

Cultural and Technical Transformations: Benefits from Yale's Linked Data Cross-Collection Discovery Platform

<https://lux.collections.yale.edu/>

Susan Gibbons

Vice Provost for Collections and
Scholarly Communication

Robert Sanderson

Director for Cultural Heritage
Metadata Projects

CNI Fall Forum, 12th December 2022



Yale's Cultural Heritage Units



Yale University Art Gallery



Yale Center for British Art



Yale Peabody Museum of Natural History



Yale University Library

*susan.
gibbons,
robert.
sanderson
@yale.edu*

Teaching with Collections



susan.
gibbons,
robert.
sanderson
@yale.edu

susan.
gibbons,
robert.
sanderson
@yale.edu

Lost Opportunities



susan.
gibbons,
robert.
sanderson
@yale.edu

Lost Opportunities

Monday January 5th 1767. * Mr. Cook & I had been getting
the other, &c. * gave Cook & I each a new Bill.
Cleared out the Company, &c. * P.M. * Part of the
had the Company's Truck Cleared out, &c.
about Noon sent the 2 Bbls of Wine, by one Cook &
* G.M. * Mittella & Moria or Miah, came here to work
and brought a Note of yesterday, date from Mr. Johnson
* G.M. * from Mr. Magratta, in the Cook's Office, *
* gave Dr. 2 Bbls.
* Shores having trading, &c. but both of us in my
Occupies frequently. * Mr. Cook's Journey in the
* with a little Copper Wood, I had a load of some = 1000.
* sent Drago to the other side for Jones (had to send
* some of them.
* P.M. * went to Mr. Lavelley's Shop, by Drago, & sent a
* enough to have the Neck of the Cook's Bill straight, which
* Robt. has kept away, at Night and it, a little better.
9



Tried this before...

Yale

Model for Campus-Wide Content Strategy

Office of Digital Assets & Infras
CNI-12.2009

Meg Bellinger, Director
Ann Green, Digital Strategic A
Louis E. King, Digital Information

odai@yale.edu

The screenshot shows the 'Discover Yale Digital Content' search page. At the top right, there are links for 'Yale CAS Login' and 'Language'. The main heading is 'Discover Yale Digital Content'. Below this, a search bar is present with a dropdown menu set to 'All Fields'. A blue 'Find' button with a magnifying glass icon is next to the search bar, followed by a link for 'Advanced'. The text below the search bar states: 'Search 1,500,000 records; 250,000 digital assets from across Yale's collections'. At the bottom, a list of participating institutions is shown: Yale Center for British Art, Yale Peabody Museum, Yale University Art Gallery, and Yale University Library.

What's Changed?

- 2016: est. deputy/vice provost for collections position
- 2017: CIO, John Barden, created cultural heritage pillar in IT
 - Pathway to central funding
- 2018-20: New leadership in collections
- Commitment to IIF and desire for cross-collection discovery
 - \$\$, new FTEs, and in-kind contributions by over 100 staff
- Awarded Andrew W. Mellon Foundation grant 2020-2022
 - Rob Sanderson

Principles and Culture

- Units retain ownership of their data
 - No new systems of record, including LUX
- Good enough is better than perfect
 - Collaboratively and incrementally work towards our goals
- Expertise exists throughout the institution
 - Development process promotes engagement and growth
- Collaboration takes us farther
 - And is rewarded by the university

Process

- Build consensus through engagement and experimentation
- Maintain an exciting vision and clear direction through project management and leadership at all levels
- Establish shared priorities and values, rather than trying to address all the concerns of all units
- Agreed-upon governance and cost sharing models
- Establish and formalized norms for collaboration

Results: Cultural Transformation

- Collaboration is contagious
 - Shared collection storage, positions, conservation labs, digital preservation, provenance expertise, diversity, equity & inclusion plans...
 - Peer connections forming
- Trust has led to generosity

