



PLANETS Data Models

Mark Evans, Tessella

mark.evans@tessella.com

<http://www.tessella.com>

Contents

- ❑ Planets Overview
- ❑ Why a data model?
- ❑ Conceptual data model
- ❑ The future

Planets Overview

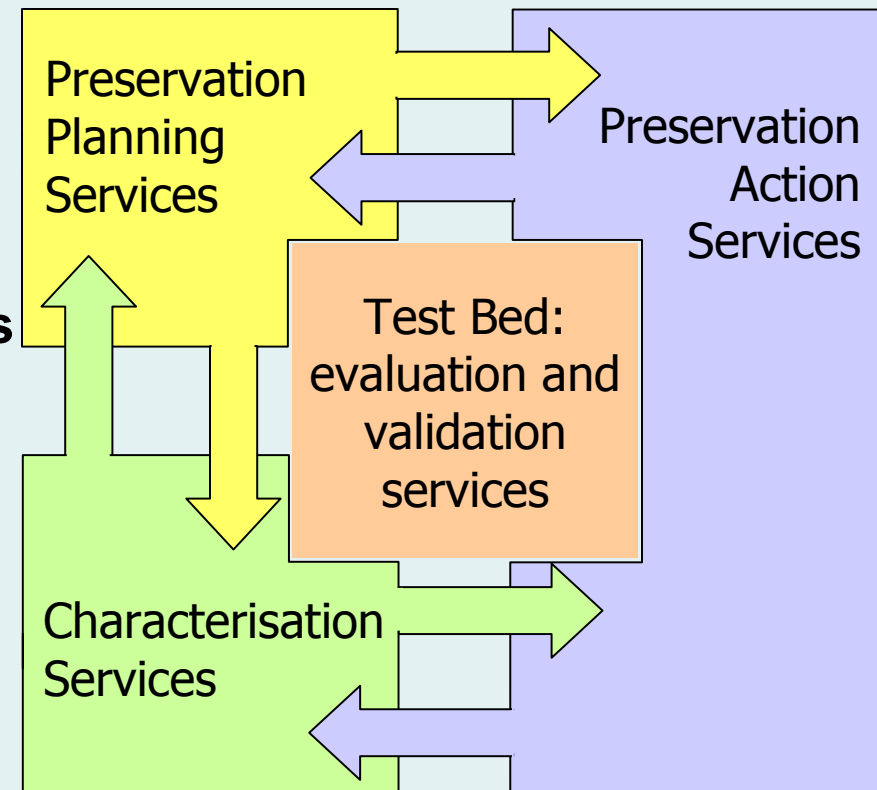
- ❑ PLANETS - Preservation and Long-term Access through NETworked Services,
- ❑ Four-year project co-funded by the European Union under the Sixth Framework Program
- ❑ Primary Goal to build practical services and tools to help ensure long-term access to our digital cultural and scientific assets
- ❑ Started on 1st June 2006.
- ❑ Other digital preservation projects funded under the FP6 Program include CASPAR and DPE.

Planets Overview - Partners

- ❑ The British Library
- ❑ The National Library of the Netherlands
- ❑ Austrian National Library
- ❑ The Royal Library of Denmark
- ❑ State and University Library, Denmark
- ❑ The National Archives of the Netherlands
- ❑ The National Archives of England, Wales and the United Kingdom
- ❑ Swiss Federal Archives
- ❑ University of Cologne
- ❑ University of Freiburg
- ❑ HATII at the University of Glasgow
- ❑ Vienna University of Technology
- ❑ Austrian Research Centers GmbH
- ❑ IBM Netherlands
- ❑ Microsoft Research Limited
- ❑ Tessella

Planets Overview – Activities

- ❑ **Preservation Planning** services
- ❑ Tools and Services for the **Characterisation** of digital objects
- ❑ Solutions for **Preservation Actions** tools which will transform and emulate obsolete digital assets
- ❑ An **Interoperability Framework** to integrate tools and services
- ❑ A **Testbed** to provide objective evaluation of different protocols, tools, services and complete preservation plans

