LIBIS/Aware conformance checker

LIBIS
INTO INFO

Agenda

- Functional analysis
- Technical design
- Research TIFF

Conformance checker

- Checks conformance with TIFF
  - Embedded metadata
  - TIFF files
- Every error can be fixed one by one
- Fix all option as well
  - For fixing files: convert to preservation safe TIFF
- Export SIP
  - With audit log of changes (PREMIS) included in METS file

Policy checker

- Supports extra validation step based on policy
  - = requested functionality
  - Based on file characteristics
- Extra functionality added
  - Supports SCAPE preservation policy Framework
  - Help users create a policy
    - Import existing SCAPE policies as example
      - Use exisiting community efforts
  - Easier to create and maintain the policy
  - Flexibel
    - Topdown or bottom-up (what ever the user prefers)

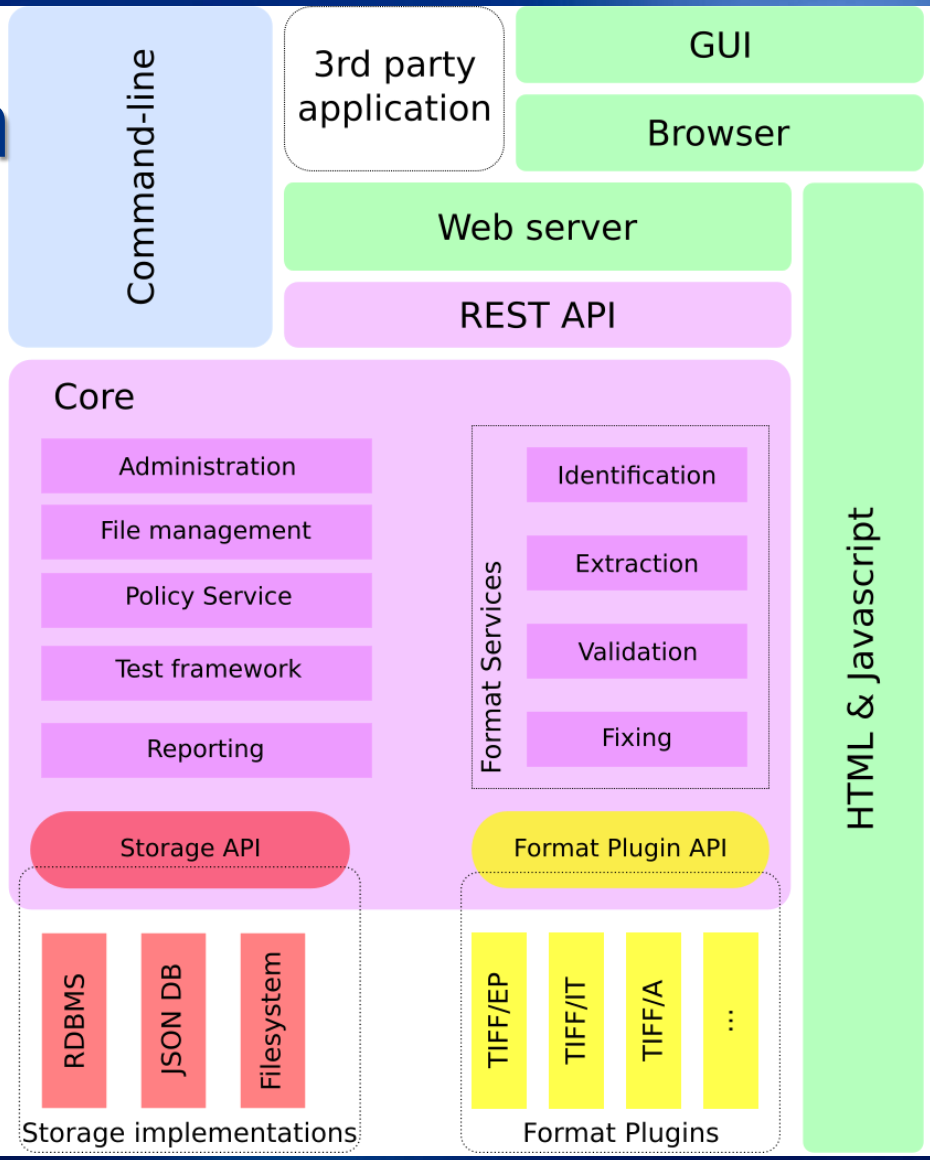
Test Framework

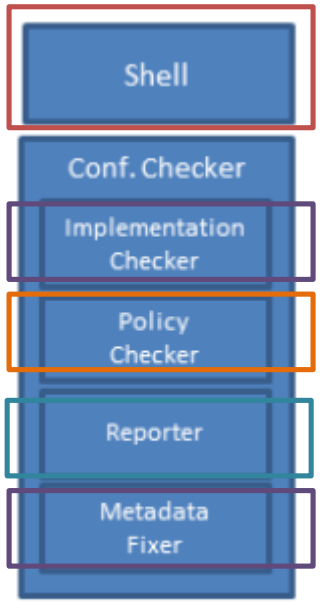
- Test the robustness of the conformance checker
- Trust the output of used tools, for example
  - memory institution wants to test new tool for creating TIFF files. They can use the test framework to see if the new tool also creates the same output
  - Developer wants to test if the change made to the TIFF output doesn't create problems in the TIFF
- Ground truth file describes the TIFF file
  - File characteristics with the values to check and errors contained in the file

