CeBIT 2008
4-9 March 2008, Hannover, Germany

Planets: Integrated Digital Preservation Services

Planets overview

- 4-year project co-funded by the European Union
- Under the same programme as CASPAR and DPE
- Started in June 2006 with €15m budget
- Coordinated by the British Library
- 16 partners including national libraries and archives, technology companies and research universities
- Focuses on the needs of libraries and archives

Aims and objectives

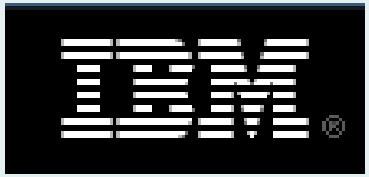
- Help ensure long-term access to cultural and scientific heritage
  - Improve decision-making
  - Ensure long-term access
  - Control costs
  - Ensure wide adoption across user community
  - Establish market place for preservation services and tools
- Build practical solutions
  - Integrate existing expertise, designs and tools
  - Deliver tools and services that can be used in an operational environment

Planets partners

- The British Library
- National Library, Netherlands
- Austrian National Library
- State and University Library, Denmark
- Royal Library, Denmark

- National Archives, UK
- Swiss Federal Archives
- National Archives, Netherlands

Planets partners



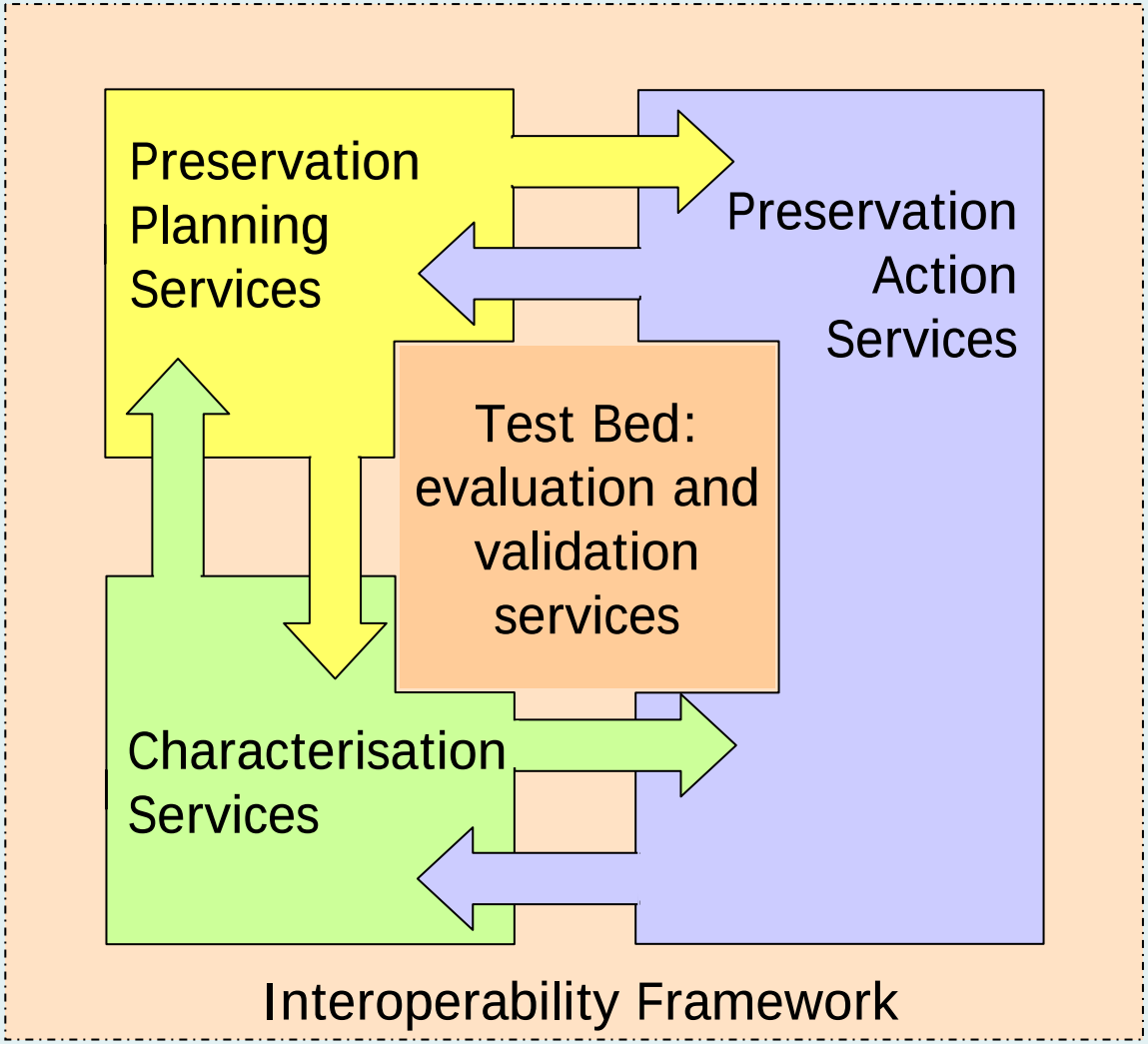
- Tessella Plc
- IBM Netherlands
- Microsoft Research
- Austrian Research Centers GmbH

- Hatii at University of Glasgow
- University of Freiburg
- Technical University of Vienna
- University of Cologne

The Planets team

All Staff Meeting, Feb 2007

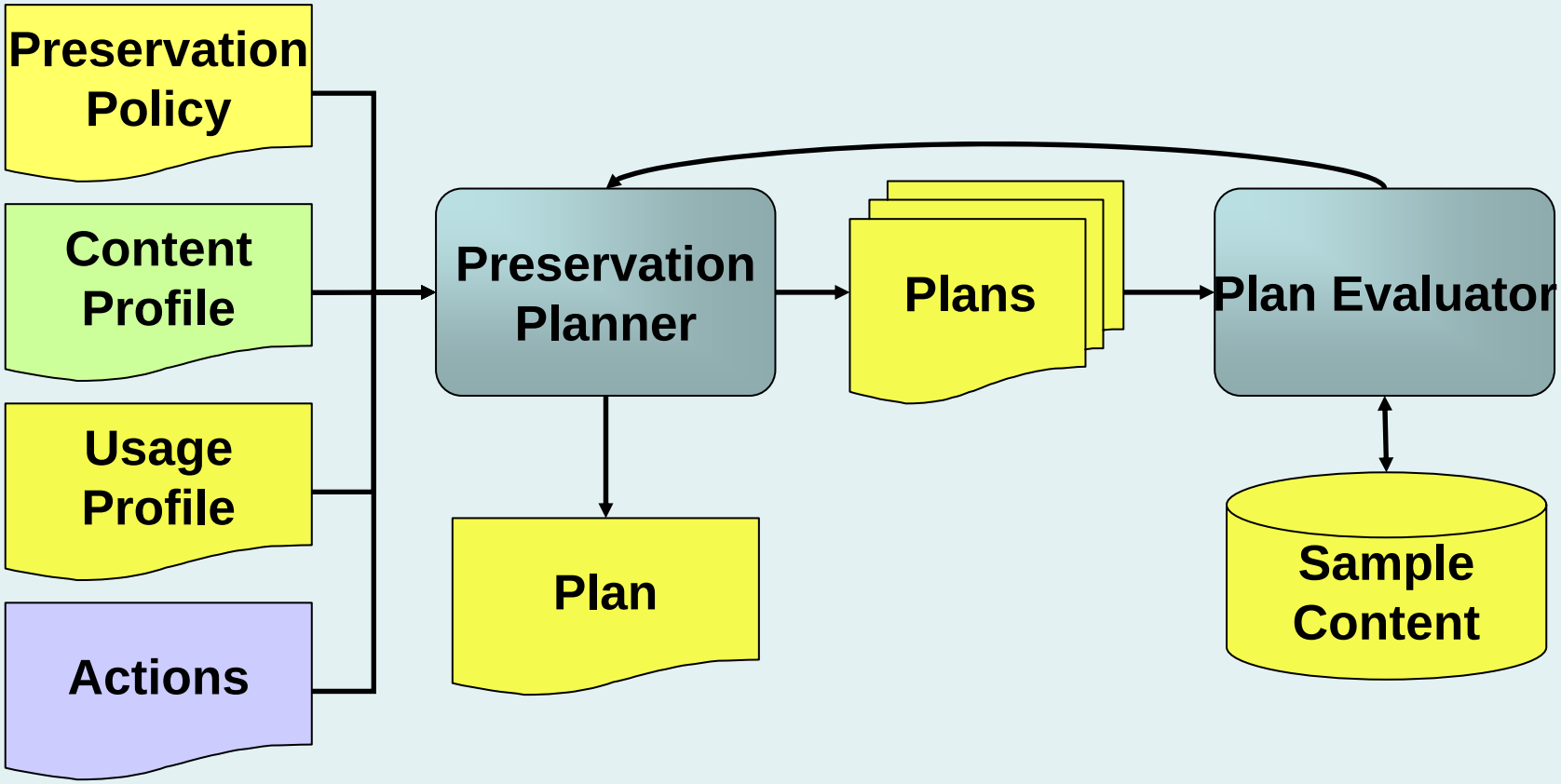
Planets architecture



Planets architecture: key components

- **Preservation planning**
  - tools and services for formulating and selecting preservation plans
- **Preservation characterisation**
  - tools and services for automatic analysis of digital objects' technical and intellectual characteristics
  - supporting registry of characterisation information
- **Preservation action**
  - methodology for describing preservation action tools
  - supporting registry
  - migration and emulation tools
- **Testbed**
  - hardware and software environment for comparing digital preservation tools and assessing their effectiveness
- **Interoperability framework**
  - service-oriented architecture
  - provides shared functions
  - integrates the Planets tools and services

