

# **Philadelphia Architects and Buildings Project**

Walter Rice  
R&R Computer Solutions  
For the Athenaeum of Philadelphia

November 7, 2007

# Philadelphia Architects and Buildings Project

- Regional initiative
- Improves access to information on the built environment
- Web-based
- Image-rich

# Basic Components

- Building Information Database
- Architect Biographies
- Digital Image Library
- The Future – Geographic Information Layer

# Access vs. Preservation

- Preservation is not the top priority – Access is most important goal
- Focus first on rich, existing data sources
- Images are scanned to increase accessibility of archives – preservation is a side benefit

# Sources of Data

- **Architectural Archives**
  - MARC Records
  - Electronic Finding Aids
- **Private Collections**
  - Insurance Policies and Surveys
  - Building Owner Drawings and Photographs

# Sources of Data (cont'd)

- **Historic Building Registries**
  - National Register of Historic Places
  - Pennsylvania Cultural Resources Database
  - Philadelphia Register of Historic Places
- **Building Surveys**
  - Historic American Building Survey (HABS) / Historic American Engineering Survey (HAER)
  - Census of Stained Glass Windows

# Challenges of Using Extant Data Sources

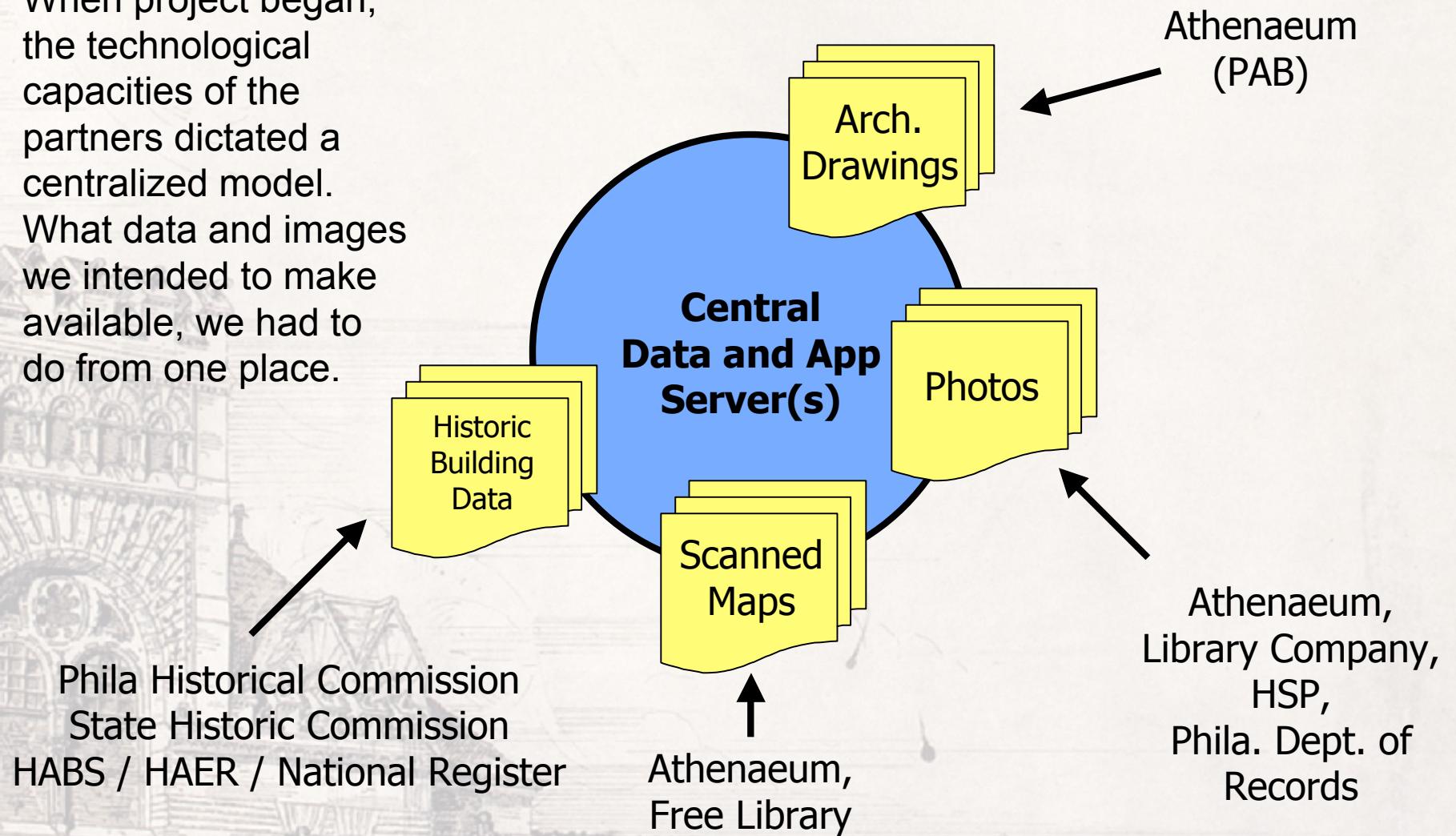
- Data accuracy and timeliness questionable
  - particularly true for governmental information on historic resources in the United States
- Information from other sources needs to be matched up with PAB records, almost always by hand
- Difficulty updating imported data without reintroducing previous errors

