

Preservation
as
a
Service for
Trust

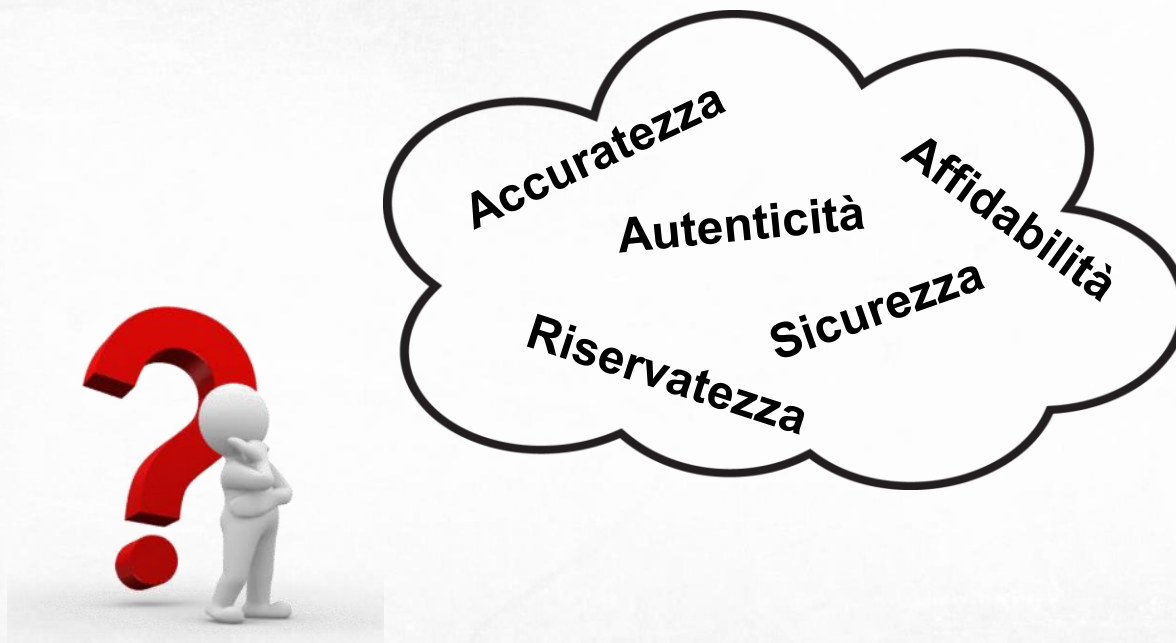
Giovanni Michetti

Università di Roma “La Sapienza”
InterPARES Northamerican Team

InterPARES Trust

<http://interparestrust.org/>

- Progetto di ricerca internazionale e interdisciplinare che esplora il tema della fiducia (trust) in relazione ai dati/documenti digitali affidati a Internet e al Cloud



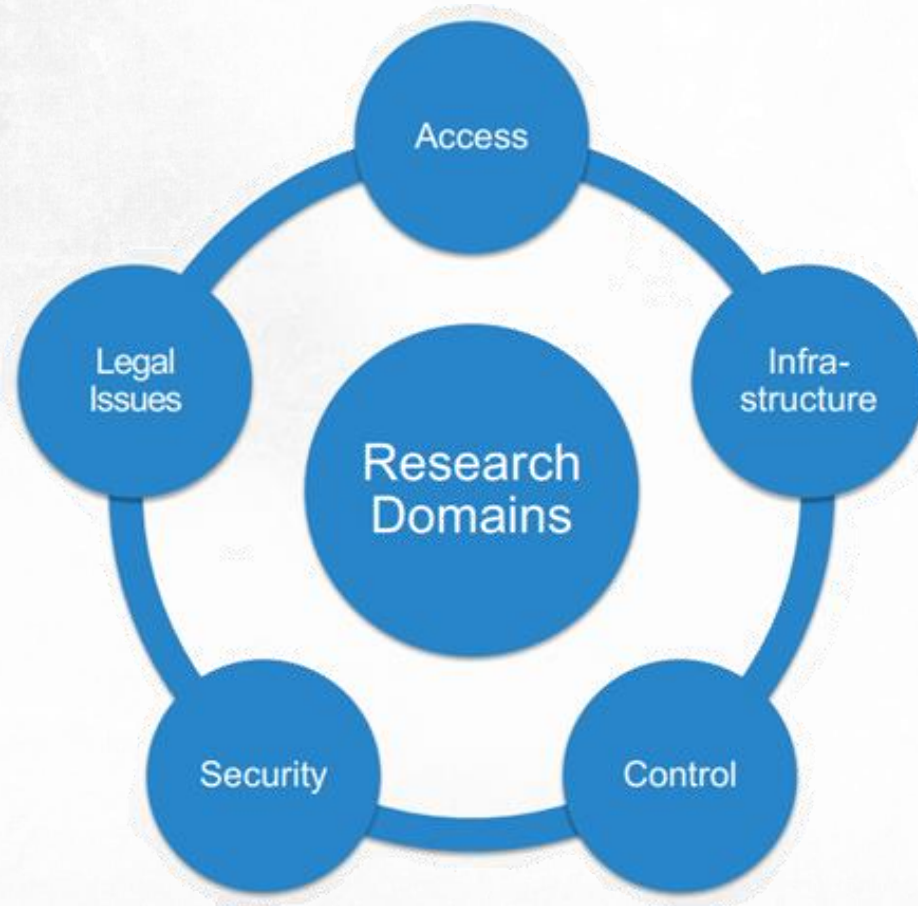
Partecipanti

- Team:
 - North America
 - Latin America
 - Europe
 - Asia
 - Australasia
 - Africa
 - Organizzazioni transnazionali
- Partner
- Consulenti
- Team gestionale

Tipologie

- Università
- Archivi nazionali e locali
- Biblioteche
- Organizzazioni internazionali
- Imprese
- Studiosi ed esperti
- ...

Aree di ricerca



What is trust?

