

A framework supporting the shift from traditional digital publications to enhanced publications

Alessia Bardi

alessia.bardi@isti.cnr.it

Paolo Manghi

paolo.manghi@isti.cnr.it

Institute of Information Science and Technologies
National Research Council (CNR)
Pisa, Italy



Rich Internet Publication

Executable Digital Object

Live

Scientific Publication Package Document

Article of the future

EXECUTABLE PAPER

RESEARCH
OBJECT

Utopia Document

Modular Article

Scientific Compound Object

Executable Digital Object
Live
Scientific Publication Pack Document

Enhanced Publications

OBJECT

Utopia

Scientific Compound Object

Enhanced Publication (EPs)

- Enhanced publications are a new research result dissemination mean
 - try to address the novel requirements of modern science
- Enhanced publications are digital objects with:
 - an **identifier**
 - **metadata** (descriptive, provenance, etc.)
 - a **narrative part** (the description of the research in natural language)
 - a set of parts, each part represents a relevant **contextual resource** of the research, e.g. datasets, images, tables, devices, workflows.
- A part of an enhanced publication
 - may have its own identifier
 - may have its own metadata
 - may be linked to other parts via semantic relationships

Enhanced Publication Information Systems (EPIs)

- Information systems for the management of enhanced publications
- Existing EPIs are typically:
 - tailored to a specific user community
 - realised “from scratch”: functionalities that are shared across disciplines and user communities are re-implemented every time
- Effective for the specific use case
- Entail non-negligible realisation and maintenance costs

Constructing a new EPIS

I wish I had EPIS
Management
Systems

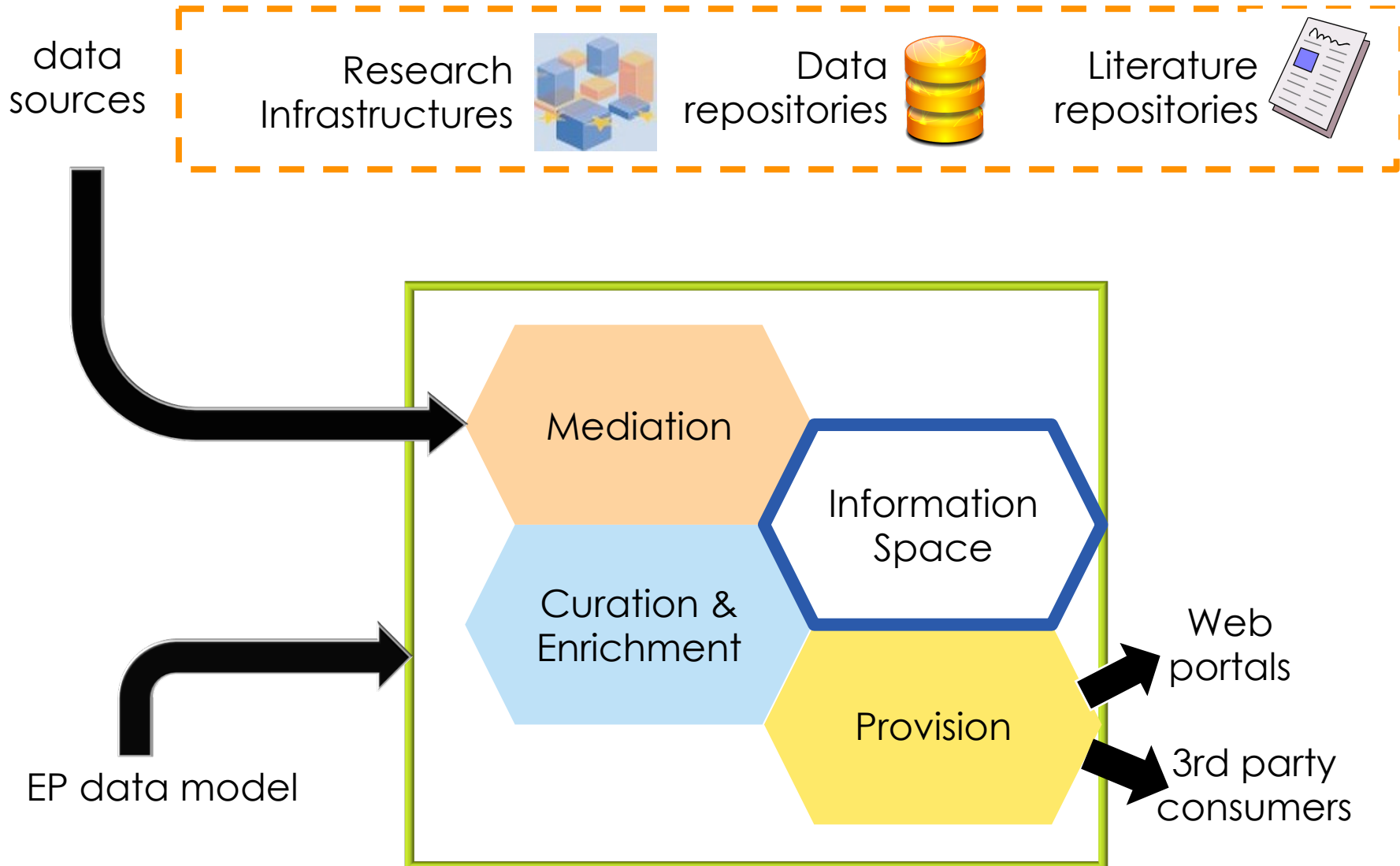
- Existing EPISs are typically:
 - hardly re-usable and configurable for other communities
- Construction from scratch is expensive