*susan.
gibbons,
robert.
sanderson
@yale.edu*

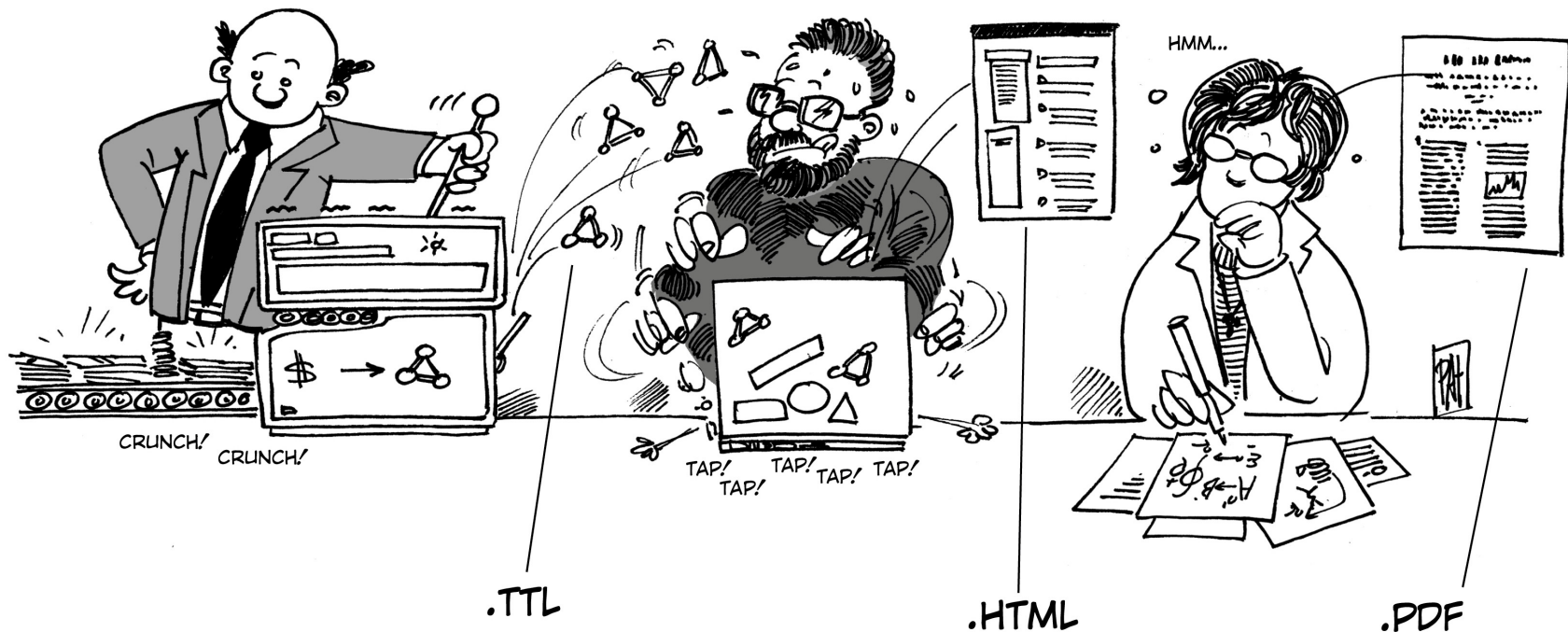


LUX Technical Principles: LOUD

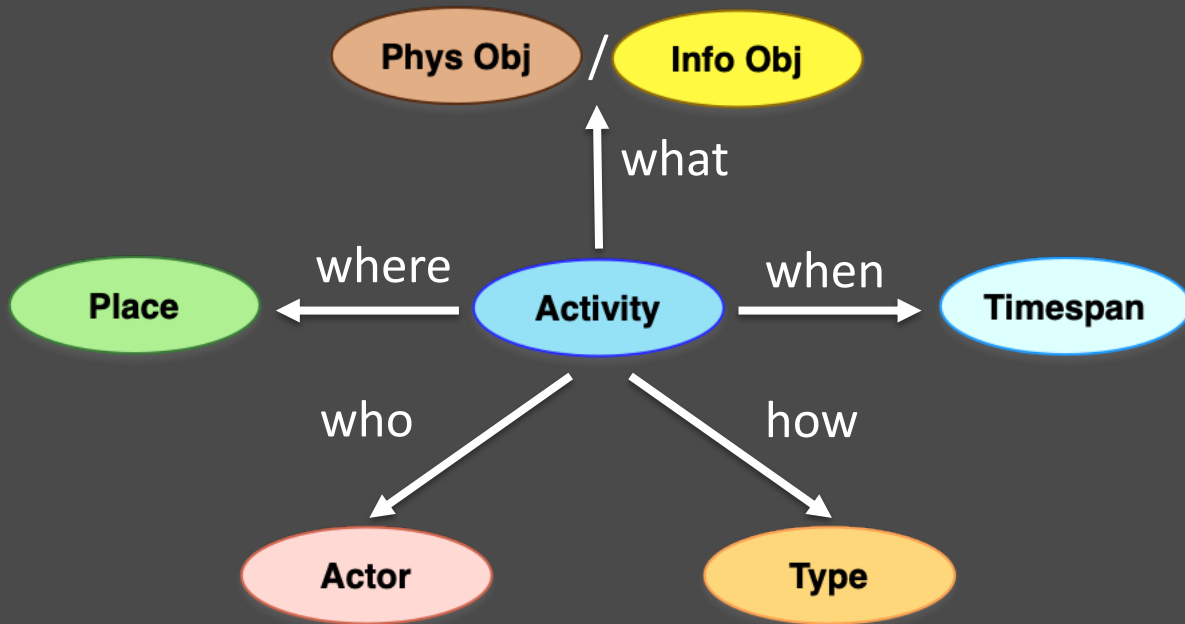
- LUX is LOUD: Linked Open Usable Data
- Standards are important
 - Too many to choose from with multiple domains!
 - Accurate semantics for cross-domain interoperability
 - JSON-LD syntax to ensure usability
- Using Linked Art: <https://linked.art/>
- Using IIIF: <https://iiif.io/>

