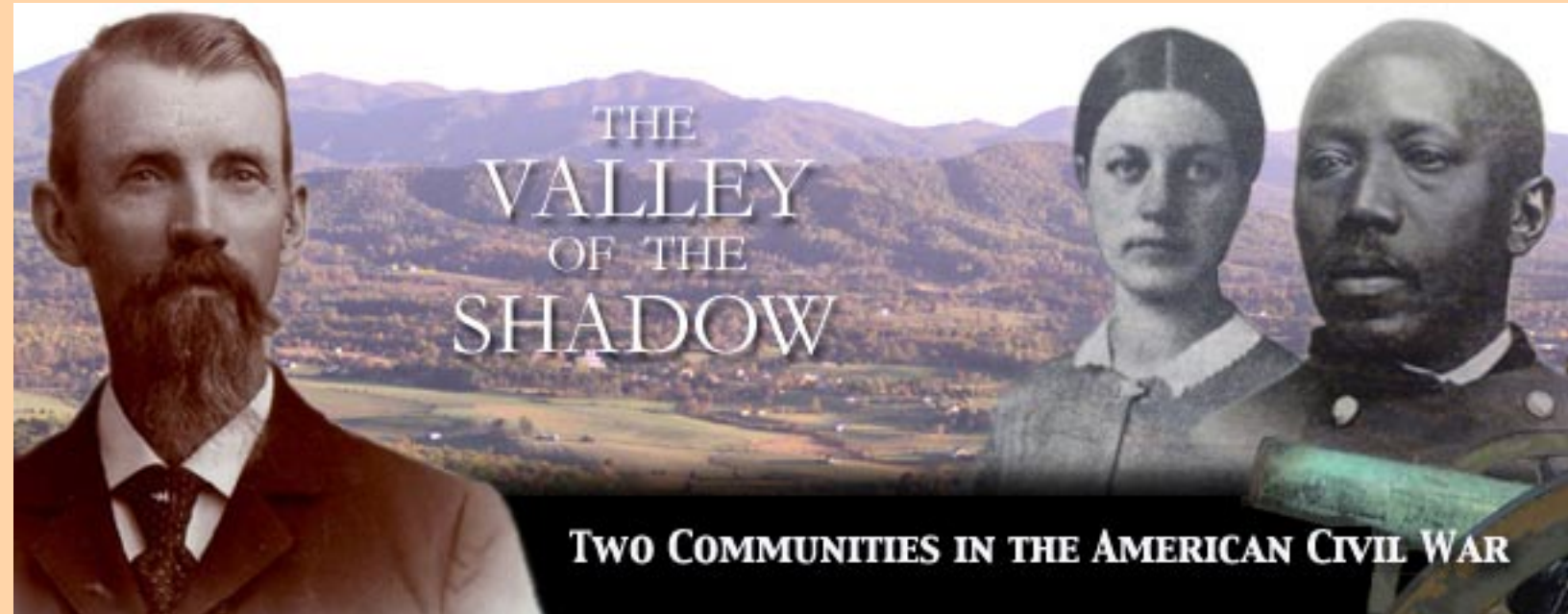


Practice Innovation: Moving a large, well-known web exhibit into a Library-managed content environment.

The Valley of the Shadow: Two Communities in the American Civil War was an early venture in web-based exhibition, and is still a heavily used online resource. As part of its Sustaining Digital Scholarship Initiative, the University of Virginia Library undertook the migration of this large collection from a stand-alone web site to part of the University's Library-managed content. How was this done? What techniques will make the end product a more durable version of what came before? Can older websites be modernized without undue effort? The University of Virginia Library's Digital Curation Services team offers their shared experience.

