24 Weeks at the GRI

Mark Pyzyk
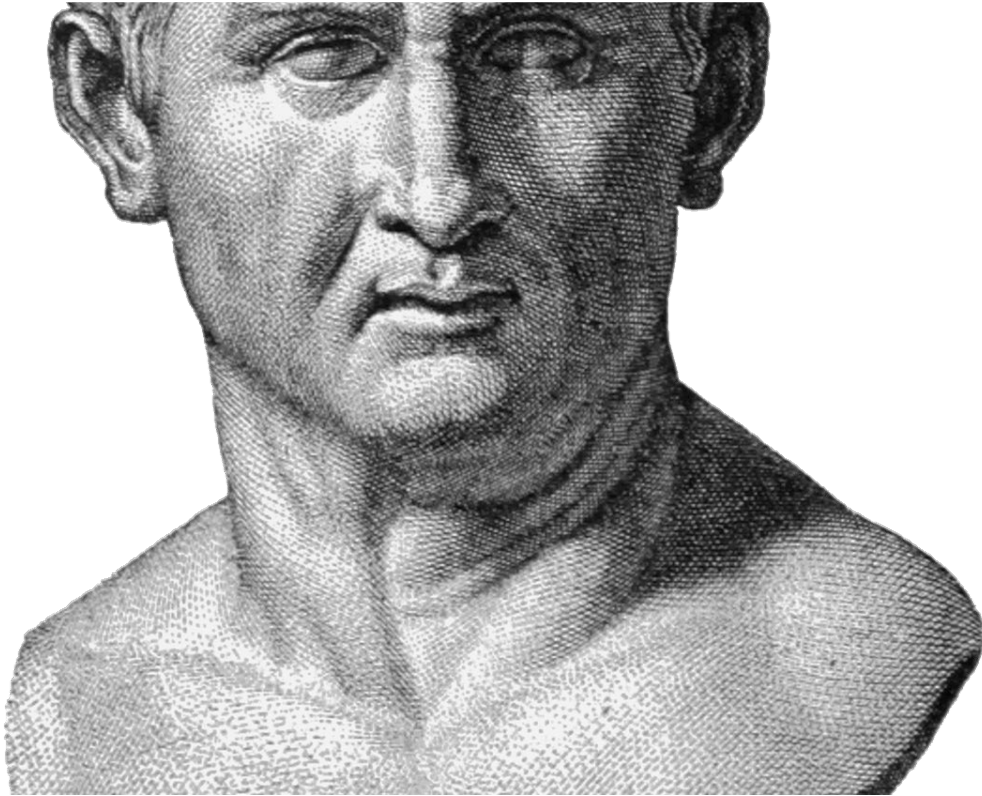
Data Analyst and Facilitator
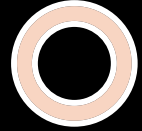
Getty Research Institute

October 2024

Terminology Matters

Marcus Tullius Cicero, on matters of terminology:

"A great many accomplished students of Greek learning were unable to share their acquisitions with their fellow-citizens, on the ground that they doubted the possibility of conveying in Latin the teachings they had received from the Greeks. In the matter of style however I believe that we have made such progress that even in richness of vocabulary the Greeks do not surpass us."

Cicero, *De Natura Deorum* 4 (trans. Miller)

Welcome to the GRI!

Getty Vocabularies Project

•"Catherine wheel or rose window? Cohuacalco or Snake House? Gaochang or Qara-Hoja? Getty Vocabularies are structured resources for the visual arts domain, including art, architecture, decorative arts, other cultural works, archival materials, visual surrogates, and art conservation. Compliant with international standards for structured and controlled vocabularies, they provide authoritative information for catalogers, researchers, and data providers."

Getty Provenance Index

•A large-scale database of provenance records regarding the sale and inventory of works of art. It digitizes many thousands of records (22,000 sales catalogues, for example), containing information on millions of records (1.8 million sales, for instance).

