Using Digitized Medieval Manuscripts

Cross-collection comparisons, annotations, tools, interoperability and related matters

John Haeger
for the Mellon Research Cluster Workshop
18 August 2011
A bit of background…

- **An assortment of digitization projects:** *Roman de la Rose Digital Library* and *Parker on the Web*, etc. (ca. 1996-)

- **Inter-project communication?** Mellon Foundation “All Projects Meeting” (2006)

- **Inter-project operation?** *Rose* and *Parker* interoperability “demonstration” (2007)

- **Scholarly uses of digitized materials:** Paris Workshop on Uses of Digitized Collections of Medieval Manuscripts and Interoperability (2010)
Until Paris…

- Manuscript digitization was target-driven, not use-driven.
- For each project, a self-contained environment was built.
- The array of environments was unfriendly to comparative studies.
- “Tools” evolved in environments of their own.
- Ingesting data into tools required off-line transactions.
- Users worked with digitized materials as they always had -- by hand.
After Paris: rethink the status quo

- Job 1: Re-examine costs, and sustainability
- Job 2: Redefine interoperation
- Job 3: Re-imagine manuscript scholarship when the objects of research are available in digital form
Job 1 - Costs and Sustainability

- Building self-contained “silos” is expensive
- Maintaining silos is even more expensive
- Maintaining code is costlier than reformatting data
- Ergo…
Job 2- Interoperability

- Cross-silo transactions must be designed individually for each silo pair
- Using a resource outside its silo is messy business
- But if
  - repositories were *just* repositories…
  - …and tool-equipped users inhabited spaces *outside* repositories…
  - …and tools and repositories used/recognized a *shared model* to represent [manuscript] data
- A *system* of repositories and tools could function on the basis of *common* web services
Job 3 – Scholarship redefined

- Then...
- Users could work with digitized manuscripts as digital documents
- Any tool could operate on any manuscript in any repository
- Cross-repository comparisons could be made easily
- Annotations (etc.) could be aligned with the manuscript/page/region annotated
- Annotations (etc.) could be reused by third parties
- Links and citations could be durable
A interim approach for 2010-2013

Research projects

- coordinated at JHU and the University of Toronto
- ...addressing consequential questions
- ...methodologically committed to the analysis of data in digital form
- supported by two Mellon grants

Technical work

- coordinated at Stanford
- ...to build the infrastructure for interoperability
- ...and create a “sandbox” environment for use by the research projects
- supported by a third Mellon grant
### Dramatis personae

**7 repositories:**
- Bibliothèque nationale de France
- British Library
- e-codices
- Roman de la Rose Digital Library
- Oxford University Libraries
- Stanford University Libraries
- ...and more to come

**5 tool projects:**
- Digital Mappaemundi
- MARGOT Annotation Tool
- SALSAH
- TILE
- T-PEN

**9 research projects:**
- in a cluster overseen by *Stephen Nichols* at The Johns Hopkins University
  - Medieval Latin Manuscript Transmission in a Digital Environment
  - Machaut in the Book
  - A Multidimensional Database for Vernacular Bible Production in the Middle Ages
  - Dynamic Modeling and Differential Imitation in 14th-century Manuscripts
- in a cluster overseen by *Alexandra Gillespie* at the University of Toronto
  - Palaeographical Cruxes in Old English Manuscripts
  - Manuscripts of Aelfric’s Catholic Homilies
  - Networks of Book Makers, Owners and Users in Late Medieval England
  - Parker’s Scribes
  - The Writing of Petitions in Later...
Features of the interoperable infrastructure

- Eliminates obligatory co-location of data resources and applications
- …but keeps both resources and applications web-based
- Resources and applications share a single data model
- Establishes a distinct and persistent ID for each resource
- Links the product of applications (“annotations”) to their “targets” persistently
- Web services enable automatic, real-time interaction between repositories and presentation/annotation systems
- Accommodates access limitations based on
Today…

USER

Repository discovery apparatus

Repository metadata

Repository images

Repository image viewer

TOOL #1

Mediated offline transaction
In an interoperating world…
Next up:

- More about the research projects (Steve and …)
- More about the interoperable data model, the Medieval Manuscript Manifest and web services (Ben and Mark)
- An introduction to DM and image annotation (Martin and Shannon)
- An introduction to T-PEN and text transcription (Jim)
- Lunch….