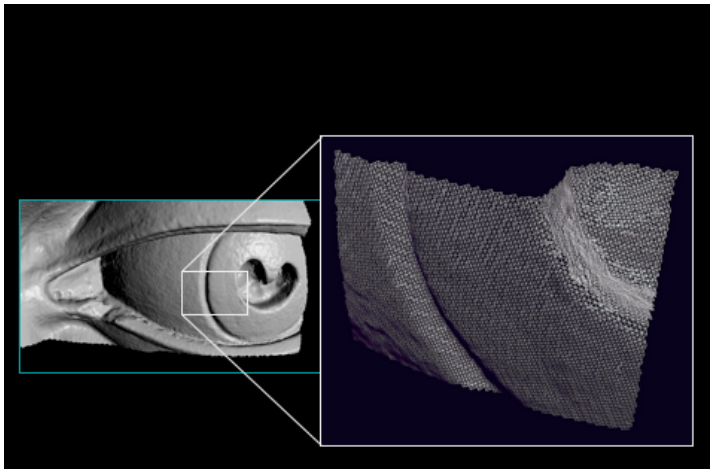
Digital Representation of Cultural Heritage Material
- new possibilities for enhanced access -



Stephen M. Griffin
Program Director
National Science Foundation (USA)

Recovering Lost Writings

OCR of non-Roman Characters and Cursive Scripts

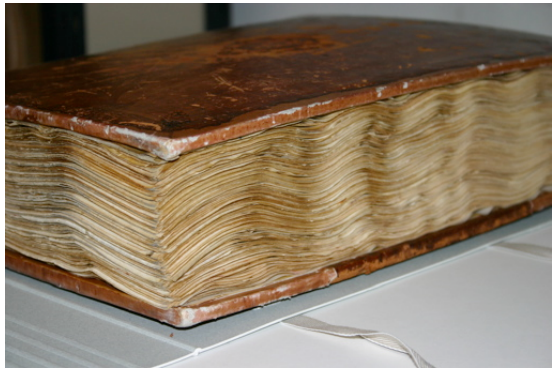
Reconstruction of Incomplete Artifacts

Recovering Lost Writings

EDUCE: Enhanced Digital Unwrapping for Conservation and Exploration

W. Brent Seales
University of Kentucky

A Many Step, Meticulous Process



Digital Restoration Using Volumetric Scanning

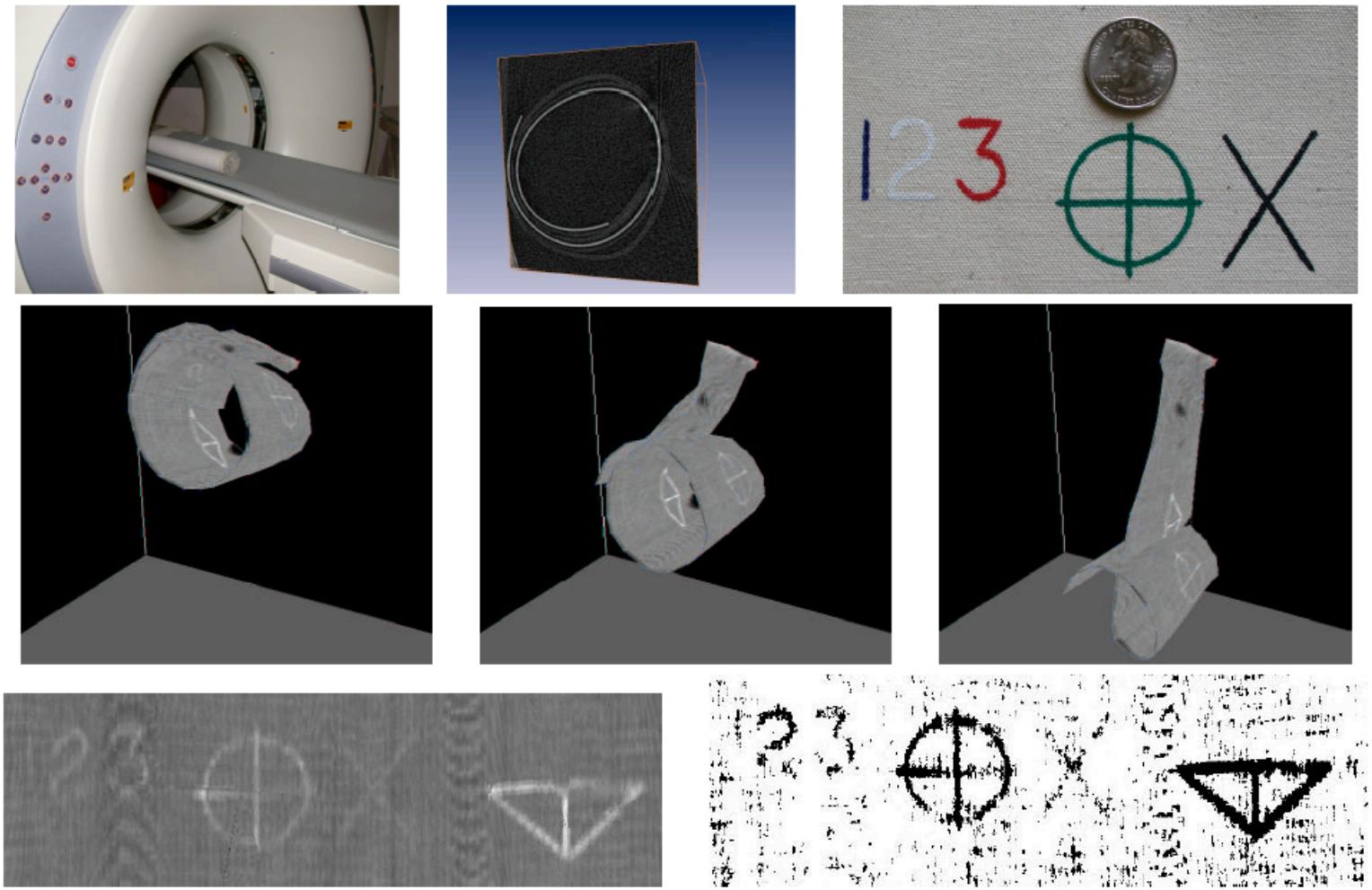


Figure 3: Top (a): Rolled canvas strip in CT scanner. (b): One slice from CT-scan of canvas strip. (c): Digital photo of original, unrolled test strip. Center (d,e,f): Three progressive views of the digital unwrapping. Bottom (g): Unrolled result. (h): Thresholded result.

Source: Seales & Lin, JCDL 2004

Digitally Unrolling Scrolls

Applications: Egyptian Papyrus (layers fused)

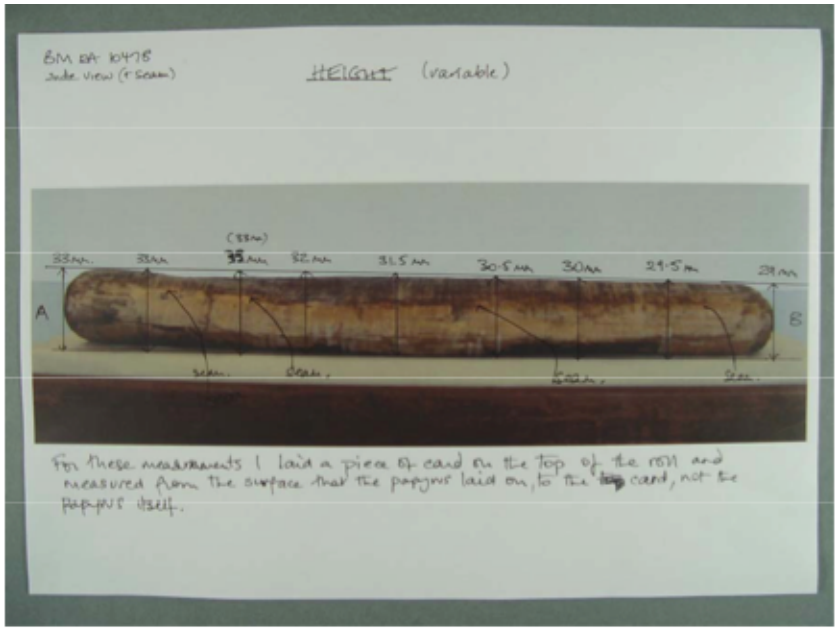


Photo D: Egyptian papyrus roll (British Museum)

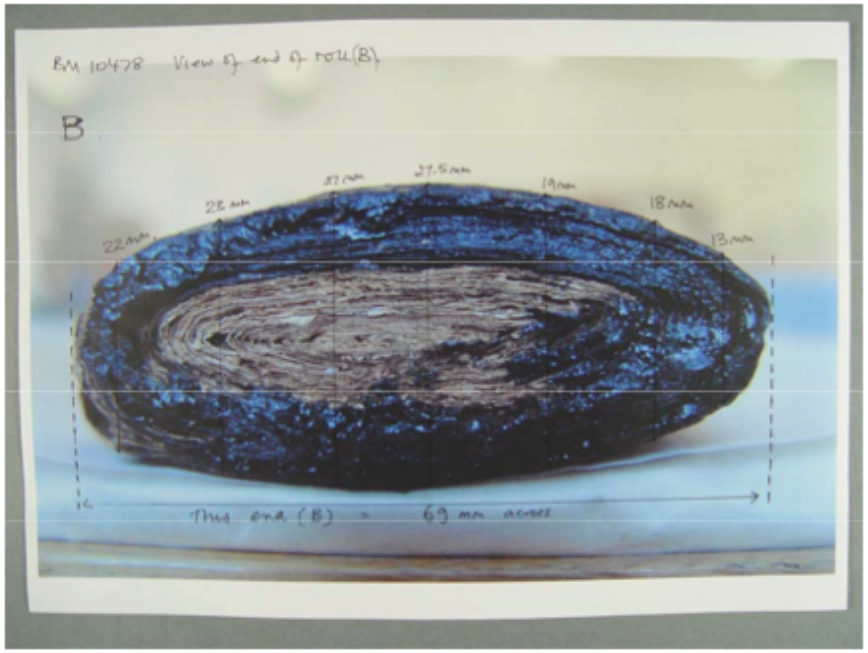
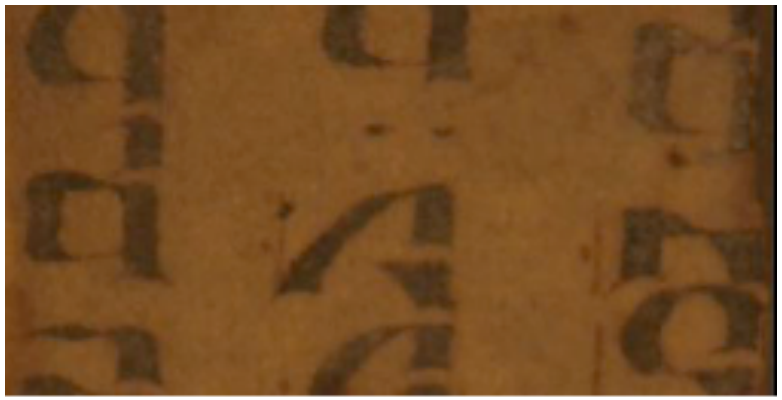
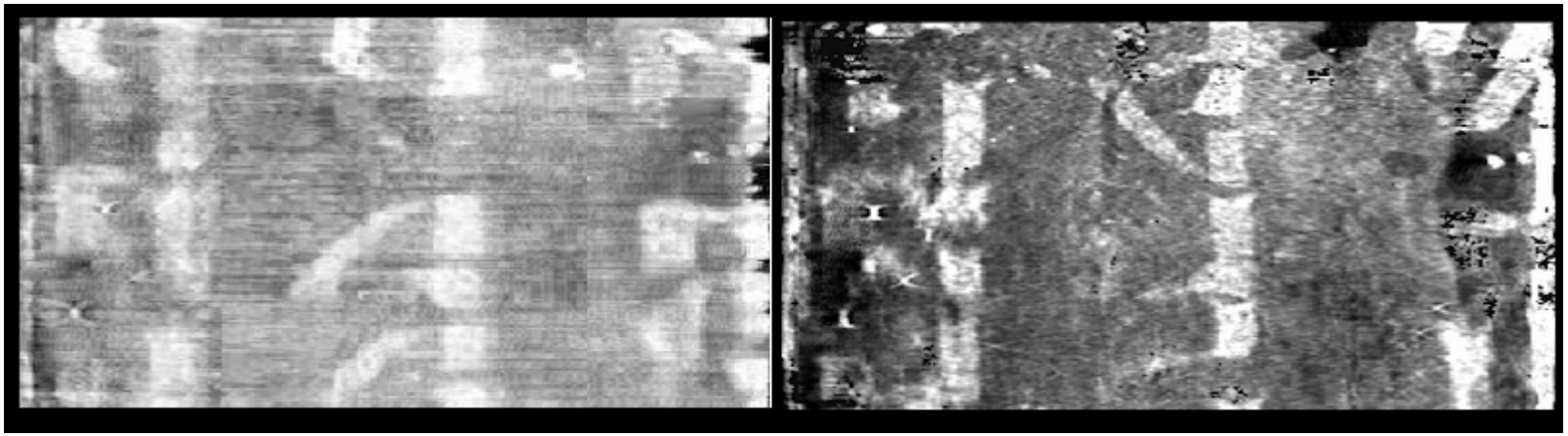


Photo E: Egyptian roll (end view)

Recovering Hidden Text





Hebrew Manuscript of Ecclesiastes

Script on the Back Side of the Page

Photo A: A conventional photograph of the fragment (Ecclesiastes) from the University of Michigan's collection

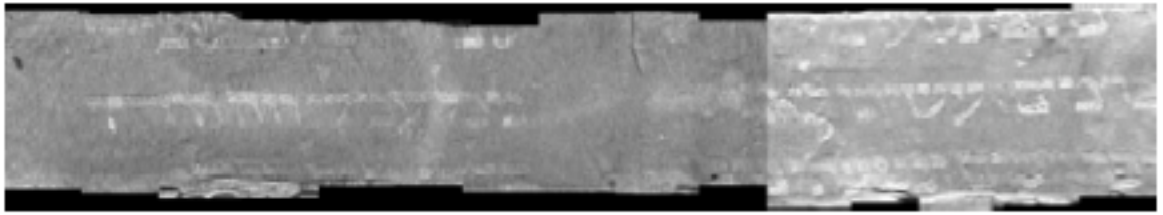


Photo B: The visible layer recovered from the noninvasive volumetric scan

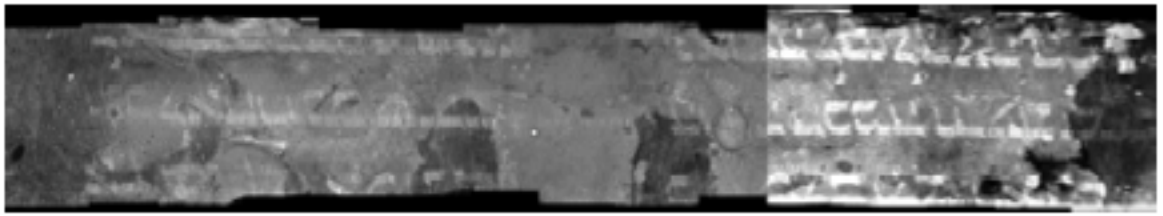
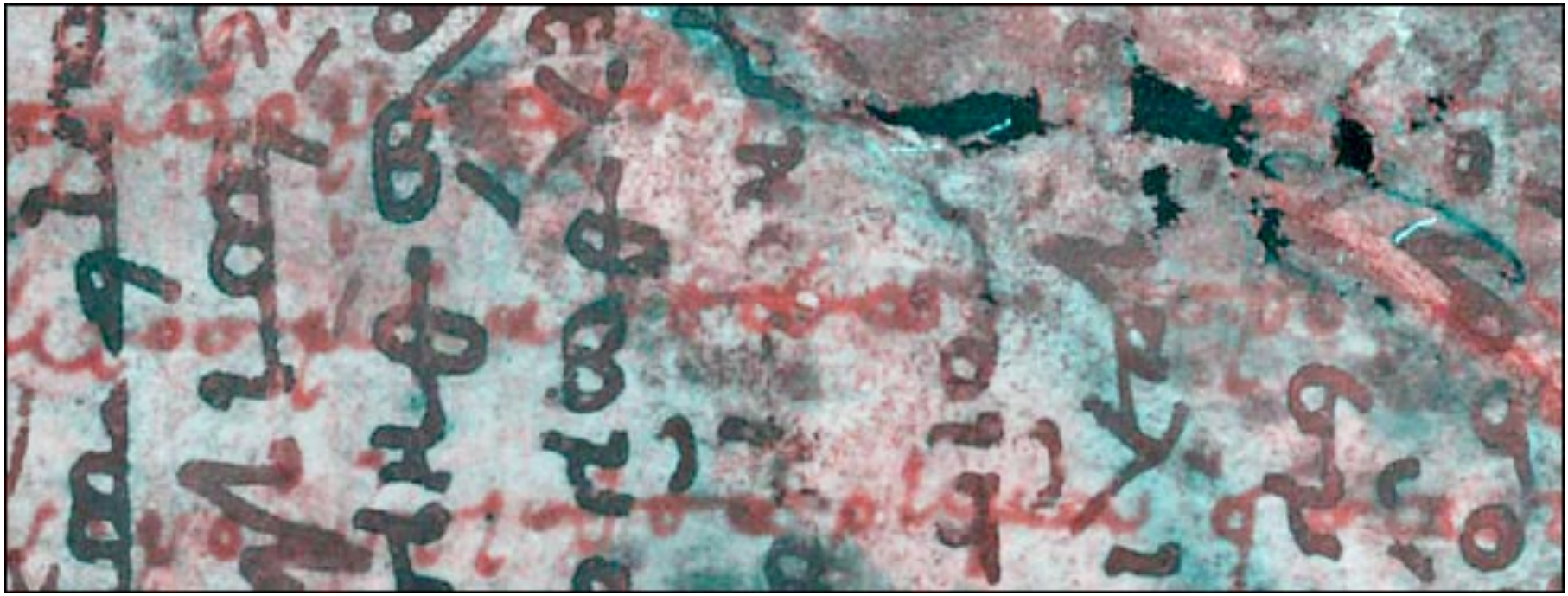


Photo C: The backside (hidden side) of the top layer. The text is distinct and from a different part of the book of Ecclesiastes.

