



EXCELERATE Deliverable D1.5

Project Title:	ELIXIR-EXCELERATE: Fast-track ELIXIR implementation and drive early user exploitation across the life sciences	
Project Acronym:	ELIXIR-EXCELERATE	
Grant agreement no.:	676559	
	H2020-INFRADEV-2014-2015/H2020-INFRADEV-1-2015-1	
Deliverable title:	Incorporation of monitoring statistics and benchmarking results in registry releases	
WP No.	WP1	
Lead Beneficiary:	38 - DTU	
WP Title	Tools Interoperability and Service Registry	
Contractual delivery date:	30 June 2018	
Actual delivery date:	26 June 2018	
WP leader:	Søren Brunak and Alfonso Valencia	38 - DTU; 12 - BSC
Partner(s) contributing to this deliverable:	38 - DTU; 5 - UTARTU, 7 - CNIO (partner terminated 31/01/2018), 10 - IRB, 21 - UiB, 25 - SIB, 29 - IP, 35 - MU	

Authors and Contributors:

Jon Ison, DTU; Hans Ienasescu (DK), Matúš Kalaš (NO), Hervé Ménager (FR), Piotr Chmura (DK), Veit Schwämmle (DK), Magnus Palmblad (NL), Henriette Husum Bak-Jensen (DK), Carlos Oscar Sorzano (ES), Tomáš Raček (CZ), Séverine Duvaud (CH), Marta Villegas (ES), Wojtek Dabrowski (DE), Olivier Sallou (FR), Juergen Haas (CH), Ahto Salumets (EE), Hedi Peterson (EE), Jacques van Helden (FR), Peter Løngreen (DK), Søren Brunak (DK)

Reviewers:

None

1. Table of contents

Table of contents	2
Executive Summary	2
Impact	3
Project objectives	4
Delivery and schedule	4
Adjustments made	4
Background information	4
Appendix 1: Incorporation of monitoring statistics and benchmarking results in registry releases	9
8.1 Technical overview	9
8.2 Review of software quality metrics	9
8.3 Tool Information Standard	15
8.3.1 Description of the standard	15
8.3.1 Application of the standard to software quality metrics	17
8.4 Implementation & plans	19
8.4.1. bio.tools curation	19
8.4.2. bio.tools features	20
8.4.2.1 Ranking and filtering of bio.tools search results	20
8.4.2.2 Metrics cards	20
8.4.3 Integration with OpenEBench	21
8.4.3.1 OpenEBench widgets	21
8.4.3.2 OpenEBench API	22
8.4.4 Challenges	23
8.5 Summary of future work	23

2. Executive Summary

The objective of EXCELERATE Deliverable 1.5 is to review software metrics and assess their usefulness in the context of the bio.tools (<https://bio.tools>) registry, with view to expose in bio.tools, metrics from scientific performance benchmarking, technical monitoring, and of software or service quality in general. The core aim is to provide the end-user with robust, scientifically relevant measures of quality and performance, to guide them in their comparison and selection of tools through: 1) enabling end-users to sort and filter bio.tools search results by various criteria, and 2) to render labels on bio.tools Tool Cards, perhaps even (eventually) a formal ELIXIR quality stamp.

This deliverable report describes research into the problem and implementation of prototype solutions towards the stated objectives and core aim. This report describes work done with ELIXIR-EXCELERATE resources:

- scoping the metrics landscape, including types of metrics, key initiatives and stakeholders.
- a thorough **review** of aspects of **software quality**, corresponding **metrics** or indicators, and their implementation **status or feasibility** from a bio.tools perspective. We identified 9 core aspects of software quality and 61 atomic metrics or indicators, of which 33 are supported by biotoolsSchema and thus calculable from software annotations in bio.tools; the rest could be provided by OpenEBench or third parties.
- publication of a “respectable beta” **Tool Information Standard** (developing the candidate described in D1.2 and D1.3); this provides a framework for improvement of software metadata necessary for the calculation of many types of metric, and is an important component in branding bio.tools as a trusted source of quality tool information.
- proposals for **labelling bio.tools Tool Cards**, based on the Tool Information Standard and aggregated metrics that are calculable from software annotations in bio.tools.
- incorporation in bio.tools of a prototype **widget** for **rendering metrics** provided by OpenEBench (described in D2.2) and suggestions on how the bio.tools:OpenEBench integration can be improved.

The foundation is now mostly in place to calculate and render in bio.tools, metrics reflecting various aspects of software quality and performance. We are, however, still some way off achieving the core aim which turns out to be a hard and multi-faceted problem: a consensus upon exactly what (aggregated) metrics and corresponding labels to use in bio.tools is needed, amongst the various stakeholders who are approaching this problem from different angles. We anticipate however, as a matter of policy, for any metric to be rendered in bio.tools it must be both *objective* and *transparent*. Work is ongoing in all areas, to provide the end-user with useful measures of tool quality and performance. The report describes progress made thus far and includes a summary of anticipated developments.

3. Impact

The impact is that there is now an emerging, community-defined information standard for software information, integrated within an infrastructure spanning rigorous data syntax and semantics, through to curation and best practice guidelines. This solid technical foundation already supports over half of 61 key metrics or indicator of software quality. It can, in due course, be availed upon not only by bio.tools to provide useful ways to sort, filter and label tools, but by the 762 individual contributors of bio.tools content, and all ELIXIR nodes and major European service providers, to improve the quality of software and software metadata served to Life Scientists globally.

4. Project objectives

With this deliverable, the project has reached or the deliverable has contributed to the following objectives:

No.	Objective	Yes	No
1	Deliver a discovery portal built upon a federated curation of a wide range of key resources for bioinformatics resources world-wide.	x	
2	Service monitoring, resource integration, interoperability aspects, and community centred benchmarking efforts.	x	
3	Deliver impact for end-users across academia, health organizations, and industry.	x	

5. Delivery and schedule

The delivery is delayed: Yes No

6. Adjustments made

N/A

7. Background information

Background information on this WP as originally indicated in the description of action (DoA) is included here for reference.

Work package number	1	Start date or starting event:	month 1
Work package title	Tools Interoperability and Service Registry		
Lead	Søren Brunak (DK) and Alfonso Valencia (ES)		
Participant number and person months per participant 1 – EMBL 12; 2 – UOXF 6; 5 – UTARTU 43; 10 - IRB 13; 12 - BSC 11; 17 - INESC-ID 1.24; 21 – UiB 18.00; 25 – SIB 9.5; 26 – CNRS 9; 29 – IP 12; 35 – MU 18.2; 38 - DTU 26 (+ UCPH 25 + AU 25 (LTPs))			

Objectives

WP1 will deliver a discovery portal built upon a federated curation of a wide range of key resources for bioinformatics resources world-wide.