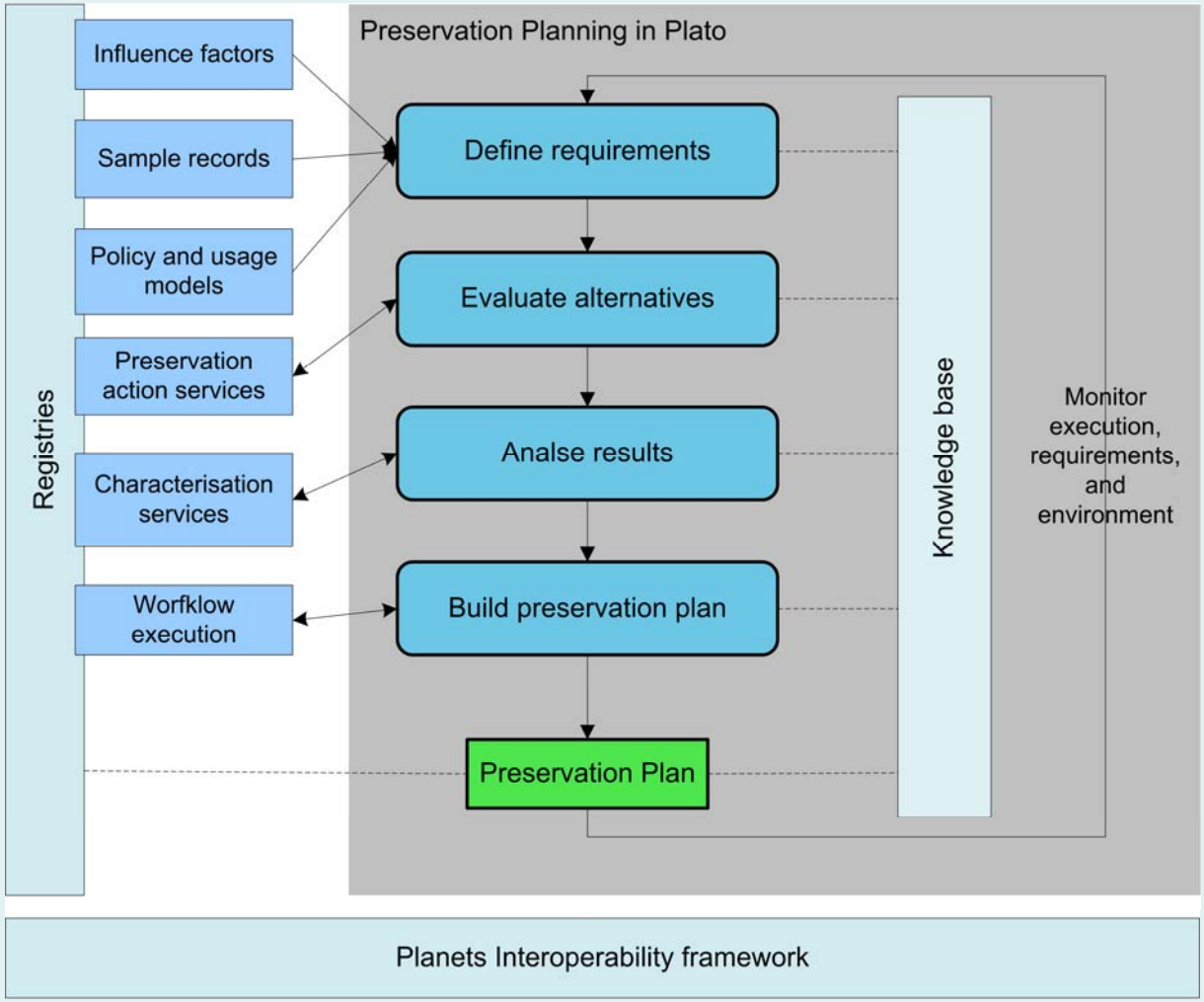
Preservation planning



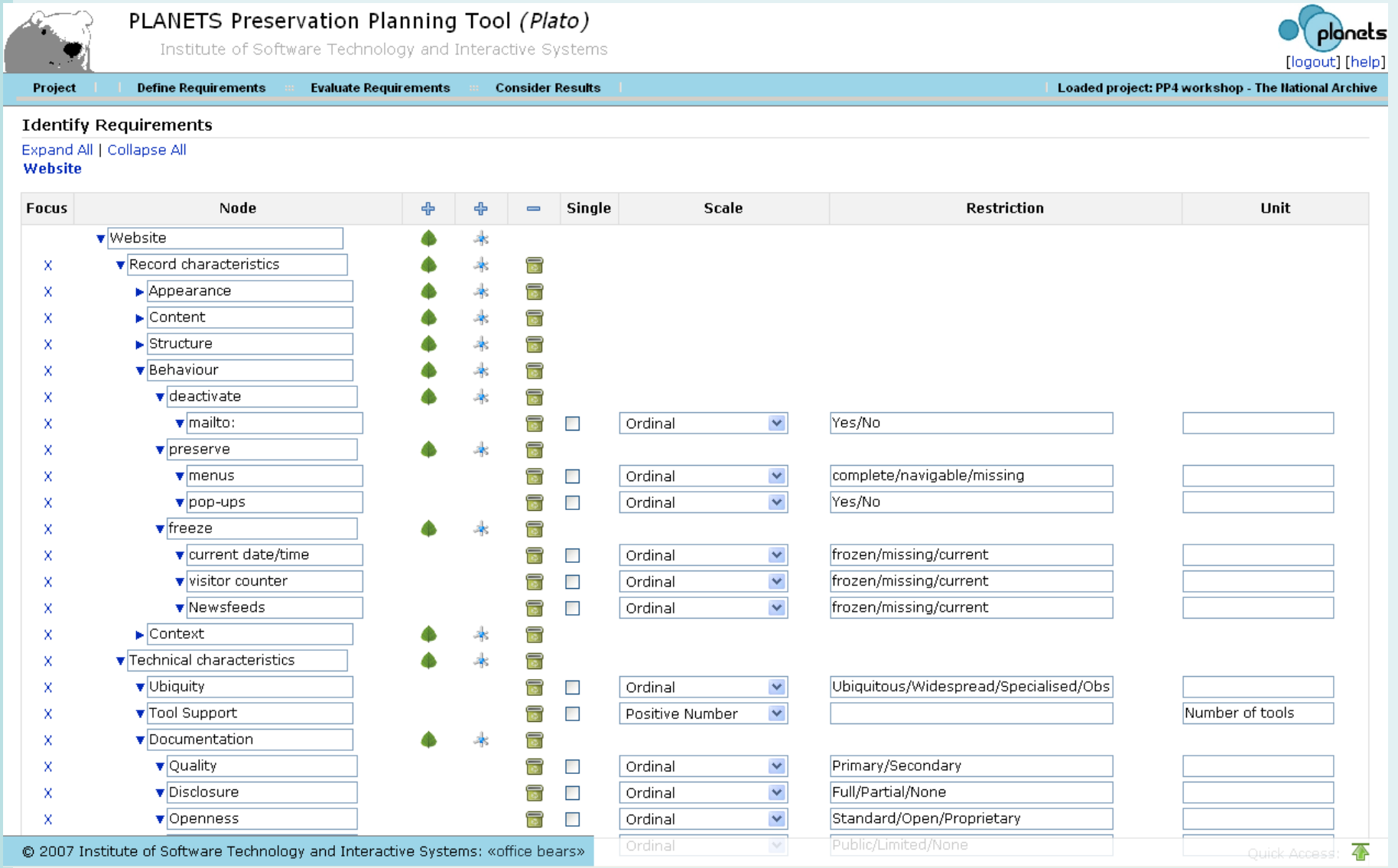
Planning for a desired future

- Parameters for successful preservation strategy
  - Organisational policy
  - Usage pattern of digital collection
- Methodology and workflow for identifying preservation requirements, defining potential solutions and selecting the most appropriate
  - Plato, workflow-based decision support tool
- Validation framework for measuring and comparing characteristics (in a typical before and after scenario)
- Proactive preservation planning including collection profiling, technology watch and advice services
- Implementing the OAIS Preservation Planning function

Preservation Planning using Plato



Planets Preservation Planning Tool (Plato)



PLANETS Preservation Planning Tool *(Plato)*

Analyse Results

Sum

☑	PDF/A (Tool A)
☑	PDF/A (Tool B)

Show

Expand All | Collapse All

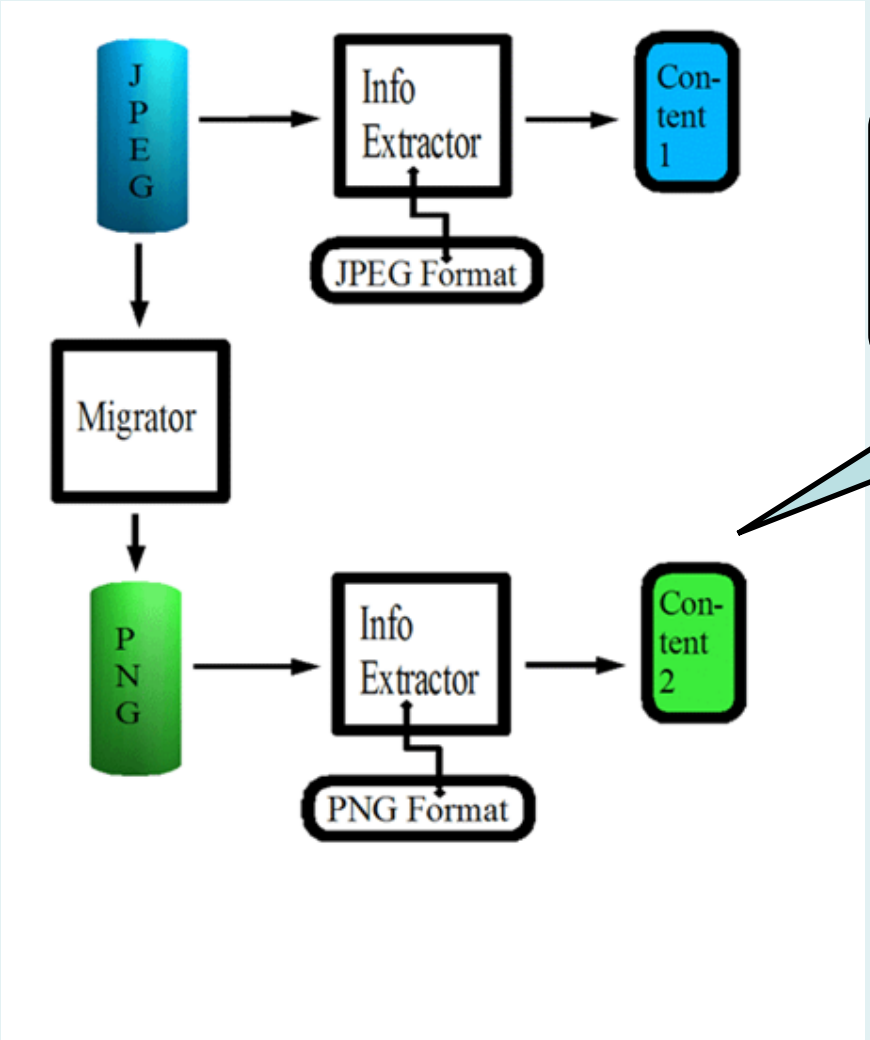
Minimalist root node

Focus	Name	Result
	▼Minimalist root node	PDF/A (Tool A): 2,98 PDF/A (Tool B): 3,19
X	▶Image properties	PDF/A (Tool A): 0,70 PDF/A (Tool B): 0,80
X	▼Karma	PDF/A (Tool A): 0,40 PDF/A (Tool B): 0,00
X	▼Filesize (in Relation to Original)	PDF/A (Tool A): 0,78 PDF/A (Tool B): 0,99
X	▼A Single-Leaf	PDF/A (Tool A): 0,40 PDF/A (Tool B): 0,80
X	▼IntRange 0-10	PDF/A (Tool A): 0,70 PDF/A (Tool B): 0,60