*susan.
gibbons,
robert.
sanderson
@yale.edu*

LOUD: Easy to Use ... by Developers!



Linked Art Model From 50,000 Feet



*susan.
gibbons,
robert.
sanderson
@yale.edu*

LUX Technical Principles: FAIR

- LUX is FAIR: Findable, Accessible, Interoperable, Reusable
- Every attempt made for:
 - Persistence of identifiers and sustainable system
 - Open licenses for images and content
 - Use of standards throughout
 - Availability of the data

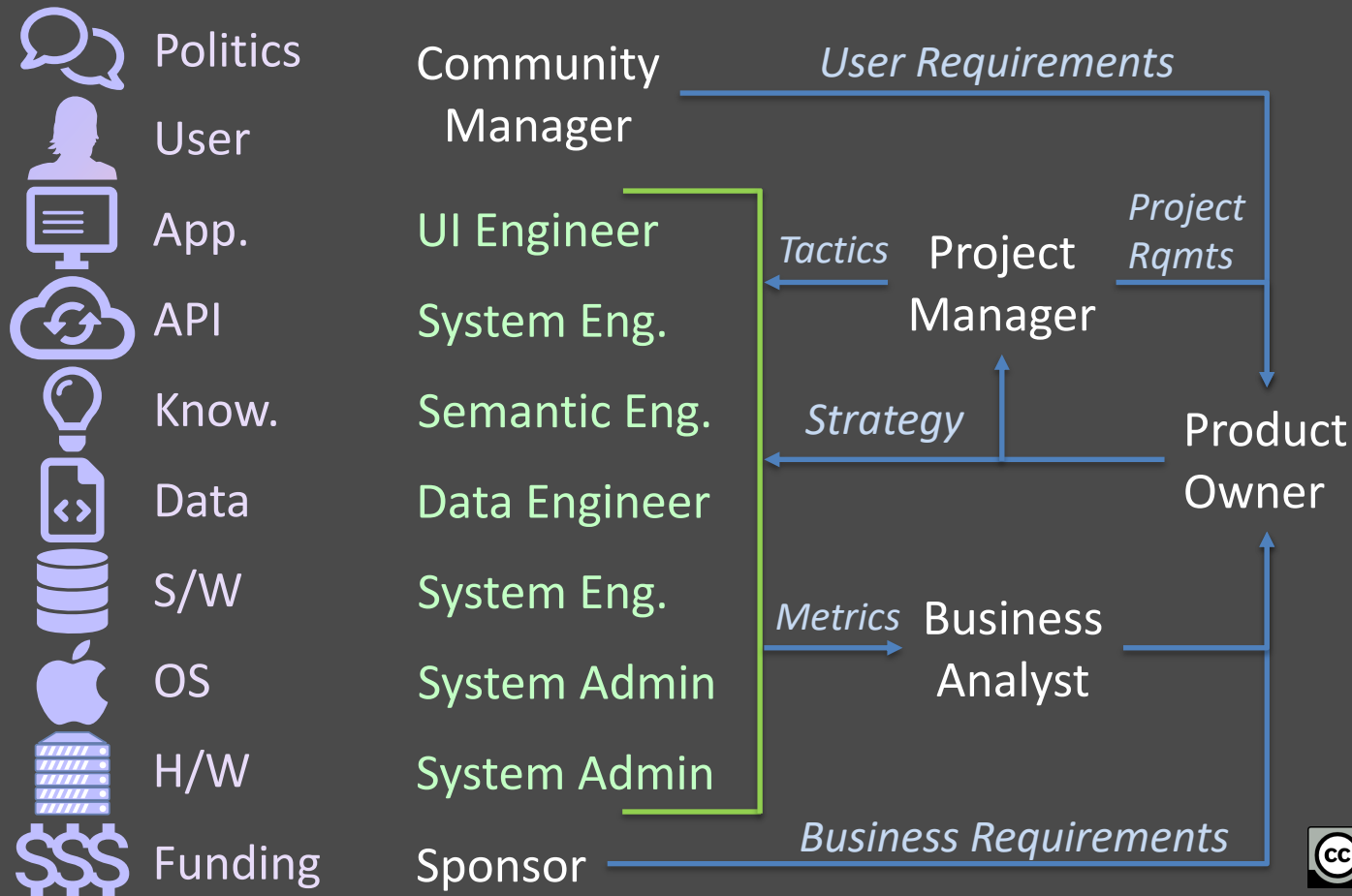
*susan.
gibbons,
robert.
sanderson
@yale.edu*