It will involve service monitoring, resource integration, interoperability aspects, and community centred benchmarking efforts. All activities, including intensive user support, are focused around delivering impact for end-users across academia, health organizations, and industry. The ELIXIR Tools and DataServices Registry is the cornerstone of the WP.

Work Package Leads: Søren Brunak (DK) and Alfonso Valencia (ES)

Description of work and role of partners

WP1 - Tools Interoperability and Service Registry [Months: 1-48]

DTU, EMBL, UOXF, UTARTU, IRB, BSC, INESC-ID, UiB, SIB, CNRS, IP, MU

Based on its first release in January 2015, WP1 will further develop the ELIXIR registry mechanism, interfaces and content upkeep strategy. The WP contains plans for the development and extension of its functionality and scope (Tasks 1.1, 1.2 and 1.5). The federated curation of the registry will ensure comprehensive content and high quality annotations, both of which are essential for the sustainable impact of the registry in the community. Scientific and technical consistency and utility will be achieved by using the EDAM controlled vocabulary. Exposing the results of efforts addressing tool benchmarking and monitoring of the resources listed in the registry will provide the end-user with a robust, scientifically relevant measure of tool quality and performance. Furthermore, the work on workbench integration and interoperability will lower the cost to developers of integrating their resources in key workflow environments, and assist the users with establishing and updating their day-to-day workflows. Finally, WP1 contains plans for comprehensive, registry related user support, which will ensure impact for users, and a dynamic management element, including marketing and community development to build the federated organization behind the registry.

The user-centric approach will thus stand as the guiding principle for the entire portal and guard its relevance to the community.

Task 1.1: Federated Registry Curation (96PM)

This task will deliver essential scientific and technical coverage in the registry and the vocabulary (EDAM) that underpins registry consistency and utility. A major community curation effort is required, including vocabulary development, resource annotation and registration. To ensure that the curation is high quality and sustainable, it must be federated across registry stakeholders, hence a major priority is building and supporting the community of federated curators. In tandem, the curation will be accompanied by focused software and other technical developments, that automate, validate and embed the curation process in relevant software systems; the essential underpinning of sustainability.

The registry has two primary purposes; to help discover tools and services and use them. Discovery means to find, understand, compare and select. It is a prerequisite to (inter)operability, which demands a precise understanding of software dependencies. Our approach is based on the acceptance that software interoperability will, for the foreseeable future, be implemented primarily by developers rather than intelligent

software agents. We will therefore, once a comprehensive set of ELIXIR Node resources are described in basic detail, extend the curation of the registry to annotate, using EDAM Format URIs (unified resource identifiers), the data formats that are supported by tools and data services.

From this, we will analyse the format-usage landscape to provide a basis for targeted software developments to improve interoperability of registered resources. We foresee these developments, which might include conversion of tools to use common formats, and development of format- converter software where needed, to be facilitated via the Matchmaking Service mechanism (D1.5).

The registry scope will be: 1. Comprehensive coverage of ELIXIR Node resources, including tools, data services (APIs) and host databases, prioritising ELIXIR-badged services and new resources from the Use Cases. 2. Coverage of other biomedical science Research Infrastructures (RIs), and key resources beyond ELIXIR (European and non-European).

A task force will be comprised of ontology developers, curators, scientific domain experts and relevant technical experts. It will run Curation and Usability hackathons with the recurrent theme of curation: resource annotation and registration, with necessary EDAM development. To facilitate networking and community build-up, two types of social event will be combined with the hackathons: 1. Knowledge Exchange Workshops, including representatives of relevant

infrastructures, institutes and projects, on themes related to the registry suggested by the community. 2. Cross-domain

Strategy Workshops to gather technical officers from ELIXIR Nodes, RIs, key resources, and other key initiatives, to discuss and develop common approaches for registry curation across RIs internationally.

EDAM provides the registry with a consistent vocabulary for topics (general scientific and technical disciplines), operations (tool functions), types of data, and specific data formats and data identifiers. Task 1.1 will work with the existing EDAM community, develop its open governance and contribution mechanisms and deliver essential utilities to ensure that maintenance, validation and community development is sustainable in the long term. We will assess and validate coverage by correlating EDAM concepts to terms used for curation, which will then inform and drive necessary additions and desirable clean-ups (removal of concepts). We will develop focused essential utilities for EDAM maintenance including automation of the release process, basic validation of content, reporting of changes between versions, deployment to ontology browsers such as BioPortal and OLS, technical integration of EDAM with applications including the registry and others, mapping of provider-supplied terms and phrases to EDAM, and revise annotation upon new EDAM releases.

To underpin the sustainability of the federated curation, this task will deliver focused software and other technical developments that will automate the registration and update of provider-supplied information, leveraging their own local software infrastructure where possible. We will work with providers to support them in doing this, and, where possible, adapt technically the local solutions to make them more broadly applicable to others. Further, in order to facilitate coverage, all relevant resource providers will be given smooth and convenient access to resource registration. This will be achieved by a combination of simple-to-obtain local login accounts and opening for using eduGAIN authentication to register resources.

Finally, this task will ensure that registered resource are citable, discoverable by the major search engines, and are placed in scientific context. It will also include technical

mark-up to support “Semantic Web” applications, e.g. Schema.org-compatible microdata or RDFa to support Google “rich snippets” and other structured search results in the major browsers. Hence, the registry will promote the registered resources and deliver impact for developers and institutes by making resources rank higher in search results and hence more findable.

Task 1.1 partners: DK, NO, FR, CH, CZ, EMBL-EBI, PT

Task 1.2: Benchmarking and Monitoring (15PM)

This task will support the monitoring and community benchmarking of analytical tools, in a systematic and sustainable way e.g. based on the efforts in WP2. Firstly, it will review the existing service quality and performance metrics and assess their usefulness in the context of a registry. This may require development of a light-weight controlled vocabulary capturing the concepts distilled from the preparatory activities above and those of WP2.

Task 1.2 partners: DK, ES, CZ, CH

Task 1.3: Workbench integration and interoperability (36PM)

There is general trend towards the use of workflows as a preferred environment for the convenient use of tools and data access, especially when resources must be used in combination with one another. This task will boost convenience and resource interoperability by implementing a Workbench Integration Enabler service that will develop the vision “register your software once - get it supported everywhere”. Technically, this service will translate the description of any tool or service that is registered in the Tools and Data Services Registry into the metadata format required by the existing major workbenches, including Mobyle, Galaxy and Taverna. Furthermore, we will develop a new, lightweight Service Launchpad for running tools and services which have programmatic access and which can be invoked using information available in the registry.