Legacy Site (Developed 1993 through 2007)



An early foray into web delivery, the Valley site was developed collaboratively over many years, many components added and changed in that time

The Valley site featured dozens of exhibit collections spread across multiple servers and domain names and was reliant upon software run by others, on (now) unsupportable systems

Database-driven search; distributed; custom search forms



38 different databases, queries via web forms (49 PHP & 189 Perl .cgi files)
(full-text document searching handled by OpenText, on separate server)

Specialists, Collaboration, Coordination

Lorrie Chisholm lead a team of student employees in re-digitizing all available artifacts featured in the site. Rights and provenance confirmed for thousands of digital images. Master images retained in central repository.
Brantley Craig undertook the rationalizing of the static web pages, resulting in the removal of approximately 35,000 pages of HTML. Working with Library developer **Ethan Gruber**, they re-engineered the site's search and presentation software.
Matthew Stephens consolidated the work of the development and image teams, deployed the new site on a test server and worked with Library IT personnel to finalize the migration and domain switch-over. Subject librarians and staff members of the **Virginia Center for Digital History** consulted regularly throughout the migration.

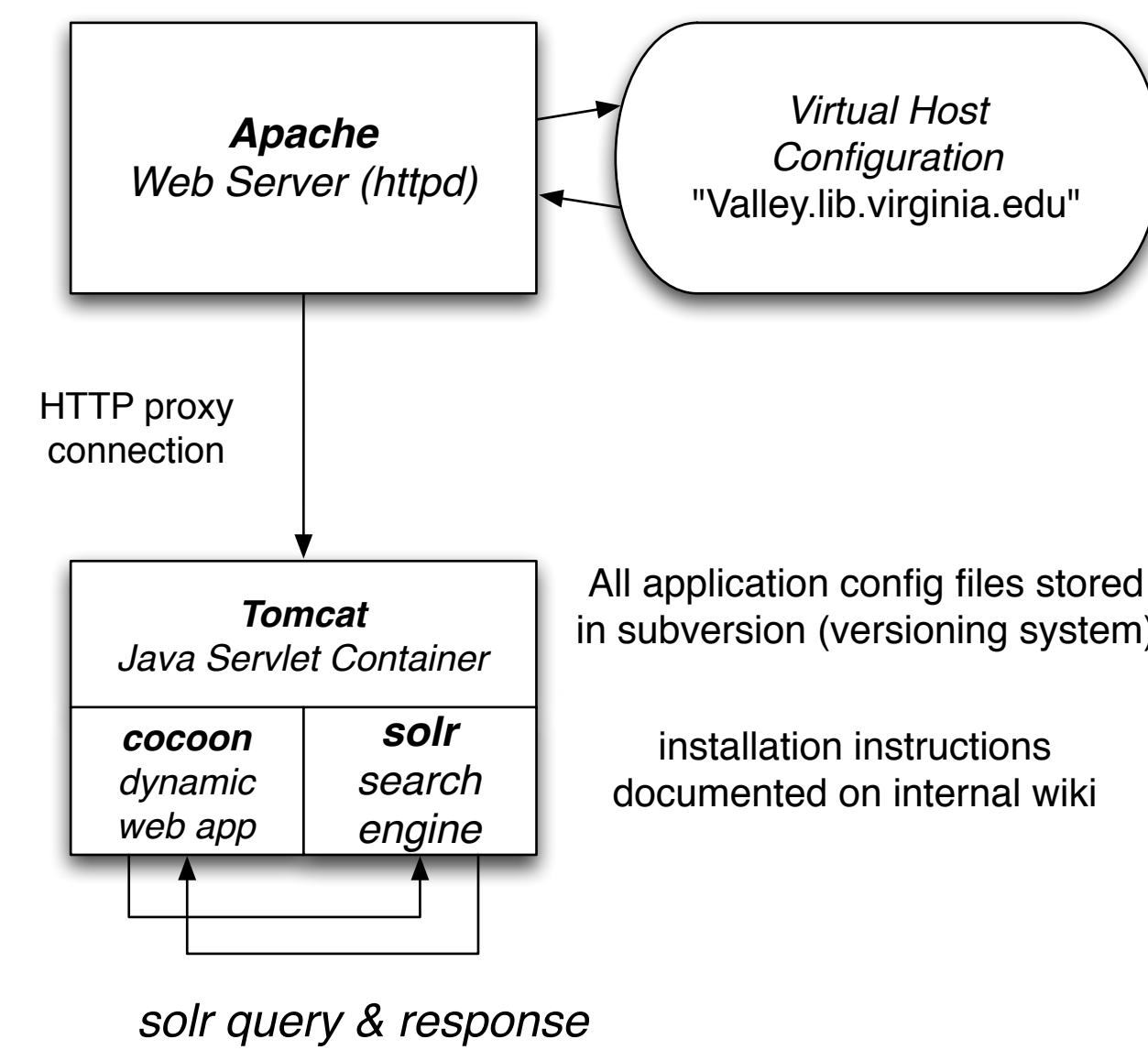
XML transcripts edited to match Library DTD, preservation standards and metadata requirements