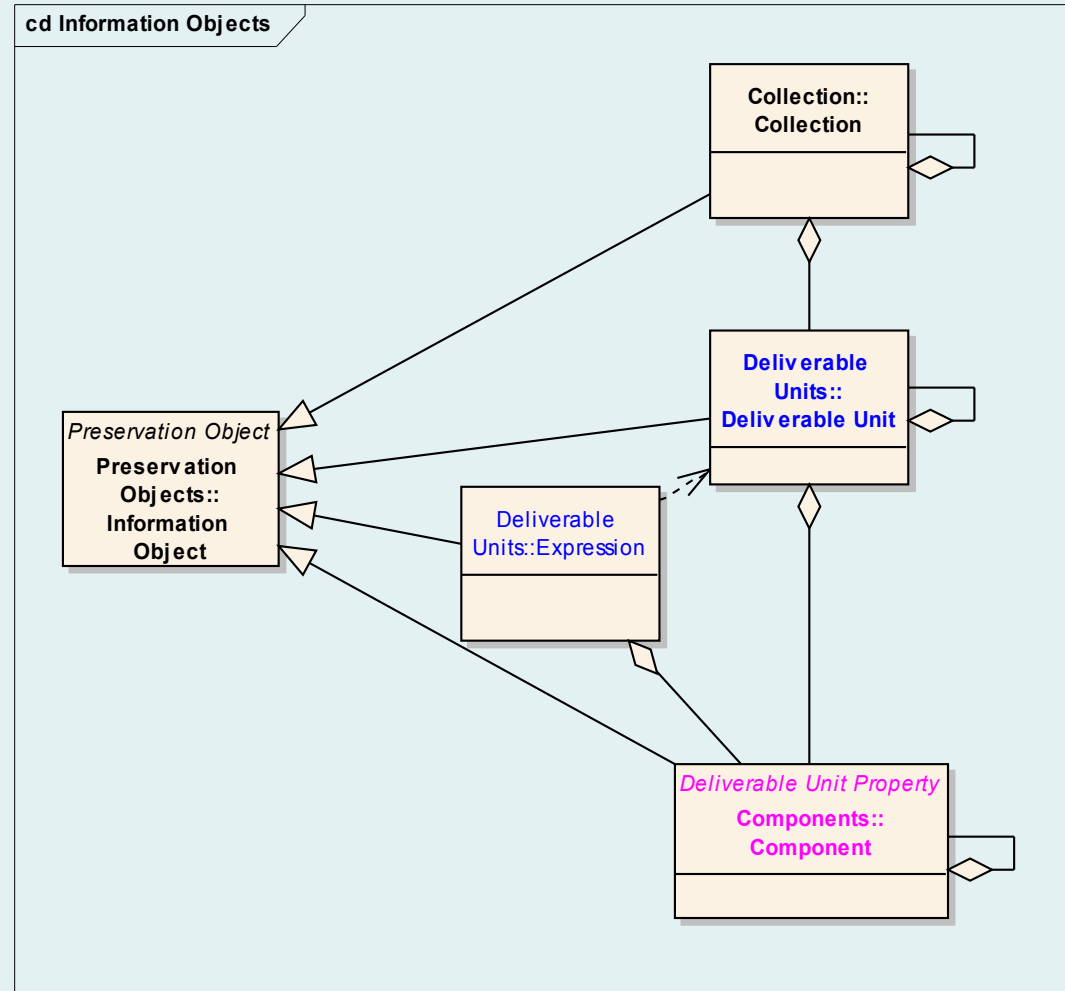


Why a Data Model?

- ❑ Need to understand same things:
 - Core input:
 - Information Objects / Digital Objects / Relationships
 - Characterisation output:
 - Digital Object Properties / Information Object Properties
 - Preservation planning output:
 - Approved action pathways
 - Essential characteristics
 - Migration output:
 - New files / New manifestations
 - Emulation:
 - New technical environment
- ❑ **MUST** be able to map between each other's inputs and outputs!

