Modeling Authenticity

Mariella Guercio & Giovanni Michetti

University of Urbino

Carlo Meghini

CNR ISTI

Outline

- Introduction
- Critical issues
- Approach
- Authenticity in CASPAR
- Conclusions

Outline

- Introduction
- Critical issues
- Approach
- Authenticity in CASPAR
- Conclusions

Authenticity:
A Key Component in the Preservation Process

- The preservation as developed by the main international projects in the sector (InterPARES and OAIS) requires that the elements related to the accuracy, the reliability and the completeness of the information objects are captured and maintained in the repositories to allow the users to evaluate their **identity** and their **integrity** (InterPARES project)

Authenticity:
A Key Component in the Preservation Process

- These elements have to be organized according to a **conceptual model** (OAIS compliant) that is able to describe the dynamic profile of the authenticity as a **process** aimed at gathering, protecting and/or evaluating information about identity and integrity

The need for an Authenticity Management Tool (AMT)

- The complexity of the preservation function in the digital area requires the development of specific **tools** able to ensure that the main elements and procedures relevant for the quality of the preservation are maintained, and the authenticity of the preserved information objects can be presumed

The need for an Authenticity Management Tool (AMT)

- The CASPAR project has identified the need for an Authenticity Management Tool with the capacity of **monitoring and managing protocols and procedures across the custody chain** in order to deliver the benefits of authenticity into information systems, from the creation to the preservation phase

Outline

- Introduction
- Critical issues
- Approach
- Authenticity in CASPAR
- Conclusions

Critical issues

(1) Integrity and Identity

Integrity

- The **integrity** of a resource refers to its wholeness.
- A resource has *integrity* when it is complete and uncorrupted **in all its essential respects**.
- The verification process should analyse and ascertain that these essential aspects are consistent with the inevitable changes brought about by technological obsolescence

Integrity

- While **the maintenance of the bit flow is not always necessary**, the completeness of the ‘intellectual form’ is required, especially with respect to the original ability to convey meaning
  - colours in a map
  - columns in a spreadsheet
- The physical integrity of a resource i.e. the original bit stream can be compromised, but the content structure and the essential components (significant properties) must remain the same

Identity

- *identity* must be understood in a very wide sense: the identity of a resource refers not only to its unique designation or identification, but
- to *the whole* of the characteristics of a resource that uniquely distinguish it from any other resource
- it refers not only to its internal conceptual structure but also to its **general context (administrative, legal, documentary, technological, some could even add social)**

Need to cope with authenticity

- An OAIS needs to have **tools** and **methods** that ensure authenticity of objects information along the preservation process

- The main issue is to **document** objects as automatically and neutrally as possible on the basis of an adequate (OAIS compliant) methodology

Critical issues

(2) Tools for Managing Authenticity

Requirements

- **Authenticity cannot be evaluated by means of a boolean flag** telling us whether a document is authentic or not

- **There are degrees in the capacity of presuming the authenticity of the digital resources**: the certainty about authenticity is a goal

Requirements

- We have to design all the mechanisms and tools keeping in mind that
  - we could have alteration, corruption, lack of significant data etc.
  - we need changes to ensure accessibility
  - we need tools, mechanisms and *weights* to understand their relevance and their impact on authenticity

Requirements

- Authenticity Management Tools have to identify mechanisms for ensuring the maintenance and verification of the authenticity in terms of identity and integrity of the digital objects

Requirements

- These tools have to provide **content and contextual information relevant to authenticity**, i.e. to the identity and integrity profile, **all along the whole preservation process by capturing and making understandable over time all the required information**

Requirements

- The main issues for the AMT are:
  - the right attribution of authorship
  - the identification of provenance in the life cycle of information objects
  - the insurance of content integrity of the whole relevant digital components and their relevant contextual relationships

Requirements

- The main issues for the AMT are:
  - the provision of mechanisms to allow future users to verify the authenticity of the preserved information objects or at least to provide the capability of evaluating their reliability in term of authenticity presumption

Requirements

So these requirements imply working on:

- authorship attribution mechanisms and provenance control
- content and contextual relationships
- integrity control mechanisms
- annotation process

Requirements

- **Every relevant aspect has to be described and documented at every stage in the life cycle** so to have, any time is needed, a sort of ‘Authenticity Card’ for any object in the repository

Outline

- Introduction
- Critical issues
- Approach
- Authenticity in CASPAR
- Conclusions

Approach

- **A conceptual model** for describing the dynamic profile of authenticity i.e. **to describe it as process** aimed at gathering, protecting and/or evaluating information mainly about identity and integrity

Methodology

- **Authenticity Team has considered PREMIS, ISAD and other descriptive standards** in order to have a very general idea of some fundamental information elements which are to be preserved for ‘authenticity purposes’
- This was assumed **as a starting point** to find some more elements by taking into account other resources (i.e. ISAAR, EAD, EAC, InterPARES, …)

Methodology

- **CIDOC CRM** was assumed as a suitable means of expressing concepts and as a resource giving us clues about relevant aspects needed for consideration, especially about dynamic aspects (temporal entities)

Problems

level of granularity

- Authenticity fundamental requirements must be clearly identified in order to avoid at the same time overload and lack of information
- This is a relevant aspect for the scientific but also for the cultural domain, intended as dynamic environment with significant values in the current life of the creators and preservers like performing arts, digital music, protecting memory institutions.