A photograph of five wooden blocks arranged in a row on a wooden surface, spelling out the word "TRUST". The blocks are light-colored with dark letters. The background is a blurred wooden surface, and the lighting is warm, creating a soft glow around the blocks.

TRUST

Trust is about vulnerability



Trust is about risk management

- La fiducia si situa fra la **speranza** e la **certezza**



- Come ogni rischio, richiede la gestione di una **deviazione** dalle aspettative




Trust is a process



Trust is contextual



Trust is about culture

- **65%** dei Norvegesi  ritiene che ci si può fidare delle persone. Solo il **3%** dei Brasiliani  ritiene che ci si può fidare delle persone

- Studi evidenziano che la fiducia è associata ad aspetti di dimensione nazionale come la ricchezza, l'educazione e l'omogeneità etnica

Trust is an economic asset



- L'informazione è diventata una **commodity** con un valore economico. La fiducia è il **framework** che permette a tale valore di crescere, muoversi ed essere scambiato

Trust and records

- I documenti sono un'**evidenza** di atti e fatti, e ci consentono di comunicare attraverso spazio e tempo
- **Tale comunicazione è basata sulla fiducia**
 - la fiducia è parte integrante dei documenti (*embedded*)
 - la fiducia è coinvolta nei processi di trasmissione
- L'ambiente digitale rende tutto più **complesso**

Né SaaS né IaaS né PaaS

SaaS (Software as a Service)

(CRM, email, virtual desktop, communication, games, ...)

IaaS (Infrastructure as a Service)

(virtual machine, server, storage, network, ...)

PaaS (Platform as a Service)

(execution runtime, database, web server, development tools, ...)



Cloud Clients

(Web browser, mobile app, thin client, terminal emulator, ...)

PaaS

Preservation. as a Service. for Trust.

Obiettivo

- Creare un **modello** per la conservazione dei documenti digitali nel Cloud mantenendo la loro autenticità
- PaaS definisce i modelli dati e i requisiti funzionali relativi ad azioni, agenti e oggetti che interagiscono in un **preservation space** non limitato a un singolo deposito o sistema

Criticità

- **Responsabilità** operative della conservazione **distribuite**, appaltate o subappaltate



- Documenti digitali spostati attraverso **diversi sistemi, depositi e tecnologie**



- **Vulnerabilità** crescente

Approccio

- La conservazione è realizzata attraverso un insieme di attività correlate che possono essere svolte da **diversi service provider** usando diverse tecnologie, con controlli operativi separati e indipendenti



Sfide della modellazione

Il modello è stato progettato per essere applicabile in una **varietà di contesti**:

- **tipi** di oggetti informativi
- **direttive** applicabili
(laws, regulations, standards, policies, business rules, agreements,...)
- **condizioni** relative a proprietà, accesso, utilizzo...
- **accordi** e **relazioni** tra le parti
- ...

Sfide della modellazione

Il modello è stato progettato per essere

- **indipendente** da specifiche tecnologie
- **implementabile**

PaaS usa UML

- **UML** = Unified Modelling Language
- I requisiti modellati in UML possono essere usati per la **generazione automatica di codice**

PaaST si basa su concetti archivistici

- Il quadro teoretico dell'**Archivistica** fornisce la conoscenza di base necessaria per modellare la funzione conservativa nel Cloud
- In aggiunta, due fondamentali modelli per la conservazione:
 - **OAIS** (Open Archival Information System)
 - **CoP** (Chain of Preservation)

Chain of custody

Ambiente tradizionale

- I documenti conservati si **presumono autentici** se sono stati conservati attraverso un'ininterrotta sequenza di conservatori affidabili per tutta la loro esistenza



Chain of custody

Ambiente digitale

- La custodia affidabile non basta: i **rischi** di corruzione o di identificazione erronea sono **troppo alti**



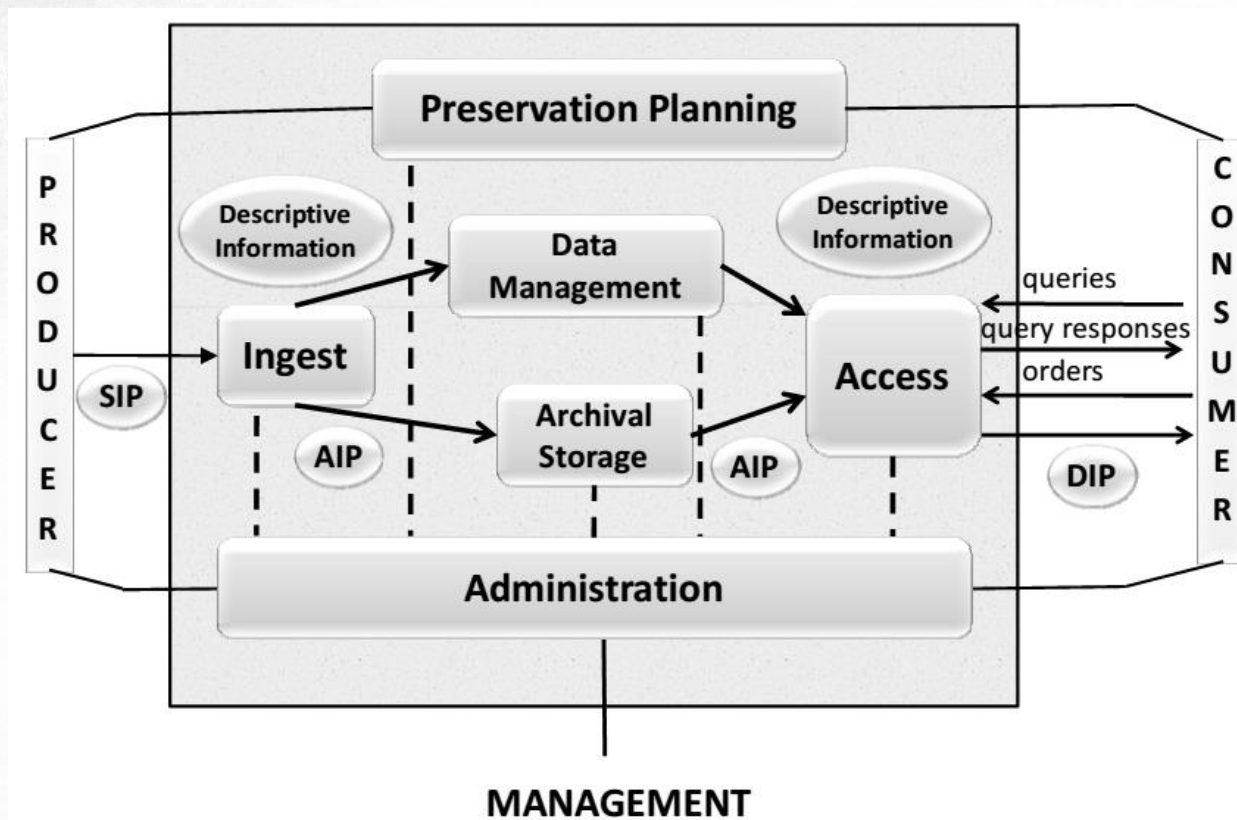
- Serve **ulteriore evidenza** per presumere che l'informazione digitale sia autentica