•Currently, in the midst of a large-scale transformation of data from a traditional relational database model to a linked open data (LOD) model.
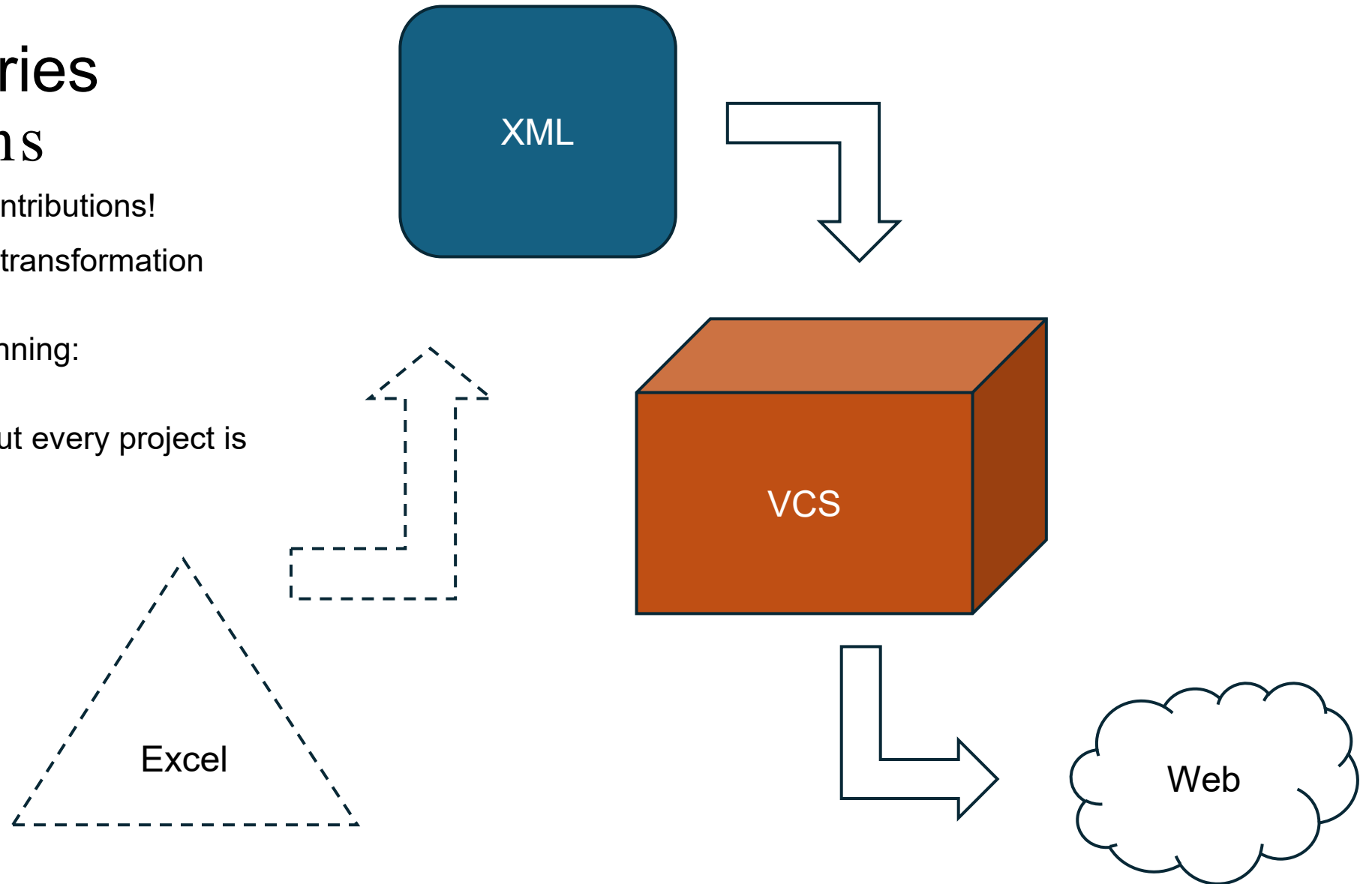
So, what are you working on?

1. Vocabularies
Contributions

- We welcome your contributions!

- Straightforward data transformation pipeline

- Except the very beginning: Contributions

- Excel is standard—but every project is a little different

XML

VCS

Excel

Web

Blacksmithing Data

- Hammering Excel data into well-formed XML

- Problems (i.e., "did you, the contributor, do one of the following?"):
    - Leave a space after a column name?
    - Slightly misspell the name of the column?
    - Include more of a value than is allowed in our presets?
    - Include values in your data that are not already present in Getty Vocabularies?

- Problems (i.e., "are we, the staff at Getty Vocabs, doing this the best way?"):
    - Our smiths: Antonio Beecroft, Jonathan Ward
    - Uh-oh. Unsynchronized data
    - Whoops! Non-version controlled, bespoke software

- Solutions
    - Github
    - Standardized documentation, workflow

2. CDWA on Quire

- Categories for the Description of Works of Art (CDWA)

- A basic resource for scholars in the art historical, curation, and museums space

- Updating the platform:
  - HTML to Quire
  - Website to online publication
  - Increasing legibility and navigability

- Emily Benoff did 99.9% of the work on this transformation

CDWA is live, right now

- Please visit the site!
- Some features missing, but COMING SOON
- Let us know what you think! mpyzyk@getty.edu

