

*Tools and Trends, 1-2 November,  
Koninklijke Bibliotheek, Den Haag*

---



## **Planets: Towards Infrastructure for Digital Preservation Services**

**Adam Farquhar  
Helen Hockx-Yu  
The British Library**

# Outline of presentation

---

- ❑ Other talks today have highlighted
  - The project goals
  - Preservation planning
  - Preservation action
  - Characterisation
- ❑ This talk
  - Brief aside
  - Planets architecture and conceptual model
  - The Planets testbed
  - The Planets interoperability framework



## (An aside: The simple office document myth)

---

- ❑ Are office documents are simple?
  - Office documents can contain:
    - Multiple character sets
    - Left-to-right, right-to-left, bi-directional text
    - Images, sound, video, vector graphics
    - Annotations and changes from multiple authors
    - Arbitrary metadata and XML components
    - Complex mathematical equations
    - Animated transitions
    - Embedded data, database connections, queries, cached data
    - Embedded components from other applications
  - Office documents have complex syntax that matches some of their complex semantics
- ❑ Archival practice long recognises the need to represent context and sufficient information to understand the semantics of a record



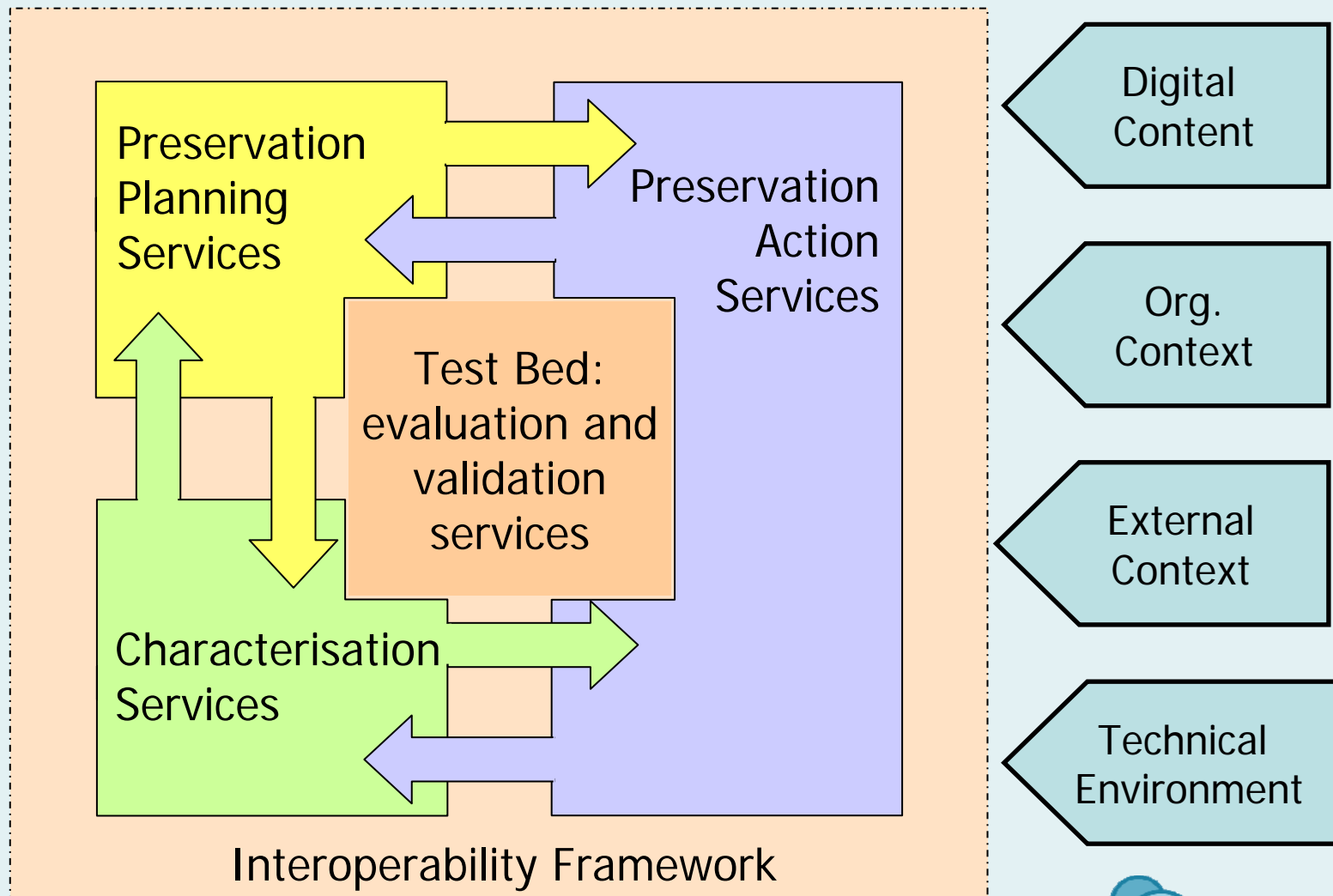
# Planets aims

---

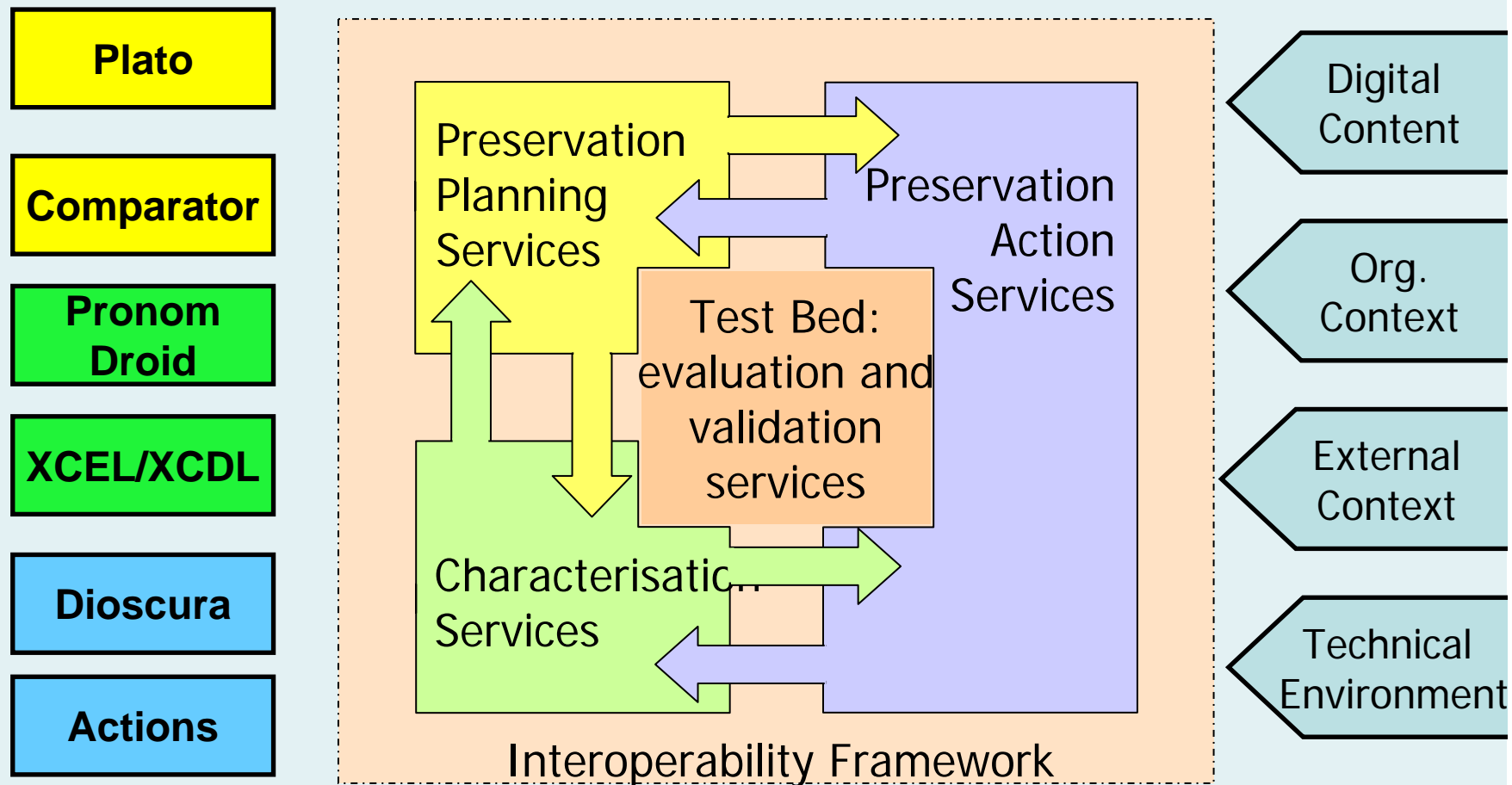
- ❑ Increase Europe's ability to ensure long-term access to its cultural and scientific heritage
  - Improve decision-making about long term preservation
  - Ensure long-term access to valued digital content
  - Control the costs of preservation actions through increased automation, scalable infrastructure
  - Ensure wide adoption across the user community and establish a market place for preservation services and tools
- ❑ Planets methods, tools, and services will enable organisations to diagnose, treat, and validate problems with their digital objects