Content characterisation

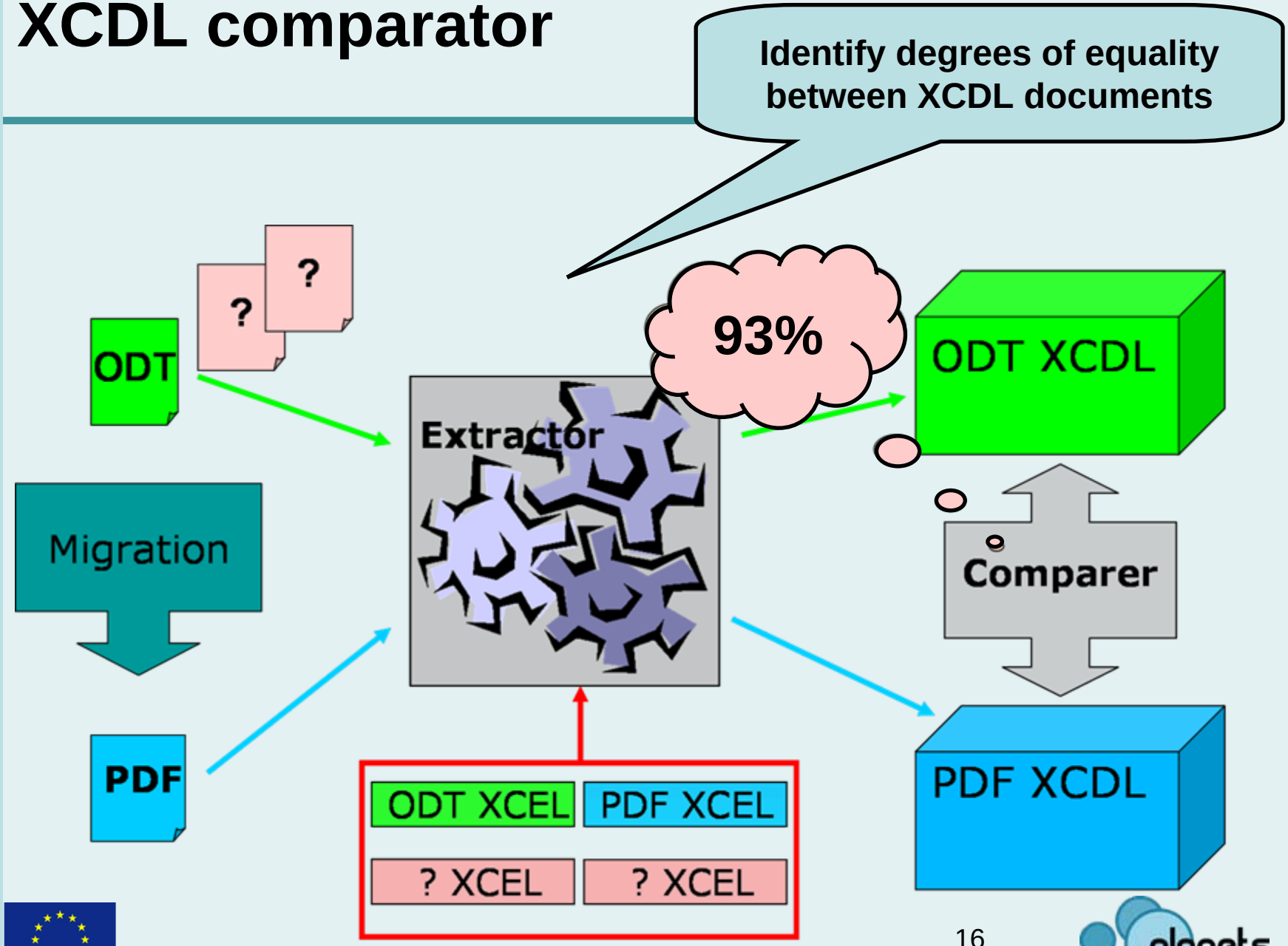
- Automating the process of identifying, describing and extracting the essential characteristics of digital objects
- Supporting registry with technical information about significant properties of digital object types
- Generic XML language supporting automatic validation of document conversions and evaluation of migration quality
  - *eXtensible characterisation definition language (XCDL)* and *eXtensible characterisation extraction language (XCEL)*
  - Hierarchically decompose and represent documents in an abstract XML language
- Extractor and Comparator software

XCEL extractor



XCEL processing software which extracts content of a file and express it in XCDL

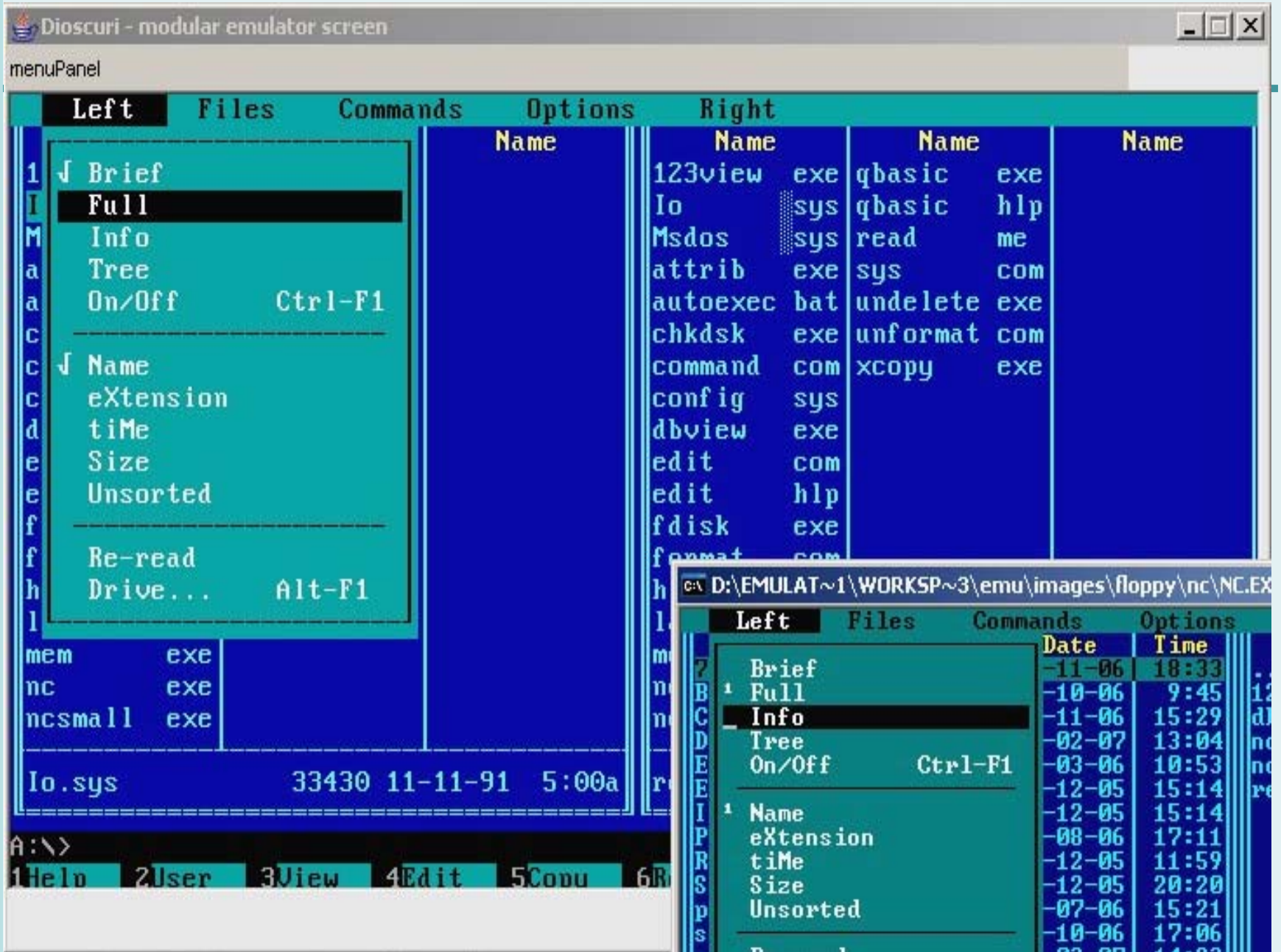
XCDL comparator



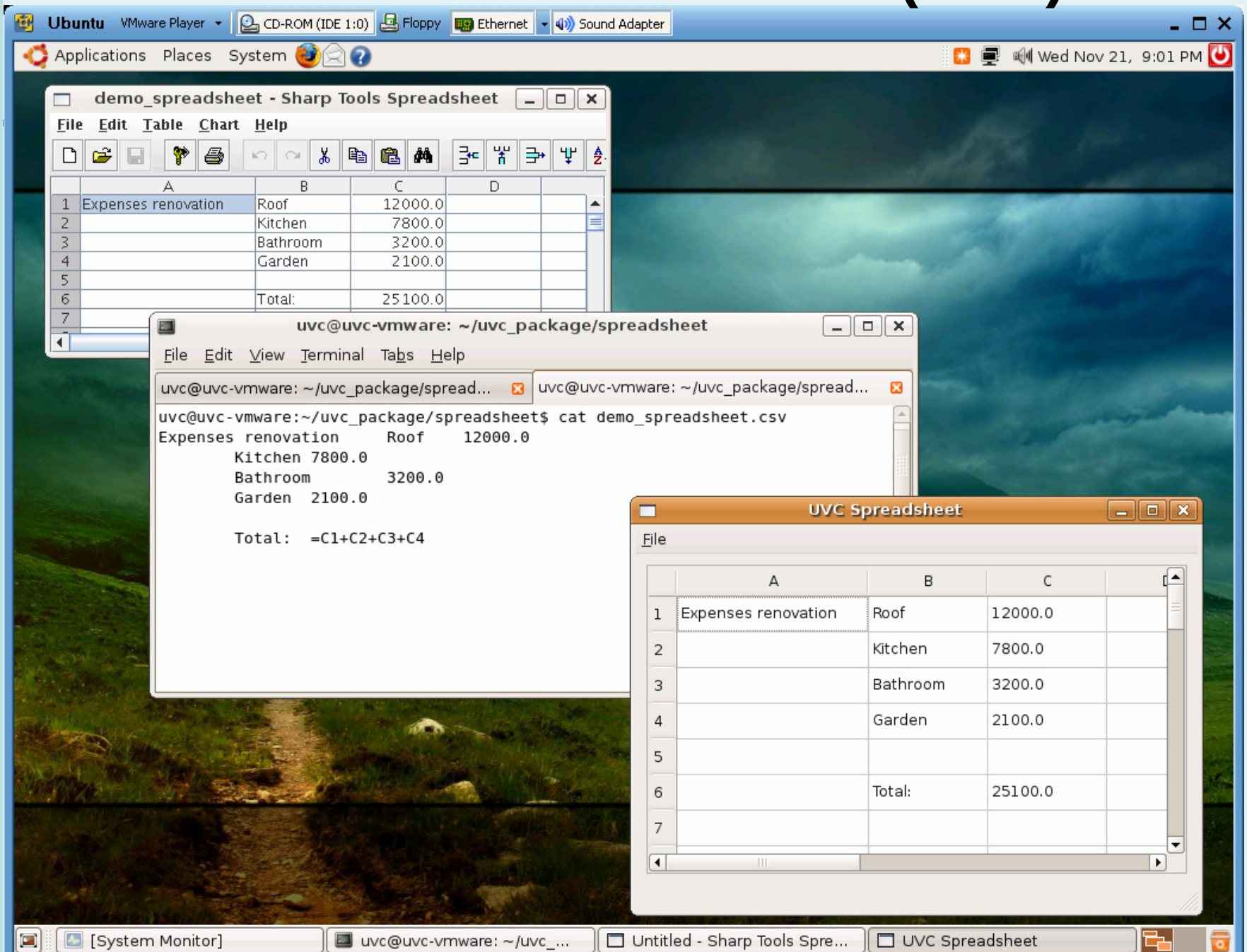
Preservation action

- Inventory of commonly used file formats
- Gap analysis of current tools
- Formal language for describing preservation action tools
- Developing the most wanted tools for digital objects
- Supporting registry with technical descriptions of tools
- Migration tools to convert digital objects into newer formats
- Emulation tools enabling interaction with digital objects using original software applications
  - Dioscuri – the modular emulator
  - Universal Virtual Computer (UVC)

