

DMAOnline Project Plan (version 0.1)

[Executive Summary](#)

[Mockup for DMAOnline](#)

[Goals and Objectives](#)

[Stakeholders](#)

[Needs of stakeholders](#)

[Use case 1](#)

[Use case 2](#)

[Use case 3](#)

[Use case 4](#)

[Use case 5](#)

[Objectives](#)

[Deliverables](#)

[Schedule and Milestones](#)

[Phase I - March 2015 to May 2015](#)

[Content/Information](#)

[Plugins](#)

[Phase II - June 2015 to September 2015](#)

[Plugins](#)

[Phase III - October 2015 to March 2016](#)

[A visual approach for phased development](#)

[Resources](#)

[Financial](#)

[Technical](#)

[Human](#)

[Communication](#)

[Conclusions](#)

Executive Summary

While a lot of focus is given to research data management activities for researchers, the key component of research data administration analytics for managers and administrators is often overlooked. As RDM is typically a cross-departmental initiative (Library, Research and Enterprise Offices, IT services, etc), a single dashboard view of how these departments contribute to RDM activities is crucial. We propose that a dashboard solution is developed which can harvest data from various sources, including RIM/CRIS systems, repositories (EPrints, DSpace, Hydra), IT services storage space allocation, DMP creation statistics (per month/year, per faculty/department, per funding council), typical dataset sizes (per faculty/department), internal vs external repository usage, etc. This solution has the potential to scale very well and solve multiple real-life use cases while providing a factual way of creating and establishing on-going business cases for RDM. We call this solution “Data Management Administration Online” (DMAOnline).

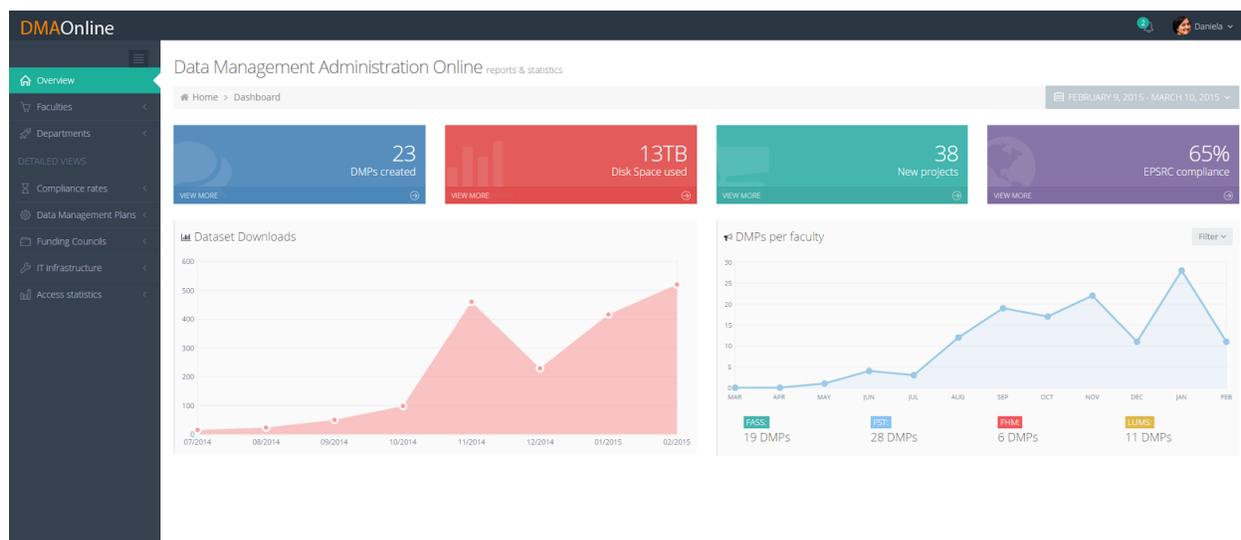
We are looking to develop DMAOnline as a simple solution that links with a single RIM/CRIS system and a single repository, and then grow it over time. The solution will be developed using a standards based approach allowing maximum interoperability options. The uses cases are explained later on in the project plan and they all solve real actual use cases that happen in universities.

