

# Integrating Human Science Data using CIDOC-CRM as Formal Ontology: a practical approach

Nicola Carboni & Leo Zorc

# Introduction

- Introductions
- Expectations
- Overview

**It's a match**  
but tell us more about  
yourself,  
your project and  
what you expect to learn  
in 5 days



# My daily job



# My daily job

# Integrating Human Science Data using CIDOC-CRM as Formal Ontology: a practical approach

*Working on different datasets...*

# Integrating Human Science Data using CIDOC-CRM as Formal Ontology: a practical approach

*...with very bad data quality...*

# Integrating Human Science Data using **CIDOC-CRM** as Formal Ontology: a practical approach

*...using a huge but public standard...*

# Integrating Human Science Data using CIDOC-CRM as Formal Ontology: a **practical approach**

*...in our daily jobs.*

# Why photo archives?



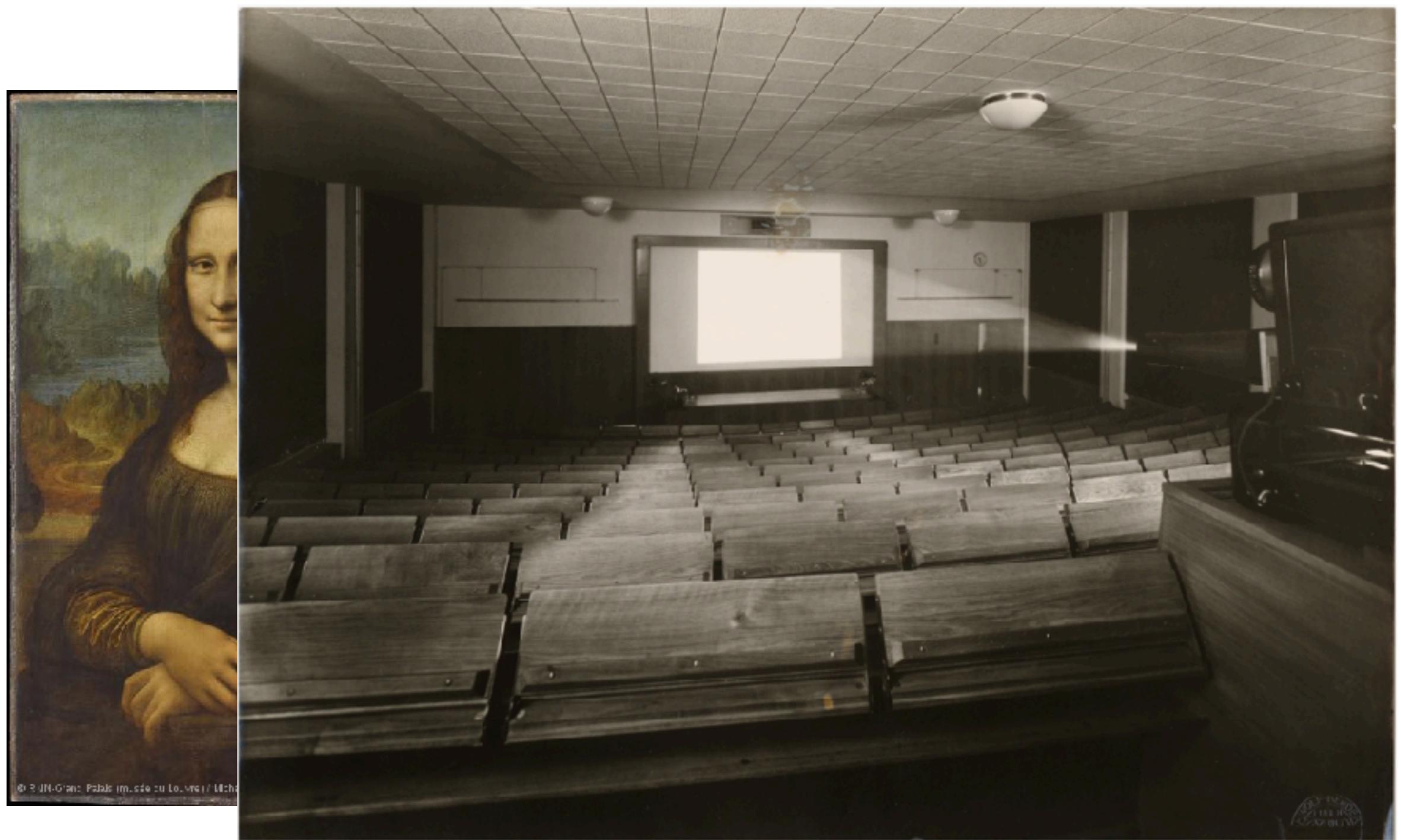
# Why photo archives?



