LOCKSS Software Re-Architecture

34th International Conference on Massive Storage Systems and Technology
May 14, 2018
Santa Clara, California

Thib Guicherd-Callin <thib@cs.stanford.edu>
Technical Manager, LOCKSS Program
Digital Library Systems and Services, Stanford University Libraries

"Lego Building Blocks" by Traumsicherheit (CC0 1.0 https://pixabay.com/en/lego-building-blocks-colorful-2285065/)
Overview

1. Shifting Sands
2. Legacy
3. Changing LOCKSS
4. Code and APIs
Shifting Sands

"Shifting, Howling Sand" by Bill Morrow (CC BY 2.0 https://www.flickr.com/photos/billmorrow/37520367760)
On the Web

- Dynamic Web
- Web replay
In IT

- Appliance model
  - Storage
  - Computation

- Management
  - Machine environment
  - Virtualization
  - Outsourcing
Legacy

"Love Padlocks on the Corktown Footbridge in Ottawa, August 14th, 2017" by William Chen (CC BY-SA 4.0 https://commons.wikimedia.org/wiki/File:Love_Padlocks_on_the_Corktown_Footbridge_in_Ottawa,_August_14th,_2017.jpg)
What's In the (LOCKSS) Box?

- 1,100+ Java classes
- 750+ test classes
- Single codebase
- Single stack
Custom Implementations

- LOCKSS Web Crawler
- Web replay
- Repository (storage)
Things Change

- Project hosting
- Source control management
- Build system
- Unit testing
- Continuous integration
- Binary distribution
- Development environment
Changing LOCKSS
LAAWS Initiative

- "LOCKSS Architected As Web Services"
- Mellon Foundation grant
- Modernization effort
Goals

● Re-architect LOCKSS daemon as collection of Web Services
  ○ WARC storage
  ○ Memento access
● Use off-the-shelf open source components
● Minimize LOCKSS-only components
Desiderata

- LOCKSS protocol compatibility
- LOCKSS plugin compatibility
Splitting Up Functionality

- Components -> services
- REST API
- Considerable refactoring
Technologies

In progress

- Maven, Maven Central
- JUnit 4, JUnit 5, Mockito
- Spring Framework 4, Spring Boot 1
  - Spring Framework 5, Spring Boot 2
- Swagger 2
  - OpenAPI 3
- OpenWayback, Pywb
- Solr
- Docker, Docker Hub

Under consideration

- Orchestration
Repository Architecture
Code and APIs
LOCKSS Developer Pages

- [https://lockss.github.io/](https://lockss.github.io/)
- [https://lockss.github.io/laaws/code-artifacts/](https://lockss.github.io/laaws/code-artifacts/)
Core

- LOCKSS Core (org.lockss.laaws:lockss-core)
  - Stable version: 1.1.1
  - Development version: 1.2.0-SNAPSHOT
  - GitHub: https://github.com/lockss/lockss-core
Repository

- **LOCKSS Repository Core (org.lockss.laaws:laaws-repository-core)**
  - Stable version: 1.11.4
  - Development version: 1.12.0-SNAPSHOT
  - GitHub: [https://github.com/lockss/laaws-repository-core](https://github.com/lockss/laaws-repository-core)

- **LOCKSS Repository Service (org.lockss.laaws:laaws-repository-service)**
  - Stable version: 1.7.0
  - Development version: 1.8.0-SNAPSHOT
  - GitHub: [https://github.com/lockss/laaws-repository-service](https://github.com/lockss/laaws-repository-service)

- **LOCKSS WARCImport Client (org.lockss.laaws:laaws-repository-client)**
  - Stable version: 1.5.0
  - Development version: 1.6.0-SNAPSHOT
  - GitHub: [https://github.com/lockss/laaws-repository-service](https://github.com/lockss/laaws-repository-service)
Metadata

● LOCKSS Metadata Extraction Service
  (org.lockss.laaws:laaws-metadata-extraction-service)
    ○ Stable version: 1.0.0
    ○ Development version: 1.1.0-SNAPSHOTS
    ○ GitHub: https://github.com/lockss/laaws-metadataextractor

● LOCKSS Metadata Service (org.lockss.laaws:laaws-metadata-service)
    ○ Development version: 1.0.0-SNAPSHOTS
    ○ GitHub: https://github.com/lockss/laaws-metadataservice

● LOCKSS Metadata Extractor Common Code
  (org.lockss.laaws:laaws-metadataextractor-common)
    ○ Stable version: 1.2.0
    ○ Development version: 1.3.0-SNAPSHOTS
    ○ GitHub: https://github.com/lockss/laaws-metadataextractor-common
Polling and Repair

- LOCKSS Poller Service (org.lockss.laaws:laaws-poller)
  - Development version: 1.0.0-SNAPSHOT
  - GitHub: https://github.com/lockss/laaws-poller
Infrastructural

- LOCKSS Configuration Service
  
  (org.lockss.laaws:laaws-configuration-service)
  
  ○ Stable version: 1.0.0
  ○ Development version: 1.1.0-SNAPSHOT
  ○ GitHub: https://github.com/lockss/laaws-configservice