# Planets architecture



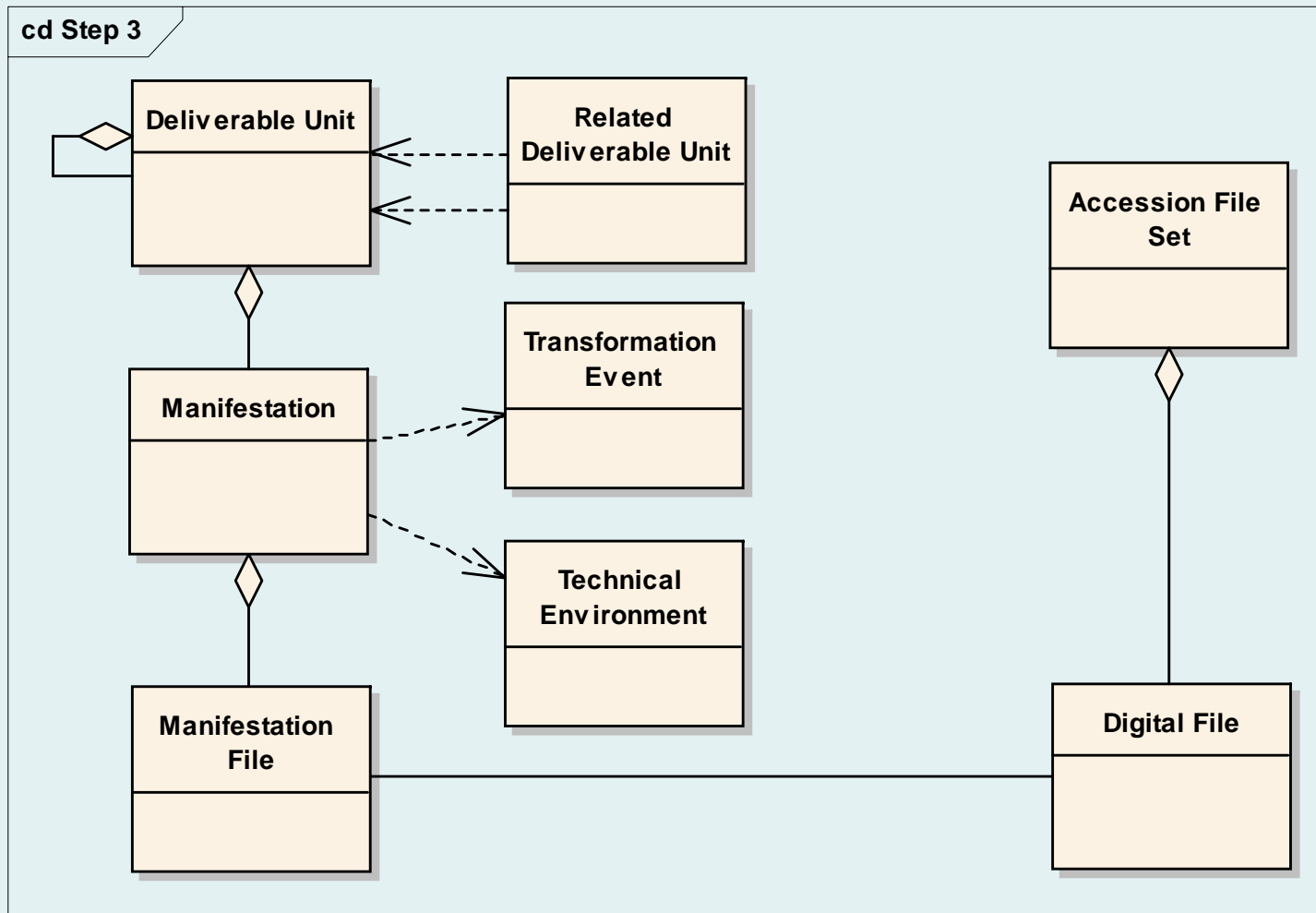
# Planets components in context



(listed are only those from today's presentations)



# Planets conceptual model – key classes



# The Planets digital preservation Testbed

---

- ❑ Digital preservation practice is still emerging
  - Substantial conceptual work
  - Ad hoc project approaches using locally selected tools
  - A craft!
- ❑ There is no systematic analysis of preservation strategies or tools and services
- ❑ Result
  - Poor and inconsistent decision making
- ❑ Planets approach
  - Provide systematic evaluation, benchmarking, assessment
  - Planning, characterisation, migration, emulation tools
- ❑ Move from craft to science



# Why do we need Testbeds in Digital Preservation?

---

- ❑ If we want to
  - Perform scientific research in digital preservation
  - Evaluate preservation approaches in diverse “real life” settings
  - Avoid duplication of work
- ❑ We need a dedicated research environment
  - Systematic execution of experiments by different institutions
  - With experiments that
    - Follow a formal methodology
    - Are reproducible
    - Are documented and accessible for analysis and comparison



# A Testbed for Digital Preservation

---

Planets definition

**Testbed:**

“A controlled environment for experimentation and evaluation, with metrics and benchmark content that allow comparison of preservation tools and strategies”



# Role of Testbed in Planets

---

- ❑ Test and validate Planets technical solutions and approaches:
  - Provide a controlled hardware and software environment for testing and evaluating **preservation action** (migration, emulation) and **characterisation** tools and services
  - Record experiments data in registries for further analysis and comparison
  - Assist the validation of the effectiveness of different digital **preservation plans**
- ❑ Improve preservation plans with empirical evidence
- ❑ Assess the suitability of the approaches across „real life“ scenarios in various organisations
  - Analyse applicability of the outcomes of Planets in existing workflows and organisational contexts
  - Evaluate their efficiency in providing practicable solutions for organisations engaged in digital preservation



# Role of Planets Testbed in the Digital Preservation Community

---

- ❑ In a second phase, the Planets Testbed will
  - Offer services to organisations outside Planets:
    - Support institutions to test preservation tools and services against benchmark content
    - Assist institutions to validate their preservation plans against their policies & content profiles
  - Enable developers and third party vendors to submit tools for benchmarking and certification:
    - Validate the suitability of their tools in preservation workflows



# Planets Testbed Application

---

## Testbed application provides:

- Clearly structured and formal process for preparing, executing, and evaluating experiments
- Repeatability of experiments, comparability and traceability of results
- Benchmarking of services
- Access to evidence base of previous experiments

