**Project:** National eResearch Collaboration Tools and Resources

**Project #:** 2179

**Contents:** Part A Brief

Part B Program Documentation

B1 Virtual Laboratories

B2 eResearch Tools

B3 Research Cloud

B4 National Servers Program

Part C Proposed Sub-Contractor Agreement

Part D **Proposal Submission *(this document)***

*Attachment 1 Research Cloud Specifications*

*Attachment 2 National Servers Program Specifications*

**Issue date**: 20th September 2011

**Responses must be received by NeCTAR by:** 4:00 pm AEST Wednesday 02nd November 2011

Copyright of this document remains the property of The University of Melbourne

|  |  |
| --- | --- |
|  | ABN 84 002 705 224  The University of Melbourne  Parkville, Victoria 3010 |
|  | NeCTAR is supported by the Australian Government (the Commonwealth) through the Super Science Initiative and the Education Investment Fund (EIF). |

Table of Contents

[Section 1 RFP 1](#_Toc303782264)

[1.1 RFP Contact Details 1](#_Toc303782265)

[1.2 RFP Timeline 1](#_Toc303782266)

[1.3 RFP Checklist 1](#_Toc303782267)

[1.4 Submission Instructions 1](#_Toc303782268)

[1.5 Late Submission 2](#_Toc303782269)

[Section 2 Contact Details of the Proposer 3](#_Toc303782270)

[2.1 Proposer Contacts 3](#_Toc303782271)

[Section 3 Compliance Statement 4](#_Toc303782272)

[3.1 Proposed Sub-Contract Compliance 4](#_Toc303782273)

[3.2 RFP Compliance 4](#_Toc303782274)

[3.3 Conflict of Interest 5](#_Toc303782275)

[3.4 Statement of Departures 5](#_Toc303782276)

[3.5 Conflict of Interest 5](#_Toc303782277)

[Section 4 Fields of Research 6](#_Toc303782278)

[Section 5 Response Template 6](#_Toc303782279)

[Section 6 Selection Criteria 18](#_Toc303782280)

[6.1 Criteria for VL and eRT 18](#_Toc303782281)

[6.2 Criteria for RC and NSP 20](#_Toc303782282)

[Section 7 Milestone and Funding Milestone Template 21](#_Toc303782283)

[7.1 Funding Estimate 21](#_Toc303782284)

[7.2 Milestone Template 22](#_Toc303782285)

# RFP

## RFP Contact Details

|  |  |
| --- | --- |
| **RFP Proposals ONLY** | [proposals-rfp-nectar@unimelb.edu.au](mailto:proposals-rfp-nectar@unimelb.edu.au) |
| **RFP Questions ONLY** | [questions-rfp-nectar@unimelb.edu.au](mailto:questions-rfp-nectar@unimelb.edu.au) |
| **General Queries**  **Questions relating to the RFP**  **should ONLY be delivered via the**  **appropriate email addresses above.** | The NeCTAR Directorate  Room 3.11, Level 3  Doug McDonell Building  The University of Melbourne, Vic 3010  Contact: (03) 8344 1277 |

## RFP Timeline

The full timeline is published and maintained on the NeCTAR website at (http://www.nectar.org.au)

|  |  |
| --- | --- |
| Request For Proposal issued | 20th September 2011 |
| Close for queries regarding proposal preparation | 5 business days before the Closing Time |
| Responses to be received by (Closing Time) | 04:00pm AEST 02nd November 2011 |

## RFP Checklist

|  |  |
| --- | --- |
| 1. Have you registered online at http://www.nectar.org.au? |  |
| 1. Have you read and understood Part A? |  |
| 1. Have you read and understood the relevant project Part B documentation? |  |
| 1. Have you read and understood Part C? |  |
| 1. Have you completed all sections of Part D? |  |
| * Section 2 Contact Information |  |
| * Section 3 Compliance Statement and Departures |  |
| * Section 4 Fields of Research (as appropriate) |  |
| * Section 5 Response, noting the selection criteria in Section 6 |  |
| * Section 7 Milestones and Deliverables |  |
| 1. Have you asked any questions you needed to, and received sufficient answers? |  |
| 1. Have you returned the pack, Part D, to [proposals-rfp-nectar@unimelb.edu.au](mailto:proposals-rfp-nectar@unimelb.edu.au)? |  |

## Submission Instructions

Proposals shall be submitted:

* electronically (as per Section 1.1);
* in English;
* in a legible font and size (suggested minimum 10pt);
* in text-searchable PDF; and
* a 10 Megabyte (including attachments) limit.

On Closing Time, the University of Melbourne Tender Board will issue all submissions to NeCTAR for review.

Proposals must be received no later than Closing Time specified in the Timeline or they will be treated as a Late Submission as described below.

NeCTAR reserves the right to change the Closing Time for any reason, in which event written notice of the change will be provided.

## Late Submission

Proposals lodged after the Closing Time or lodged at a location or in a manner that is contrary to that specified in this RFP will be disqualified from the selection process and will be ineligible for consideration, except where the Proposer can clearly demonstrate, to the reasonable satisfaction of NeCTAR, that late lodgement of the Proposal:

1. resulted from the mishandling of the Proposal by NeCTAR; or
2. was hindered by a major incident and the integrity of the selection process will not be compromised by accepting a Proposal after the Closing Time.

The determination of NeCTAR as to the actual time that a Proposal is lodged is final. Subject to paragraphs (a) and (b) above, all Proposals lodged after the Closing Time will be recorded by NeCTAR, and will only be opened for the purposes of identifying a business name and address of the Proposer. NeCTAR will inform a Proposer whose Proposal was lodged after the Closing Time of its ineligibility for consideration. All such Proposals will be returned at the conclusion of the Selection Process.