Recovering Overwritten Text
Imaging the Lost Writings of Archimedes



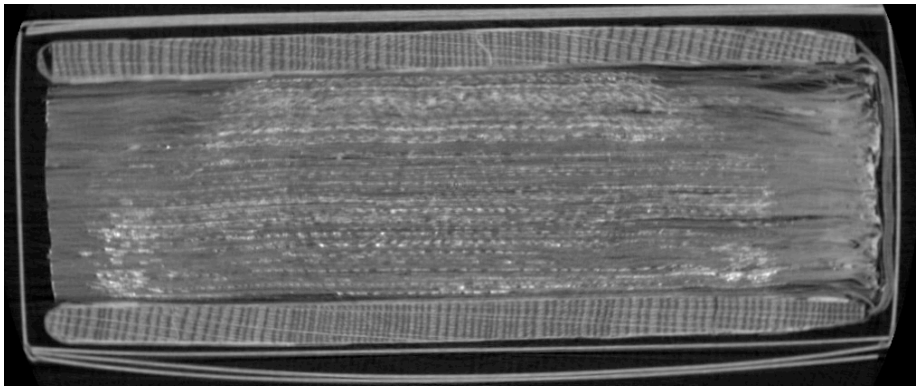
Keith T. Knox, Boeing

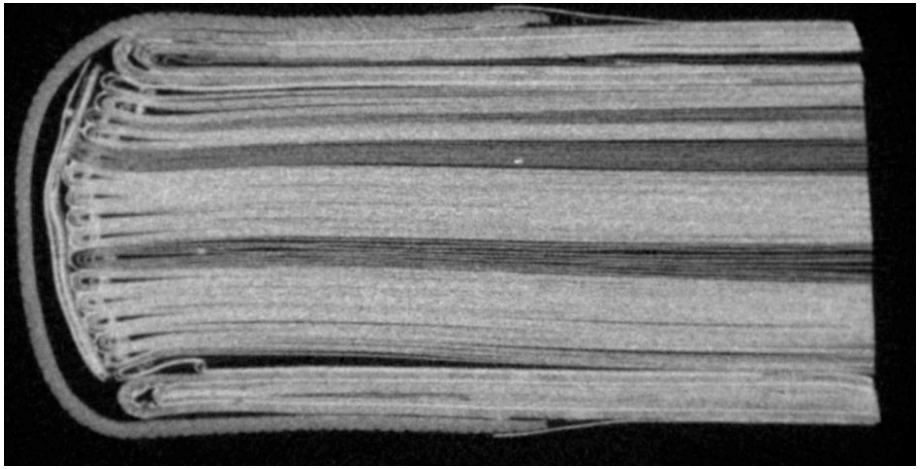
http://www.archimedespalimpsest.org/

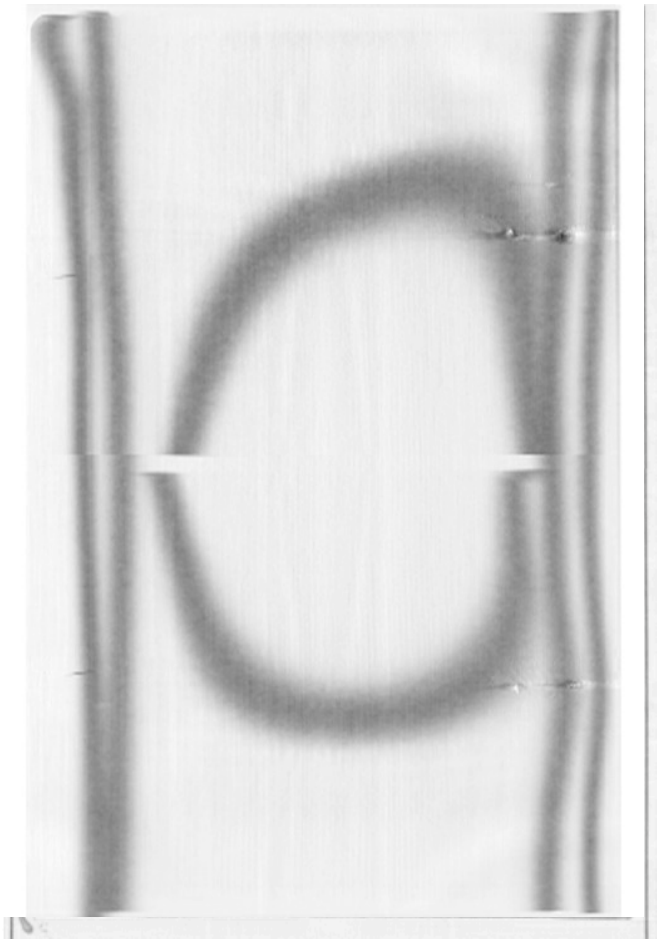
Page Imaging Without Opening the Book

High Resolution X-ray CT

16th-century Mexican Psalterium



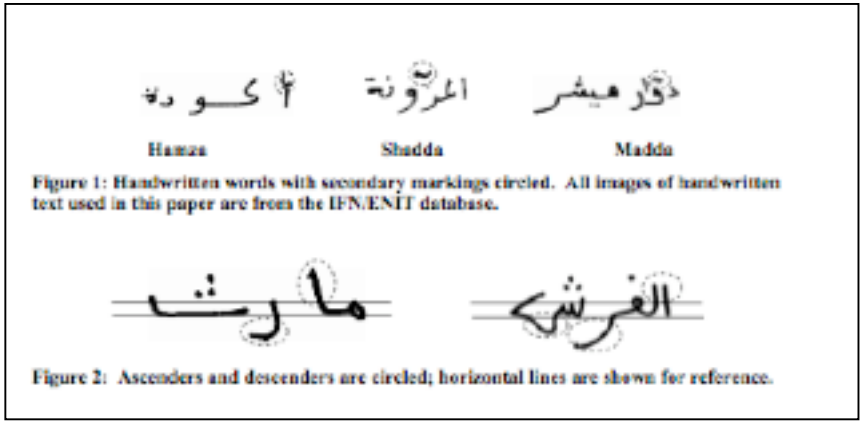


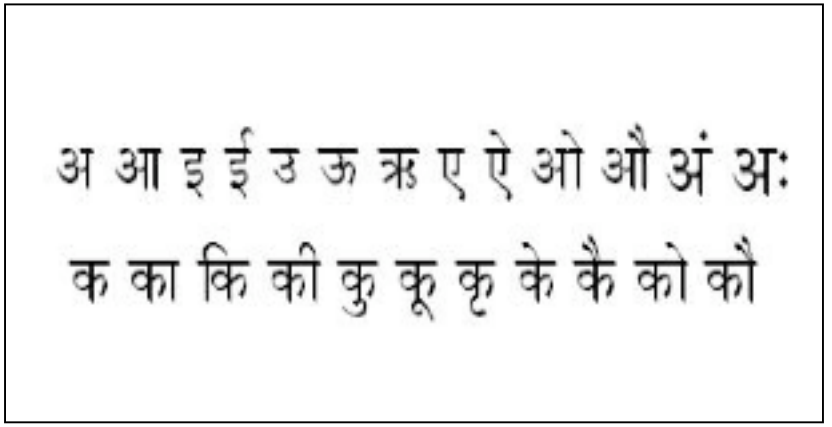


OCR of Scripts and Texts

The proposal to the Unicode Consortium recommends the addition of Vedic characters, such as pitch marks in red in this birch bark Vedic manuscript, to the Unicode standard.

OCR of non Roman Scripts



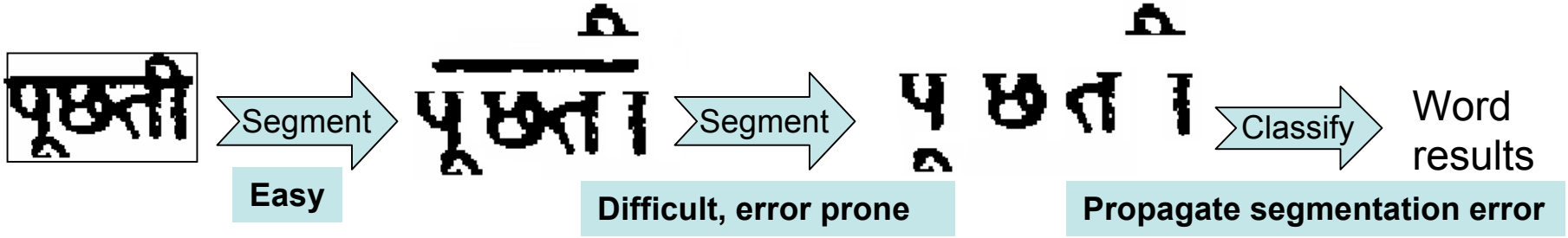


Arabic, Devanagari

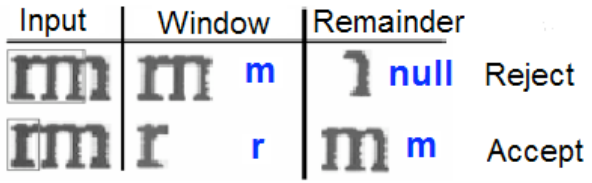
Center of Excellence in Document Analysis and Recognition (CEDAR)
SUNY, Buffalo

http://www.cedar.buffalo.edu/script/Dscript.html

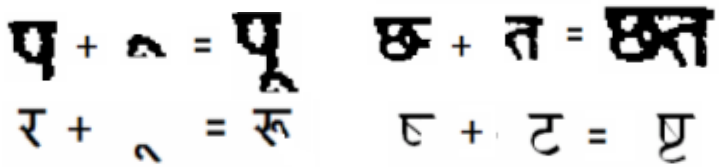
Background
(Prior techniques)



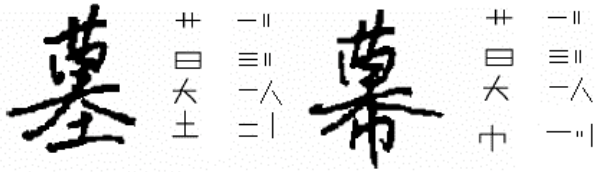
- Prior Devanagari OCR is segmentation-driven [Pal 97, Bansal 00]
- Techniques not robust for all character shapes



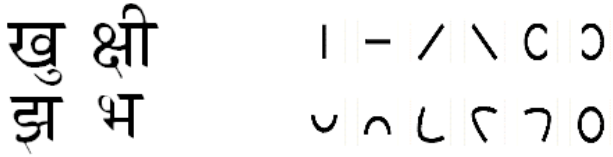
Linear window to segment Latin characters [Song et al 05]



Devanagari characters which cannot be segmented using linear window in single direction



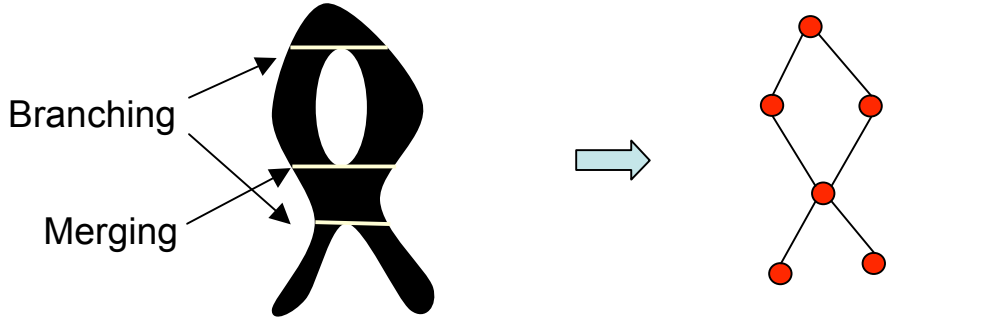
Chinese characters modeled as radicals and straight lines [Suen et al 03]

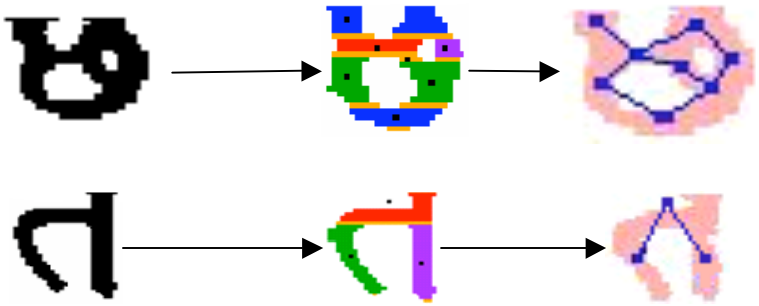


Most Devanagari primitives are non-linear

Proposed Methodology

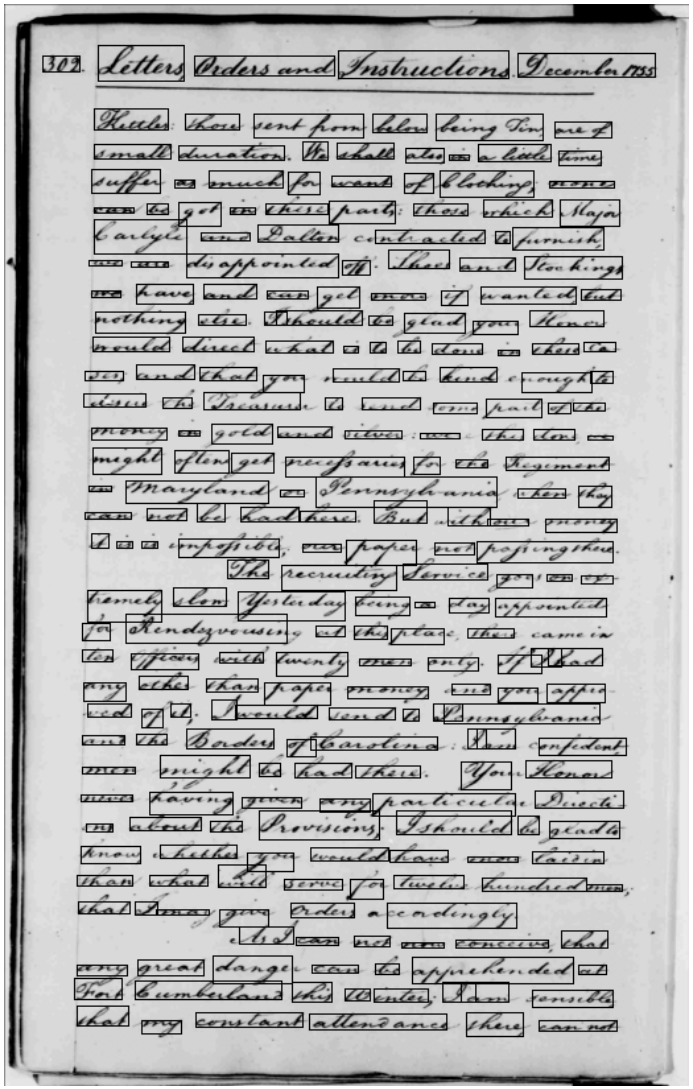
(Block Adjacency Graph – BAG for recognition)





- Natural breaks and joins of characters can be represented
- Connectivity and relative positions of nodes is embedded in the BAG
- Shape similarity is made more apparent
- Selecting sub graphs from the BAG can lead to character segmentation