Chain of Preservation

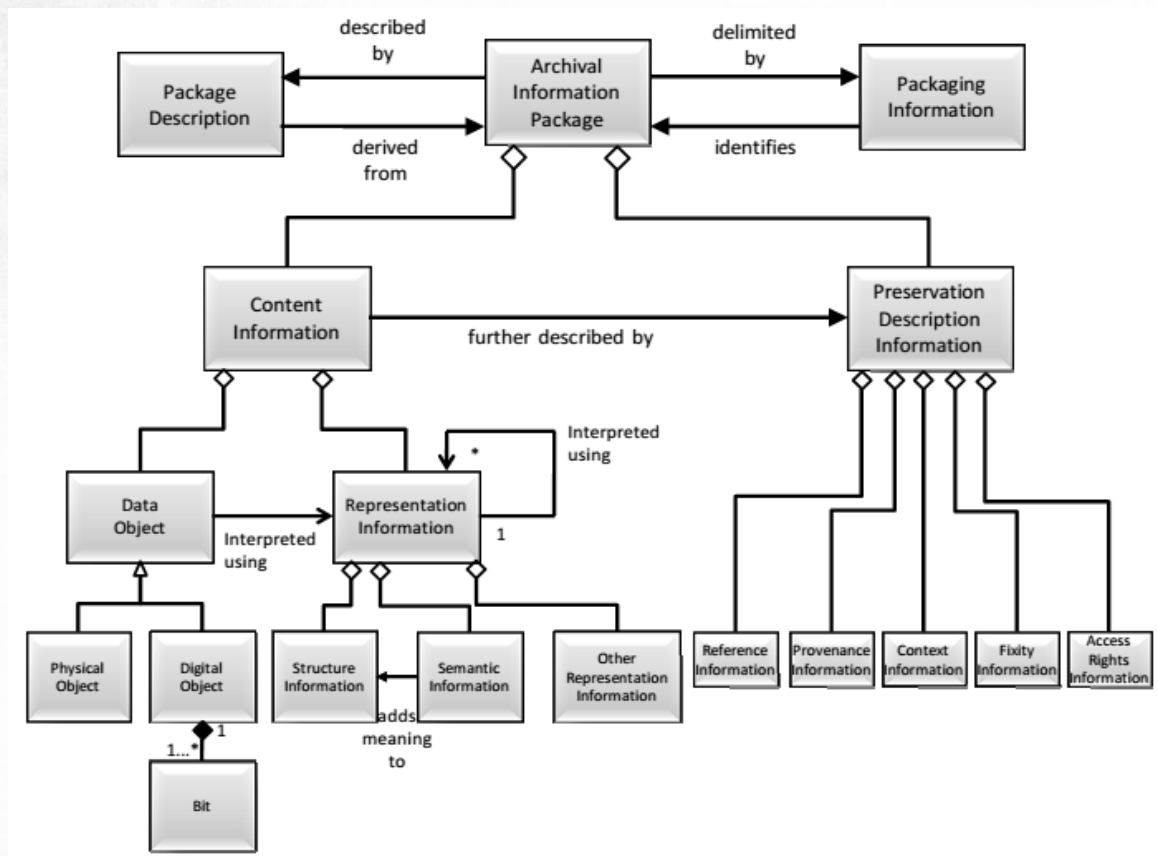
Ambiente digitale

- Dobbiamo provare che i documenti sono stati soggetti ad un **regime ininterrotto di controlli e misurazioni** sufficienti a mitigare tali rischi
- **CoP**: Un sistema di controlli che si estende per l'intera vita dei documenti e garantisce la loro identità ed integrità nel corso di qualunque azione che in qualunque modo li coinvolga

OAIS. Modello funzionale



OAIS. Modello informativo



Tweaking OAIS - #1

OAIS

Sistema singolo che ricomprende tutte le attività fondamentali ed include tutte le informazioni

PaaST

Oggetti e relative informazioni sono sparse tra diversi service provider

Tweaking OAIS - #2

OAIS

Oggetti digitali e non-digitali

PaaST

Solo oggetti digitali

Tweaking OAIS - #3

OAIS

Si occupa di *informazioni*

PaaST

Si occupa di *informazioni*, ma è progettato per *documenti*

Tweaking OAIS - #4

OAIS

Non specifica un tipo di architettura o di implementazione

PaaST

È formulato per facilitare l'implementazione

Tweaking OAIS - #5

OAIS

- Si occupa di informazioni che non necessariamente sono *machine actionable* (ad esempio, dizionario e grammatica inglese)

PaaST

Mira a facilitare l'uso di metodi automatici per la conservazione di informazioni digitali