# Contact Details of the Proposer

## Proposer Contacts

The Contact Details of the Proposer are to be detailed in section 2.1.1 below.

Please add the details of any anticipated participating organisations in section 2.1.2. Add extra lines as required.

### Proposer

|  |  |
| --- | --- |
| Organisation Name | Centre for Genetic Epidemiology & Biostatistics, University of Western Australia |
| Contact Name | Paul White |
| Position | Manager Informatics Systems Development |
| Business Address | 35 Stirling Highway, Nedlands, WA, 6009 |
| Postal Address | 35 Stirling Highway, Nedlands, WA, 6009 |
| Telephone | 08 6488 6733 |
| Facsimile | 08 6488 6750 |
| Mobile Phone | 0412 352 883 |
| E-mail | paul.white@uwa.edu.au |

### Participating Organisations

|  |  |  |
| --- | --- | --- |
| Organisation / Group Name | Location | Role |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

# Compliance Statement

## Proposed Sub-Contract Compliance

**Are there any Departures from the Contract (Part C) Terms and Conditions?**

**No** There are no departures from the terms and conditions (i.e. Full Compliance)

**Yes** There are departures from the terms and conditions

Detail the departures in Section 3.4 of this document.

The proposing organisation warrants that except for the departures listed in Section 3.4, the response is in full compliance with the Contract terms and conditions and no further contractual issues will be entered in to.

Signature of authorised person making the statement Name and role (printed) Date

## RFP Compliance

**Are there any Departures from the RFP Terms and Conditions (Part A)?**

**No** There are no departures from the terms and conditions (i.e. Full Compliance)

**Yes** There are departures from the terms and conditions (i.e. Does not Fully Comply)

Detail the departures in Section 3.4 of this document.

The proposing organisation warrants that except for the departures listed in Section 3.4, the response is in full compliance with the RFP terms and conditions.

Signature of authorised person making the statement Name and role (printed) Date

## Conflict of Interest

**Are there any known or potential conflicts of interest responding to the RFP and its Terms and Conditions or in delivering the proposed works?**

**No** There are no conflicts of interest

**Yes** Describe the conflicts in Section 3.5 of this document.

Do you commit to inform the University of Melbourne of any future conflicts or potential conflicts as they arise?

**Yes**

Signature of authorised person making the statement Name and role (printed) Date

## Statement of Departures

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Clause or Reference** | **Nature of Compliance** | **Proposed wording of amendment** |
| **Proposed Sub-Contract** |  |  |  |
| **RFP Terms and Conditions** |  |  |  |

## Conflict of Interest

|  |  |  |
| --- | --- | --- |
| **The Nature of the Conflict** | **Implications of the Conflict** | **How the Conflict is to be managed?** |
|  |  |  |

# Fields of Research

For RC and NSP proposals, this section is optional. RT and VL proposals must complete this section. Select up to five disciplines using either the two or four digit codes, or a mixture of both, and allocate a percentage score or weight against how closely the Proposal is aligned to a particular community discipline or Field Of Research. The FOR codes are available at: <http://www.arc.gov.au/applicants/codes.htm>

|  |  |
| --- | --- |
| DISCIPLINE/FOR Code | Weight (percent) |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  | 0% |

# Response Template

Complete the following table, ensuring a response to all headings and statements are provided. Attach this as a separate document to your proposal. Not all elements apply to every program (RT, VL, RC, NSP), please indicate “not applicable” as appropriate.

Any additional material or brochures can be added as attachments to the Proposal. The page counts are an indicative guideline for responses.

**SUMMARY**

## Program and Proposal Title (0.1 pages)

* *Which program is this proposal addressing? (RT, VL, RC, NSP).*
* *Provide a short title for the proposal to use as a reference in communications.*

Cloud-based Bioinformatics Tools

This proposal is addressing the eResearch Tools program.

## Executive Summary (1 page)

*Summarise the context that leads to this project and briefly outline the vision for the outcomes.*

Data is the primary asset of biomedical researchers, and the engine for both discovery and translational medical research. As the volume and sensitivity of research data increases, due to technologies such as next generation sequencing of human specimens linked to clinical datasets, so too does the requirement for access to application software for integrating and interrogating the different types of research data. Researchers often need to execute complicated queries and conduct analysis across multiple data types, such as phenotypic, genotypic, pedigree and biospecimen data.

The Ark project was established within the Centre for Genetic Epidemiology & Biostatistics at the University of Western Australia in late 2009 to develop an open-source platform to support Australian and International biomedical researchers. The objective of The Ark project is to provide a suite of secure, integrated web-based applications that incorporate the majority of the functionality required to conduct a complex study or clinical trial.

The project has now delivered a Java-based solution that delivers the following capabilities:

1. Create and a configure a study;
2. Define users and manage their roles and access permissions
3. Define and manage a research participant pool;
4. Define and manage the collection of phenotypic research data;
5. Dynamically generate Electronic Data Capture (EDC) forms for collecting most types of textual research data;
6. Manage physical biospecimens and the associated data (LIMS); and
7. Generate reports.

This functionality has now been released as Version 1.0 and is being used in production. The project team is currently focussed on supporting Version 1.0 and completing the first version of the following functionality:

1. A flexible data extraction tool to extract de-identified data for analysis using software applications like SAS, SPSS or Stata;
2. Additional reporting functionality to support cross-module reports.

The software has been designed and developed in accordance with the following guiding principles:

1. The software should be web-based;
2. One instance of the software should securely support multiple research studies and multiple users;
3. The core application software should not be developed to suit any one study/project but rather should be configurable to suit a broad range of studies;
4. The software should be as user-configurable as feasible through a browser-based graphical user interface;
5. The software must be highly secure; and
6. The software modules should be well integrated.