Problems

variety of domains

- Authenticity methodology and concepts are cross-domain but their deployment is strongly dependent on specific environments.
  - the Reference Information for a book could be ISBN, very specific and not suitable for other typologies
  - the authorship concept is quite 'easy' for a book but what about the author of a movie , or other cultural products in the performing arts?

Problems

Integration of concepts coming from different ontologies

- Some concepts have a partial overlapping
- It is not easy to decide whether an element has to be mapped onto either this or that OAIS conceptual element (e.g. whether the ISAD element “System of arrangement” belongs to either OAIS Provenance or OAIS Context).

Outline

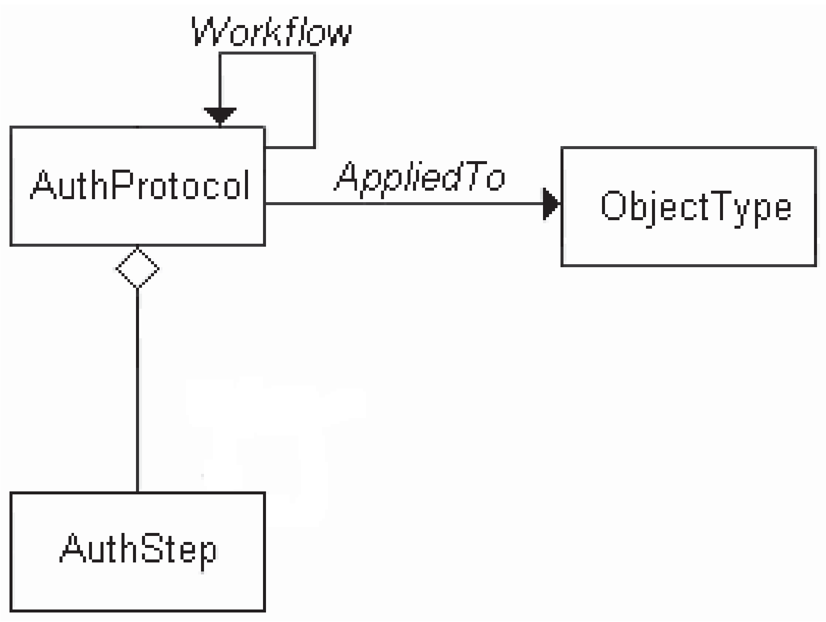
- Introduction
- Critical issues
- Approach
- Authenticity in CASPAR
- Conclusions

Authenticity in CASPAR

(1) The Conceptual Model

Authenticity Protocol

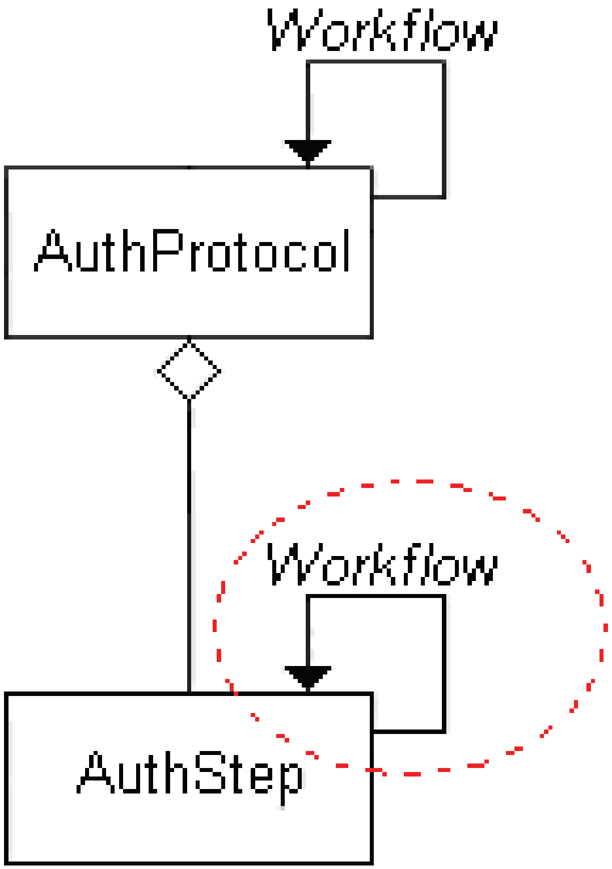
- The protection and assessment of the authenticity of digital object is a **process**.
- In order to manage this process, we need to define the procedures to be followed
- We call one of these procedures an **Authenticity Protocol** (abbreviated as **AP**)