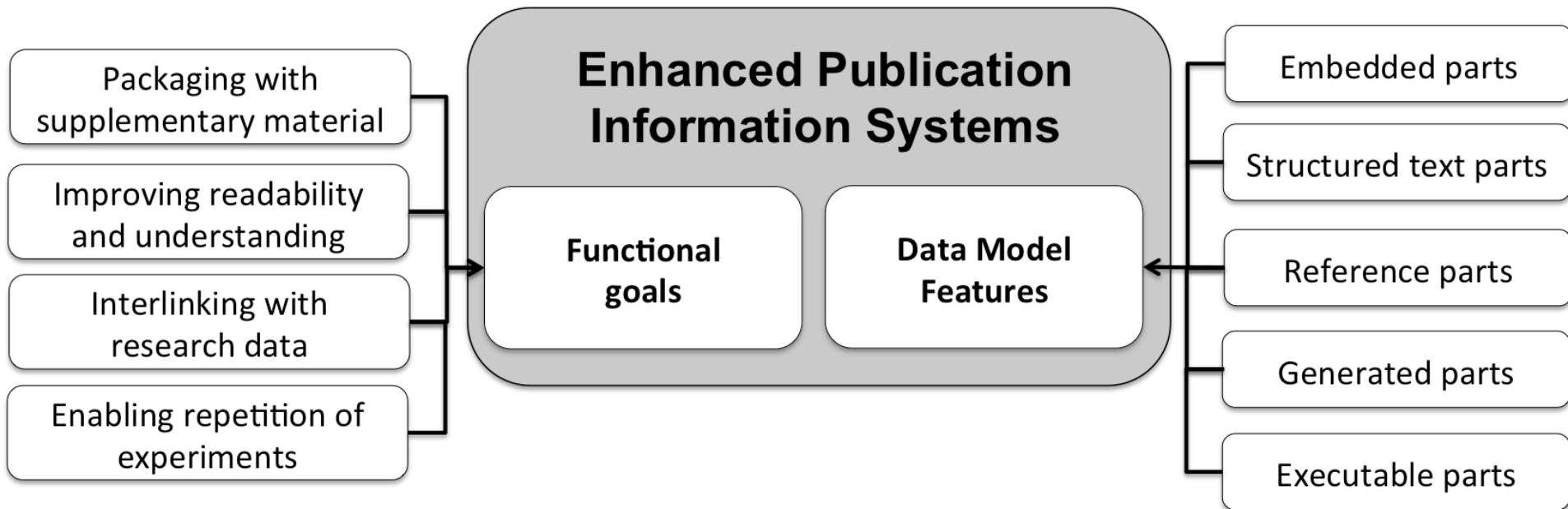


EPIS Management Systems (EPMS)