Dioscuri – the modular emulator



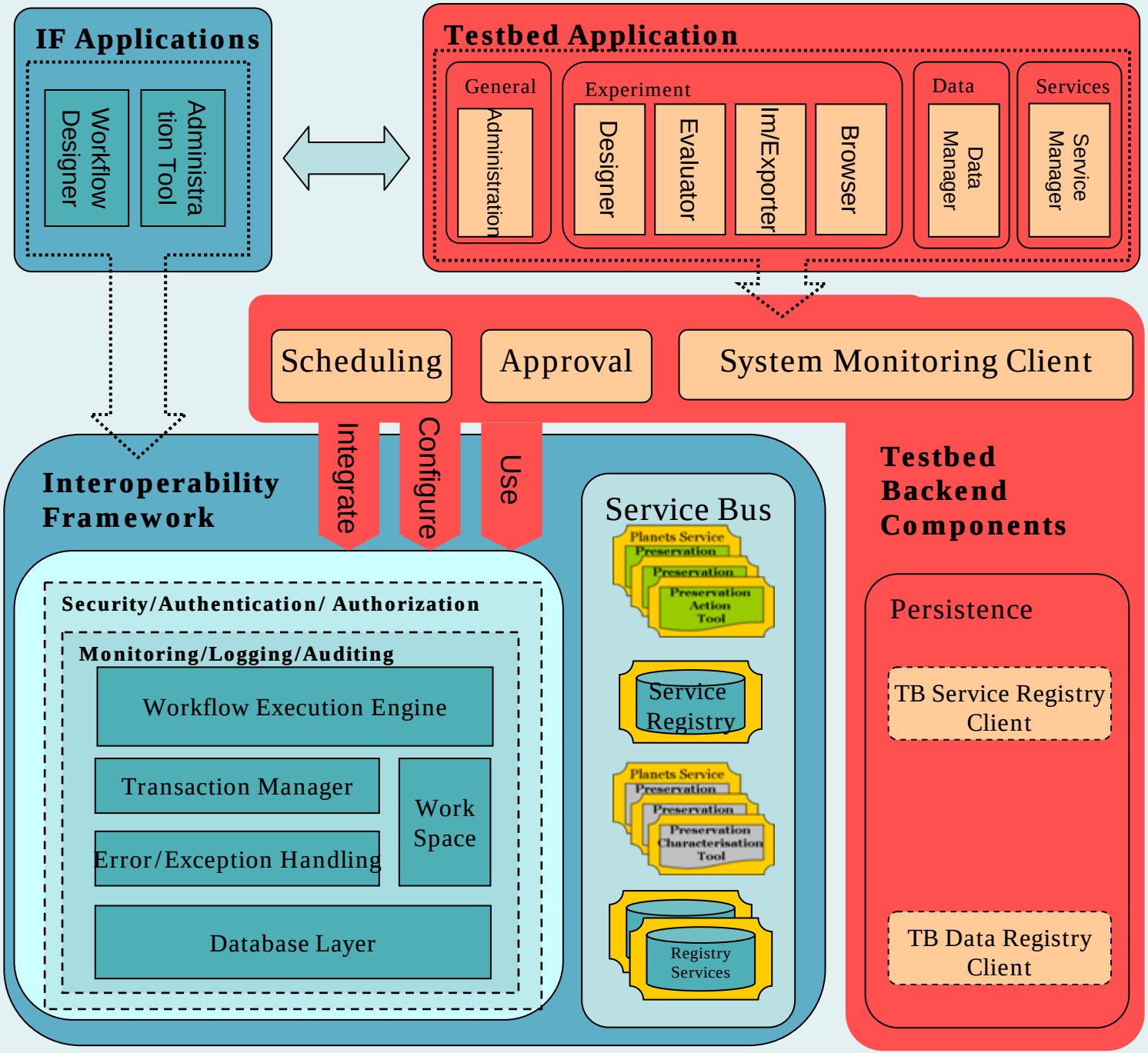
The Universal Virtual Machine (UVC)



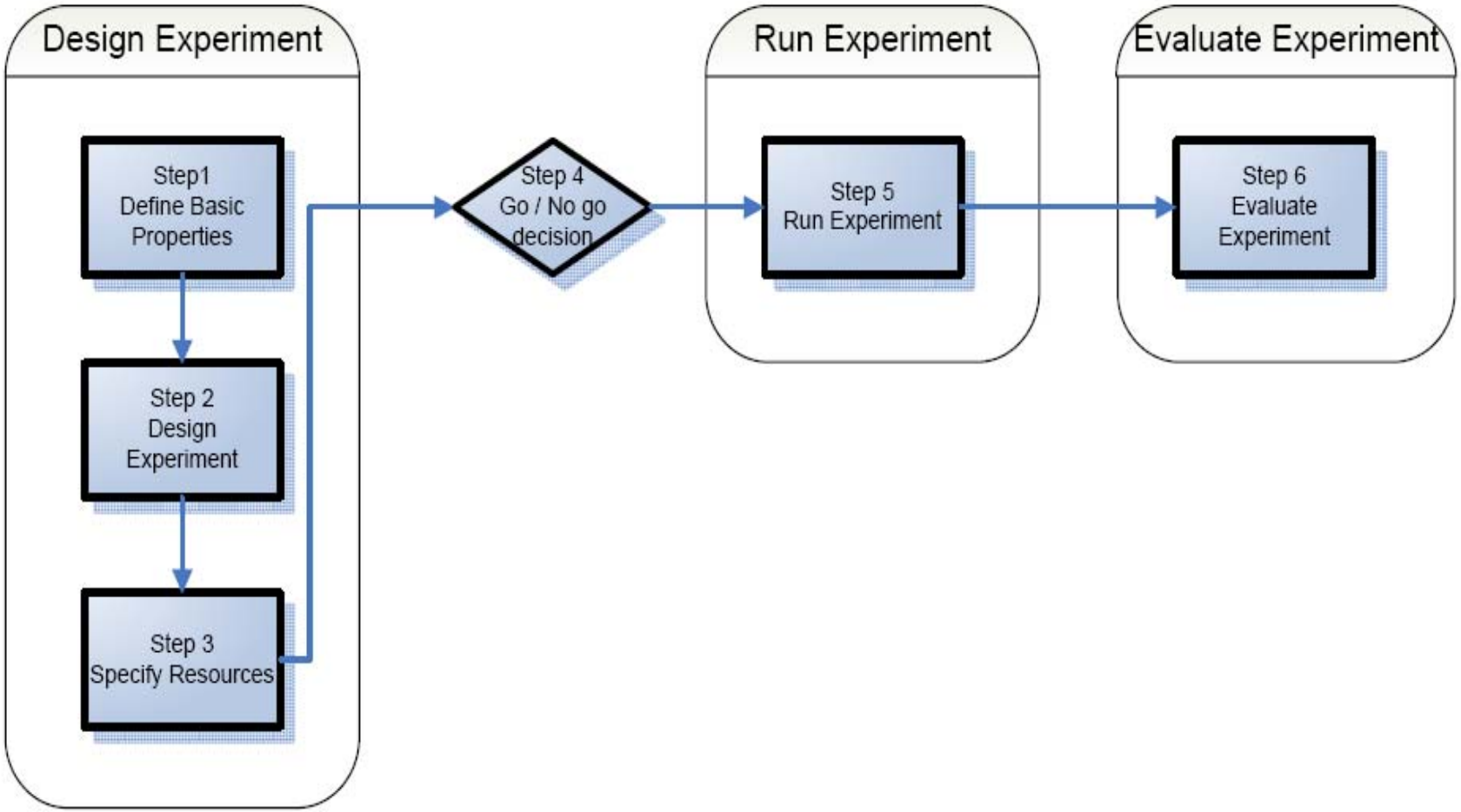
Planets Testbed: learn what works

- “A controlled environment with metrics and benchmark content that allows the objective comparison of preservation tools and strategies through experimentation”
- A dedicated scientific research environment
- Consists of
  - Data storage, hardware, Planets software, Testbed application
  - Benchmark and other content
- Use well known and widely spread technologies
- Available for external organisations in early 2009

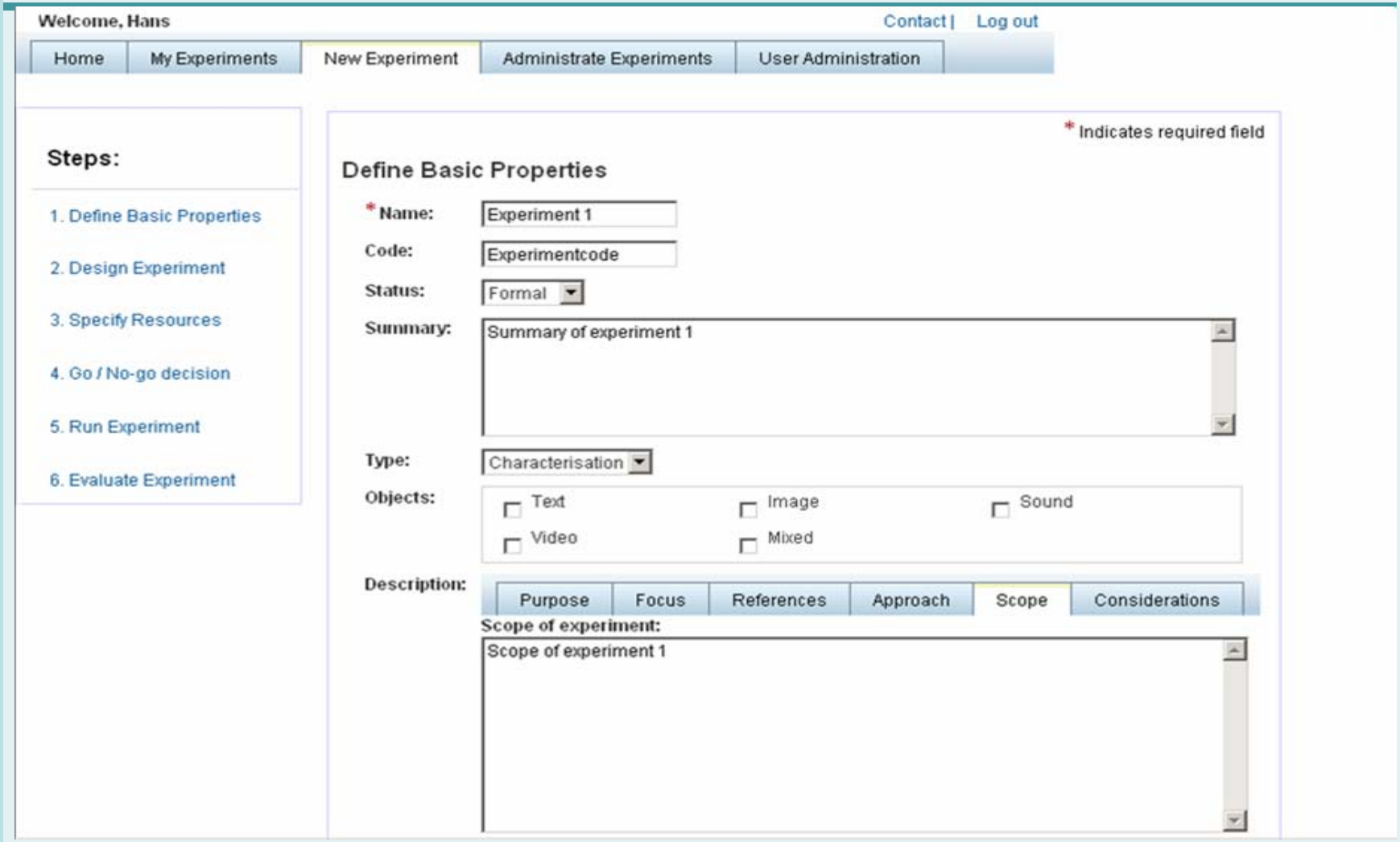
Planets Testbed



Testbed experiment process



Prototype Testbed



Planets Software: Vision

- A single downloadable software package
- Simple to install, configure and administer
- When deployed, a **Planets instance** is created, in which
  - 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)