- An AP is a set of interrelated steps, each called **Authenticity Step** (abbreviated as **AS**)

- An AP is applied to an **Object Type**, i.e. to a class of objects with uniform features for the application of an AP

- Any AP consists of a set of (Authenticity) Steps (Ass)

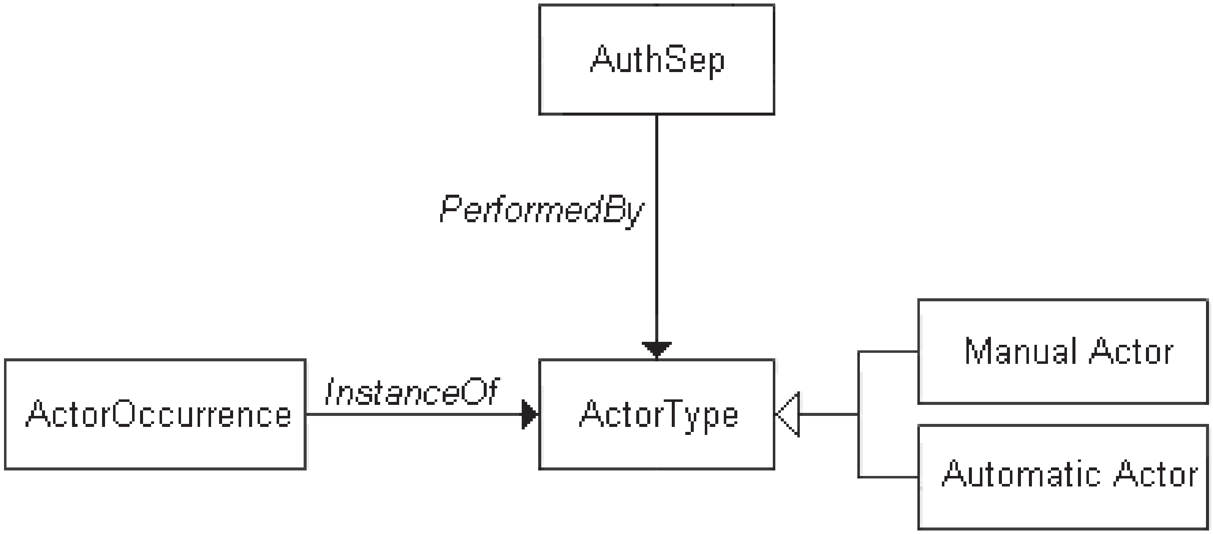


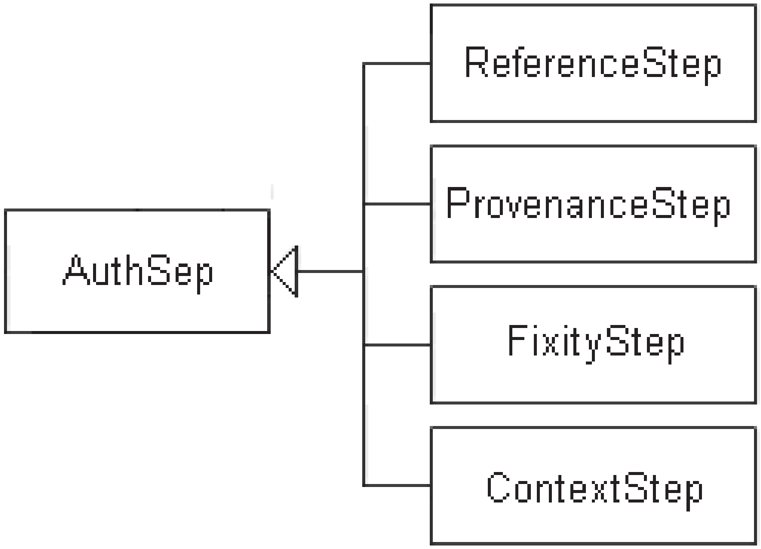
- Every AS models a part of an AP that can be executed independently as a whole, and constitutes a significant phase of the AP from the authenticity assessment point of view.

- The relationships among the steps of an AP establish the order in which the steps must be executed in the context of an execution of the protocol

- To model these relationships we can use any workflow model. We do not enter into the details of this modeling here, and simply denote as ***Workflow*** the set of required relationships

- An AS is performed by an **Actor Type**, a class that generalizes both **Automatic Actor** and **Manual Actor**, the former performing tasks in an automatic way (hardware/software), the latter using human intervention