all available original documents re-digitized to current preservation standards

UVA Library undertakes responsibility to exhibit collections as rights transferred from VCDH

Site Redesign & Migration (2008-9)

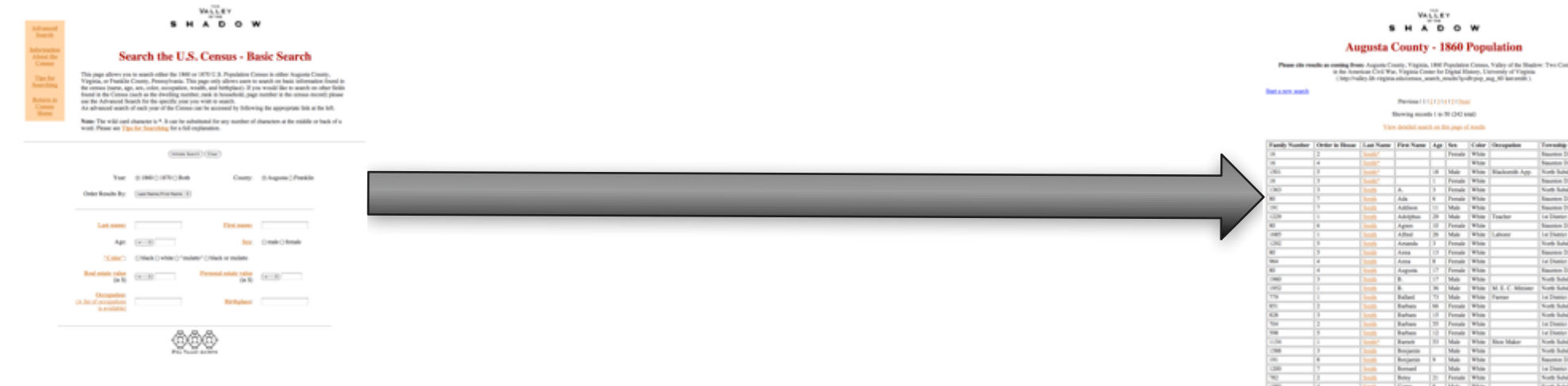
Architecture



development of an application bundle to serve all exhibit functions, using tools in use elsewhere in Library Online Environment

Search & Navigation

Strategy: Consolidate & Simplify datastore and web forms



existing web forms manipulated via a JQuery add-on, sending queries to a single Solr Index

1 library file (JQuery) + 1 .js file for each search form (28 files total)

Content Delivery

Strategy: use dynamic web software already used by Library



- open-source java-based web servlet
- deploys in same environment as Solr
- quick and easy XML/HTML manipulation

XSLT used to transform MySQL dumps to SOLR index files, AND to style all dynamic result pages; friendly learning curve increased collaboration

Content Management

Strategy: Preserve apparent structure, cull underneath

5.1GB of static web material purged (319MB kept)

re-organized HTML directories and edited links accordingly

re-organization documented via log records in versioning software

Use of cocoon's 'sitemap' permits reorganization of files while preserving old URLs

Library-Managed Site (Deployed July 2009)