To develop the Enabler Service, we will align the registry software description model and the schemas used by the workbench systems or required by the Launchpad, and subsequently revise the model and schemas to facilitate the metadata transfer. Furthermore, to prove the principle, new high priority tools and services, including those developed in the Use Cases.

Task 1.3 partners: DK, EE, FR, CH, PT

Task 1.4: User support and derived registry development (36.7PM)

This task will provide direct and indirect user support to deliver impact for ELIXIR end-users. Direct support will be achieved primarily by leveraging the existing and highly popular user bioinformatics forums (BioStars, BioPlanet etc.).

A User-support specialist will patrol such forums and respond to questions in one of four ways: 1) Where resources answering to the Users needs exist in the registry, a link to them in the registry will be provided via our API. 2) Where resources exist in the registry, but the registry API cannot be used to answer the question directly, they will request new features of the API and in so doing drive development of the Query Interface. 3) Where an appropriate resource exists but has not been registered, they will request the appropriate registry curator add it to the registry. 4) Where a registered resource exists that is close, but not quite what is required, they will forward feature requests to the appropriate developers, possibly via the Matchmaking Service (D1.5).

Indirect user support will be achieved primarily by ensuring the registry interfaces are highly usable and match very closely the needs of the user. To achieve this, we will run user experience sessions during the Curation and Usability community. Scientific and technical consistency and utility will be achieved by using the EDAM controlled vocabulary.

Exposing the results of efforts addressing tool benchmarking and monitoring of the resources listed in the registry will provide the end-user with a robust, scientifically relevant measure of tool quality and performance. Furthermore, the hackathons (see Task 1.1) in order to evaluate usability. We will develop comprehensive Good Practice Guidelines for the curation of the registry in all aspects, but in particular the annotation of common types of resources using EDAM.

We will also participate in the development of an ELIXIR Experts Registry where users can discover relevant expertise within the ELIXIR network, and an ELIXIR User Helpdesk to answer general questions concerning use of the registry, forwarding specialised scientific and technical enquiries to relevant experts.

Task 1.4 partners: DK, CH

Task 1.5: Management, marketing and community build-up (46PM)

This task will build the federated organisation primarily by identifying and facilitating key collaborations between registry stakeholders. This will be achieved by organising 'Resource Synergy Meetings', where we will identify and encourage targeted software developments, e.g. to coordinate curation and data sharing. We will also promote resource integration and usability, e.g. by cross-linking resources and through API harmonization. As a prerequisite to these Synergy Meetings, a Resource Metadata Catalogue, listing all relevant resources, their scientific and technical scope, and information fields (schema), will be compiled and used to compare providers and identify redundancies. We will also use these meeting to cross-link the Tools & Data Services Registry with other key ELIXIR registries, for example the Training Materials Registry, the ELIXIR Events Registry, and the Experts Registry.

This task will also develop an oversight and management strategy and leverage partners within and beyond the ELIXIR organisation to implement strategy. To drive delivery, it will identify and encourage collaboration, monitor actions, identify delays, and intervene where necessary. It will raise community awareness and therefore impact by contributing to a forceful marketing campaign via all appropriate marketing channels, including popular social media. It will provide support to funders, publishers and others at the EU and national level, that policy is aligned with the aims of the registry organisation.

Task 1.5 partners: DK

8. Appendix 1: Incorporation of monitoring statistics and benchmarking results in registry releases

8.1 Technical overview

This deliverable report is structured as follows:

- **Review of software quality metrics** (Section 3.2) considers types of metrics, key metrics initiatives and stakeholders
- **Tool Information Standard** (Section 3.3) summarises the publication of a “respectable beta” Tool Information Standard (developing the candidate described in D1.2 and D1.3): this provides a framework for improvement of the software metadata necessary for the calculation of many types of metric. We propose metrics to include in bio.tools, based purely on the Tool Information Standard and software annotations in bio.tools
- **Implementation & plans** (Section 3.4) summarising incorporation in bio.tools of a prototype widget for rendering metrics provided by OpenEBench (described in D2.2) and suggestions on how the bio.tools:OpenEBench integration can be improved. Ongoing and future work on bio.tools is also described.
- **Future work** (Section 3.5) summarises work left to do, described in Sections 3.2 - 3.4.

8.2 Review of software quality metrics

The 2017 EXCELERATE midterm review recommended “*devotion of effort to ensure that inclusion of tools in the catalogue is perceived as an important label of quality and value*”. In practical terms, this boils down to implementing rigorous methods for assessing tools and services that are registered in bio.tools, and labelling entries, perhaps (eventually) providing a formal ELIXIR quality stamp; an impactful “seal of approval”. The value here is corroborated by many end-user requests for flexible ways to sort and filter entries by different facets of the quality of the tool and service provided.

Metrics of software quality, and whether and how these should be used in bio.tools, has been subject to a great deal of intensive discussion, in different contexts, within ELIXIR and more broadly. We recognise various initiatives that are a potential source of metrics, variously with a similar, overlapping or different scope and perspective to bio.tools:

- **scientific performance** metrics¹ which assess the scientific performance of a tool according to some formalised benchmark (benchmark data, tool parameters, execution environment *etc.*): produced by various scientific communities and benchmarking competitions: in scope of EXCELERATE WP2
- **technical performance** metrics² assess the purely technical performance of a tool, *e.g.* when a tool was last tested on a test dataset, whether it passed the test, how long an online service has been available *etc.*: produced by some software groups: in scope of EXCELERATE WP2

¹ <https://biotools.sifterapp.com/issues/127>

² <https://biotools.sifterapp.com/issues/217>

- **data resource** metrics³ assessing a biological database and its provision as an online service, including *e.g.* the ELIXIR Core Data Resource⁴ criteria, and the proposed NIH metrics to assess the value of biomedical digital repositories⁵: in scope of EXCELERATE WP3
- **interoperability** metrics for assessing candidate ELIXIR Recommended Interoperability Resources (RIR)⁶: in scope of ELIXIR Interoperability Platform
- **software good practice** metrics⁷ assess compliance of a tool to recommended practice, *e.g.* top 10 metrics for life science software good practices⁸ and four simple recommendations to encourage best practices in research software (4OSS)⁹
- **software assessment criteria** of software sustainability, maintainability, and usability proposed by the Software Sustainability Institute¹⁰
- **national assessment criteria** of resources including software for inclusion in ELIXIR Node portfolios (the UK node being a good example of a Node with rigorous criteria)
- **FAIR principles** metrics¹¹ assess whether a tool is FAIR¹² (Findable, Accessible, Reusable, Interoperable): under consideration by many groups within and beyond ELIXIR: in scope of EXCELERATE WP5
- **social** metrics assess a tool's preferences of end-users (*e.g.* "likes", 5-star ratings *etc.*)