Si occupa di informazioni *machine actionable*

Tweaking OAIS - #6

OAIS

- Designated Community:
gruppo target di utenti

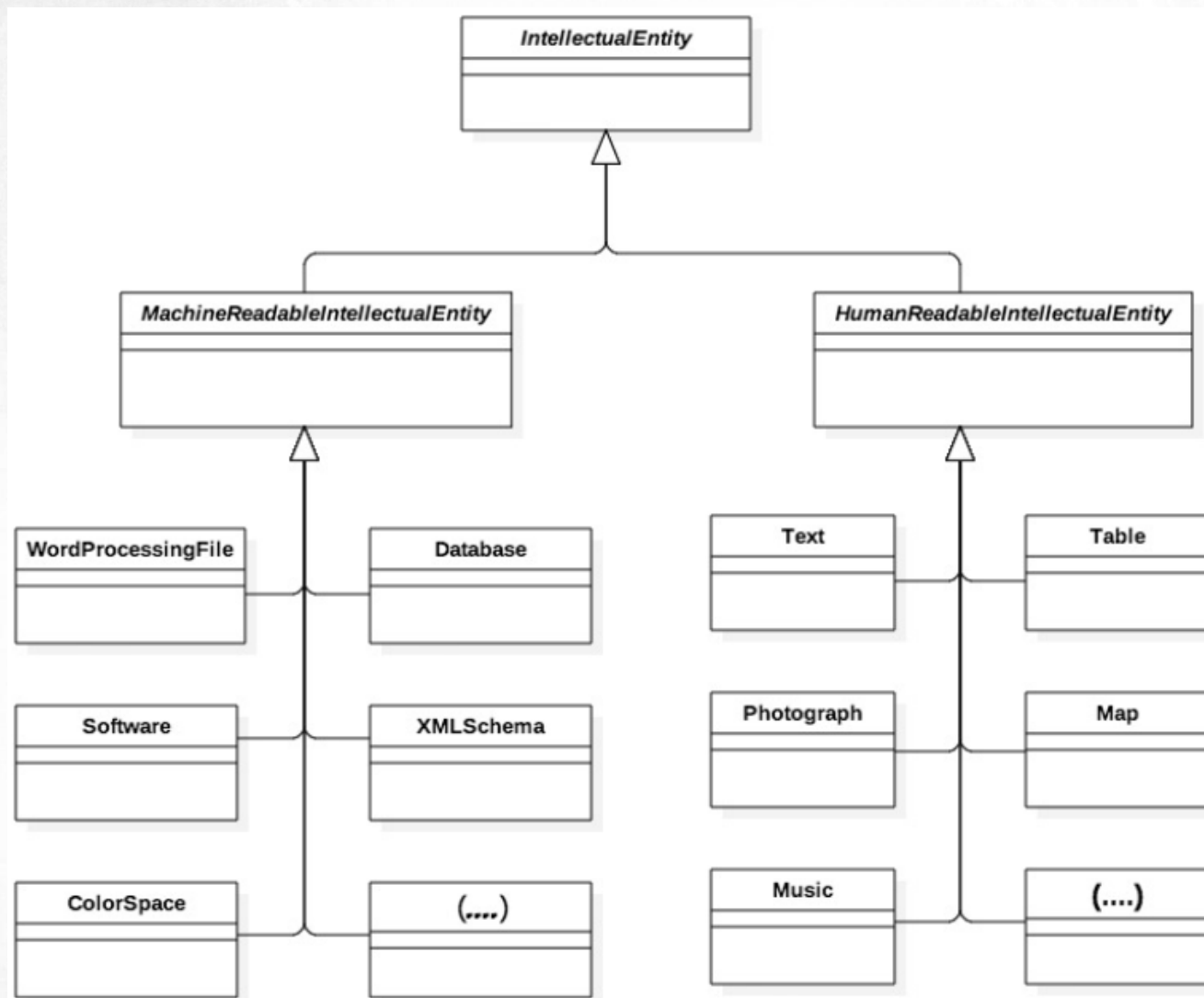
PaaST

