History of IIIF at Harvard Library

- **2000–2014:** Still images and page-turned objects in Library’s preservation repository (DRS) delivered by custom local solutions.
- **2015:** Page-turned objects migrate to IIIF/Mirador instance.
- **2018:** Still images migrate to separate IIIF/Mirador instance.
Digital assets at Harvard Library

7,324,247 digital objects total

- 7,132,049 images
- 156,346 page-turned objects
- 14,437 audio files
- 1,761 videos
Why the Redesign?

- Seeing increased use of our digital collections in the context of remote research and teaching
- Modernize IIIF-powered image delivery infrastructure
- Support multi-tenant storage of assets and manifests
- Decouple delivery of assets from our preservation repository
- Resolve performance issues in our current infrastructure
The Elevator Pitch

As the use of Harvard’s digitized content continues to grow, the IIIF infrastructure redesign project will enable current and prospective users to access these resources quickly, reliably, and delivered with the latest IIIF tools. This project will also enable our stakeholders to utilize our centralized infrastructure so they do not have to maintain their own.
Project Structure

- Agile Scaled Scrum Program
  - Two active teams working in sync (IIIF assets and IIIF services)
  - Daily stand-ups for each team
  - Weekly standup for team leads (Scrum of Scrums)
  - Backlog population and sprint planning
  - Sprint retrospectives
  - Open Demos

- Later
  - Spin up third team for IIIF viewer
Phased Implementation

● Phase 1 (MVP) - End of 2020
  ○ Fast, reliable delivery
  ○ Multi-tenant asset and manifest support
  ○ IIIF 2.1 and 3.0 presentation support
  ○ Persistent IIIF unique identifiers
  ○ Descriptive metadata (stretch goal)
  ○ IIIF full-text search and display (stretch goal)
Phased Implementation

- Phase 2 and Beyond (2021+)
  - Mirador 3
  - HL-provided annotation tools
  - Digital lending
  - Delivery AV resources (and eventually 3D)
  - Ability to download highest res version
  - Right to left viewing support
  - Export of resources for collections-as-data
Architecture Principles

Separation of Concerns
- Modularity - Isolate business logic into modular components

Don’t Repeat Yourself (DRY)
- Avoid repetition and overlapping implementation of the same functionality

Single Responsibility Principle
- A module has a single responsibility, avoid monolith modules and functions

Principle of Least Knowledge
- Encapsulation of implementation details
- Low coupling: A component should have the least amount of knowledge about implementation details as possible, to avoid tightly coupled interdependencies
Microservices

- A microservice is a modular service that has a single responsibility
- Microservices achieve many architecture principles such as Separation of Concerns, DRY, Single Responsibility, and Principle of Least Knowledge
- Low coupling
  - APIs to abstract away implementation details of other components
- Examples:
  - Asset Lookup is a single microservice responsible for lookups only
    - Translates a requested URL to an IIIF url to find the image location
  - Authentication is a single microservice responsible for granting tokens
  - Converter service is a single microservice responsible for converting manifests
    - API can be used by other services for manifest conversion
    - This converter will be contributed back to the community
Feature Highlights

- Specification for persistent identifiers for assets
- Auth app
- Asset database
- Image server
- Asset lookup
Feature Highlights

- Manifest validator microservice
  - Added complete field support, will contribute back
- Manifest converter microservice
  - Added full version 3 support, will contribute back
- Administration API for ingest of manifests
- Presentation API for optimized delivery of IIIF manifests and collections
Technologies

Languages

- Python (Flask) and JavaScript (NodeJS)

Infrastructure

- Docker and Ansible

Image Server

- Cantaloupe

Databases

- PostgreSQL and MongoDB
Work in the current sprint

IIIF Services Team

- **Sprint goal**: solution deployed to development environment
- Progress on productionizing solution deployment:
  - IIIF Services API moved to development environment
  - Storing canvas information in development environment
- Enhancements to manifest ingest
- Enhancements to the Presentation API
- Data model enhancements for administrative metadata
- Requirements gathering for authorization
Work in the current sprint

IIIF Assets Team

- **Sprint goal**: query the lookup service from the image server
- Completed: image server now queries the lookup service directly, lookup service verifies tokens and returns a policy statement
- In flight: working on the Asset Delivery app and infrastructure orchestration
- Policy work in progress
  - Maximum size of image thumbnails
  - Recommendations for handling non-compliant JPEG2000 images
  - Recommendations for tiered storage of delivery images
Demo

Asset Lookup

- Lookup service now checks for the presence of an auth token and returns a policy statement based on whether or not the user is authenticated or not.
Demo
Questions?

Harvard Library IIIF Redesign project wiki:  
https://wiki.harvard.edu/confluence/x/ABRiDw