- A framework for the construction of configurable EPISs:
 - generic
 - domain independent
 - modular, customizable, extensible
- Developer tools for the definition and configuration of:
 - the EP data model i.e. the type of the entities that form an enhanced publication, in terms of structure and semantics
 - the functionalities for the consumption of EP (e.g. Linked Data, OAI-ORE and OAI-PMH exports)
 - the algorithm for content processing (e.g. mining and inference to extract knowledge from articles' full-texts)
 - the datasources to be used as entity providers

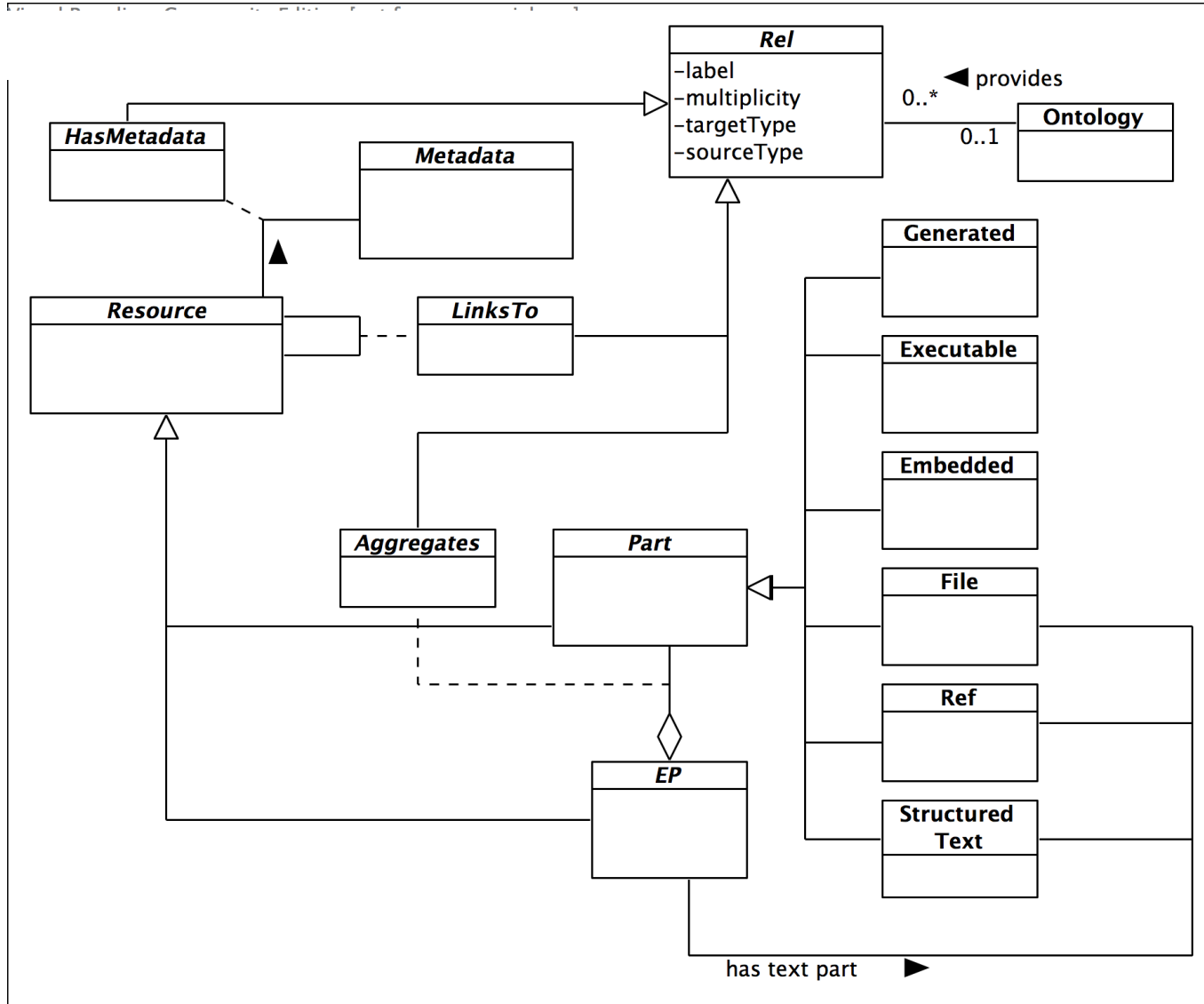
Functional areas of an EPMS





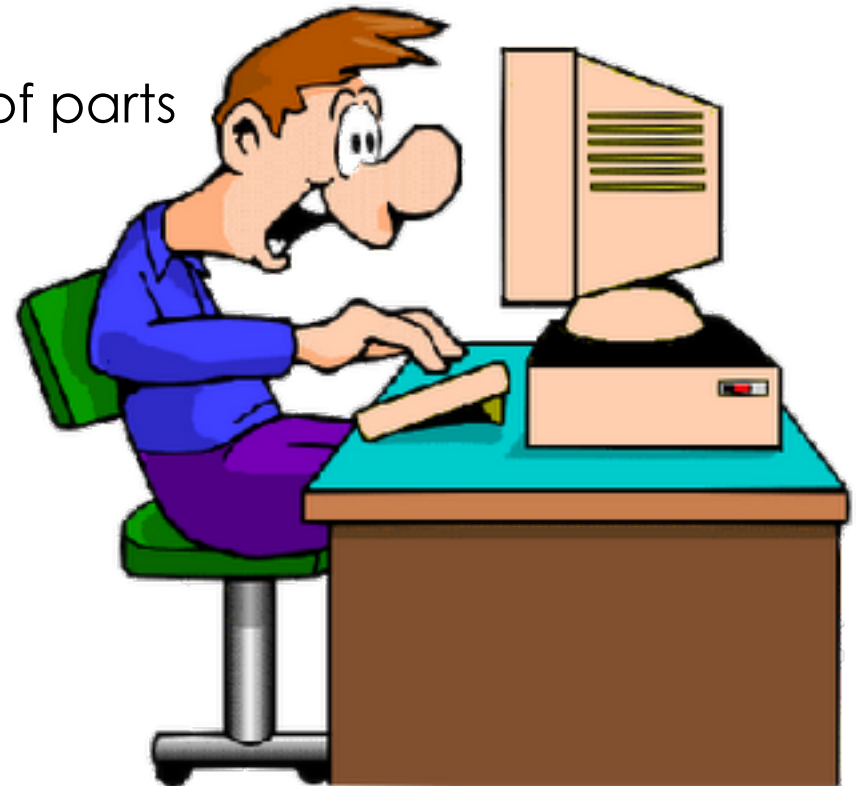
Alessia Bardi and Paolo Manghi. Enhanced Publications: data models and information systems. *LIBER Quarterly*, 22(0), 2014.

The EP “meta-model”