Designated Community:
trattata in piani, politiche, procedure

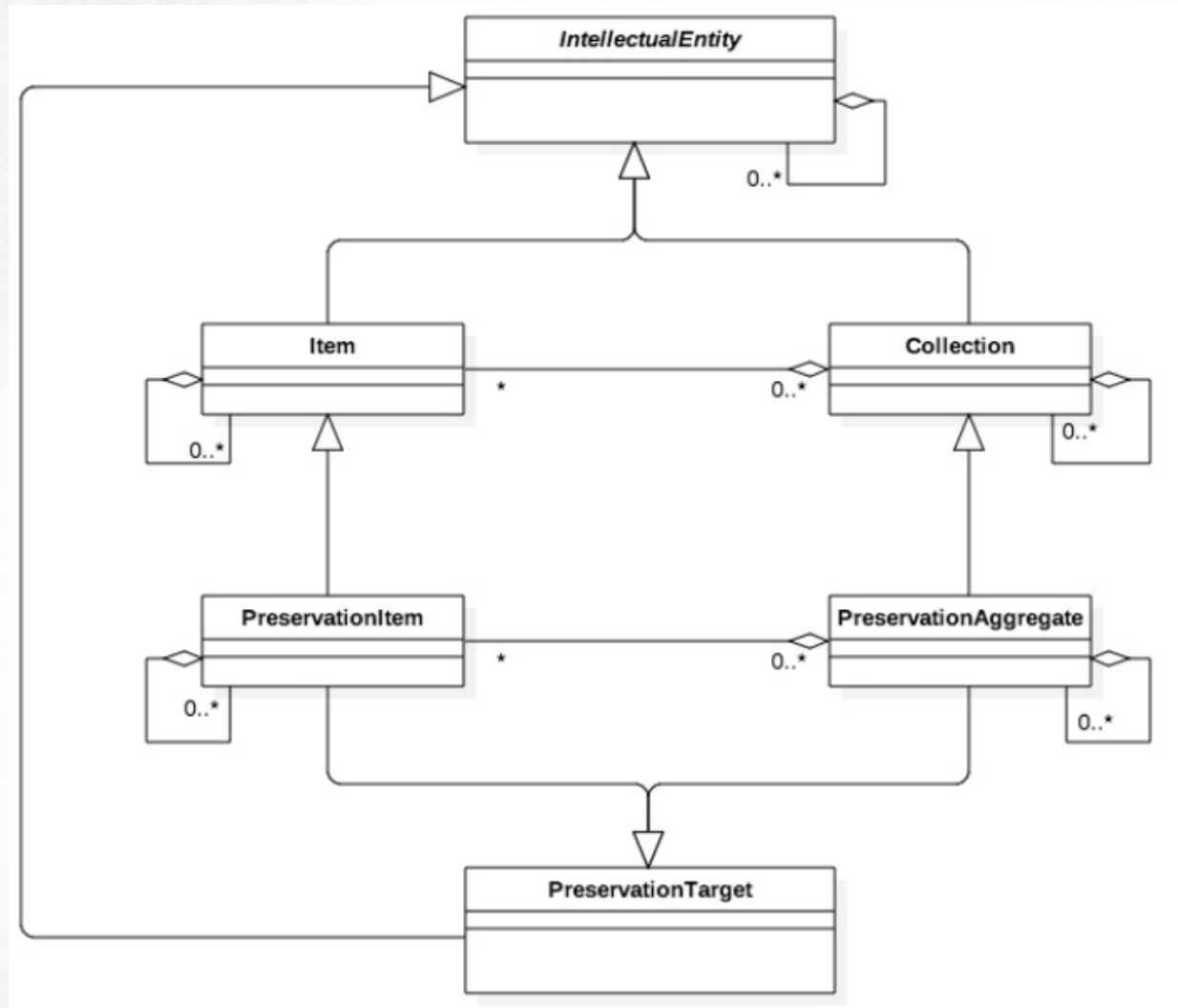


Eventualmente rappresentata e
implementata come Preservation
Rules