From a service perspective, DMAOnline will be an online, platform independent and scalable dashboard solution with a plug-n-play architecture for data sources. DMAOnline will be able to provide all your Research Data Management (RDM) information in a single place. We believe that a solution like DMAOnline can play a significant role in providing a central and core interface to identify upcoming needs, develop ongoing business cases, and report compliance amongst many other uses. Furthermore, DMAOnline will be able to visualise complex and disparate research data information in way that it can be understood and processed by research administrators and HEI management.

In the long-term, we believe that DMAOnline has the potential to be a sector wide dashboard solution that can be extended beyond RDM and used by any HEI easily.

Mockup for DMAOnline

A mock up design for the dashboard may look like:



Goals and Objectives

Stakeholders

We have identified the following stakeholders that will be directly/indirectly affected by DMAOnline:

- Research Data Managers
- Repository Managers
- Research Support Officers
- IT Services (Storage) Managers
- Pro-Vice-Chancellor for Research
- Deans for Research (Faculty level)
- Head of Departments

Needs of stakeholders

Using the concepts from agile methodologies, we will define the needs as user stories. Following are a few examples (out of many) of the needs of the stakeholders.

Use case 1

As a Research Data Manager, I would like to know the number of Data Management Plans (DMPs) created against funding council grants (or beyond) so that support can be provided at point of need.

Use case 2

As a Repository Manager, I would like to see how many publications have been deposited which are RCUK funded and how many of them have had acknowledgements so that I can report compliance.

Use case 3

As an IT Services (Storage) Manager, I would like to know the trend of dataset generation and disk space usage so that I can incorporate additional storage purchase in next year's plan.

Use case 4

As a PVC(R), I would like to have an overview of my institution's response to the EPSRC expectations so that we can achieve compliance.

Use case 5

As a Dean for Research for Faculty X, I would like to identify the compliance rates of my faculty for funding council requirements on research data so that I can take appropriate steps if needed.

Objectives

The following objectives are required to be met to accomplish the core deliverables from this project.

ID	Title	Phase
1	Core architecture of DMAOnline (dashboard development)	I
2	Interface to EPrints for publications and project information	I
3	Schema definition of what CSV files can be ingested and in what format	I
4	Interface to ingest CSV files	I
5	Interface to the storage APIs where available	I
6	Interface to Pure for publications and project information	II
7	Interface to DMPOnline API (if ready by then)	II
8	Interface to Figshare for harvesting information about datasets	II

9	Interface to DSpace for harvesting publications information	II
10	Optional Interface to dropbox or box.net for remote storage information	II
11	Source code submitted to GitHub as Open Source	II
12	Technical documentation provided via GitHub	II
13	Supporting documentation provided via website	III
11	Report on priorities for Phase III developments from a community workshop	III

