

Project Acronym:
Version:
Contact:
Date:



JISC Project Plan

Before completing this template please note:

- *The Project Management Guidelines have detailed instructions for preparing project plans.*
- *Please expand tables as appropriate.*
- *Fill in the information for the header, e.g. project acronym, version, and date.*
- *Prepare a cover sheet using the cover sheet template and attach to the project plan.*
- *This template is for completion by JISC funded project managers*
- *Text in italics is explanatory and should be deleted in completed documents.*
- *Please check with your programme manager before completing this form whether they would like to use a specially adapted template specific to your project.*

Overview of Project

1. Background

Digitizing data invariably provides opportunities for a wealth of communities to make use of new resources. Making the most of such data sets necessitates working closely with each community to ensure that their requirements are met, but this task is made harder when the potential users are from different fields, or when they include both academics and the general public. A particularly rich example lies in naval records; while of obvious interest to naval historians, both professional and amateur, these logs have recently been recognised by meteorologists and climate scientists as the ideal source of historical records. We propose to develop a tool to allow citizen scientists to participate in the digitization of a unique data set from World War I Royal Naval ships, bringing together our existing community of volunteers with climate scientists from the Met Office and beyond and with the team behind the popular Naval-History.net and their counterparts at the National Maritime Museum. By bringing together these disparate groups, we will thus not only produce new observations of historical climate for the period around the First World War, enriching and expanding the online availability of the historical records of the Royal Navy, but will produce the tools needed for a future worldwide effort. This project will expand the reach of our Zooniverse network of citizen science projects to new areas of science, and to new communities.

2. Aims and Objectives

The aim of the project is to demonstrate the viability of a distributed approach to the digitization of naval records. Specifically, we will produce an online interface that will allow the transcription of climate data and historical information from Royal Naval logbooks from the period of and around the first world war.

3. Overall Approach

By reusing the existing Zooniverse framework we will be able to take advantage of a suite of tools to support online citizen science. The main task is therefore to produce an engaging web layer as an interface to the digitized data, capable on one hand of allowing free response while also ensuring that searchable metadata fields are populated.

An initial interface and site will quickly be produced and used to make a proportion of the data available to visitors to the Naval-History.net website. The requirements of the ACRE climate science team are well established; a series of standard meteorological indicators to be transcribed (if present) whenever a page is viewed. The first interface will request this information from viewers, alongside essentials such as ship name and location. Building upon previous work by the ACRE5 team we will

Project Acronym:
Version:
Contact:
Date:

also offer at this stage, basic overlays on a Google Maps interface displaying the current position of the vessel and positions prior to the date entered. Crucially, freeform tagging will be allowed to encourage trial users to record what they consider important. The existing Zooniverse framework can provide a continually updating record of the most popular tags and later interfaces will be informed by the tagging efforts of the early users.

4. Project Outputs

A. A community-driven interface for future transcription efforts

This proposal is aimed at developing the community links and tools required to allow for a mass-transcription effort. The tools developed for this prototype will obviously be extendable to other navy and marine logbook records with the same basic structure, however we believe that by building flexible interface elements it would be possible to repurpose the tools for a wide variety of data sources. All interfaces developed for this project will be freely available as open-source tools for use by the wider academic community.

B. A wealth of data for the Naval-history.net community

The data collected during this project of historical value is expected to be of great interest and value to the Naval-History.net community. We will liaise with Naval-History.net to ensure that we provide the transcribed data in a convenient format and structured appropriately for their needs. The Naval- History.net will of course have full access to all of the transcribed data via the public data API.

C. Extending the timeline of measurements for current climate measurements.

For the extended First World War (1914-23) period that will be addressed in this pilot project, we will make a database of all the observations in the 200,000 logbook pages to be analysed. A typical logbook page contains six observations, each including air pressure; dry-bulb, wet-bulb, and sea temperatures; and wind speed and direction. The resulting 1.2 million new weather observations will be converted into International Maritime Meteorological Archive (IMMA) format, and then included in the International Comprehensive Ocean-Atmosphere Data Set.