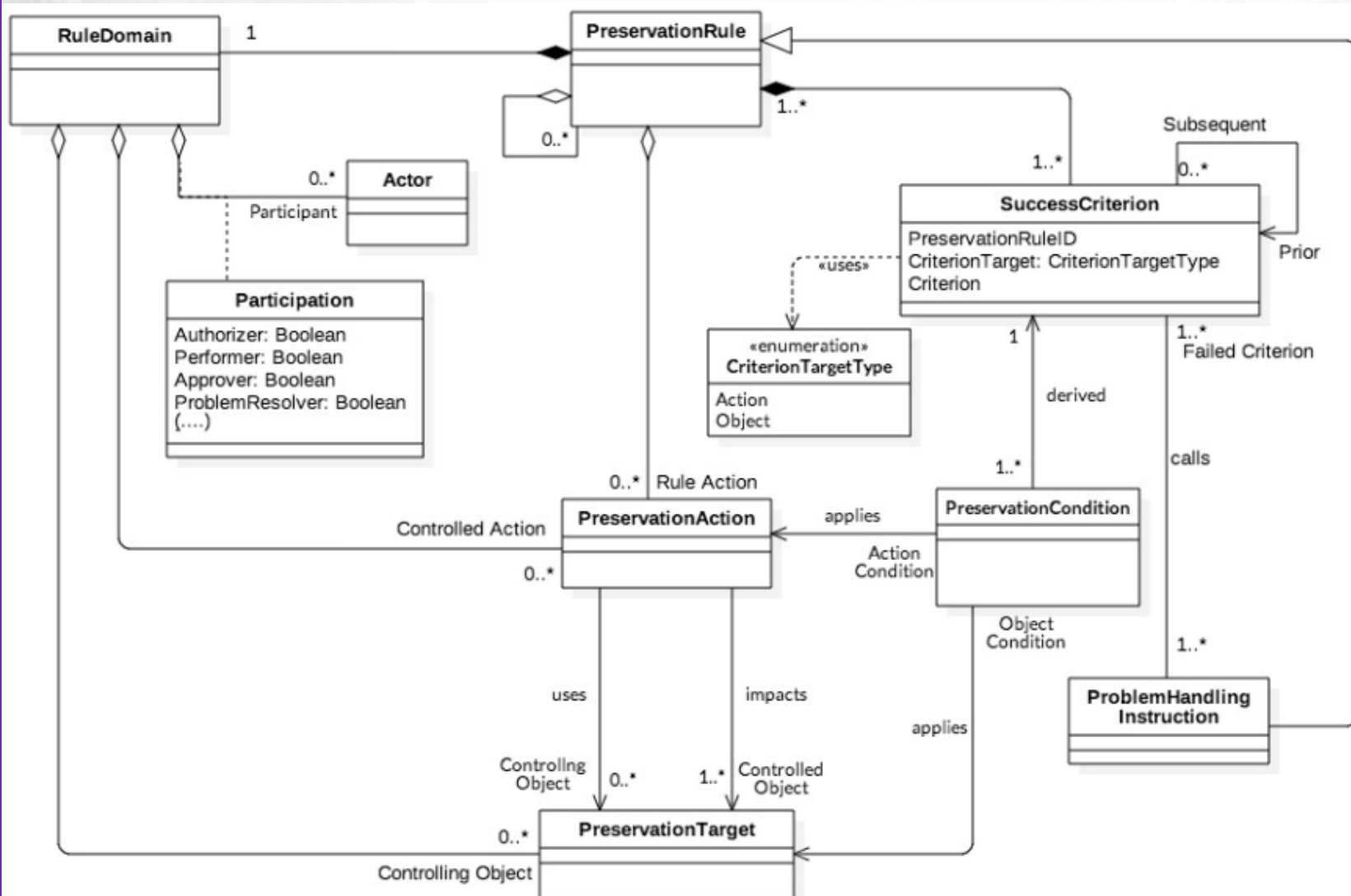
PaaST. Intellectual Entity



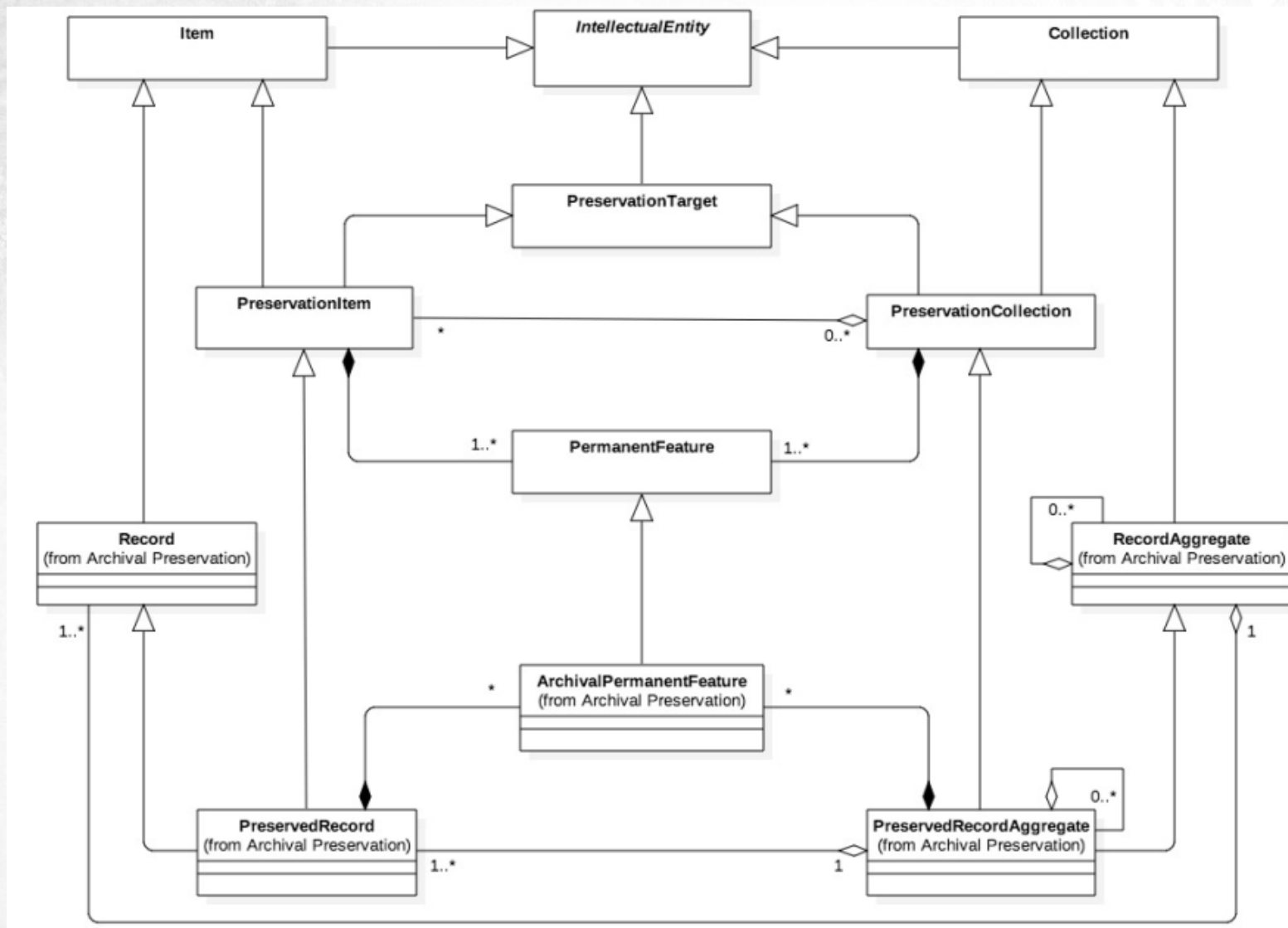
PaaS. Intellectual Entity Composition



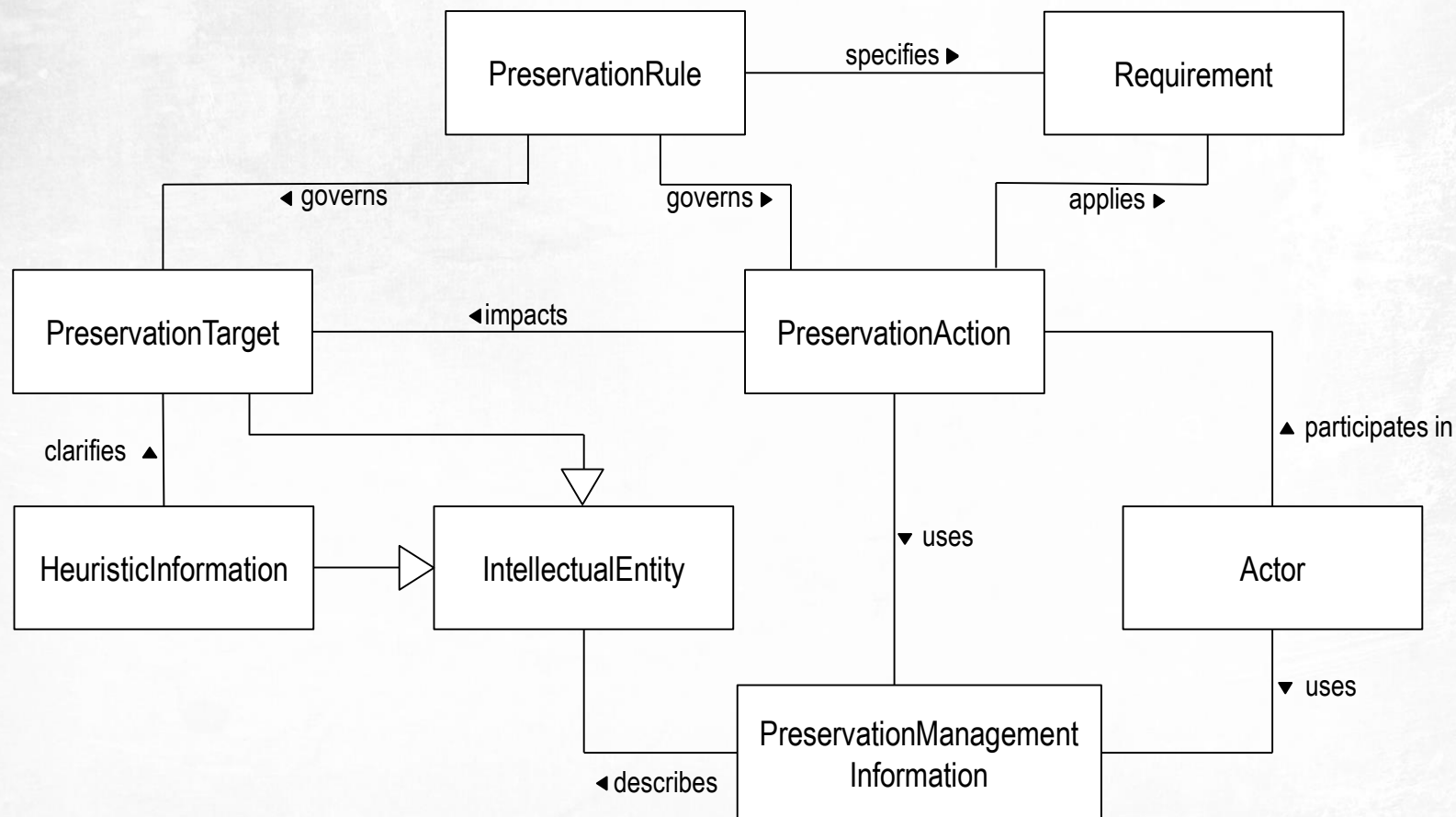