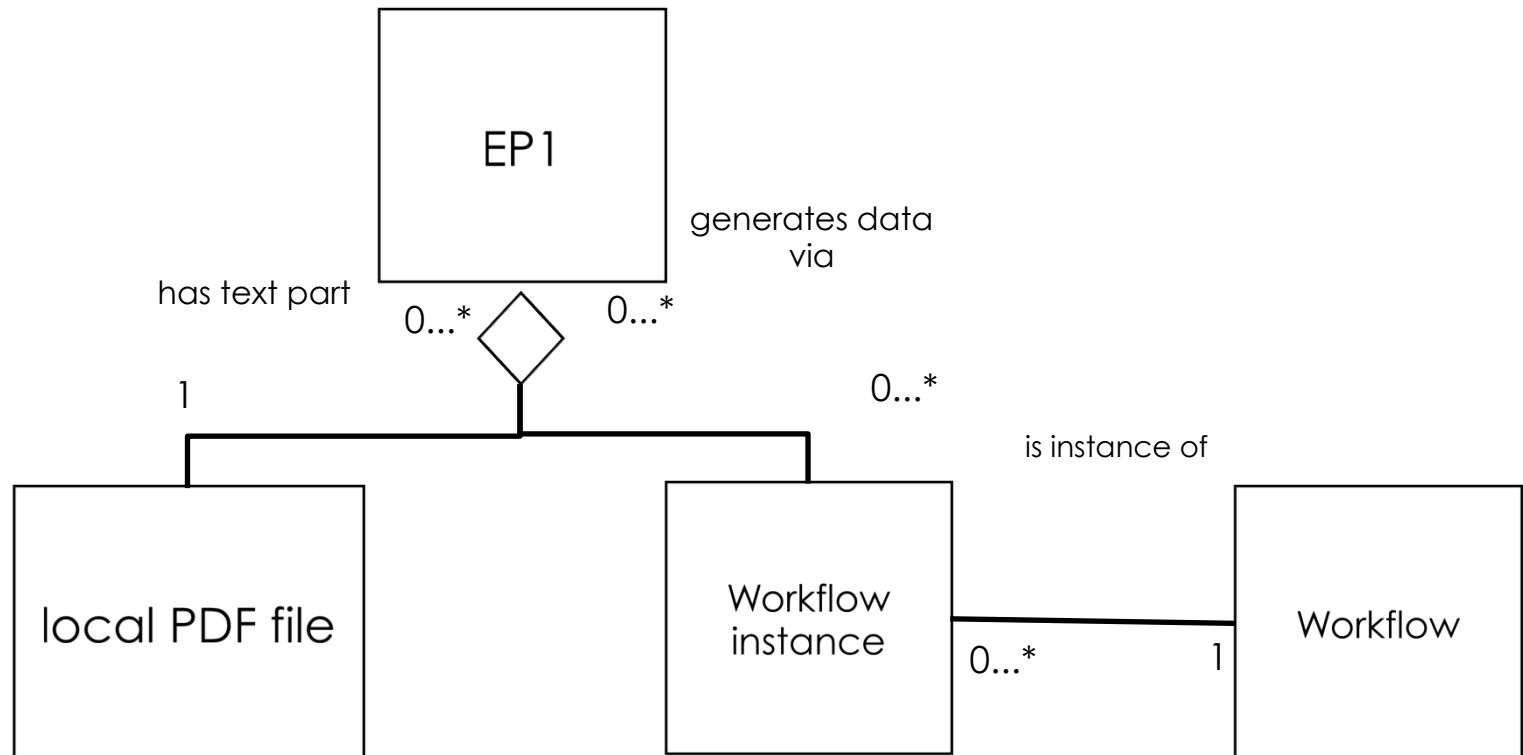


Defining your EP data model

- Language to define EP data models
 - metadata configurability (standard and/or idiosyncratic)
 - relationships configurability
 - extensible to additional types of parts
- Dev tools to configure the system based on the defined model



Example



Example

```
EP EP1 = new EP();
File Narration = new File(mime:"application/pdf");
Ref WfInstance = new Ref();
Ref Workflow = new Ref();
Metadata dcMetadata = Metadata.fromXSD("http://dublincore.org/schemas/xmls/simpledc20021212.xsd");
Metadata wfMetadata =
    Metadata.merge(dcMetadata, [{name:executionTime, type:long, repeatable:0, required:0},
                                {name:executedBy, type:String, repeatable:0, required:1} ] );
EP1.hasMetadata(targetType:dcMetadata, label:"descriptiveMetadata")
Narration.hasMetadata(targetType:dcMetadata, label:"descriptiveMetadata")
Workflow.hasMetadata(targetType:dcMetadata, label:"descriptiveMetadata")
WfInstance.hasMetadata(targetType:wfMetadata, label:"descriptiveMetadata")
EP1.hasTextPart = Narration;
EP1.aggregates(targetType:WfInstance, label:"generates data via", multiplicity:*)
WfInstance.linksTo(targetType:Workflow, label:"is instance of", multiplicity:1);
Workflow.linksTo(targetType:WfInstance, label:"has instance", multiplicity:0..*);
```

Conclusion

- High costs of EPIs are excluding many research institutions from the shift from traditional digital publications to enhanced publications
 - Main agents realising running EPIs are those who have human and economic resources to afford such an investment
 - e.g. PLoS, Elsevier, Nature
- Cutting the realisation and maintenance cost of EPIs is a first step to help institutions in their way towards the modern scholarly communication

Thank you