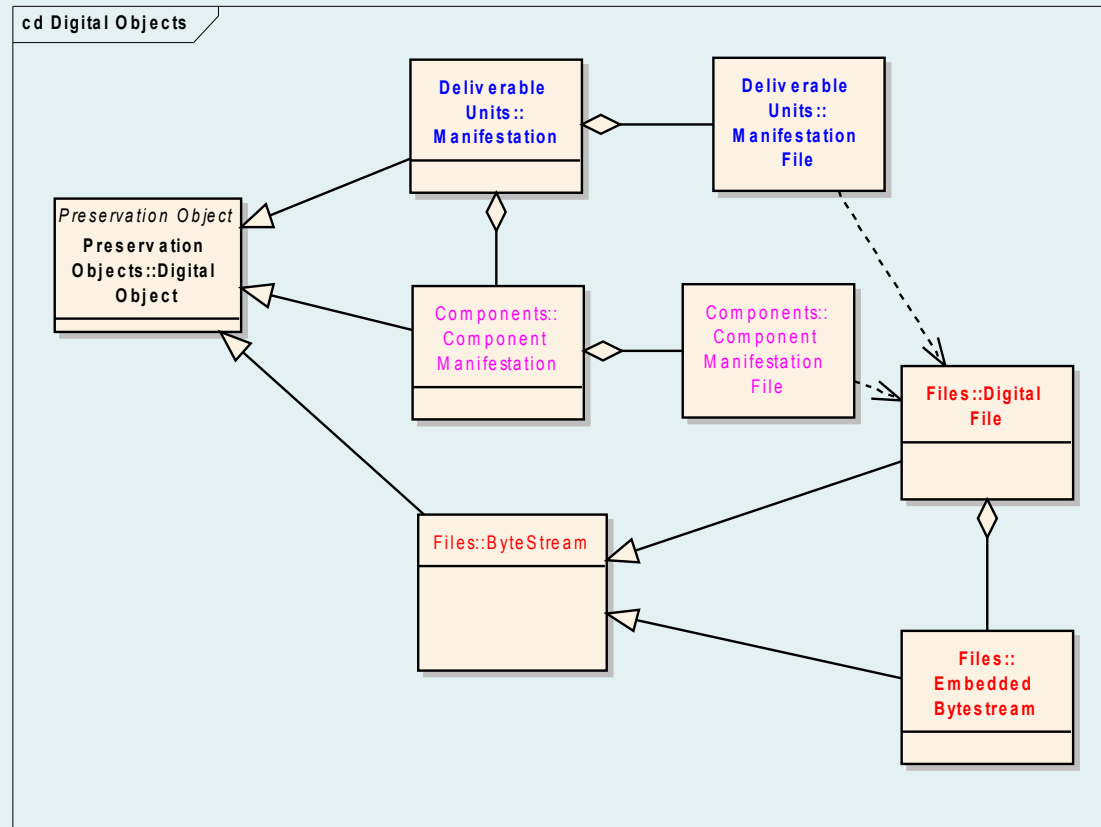
Information Objects

- ❑ The conceptual entities that need to be preserved
- ❑ 4 types
 - Collections
 - Deliverable Units
 - Expressions
 - Components
- ❑ Hierarchical structure