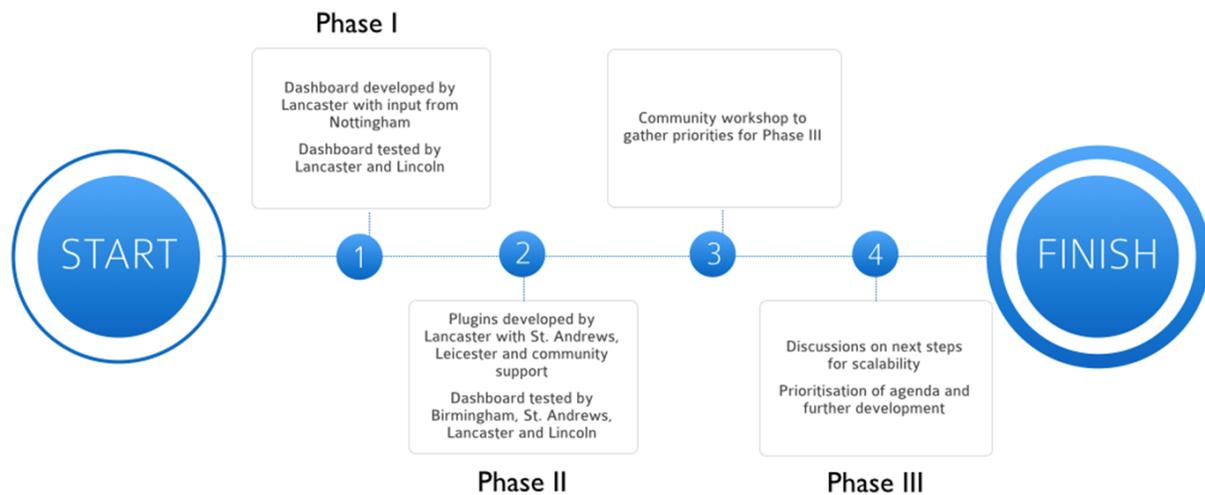
We believe that this list may expand over time depending on the success of the project.

Deliverables

We expect the following core deliverables from the project.

- Core dashboard with a plug-n-play architecture
- Business use cases
- Source code for the backend development, including all the plugins
- Prioritisation output report (from the community prioritisation workshop)
- Technical documentation
- Supporting documentation
- Website (with details of DMAOnline)
- Blog (for project activities)

Schedule and Milestones



Phase I - March 2015 to May 2015

Phase I will incorporate the core dashboard architecture and development, and the following content/information and plugins, directly satisfying some of the core needs of institutions at the time when EPSRC requirements are coming in action.

Content/Information

Data Management Plans (DMPs), Project Information, Active storage information, Publications per project, and Acknowledgement per publication.

Plugins

As a generic type, there will be a CSV plugin with well-defined schema. There will be plugins for EPrints, and Storage Infrastructure.

Phase II - June 2015 to September 2015

Phase II will enhance the outcomes from Phase I, making sure the plugins are plug-n-play (widgets) with the development of the following additional plugins.