# Why photo archives?



# Why photo archives?



# Why photo archives?



# Overview

|             | Monday                                 | Tuesday                               | Wednesday               | Thursday | Friday |
|-------------|--|---------------------------------------|-------------------------|----------|--------|
| 1st Session | Intro<br>-<br>Formal<br>Ontologies     | CIDOC-CRM                             | (Teaser<br>Session)     | Karma    | 3M     |
| 2nd Session | Formal<br>Ontologies<br>-<br>CIDOC-CRM | CIDOC-CRM<br>-<br>other<br>Ontologies | Workflows<br>-<br>Karma | 3M       | Finale |
| 3rd Session |  | Hands-On                              |                         |          |        |



# CIDOC-CRM

## Introduction

# What is CIDOC-CRM?

- CIDOC stands for International Committee for Documentation of ICOM the International Council of Museums
- CRM is the Conceptual Reference Model of CIDOC
- in short: a model to describe cultural heritage that started in paper format in the late 80s and went through several technological advancements

# In more detail

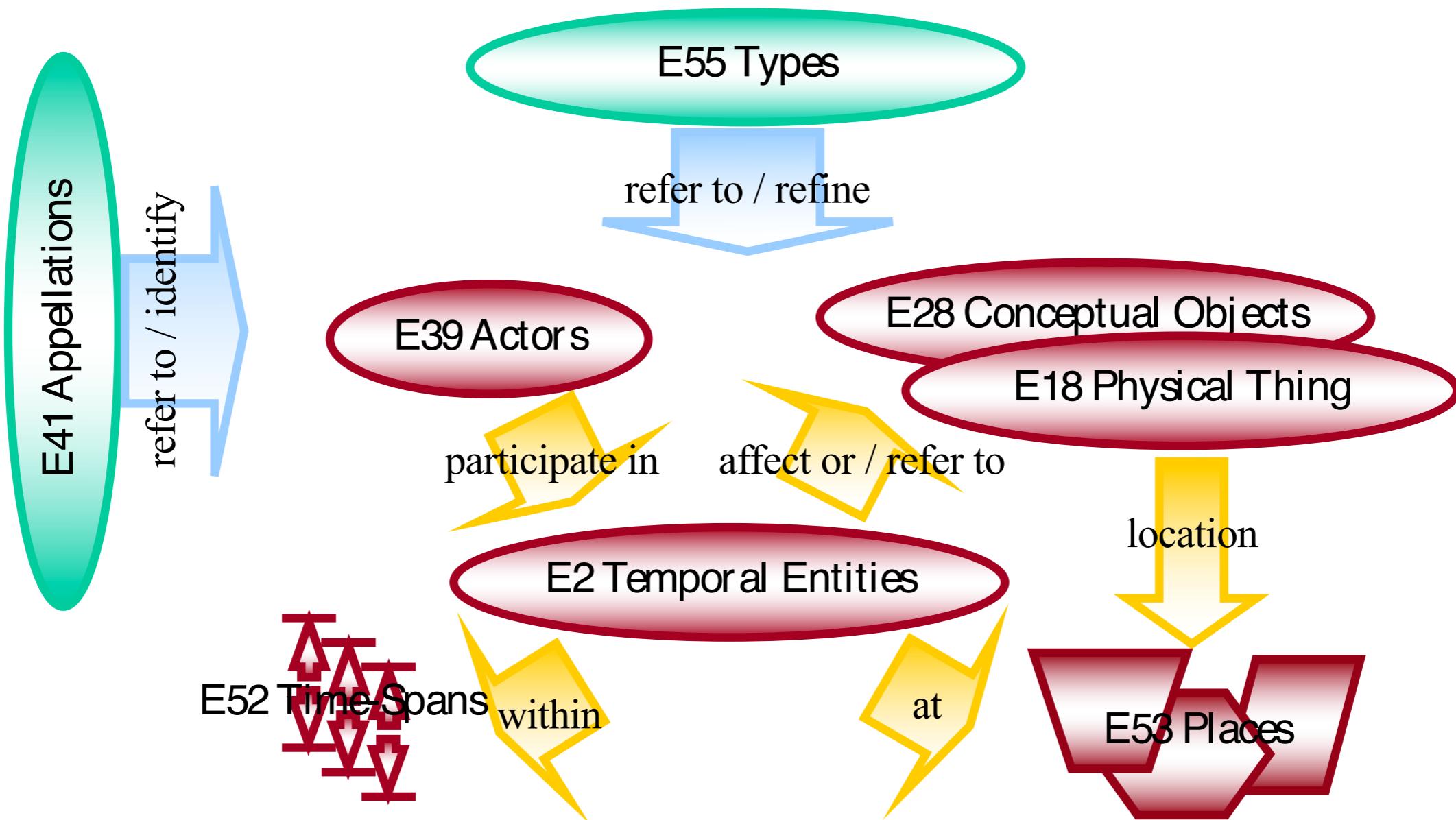
- SIG, the Special Interest Group, was tasked to standardise the model in order to make it ISO compliant.
- Nowadays the SIG is the main body and discussion forum of CIDOC-CRM, that works on the content, the publication as well as the technology involved.
- Main objective: connect data from the cultural heritage domain coming from different institutions in order to be able to make an informed scientific argument.