5. Project Outcomes

The most important outcome of this project will be a community who are actively engaged in data recovery. The project outlined in this proposal will bring together three disparate communities: naval historians, climatologists and citizen scientists. The end result will be an active group of people working together to exploit historical records for a variety of ends. We will thus bring together three existing digitization efforts – ACRE, the National Maritime Museum’s project to digitize their archives, and that of Naval- history.net.

6. Stakeholder Analysis

Stakeholder	Interest / stake	Importance
Naval-history.net	Expert community and audience for final data	High
ACRE	Audience for final climate data	Medium to High
Citizen Science Alliance	Providers of the Zooniverse framework	Medium
National Maritime Museum	Interested in extending their digitization effort, showcasing project in galleries	Low

Project Acronym:
Version:
Contact:
Date:

7. Risk Analysis

Risk	Probability (1-5)	Severity (1-5)	Score (P x S)	Action to Prevent/Manage Risk
Staffing	2	4	8	Using existing Oxford staff
Organisational	3	4	12	Regular use of email list and meetings to ensure disparate parties remain informed
Technical	2	5	10	Reuse of existing framework were possible, choice of technologies were expertise is available in-house.
External suppliers	4	2	8	Early action needed to recruit external suppliers.
Legal	1	5	5	Confirmation from National Archives that we have copyright clearance received, advice taken from Oxford data protection staff.
Need for publicity	5	2	10	Media campaign needed to take into account sensitivity of climate change as a topic; recruiting external writers and reviewers for site text and press materials.

8. Standards

Name of standard or specification	Version	Notes
XHTML Translational	1.0	Used for all views
REST XML API	1.0 (XML)	User for public and private APIs
CSS		

9. Technical Development

We will be developing using a combination of specification-based and test-driven approaches. New (API) functionality will be specified based upon the interface and data access requirements and then built using a 'test first' approach. Code development will typically operate in cycles of between 1 and 2 weeks, with weekly builds of the current application being reviewed by team.

The Zooniverse team favour short development cycles combined with extensive unit, functional and integration tests of the application that they are building. Test-driven development (TDD) is a well-established method of building web applications especially when using the Ruby on Rails web framework.

10. Intellectual Property Rights

The images of the ship's logs were produced by the ACRE team at the Met Office from documents held in the National Archives, and they have obtained permission for their use in this project. The Zooniverse framework is owned by the Citizen Science Alliance Institutes who are committed to maintaining it (and hence the project) until at least 2014 and then either extending this term or

Project Acronym:
Version:
Contact:
Date:

releasing the software as open source. The interface developed under this grant and the data produced by it will be placed in the public domain as soon as possible.

Project Resources

11. Project Partners

1. The Citizen Science Alliance: From Galaxy Zoo to the Zooniverse

We bring to this application the strength of an international collaboration, with the aim of increasing the involvement of the public in real research. The Citizen Science Alliance, which has developed and runs the 'Zooniverse⁴' network of web-based projects, grew out of the success of the Galaxy Zoo project. Since the launch of Galaxy Zoo 2 we have been working to build a software framework for citizen science. This flexible infrastructure, built to support the widest possible variety of projects, allows us to test, launch and refine projects with large or small audiences. The project proposed here has particular overlap with Zooniverse projects in the field of ocean science, being developed in association with our partners at the National Maritime Museum, and with a project to digitise the collection of ancient papyri held in the Sackler Library of the University of Oxford.

2. Climate Science: the ACRE project

Our climate science partners for this proposal are The International Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative. ACRE aims to extend the available time period covered by calibrated and tested climate datasets in order to improve the testing and calibration of climate models by recovering millions of historical records. ACRE is run by a consortium of three primary partners, the Queensland Climate Change Centre of Excellence (QCCCE) in Australia; the Met Office Hadley Centre in the UK; and the US National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory (ESRL) and the Cooperative Institute for Research in Environmental Sciences (CIRES) at the University of Colorado.

ACRE is the only international initiative of its type and has endorsements from organizations such as the World Meteorological Organisation (WMO), the Group on Earth Observations (GEO), the Global Climate Observing System (GCOS), wide international support and the aid of various working groups of ⁴<http://zooniverse.org> GCOS and World Climate Research Program (WCRP), ACRE provides an umbrella that links together some 30+ projects, institutions, organisations, data rescue and climate applications activities around the globe. While we have good relationships with many ACRE partners we will work directly with colleagues at the Met Office who will be responsible for collecting and reporting the views of their colleagues.