3. Getty Provenance

- "Provenance Index Remodel"
- Large-scale transformation of provenance data
- STAR data management system
- Arches

Van Gogh, Factories of Clichy

Object Title ⌄

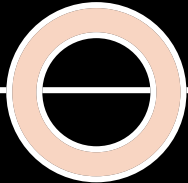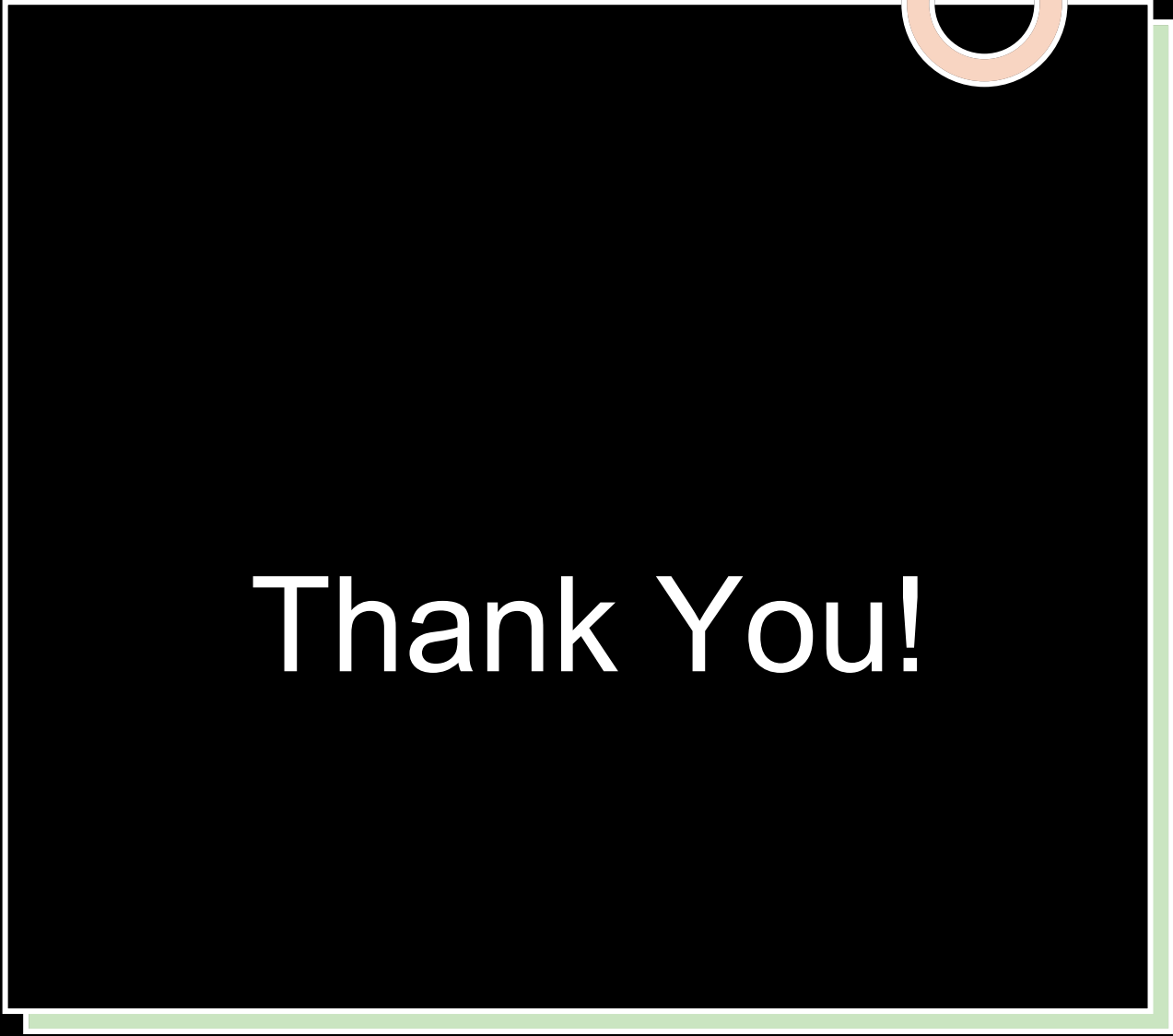|	
Title	Factories at Clichy
Title Type	titles (general, names)
Title Language	English (language)
Source Document for Title	Knoedler Stock Book 10, Page 20

Object Type ⌄

Type	paintings (visual works)

Statement about Object ⌄

Statement	P [panel]
Statement Type	materials/technique description
Statement Language	English (language)
Source Document for Statement	Knoedler Stock Book 10, Page 20

What's Next?

Thank You!