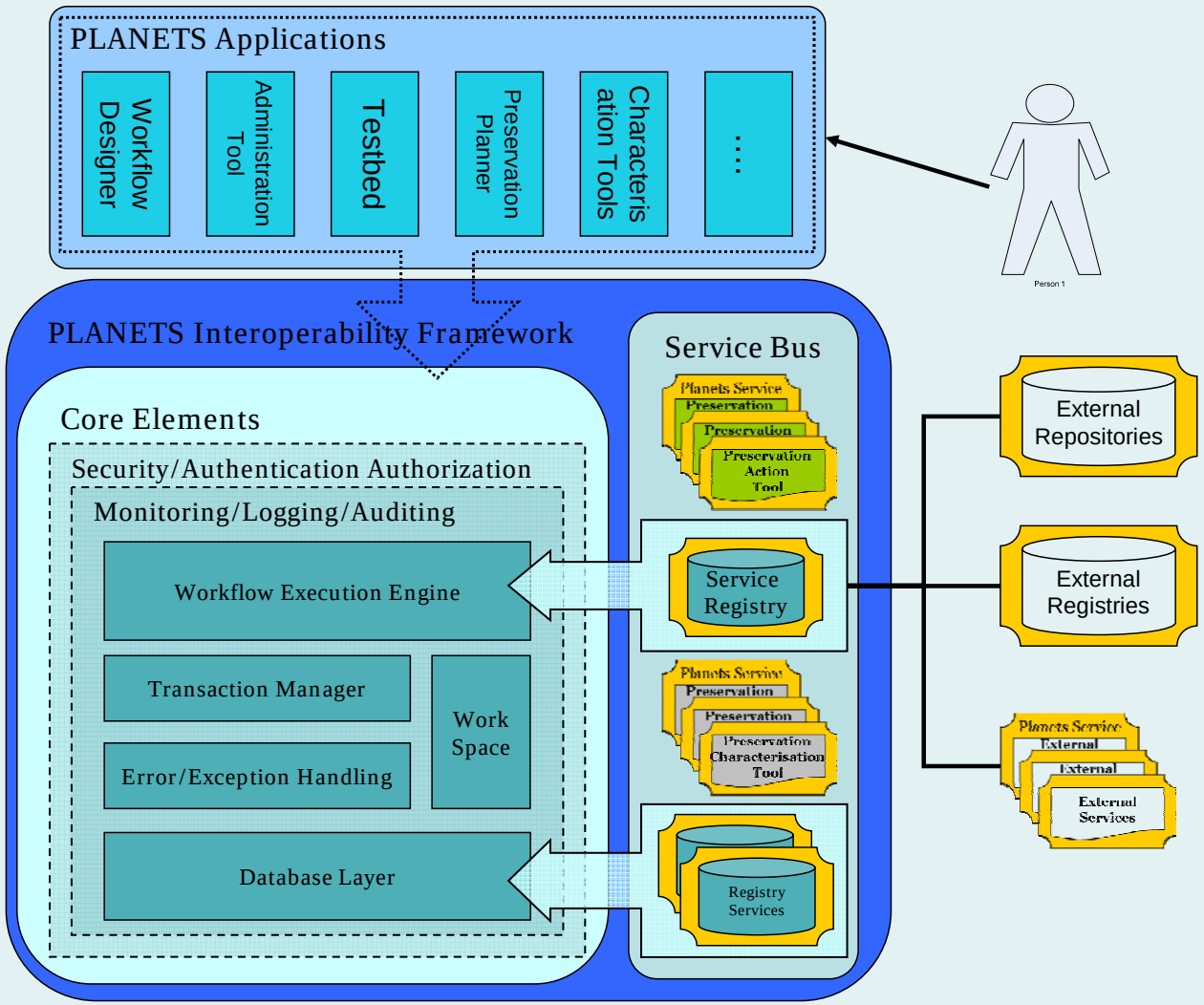
Planets Software: motivation

- Provide commonly required functions
  - Data persistence
  - User management
  - Authentication and Authorisation
  - Monitoring, Logging, and Messaging
- Meet non-functional requirements on the infrastructure
  - Robust
  - Scalable
  - Distributed

Planets software: benefits

- **Efficiency**
  - Components can be shared and only need to be developed once
  - Optimise the number of components: IF provides a single database for all components, so only one database need be installed
- **Interoperability**
  - Common components help ensure interoperability between various applications
  - Support access to remote and distributed 3rd party characterisation and migration services by enforcing Web Service standards