The scope is thus thematically broad with many stakeholders and moving parts, hence the focus of EXCELERATE WP1 is upon metrics which can be *calculated from data already available in bio.tools*, whilst remaining open to include metrics provided OpenEBench and third parties. Social metrics, whilst no doubt valuable in some contexts, are out of bio.tools scope. They could in future be provided through a "Tool Advisor" site closely linked to bio.tools, providing a social forum for tool developers and users, or perhaps by adapting an existing portal such as biostars.org¹³ or ELIXIR Virtual Coffee Room¹⁴.

Labelling of bio.tools entries could be impactful (both positively and negatively) on bio.tools content providers (individual tool developers and major service providers). The issue therefore is highly sensitive and crucial to get right: we anticipate, as a matter of policy, for any metric to be rendered in bio.tools it must be both **objective** and **transparent**. This excludes from the get-go for example user-ratings of tools, and any metric whose calculation method is not made public.

From the perspective of bio.tools, metrics fall into two basic categories:

- metrics which can be calculated automatically solely from bio.tools data: these are the focus of EXCELERATE WP1

³ <https://biotools.sifterapp.com/issues/218>

⁴ <https://www.elixir-europe.org/platforms/data/core-data-resources>

⁵ <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-16-133.html>

⁶ <https://www.elixir-europe.org/platforms/interoperability/rir-selection>

⁷ <https://biotools.sifterapp.com/issues/192>

⁸ <https://f1000research.com/articles/5-2000/v1>

⁹ <https://f1000research.com/articles/6-876/v1>

¹⁰ <https://www.software.ac.uk/sites/default/files/SSI-SoftwareEvaluationCriteria.pdf>

¹¹ <https://biotools.sifterapp.com/issues/226>

¹² <https://www.nature.com/articles/sdata201618>

¹³ <http://biostars.org>

¹⁴ <https://cafe.elixir.ut.ee/about/>

- metrics which must be calculated using data not in bio.tools: these are in scope of EXCELERATE WP2 (and others)

For purpose of discourse, we assign all metrics into one of the following primary types, reflecting different aspects of software quality:

- **Deployability** - how easily can the tool be made available for use?
- **Accessibility** - how easily can the tool be accessed?
- **Usability** - how easily can the tool be used?
- **Interoperability** - how easily can the tool be connected or integrated with other tools / resources?
- **Performance** - how well does the tool perform, both technically and scientifically?
- **Quality of service** - what is the quality of service and support around the the tool?
- **Participation & activity** - how active is the project and how easy is it to contribute?
- **Governance** - how developed and transparent is the governance of the tool?
- **Usage & impact** - how much is the tool used and what impact does it have?

In Table 1 (below) we summarise possible indicators or metrics for each of these aspects, and note the current support in biotoolsSchema¹⁵ (the information model used by bio.tools), what additional work is ongoing or required, or whether a metric is out of bio.tools scope. The list is not exhaustive, and includes mostly quantitative but also some qualitative metrics. Of the total of 61 candidate metrics/indicators, 33 are currently supported by biotoolsSchema (*i.e.* calculable from software annotations in bio.tools); the rest could be provided by OpenEBench or third parties. What remains to do, in collaboration with ELIXIR partners broadly, is to firm up and prioritise the list of software quality aspects and indicators/metrics, to guide future implementation (a few suggestions are given in Section 3.3.1).

Aspect	Indicators / metrics	Status
<i>Deployability - how easily can the tool be made available for use?</i>		
Administrator documentation	<ul style="list-style-type: none"> • Are there instructions on how to build, install and configure the tool? • Are the instructions adequate, including necessary details <i>e.g.</i> dependencies, operating system requirements <i>etc.</i>? 	biooolsSchema supports links to docs, but extension of controlled vocabulary (CV) for link types is needed. Adequacy of instructions is out of scope.
Dependency management	<ul style="list-style-type: none"> • Are dependencies managed automatically during the build process? 	Out of scope currently.
Packaging	<ul style="list-style-type: none"> • Has the tool been packaged or containerised and made publicly available, <i>e.g.</i> via BioContainers (DockerHub, QUAY.io), bioboxes, Debian/RPM, CPAN, BioConductor <i>etc.</i>? 	Crosslinking bio.tools and BioContainers is supported by ELIXIR Implementation Study (work ongoing).
Version information	<ul style="list-style-type: none"> • Is the tool versioned? • Is information on available (legacy) versions available? 	biotoolsSchema supports version information but exhaustive annotation of all tool versions in bio.tools is

¹⁵ <http://github.com/bio-tools/biotoolsschema>

		currently out of scope.
Accessibility - how easily can the tool be accessed?		
Licensing	<ul style="list-style-type: none"> Is the software licensed, and is the license a standard one? Is the software source code license recognised as “Open source”¹⁶ and/or “Free software”¹⁷? 	biooolsSchema provides a CV for license that supports such labels. License is extensively annotated in bio.tools.
Availability	<ul style="list-style-type: none"> Is the availability of the tool (source code and binaries) in a repo such as GitHub, GitLab or BitBucket documented? 	biooolsSchema supports links to downloads.
Restrictions	<ul style="list-style-type: none"> Are access or usage restrictions documented, <i>e.g.</i> limitations on who or how a tool may be used, cost, or general terms of use? 	biooolsSchema provides CVs for terms of use, accessibility and cost.
Usability - how easily can the tool be used?		
End-user documentation	<ul style="list-style-type: none"> Is there documentation for end-users <i>e.g.</i> tool homepage, readthedocs <i>etc.</i>? Is there a concise and clear description of what the tool does? Are all the tool parameters/options documented? Are there more detailed usage instructions, use-cases, examples <i>etc.</i>? Is the documentation correct and up-to-date? 	biooolsSchema supports links to documentation of various types. bio.tools tool “descriptions” are being improved according to Curation Guidelines (work in progress). Tool parameters / options <i>per se</i> are currently out of scope, but could extend documentation CV to support this. Correctness / up-to-dateness would need manual verification.
Deployments	<ul style="list-style-type: none"> Is there information on where the tool is deployed for use, <i>e.g.</i> as an online service (Web applications, Galaxy <i>etc.</i>)? Is there information on integrated systems <i>e.g.</i> workbenches and suites, that include the tool? 	biooolsSchema supports links but extension of CV for link types is needed. Tooling for crosslinking bio.tools and Galaxy is developed, and funding for it’s application sought from ELIXIR Tools Platform. biotoolsSchema extension for tool relationships to workbenches, suites <i>etc.</i> is planned (work in progress).
Training	<ul style="list-style-type: none"> Are there training materials about the tool <i>e.g.</i> training courses, tutorials <i>etc.</i>? 	biooolsSchema supports links to training materials.
Interoperability - how easily can the tool be connected or integrated with other tools / resources?		