PaaST. Preservation Rule



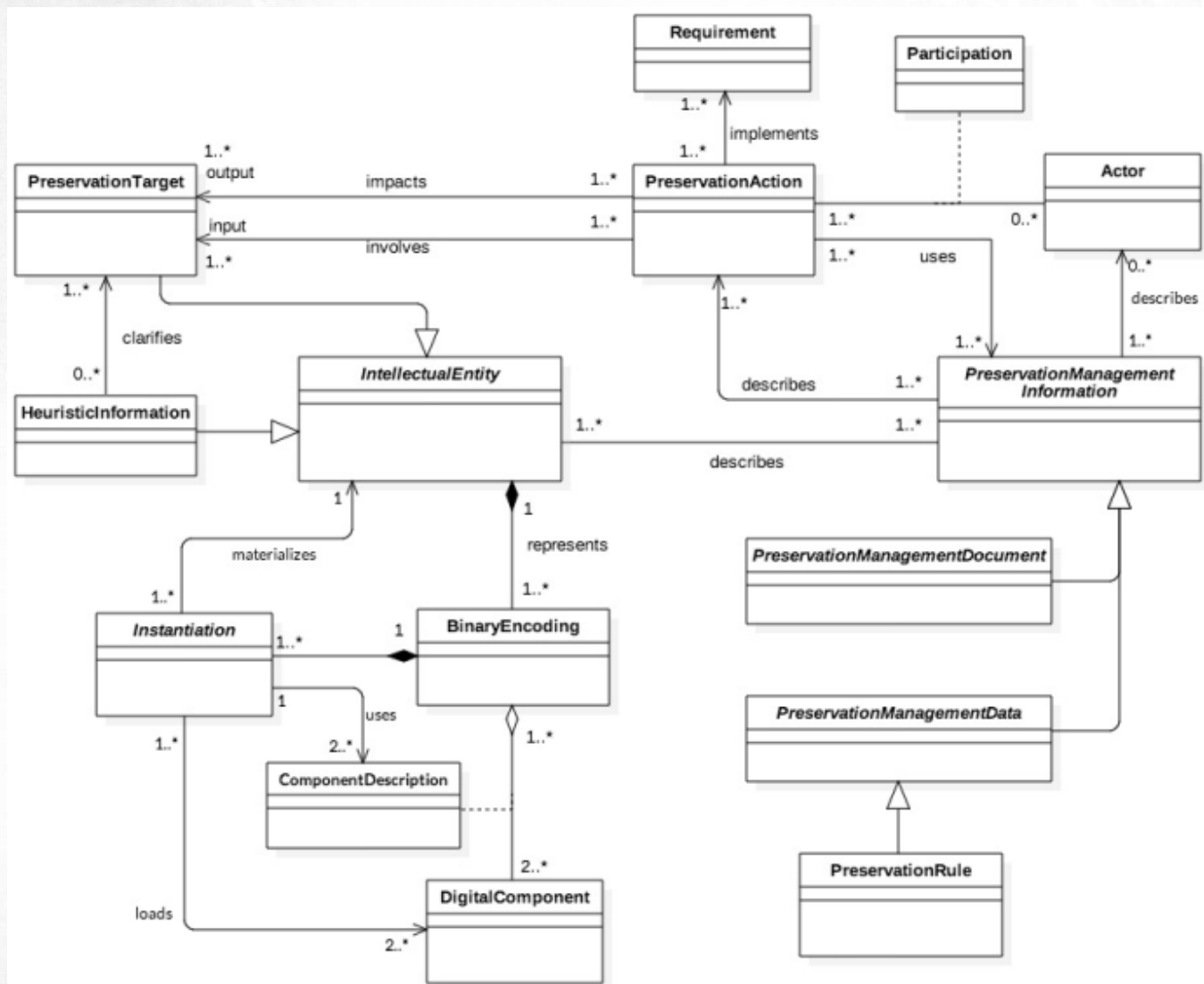
PaaST. Archival Preservation Target



PaaST. Preservation (overview)



PaaST. Preservation (overview)