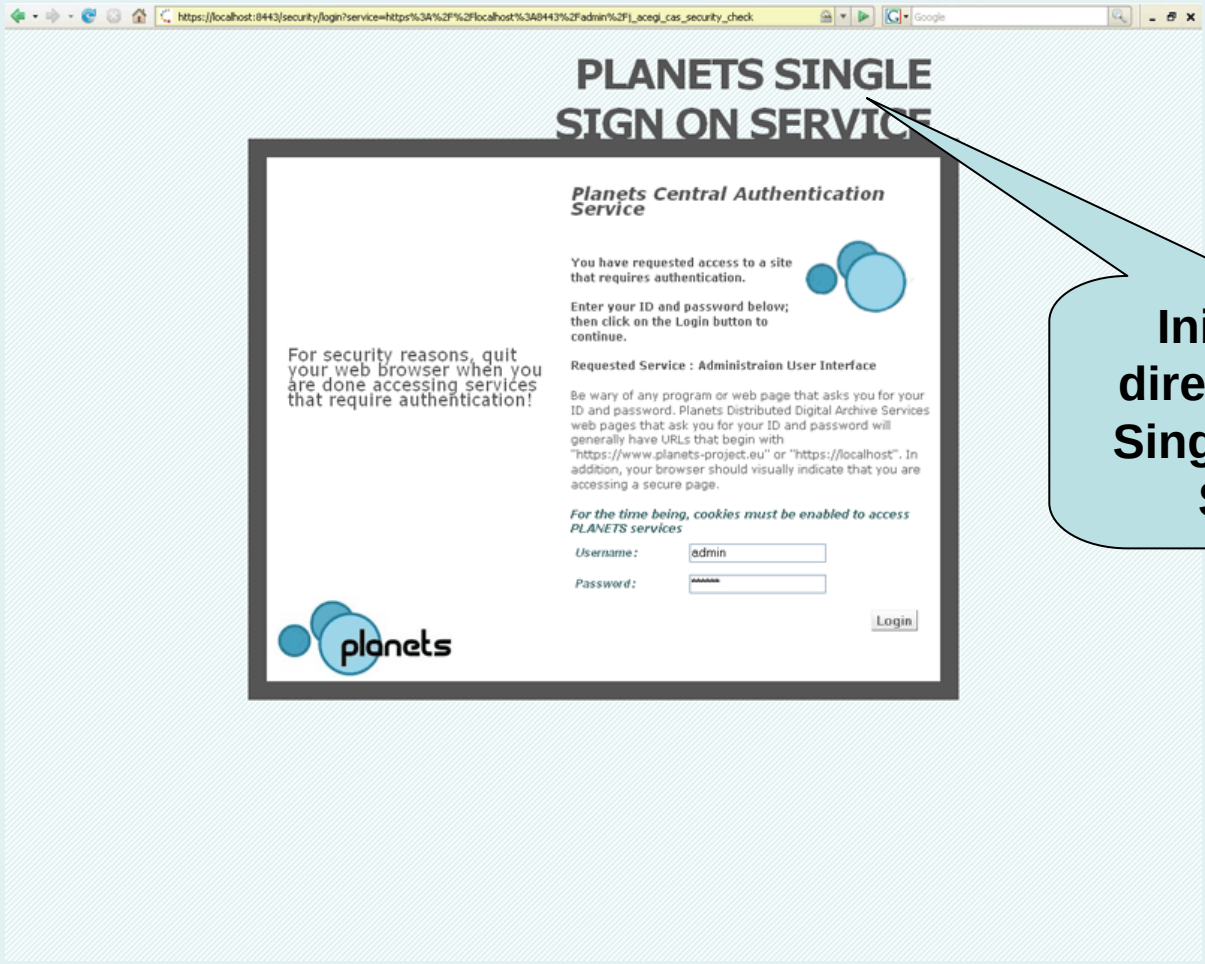
Planets software: architecture



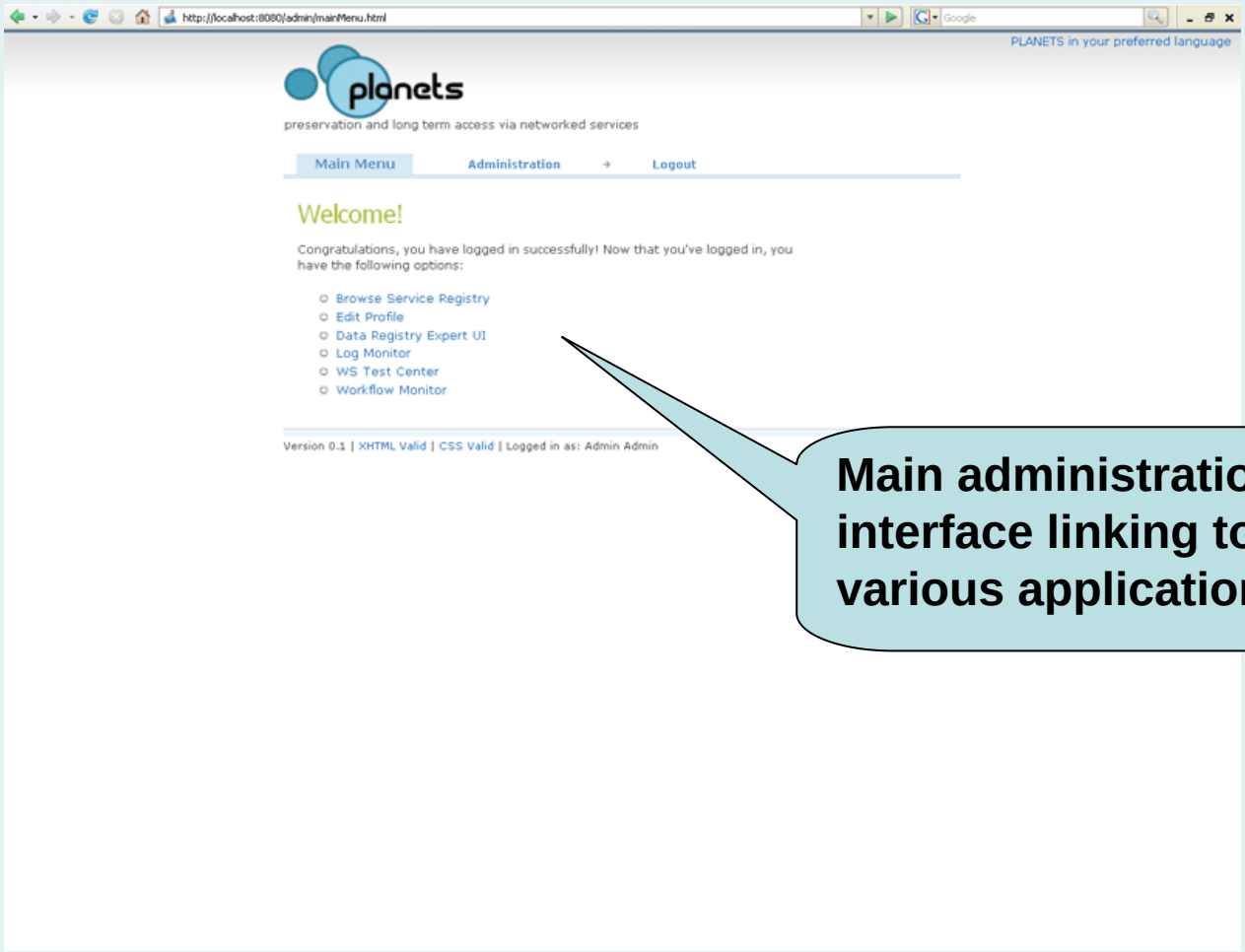
Interoperability Framework components

- Application server (JBoss)
- Relational Database (Apache Derby)
- Service Registry (jUDDI)
- Data Registry (JCR implementation on Jackrabbit)
- Workflow Design Tool
  - Workflow expressed in BPEL
  - Expert design tool based on the Eclipse BPEL Plugin

Single sign-on service



Administration interface

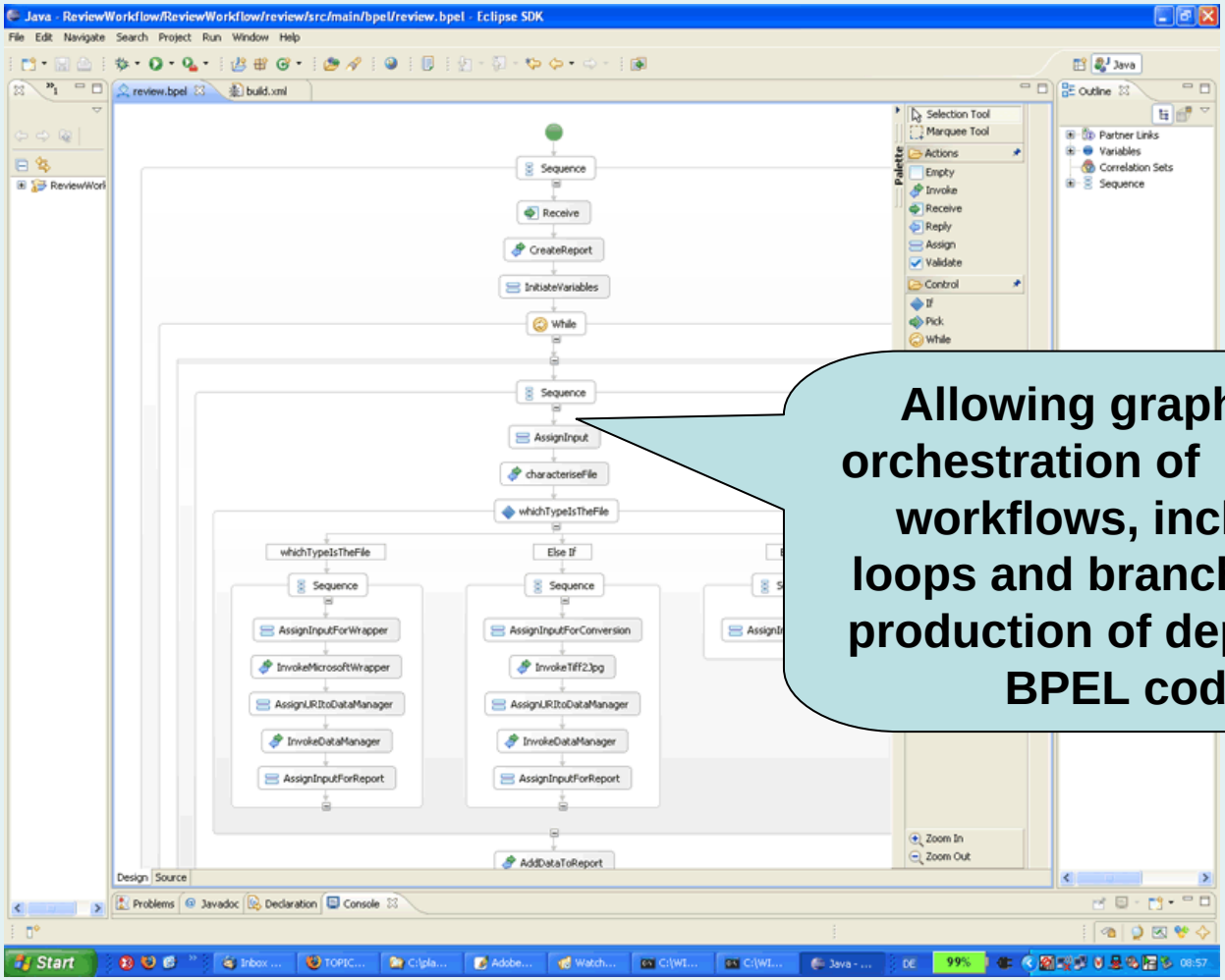


Planets Web Service registry

Admin interface linking to Planets Web Service registry, where services can be manually registered and searched