# Pragmatics 101: Metadata

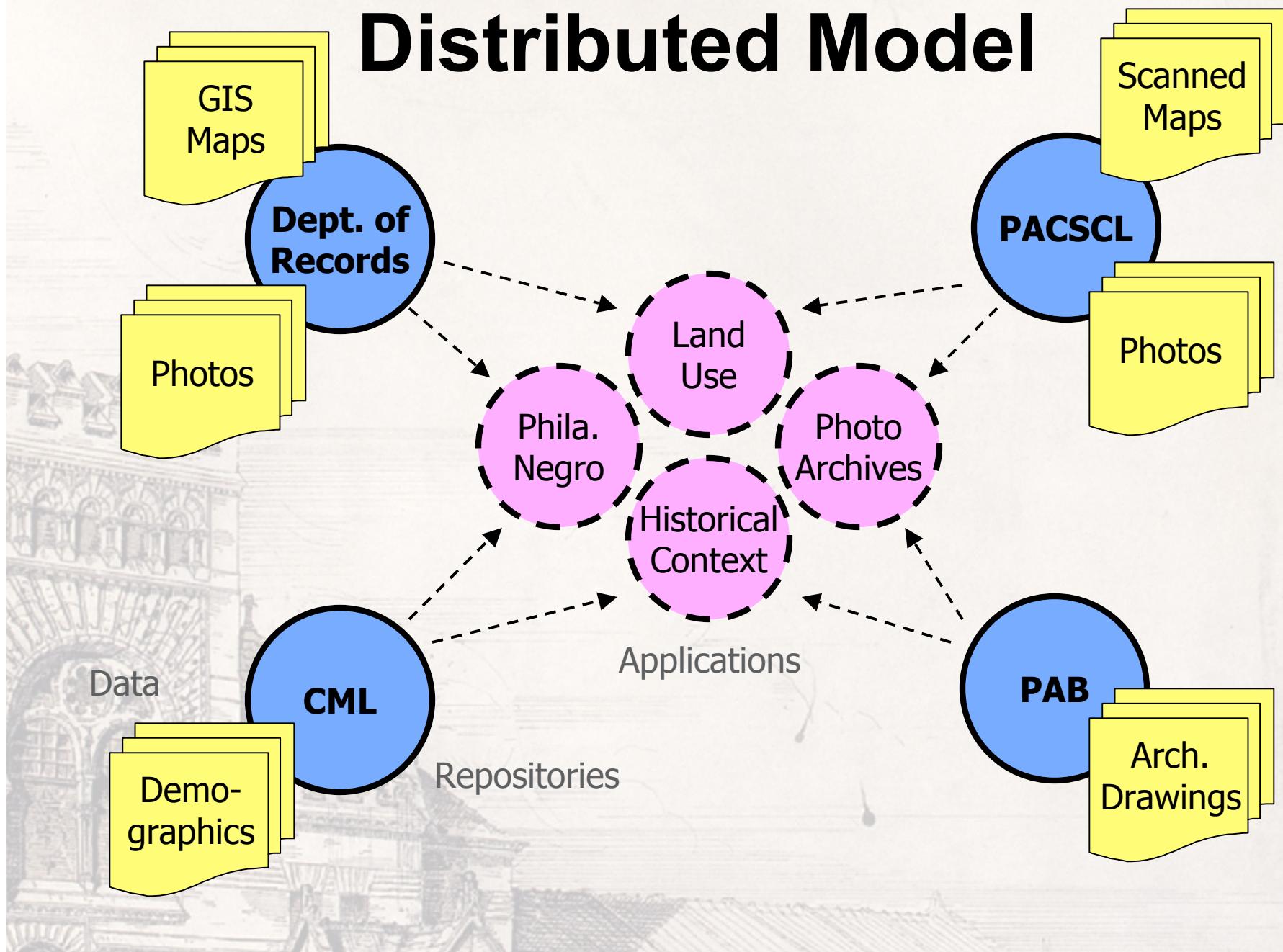
- Gathered data in very different formats, considering different information
- Necessary structure for building history not provided by any developed metadata standard
- Metadata inspired by
  - Existing legacy datasets
  - Standards: Dublin Core, MARC
  - Primarily access needs

# Centralized Model

When project began, the technological capacities of the partners dictated a centralized model. What data and images we intended to make available, we had to do from one place.



# Distributed Model



# Distributed Model

- Data stored by each institution (or small group of institutions)
  - or in a distributed set of centralized systems
- Can be built using existing systems, or systems designed for other purposes
- Scales by institution

# Challenges

- Metadata Standards – Dublin Core?
- Geospatial and Temporal Metadata
  - Map extents, feature points
  - Date and time, both instant and spanned
- Precision, Specificity, and Uncertainty
  - “702 Chestnut St.” vs. “7<sup>th</sup> & Chestnut”
  - “Chestnut St.” vs. “Philadelphia”
  - “1902” vs. “c. 1900” vs. “Early 20<sup>th</sup> Century”

# Challenges (cont'd)

- Aggregation / Correlation / Change
  - Changing place names, boundaries, addresses, and uncertain correlations (city directories)
  - Re-aggregating data sets (e.g., census data) to provide useful comparisons over time
- User Interface
  - Providing useful interaction both to experience GIS experts, but also to scholars and hobbyists, and even tourists.

# GeoHistory Next Steps

- Technology Infrastructure Devel.
  - Conceptual foundation, model
  - Time-enable geographic tools
  - Geo-temporal metadata tools
- Data Development
- Training, Evangelism, Support

# Sustainability

- Three main sources:
  - Grant funding
    - Data/image expansion
    - Congruent projects
  - Subscriptions and reproductions
    - System maintenance, hardware,
    - Internal collections processing
  - Digitization services

# **Digitization Services**

- Regional Digital Imaging Center
  - Digitization of large format and delicate items
  - Outsourced options for online content delivery
- “Relevant content” projects
  - Synergy provides savings to client institution and grows PAB

# Conclusion

- Content
- Metadata and Interoperability
- Sustainability

[www.PhiladelphiaBuildings.org](http://www.PhiladelphiaBuildings.org)

[www.PhiGeoHistory.org](http://www.PhiGeoHistory.org)