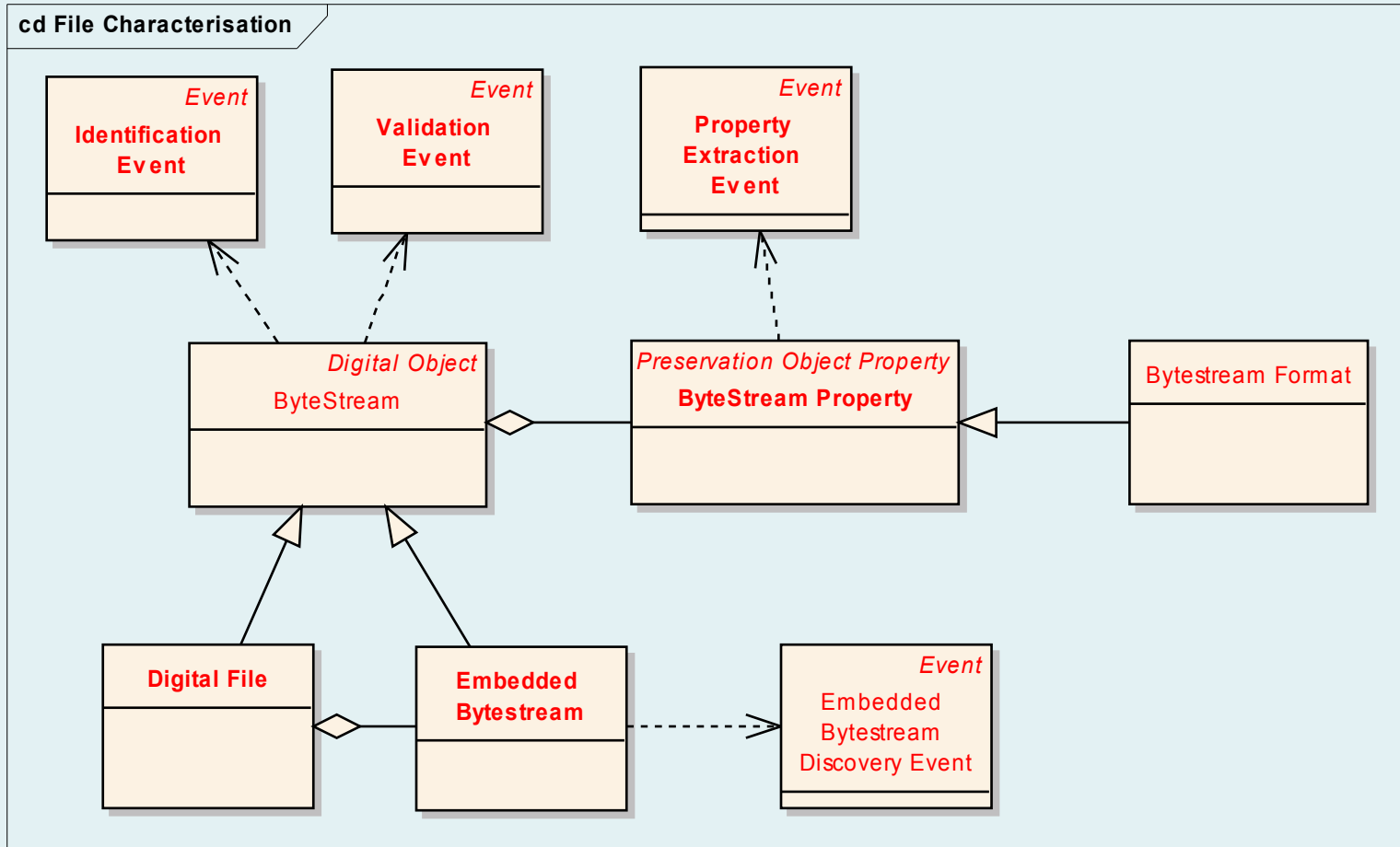


Digital Objects

- ❑ Represent the manifestation of the information objects
- ❑ Deliverable Units and Expressions are manifested through **Manifestations**.
- ❑ Components are manifested through **Component Manifestations**.
- ❑ Each Manifestation or Component Manifestation is the collection of **Digital Files**.