# In more detail

- capture not only descriptive data but also the underlying semantics
- organized in a formal structure, with „clearly“ defined concepts and relationships (concepts are often called classes and relationships properties)
- understood by humans and machines, therefore it needs to involve computer scientists and domain experts
- event based model
- connecting information about things

# The CIDOC CRM

## Top-level classes useful for integration

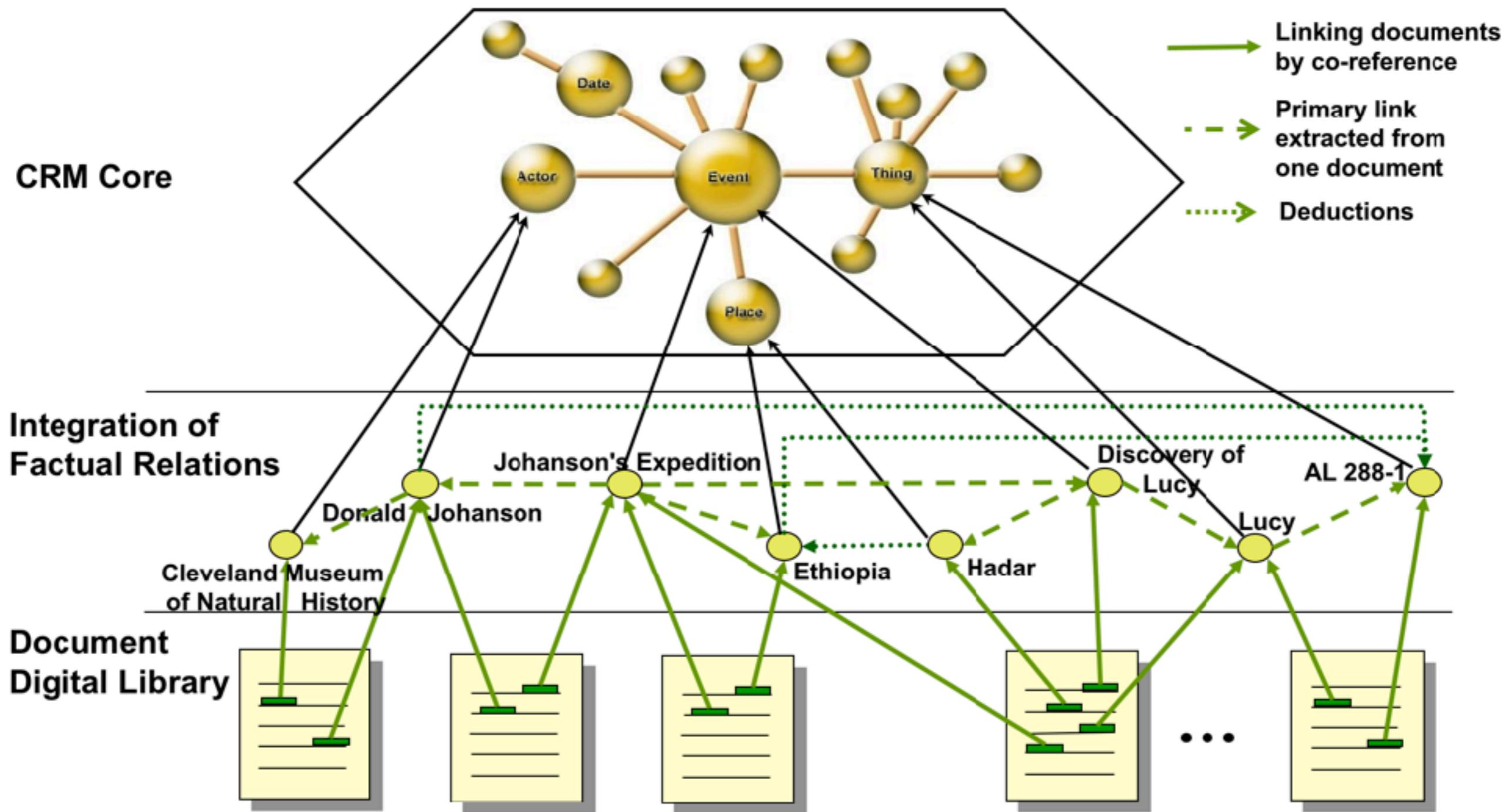


# The CIDOC CRM

## The types of relationships

- ◆ **Identification** of real world items by real world names
- ◆ **Observation** and **Classification** of real world items
- ◆ **Part-decomposition** and structural properties of Conceptual & Physical Objects, Periods, Actors, Places and Times
- ◆ **Participation** of persistent items in temporal entities
  - creates a notion of history: “world-lines” meeting in space-time
- ◆ **Location** of periods in space-time and physical objects in space
- ◆ **Influence** of objects on activities and products and vice-versa
- ◆ **Reference** of information objects to any real-world item

# The CIDOC CRM Documents and Factual Knowledge



Nicht sicher | www.cidoc-crm.org/Entity/E21-Person/Version-6.2

ICOM International council of museums CIDOC International committee for documentation

