

# **Sprint 3 Work Plan**

Edited by David Kay – Final - 30 January 2015

## **Table of Contents**

<b><u>SPRINT 3 WORK PLAN</u></b>	<b>1</b>
<b>1 - BACKGROUND</b>	<b>2</b>
<b>2 - DEVELOPMENT OUTCOMES</b>	<b>3</b>
<b>3 - SERVICE OPPORTUNITIES</b>	<b>3</b>
<b>4 - CURRENT POSITION AFTER SPRINT 2</b>	<b>3</b>
<b>5 - SPRINT 3 WORK PLAN FOCUS</b>	<b>5</b>
<b>6 - SPRINT 3 WORK PLAN DETAIL</b>	<b>6</b>
<b>7 - INTEGRATION MATRIX</b>	<b>8</b>
<b>8 - ADDITIONAL UNDERPINNING ACTIVITIES</b>	<b>9</b>

## 1 – Background

For the purposes of the mid-point review (18 December), Jisc Monitor development work was broken down into ten areas – eight areas that had been progressed in Sprints 1 and 2, and two integration areas tentatively scheduled for Sprint 3. The mid-point review (18 December)

- Reported on progress under each use case
- Reviewed design and development priorities for the final Sprint
- Considered ongoing underpinning and externally facing activities regardless of Sprint 3 focus

In this review the Project Board weighed potentially conflicting strategies, favouring Strategy 'A' on account of the time pressures faced by UK HEIs to address funding body and REF requirements:

- Strategy 'A' - Getting as close to service as possible in areas that have most clearly indicated immediate and tangible potential
- Strategy 'B' - Gaining a clearer assessment of service opportunities and other recommendations in areas that are least clear at present

Commitments to dialogue with specific projects and with the wider service development community were also noted, whilst recognizing they could be reversed in favour of short term objectives.

The following were regarded as immediate priorities:

- **CORE** – to investigate the potential linkages around people and compliance data, taking account of the new CORE API
- **Use Cases** – to publish the use cases gathered from the community as a 'website'
- **Data Model** – to consolidate the logical data model work from across the project; collaborating with Glasgow (E2E), to contribute the findings in to the CASRAI process

Whilst important in the wider development of the 'Jisc Open Access offer', the following were excluded from the Sprint 3 plan:

- **Publications Router** – to test the forthcoming Router Notification service as a feed for the Monitor Publications database
- **SHARE** – to review the opportunities opened up by the GUIDE approach; to test the SHARE notification service
- **SHERPA** – to specify potential interactions with the SHERPA team

## 2 - Development Outcomes

To determine the feasible development outcomes for the final Sprint, each area of development has been assessed and scheduled against the following potential outcomes (listed in a developmental order):

- Use Cases
- Data Model (including input to standardization)
- Specification
- Prototype development
- Design harmonization (with other Monitor prototype components)
- Monitor integration (between use cases)
- External integration (with other services)
- User review and testing
- Service opportunities linked to this work

The scale for assessing the outcomes is as follows:

Not Applicable (N/A)	Nothing delivered	Outline indication (POC)	Clear indication	Comprehensive approach
----------------------	-------------------	--------------------------	------------------	------------------------

## 3 - Service Opportunities

At the mid-point review, the priority service opportunities were as follows:

- **Option A – A common database** for collecting, managing and reporting on data about research outputs, used by data acquisition (B-D), local management (H) and aggregation (I) services
- **Option H – ‘Monitor Local’, an institutional OA management application** (previously referenced as MOAP [Managing Open Access Publication], based on the Option A database, providing a data feed to Option J)
- **Option J – A UK OA data aggregation**, based on the Option A database and which could take data feeds from Option H

Subsidiary service opportunities were noted:

- Options B through D – Three Monitor utilities for harvesting and mapping, providing publication, compliance and people data for other components in the UK OA ecosystem (for example Options A, H, J)
- Options E through G – Mechanisms for consuming data from and pushing data to other open services in the ecosystem (Core, Router, SHARE, SHERPA); subject to quality of data; these services might add value to Option A

#### 4 - Current Position after Sprint 2

This table illustrates how far each area had been progressed by the end of Sprint 2 (December 2014).