3. Naval History: Naval-History.net and the National Maritime Museum

Naval-History.net was started in 1998 by Gordon Smith with the content of two naval history books. It aims to provide a scholarly source of Royal Navy and 20th century naval history. Highly recommended by the Naval Historical Branch of the MOD, Naval-History.net now receives over 5000 visitors per day and has amongst its records, a complete record of all RN casualties from 1914 to present and a day-by-day account of all naval activities from 1939 to 1942. The information collected during this project will provide naval historians and family history researchers with a wealth of new content for a period currently severely lacking in information.

The National Maritime Museum is working with Naval-History.net to publish approximately 30,000 ship histories online. We will be consulting with their naval history curator to ensure that the interfaces developed for this project provide information of historical value.

We will also work with Oxford University's community engagement department to pilot a school's project using the site, alongside existing Zooniverse teaching resources.

No formal project agreement is believed to be necessary.

12. Project Management

Project management will be the focus of grant PI and Citizen Science Alliance chair Chris Lintott, who will be responsible for assuming that our different partners remain informed of progress and appropriately consulted. Technical lead Arfon Smith will be responsible for coordinate internal and external software development and design.

Project Acronym:
Version:
Contact:
Date:

13. Programme Support

Advice on other program areas that may benefit from what we're developing would be very welcome.

14. Budget

Following the initial meeting with all partners, it became clear that the interface required could easily be built by the Oxford postdoc (Stuart Lynn) without impacting the other work that is necessary. We therefore propose to spend the money allocated for external development on external design by someone with strong experience in usability, and – if possible – on content to explain the site's operation and purpose to our disparate communities.

Detailed Project Planning

15. Workpackages

See enclosed document and GANTT chart.

16. Evaluation Plan

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
19 July – 6 August	Interface suitability	Does this interface record the necessary data.	Beta test with existing Zooniverse volunteers.	Naval-history and ACRE satisfaction.
9 August – 16 August	External evaluation of interface	Is the interface serving the needs of the community?	External evaluator	Sign off by external evaluator

17. Quality Plan

Output Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
July/August	Data set acceptability	Assessment by ACRE and naval-history teams	Sign-off by ACRE/Naval-history	Mechanism for data reduction	N/A

18. Dissemination Plan

Timing	Dissemination Activity	Audience	Purpose	Key Message
	First test.	Naval-history.net community	Identify likely extent and direction of use of freeform tagging.	'What would you want to say if presented with this image?', project

Project Acronym:
Version:
Contact:
Date:

				awareness.
	Beta testing	Existing Zooniverse audience	Beta testing; can non-experts produce useful data with this interface?	The Zooniverse is expanding to naval history and climate change.
	Public release	General Public	Launch	Explore WWI naval history and help science.

19. Exit and Sustainability Plans

Project Outputs	Action for Take-up & Embedding	Action for Exit
An operating 'oldweather website'	To be maintained by Citizen Science Alliance; ongoing community engagement required.	Site archived online by CSA once digitization is complete.
Data release to ACRE	Validation.	Submission to international databases
Data release to naval-history.net	Creation of online interface to 'reduced' data.	Release of 'final' dataset.

Project Outputs	Why Sustainable	Scenarios for Taking Forward	Issues to Address
Interface for digitization	Can be applied to other naval data sets.	Need to work with ACRE/Naval-history/new partners to identify and (perhaps) to image other logs.	Ongoing funding for digitization, site operation and development.
A community of volunteers engaged with the project	Self-sustaining given suitable space and encouragement.	Provision of more advanced tools; for example, allowing volunteers to collate information in a way that is useful to each other and the academic community.	Continuing integration of this project with new Zooniverse features.