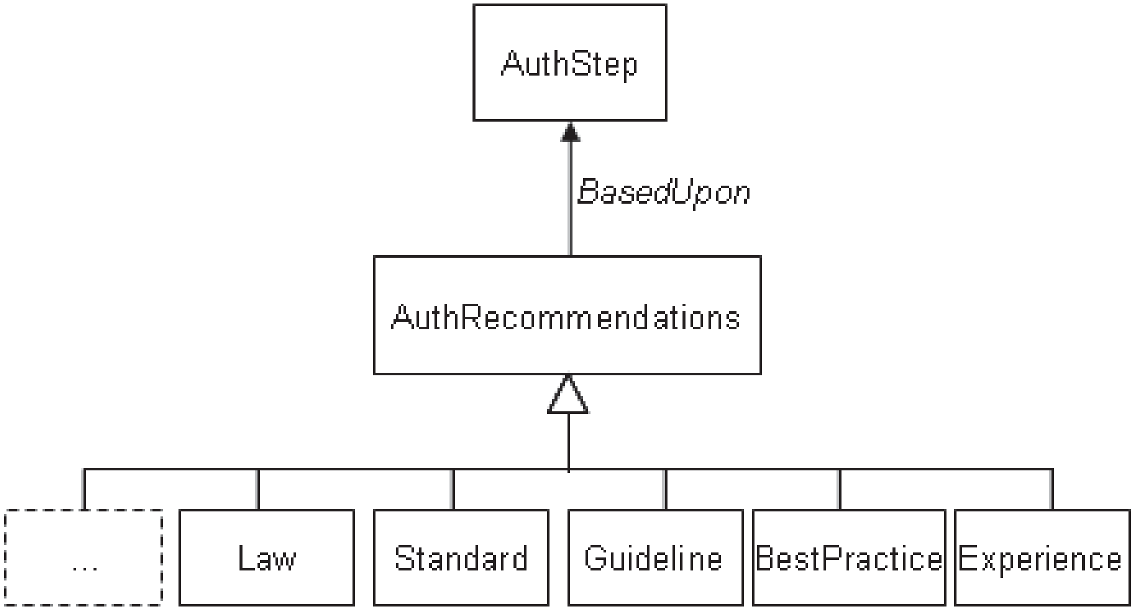
- There can be several types of ASs.
- We classify Steps based on the kind of PDI required to carry out the AS:
  - Reference Step
  - Provenance Step
  - Fixity Step
  - Context Step

Authenticity Recommendations

- Since an AS involves a decision, it is expected that it provides at least information about:
  - good practices, methodologies and any kind of regulations that must be followed or can help in the analysis and evaluation
  - possibly the criteria that must be satisfied in the evaluation

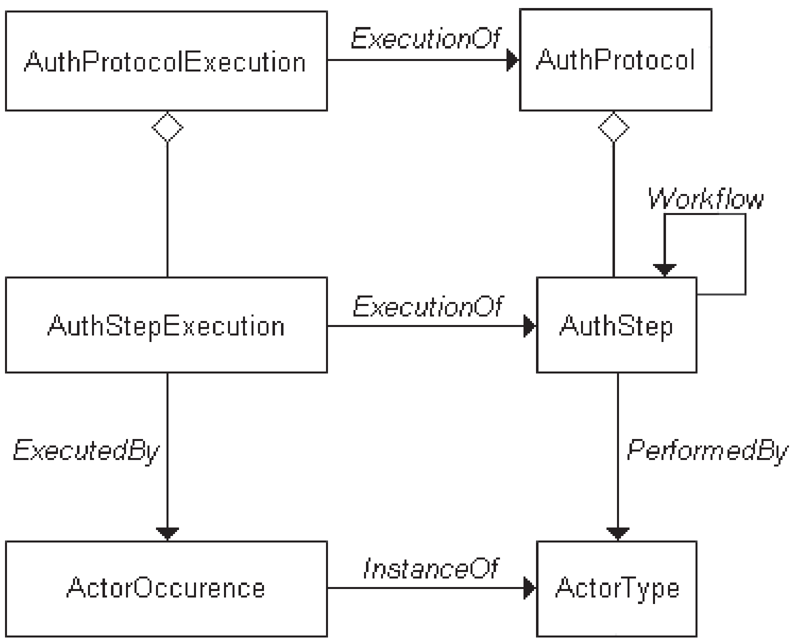
- This information is modeled under the general class of Authenticity Recommendations.

Authenticity Recommendations



Authenticity Protocol Execution

- APs are defined in order to be executed by an actor on objects belonging to a specific typology.
- The execution of an AP is modelled as an **Authenticity Protocol Execution** (**APE**).
- To execute an AP means to execute its steps.
- The execution of AS is modelled as an **Authenticity Step Execution** (**ASE**).