We are not aware of any equivalent open source software applications available nationally or internationally and we are unaware of any viable alternatives for researchers that do not have the required funding or expertise to purchase commercial solutions or develop their own solutions. Smaller studies are often able to operate using simple tools, such as Microsoft Access or Microsoft Excel, but there is no doubt that the research being undertaken by these studies is being compromised by the lack of suitable informatics support.

## Research Community Profile (1 page)

*Provide a profile of the research community that is sponsoring the proposal; include the aims of the community, geographic spread or location and membership size.*

Users are both clinical and basic science researchers who use clinical specimens and health information from any disease type for molecular and cellular analyses. Current users of the Ark cover cancer, diabetes, cardiovascular, obesity, eye diseases, mental health and infant and child health. The platforms would also support use by those collecting specimens from clinical trial, both for academic and commercial purposes.

## Development Organisation Profile (2 pages)

*Provide details of the organisations that will contribute to the development of the proposed infrastructure; include information about their capacity and capabilities, their track record and relevant experience in this role, their approach to quality standards, support and warranty mechanisms, etc. Include any supporting statements from Research Users. Where specific projects are mentioned, indicate specifically the aspects which support your track record or experience.*

The software development will be undertaken primarily by the following organisations:

1. The Centre for Genetic Epidemiology & Biostatistics – genomic capabilities and analysis
2. The Ark ([www.the-ark.org.au](http://www.the-ark.org.au)) project team at the Centre for Genetic Epidemiology & Biostistics, UWA, consists of three full-time professional software developers and one part time project manager. The Ark team has been developing and supporting informatics software to support the Centre for Genetic Epidemiology and Biostatistics and collaborating institutions since 2004. Software tools developed by the team has been in production since 2005 and currently provides support to a number of research groups, including:
   * The Western Australian Research Tissue Network at St John of God Hospital
   * The Busselton Healthy Aging Study (The Busselton Foundation)
   * The Western Australian DNA Bank (which in turns supports over 40 studies)
   * The National Breast Cancer Foundation
   * The Western Australian Institute for Medical Research
   * The Lions Eye Institute

The earlier software tools were developed using a mix of C, Oracle 4GL and Java development languages. Two years ago The Ark project was initiated to develop the Java-based research tools that form the basis of this bid.

All software development has always and continues to be the result of close collaboration with a number of key research groups. This helps ensure that the software is always fit for purpose and is always tested using real data by future users of the applications before production deployment. Continual feedback is encouraged and managed by providing most researchers with online access to Jira ([www.atlassian.com/en\_AU/software/jira/](http://www.atlassian.com/en_AU/software/jira/)), the web-based issue tracking system used by The Ark team. Researchers are able to submit issues and enhancements online and track their progress online. All issues are reviewed weekly and prioritised. Authorised researchers have access to view the current status of any of the issues associated with The Ark Jira instance.

All researchers are also able to access The Ark’s wiki ([wiki.genepi.org.au](http://www.wiki.genepi.org.au)) to view ongoing analysis, design and user documentation.

Primary responsibility for supporting a The Ark module is assigned to a specific developer but all developers have sufficient familiarity with all of the software to provide backup support. Care has also been taken to ensure that a shared set of design patterns is used by all of the developers.

1. The Ark project team at the Centre for MEGA Epidemiology, Melbourne University, consists of one developer with responsibility for contributing to the application development as well for providing technical and implementation support to Eastern States studies.
2. OBiBa ([www.obiba.org](http://www.obiba.org)) is an international software development project based at at McGill University, Montreal, Canada. OBiBa is committed to building a full suite of open source software for biobanks. It is comprised of several independent and self-funded teams around the world, each of which is producing stand-alone applications that support particular biobank activities. The applications can be customized and integrated to create a complete biobank information management system.

OBiBa offers a collaborative infrastructure to its teams and to other developers who may wish to join the OBiBa community. The infrastructure for developers includes integrated tools for documentation, issue tracking, deployment, and project management. The Ark project is a member of OBiBa. Paul White is one of the Principal Investigators of the OBiBa project.

1. The Australian Twin Registry at Melbourne University has been developing software to support the operation of The Australian Twin Registry for several years. The ATR’s development and operational staff have and will continue to contribute to the specification and testing of The Ark software.
2. Texas University – genomic capabilities

Include letters of support

## Operational Organisation Profile (2 pages)

*Provide details of the organisations that will contribute to the development of the proposed infrastructure once they have been commissioned; include information about their capacity and capabilities, their track record and relevant experience in this role, their approach to quality standards, support and warranty mechanisms, etc. Include any supporting statements from Research Users.*

*Where specific projects are mentioned, indicate specifically the aspects which support your track record or experience.*

The primary support organisations will be:

* The Centre for Genetic Epidemiology & Biostatistics at the University of Western Australia
* The Centre for MEGA Epidemiology at the University of Melbourne
* OBiBa

## Other Participants (0.5 pages)

*Name any other institutions or groups that will need to be involved in the project planning and execution and their roles.*

St John of God Health Care (Private partner)

Clinical Oncological Society of Australia- cooperative cancer trials groups

WAIMR

DoHWA – Di Rosman

Curtin – James Semmens

Busselton Health Study

Cancer Council of New South Wales

Cancer Institute of New South Wales

National Breast Cancer Foundation

Foundation Health Consumers Council of WA- consumer engagement: In WA there is a strong history of consumer engagement with clinical datasets through the WA Health Data Linkage Unit. These pathways would be used to ensure that the governance of the proposed work is consistent with consumer expectations.