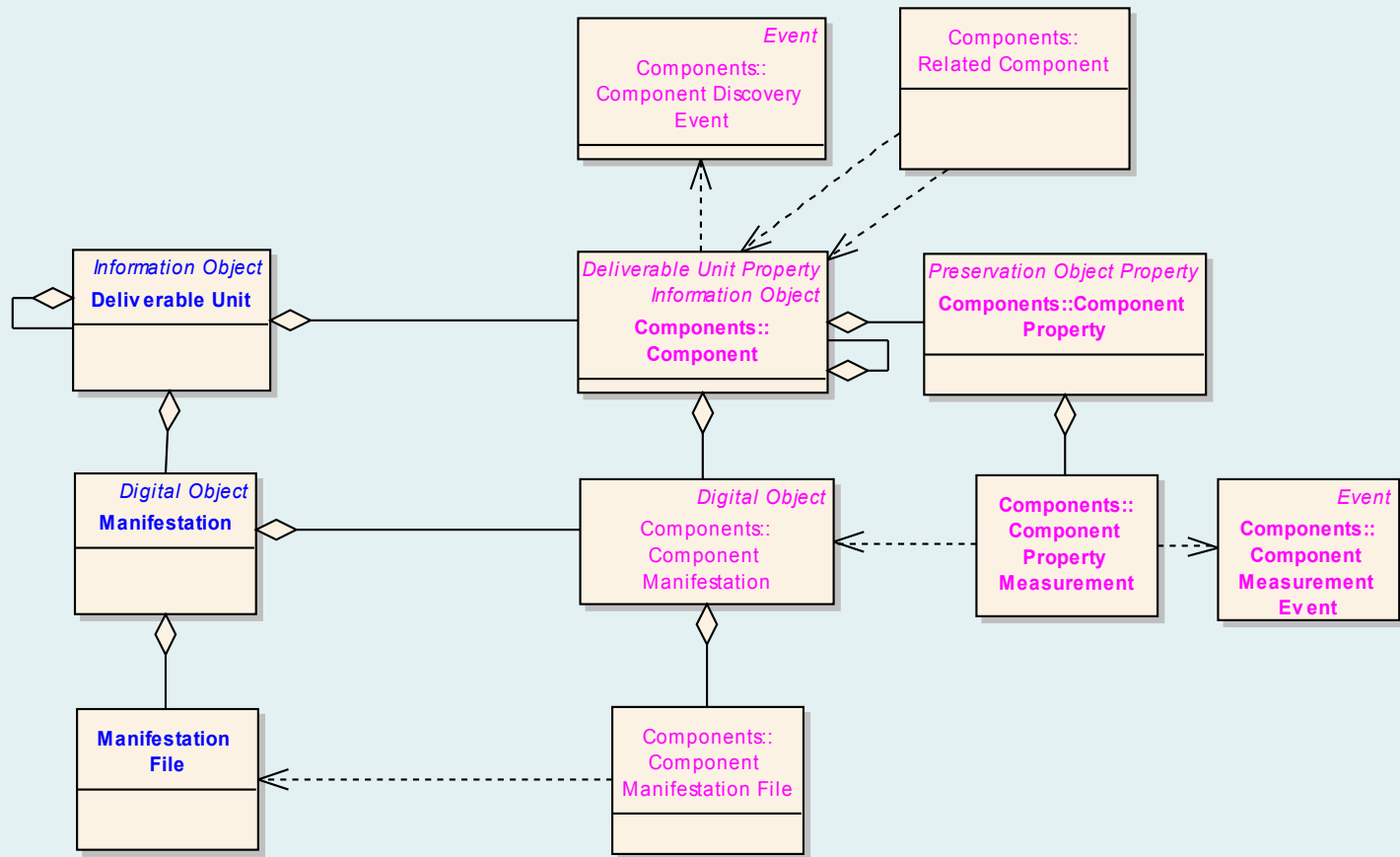


File-level characterisation

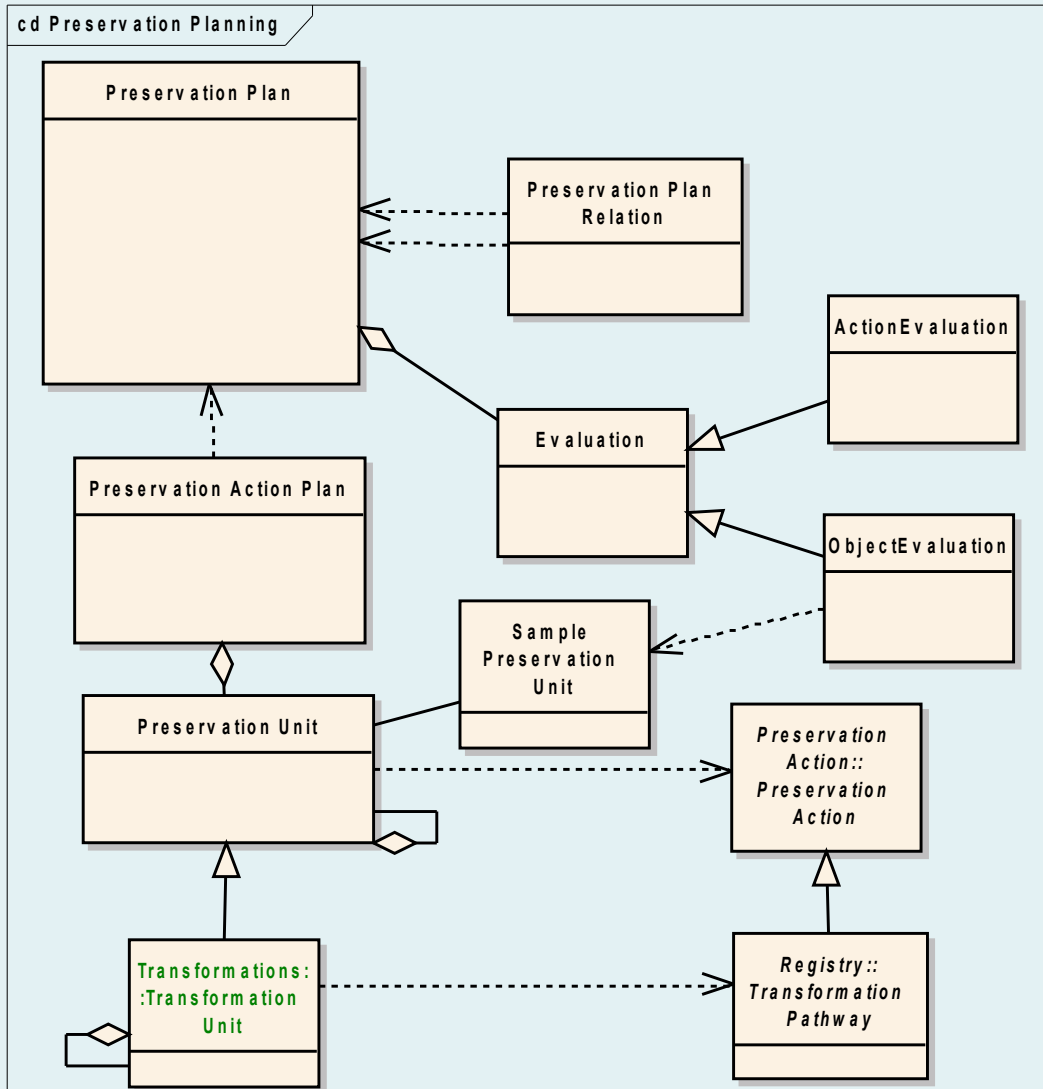


Record characterisation

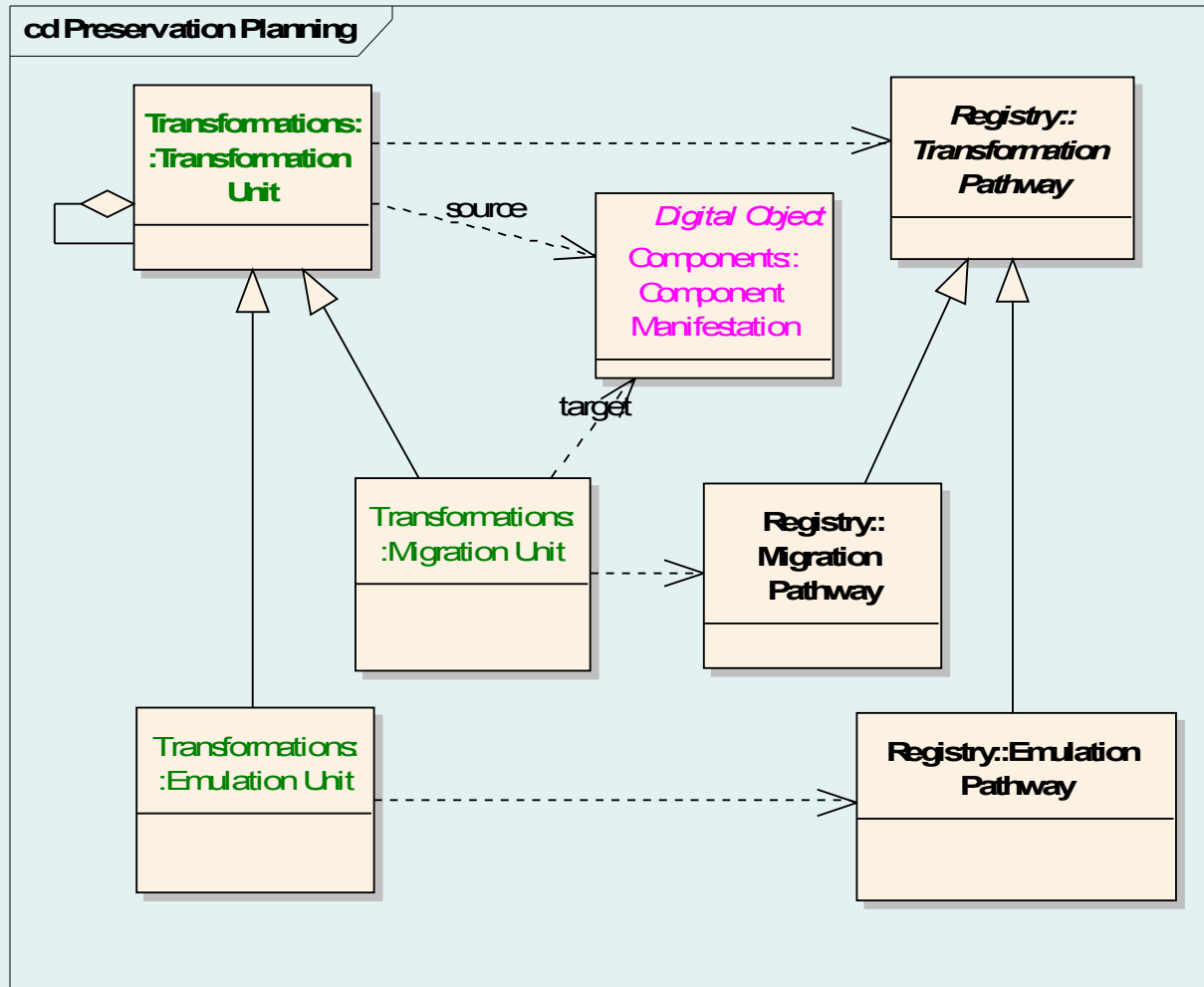
cd Deliverable Unit Characterisation



Preservation Planning 1