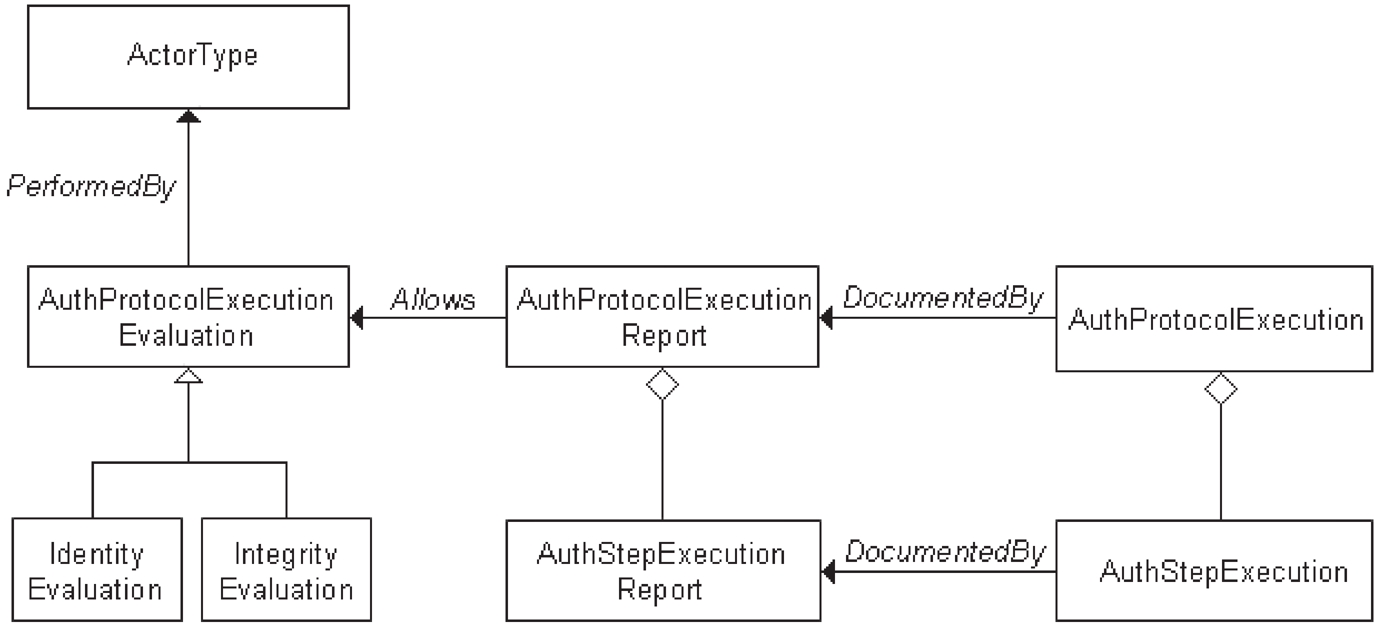
- APE and ASE are related to AP and AS via the ***ExecutionOf*** association, which also gives some information about the execution, including:
  - the actor who did the execution
  - the information which was used
  - the time, place, and context of execution
- Every ASE is executed by an **Actor Occurrence**, i.e. an instantiation of the **Actor Type**

Authenticity Report

- The outcomes of executions must by documented in order to gather information related to specific aspects of the object, e.g. title, extent, dates, and transformations

- An **Authenticity Step Execution Report** documents the step has been done – via the ***Documented By*** relation – and collects all the values associated with the data elements analysed in a specific **Authenticity Step Execution**