*Although not necessarily a ‘need’ it will be ideal (assuming the groups are successfully funded) if other groups submitting proposals that closely align with this proposal are involved with our project planning.*

*UQ- submission-other groups.to optimise utility*

## Key Personnel (0.5 pages)

*State any key individuals that are required for specific project activities and their availability.*

*Provide names, organisational locations, and their expected roles.*

*For example; Project Managers, designers, technical experts.*

**Name:** Paul White

**Organisation:** University of Western Australia

**Role:** Project Manager

**Availability:** 60%

**Name:** Christopher Ellis

**Organisation:** University of Western Australia

**Role:** Technical Architect and Senior Developer

**Availability:** 100%

**Name:** Eric Lam

**Organisation:** University of Western Australia

**Role:** Technical Architect and Developer

**Availability:** 100%

**Name:** Adrian Bickerstaffe

**Organisation:** Melbourne University

**Role:** Software Developer

**Availability:** 50%

**Name:** Professor Eric Moses

**Organisation:** Centre for Genetic Epidemiology and Biostatistics, University of Western Australia

**Role:** XX

**Availability:** 20%

**Name:** Professor John Hopper

**Organisation:** Australian Twin Registry, Melbourne University

**Role:** XX

**Availability:** 10%

**Name:** Dr Nik Zeps

**Organisation:** Western Australian Research Tissue Network, St John of God Hospital

**Role:** XXEthics, pathology based -biobanking, clinical trial biobanking

**Availability:** 20%

**Name:** Philippe Laflamme

**Organisation:** OBiBa Project

**Role:** Technical Architect

**Availability:** 5%

## Infrastructure (2 pages)

*Describe the proposed infrastructure to be developed. Provide supporting documentation, specifications etc as required.*

In addition to ongoing enhancements to the existing software the following capabilities have been identified as necessary to support the medical research community:

* A contact management and pedigree management module;
* A genotypic data management module. Note that we have developed a prototype in conjunction with researchers and developers from the OBiBa project at McGill University in Montreal; COMMENT ON GIN SUBMISSION AND TEXAS
* Two and three dimensional pedigree visualisation and exploration software. Note that we have developed a prototype for a 3 dimensional pedigree visualisation tool that may be used as the basis for this module;
* Integrated invoicing and billing capabilities;
* Registry Management functionality for managing participant registries, such as The Australian Twin Registry; and
* HPC integration.

Planned enhancements to the existing and new software will include:

* Increased scope of web services to support loosely coupled integration with third party applications;
* Increased support for localisation; and
* Online training and video guides.

May need to write some functional specifications to submit as attachments.

**RESEARCH COMMUNITY NEEDS & BENEFITS**

## Target Research Community (0.5 pages)

*Identify the research communities and the expected number of users.*

All clinical and basic science research that uses human biospecimens linked to health information. In WA is at least 500 researchers. Nikalso include the international use that will be made. It is difficult to estimate just how many people this will represent but based upon those submitting there is a network of other 1000 researchers that will benefit from the the initial proposal.

## Needs and Impact (2 pages)

*Describe the needs of the research community. Describe the impact on current research practices and any opportunities the new infrastructure will provide to those communities.*

*List the benefits to be derived from the delivered infrastructure, describe them in quantifiable terms where possible.*

*Outline how the impacts and benefits will be tracked, managed and measured.*

Molecular research using human biospecimens needs to be associated with patient information. There are very few IT solutions to manage such data and it is a significant rate-limiting step to research outcomes. Database maintenance includes writing searches/reports and creating appropriate links between different datasets. There is often no means to fund this core infrastructure and typically solutions include writing ‘databases’ in lower end packages like access, file make pro or even excel. A means to interrogate multiple pieces of data across different databases holding information about a single person will significantly enhance research capabilities and allow more complex statistical analyses to detect associations in complex diseases.

## Broader adoption (0.5 pages)

*State which additional communities, resource providers or organisations would also be expected to benefit from the use of the new infrastructure and services should the project succeed.*

The Ark software is designed to be extremely flexible and user-configurable. As a result the software will support a range of different communities where multiple data types need to be associated with a single physical entity. For example:

1. We are currently working with a commercial partner, Patrick Rose, to configure the system to support the collection of health and incident data in remote communities in regional Western Australia;
2. The UWA Crime Research Centre has approached us to discuss using The Ark to manage crime research data;
3. eHealth-the software and linkages are written in such a way that they can readily link to electronic health record systems as they come online.

## Value adding (1 page)

*Identify the components of the project that are adding value to existing research infrastructure investment. For example; building new services on top of and using existing research infrastructure.*

*Describe the alignment with national research infrastructure and eResearch infrastructure priorities. See* [*http://www.innovation.gov.au/Science/ResearchInfrastructure/Pages/default.aspx*](http://www.innovation.gov.au/Science/ResearchInfrastructure/Pages/default.aspx)*.*

*Note how this project will engage with, and leverage off, the other national infrastructure programs such as ANDS, AAF, NCI, and RDSI, as well as any other NCRIS Capability Platforms.*

*Identify components that could be used by other research communities and organisations to resolve their problem. For example; a part of the workflow that could be used by other domains and research communities.*

The Ark project builds upon a software development and research project initially funded by an NHMRC Enabling Grant. This predecessor project, commonly referred to as WAGER, has provided a solid foundation in terms of informing the functionality, design, development and support approaches that have been adopted by The Ark team. Through the course of the WAGER project the team also built strong relationships with a number of Australian research groups, some of which are collaborators and supporters of this grant application.