Plugins

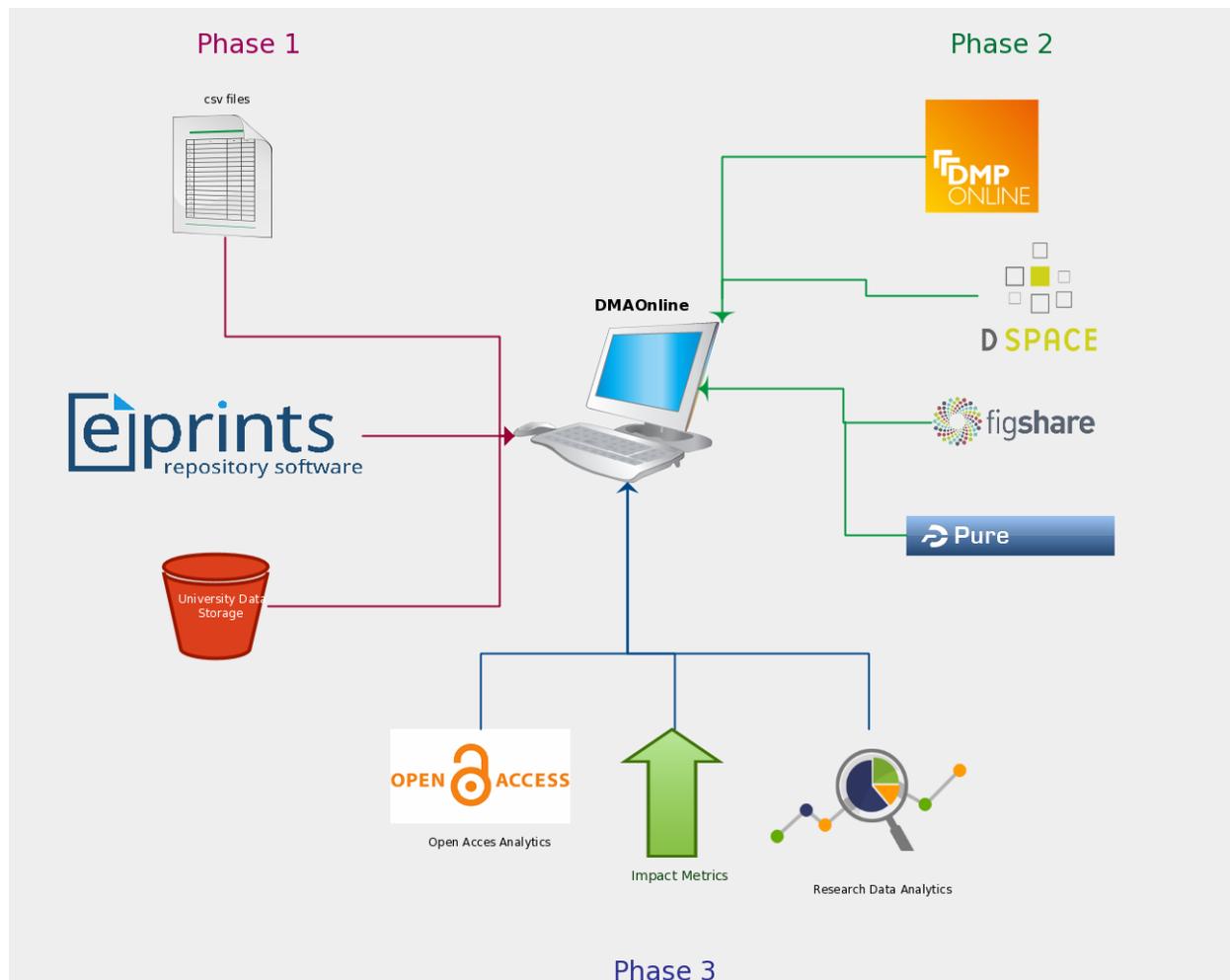
DMPOnline, Pure, Figshare, DSpace and optionally dropbox or box.net depending on requests.

Phase III - October 2015 to March 2016

Phase III will further investigate the scalability phase of the project. It is here that we discuss other potential uses of the DMAOnline, and possibilities of extending it for Research Data

Analytics, Open Access Analytics, Metrics for Impact or approaches towards collaborative benchmarking and cost analysis. We are proposing a community workshop to bring the relevant parties in a single room, show them the outputs of DMAOnline after phase II, and ask them to prioritise what's important for the future.

A visual approach for phased development



Indicators of success

For phase I, the following indicators of success will apply:

- A scalable plug-n-play based dashboard solution is designed and approved by all project partners.
- A plugin for EPrints developed which can harvest publication information, and a working scenario blogged about.

- A CSV plugin with a relevant schema developed for ingesting project information along with other information, approved by all project partners.
- A connection to the storage API (if available) provided for active data storage usage, and the scenario blogged about.

For phase II, the following indicators of success will apply:

- All Phase II plugins (DMPOnline, Pure, Figshare, DSpace) developed and dashboard solution adopted by project partners.
- DMAOnline to be provided as a SaaS based service within three months of Phase II.
- 20% of target audience from EPrints, Pure and DSpace communities to adopt DMAOnline within three months of completion of Phase II (depending on the schedule of their user groups).

For phase III, the following indicators of success will apply:

- A community workshop conducted to gather requirements and priorities for Phase III with positive feedback.
- A core prioritised requirement development cycle to start and finish by the end of Phase III.