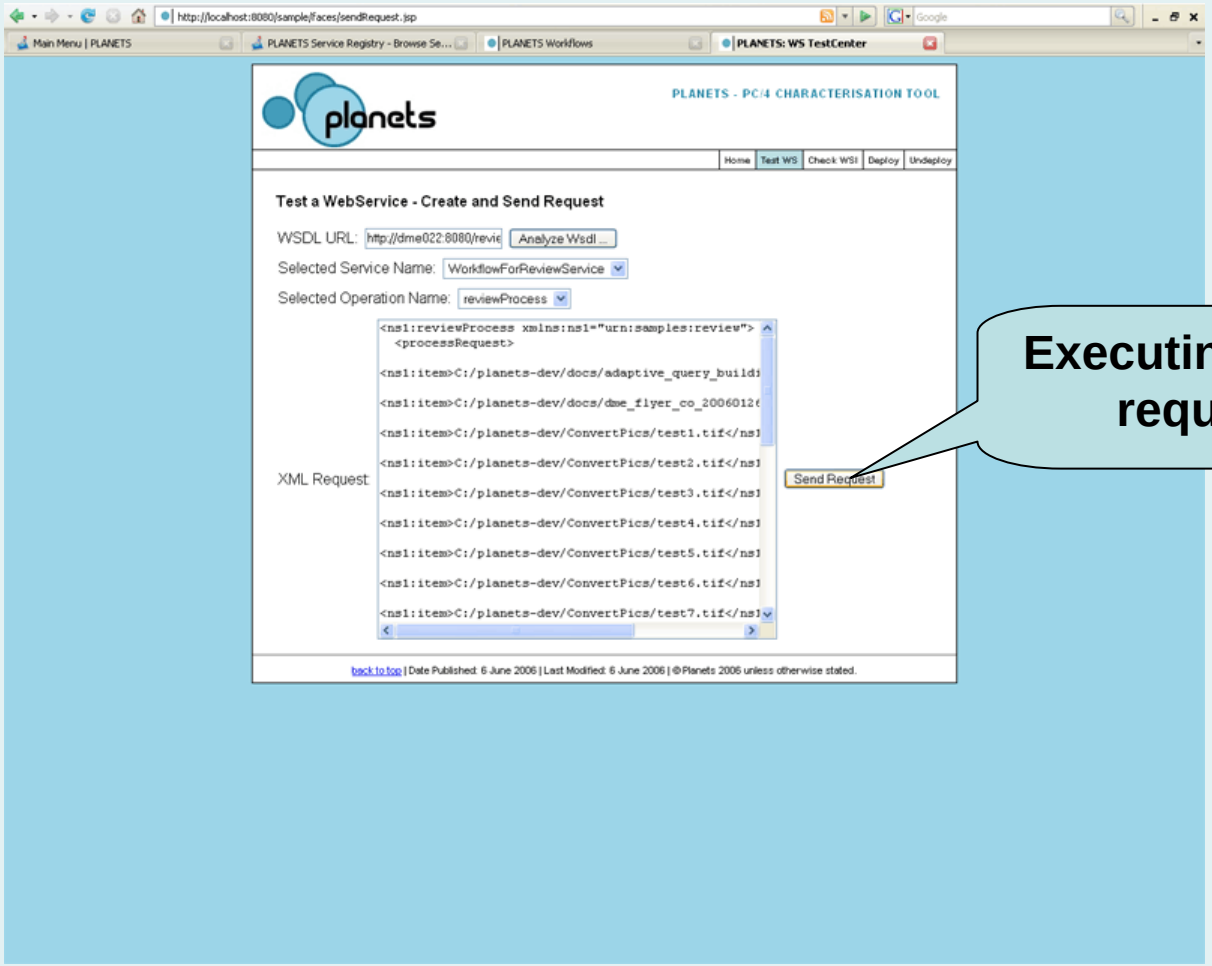
Workflow design tool

Workflow design tool

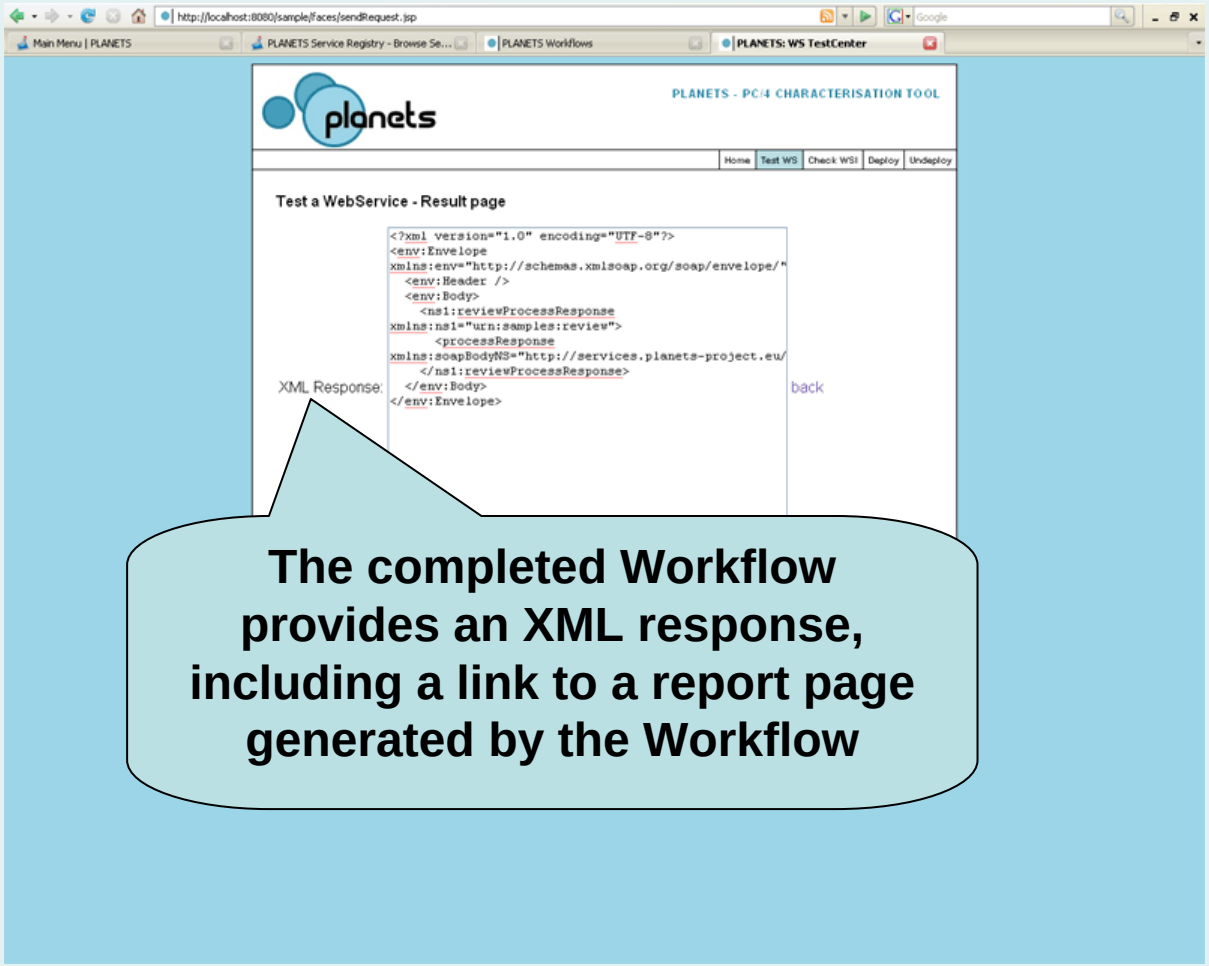


Allowing graphically orchestration of complex workflows, including loops and branches, and production of deployable BPEL code