LUX Technical Principles: SHARED

- LUX is SHARED:
 - Sustainable – Shared costs, integrated services
 - Harvestable – IIF Change Discovery
 - Available – Open licenses, links to data, standards
 - Reconciled – Entities are aligned with 12+ systems
 - Enriched – ... and enriched from their knowledge
 - Diverse – Every effort at inclusion and diversity

*susan.
gibbons,
robert.
sanderson
@yale.edu*

Project / Product Roles



Technical Design and Process

- Practical approach to development of an ambitious project
 - Incremental prototypes – always ensure there's a fallback
 - Separation of concerns via standards
- Collaboration and research for a solution architecture
 - Extensive analysis for back-end, multi-modal database
 - Must allow full-text, graph, and hybrid queries
 - Front-end interacts primarily with JSON-LD documents
 - Not via custom APIs or SPARQL

*susan.
gibbons,
robert.
sanderson
@yale.edu*

LUX: Cross-Collection
Discovery at Yale

*susan.
gibbons,
robert.
sanderson
@yale.edu*

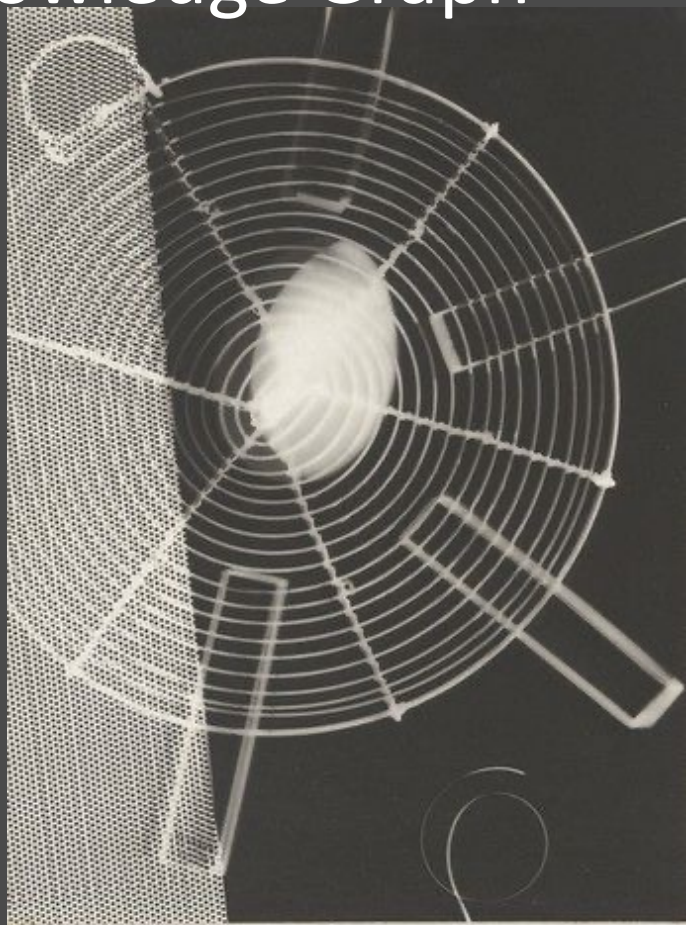
Yale

Demo!