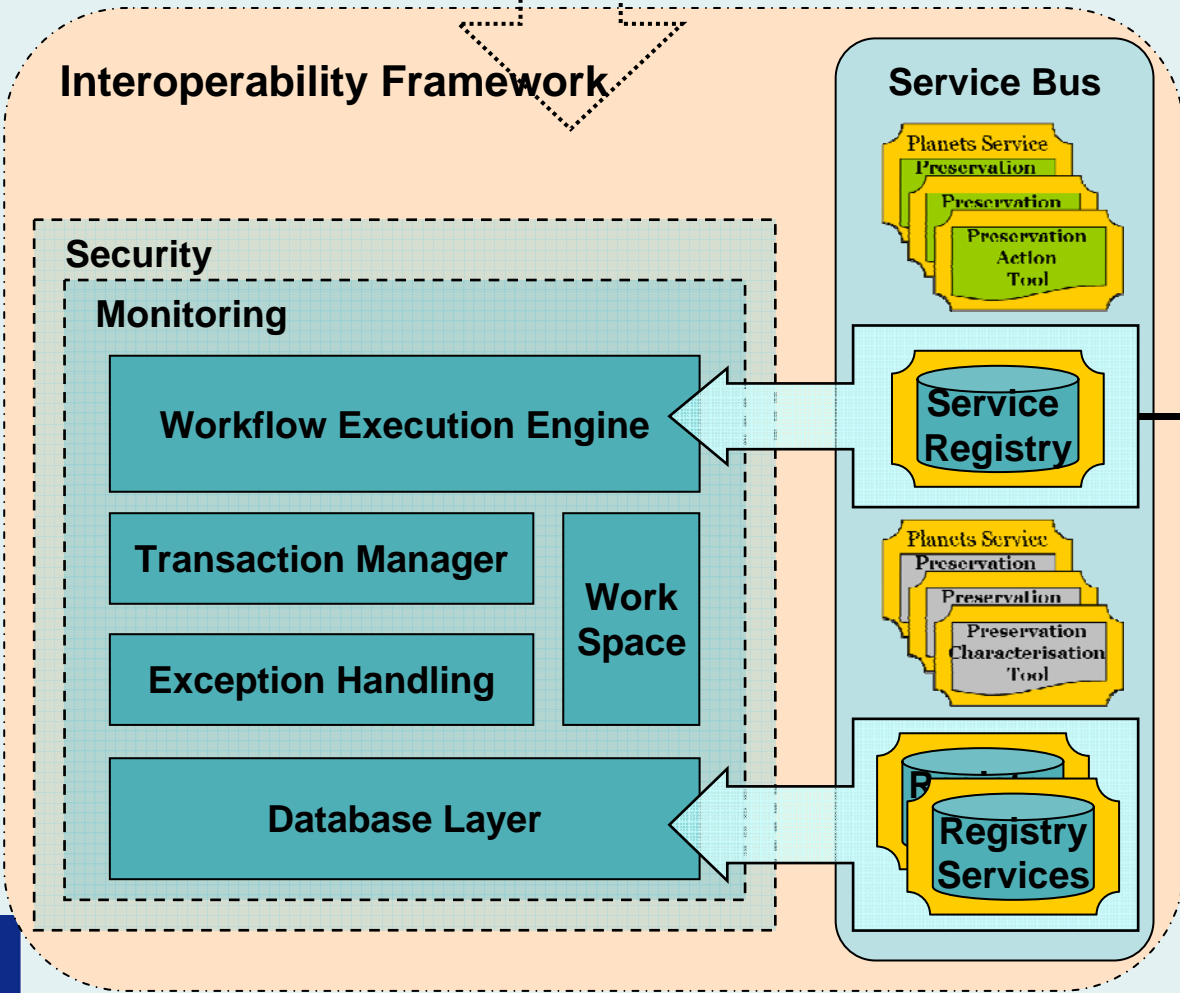
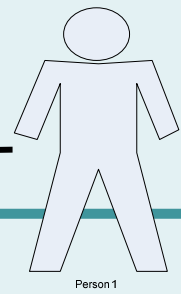
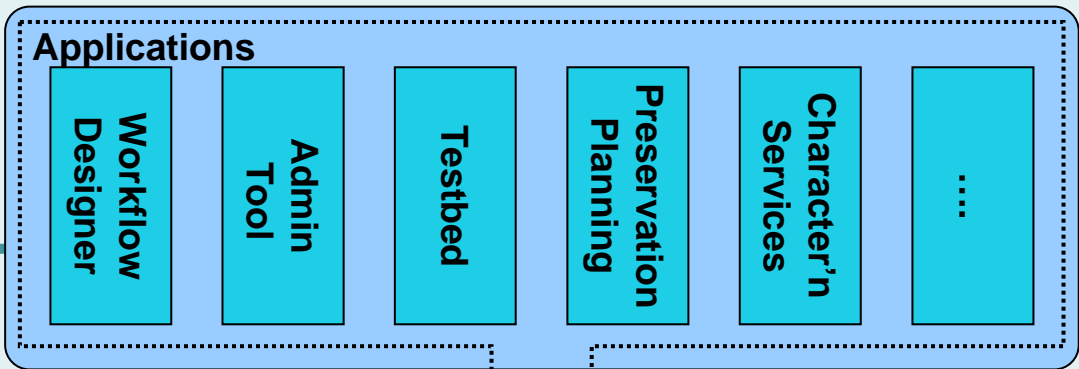