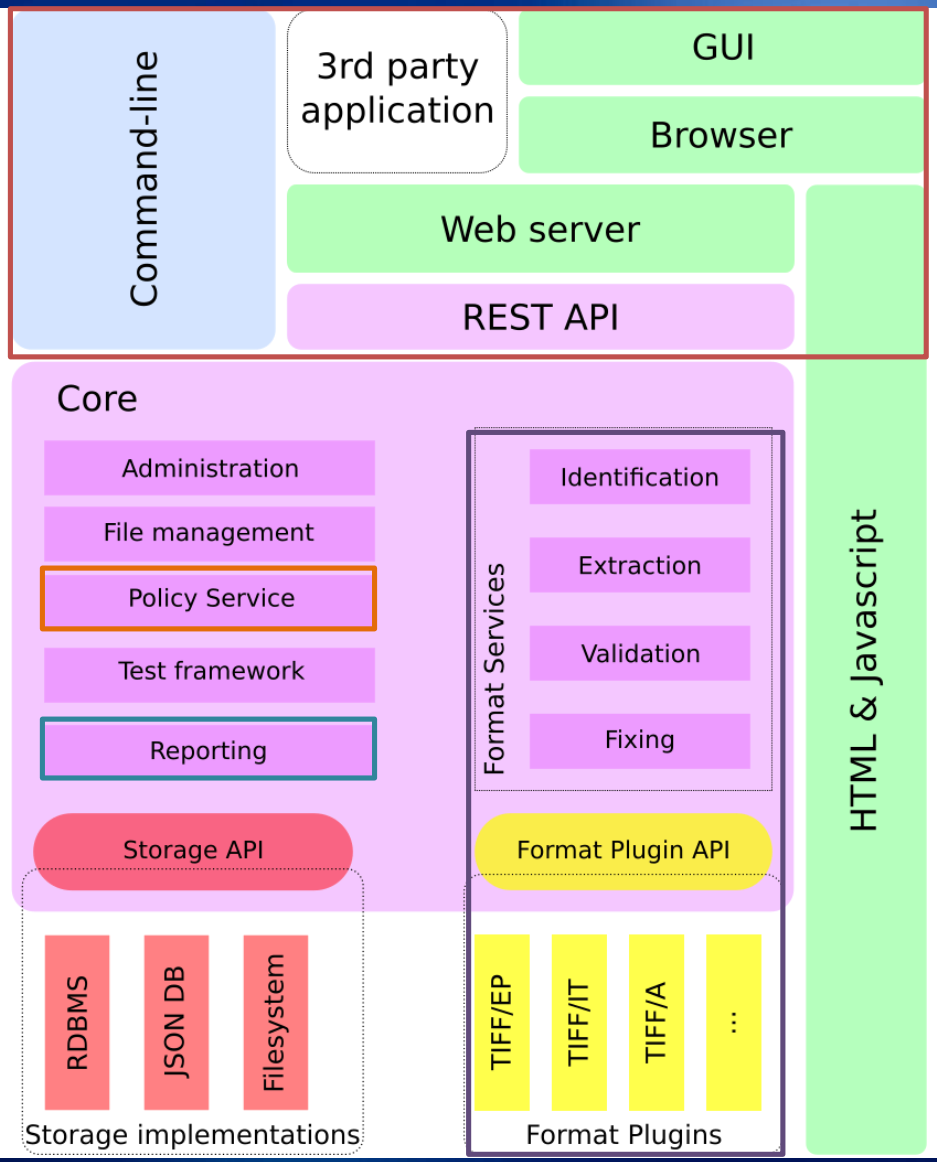
Technical design

- Flexible design:
  - Different ways of integrating:
    - API
    - Library (= Core)
    - Command line tool
  - Different User interfaces
    - Command line
    - Web interface
    - … (extensible)
  - Flexible Storage layer
    - Different options (JSON, NoSQL DB or Relational database)
    - Separate from other layers