Word Spotting: Indexing Handwritten Manuscripts



Example Manuscript and Segmentation

http://ciir.cs.umass.edu/research/wordspotting/

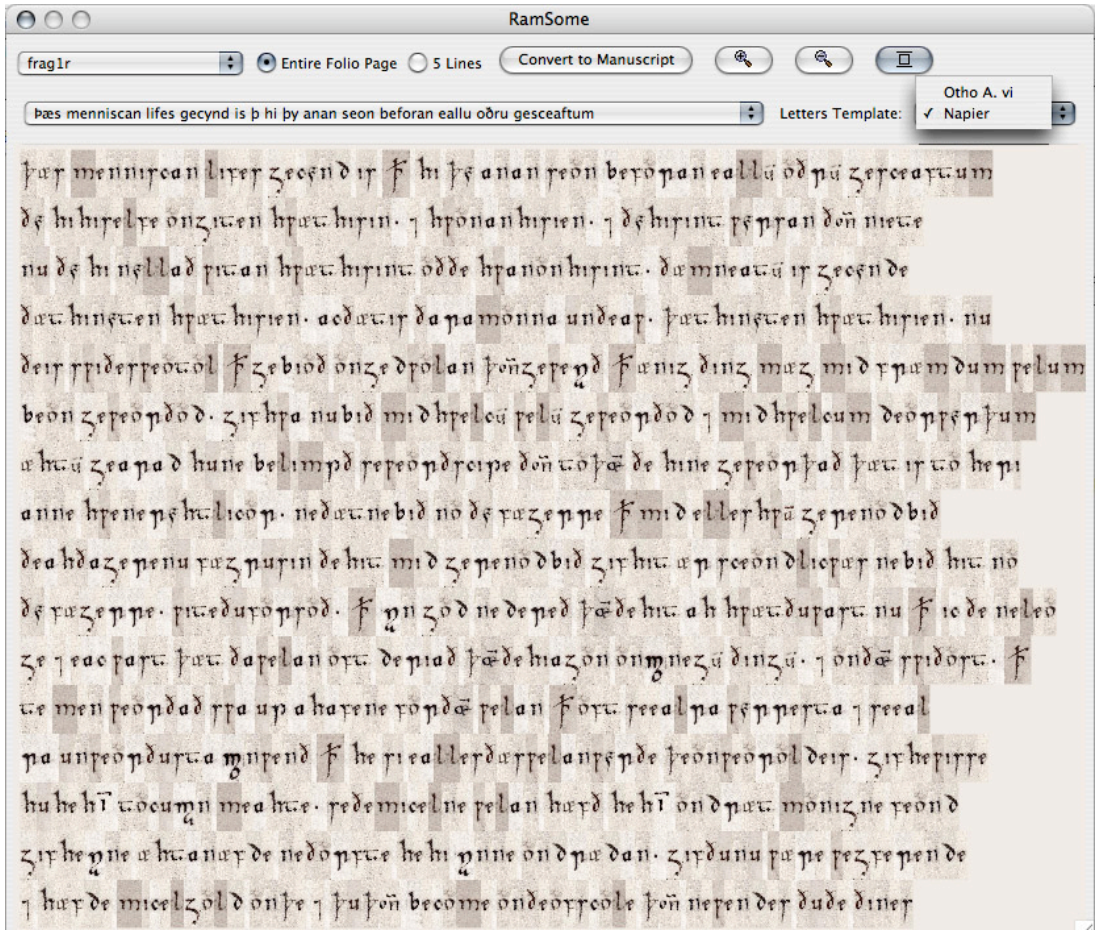
Edition Production Technology (EPT)

Analytical tools created as part of a comprehensive "edition production technology" (EPT) for image-based electronic editions can help editors reconstruct folios from lost or damaged manuscripts. A case in point is the Napier fragment of the Alfredian Boethius, the bottom portion of a MS leaf found and lost by A. S. Napier in 1886. Assembling and displaying Napier's detailed descriptions, digital tools can not only recreate a plausible reconstruction of the lost leaf, but also throw legitimate doubt on its authenticity.

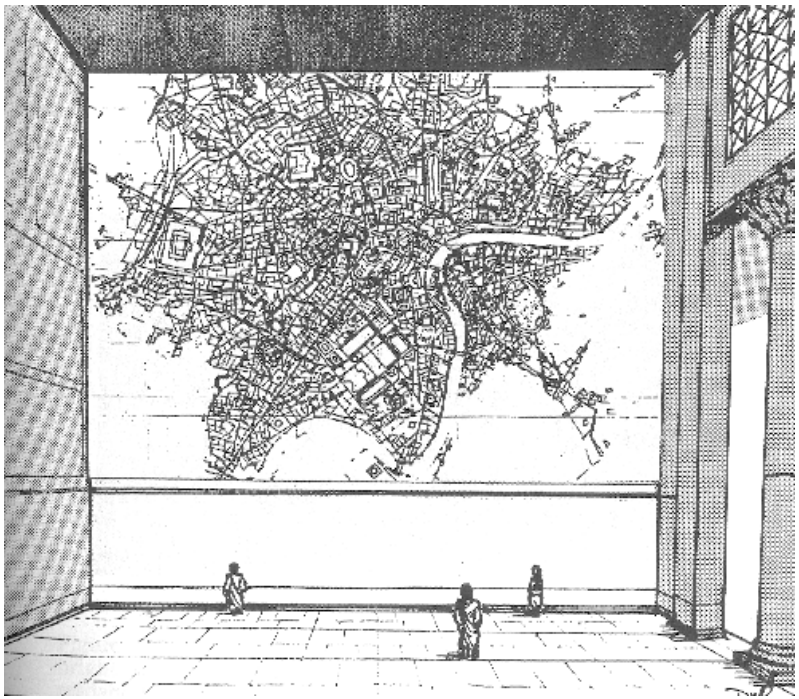
Figure 10: Detail from RamSome reconstruction of Napier fragment

Digital Medievalist 1.1 (Spring 2005). ISSN: 1715-0736.

Edition Production Technology
-individual letter recognition-



Reconstruction of Large and Incomplete Artifacts: Forma Urbis Romae



David West Reynolds

Digitization Goals

Prof. Emilio Rodriguez-Almeida

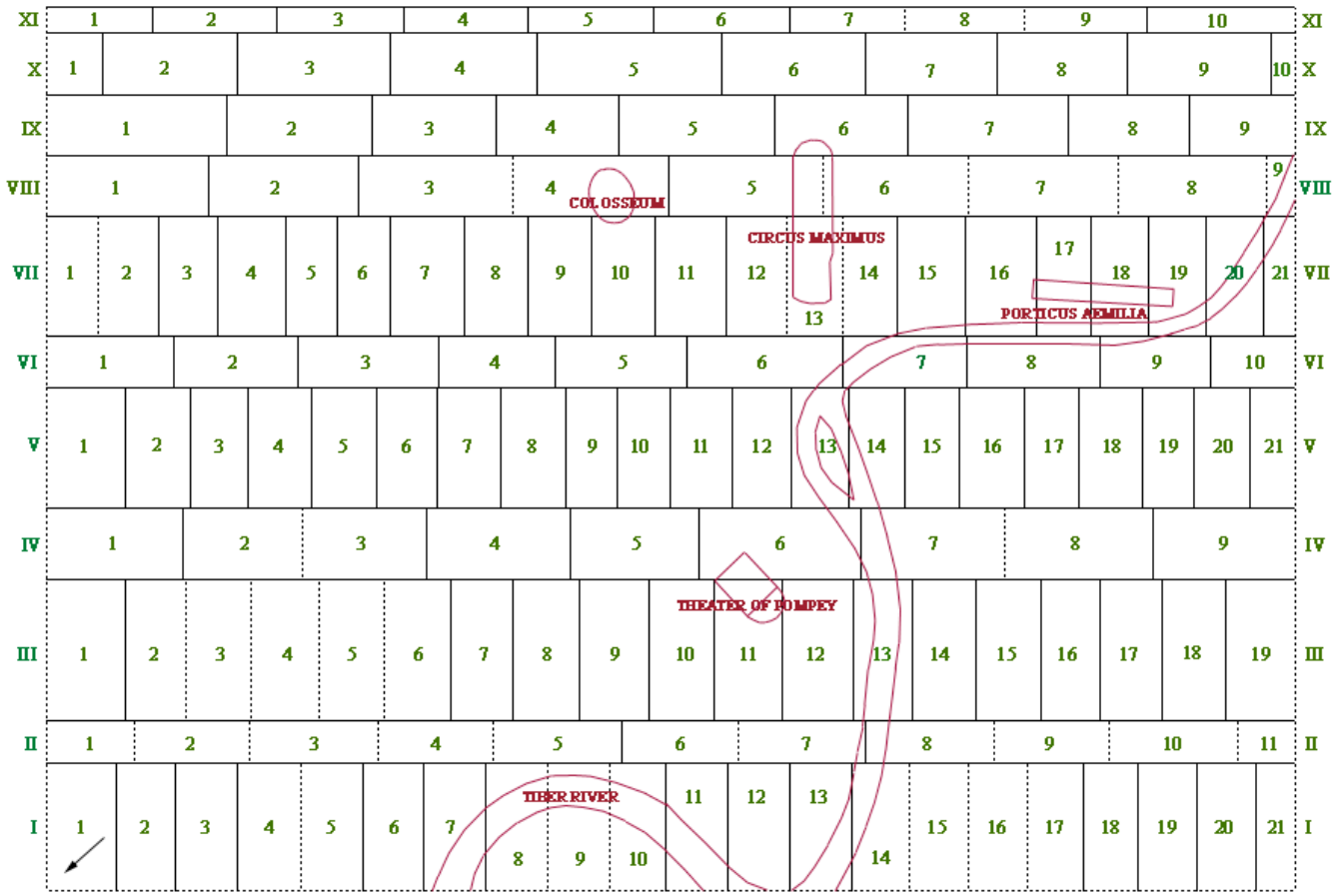
- Complete 3D archive of the extant fragments (1,100+ fragments)
- Adequate spatial sampling resolution for 3D matching and analysis (0.25 mm)
- High resolution (100 dpi) color photographs of top/bottom surfaces

http://graphics.stanford.edu/projects/forma-urbis/

Stanford∧Digital∧Forma∧Urbis∧Romae∧Project

DIGITAL MAP OF THE SLABS OF THE FORMA URBIS ROMAE

Click on a slab to view fragments associated with it.



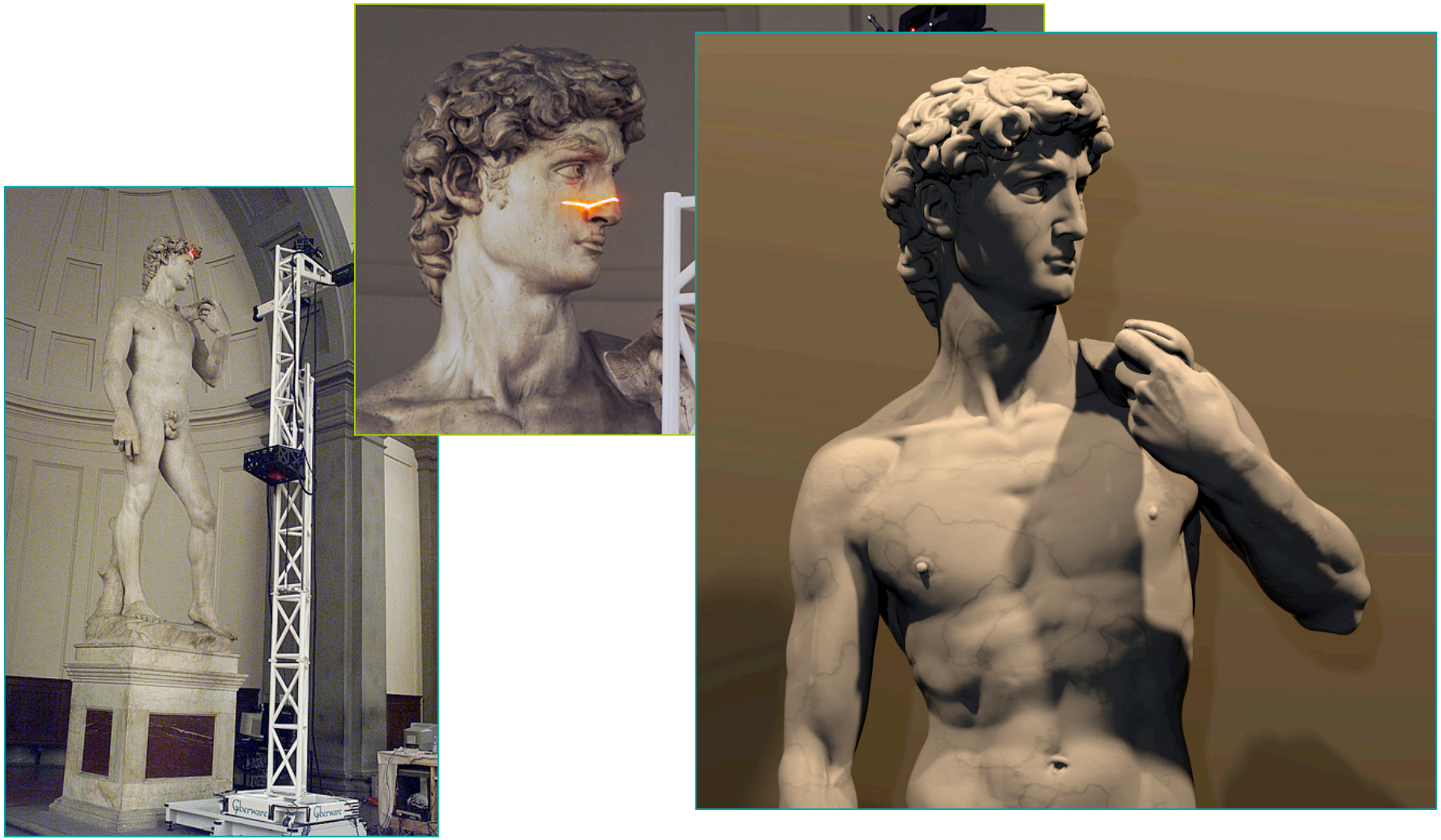
http://formaurbis.stanford.edu/docs/FURslabmap.html

FRAGMENTS ON SLAB VIII-4

FRAGMENTS ON SLAB VIII-4

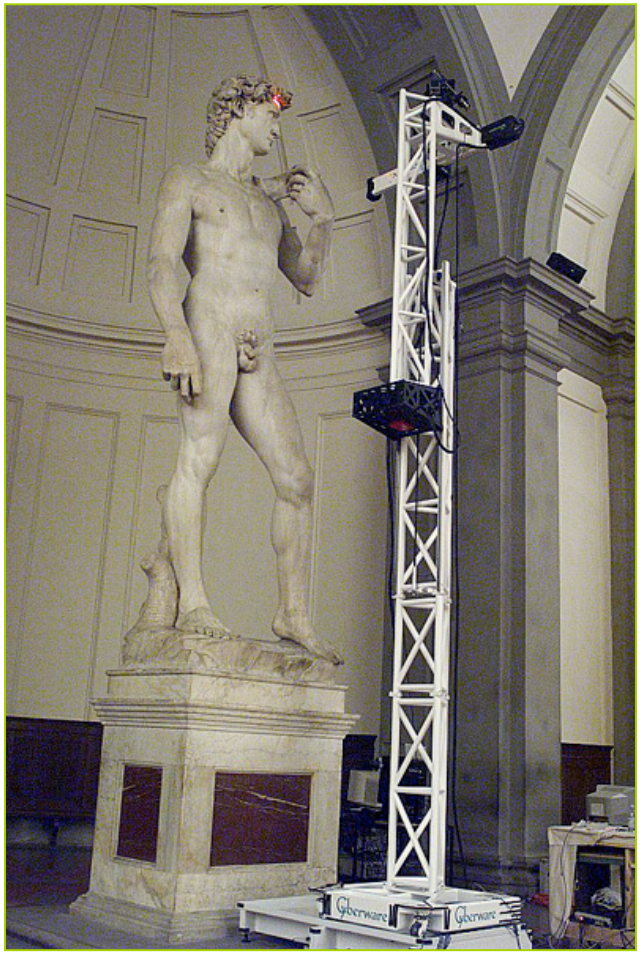
Stanford #	AG1980 #	PM1960 #	Slab	Back Surface	Identification
13ac	13a,c	13 a c	VIII-4	rough	Flavian Amphitheater, Colosseum (*Amphitheatrum*)
13b	13b	13 b	VIII-4	rough	Flavian Amphitheater, Colosseum (*Amphitheatrum*)
13de	13d-e	13 d e	VIII-4	rough	Flavian Amphitheater, Colosseum (*Amphitheatrum*)
13f	13f	13 f	VIII-4	rough	Flavian Amphitheater, Colosseum (*Amphitheatrum*)
13g	13g	13 g	VIII-4	rough	Flavian Amphitheater, Colosseum (*Amphitheatrum*)
13hi	13h-i	13 h i	VIII-4	rough	Flavian Amphitheater, Colosseum (*Amphitheatrum*)
13l	13l	13 l	VIII-4	rough	Flavian Amphitheater, Colosseum (*Amphitheatrum*)
13m	13m	13 m	VIII-4	rough	Flavian Amphitheater, Colosseum (*Amphitheatrum*)
13n	13n	13 n	VIII-4	rough	Flavian Amphitheater, Colosseum (*Amphitheatrum*)
13o	13o	13 o	VIII-4	rough	Flavian Amphitheater, Colosseum (*Amphitheatrum*)
13p	13p	142	VIII-4	rough	The Dacian(?) Training School for Gladiators (*ludus Dacicus*)
13q	13q	136	VIII-4	rough	Baths of Trajan (*thermae Traiani*)
13r	13r	135	VIII-4	rough	Baths of Trajan (*thermae Traiani*)
13s	13s	109	VIII-4	rough	Baths of Trajan (*thermae Traiani*)
110	110	110	VIII-4	rough	Baths of Titus (*thermae Titi*)?
113	113	113	VIII-4	rough	Rooms between the Baths of Trajan (*thermae Traiani*) and the Baths of Titus (*thermae Titi*)