	Nothing delivered	Outline indication (POC)	Clear indication	Comprehensive approach						
	Use Cases	Data Model	Spec	Proto Type	Harm-onise Design	Monitor Integr-ation	External Integr-ation	User Review	Service Opps	
1 - Publications Harvesting										
2 - Compliance Checking										
3 - Publications and Compliance Database										A
4 - Metadata enhancement										
5 - GUIDE - Enabling Identification										
6 - Institutional OA Management (MOAP)										H, J, A
7 - Above-Campus Costs Aggregation										J, H, A
8 - Notifications to SHERPA										
9 - Notifications from ROUTER / SHARE										
10 - Compliance data from CORE										

## 5 - Sprint 3 Work Plan Focus

Areas to be progressed over the available 3 months (mid-January to mid-April 2015) principally address the following priorities:

- Getting as close to service as possible in areas that have most clearly indicated immediate and tangible potential
- Adding value to the known publications data that HEIs are already processing
- Enhancing direct impact on HEIs through engagement, collaboration and Pathfinder project liaison

It is emphasised that none of the areas will be ready for service.

This table illustrates the plan to progress priority areas by the end of the project (May 2015).

Not Applicable (N/A)	Nothing delivered	Outline indication (POC)	Clear indication	Comprehensive approach					
	Use Cases	Data Model	Spec	Proto type	Harm-onise Design	Monitor Integr-ation	External Integr-ation	User Review	Service Opps
	1 - Publications Harvesting								A, B
	2 - Compliance Checking								A, C
	3 - Publications and Compliance Database						n/a		A
	4 - Metadata enhancement								A
	5 - GUIDE - Enabling Identification							POC	A, D
	6 - Institutional OA Management (MOAP)							UX	H, J, A
	7 - Above-Campus Costs Aggregation							UX	J, H, A
	10 - Compliance & People from CORE				n/a		Depends		A, G

## 6 - Sprint 3 Work Plan Detail

Adoption of this work plan is subject to budget related options (subject of separate note to Ross).

### WP1 - Publications Harvesting

- No further work will be done on publications harvesting.
- A baseline understanding of the issues has been established (confirmed by dialogue with other projects and services) and some software methods have been usefully tested.
- Sprint 3 will concentrate on the (potential) published outputs known to and being processed by institutions.

### WP2 - Compliance Checking

- This service will feed data in to Monitor Local
- Develop processes that can be used to check aspects of 'compliance' for known published outputs recorded in the Monitor database
- Any interfaces (APIs) should not be unique to Monitor - we should build assuming we can feed in data from arbitrary sources
- Definition of what is to be checked will come out of existing work and sources such as funder policies and the UCL Pathfinder (<http://blogs.ucl.ac.uk/open-access/files/2014/12/UCL-Pathfinder-baseline-report-Dec-2014.pdf>).
- The results of compliance checking will be made available to the database via an API; it should be possible for other services to consume these results
- Some of this work is likely to be exploratory in the sense of finding out how easy/hard some aspects of compliance are to check automatically and identifying which checks can be done real-time and which need to be batch processed (as per Sprint 1)
- This work will involve dialogue with Core and ContentMine

### WP3 - Database

- This is part of 'Monitor Local'
- The database model will be enhanced as required to support the developments required in the other work packages
- A one-off set of Journal Titles will be seeded by import from KB+; this would need to be a repeatable update for a real service
- The database implementation will inform the logical data model deliverable

### WP4 - Metadata enhancement

- This is part of 'Monitor Local'
- Institutions will be able to use web forms to edit metadata regarding local publication ('Research Outcome') records
- Where updated publication information becomes available via such as notification services or trusted datasets (such as CrossRef and GtR), institutions will have the option for that data to overwrite their current local record
- This should build on the GUIDE / 'Did you know' approaches (WP5)

### **WP5 – GUIDE prototype**

- This service will feed data in to Monitor Local
- GUIDE (Getting Useful IDs Early) is an ID mapping (or ‘Co-referencing’) service idea based on the premise that, in the short and medium term, there will be no single or pervasive method for using consistent organisation, people and grant IDs throughout the supply chain
- This is envisaged as an open ‘co-referencing’ service that can be used by ‘Monitor Local’ and other applications
- Liaise with HEIs able to collaborate (e.g. Birmingham, Glasgow, Imperial, Kent) on a contained prototype for claiming and mapping IDs to enhance knowledge of institutional affiliation that can be used in processing notifications and Research Outcome (RO) records entered by others
- Author email addresses and ORCIDs will be harvested from CORE and provided by the participating institutions
- This data will be enable a ‘Did you know?’ service where institutions can be alerted about ROs which have been updated or about which they may not be aware involving their authors
- Candidate ‘Did you know?’ RO records will be visible for claiming and local enhancement in the MOAP web interface