Technical design

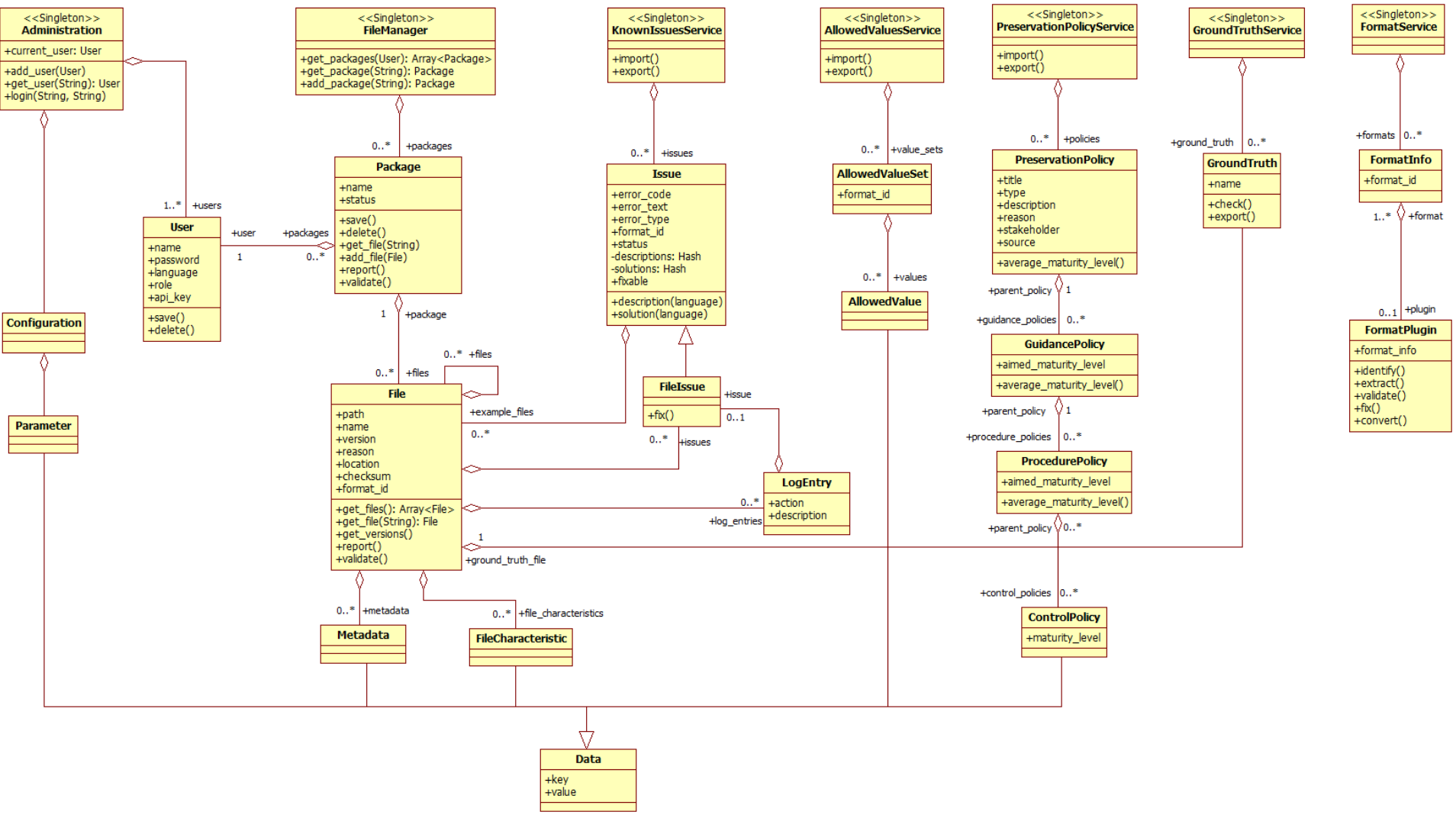



Shell
Conf. Checker
Implementation Checker
Policy Checker
Reporter
Metadata Fixer


Command-line
3rd party application
GUI
Browser
Web server
REST API
Core
Administration
File management
Policy Service
Test framework
Reporting
Storage API
Format Services
Identification
Extraction
Validation
Fixing
Format Plugin API
HTML & Javascript
RDBMS
JSON DB
Filesystem
Storage implementations
TIFF/EP
TIFF/IT
TIFF/A
...
Format Plugins

LIBIS
INTO INFO

Core Class diagram



Validation

- Databases to store validation configuration
  - Database of allowed values
    - = allowed values according to the specification
    - Allowed file characteristics
    - Can also be a function in the plugin for more complex validation
    - Provided by file format plugin
  - Database of known issues
    - Known errors in files
    - Provided by file format plugin
- Not in code, but as configuration
    - => more flexibility
    - Changeable without changing the code

Technology used

- JRuby
  - can integratie with Java and Ruby programming languages
- Third party libraries
  - Most licenses are MIT and this can be relicensed
  - Some are GPLv3 and MPLv2- compatible licenses

TIFF/A specification document

- Specification that discusses:
  - Problems in the TIFF formats, with example files to illustrate the problem
  - The solutions for the identified problems
  - The difference between the open source and Pro version of the TIFF file format plugin in a technical detail

TIFF/A specification document

- Preservation-safe TIFF
  - Not a new format
  - Subset of TIFF
  - Clearifies the current TIFF specification and makes improvements on real-life practices
  - It also contains advice for developers of decoders on how to implement the specification
  - Solves preservation and data exchange issues
  - Also excludes licensing issues, it is ‘license safe’

TIFF/A evaluation

- Evaluation is on going:
  - other TIFF experts (one from Adobe)
  - On the libtiff mailinglist
- Evaluation by the experts
  - Is the text clear?
  - Missing components?
  - Are the choices we made good?
    - BigTIFF is added
    - Only uncompressed en lossless compression (flate) is allowed
- Will be input for discussion with ISO in phase 2
  - Evaluation is important input for ISO

Questions?

- Documents and test files: http://preformatiff.github.io/TiffGuide/