# CIDOC CRM CONCEPTUAL REFERENCE MODEL

Home The Model Activities Resources Community News

Search

## Resources

### MODEL RELATED MATERIAL

- Versions of the CIDOC-CRM
- References
- Presentations
- Technical Papers
- Tutorials
- Critics
- Important Theories
- Publications
- Mappings
- Compatible Models
- Translations
- FAQ

### MEETING PRODUCTS

- Minutes
- Issues
- CRM SIG archive
- Meeting Contributions

## E21 Person In Version 6.2

**Subclass of :** [E20 Biological Object](#), [E39 Actor](#)  
**Superclass of:** No superclasses found  
**Scope Note:**  
This class comprises real persons who live or are assumed to have lived.  
Legendary figures that may have existed, such as Ulysses and King Arthur, fall into this class if the documentation refers to them as historical figures. In cases where doubt exists as to whether several persons are in fact identical, multiple instances can be created and linked to indicate their relationship. The CRM does not propose a specific form to support reasoning about possible identity.

**Examples:**

- Tut-Ankh-Amun
- Nelson Mandela

**In First Order Logic:**  
 $E21(x) \supset E20(x)$ ,  $E21(x) \supset E39(x)$

**Properties:**  
[P152 has parent \(is parent of\) : E21 Person](#)

**Inherited properties:**

- [List of properties](#)

**Inherited references:**

- [List of references](#)

Intro Learning Material Look up