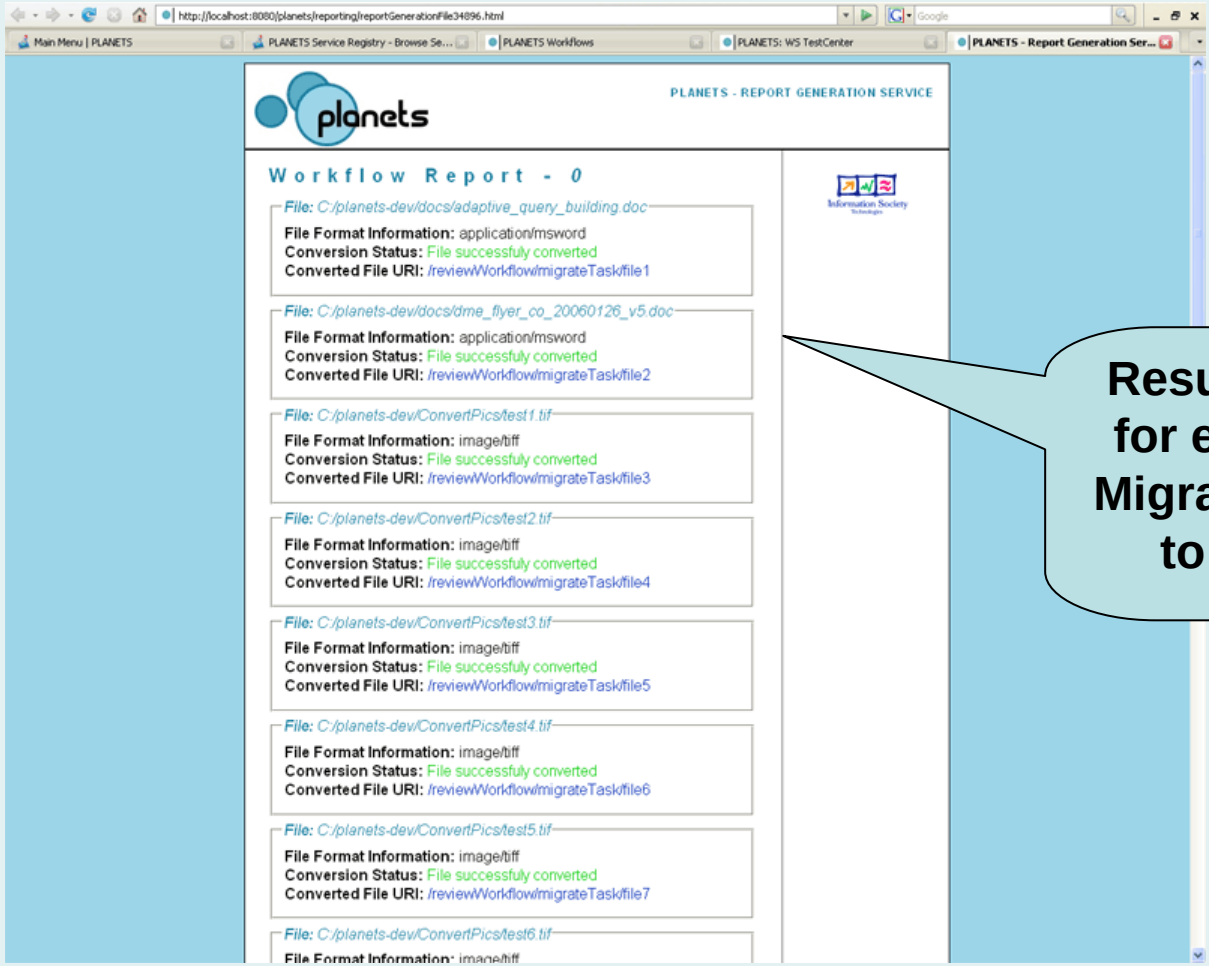
Workflow execution



Workflow execution

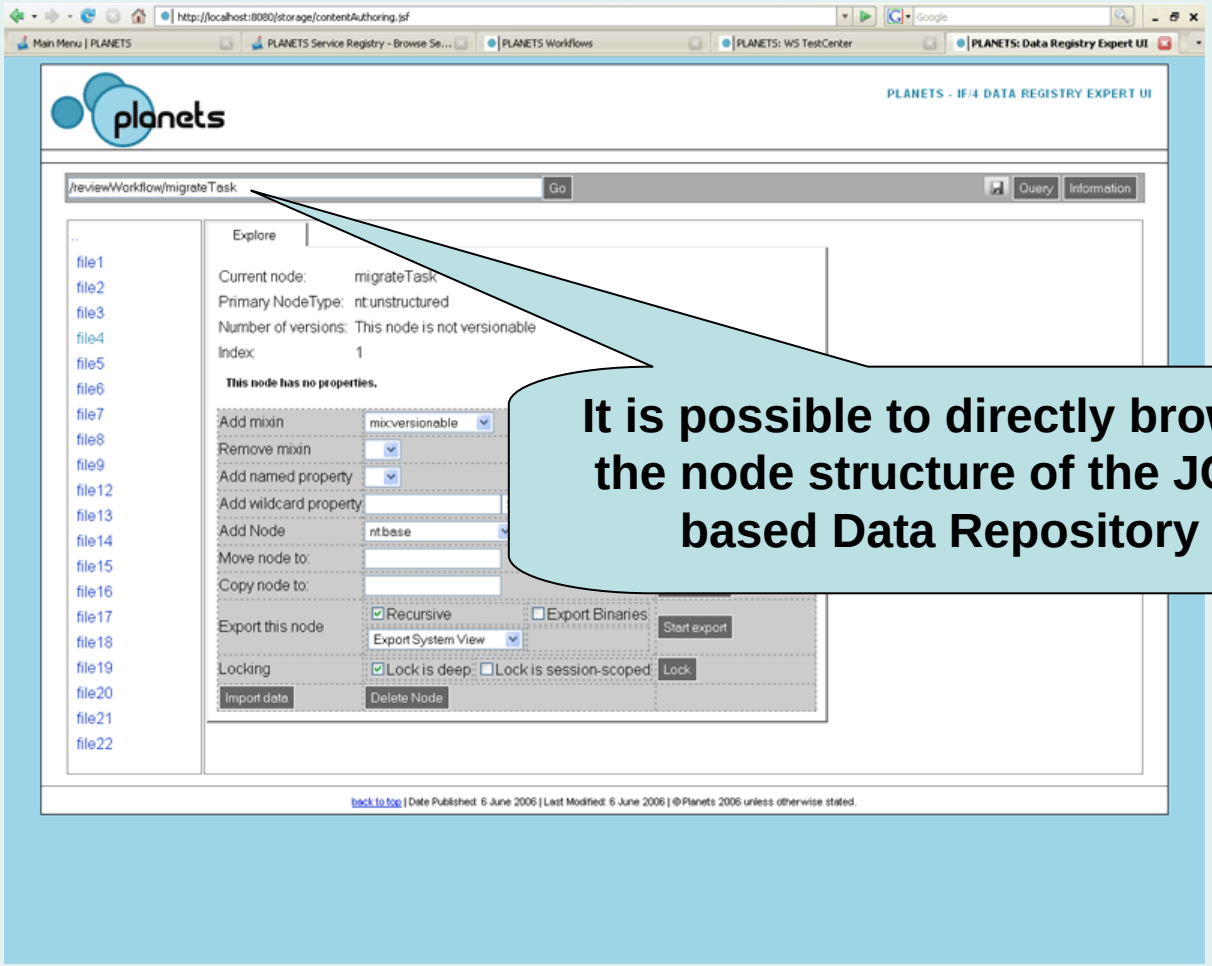


Workflow execution

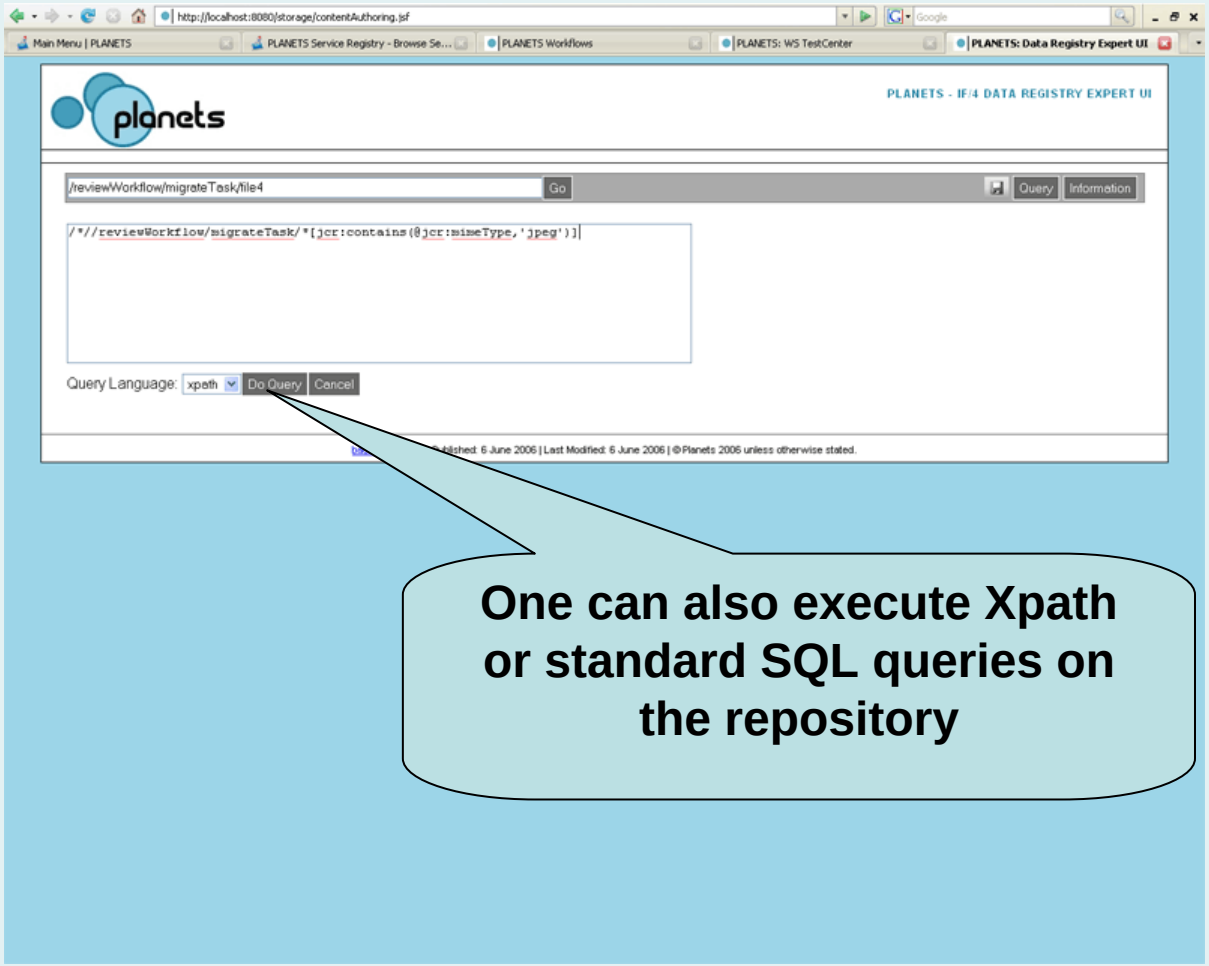


Results of the workflow for each submitted file. Migrated files are written to the data registry

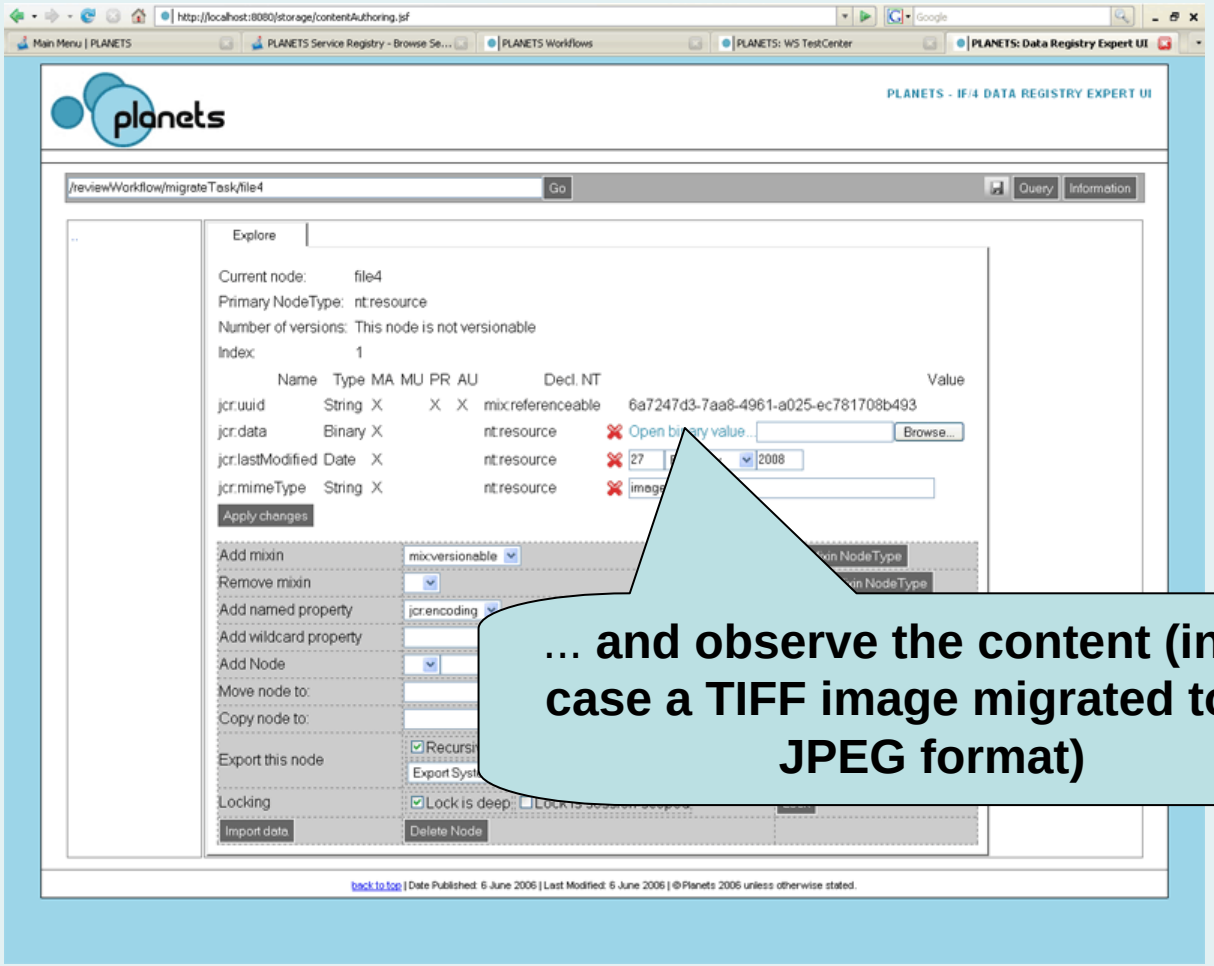
Data registry



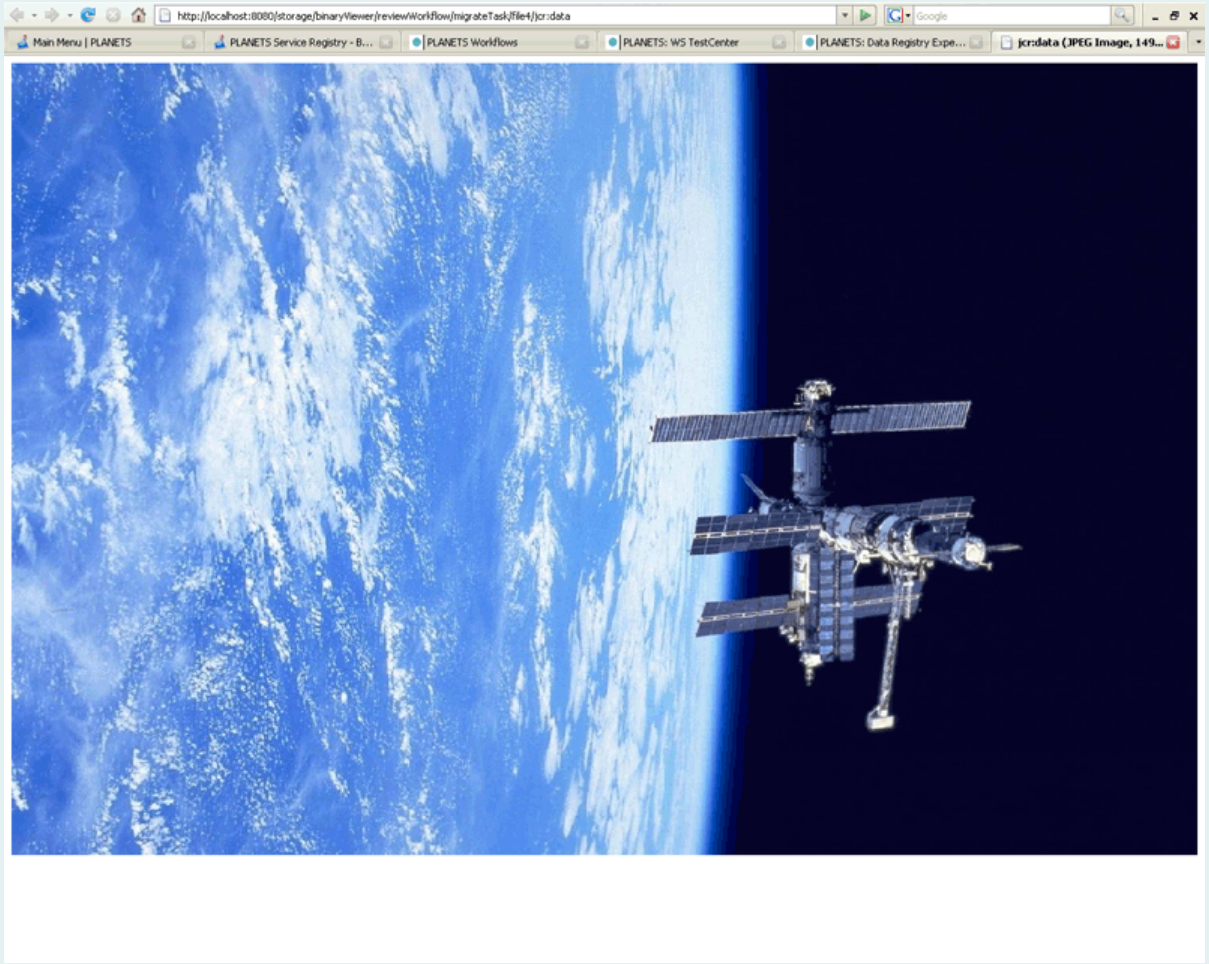
Data registry



Data registry



Data registry



Conclusion

- Planets will help organisations diagnose and treat problems with their digital collections
- High levels of automation and scalable components will reduce costs and improve quality
- Empirical data will enable improved decision making

Find out more about Planets

- **Contact us**: info@planets-project.eu
- **Website**: www.planets-project.eu
- **Planets news**: http://www.planets-project.eu/news/
- **Workshops, conferences and training events**: http://www.planets-project.eu/events/
- **Quarterly newsletters**: http://www.planets-project.eu/publications/?search[0]=10
- **Scientific publications**: http://www.planets-project.eu/publications/?search[0]=12