¹⁶ OSI Open Source Definition (<https://opensource.org/docs/osd>)

¹⁷ FSF Free Software Definition (<http://www.gnu.org/philosophy/free-sw.html>)

Identifier	<ul style="list-style-type: none"> Does the tool have a unique identifier assigned by a recognised tool ID authority, <i>e.g.</i> toolID provided by bio.tools? 	All bio.tools entries are assigned a unique ID; a URL-safe version of the supplied tool name.
Formalised description	<ul style="list-style-type: none"> Is a formalised (machine-readable, schema-compliant) description of the tool available, <i>e.g.</i> biotoolsSchema (bio.tools), OpenAPI/Swagger, WSDL <i>etc.</i> ? Is the tool description formatted for a documentation generator system <i>e.g.</i> pydoc, JavaDoc, Doxygen <i>etc.</i>? 	biotoolsSchema provides a formalised tool description and tooling is developed to import OpenAPI descriptions. Tooling for translation/import from document generators is partially implemented, but annotation of such is currently out of scope.
Use of standards	<ul style="list-style-type: none"> Does the tool use community-recognized standards, <i>e.g.</i> standard file formats? Is the tool function (operations, I/O data types and formats) annotated using EDAM ontology? 	biotoolsSchema through its use of EDAM supports this directly.
Performance - how well does the tool perform, both technically and scientifically?		
Test framework	<ul style="list-style-type: none"> Are test data and scripts publicly available? Is the tool integrated into an automated test framework, <i>e.g.</i> with compilation tests, unit tests, functional tests, regression tests, system tests <i>etc.</i>? 	biotoolsSchema supports links to downloads of test data and scripts. Annotation of test framework is currently out of scope.
Technical performance	<ul style="list-style-type: none"> Are test results publicly available? Are data for online services publicly available, <i>e.g.</i> is the service available (up), uptime, response time? 	This can be delivered by linking to or rendering in bio.tools, data maintained by OpenEBench (work in progress). Partially implemented by OpenEBench widget (Section XXX).
Scientific performance	<ul style="list-style-type: none"> Has the tool been scientifically benchmarked? Is it integrated into a continuous benchmark framework? 	This can be delivered by linking to or rendering in bio.tools, data maintained by OpenEBench (work in progress).
Quality of service - what is the quality of service and support around the the tool?		
End-user support	<ul style="list-style-type: none"> Is support in using the tool <i>e.g.</i> via mailing list, helpdesk, issue tracker <i>etc.</i> available and documented? 	biotoolsSchema supports links to types indicated.
Quality assurance	<ul style="list-style-type: none"> Is there a rigorous (and documented) quality assurance process for software development, releases <i>etc.</i>? 	Out of bio.tools scope currently.
Responsiveness	<ul style="list-style-type: none"> Is there a publicly accessible issue tracker? How many resolved / unresolved issues? 	biotoolsSchema supports links to issues tracker, but rendering issue metadata

	<ul style="list-style-type: none"> • Time taken to resolve tracked issues? • How much recent activity in the issue tracker? 	(available from e.g. GitHub) is currently out of scope.
Participation & activity - how active is the project and how easy is it to contribute?		
Transparency	<ul style="list-style-type: none"> • Is there a public roadmap for the tools development? • Are there instructions on how to get involved? • Are there regular developer meetings and can potential contributors participate? 	Extension of biotoolsSchema CV for documentation types is needed. Meetings and openness are out of scope.
Activity	<ul style="list-style-type: none"> • Is there a public changelog? • Number and frequency of new features? • How many code commits, pull requests? 	Extension of biotoolsSchema CV for documentation types is needed. Rendering data on features, commits etc. (available from e.g. GitHub) is currently out of scope.
Community	<ul style="list-style-type: none"> • How many people are involved in the project, and from different sites? • How many external code commits, forks, derived projects etc.? 	Out of bio.tools scope currently.
Developer documentation	<ul style="list-style-type: none"> • e.g. contributions and communications policies for potential contributors? • Is there technical documentation to help potential contributors, e.g. architecture overview, code conventions, process etc.? 	Extension of biotoolsSchema CV for documentation types is needed.
Credits	<ul style="list-style-type: none"> • Are contributions properly acknowledged? 	biotoolsSchema provides extensive support for credits.
Governance - how developed & transparent is the governance of the tool?		
Governance model	<ul style="list-style-type: none"> • Is the governance of the project providing the tool or service documented? 	Extension of biotoolsSchema CV for documentation types is needed.
Advisory board	<ul style="list-style-type: none"> • Is the tool in the review remit of an independent advisory board? 	Out of bio.tools scope currently.
Sustainability plan	<ul style="list-style-type: none"> • Is a plan in place for long-term sustainability of the tool? 	Out of bio.tools scope currently.
Legal framework	<ul style="list-style-type: none"> • Is a legal framework e.g. for access, security, licensing etc. in place and documented? 	Out of bio.tools scope currently.
Usage & impact - how much is the tool used & what impact does it have?		
Downloads	<ul style="list-style-type: none"> • How many downloads of the source code, binaries, packages, containers etc.? 	Out of bio.tools scope currently - but attractive if it can be harvested automatically..