Digital Michelangelo Project

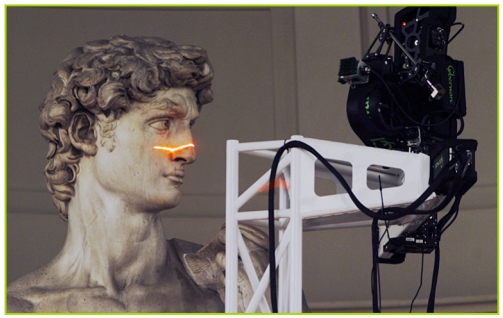


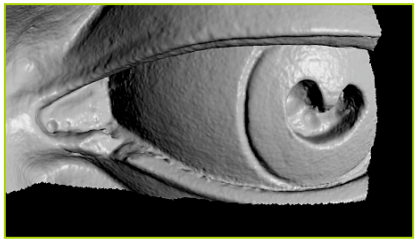
Large Data Objects
Michelangelo's David

Laser scanning, digitization and computational rendering of Michelangelo's David

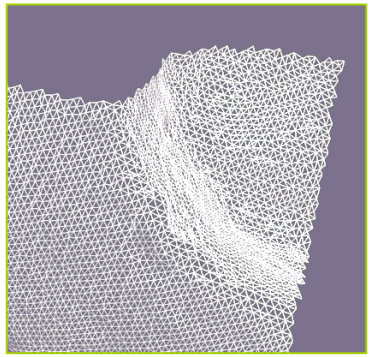


By Professor Marc Levoy of Stanford University





David's left eye



The scanning gantry is 7.5 meters high. 480 individually aimed scans were performed. The resulting 3D raw data set exceeds 250GB. The digital model, accurate to .25mm, contains 2 billion polygons and 7 thousand color images.

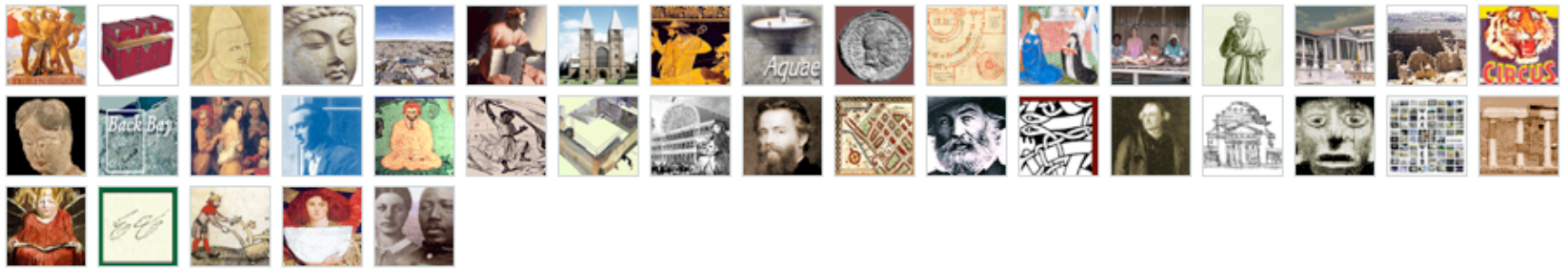
http://graphics.stanford.edu/projects/mich/

Institute for Advanced Technology in the Humanities
University of Virginia

IATH

The Institute for Advanced Technology in the Humanities

Current Research Projects



The Walt Whitman Archive

Folsom, Ed, The Carver Professor of English at The University of Iowa
Price, Kenneth M., The Hillegass Professor of American literature at the University of Nebraska, Lincoln
Networked Associate Fellows 1997

http://iath.virginia.edu/

Rome Reborn 1.0

Institute for Advanced Technology in the Humanities at the University of Virginia

1 2 3 4 5

History. From 1997 to 2007 the UCLA Cultural Virtual Reality Laboratory (CVRLab), the UCLA Experiential Technology Center (ETC), the Reverse Engineering (INDACO) Lab at the Politecnico di Milano, and the Institute for Advanced Technology in the Humanities (IATH) of the University of Virginia collaborated on a project to create a digital model of ancient Rome as it appeared in late antiquity. The notional date of the model is June 21, 320 A.D.

The name. The project is called "Rome Reborn" in homage to the founding text of the field of Roman topography, the De Roma instaurata of Flavio Biondo (1444-46).

Stanford Forma Urbis

St. Gall Plan

A Correct 10 Fragment Reconstruction

The reconstruction approach is to add one fragment at a time and is search-intensive because sherds are chipped, some are missing, and data is noisy.

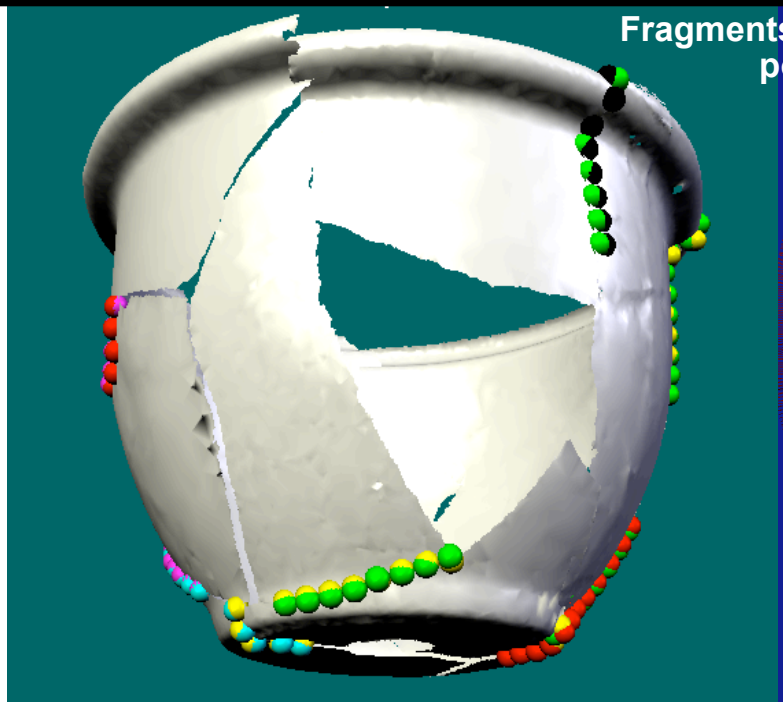
Fragments are matched and aligned using a probabilistic Maximum Likelihood Estimation measure

$$\max_{all\ geometric\ parameters} \mathbf{P}\{all\ sherd\ data \mid geometric\ parameters\}$$

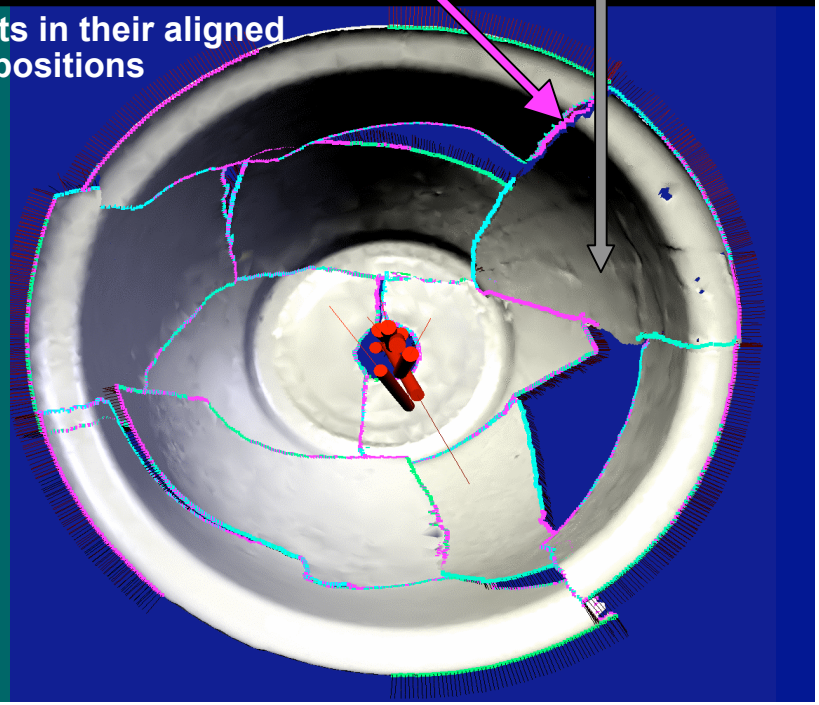
$$= \max_{\alpha, \beta, \mathbf{T}_1, \mathbf{T}_2, \ldots \mathbf{T}_\mathbf{I}} \prod_{i=1}^{\mathbf{I}} P(break\ curve\ data_i | \mathbf{T}_i, \beta)\, P(surface\ data_i | \mathbf{T}_i, \alpha)$$

α - Axis/profile-curve for the pot
β - Break curves for all fragments
T_i - Euclidean transformation for moving sherd *i* to correct position to reconstruct pot.

Global Pot Profile-Curve



Fragments in their aligned positions



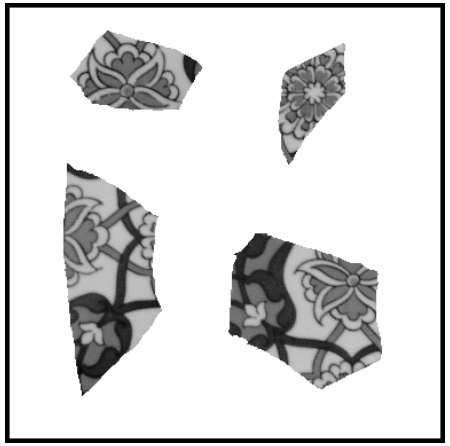
A Texture Based Matching Approach for Automated Assembly of Tiles and Patterned Pottery

Goal: To reconstruct unknown largely flat objects that have been broken into a large number of irregular fragments

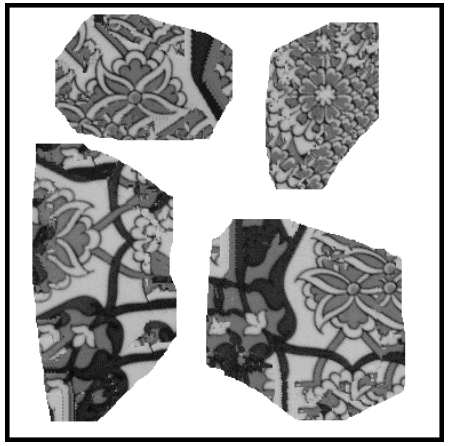
Solution: A new approach based on

1. the pixel color values in a band outside the border of the pieces

An example:



Original picture



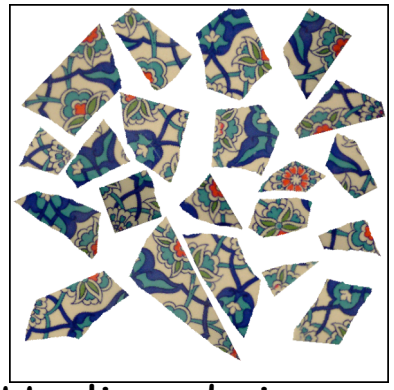
Enlarged images

Confidence images
(local affinity cost per pixel)

2. Define an affinity cost of corresponding pieces that utilizes all kinds of image information, such as continuity of edges, textural patterns, and color similarities
3. Treat the assembly problem as the minimiation of this affinity cost.

Fragment alignment is based on minimizing an affinity cost:

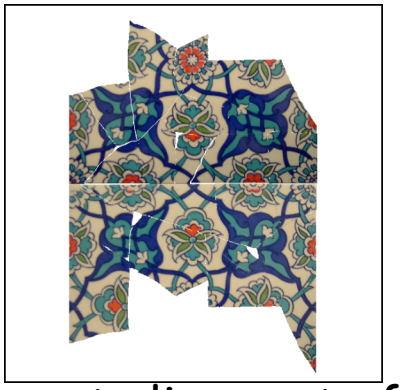
$$\frac{\sum_{i}^{n_p-1}\sum_{j}^{n_p}\left[\sum_{k}^{n_j} w_{f^k} D^k(f_i^k, f_j^k) - 1\right] G_i^0 G_j^0}{\sum_{i}^{n_p} G_i^0} + \sum_{i}^{n_p-1}\sum_{j}^{n_p} w_o L(G_i^o) L(G_j^o), \quad L(x) = \begin{cases} 1 & x = 1 \\ 0 & x \neq 1 \end{cases}$$



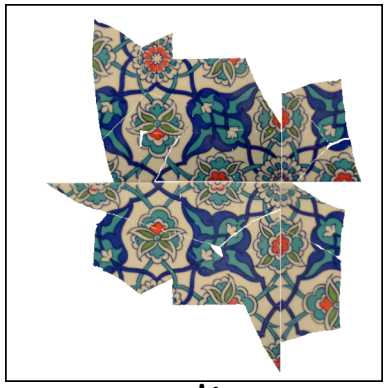
Unaligned pieces

F_{cost}=0

F_{cost}=-17841



An incorrect alignment of pieces



An incorrect alignment of pieces

F_{cost}=-18577

F_{cost}=-20250

Correctly aligned pieces

Two incorrect alignments that superficially look correct and the least cost alignment which is also the correct alignment

The general problem is that pieces from individual tiles are incomplete and pieces from more than one tile are present.

The final automatic alignment of pieces from two different ceramic tiles.

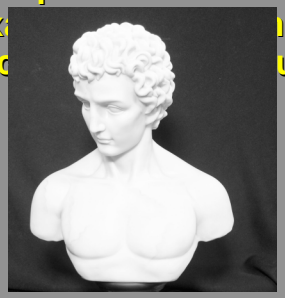
Shape from Silhouettes

Goal: Accurate reconstruction of freeform 3D surfaces or eroded surfaces from a sequence of images taken by a moving handheld camera or video camera.

Problem: Archaeological site objects, both small fragments and large objects, often have spatially slowly varying intensities (i.e., no patterns) in their images. This makes accurate 3D reconstruction of surfaces by standard stereo reconstruction through matching of image intensities in two or more images impossible.

Solution: Do 3D surface reconstruction using silhouettes. Object silhouettes in images are relatively easy to estimate and are independent of object internal appearance. Downside is that most concavities cannot be reconstructed from silhouettes alone.

Contribution: A new conceptual approach to reconstruction based on the space of measured tangent planes; computation is fast and low computational cost; only 21 images are used in this ex[illegible] but then with inc[illegible]



Two images of the David bust

Top row shows *three* out of *twenty one silhouettes* used for reconstruction.

