DPLA Technical Development Plan

# Background

This document provides a high-level view of the integrated development plan for the DPLA platform and front-end.

# High Level Goals

* Provide a technology platform that supports the DPLA front-end
* Build a compelling front-end that demonstrates the potential of the DPLA platform
* Provide an API for open access to meta-data within the repository
* Deliver by April 2013

# Approach

We are following an iterative development approach for the platform with the goal of publicly deploying an updated version every four weeks between now and the beta launch in April 2013. All development will take place in the open, code will be available on github throughout the process, and issue trackers will be public. The platform development team will work with the front-end design and development teams to ensure that dependencies and deliverables are clearly defined.

# Scope

The high-level platform scope for the beta release is split into two phases based on delivery date. The scope will be updated as necessary to meet the needs of the front-end development.

## October 2012

* Metadata API for items, collections, contributors, and events
* Metadata API technical infrastructure (monitoring, logging, versioning)
* Metadata API file formats: JSON and XML
* Metadata ingestion API
* Metadata ingestion via one-time file load (MARC21, EAD)
* Metadata ingestion via OAI-PMH
* Metadata enrichment: work-level entity clustering, authorities resolution
* Schema and repository for items, collections, contributors, and events
* Ongoing import of metadata from existing repositories

## April 2013

* Metadata API for creators and users
* Metadata API technical infrastructure (authorization, authentication, user management)
* Metadata ingestion via incremental or updated file load (MARC21, EAD)
* Metadata enrichment: de-duplication, normalization, remediation process, Wikipedia integration
* Framework for ingestion of metadata from custom schemas
* Schema and repository for creators and users
* Ongoing import of metadata from existing repositories
* Javascript widgets for embedding DPLA data on third-party websites
* Documentation for deploying a local instance of the DPLA platform

## Front-end Scope

The front-end scope will be informed by the front-end use cases and fully defined through the front-end design process

# Key Internal and External Dependencies

|  |  |  |
| --- | --- | --- |
| **Dependency** | **Owner** | **Required By** |
| Complete the use cases that will inform the front-end requirements and design | Audience and Participation Workstream | July |
| Select a team to design and build the front-end | Secretariat / Steering Committee | July |
| Assemble team to build the platform | Technical development | July |
| Define the nature and scope of content to be included in the repository for launch. This information is required to scope, design, and implement the front-end and the content ingestion process. | Content Workstream | August |
| Content to support the front-end is available and ready to be loaded | Content Workstream | October |
| Identify any additional technical or content requirements for the platform resulting from the design of the front-end | Front-end design team | October |

# Key Risks

|  |
| --- |
| **Risk** |
| The nature and quality of the metadata to be loaded for the Beta launch is not known. Inconsistency in the metadata, or variations in metadata quality could degrade the usefulness of the API. |
| The development team is not yet staffed to successfully deliver against this plan. |
| The content providers for the metadata for the beta launch are not yet defined. If sufficient time is not available to load and QA the content, the quality of the metadata returned through the API could be degraded |
| Aggregation and enrichment of metadata may be more complicated than expected, leading to poor search results through the API |

# Timeline