Usage statistics	<ul style="list-style-type: none"> • Are stats on tool usage available? • How many jobs run, frequency of usage <i>etc.</i>? 	Out of bio.tools scope - probably unfeasible.
Community served	<ul style="list-style-type: none"> • What is the size and measured demand of the community served, in context of the scientific field the tool serves? 	Out of bio.tools scope.
Publications	<ul style="list-style-type: none"> • Is there a publication that primarily describes the tool? • Are there other publications about the tool, <i>e.g.</i> comparisons, benchmarks, applications, reviews <i>etc.</i>? 	biotoolsSchema supports annotation of publication identifiers and types. Most bio.tools entries have a publication.
Citations	<ul style="list-style-type: none"> • How many citations of the tools primary publication? 	bio.tools renders citation information for publications about tools, and allows sorting by citation counts.
Altmetrics	<ul style="list-style-type: none"> • Are Altmetrics data <i>e.g.</i> attention score, "mentions", geographical breakdown <i>etc.</i> of the primary publication readily available? 	bio.tools Tool Cards include the Altmetrics widget, which links out to such data.
Use cases	<ul style="list-style-type: none"> • Are there documented use cases of the tool, <i>e.g.</i> in published research studies, production environments (major service centres) <i>etc.</i>? 	Out of bio.tools scope.
Case studies	<ul style="list-style-type: none"> • Are there documented case studies demonstrating the tool value and assessing <i>e.g.</i> what the impact would be if the tool wasn't available, whether there are alternatives and the implications of using those? 	Out of bio.tools scope.

Table 1. Indicators and metrics for aspects of software quality.

8.3 Tool Information Standard

8.3.1 Description of the standard

The **Tool Information Standard** is one component of a stack of interrelated technologies and guidelines (Figure 1) used by bio.tools:



Figure 1. Technology and guidelines used by bio.tools.

These components are (or will) be described in detail in other EXCELERATE deliverable reports. In summary:

- **biotoolsSchema**¹⁸ is a formalised XML schema (XSD) which defines a description model for bioinformatics software. It defines a **syntax** for software attributes that underpin many metrics.
- **EDAM ontology**¹⁹ is an ontology of well-established, familiar concepts that are prevalent within bioinformatics and computational biology. It defines the **semantics** - a controlled vocabulary - for software in scientific terms, and is essential for certain types of metric.
- **Tool Information Standard**²⁰ lists software attributes (from biotoolsSchema) that must be specified within a 5-tier scale of description completeness and quality. It provides a practical framework and metrics for description of individual tools, and the curation and improvement of collections such as bio.tools (see below).
- **Curation Guidelines**²¹ describe conventions for how each attribute should be specified when registering a tool in bio.tools. These human-readable and user-friendly guidelines provide information that goes beyond syntax and semantics provided by biotoolsSchema and EDAM.
- **Software Best Practice** can be abstracted from the Curation Guidelines, for example, general recommendations for the description / documentation of software in contexts other than bio.tools.

The Tool Information Standard - summarised as a work in progress in EXCELERATE D1.2 - has been significantly revised and simplified in light of community feedback²², producing a “respectable beta”. The standard is comprehensively documented, and now includes links to biotoolsSchema and the emerging Curation Guidelines. We have published it online:

<https://bio-tools.github.io/Tool-Information-Standard/>

The standard (Figure 2) consists of five lists software attributes that must be specified for a software description (e.g. a bio.tools entry) to be assigned in a 5 tier rating of *completeness*. Some attributes are grouped (see tables on right) for purposes of determining adherence to the

¹⁸ <https://github.com/bio-tools/biotoolsschema>

¹⁹ <https://github.com/edamontology/edamontology>

²⁰ <http://github.com/bio-tools/tool-Information-Standard>

²¹ http://biotools.readthedocs.io/en/latest/curators_guide.html

²² <https://github.com/bio-tools/biotoolsSchema/issues/77>

standard. In Section 3.3.1 we summarise a few ways of how the standard might be used for labelling various facets of tool quality.

	SPARSE	BASIC DETAILS	DETAILED	HIGHLY DETAILED	COMPREHENSIVE	
Name	✓	✓	✓	✓	✓	
Description	✓	✓	✓	✓	✓	
Homepage	✓	✓	✓	✓	✓	
Unique ID	✓	✓	✓	✓	✓	
Tool type		✓	✓	✓	✓	
Scientific topic(s)		✓	✓	✓	✓	
Publication (1)		✓	✓	✓	✓	
Support → (2)		✓	✓	✓	✓	
Scientific operation(s)			✓	✓	✓	
Documentation (3) →			✓	✓	✓	
Operating system			✓	✓	✓	
Language			✓	✓	✓	
License (4)			✓	✓	✓	
Type of input & output data (5)				✓	✓	
Accessibility →				✓	✓	
Code availability (3) →				✓	✓	
Downloads (3) →				✓	✓	
Supported data format(s) (6)					✓	
Scientific benchmark (3)					✓	
Technical monitoring (3)					✓	

Category	Sub-category	Requirement
→ Documentation	General	
	Manual	At least one
	API documentation	
	API specification	
→ Code availability	Repository	
	Source code	At least one
	Source package	
→ Accessibility	Terms of use	
	Accessibility	At least one
	Cost	
→ Support	Helpdesk	
	Issue tracker	At least one
	Mailing list	
	Contact person	
→ Downloads	Biological data	
	Binaries	
	Binary package	
	Container file	
	CWL file	
	Ontology	At least one
	VM image	
	Tool wrapper (galaxy)	
Tool wrapper (taverna)		
Tool wrapper (other)		

(1) A valid identifier (DOI, PMID or PMCID) or “Unpublished” must be specified.
 (2) An appropriate link or email annotated as the “primary contact” must be specified.
 (3) An appropriate link must be specified.
 (4) Either a valid license or “Unlicensed” must be specified
 (5) At least one input or output must be specified.
 (6) At least one data format for each input and output must be specified.

Figure 2. Attributes in the Tool Information Standard.

8.3.1 Application of the standard to software quality metrics

The Tool Information Standard provides a framework for computing aggregated metrics, purely from data in bio.tools, for purposes of labelling bio.tools entries in various ways. A few examples are given below (Figure 3 and Table 2) for illustration purposes. Exactly if and how to use such labels is currently under discussions, which must be broadened to include more ELIXIR stakeholders broadly. These discussions likely will inform a revision of standard to better reflect the software quality aspects and metrics (Table 1) once these are more settled.



Figure 3. Labels of software metadata (bio.tools entry) quality based on the Tool Information Standard.