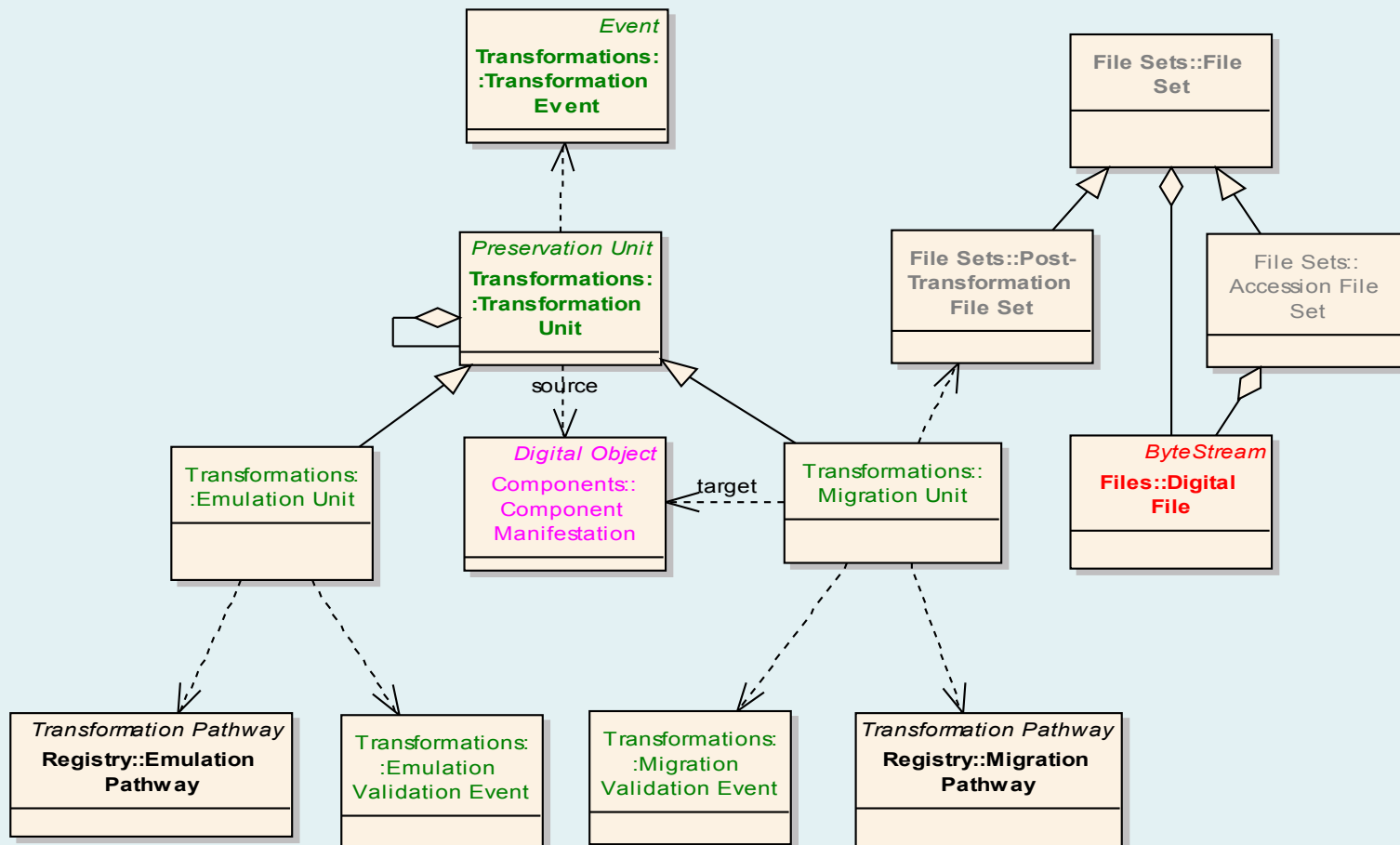


Preservation Planning 2



Preservation Action

cd Transformation



Registry interaction

- ❑ Separate model for the Technical Registry
- ❑ Conceptual Model assumes there will be references for:
 - Formats
 - Software
 - Technical Environments
 - Properties
 - Migration pathways
 - Emulation pathways
- ❑ All these will have PUIDs in new Registry:

Data Models : The Future

- V1R4:
 - Approved
- V1R5 changes in:
 - New preservation planning
 - Emulation
 - Various other changes
 - Not approved
- Other models under development:
 - Conceptual model
 - Physical model
 - Service definitions

Links

- Planets homepage:

<http://www.planets-project.eu>