Appendixes

Appendix A. Project Budget

Appendix B. Workpackages

Project Acronym:
 Version:
 Contact:
 Date:



JISC WORK PACKAGE

Project start date: 15 April 2010

Project completion date: 15 October 2010

Duration: 6 months

				Milestone	Responsibility
YEAR 1					
WORKPACKAGE 1: Requirements gathering <i>Objective: To understand interface/data requirements</i>	15 April	6 May	Definition of transcription entities from Naval-History.net & ACRE		Brohan, Allan, Smith (G)
1. Define historical transcription entities	15 April	22 April	List of desired output values		Smith (G)
2. Define meteorological transcription entities	15 April	22 April	List of desired output values		Brohan, Allan
3. Evaluate interface options (JS vs Flash etc)	22 April	6 May			Smith, Lynn
WORKPACKAGE 2: Prototyping interfaces <i>Objective: To test a variety of development tools for building interfaces and select most appropriate</i>	22 April	25 June	A working first beta interface		
4. Develop initial website/interface design concepts	22 April	11 June	Photoshop (PSD) visuals		Designer
5. Initial build of transcription tool	22 April	11 June	Functional development interface		Lynn

Project Acronym:
Version:
Contact:
Date:

Workpackage and activity	Earliest start date	Latest completion date	Outputs (clearly indicate deliverables & reports in bold)	Milestone	Responsibility
6. Additional design work for development interface	11 June	18 June	2 nd generation of transcription interface		Designer
7. User testing at NMM/Oxford	18 June	25 June	Feedback from members of the public		Lynn/Smith (A)
WORKPACKAGE 3: API modifications to support interfaces <i>Objective: To update/modify the Zooniverse API software to support the project</i>	22 April	11 June	An API layer able to support the transcription interface		Smith (A), Lynn
8. Evaluate data capture requirements	22 April	29 April			Smith (A)
9. Make necessary modifications to support data structures	22 April	11 June			Smith (A)
10. Test development transcription interface against updated API	11 June	25 June			Smith (A), Lynn
WORKPACKAGE 4: Interface updates <i>Objective: To incorporate feedback from initial user testing into transcription interface</i>	28 June	16 July	An updated transcription tool		Lynn, Smith (A)
11. Gather feedback from beta	28 June	2 July	Consolidated list of feedback from user testing		Lynn
12. Make changes to transcription tool	2 July	16 July			Lynn

Project Acronym:
Version:
Contact:
Date:

Workpackage and activity	Earliest start date	Latest completion date	Outputs (clearly indicate deliverables & reports in bold)	Milestone	Responsibility
WORKPACKAGE 5: Public beta <u>Objective:</u> <i>To refine/finalise transcription interface. Update site content to reflect feedback from community</i>	19 July	16 August	Feedback from public beta, updated interfaces, site copy etc.		Lynn, Smith (G), Smith (A), Brohan, Allan, Charman-Anderson, Lintott
13. Support public beta - communication with community. Respond to questions.	19 July	16 August			Lynn, Smith (G), Brohan, Allan
14. Review initial beta data	26 July	9 August	Processed transcribed data		Smith (G), Brohan, Allan
15. Evaluation of project by external assessor	9 August	16 August	Report/summary of success of project		Charman-Anderson
WORKPACKAGE 6: Beta-led interface refinements <u>Objective:</u> <i>To have a final build/design of the transcription interface and website</i>	23 August	6 September	A final version of the transcription interface		Lynn, Designer
16. Interface modifications	23 August	6 September			Lynn
17. Design modifications	23 August	30 August			Designer
WORKPACKAGE 7: Launch preparation/project launch <u>Objective:</u> <i>Final signoff of interfaces/content.</i>	6 September	20 September	A final version of the transcription interface		Lynn, Smith (A)
18. Final review of interface, record screencasts etc.	6 September	13 September	Tutorial material (screencasts)		Lynn, Smith (A)
19. API documentation	6 September	20 September	API documentation		Smith (A)

Project Acronym:
 Version:
 Contact:
 Date:

Workpackage and activity	Earliest start date	Latest completion date	Outputs (clearly indicate deliverables & reports in bold)	Milestone	Responsibility
--------------------------	---------------------	------------------------	--	-----------	----------------

20. Release of open source transcription tool	20 September	15 October			Smith (A)
21. Project launch	20 September	15 October			