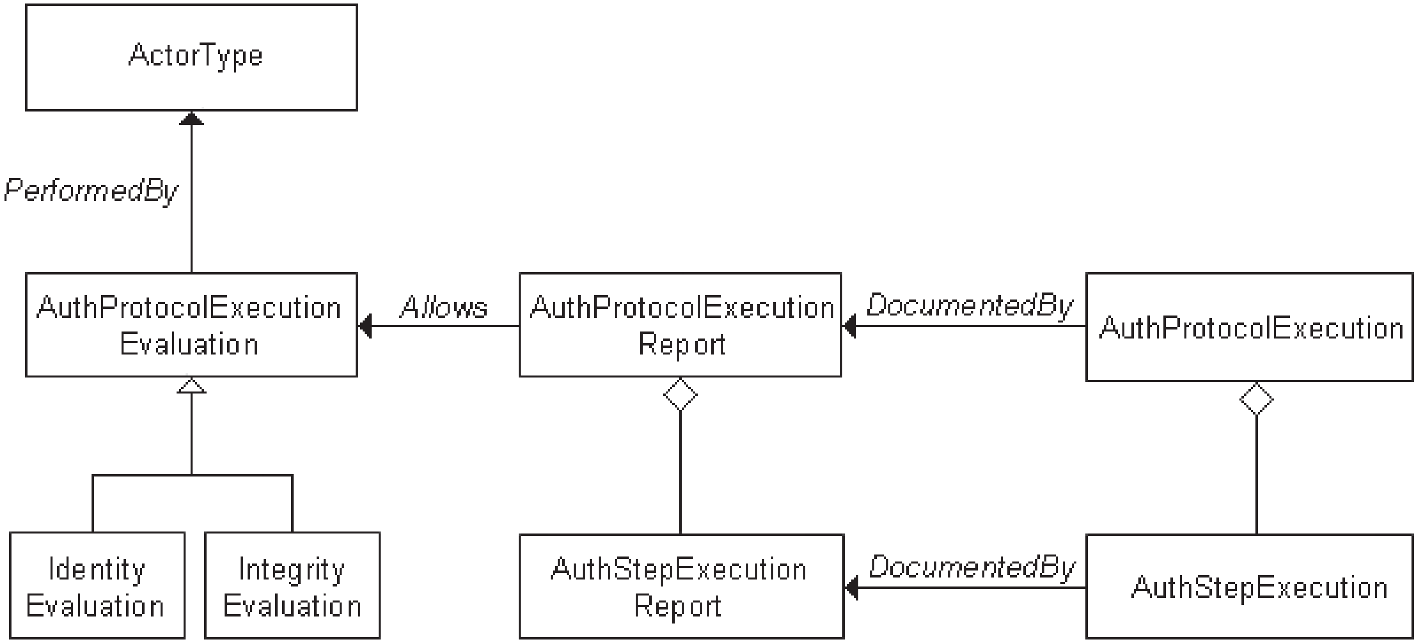
Authenticity Report



Authenticity Evaluation

- The report provides a complete set of information upon which an entitled actor (manually, or automatically by means of a metric) can build a judgment, an **Authenticity Protocol Execution Evaluation** which states an evaluation about the authenticity of the resource referring to both the identity and the integrity profile

Authenticity Evaluation

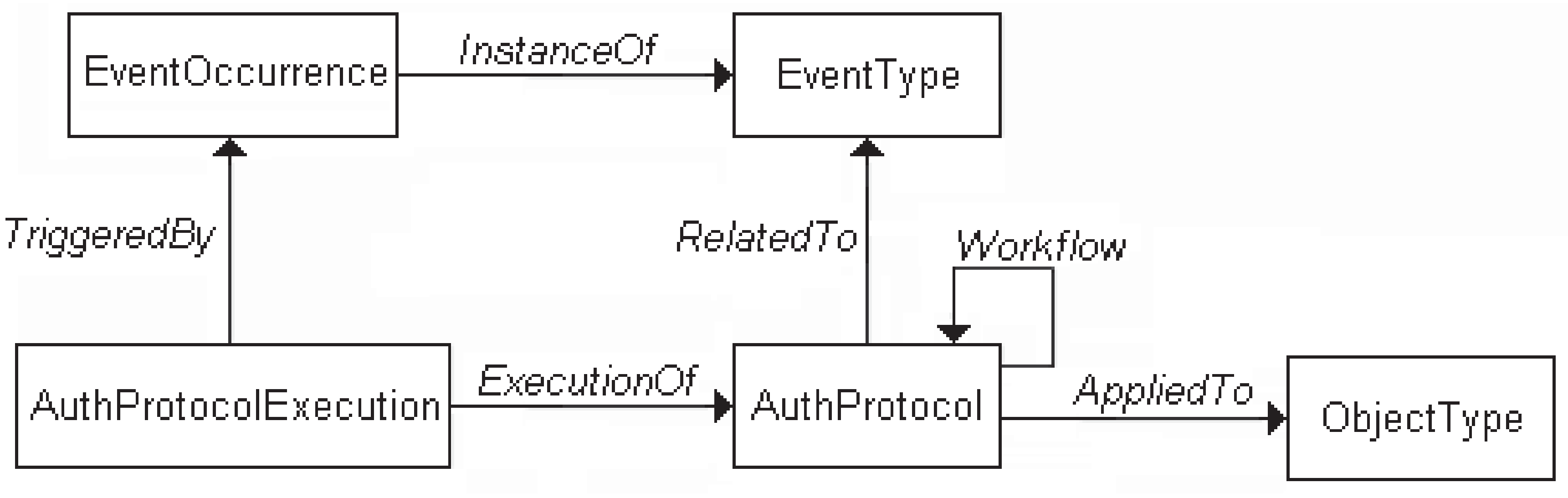


Authenticity Event

- Authenticity should be monitored continuously so that any time a resource is somehow changed or a relationship is modified an Authenticity Protocol is activated and executed in order to verify the permanence of the resource’s relevant features that guarantee its authenticity

- Any event impacting on a certain type of a resource should trigger the execution of an appropriate protocol: the Authenticity Protocol Execution is ***triggered by*** an **Event Occurrence**, i.e. the instantiation of an **Event Type** that identifies any act and/or fact related to a specific Authenticity Protocol