### **WP6 – Managing Open Access Publications (MOAP)**

- This is the user facing part of ‘Monitor Local’
- Implement forms that demonstrate the value of this application to HEIs, relating to management of people links, compliance and costs, and demonstrating ease of entry and exploiting Monitor sourced data
- Development of the key Financial Data form will build on work already scheduled for KB+; note that it is intended that the form will be suitable for use by libraries across both systems
- Do not implement Task Management, though this will be a key utility for a worthwhile service
- As above (WP4, WP5), candidate ‘Did you know?’ Research Outcome records will be visible for claiming and local enhancement in the MOAP web interface

### **WP7 - Costs aggregation**

- This service will consume data from Monitor Local or alternative sources
- Work with Stuart Lawson to add all currently available data to the aggregation
- Extend the aggregation data model to encompass recording ‘logical’ compliance information for articles
- Add the ability to identify and report on duplicate records using DOIs (this is the minimum requirement - we might consider other methods for identifying duplicates and what we do with identified duplicates)
- Support data available from MOAP application
- Work with RCUK, COAF and Jisc to develop more reports (or refine existing reports) from the aggregation – for example, with date constraints

### **WP8 – Interactions with SHERPA**

- No development work will be prioritised in this area in Sprint 3

## WP9 – Incoming Notifications (Router, SHARE)

- No development work will be prioritised in this area in Sprint 3

## WP10 – Interactions with CORE

- This will feed data in to Compliance Checking and GUIDE
- Use the new Core API to extract compliance (WP2)
- Use the new Core API to extract person ID data (WP5)
- Use the new Core API to extract latest deposits (WP5)
- Note that this might require some small API modification by CORE, regarded as generally beneficial

## 7 – Integration Matrix

Integration is a key focus for Sprint 3 of the Monitor plan.

This matrix illustrates integrations where there is a dependency across the work undertaken within Monitor (as opposed to dependencies on external services).

From \ To	Compliance	GUIDE / DYK	MOAP	Aggregation
WP2 Compliance		As per MOAP	API - Results	n/a
WP5 GUIDE Did you know?	As per MOAP		Display DKN AOs	n/a
WP6 MOAP	API - Targets	Push Affil'd IDs		Enhance API
WP7 Aggregation	n/a	n/a	n/a	

The following interactions with external services are identified in the work plan.

From \ To	Compliance	GUIDE / DYK	MOAP	Aggregation
From CORE API	Evidence	IDs	n/a	n/a
From Router Notifications	n/a	Publication events	Via DYK	n/a
From SHARE Notifications	n/a	Publication events	Via DYK	n/a
From / To SHERPA	Send Real compliance	n/a	Get FACT guidance	Informs SHERPA



## 8 - Additional Underpinning Activities

- 1 - Online progress sessions for HEIs
  - Publicise Sprint 3 direction (Week 2 Feb)
  - Workshop (March 26)
  - Two / Three online sessions (Feb 10, Feb 24)
  
- 2 - Publish database ERA model
  - Workshop (Jan 16)
  - Feedback to CASRAI
  - Possible link via PLOS to ESAC
  
- 3 - Publish web version of Requirements Catalogue (Week 9 Feb)
  - Liaise with PLOS on this
  
- 4 - Publish IDs Report
  - Liaise with Birmingham, Glasgow, Imperial and Kent subject to their availability
  - Could liaise with ARMA, ORCID pilots and SHARE on this
  
- 5 - Liaison with specific Pathfinders
  - Glasgow – Database, IDs
  - Edinburgh - Costs
  - UCL - Costs
  - Hull – Database, Costs
  - Bath, Manchester – Local processes and workflows
  
- 6 – Participate in Jisc dialogue with
  - Submission Systems – notably Aries / ScholarOne, potentially with Wiley
  - CrossRef
  - SHARE
  - CHORUS
  
- 7 - Subject to overarching Jisc OA plan
  - Concluding event and/or workshop(s) for the HE community?
  - Contribute to OA blog as scheduled
  - Input to Jeremy Atkinson review
  - Update Jisc strategic groups as applicable – Funders, Gold OAIG, SCAG, etc