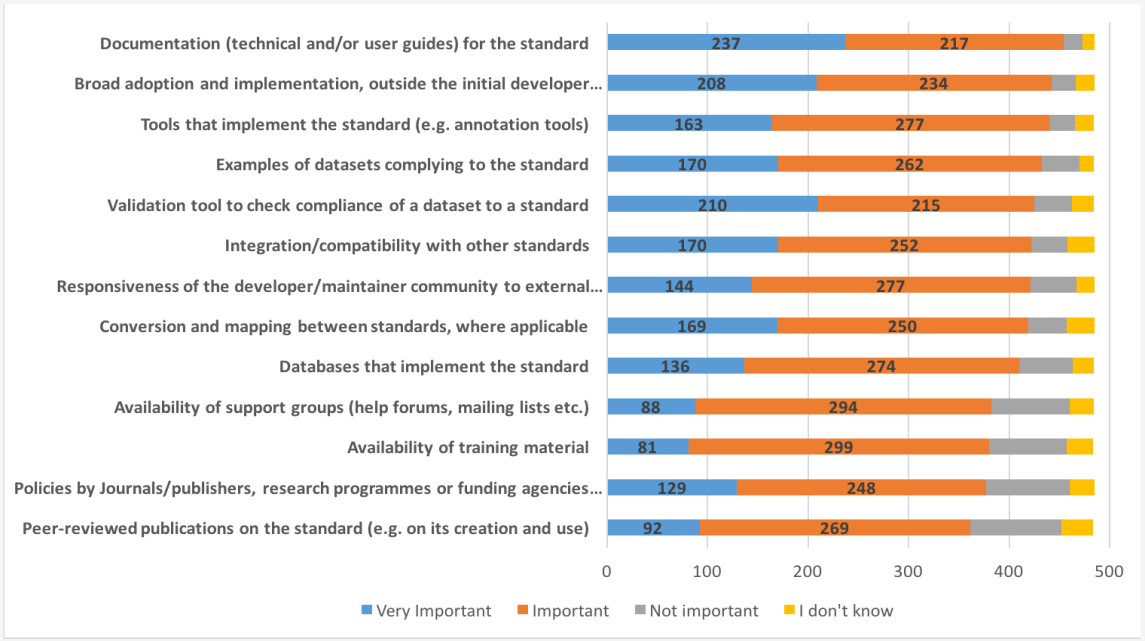
This question gives a clear indication of the most sought after features, and equally, of those which are considered less necessary. Of these features, BioSharing already captures 81% (13/16) and highlights the value of cross-linking standards to policies and databases, and of course cross-referencing other resources with training material, tools and so on.

4. **What functionality should a standards registry have?** Respondents were allowed to choose from: 'Very important', 'Important', 'Not important', and 'I don't know'.



This question gives an indication on the kind of interactions users would like to have with the registry. BioSharing already accommodates 64% (7/11) of the requested functionality.

5. **Which indicators should a standards registry capture to assess the maturity, adoption and use of a standard?** Respondents were allowed to choose among: 'Very important', 'Important', 'Not important', and 'I don't know'.



All the indicators listed in this question were considered either very important or important by over 2/3 of the respondents. Approximately 40% of the indicators are already implemented in BioSharing, with the missing indicators overlapping

with some of the features already requested in earlier questions. Once more, this indicates the value that BioSharing brings by not just focusing on standards but cross-linking them to policies and databases.