Metric	Label	Status
4OSS Recommendations	4OSS logo ²³ attached if all 4OSS if verified as adhering to all 4OSS recommendations	Supported by biotoolsSchema: requires the tool to be registered in bio.tools with annotations for license, link to software repo, and documentation. Extension of biotoolsSchema CV for documentation type “Communications” is needed.
Open Source	One of four labels attached if license is annotated: “Open-source”, “Free software”, “Free and open source”, “Copyleft” ²⁴ .	Requires license annotation to be cross-referenced to the license lists maintained by the OSI ²⁵ and FSF ²⁶ initiatives. The “Proprietary”, “Other” and “Unlicensed” are alternatives supported in biotoolsSchema, but would not be rewarded with a label (to encourage good practice).
Software metadata quality	Conceivably up to four labels: <ul style="list-style-type: none"> bio.tools entry completeness (“SPARSE” through to “COMPREHENSIVE”) “Manually inspected” (if an entry was manually inspected by entry author, official bio.tools curator, or someone else) 	Requires implementation of logic to compute the metrics and significant UI development for the verification process. Bio.tools entry completeness could alternatively be handled by a

²³ A logo is needed.

²⁴ See EXCELERATE D1.1 report for definitions.

²⁵ OSI Open Source Definition (<https://opensource.org/docs/osd>).

²⁶ FSF Free Software Definition (<http://www.gnu.org/philosophy/free-sw.html>).

	<ul style="list-style-type: none"> • “Verified” <i>i.e.</i> confirmed (from manual inspection) as adhering to the Curation Guidelines • overall entry quality (“NEEDS TO IMPROVE” through to “EXCELLENT”) - conflating the notions above 	LinkedIn-style (%) approach.
Accessibility	Label design must reflect degree of accessibility, or perhaps only rendered if all criteria are met.	Requires license, cost (“Free of charge”) and accessibility (“Open access”) annotations, and links to repository and terms of use. License would have to be a recognised open source one.
Usability	Label design must reflect degree of usability, or perhaps only rendered if all criteria are met.	Requires description annotation, links to documentation (“Manual”, “API documentation”, “Training material”), and links to where the tool is deployed for use. Description would have to be verified as adhering to Curation Guidelines. Some extension of CVs needed (“Command-line docs” for documentation and “Web server”, “Galaxy” etc. for links), plus extensions for tool relationships.
Interoperability	Label design must reflect degree of interoperability, or perhaps only rendered if all criteria are met.	Requires the tool to be registered in bio.tools and complete / verified EDAM annotations (operations, I/O data types and formats).

Table 2. Candidate metrics and labels computable from bio.tools data.

8.4 Implementation & plans

8.4.1. bio.tools curation

The Tool Information Standard provides a basis to monitor software descriptions: its application to drive content quality improvement of bio.tools will be described in EXCELERATE D1.3 report. Here (Table 3) we show only a breakdown of the annotations by type, as evidence of major content growth and improvement, especially for documentation, contact information, operating system and code availability groups (all attributes that underpin important metrics of quality). Whilst manual curation is essential, especially for verification of information, given the available resource we cannot hope to manually curate all the different aspects of software quality. To improve sustainability we are exploring the automation of the curation, where possible (subject of D1.3 report).

Annotation type	Number of entries	Annotation type	Number of entries
Name	5952	Name	10756
Description	5952	Description	10756
Homepage	5952	Homepage	10756
Tool Type	7297	Tool Type	12595
Unique ID	5952	Unique ID	10756
Topic	14397	Topic	29570
Publication	7477	Publication	12472
Contact	4246	Contact	12045
Operation	14466	Operation	20731
Documentation	2933	Documentation	10712
Operating System	2056	Operating System	25781
Input/Output	6724	Input/Output	7132
Code availability	5014	Code availability	12504
Accessibility	393	Accessibility	500
Data format	5112	Data format	6624
Community	1436	Community	5142
Downloads	1903	Downloads	2372
Total	97262	Total	201204

Table 3. Annotations by type. The number of aggregated metrics of types defined in the Tool Information Standard is shown: from EXCELERATE D1.2 (left) and now (right). Note that for purposes of comparison, the data are formatted to the (older) version of the standard available when D1.2 was published.

8.4.2. bio.tools features

8.4.2.1 Ranking and filtering of bio.tools search results

We extended the means by which bio.tools entries can be sorted by an end-user, which now includes:

- **Citation count** (of primary publication)
- **Publication date** (of primary publication)
- **Updated** (when the bio.tools entry was updated)
- **Added** (when the bio.tools entry was created)
- **Name** (name of tool)

Many improvements are possible, for example, sorting by AltMetric attention score (planned in EXCELERATE D1.2 report), or allowing a user to select either the accumulated number of citations, or an adjusted metric to identify “rising stars” - tools which (inferred from citations) appear to be rising rapidly in prominence / usage. More broadly, new metrics (Tables 1-2) will be added as these become settled, and should provide the end-user with a flexible and useful ways to sort and filter search results by various criteria. This will involve iterations of 1. End-user requirements gathering and prioritisation. 2. Implementation. 3. Evaluation and optimisation of usability / end-user satisfaction.

8.4.2.2 Metrics cards

A “Metrics Card” (available as a tab alongside existing bio.tools Tool Cards) could display detailed information for all metrics available for a tool. That would include an assessment of the completeness and quality of the software description as per the Tool Information Standard. Importantly, metrics behind any label must be transparent to the user, such that a tool developer or service provider can easily see what they need to do, in order to improve their tool or bio.tools entry, and achieve the desired quality stamp.

8.4.3 Integration with OpenEBench

OpenEBench²⁷ is a system for technical monitoring and benchmarking of tools developed as part of the EXCELERATE WP2. OpenEBench calculates and aggregates metrics, of various types and from various sources, opening the possibility to share these data with other projects, such as bio.tools.

8.4.3.1 OpenEBench widgets

Our integration approach uses widgets, as they are easy to use, and highly reusable in sites other than bio.tools. A prototype widget for rendering metrics data from OpenEBench was developed by WP2 in collaboration with WP1. The widget is as a blob of HTML, which can be embedded in a site, and when configured with a bio.tools toolID, will render metrics data. The widget is embedded into bio.tools Tool Cards (Figure 4), for tools where monitoring data is available in OpenEBench. It currently includes four quadrants and a central motif; hovering over one of these brings up information (currently Documentation, Support, License, Buildability and Usage).