Initial risk analysis

Risk	Probability (1-5)	Severity	Score (PxS)	Action to Prevent / Manage Risk
Staffing				
Loss of key project staff	1	5	5	Extensive development team so knowledge will be shared; clear documentation
Unable to hire technical resource for backfill	2	4	8	Develop the Job Description and Person Specification and advertise as soon as possible.
Organisational				
Miss project deadlines for deliverables	2	4	8	Short project phases with clear deliverables
Technical				

Poor component design causing performance issues	1	3	3	The architecture of the dashboard designed using a standards approach, with expertise sought from project partner (University of Nottingham)
Dependency on external services which may fail	3	4	12	To develop a local caching solution which saves the last accessed value if the external services are non-responsive. To put in appropriate time outs and messages to the end user.
No infrastructure support for running DMAOnline as a service	2	4	8	ULCC is willing to host the service (which may incur additional costs). We also hope Jisc would be keen on taking control of DMAOnline as a service in the future.

Resources

Financial

Phase I

0.2 FTE x 2 Developers = £3800

0.4 FTE x 1 Developer = £3800

0.1 FTE x 1 Project Manager = £2500

0.4 FTE x 1 Research Data Manager = £3300

0.1 FTE x 1 Project Sponsor = Contribution in kind

Partner workshop = £1000 (approx)

Total cost for Phase I = £14,400

Phase II

0.2 FTE x 2 Developers = £3800

0.4 FTE x 1 Developer = £3800

0.1 FTE x 1 Project Manager = £2500

0.4 FTE x 1 Research Data Manager = £3300

0.1 FTE x 1 Project Sponsor = Contribution in kind

Total cost for Phase II = £13,400

Phase III

0.2 FTE x 2 Developers = £3800

0.4 FTE x 1 Developer = £3800

0.1 FTE x 1 Project Manager = £2500

0.4 FTE x 1 Research Data Manager = £3300

0.1 FTE x 1 Project Sponsor = Contribution in kind

Total cost for Phase III = £13,400

Other costs

Community Workshop = £1000

Attending sandpit workshops = £1500

Pure User group meeting = Contribution in kind

EPrints User group meeting = £300

Dedicated website and blog = Contribution in kind

Overall cost for the whole project = £44,000 (approx).

Technical

The project will use Microsoft Azure cloud infrastructure utilising Linux virtual machines and SQL Server data storage. The front end dashboard will use a HTML5 framework with possibly a commercial theme such as Metronic. This will allow the team to concentrate on core functionality development rather than interface design. A management application will be developed using Ruby on Rails.

Human

Core team

Name	Job title	Institution	Role in project
Masud Khokhar	Head of Digital Innovation	Lancaster University	Project sponsor
Michael Dunne	Assistant Librarian (Digital Projects)	Lancaster University	Project manager

Hardy Schwamm	Research Data and Repository Manager	Lancaster University	Use case provision
John Krug	Systems and Analytics Manager	Lancaster University	Development
Andy Hartland	Library Applications Specialist	Lancaster University	Development
Adrian Albin-Clark	Digital Developer	Lancaster University	Development

Collaborative team

Name	Job title	Institution	Role in project
Tom Parsons	Researcher Training and Development Manager	University of Nottingham	Dashboard development, Project guidance, Testing
Bev Jones	Research Repository & Info Officer	University of Lincoln	Use case provision, Project guidance, Testing
Aslam Ghumra	Research Data Manager	University of Birmingham	Use case provision, Project guidance, Testing
Anna Clements	Head of Research Data and Information Services	University of St. Andrews	Project guidance, euroCRIS

Communication

Stakeholders include all project partners, Jisc, DCC, EPrints user community, Pure user community and all of UK and wider HEI institutions. Therefore it is vital that communication happens through user groups, events, and other closed events as well as using wider open dissemination channels such as blogging and tweeting. There will be a dedicated website to support this and all tweets will include #DMAOnline hashtag as well as any #dataspring hashtags.

Conclusions

We believe that there is a clear gap in the RDM landscape of a tool such as the proposed DMAOnline tool which will aggregate and visualise research data administration for a number of use cases.

DMAOnline would also be a very timely project with research funder expectations such as EPSRC's requiring to bring together RDM information from a variety of sources.

For the long-term, Phase III of the project and beyond, we see a potential for the our flexible dashboard approach to go beyond research administration information to include related topics such as Open Access or Impact Metrics.