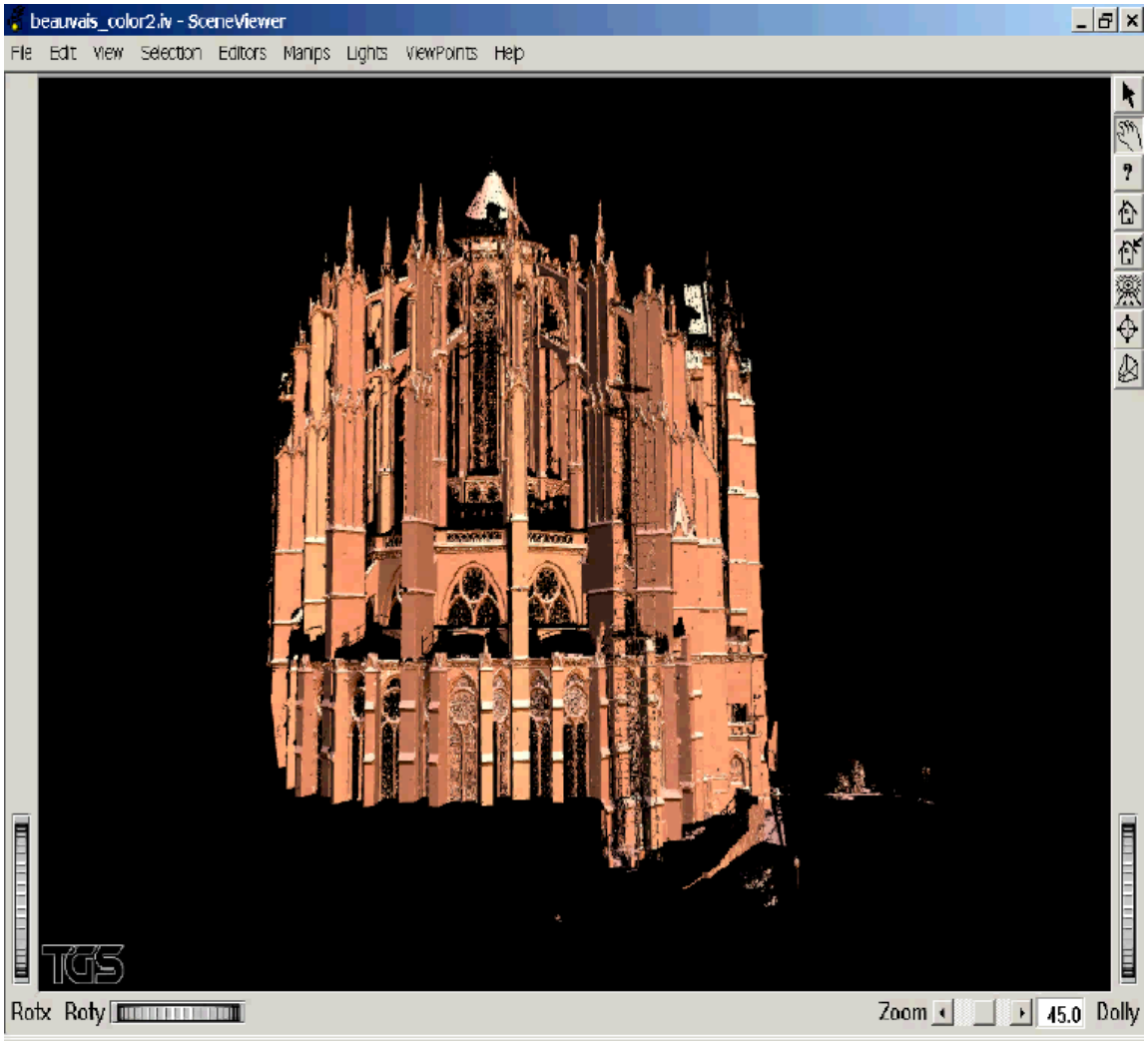
Bottom row shows two views of 3D point cloud estimated from those silhouettes

http://www.lems.brown.edu/vision/

Visualization for Archaeology

- Most analysis/interpretation happens off-site after digging when the real 3D environment is missing
- Current tools focus on 2D data and do not incorporate 3D information
- Collaboration is a must
- Goal: Develop collaborative, immersive visualization environment for analyzing 3D data off-site
  - Natural multimodal interaction: speech, gestures, touch
  - Combine 3D and 2D user interfaces
  - Real-time visualization

Modeling, Visualizing and Analyzing Historic and Archaeological Sites

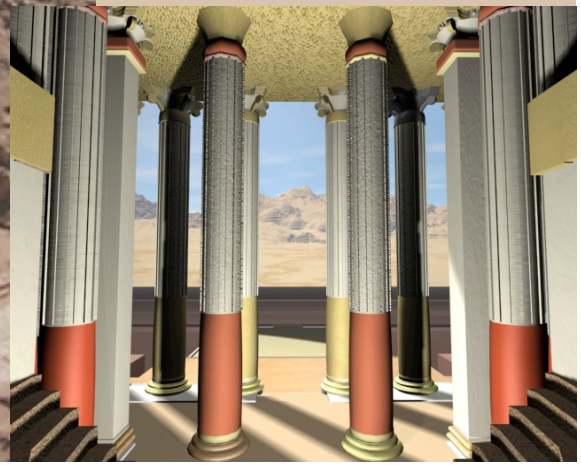


http://www1.cs.columbia.edu/~allen/ITR

Modeling Archaeological Sites

- Archaeology
  - Destructive process: models preserve state
  - Need for accurate and detailed documentation
  - Models allow off-site visualization, analysis
  - Can track changes over time
- Integrate a variety of data sources
  - Range data
  - Digital photographs
  - GIS data Archaeology

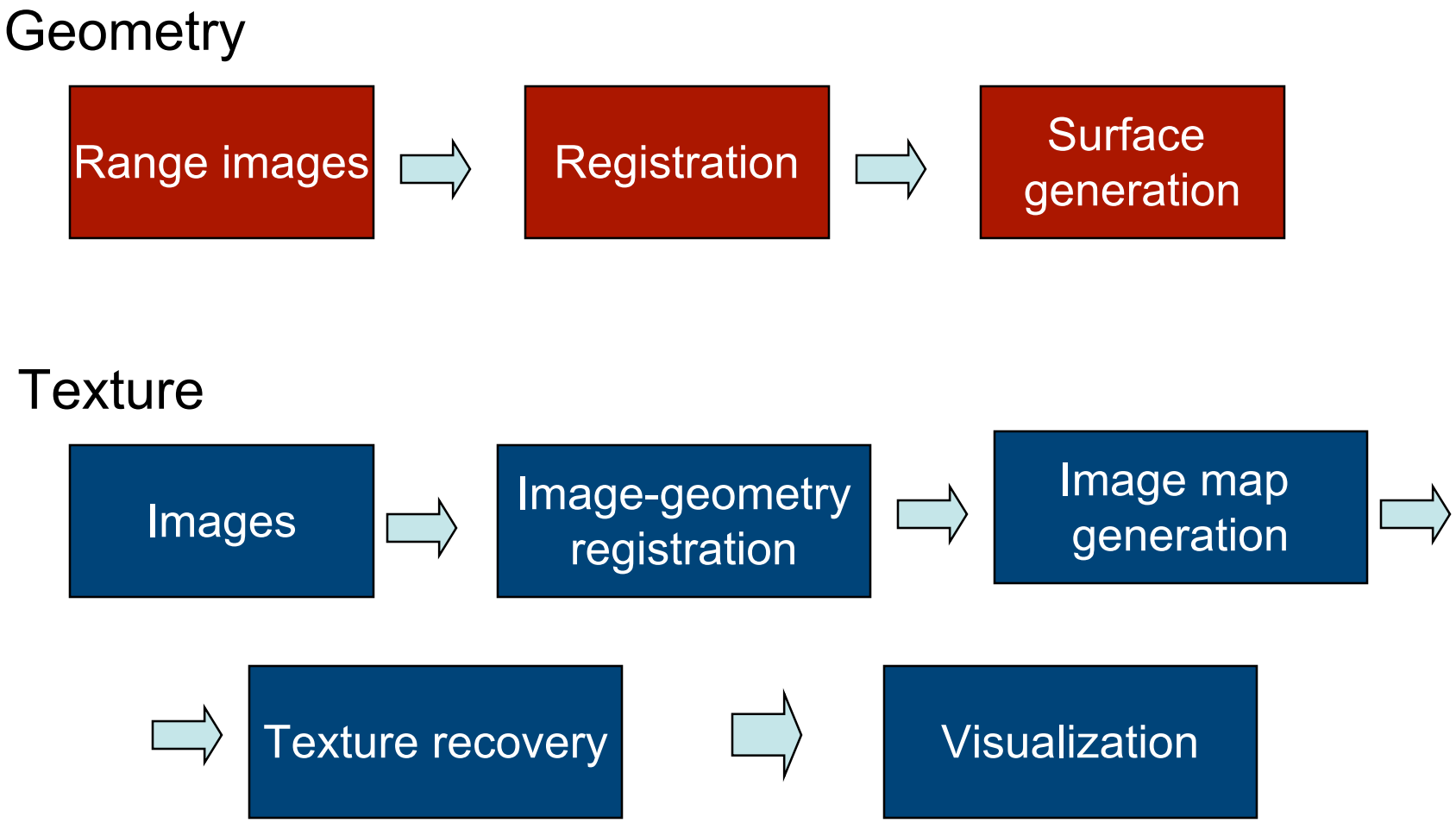
Virtual Reconstruction of Archaeological Sites
e.g., The Great Temple in Petra, Jordan



- Excavations 1993-present
- Over 70 Trenches Excavated
- Database with close to 500,000 Finds

Shown are reconstructed temple: outside views, and a 3D VR Brown-CAVE interior view.

The 3D modeling pipeline



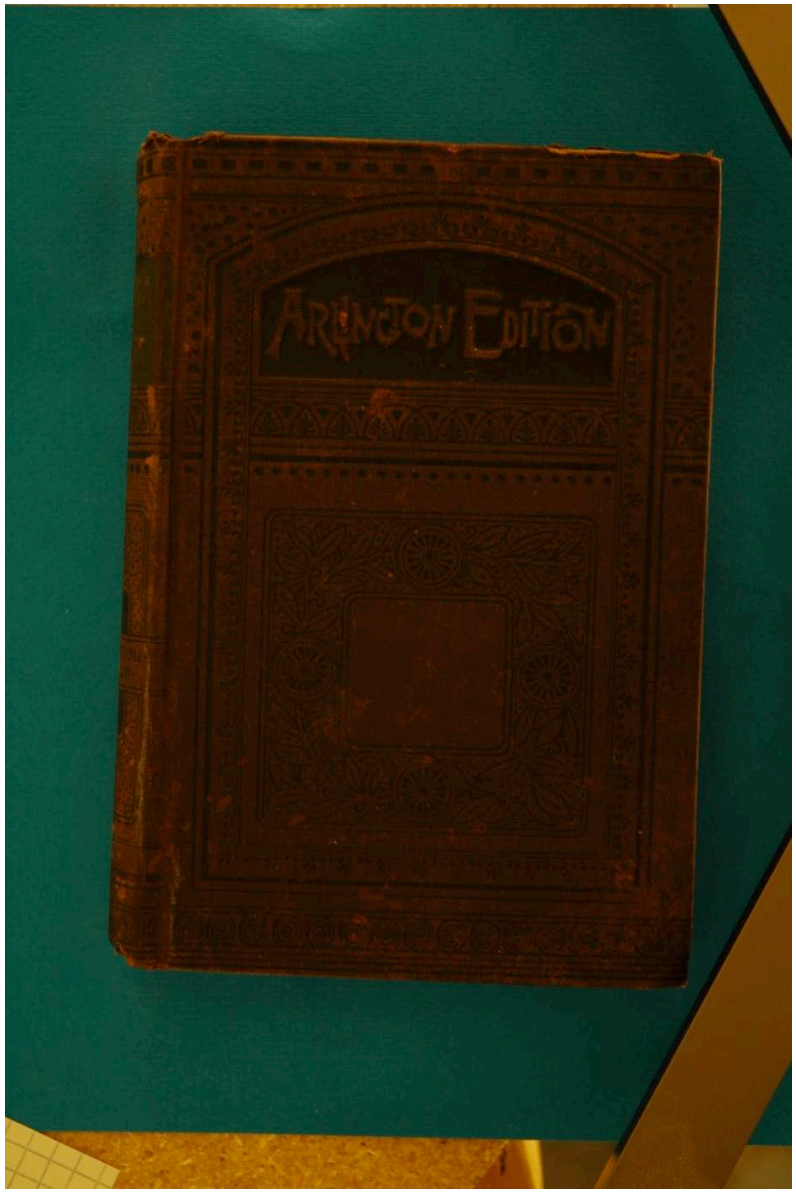
The Importance of Lighting

‘Good’ Uniform Lighting Hides Details

Fresnel Reflectance can reveal them,
but weak, mostly invisible
to overhead cameras
even with raking-angle illumination.

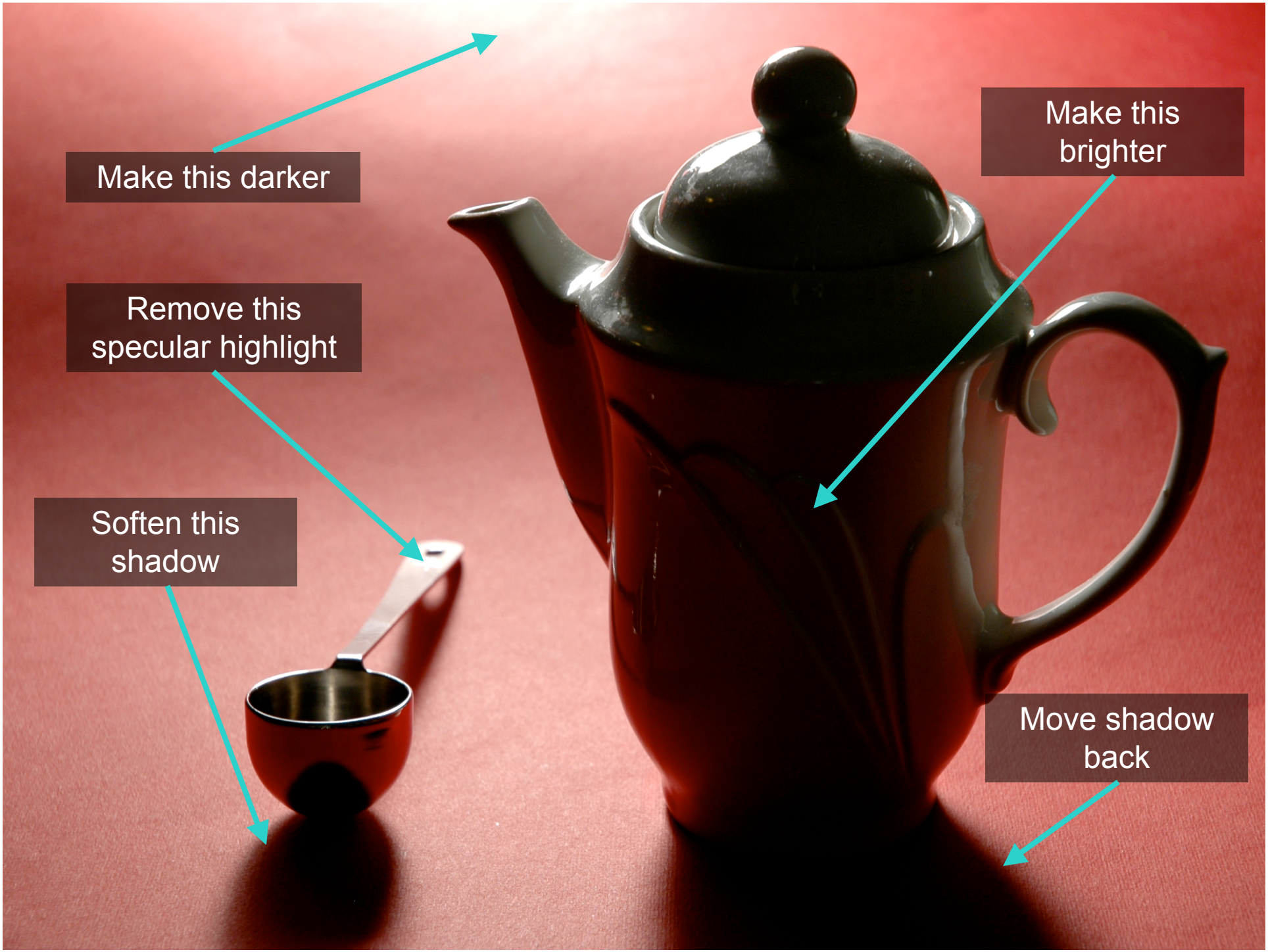
- Novel: Multiple Raking Cameras
plus continuously-variable
raking illuminant.