The Ark team have been working with the Western Australian Data Linkage Unit since 2005 to facilitate the linking of research data with public health data. The Ark team intends to continue this relationship with the Western Australian Data Linkage Unit and with the Australian-wide equivalent - The Population Health Research Network (PHRN).

The Ark project will integrate with existing capability platforms as follows:

* The existing The Ark LDAP authentication services will be integrated with AAF so that existing researcher credentials may be used to access The Ark software;
* Over the course of 2012 The Ark will retire its existing server hardware and migrate all of its existing applications to the Research Cloud. Note that The Ark already has access to two National Server Program virtual servers. The applications on these servers will be moved to the Research Cloud at the appropriate time;
* One of the primary objectives of The Ark project team is to position researchers to take advantage of the significant HPC capability currently being provided by NCI. To gain new insights from the increasingly large data sets being made available through technologies such as high throughput screening will require seamless access to high performance computing hardware and the relevant software tools;
* The Ark project team currently has its own SAN for data file storage. It is planned that during the course of the NeCTAR project this will be retired and the RDSI data storage will become the preferred storage medium for image and genotypic data files. This will not only provide access to more scalable , low cost hardware, it will facilitate the analysis of the large data sets as these data sets will no longer need to be moved over the WAN to HPC nodes for analysis; and
* There is the possibility that, if the issues associated with fully de-identifying data sets can be addressed, then some of the collaborator data sets could be moved to ANDS. This is not currently within scope for this proposal.
* Link to genomic data capabilities

There are a number of components of The Ark solution that could be leveraged by other research communities, including:

* The Laboratory Information Management System (LIMS) for biospecimen management
* The Phenotypic data management module for storing phenotypic data about an entity
* The genotypic data management module for managing any time of genotypic data

**PROJECT MANAGEMENT**

## Governance (1 page)

*State who is accountable for assessing project performance, what process will they apply.*

*Describe the authority structure over resources in the proposed project.*

*List all members of the Project Governance Body.*

*Describe the organisation’s Project Management methodology, scaled as appropriate for the proposed sub-project, and the maturity of its use within the organisation.*

*Describe the key processes and templates used internally by the organisation for governance and project management.*

For practical management the governance needs to be defined by one institution-

Steering committee-is the Independent Advisory board

Insert a diagram showing the governance structure – incl. Steering Committee

## Project Scale (0.5 pages)

*Identify the overall scale expected in the project, total effort, amount of funding required, amount of co-investment proposed and nominate any other participants that have indicated a willingness to participate through providing resources and what they are.*

## Project Approach (1 page)

*Detail how the required services will be developed and delivered. Outline the different stages of activity.*

1. Integration with existing research infrastructure – single signon etc
2. Comprehensive testing to ensure existing software is cloud ready
3. Establishment of customer support tools and processes
4. Deployment of existing functionality to the cloud in a production environment
5. Migration of existing users to the Research Cloud
6. Incremental development of functionality to support collaborative partner requirements:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Pedigree Storage & Visualisation | Genotypic Data Management | Enhanced Data Linkage & Reporting | Integrated Invoicing & Billing | Registry Management | HPC Integration |
| Clinical Oncological Society | X | X | X | X |  |  |
| Busselton Health Study | X | X | X |  |  | X |
| Australian Twin Registry | X |  | X | X | X |  |
| National Breast Cancer Foundation |  |  | X |  | X |  |
| Cancer Council of New South Wales |  | X | X | X |  |  |
| Western Australian Institute for Medical Research | X | X | X |  |  | X |
|  |  |  |  |  |  |  |

1. Staged deployment to the broader research community

MORE DETAIL

## Key Deliverables and Acceptance Criteria (1 page)

*Define the key project deliverables.*

*Deliverables include the services, infrastructure and functionality specified above for development by the sub-project as well as any required project management artefacts.*

*Show the Acceptance Criteria against each deliverable. These will be further elaborated during project delivery when Commissioning Tests are prepared.*

*Define the Acceptance Criteria specific to Commissioning.*

## Quality Control (1 page)

*Identify the personnel, processes and any special resources that will be required for Quality Control, Acceptance and Commissioning Testing activities on the proposed project.*

*State who is responsible for the completion of each deliverable.*

*State who is accountable, within the sub-contracted organisation, for the acceptance of each deliverable.*

*In contrast to “what we will test” described earlier, this is a “how we test” description.*

Members of The Ark project team will be responsible for developing the functional requirement and user interface design specifications. Representatives from the key project collaborators, St John of God Hospital, The Australian Twin Registry, and XXX will be responsible for signing off these specifications.

System specifications describing how new functionality will be developed will be developed by The Ark project team and signed off internally.

Most software code written by The Ark team undergoes peer review by other members of the development team. Programmers are responsible for their own unit testing.

Software developed within each project sprint will be submitted to uTest, a Massachusetts-based company that provides a virtual testing service with over 45,000 professional software testers worldwide.

**Functional Testing**

After reviewing each application uTest will develop a scope of work for the testers including a short description of the product, areas that are in scope and out, the kind of issues we are most interested in and any special instructions around product access, bug reporting or communication.  All of this testing will be driven by test.

**Usability Testing**

At the GUI design stage we will consult with a uTest UX Expert to generate a holistic review of the product for usability and to develop a usability survey. Then we have a number of testers from the uTest crowd complete this custom survey giving their opinion of the usability of the product. Finally, the UX Expert will submit a single report providing us with a complete product review with specific recommendations on how to better design and deploy the product for heightened usability.

**Security Testing**

We will utilise uTest’s manual penetration services to identify vulnerabilities in the software.