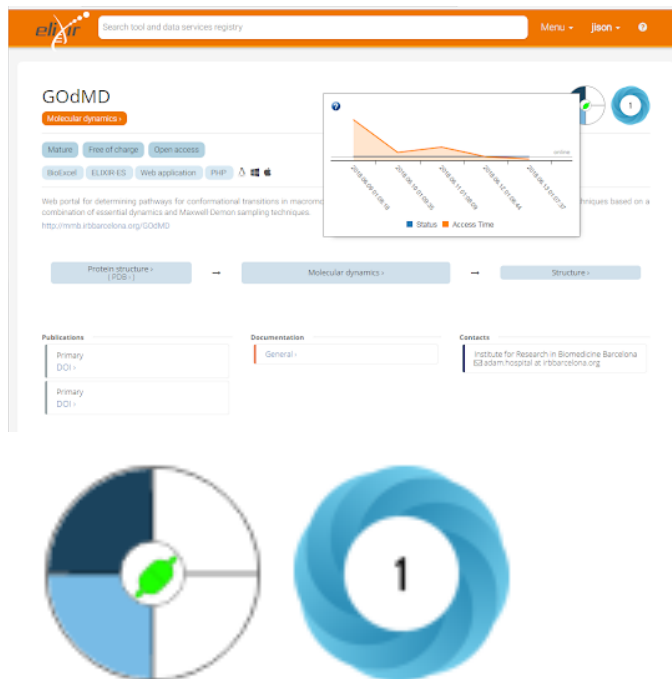


Figure 4. A bio.tools Tool Card (left) embedded with a widget rendering technical monitoring data available from OpenEBench (the widget is to the left of the Altmetric badge). When a region of the widget is clicked, some technical monitoring data is shown (tool access stats in this example).

The widget is included in bio.tools as a proof of principle. From the bio.tools perspective, the priorities to develop the integration are:

1. Reconfigure the existing widget such that it does not render information (such as “Documentation” and “Support”) that is already immediately obvious from the Tool Card, *i.e.* include only technical monitoring information that is unique to OpenEBench
2. Provide a separate widget for rendering scientific benchmarking information, something simple in the first instance, for example a “traffic-light” approach

²⁷ <https://openebench.bsc.es/>

- a. **Red**: tool has neither a benchmark publication or public benchmark data
 - b. **Orange**: tool has a benchmark publication and benchmark data (link to results of community benchmarking competition)
 - c. **Green**: tool has a benchmark publication, benchmark data, and is integrated in a continuous benchmark framework (link to results)
3. Work in collaboration with EXCELERATE WP2 to provide a gallery of widgets for rendering various other facets of software quality, including metrics calculable purely from bio.tools (see examples in Table 2), and those unique to OpenEBench.
 4. Where information is available in OpenEBench for a specific version of a tool, enable the widgets to be configurable, allowing the end-user to specify which version to render information for.
 5. Ensure that where software information are aggregated by OpenEBench, for example through automated means, that these data are made available to bio.tools, ideally by direct update of the corresponding bio.tools entries. In this way the scope and focus of bio.tools and OpenEBench is preserved, to the benefit of both projects.

8.4.3.2 OpenEBench API

The priority is to develop the OpenEBench API²⁸ to serve, for a given bio.tools toolID, atomic monitoring and benchmarking data, as available for versions of that tool. This can include things such as average response time, uptime, links to latest benchmark results, *etc.* These data can then be used by bio.tools for purposes way beyond what can be achieved by a simple widget, *e.g.* for filtering and sorting search results, comparison of tools within a certain class (EDAM Operation), and so on.

Purely for illustration purposes, an endpoint such as:
<https://openebench.bsc.es/tool/toolid={biotoolsID}>

e.g.
<https://openebench.bsc.es/tool/toolid=signalp>

could returning all available benchmarking and technical monitoring results for this tool, *e.g.*:

```
[
  toolID: "biotools:signalp"
  {
    version: "4.1",
    technical:
    {
      avg_response_time: "123",
      up: "Yes",
      ...
    }
    benchmark:
    {
      competition: "https://results.com/result1"
      continuous:
      "https://openebench.bsc.es/benchmark/toolid=signalp&ver=4.1"
    }
  },
  {
    version: "4.0",
    technical:
```

²⁸ <https://github.com/inab/elixibilias>

```

    {
      avg_response_time: "345",
      up: "No",
      ...
    }
    benchmark:
    {
      competition: "https://results.com/result1"
      continuous:
      "https://openebench.bsc.es/benchmark/toolid=signalp&ver=4.0"
    }
  },
  ...
]

```

8.4.4 Challenges

We have summarised a light-touch approach to rendering in bio.tools information about technical monitoring and scientific benchmarking; through links to sites (*e.g.* OpenEBench) where test data, scripts, benchmarks, benchmark results, relevant publications *etc.* can be researched in more detail, or by pulling from OpenEBench simple metrics where these are available. A more quantitative approach, yielding key performance indices (KPIs) for tools in a concise way, and thus allowing to render benchmarking data directly in bio.tools, is a much harder problem, in scope of EXCELERATE WP2, and must involve the scientific benchmarking communities. Ranks in benchmarks for tools would have to be attached with the specific operation performed and measure - at least until we have a better understanding of how to boil such benchmarking data down into a single KPI.

To give one example “5th in CASP²⁹” might include operations such as “Modeling”, “Refinement” and “Structure prediction”, which in turn are associated with certain biologically relevant scores, *e.g.* “Ligand accuracy”, “Oligomeric score”, “Binding site accuracy”, and so on. Doing this in a generic way, across multiple disciplines, is perhaps infeasible with the resource available to EXCELERATE, but it does point to the interesting possibility of annotating the tools, benchmarks, datasets *etc.* using the EDAM ontology to make the whole problem more tractable, especially in context of a more automated and extensible benchmarking environment. For example, consistent annotations of EDAM Operations and Formats would allow the relations between relevant entities and rendering of the data to be more conveniently handled. Discussions around this are ongoing, at a lower priority until more of the basics are in place.

8.5 Summary of future work

Table 4 (below) summarises ongoing and possible future work mentioned in the previous sections.

Action	Section
Settle and prioritise definitive list of software quality aspects and indicators/metrics (Table 1), in collaboration with ELIXIR partners broadly.	8.2

²⁹ CASP critical assessment of protein structure prediction, <http://predictioncenter.org/>

Extension of biotoolsSchema (see Table 1) to better support metrics of various types.	8.2
Revision of Tool Information Standard to reflect settled list of metrics/indicators.	8.3.1
Settle which candidate labels (Table 2) to implement in first instance.	8.3.1
Improved options for filtering and sorting of bio.tools search results.	8.4.2.1
Mock-up of bio.tools Metrics Card.	8.4.2.2
Develop the bio.tools::OpenEBench integration (widgets).	8.4.3.1
Develop and apply OpenEBench API	8.4.3.2
Discussions around quantitative metrics.	8.4.4

Table 4.