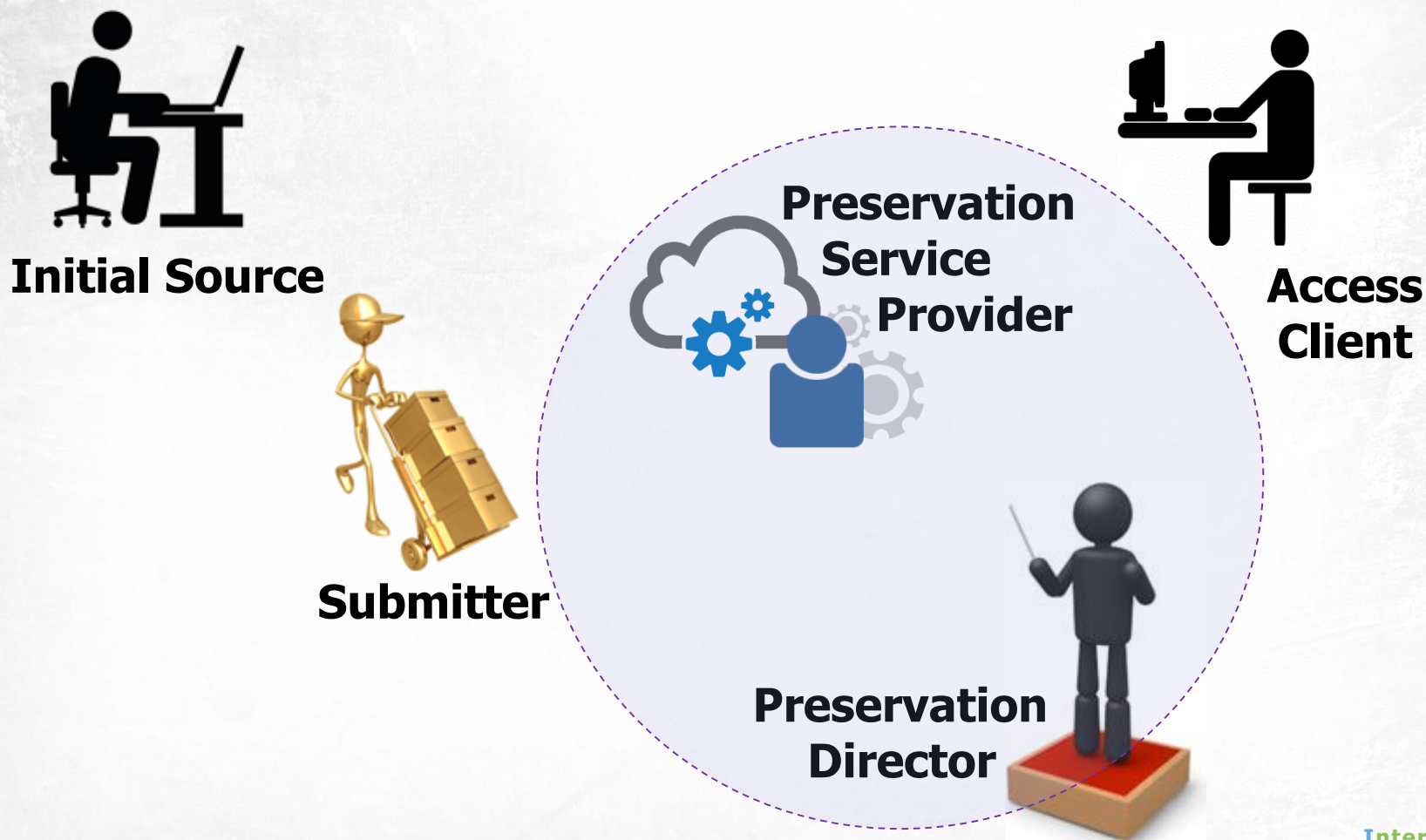
PaaST. Environment

- Diversi agenti interagiscono in maniera modulare e integrata per svolgere diversi insiemi di attività conservative
- PaaST si concentra sui tipi di attività, non sulle entità che svolgono tali attività



- PaaST distingue le parti coinvolte nella conservazione sulla base del ruolo (**capacity**) che svolgono

PaaST. Capacities



PaaST. Activities

Insieme delle azioni di conservazione coperte da PaaST:

- **negoziiazione** di termini e condizioni di conservazione
- **trasferimento** al Service Provider degli oggetti da conservare
- **gestione** degli oggetti in archivio
- **consegna** degli oggetti informativi agli utenti

PaaS.T. Requirements

- **Requirements**: ciò che un Preservation Service Provider deve garantire per svolgere un'azione a supporto della conservazione digitale
- I requisiti sono aggregati in **Capabilities** (cioè, azioni o insiemi di azioni)
- Le Capabilities possono essere combinate per formare dei **Servizi** compositi, e implementate usando diversi metodi e strumenti

PaaST. Capabilities

Preservation Action Capabilities

- Azioni correlate direttamente alla conservazione dei Preservation Targets

Preservation Management Capabilities

- Azioni che garantiscono che le prestazioni delle Preservation Actions siano tali da raggiungere gli obiettivi della conservazione

A che punto siamo?



PaaST. Requirements

1. Process a Submission

1.1 Receive a Submission

<i>ID</i>	<i>Requirement</i>
1.1.1	The service shall provide the capability to store a received Submission.
1.1.2	The service shall provide the capability to receive a PreservationManagementDocument that contains SubmissionInformation.
1.1.2.1	The service shall provide the capability to store a received PreservationManagementDocument that contains SubmissionInformation.
1.1.3	The service shall provide the capability to receive SubmissionData.
1.1.3.1	The service shall provide the capability to store received SubmissionData.
1.1.4	The service shall provide the capability to receive a DigitalComponent as part of a Submission.
1.1.4.1	The service shall provide the capability to store a DigitalComponent received in a Submission.

PaaS come standard



L'Object Management Group (OMG) ha approvato lo sviluppo di uno standard sulla conservazione digitale basato sui requisiti funzionali definiti in PaaS

The prologue is past

*«We have a **broad base of experience** from which to reach some conclusions and view the future»*

David Warner, former president of the Society of American Archivists

The prologue is past

- Non possiamo più fidarci dei documenti usando lo stesso approccio e gli stessi metodi del passato
- Tuttavia, tali approcci e metodi costituiscono un terreno solido sul quale costruire nuove prospettive, modelli e strumenti



The prologue is... PaaS

Grazie dell'attenzione!



giovanni.michetti@uniroma1.it