**User Acceptance Testing**

User Acceptance Testing will be undertaken by our partner organisations in order to test the software in real production environments using real data. Prior to each cycle of User Acceptance testing we will migrate production data into a secure testing instance for these partners to test in order to ensure that both the migration scripts and the application functionality, performance, security, etc. meet the required levels.

## Risk and Issue Management (1 page)

*Define the key risks to the successful delivery of the proposed project.*

*Define any open issues that need resolving before the proposed project can start delivery.*

*Detail any key questions that will affect the operation of the proposed project pending a decision.*

*Define how the major risks and issues to the proposed project*

**LEVERAGING**

## Standardisation and Interoperability (0.5 pages)

*Describe the global technology development or standardisation work that will be adopted, adapted or extended within the project and any risk reduction or additional value available by collaboration with similar activities occurring elsewhere in the world.*

*Identify any local or emerging standards that will be incorporated by the project.*

The Ark team is developing all of the software to be as flexible as possible so that it can be used to support studies that have chosen to adopt specific ontologies or have developed their own coding standards for data storage. The ability for researchers to define their own data dictionary and data entry forms as well as to define custom fields exists in all modules. All field definition is by way of a graphical user interface.

The Ark team will also work with the (P3G) Public Population Project in Genomics’ DataShaper group ([www.datashaper.org](http://www.datashaper.org)) to encourage the adoption of their Data Schema and Harmonization Platform for Epidemiological Research.

All software development will be done according to the relevant W3C software development standards.

CTRNet

Uppsala

**FINANCIAL**

## Budget Breakdown (1 page)

*Provide a proposal breakdown of the project budget by the milestones, which are described as an attachment to the Proposal in the format described below.*

*Include proposed staffing levels; where actual individuals have not been allocated to the sub-project, use a role name and Full Time Equivalents (FTE) to show the number and value of budgeted staff that will be working on the sub-project at that milestone. For example; Software Designer by 2 or two Software Designers working full time on the sub-project. Each individual or FTE role is to be included as a separate line item.*

*Break down the budget into EIF (NeCTAR) funding or co-investment funding.*

**SERVICES AND SUPPORT**

## Service Levels (0.5 pages)

*Specify the service level that will be offered for each service in relation to the levels discussed in the relevant Part B document.*

The Ark service desk personnel will primarily be located at the Centre for Genetic Epidemiology at the University of Western Australia with additional support being provided out of the Centre for MEGA Epidemiology at the University of Melbourne. This will ensure that support personnel will be available from 9am to 5pm Monday to Friday for all Australian locations.

All The Ark users will have access to an instance of the Jira issue tracking tool running in the Reearch Cloud to enable them to log and then follow the progress of issues, including bugs, enhancements and requests for assistance with tasks like project configuration and data migration.

The Ark team will ensure that the firewall software used provides …..

• operational scripts and tools, e.g. to support triage, monitoring, security, etc.;

The service levels provided by The Ark software will be very dependant on the service levels being offered by the NeCTAR Research Cloud as it is intended that all The Ark software will be hosted in the cloud. Independent of the Research Cloud availability, The Ark will endeavour to provide access to the application software at a 95% level between the hours of 9am EST time to 5pm WA time. All major software upgrades will be conducted after 3pm WA time to ensure maximum availability for Eastern States users.

All The Ark documentation is currently and will continue to be hosted on The Ark’s Confluence wiki. This is currently hosted on our own servers but will be migrated to the Research Cloud in the near future.

Training will be conducted using a number of mechanisms:

* Face-to-face training will be conducted by The Ark staff from either Melbourne or Perth. Where travel is required then these costs will be borne by the research institution receiving training.
* Online video presentations. The intent is to develop a number of online video training sessions that will be accessible from The Ark website.
* How-to documentation will be developed that provides a step-by-step approach to configuring and using The Ark tools.

## Operations and User Support (1 page)

*Detail the proposed operator of each service, what support will be provided to users and by whom.*

The Centre for Genetic Epidemiology and Biostatistics at the University of Western Australia will operate the software tools implemented by The Ark project. The support mechanisms provided to users will include:

* Online support by providing all users with access to the Jira issue management tracking tools;
* Telephone support as appropriate;
* Access to online documentation through The Ark’s wiki; and
* Access to answers to frequently asked questions on The Ark’s wiki.

Significant face-to-face support will also be provided to the project collaborator’s during the development and implementation stages of the project.

MORE DETAIL

## Sustainability (0.5 pages)

*Describe how the project infrastructure will be made sustainable following the completion of the project and becoming operational, when EIF funds cannot be used. Include information on the proposed business/financial model and the timeframes to which they apply.*

During the development stages of the project under the NeCTAR grant the project infrastructure will be funded by a mix of key collaborator in-kind and cash contributions and a fee-for service model. A fee-for service model is already in use with 12 research projects already paying for access to the pre-existing non open-source version of The Ark’s software. It is proposed that over the course of 2012 these projects will be migrated to the open-source version of The Ark software. The current fee for service model has the following features:

* Hosting and support charges are on a per-module basis. Researchers have access to and only pay for the modules that they require;
* Hosting and support charges are independent of the size of the research project;
* Configuration and training is charged on a per module basis;
* A certain amount of disk storage is available for each study. Additional storage attracts a per Terabyte charge;
* Fees are for access within a calendar year – charged pro rata;
* All client-specific work is charged on a per hourly basis;

Changes to the existing business model to support the open-source software include:

* Researchers may choose to host their own instance of the software and may elect to enter into a technical support contract as long as they are running a supported, non-modified version of the software;

Subsequent to the completion of the development phase of the NeCTAR project the operational costs will be funded through the fee-for-service model.