Authenticity Event

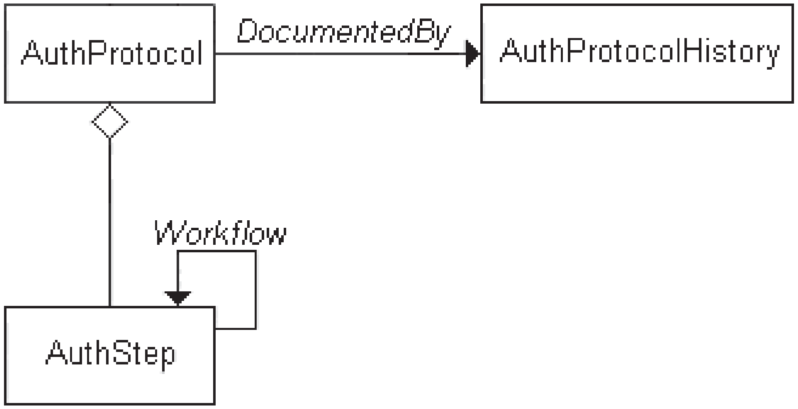


Authenticity Protocol History

- APs evolve.
- The evolution of an AP may concern the addition, removal or modification of any step making up the AP, and the change of the sequence defining the *Workflow*.
- The old AP must be retained for documentation purposes

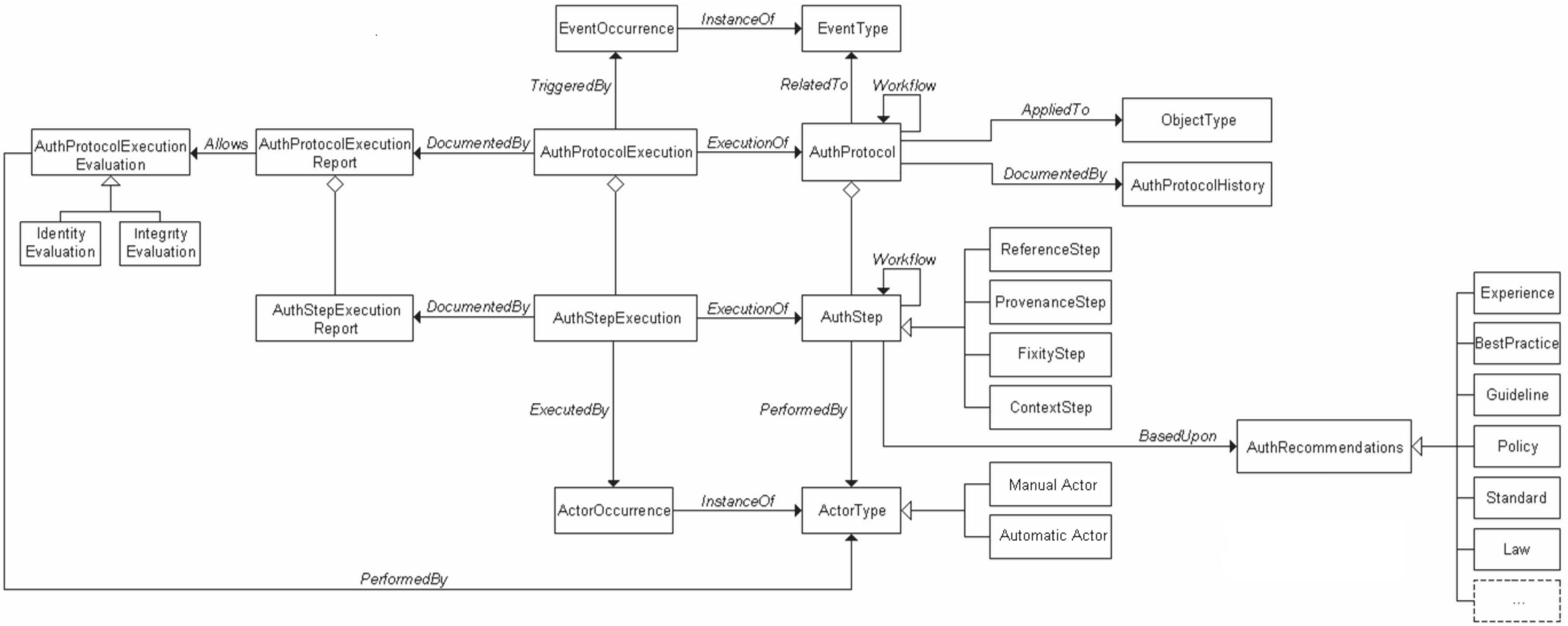
- When an AS of an AP is changed, all the active executions of the AP that include an ASE related to the changed step, must be revised, and possibly a new execution is required for the new (modified) step

Authenticity Protocol History



- The authenticity of a resource is strongly related to the criteria and procedures adopted to analyse and evaluate it: the evolution of the Authenticity Protocols over time should be documented – via the ***Documented By*** relation – in an **Authenticity Protocol History**