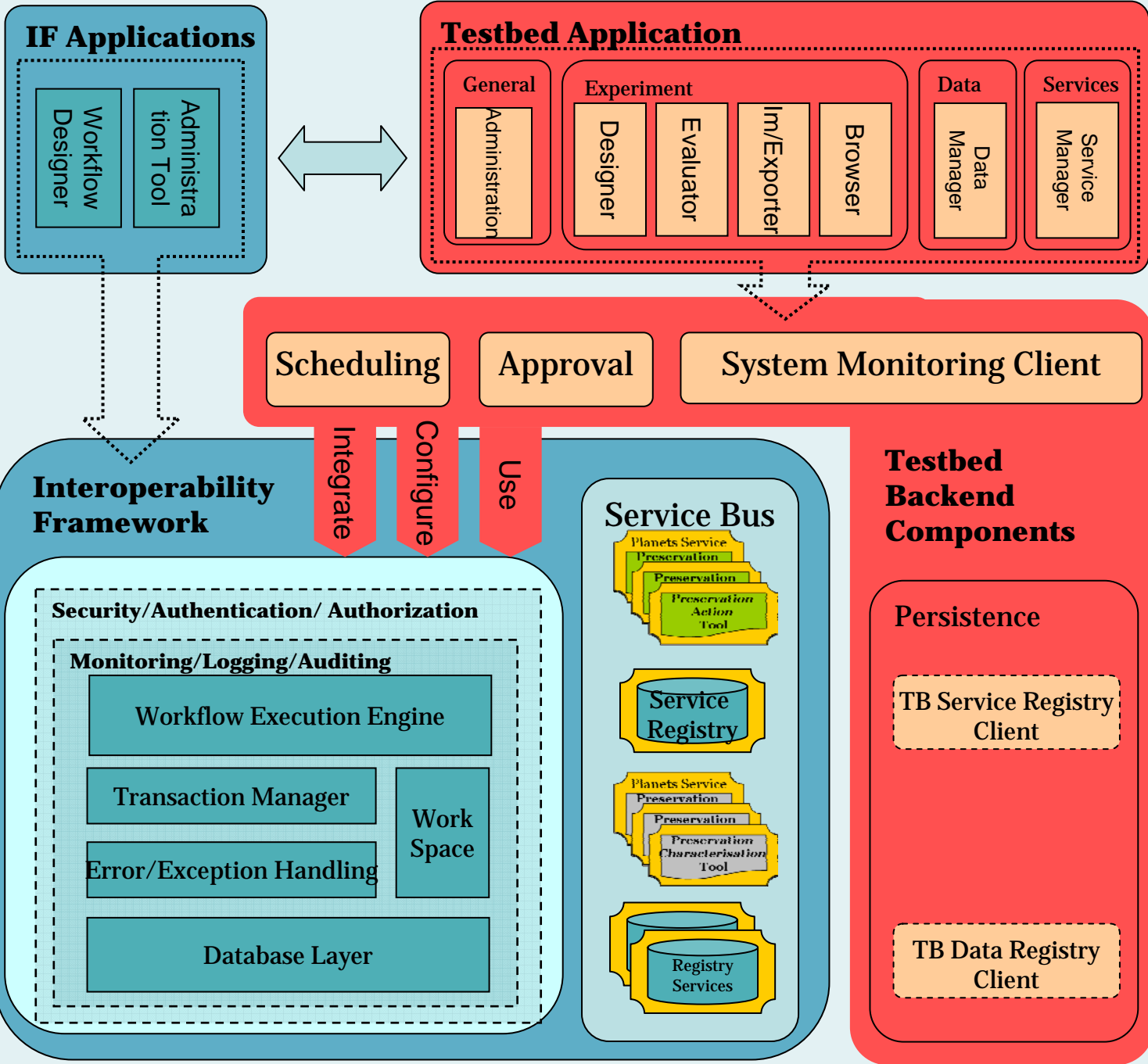
# Interoperability Framework

---

- ❑ Provides the glue to hold the Planets tools and services together
  - Provide service registries
    - Characterisation services
    - Preservation action services
  - Provide shared services
    - Security, authentication, authorisation,
    - Monitoring, logging, auditing
    - Intermediate data, repository, file system space
    - Execute and manage workflows
  - Enable third-parties to plug-in tools and services
  - Enable vendors to embed or provide preservation services







# Key technology choices

---

- Extensive use of XML and web services throughout
- Extensive use of enterprise quality components
- JSF (Java Server Faces) for user interfaces
- Workflows
  - BPEL – Business process execution language to describe experiments and plans
  - Eclipse BPEL workflow designer
- Repository and interfaces
  - JSR-170 Repository API
  - Jackrabbit to manage intermediate storage and data
  - Drivers for specific repository software
- JBoss application server



# Planets Software: Vision

---

- ❑ Integration of Planets results in a single downloadable package
- ❑ This package will be simple to
  - double-click and install
  - configure
  - administer
- ❑ When this package – a Planets instance – is deployed
  - an administrator can
    - create user accounts
    - deploy and browse services
    - browse registries
  - a preservation expert can
    - define service workflows (Workflow Design Tool)
    - define and evaluate preservation plans (Preservation Planning Application)
    - define and run experiments (Testbed)
  - a librarian or archivist can
    - define and test preservation plans
    - execute preservation processes on a repository (Online Design Tool)



# Status

---

- ❑ Testbed
  - Community Testbed Instance hosted by University of Glasgow (HATII)
  - Pilot release (Dec 07)
  - The Planets project partners (08)
    - Experiment with tools and services within Planets
    - Initial case studies
  - The digital preservation community (09)
    - Support institutional evaluation against benchmark content
    - Assist institutions to validate preservation plans against their policies & content profiles
  - Tool developers and vendors (10)
- ❑ Interoperability framework
  - Internal release (Oct 07)
    - Enables Planets application implementation
  - Integration release (May 08)
    - Click-and-install Planets software package



# Conclusion

---

- ❑ Planets methods, tools, and services will enable organisations to diagnose and treat problems with their digital objects
- ❑ High levels of automation and scalable components will reduce costs and improve quality
- ❑ Easy-to deploy software will enable organisations to implement the approach
- ❑ Pluggable service-oriented architecture supports extension
- ❑ Empirical data will enable improved decision making
- ❑ Find out more: <http://www.planets-project.eu>