- very sensitive measure of surface texture
and angular perturbation (2X)



Jack Tumblin
Northwestern University

Fresnel Reflectance Reveals Them


Make this darker
Make this brighter
Remove this specular highlight
Soften this shadow
Move shadow back

Sketch Your Desires, Optimize

Result

cdli

CUNEIFORM DIGITAL LIBRARY INITIATIVE

A joint project of the
University of California at Los Angeles *and the*
Max Planck Institute for the History of Science

[Home] DIGITAL LIBRARY Education Publications

About CDLI
Project Associates and Staff
Related Projects
Communications

Methods & Conventions
Tools & Sundries

CDLI Catalogue Search

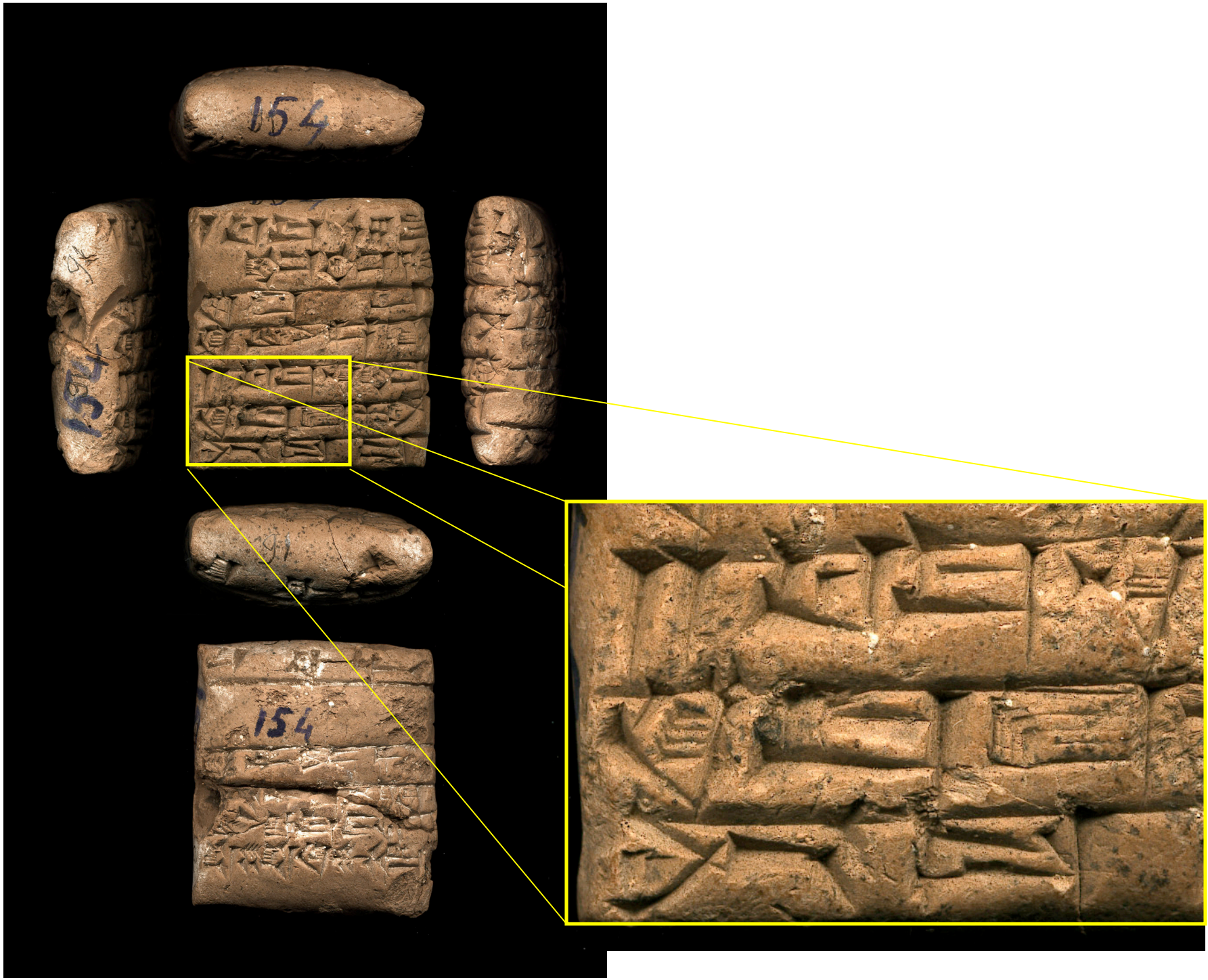
Yahoo! search go
Search the web Search this site

Last Updated: 20 February 2004

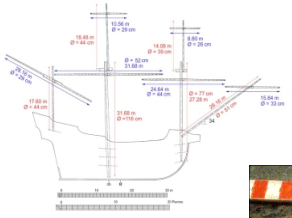
Funded by the Digital Libraries Initiative
(NSF/NEH) and the Max Planck Society

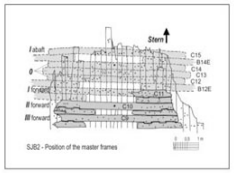
Berlin Mirror

Lead page of the CDLI


154
154

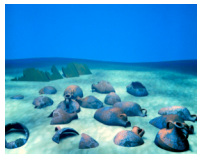
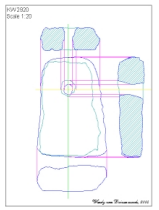
The Nautical Archaeology Digital Library







⚓ Ship reconstruction

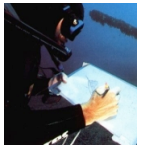



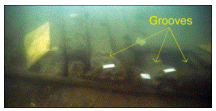

Main Goals

- Computational assistance for the identification and placement of ship fragments
- Algorithm and visualization-based mechanisms for ship reconstruction
- Heterogeneous data sources and media integration
- Mapping an underwater archaeological excavation
- Cross-linking of a dynamically-growing collection
- Digital library replication and synchronization
- Incorporating and managing uncertain data

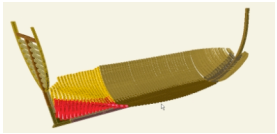


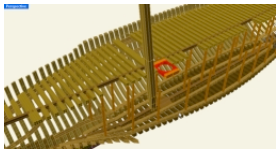
⚓ Artifact tracking/ analysis



⚓ Underwater Excavation






⚓ Ship modeling

http://nadl.tamu.ed/

Computational photography

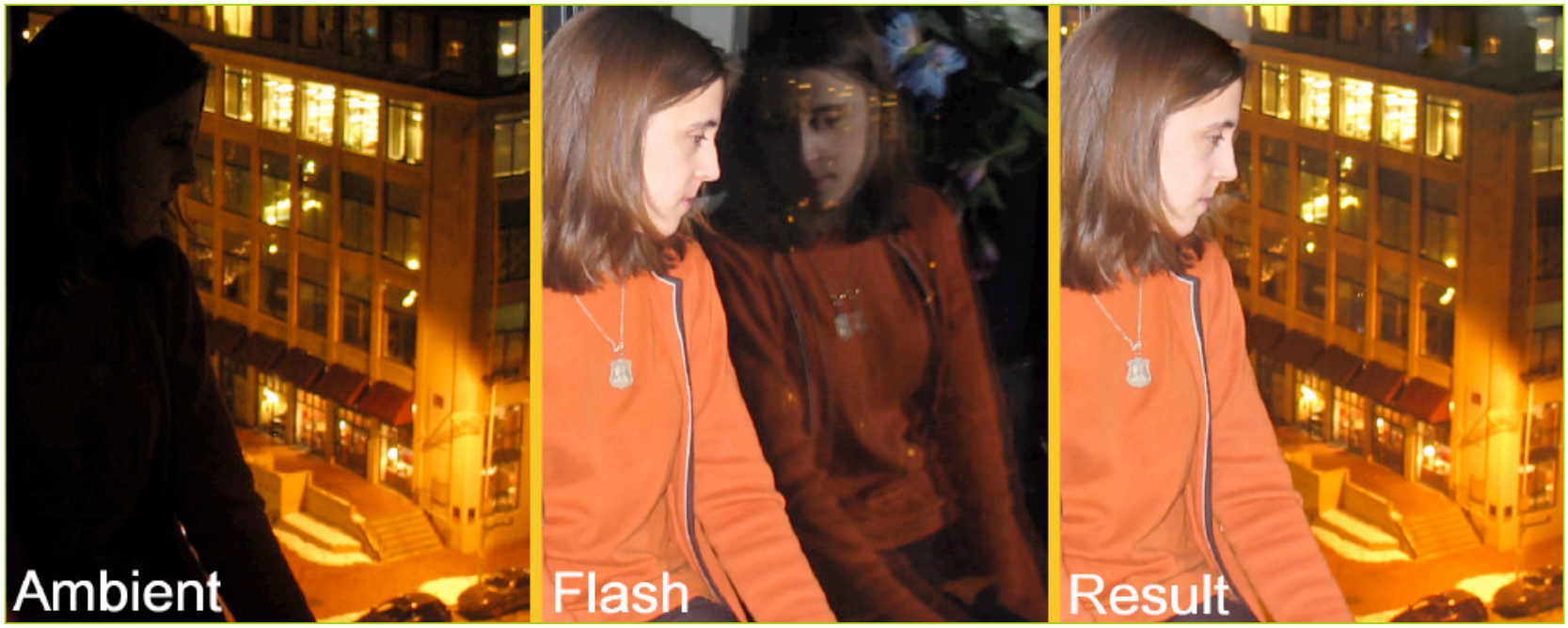
Deblurring with blurred/noisy image pairs
[Yuan SIGGRAPH 2007]

- deconvolve long-exposure (blurred) image, using short-exposure (noisy) image as prior

Example of digital refocusing

Marc Levoy, Stanford University

Flash-noflash photography
[Agrawal SIGGRAPH 2005]



- compute ambient + flash – features in sum that don't appear in ambient alone (as determined from image gradients) (except where ambient image is nearly black)

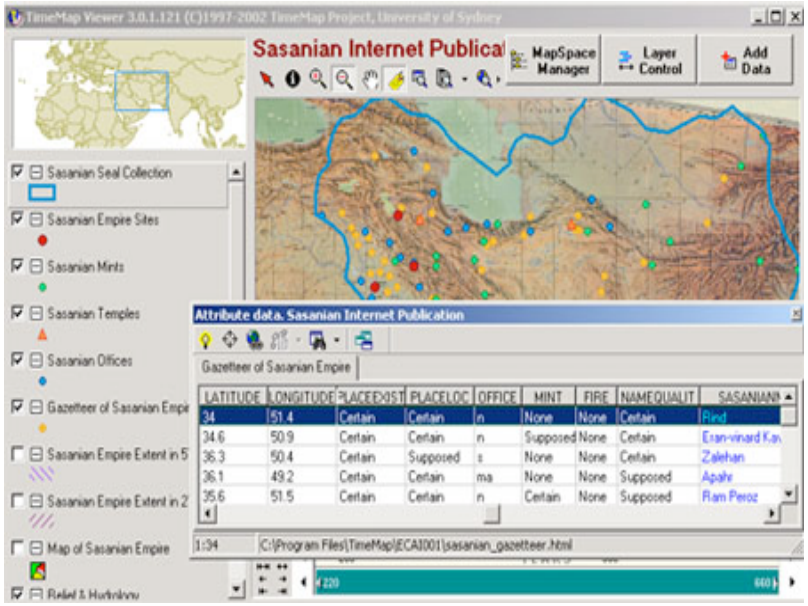
Scene completion using millions of photographs
[Hays & Efros SIGGRAPH 2007]

- search for matches from a large database
- Find least visible seams using graph-cut algorithm
- blend gradients & integrate to create image

Other applications of real-time image stabilization

remove foreground objects that don't appear in all shots
[Agarwala SIGGRAPH 2005]

Large, Complex Multinational Projects

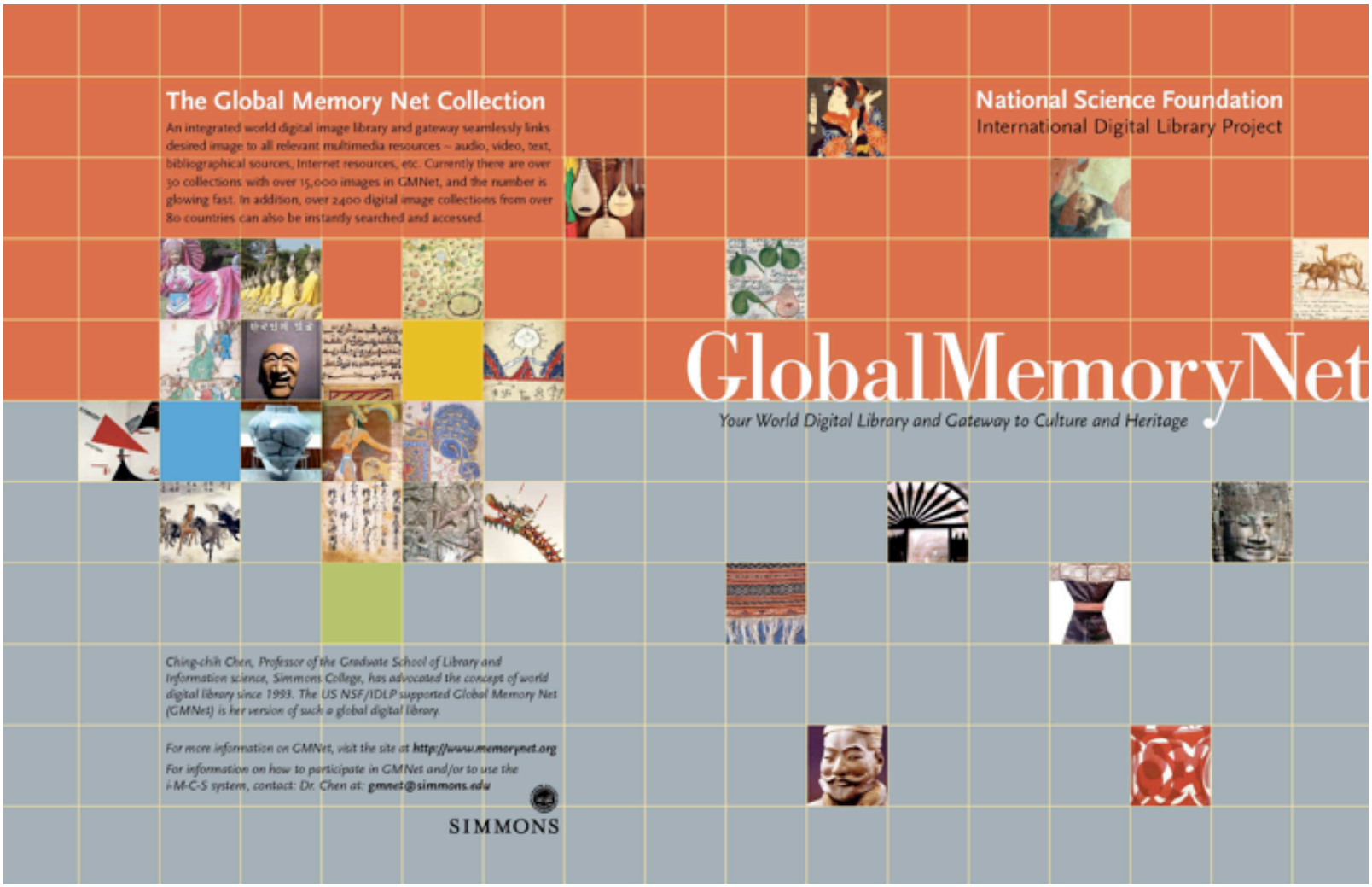


Time-enabled GIS viewing for historical and cultural gazetteer data.