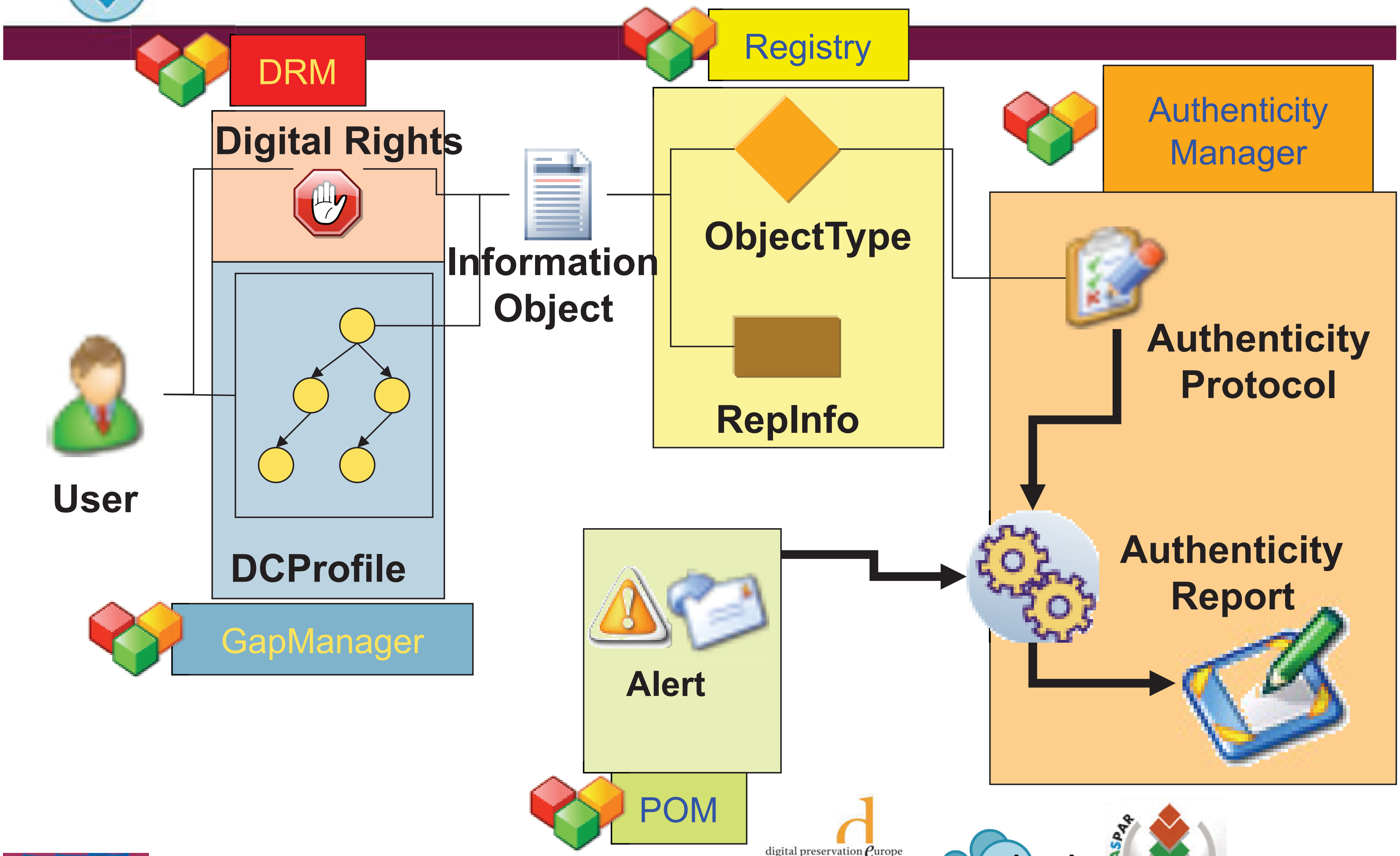
Overall Authenticity Model



EventOccurrence
InstanceOf
EventType
TriggeredBy
RelatedTo
Workflow
AppliedTo
ObjectType
AuthProtocolExecution Evaluation
Allows
AuthProtocolExecution Report
DocumentedBy
AuthProtocolExecution
ExecutionOf
AuthProtocol
DocumentedBy
AuthProtocolHistory
Identity Evaluation
Integrity Evaluation
Workflow
ReferenceStep
ProvenanceStep
FixityStep
ContextStep
Experience
BestPractice
Guideline
Policy
Standard
Law
...
AuthStepExecution Report
DocumentedBy
AuthStepExecution
ExecutionOf
AuthStep
ExecutedBy
PerformedBy
BasedUpon
AuthRecommendations
ActorOccurrence
InstanceOf
ActorType
Manual Actor
Automatic Actor
PerformedBy

Authenticity In CASPAR

(2) Authenticity within the CASPAR Framework



DRM

Digital Rights

Information Object

User

DCProfile

GapManager

Registry

ObjectType

RepInfo

Authenticity Manager

Authenticity Protocol

Authenticity Report

Alert

POM

The role of the testbed partners

IBM, IRCAM, UNESCO and ESA have been involved for the **validation of the conceptual model** and for **testing/ verifying the Authenticity Model and consequently refining it**

Outline

- Introduction
- Critical issues
- Approach
- Authenticity in CASPAR
- Conclusions

Conclusions

- CASPAR has developed a conceptual model for authenticity
  - Preliminary version to be extended and validated
- The model has been used to capture authenticity aspects in the test-beds
- Partial implementations

Thank you for your attention