The intent is to have the production and test instances of the software hosted in the Research Cloud. Costs for this are as yet unknown but if this does incur a charge then this will be passed onto the organisations using the software on a pro-rata basis.

Currently if new The Ark functionality is requested by a specific client then the work is typically only undertaken if will contribute to the main branch of the software, even if the client is willing to pay for the development. During the course of the NeCTAR project the new functionality will be driven by the requirements of the key project collaborators.

## IP, Licensing and Access (0.5 pages)

*State if any software licenses will be used for software developed by the proposed project, or other software used for the services to be delivered by the proposed project.*

*State if software developed by the project will be made available under an open-source license.*

*State if other Intellectual Property (IP) or licensing restrictions are relevant to the services that will be delivered.*

*State if there are any restrictions on access to the services that will be delivered.*

All software development will be undertaken under the GPLv3 open source software license. The GNU General Public license is a copyleft license for general use, which means that derived works can only be distributed under the same license terms. Under this philosophy, the GPL grants the recipients of a computer program the rights of the free software definition and uses copyleft to ensure the freedoms are preserved, even when the work is changed or added to.

There will be no other Intellectual Property or licensing restrictions relevant to the services that will be delivered.

There will be no restrictions on access to the services that will be delivered.

Software licenses used for software development will include:

* Macintosh Operating System
* Windows Operating System
* Centos Linux Operating System (Local & NeCTAR Research Cloud)
* Navicat Data Modelling Tool
* MySQL Database (open source license)
* Atlassian Jira issue tracking and project management (open source license)
* Atlassian Confluence wiki (open source license)
* Subversion source code management
* Drupal web content management
* Limesurvey
* Hudson Continuous Integration
* Microsoft Project
* OpenLDAP

## Communications and Engagement (0.5 pages)

*Describe the means by which customer satisfaction with the proposed project’s planning, requirements gathering, scoping decisions, progress, quality and outputs will be measured.*

An Agile software development approach will be used for the project. Software will be developed in short iterations, typically 4 weeks. Each iteration will involve a team working through a full software development cycle including planning, requirements analysis, design, coding, unit testing, and acceptance testing. At the end of each iteration a working product will be demonstrated to stakeholders. This will minimise overall risk and allow the project to adapt to changes quickly. This approach will provide the project customers the opportunity to provide timely feedback that can rapidly be integrated into the software development process.

## Constraints and Dependencies (0.5 pages)

*Define and explicitly quantify any schedule, expenditure, resource, scalability, performance, and quality constraints or limitations on the project and its deliverables.*

*State the dependencies with external parties, including other NeCTAR projects, which have been identified in planning the proposed project.*

# Selection Criteria

The following table outlines the criteria that will be used to assess proposals, based on the responses to Section 5 above. They are provided here only for the information of respondents and to ensure that responses consider the key elements being sought.

## Criteria for Virtual Laboratory and eResearch Tools Proposals

|  |  |  |
| --- | --- | --- |
| **Category** | **Weight (%)** | **Criteria** |
| **Research Community** | **20%** |  |
| Research community to benefit is well-defined (by location, institutions, size), nationally significant and the proposal is well supported by the research community | | |
| The needs of the research community to be addressed by the proposal are well defined and significant | | |
| Researcher participation in the proposal is well defined, significant and national or international in scale, supporting outreach, uptake, testing and evaluation of the project infrastructure | | |
| The proposal is aligned to national research priorities | | |
| The proposal is aligned with, and contributes to implementation of, national research infrastructure and eResearch priorities | | |
| **Research Impact** | **30%** |  |
| The benefits to be delivered to the identified research community are well described, achievable, significant and measurable | | |
| The process for tracking and measuring the benefits is defined and achievable | | |
| **Virtual Laboratories:** | | |
| The proposal integrates significant infrastructure and research capabilities on a national scale to deliver a transformative impact for the research community identified in the proposal:   * Improving access to instruments, data, compute and other research infrastructure * Enables new research practices through research workflows * Address emerging research challenges * Support cross-institutional and cross-disciplinary research workflows through the provision of integrated collaborative ICT infrastructure * Connect significant infrastructure capabilities to support discipline and problem oriented research workflows: e.g. remote laboratory access, computation, research data repositories, workflow tools and sensor networks * Provide an exemplar to research communities of the benefits of integrating significant research support capabilities into a rich online collaborative environment. | | |
| **eResearch Tools:** | | |
| The proposal delivers software infrastructure which improves existing tools to:   * Enhance support for research collaboration * Improve remote access to underlying research facilities and infrastructure * Enhance support for research workflows, including cross-institutional and cross-disciplinary workflows, and/or * Enable the connection of research data sets and repositories with research tools and workflows. | | |
| Potential benefits from re-use across research disciplines are well described, realisable and significant | | |
| **Implementation** | **30%** |  |
| An appropriate Governance structure has been defined for the Project   * A Governance body appropriate to the project has been defined * Authority structure over resources in the project has been described * Key personnel (if required) and their roles have been clearly defined | | |
| The infrastructure to be created by the proposal is well-described, achievable and will deliver the research impacts described.   * The scope is realistic and outcomes are achievable * Project Management is well-described and appropriate to the proposal scale * Key risks have been identified and are manageable * Issues that require solving have been identified * Dependencies with third parties have been listed | | |
| Capability and track record in development and operation of eresearch infrastructure   * The contribution and track record of each organisation in the proposal to development and operation of the infrastructure is well-described and appropriate to the successful delivery of the proposal infrastructure | | |
| The proposal leverages or builds upon existing research infrastructure where appropriate:   * the NeCTAR Research Cloud and NSP for infrastructure hosting and computation * existing research and eresearch infrastructure (eg. RDSI, ANDS, NCI, Pawsey, Super Science, NCRIS Capabilities, State and Institutional infrastructure)   It is expected that proposals will utilise the AAF for common authentication services. Where it is not possible for the proposed infrastructure to utilise the AAF for authentication purposes, appropriate and reasonable justification has been provided. | | |
| **Financial and Co-investment** | **20%** |  |
| The project budget is well-described, matched to appropriate milestones and appropriate to the needs of the project | | |
| The identified co-investment achieves the target level, is appropriate to the needs of the project, and adequately covers the operational requirements of the proposal | | |
| The proposal identifies an appropriate model for delivering future sustainability of the infrastructure | | |
| Proposed expenditure of EIF funds is adequately described and conforms to the EIF funding guidelines (Mandatory) | | |
| The proposal conforms to the principles on Access and Pricing as described in Part B of the NeCTAR RFP (Mandatory) | | |