www.ecai.org

Global Memory Net: A World Image Library and Gateway

Ching-chih Chen, Simmons College

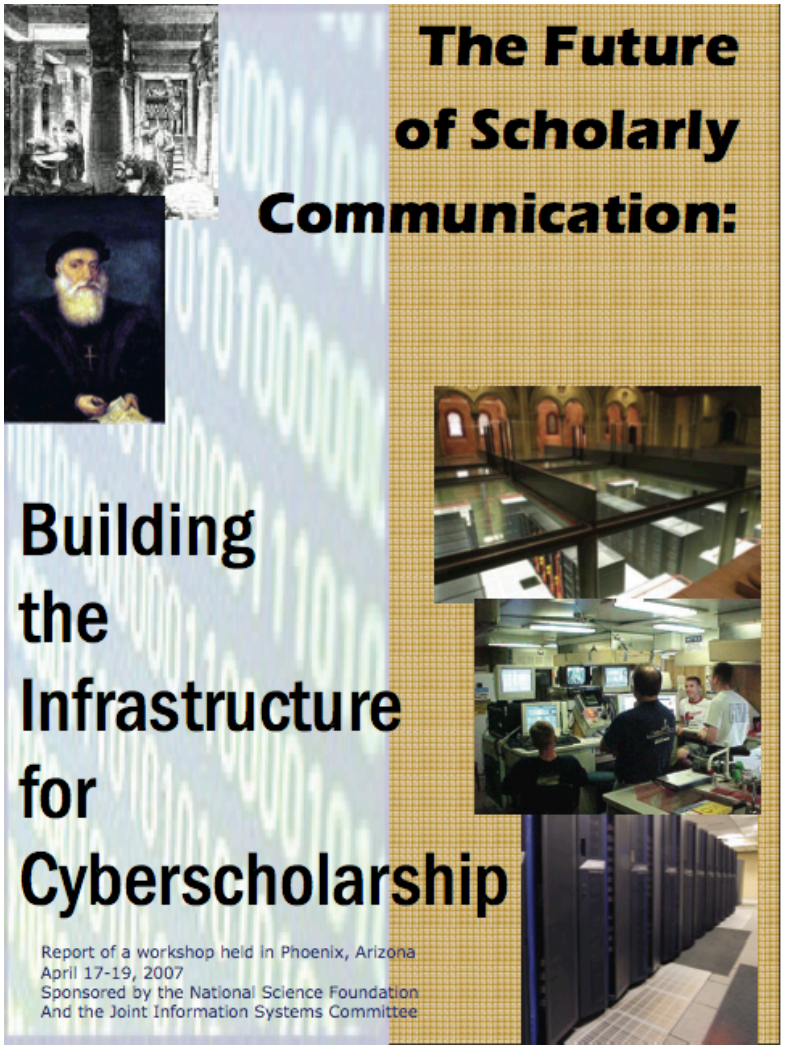


Internet Archive

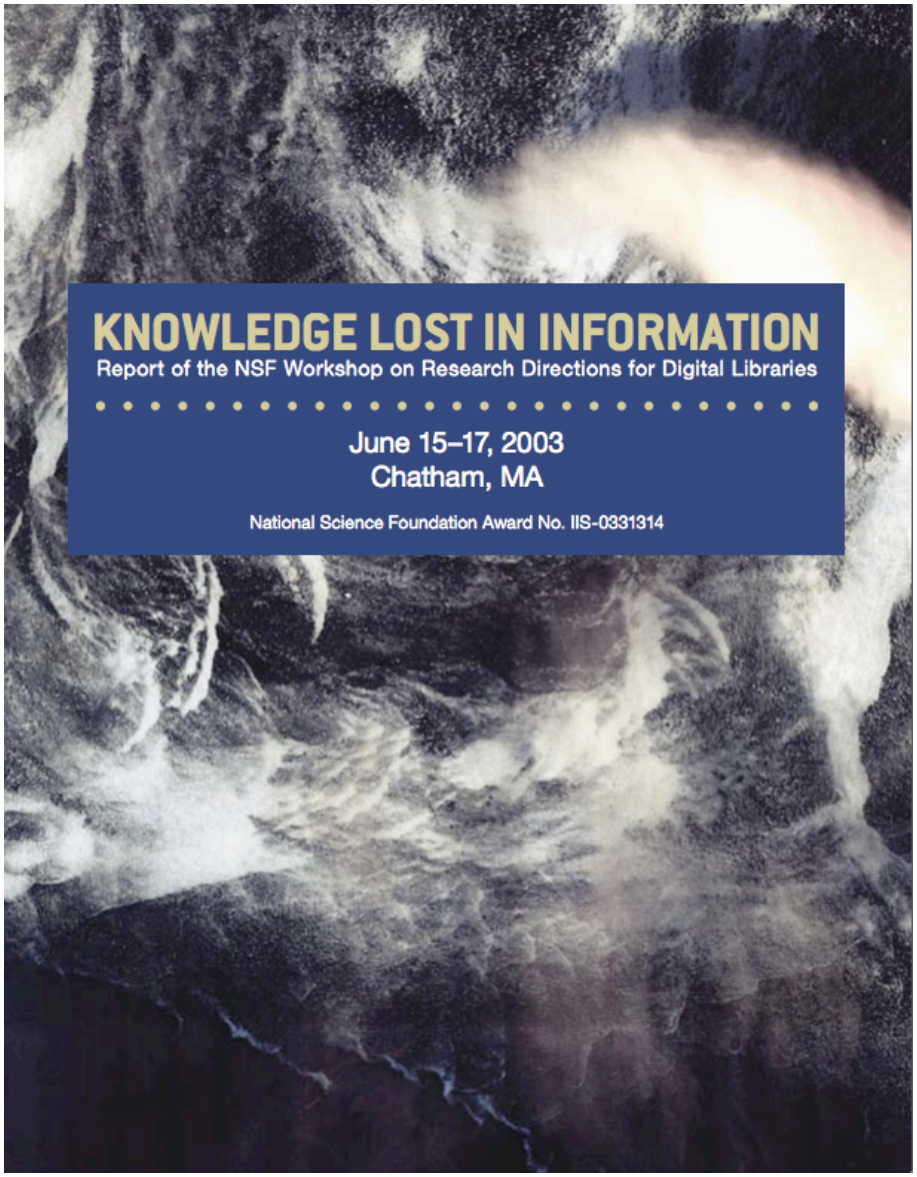
- Web: 10 billion pages from 1996-2001
- Television: 2000 hours of Egyptian and US TV
- Movies: 1000 archival films
- 100 Terabytes of data
- Storage on 200 computers

The second copy world wide, after the original copy in San Francisco

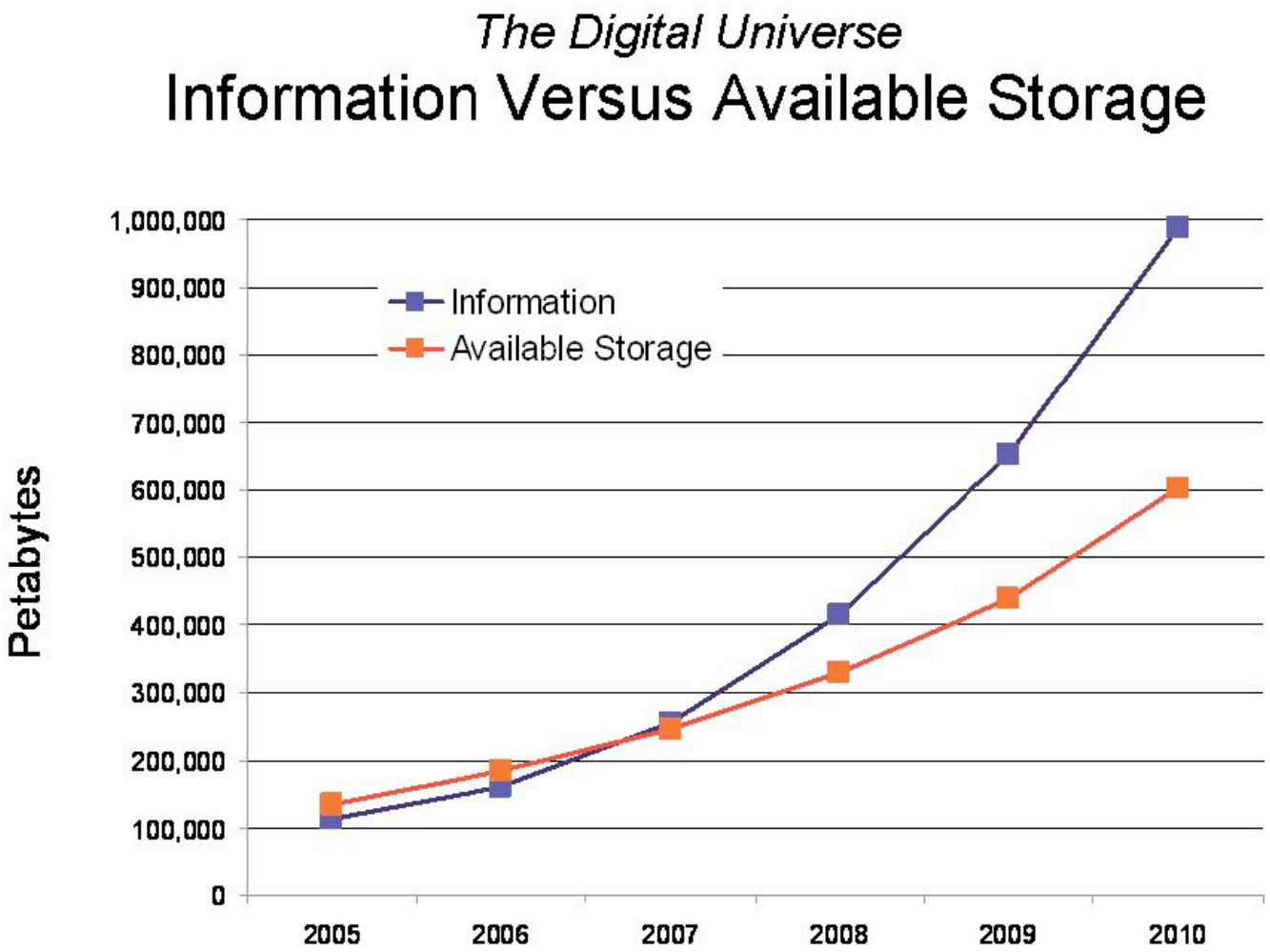
Where to go from here …



http://www.sis.pitt.edu/~repwkshop/index.html



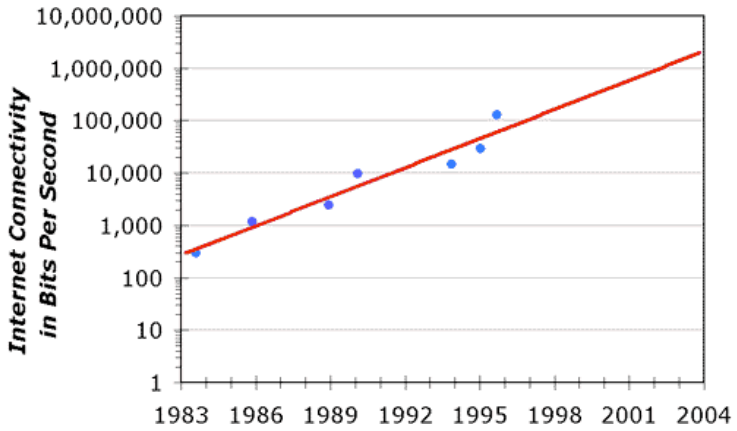
http://www.sis.pitt.edu/~dlwkshop/



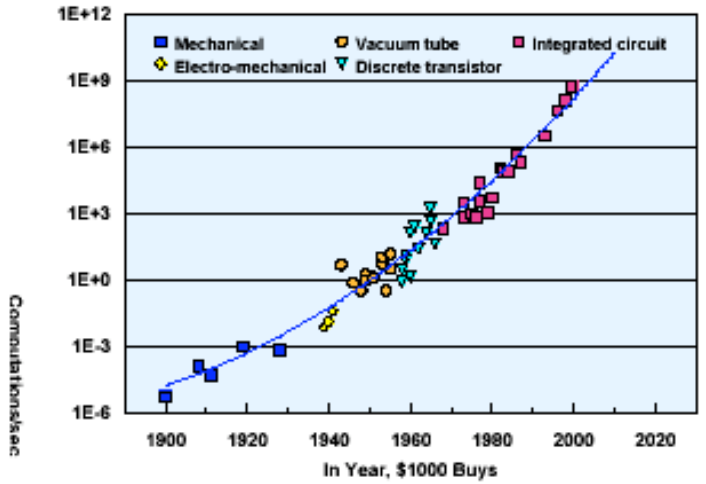
IDC White Paper: The Expanding Digital Universe

Exponential Changes in Core Technologies

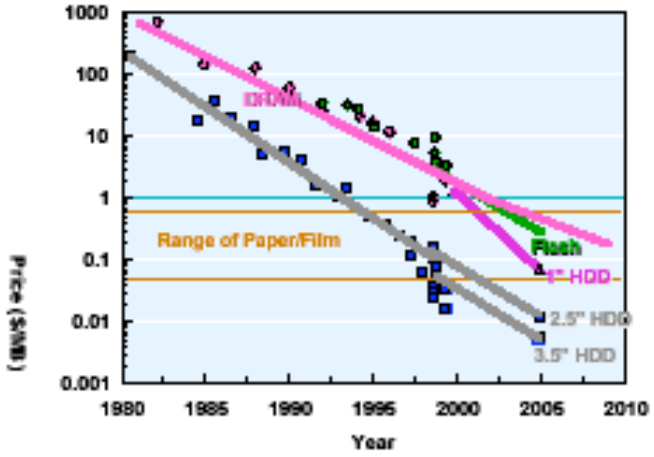
Internet Connectivity (Bits per Second)



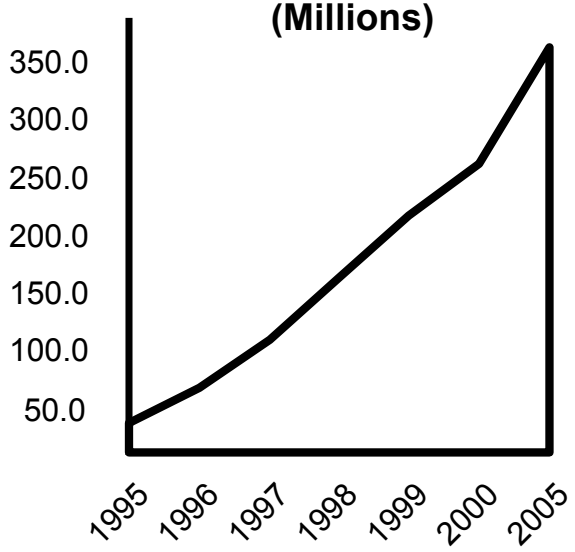
Average Price of Cycles



Average Price of Storage



People on the Internet (Millions)



Proliferation of Institutional Repositories

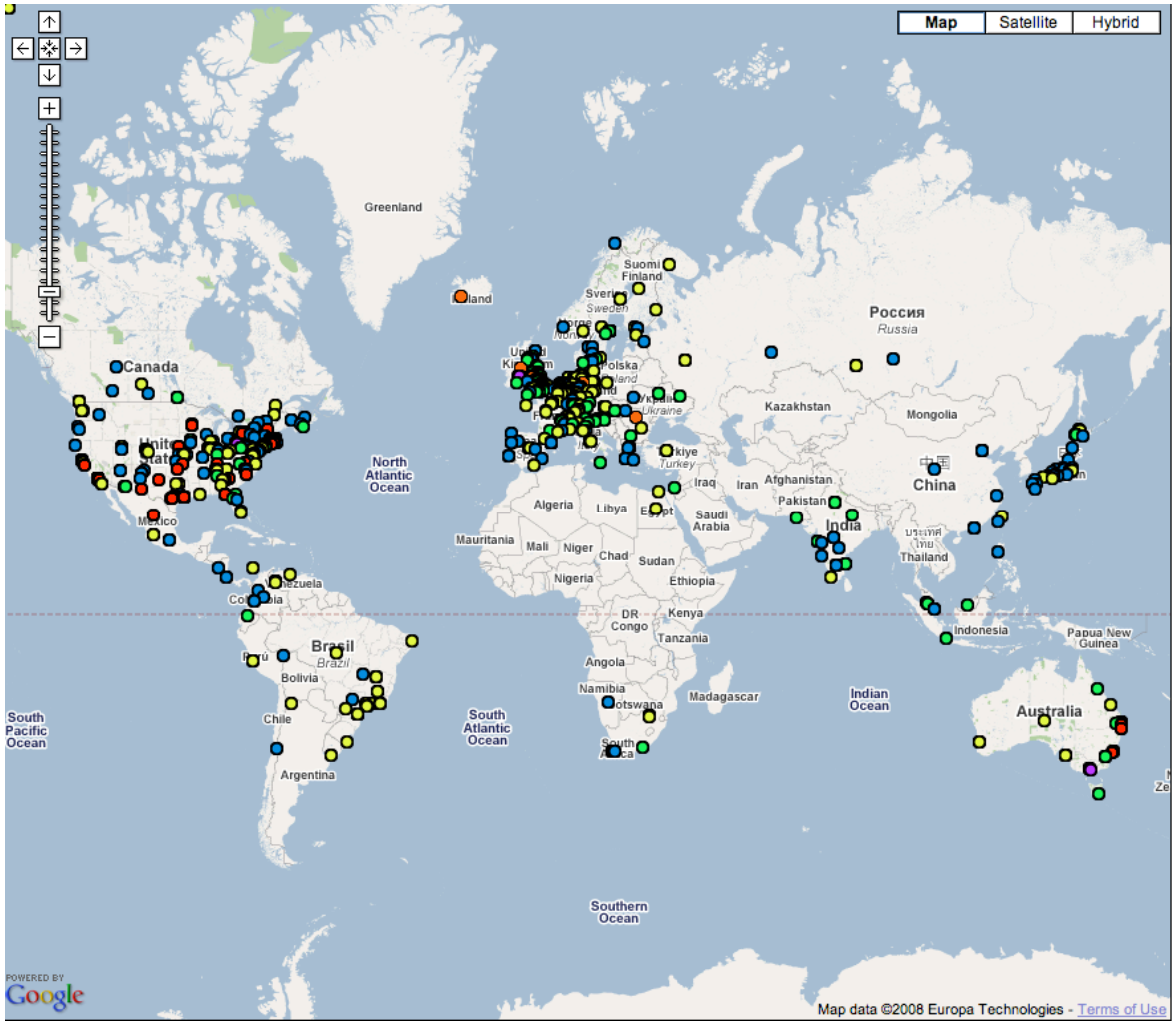
Registry of Open Access Repositories (ROAR)