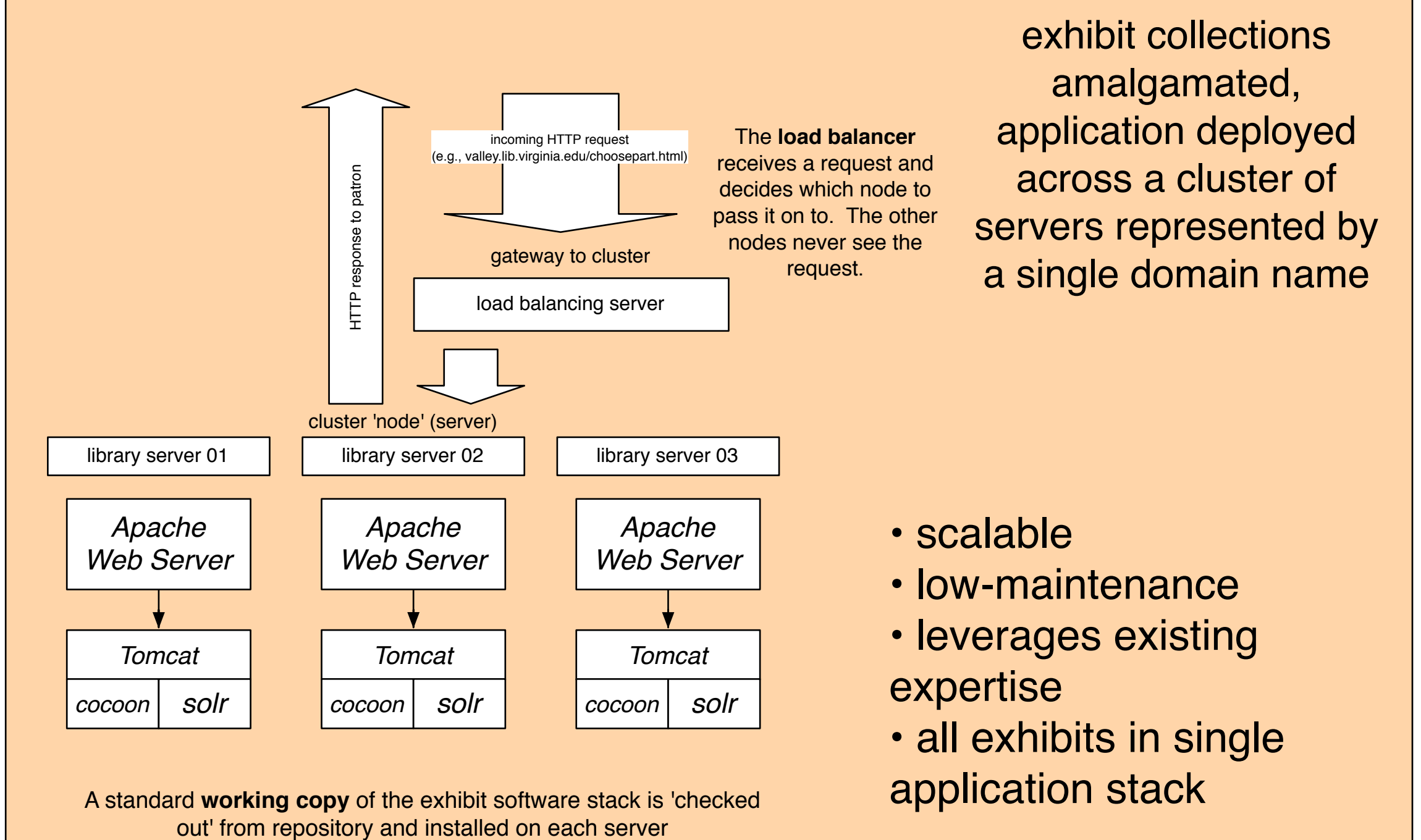
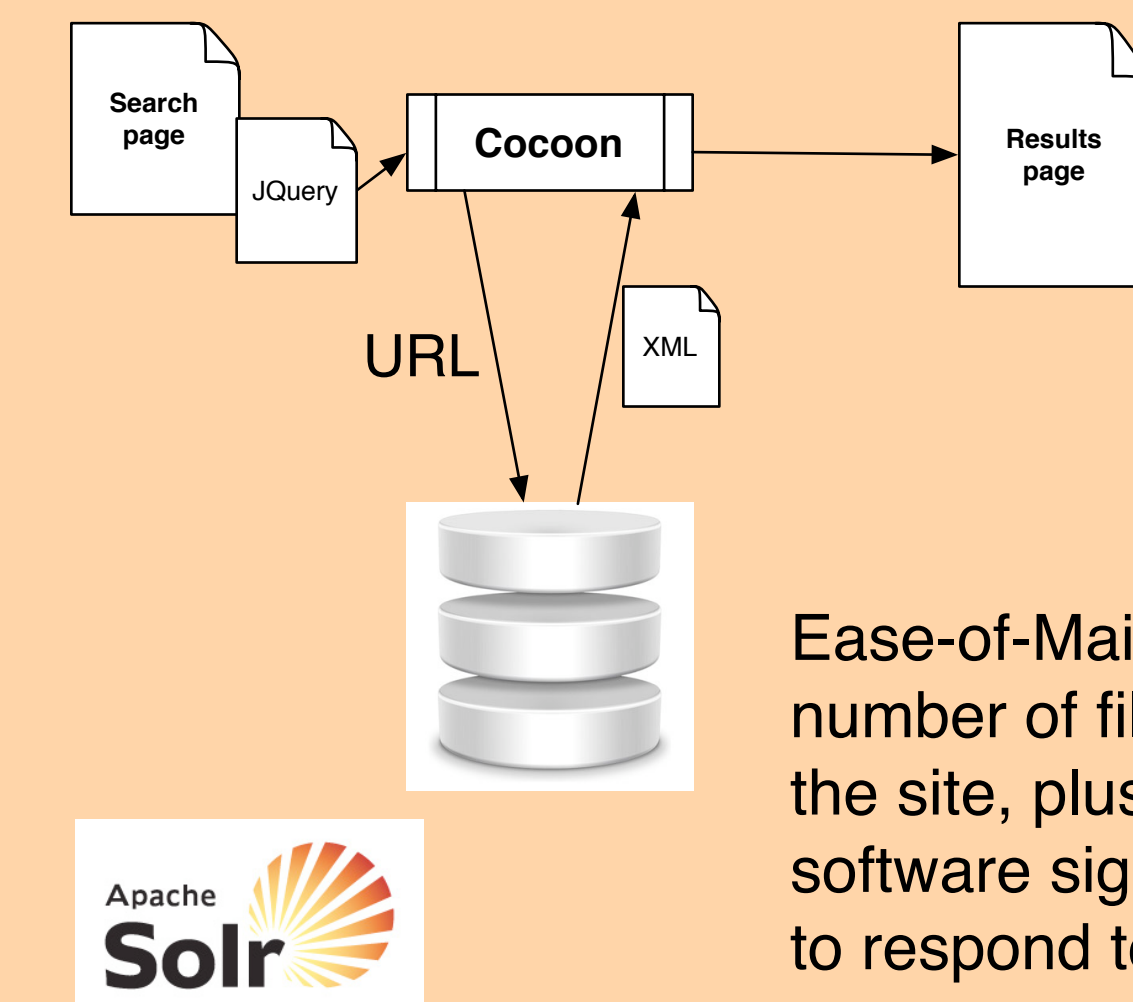


exhibit collections amalgamated, application deployed across a cluster of servers represented by a single domain name

- scalable
- low-maintenance
- leverages existing expertise
- all exhibits in single application stack

Index-driven search; consolidated; JQuery + Solr + XSLT



- single datastore
- easy to copy & migrate
- all request processed in a standard language (JQuery)

Ease-of-Maintenance Achievement: reducing the number of files constituting the dynamic portion of the site, plus logging change entries in versioning software significantly reduces time & effort required to respond to problems.

Trust, but Verify!

Open-source tools such as Integrity and Firebug provided easy testing of dynamic site's functionality

all files under version control for managed change

Memorandum of Understanding between donating and curating parties facilitated migration

Archival copy of original website stored separately

Installation documented on department wiki

same backup/restore technologies as other Library services