## Criteria for Research Cloud and National Server Program Proposals

|  |  |  |
| --- | --- | --- |
| **Category** | **Weight (%)** | **Criteria** |
| **Research Community** | **10%** |  |
| Research community to benefit is well-defined (by location, institutions, size), nationally significant and the proposal is well supported by the research community | | |
| The needs of the research community to be addressed by the proposal are well defined and significant | | |
| Researcher participation in the proposal is well defined, significant and national or international in scale | | |
| The proposal is aligned to national research priorities | | |
| The proposal is aligned with, and contributes to implementation of, national research infrastructure and eResearch priorities | | |
| **Research Impact** | **20%** |  |
| The benefits to be delivered to the identified research community are well described, achievable, significant and measurable   * Process for tracking and measuring the benefits is defined and achievable | | |
| Potential benefits from re-use across research disciplines are well described, realisable and significant | | |
| **Implementation** | **40%** |  |
| An appropriate Governance structure has been defined for the Project   * Proposers commit to operate under the Governance arrangements for the RC and NSP as described in the NeCTAR Final Project Plan (Section 4.3) * Research Cloud node proposers commit to operate in accordance with the principles described in Section 4.3.1.3 of the NeCTAR Final Project Plan * Authority structure over resources in the project has been described * Key personnel (if required) and their roles have been clearly defined | | |
| The infrastructure to be created by the proposal is well-described, achievable and will deliver infrastructure in a timely manner.   * Project Management is well-described and appropriate to the proposal scale * Key risks have been identified and are manageable * Issues that require solving have been identified * Dependencies with third parties have been listed | | |
| Capability and track record in development and operation of signifcant ICT infrastructure for research users   * The contribution and track record of each organisation in the proposal to development and operation of the infrastructure is well-described and appropriate to the successful delivery of the proposal infrastructure | | |
| The proposal leverages or builds upon existing research infrastructure where appropriate:   * Co-location with proposed nodes of the RDSI Project * Co-location of proposed Research Cloud and NSP nodes to achieve cost reductions * Co-location with other significant eresearch infrastructure (eg. HPC, Repositories Instruments) | | |
| **Financial and Co-investment** | **30%** |  |
| The project budget is well-described, matched to appropriate milestones and appropriate to the needs of the project | | |
| The identified co-investment achieves the target level, is appropriate to the needs of the project, and adequately covers the operational requirements of the proposal | | |
| The proposal identifies an appropriate model for delivering future sustainability of the infrastructure | | |
| Proposed expenditure of EIF funds is adequately described and conforms to the EIF funding guidelines (Mandatory) | | |
| The proposal conforms to the principles on Access and Pricing as described in Part B of the NeCTAR RFP (Mandatory) | | |

# Milestone and Funding Milestone Template

## Funding Estimate

Please add the details of any anticipated participating organisations in the below table along with their anticipated Funding Allocation as a percentage of the Proposer’s Total Funding estimate.

|  |  |
| --- | --- |
| Organisation / Group Name | Anticipated Distribution of EIF Funds (%) |
|  |  |
|  |  |
|  |  |
|  |  |

## Milestone Template

Complete the table overleaf with proposed milestones and the associated budgets and proposed funding amounts to be drawn down from NeCTAR. The submitted table must form an attachment to the Proposal and will be used to prepare the contract Schedules. Deliverables are to be described in as much detail as necessary to show that careful thought has been spent on planning. Example milestones shown may apply to a particular type of project, but are expected to be adapted to suit the needs of the project and NeCTAR Program.

**Note** – **Items in “Deliverables/Completed Activity” are mandatory.**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Funding Milestone**  **Yes / blank** | **Milestone Title** | **Deliverables/Completed Activity** | **Target Milestone Date** | **NeCTAR (EIF) funds ($thousands)** | | | | **Co-investment**  **(budgeted contribution value) (‘000)** |
| **Requested**  **(‘000)** | **Planned Expenditure breakdown** | | |
| **Labour**  **(‘000)** | **Equipment**  **(‘000)** | **Other**  **(‘000)** |
| 1 | Yes | Sub-contract signed |  | 31 Jan 2012 |  |  |  |  |  |
| 2 |  | Project Initiation complete | *Communications plan prepared and sent to NeCTAR (Signed contract + two months).* | 31Mar 2012 |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |
| 10 |  |  |  |  |  |  |  |  |  |
| 11 | Yes | Final Admin Closure | *Post-implementation Review (PIR) conducted and sent to NeCTAR.*  *Practical Completion Certificate accepted by NeCTAR.* | 30 Sep 2013 | 100  (last ten percent) |  |  |  |  |
| 12 |  | Operations to June 2014 | *Service Levels met and reported to NeCTAR as defined.* |  |  |  |  |  |  |