http://roar.eprints.org/

The Directory of Open Access Repositories - OpenDOAR

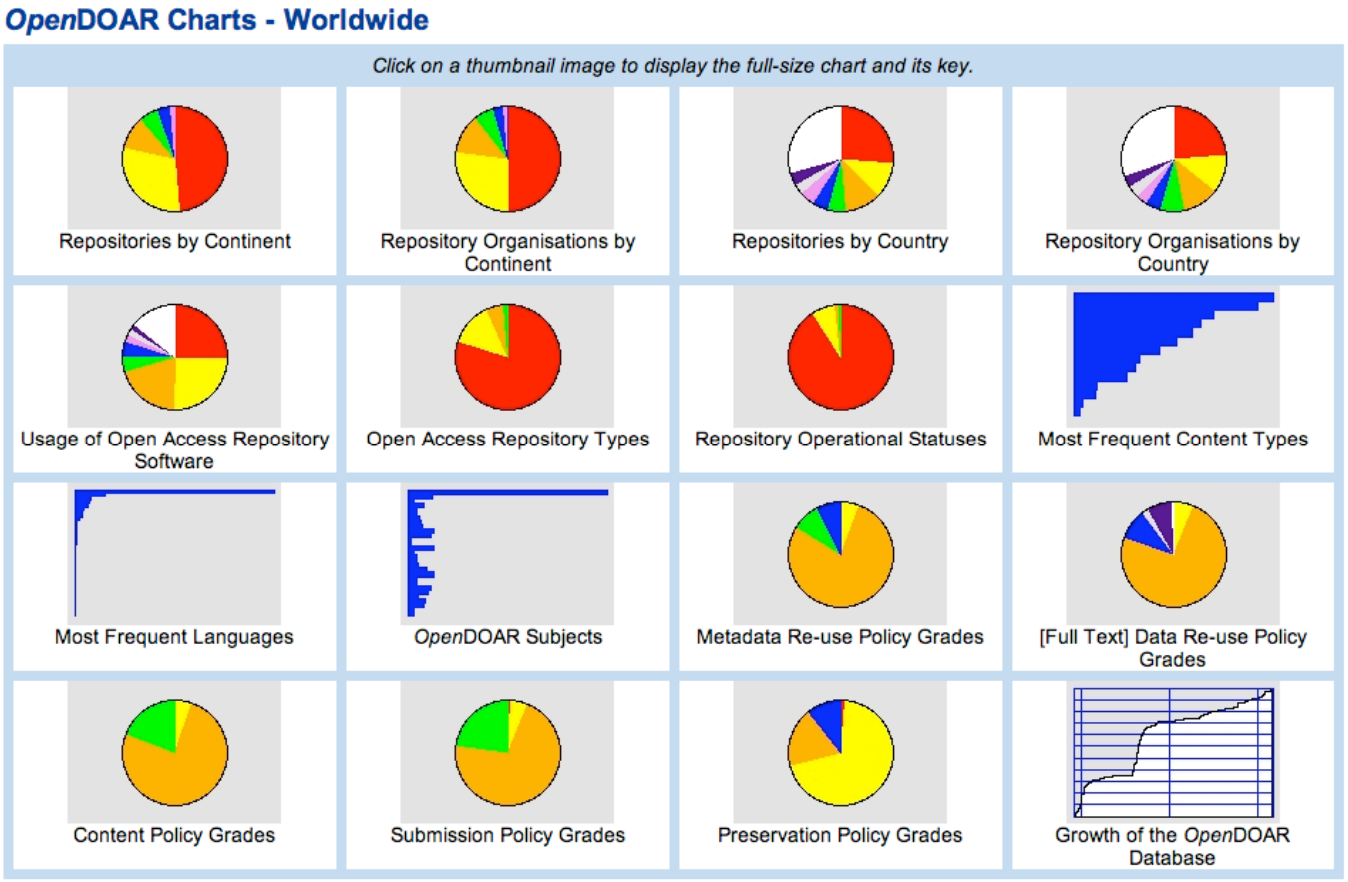
http://www.opendoar.org/

Repository Maps



http://maps.repository66.org/

The Directory of Open Access Repositories - *Open*DOAR



Thoughts about Repositories, Use, and Re-Use

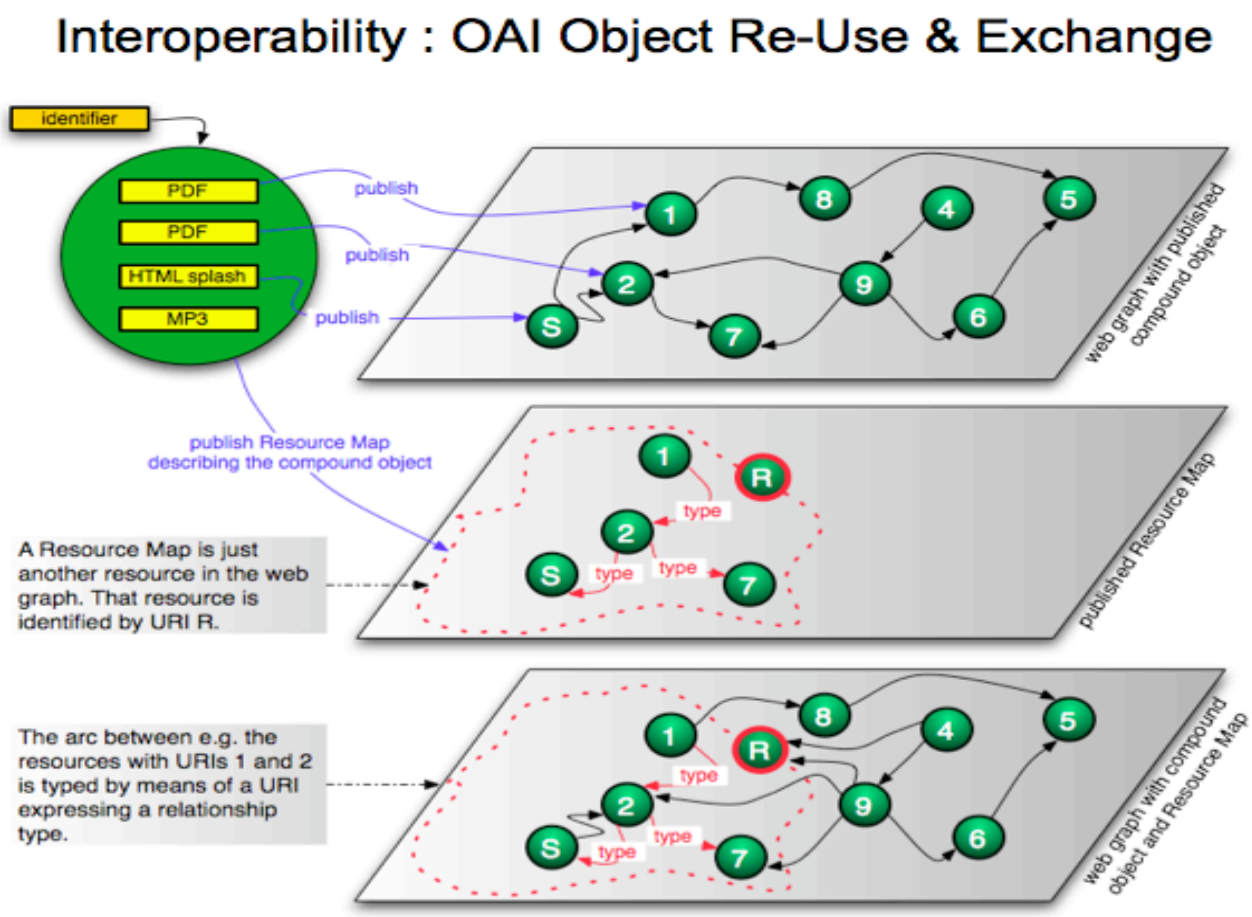
Herbert Van de Sompel

Digital Library Research & Prototyping Team
Research Library
Los Alamos National Laboratory, USA

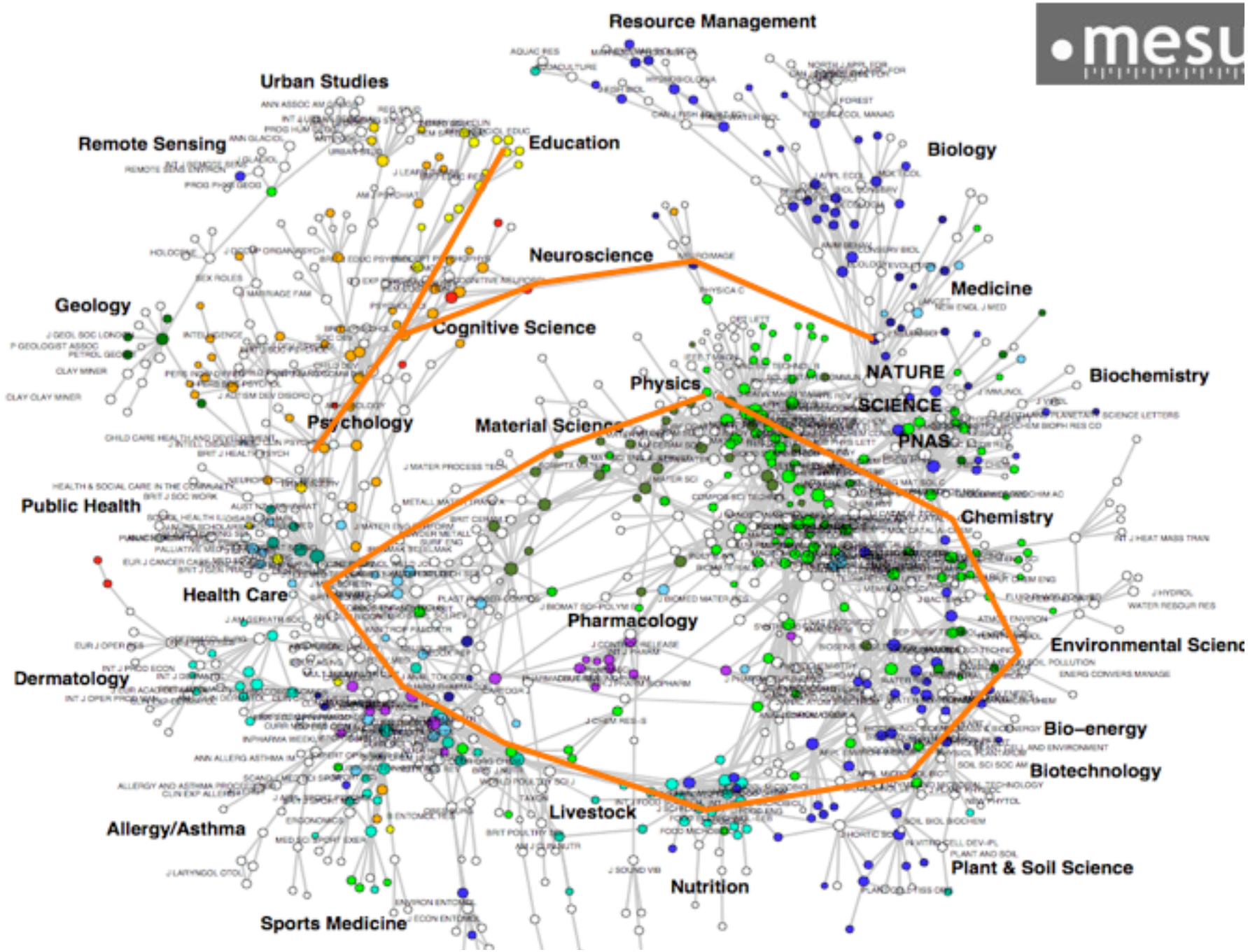
The European Library Conference
Frankfurt, Germany
January 31st 2008

With ideas from: *Carl Lagoze, Herbert Van de Sompel, Michael Nelson, Simeon Warner, Robert Sanderson, and Pete Johnston. Object Re-Use and Exchange: A Resource-Centric Approach. Submission to JCDL 2008.*

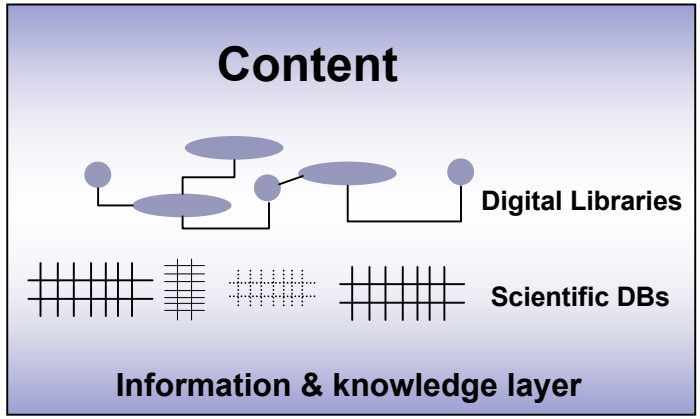
Open Archives Initiative
Object Reuse and Exchange

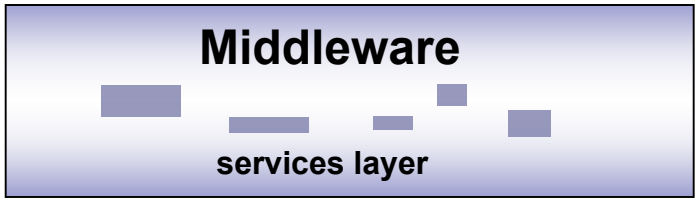


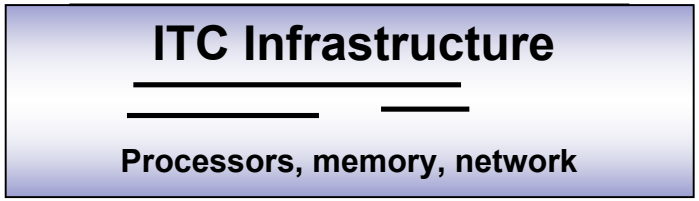
http://www.openarchives.org/ore/0.1/

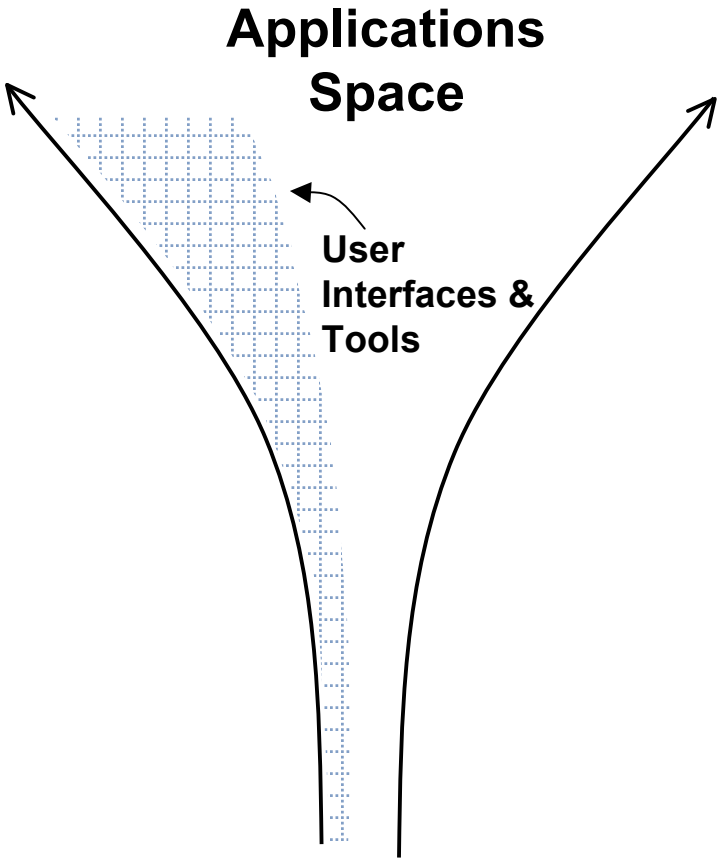

Resource Management
Urban Studies
Remote Sensing
Education
Biology
Neuroscience
Medicine
Geology
Cognitive Science
NATURE
Biochemistry
Physics
SCIENCE
Psychology
Material Science
PNAS
Public Health
Chemistry
Health Care
Pharmacology
Environmental Scienc
Dermatology
Bio–energy
Biotechnology
Livestock
Allergy/Asthma
Plant & Soil Science
Nutrition
Sports Medicine

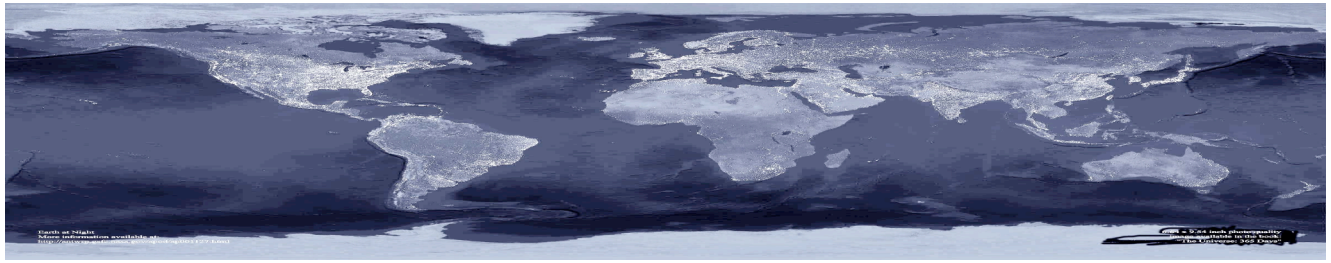
Cyberinfrastructure: An Evolving View











Thank You