Observation: LUX is a Knowledge Graph

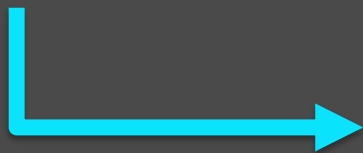
- More like the web than a database
 - Everything is connected
 - Expected search features are available, augmented by useful graph features
 - Shifts the paradigm from known-item search
 - ... to browse, explore and discover
 - ... and towards cultural research



Observation: Reconciliation is Critical

- Connections across collections require reconciliation
 - Should have one record per entity, not one per dataset
 - Allows enrichment from external sources
 - Richer records are more relevant, and rank higher in search

“Trumbull, John, 1756-1843”



John Trumbull, 1756-1843

Objects Made/Discovered Works Created/Published Works About

The Transfiguration
Agents: John Trumbull...
Supertypes: Panel painting

Joshua at the Battle of Ai - Attended by Death
Agents: John Trumbull...
Supertypes: Panel painting

The Battle at Bunker's Hill
Agents: John Trumbull, Johann Gotthard von Müller...
Supertypes: Engraving

preferred terms Trumbull, John, 1756-1843
John Trumbull
約翰·特倫布

Observation: Cross-Domain is Hard

- Different domains have different descriptive practices
 - Object vs Work vs Collection
 - Requires a core ontology to be profiled and clarified
 - Domains expect domain-specificity, even aggregated
- Labels are traditionally used as identifiers
 - A mental shift for practitioners needed for *things not strings*
 - Cross-domain reconciliation can introduce unexpected results, e.g. “landscapes” = “ecological zones” ?



Landscape

Produced By Aiseki

Production Date late 18th–early 19th century

Production Location Japan

Type Panel painting, visual artwork, hanging scroll, Ecological zones

Materials Ink

Collection Objects in the Yale University Art Gallery (YUAG), Asian Art

Observations: Resources are Needed

- Trail-blazing is costly
 - Staff time
 - Technology licenses and infrastructure
 - Professional services
 - Time
 - For discussion, consensus and understanding
 - For implementation and testing
- Hope to reduce cost for others to follow
- There are many flowers along the way!



*susan.
gibbons,
robert.
sanderson
@yale.edu*

Conclusions: Known Unknowns

- Extent of Adoption?
 - Still being developed, tested
 - Launch event in spring
 - Challenges of a new site, UI & paradigm
 - What is “intuitive” *and* “innovative”
- Ease of Sustainability?
 - Is it “mission critical”?
 - Mitigation: shared costs, automation
 - Phase 2: LUX as a service, not a site

susan.
gibbons,
robert.
sanderson
@yale.edu



Conclusions: Shareable Outcomes

- Data and Image content
- Data processing tools
- Database configurations
- Website code
- Technical documentation
- Process documentation
- Other? Ask!

The screenshot displays a GitHub repository page for 'LUX' with several sub-projects listed. Each project entry includes its name, visibility status, description, programming language, star/fork counts, issue count, and update time. A green line graph is visible on the right side of each project entry, likely representing commit activity over time.

Project Name	Visibility	Description	Language	Stars	Forks	Issues	Updated
marklogic	Private	Code, issues, and resources related to LUX MarkLogic	XQuery	0	0	64	9 hours ago
lux-web	Public	Web front end of LUX	TypeScript	0	0	99	5 days ago
lux-project	Public	Issues and resources that don't belong to a specific code repository	-	0	0	3	12 days ago
lux-crop-tool	Private	-	TypeScript	0	0	1	20 days ago
lux-web-cache	Private	Varnish cache in front of the middle tier	Shell	0	0	0	26 days ago
lux-backend	Internal	Node backend for LUX frontend (a.k.a. middle tier)	JavaScript	0	0	3	Updated on Oct 24
lux-database	Private	PostgreSQL database definitions for LUX	-	0	0	0	Updated on Sep 12
lux-harvester	Private	Harvester of LUX unit records	Go	0	0	0	Updated on Apr 5

Shifting Culture via Digital Transformation

- Moving from “Collections Discovery” to “Cultural Research”
- Improving our collection data, and hopefully others’
- Improving understanding of our internal culture
... and the originating cultures of the objects we curate
- Informing our understanding of cataloging practices
... across domains, systems, and formats
- Moving our culture towards effective collaboration

Thank You!

Discussion?