

The background features a light blue and white color scheme with a pattern of binary code (0s and 1s) scattered across it. A prominent blue chain with circular links is draped across the center, extending from the left side towards the right. The chain is set against a backdrop of overlapping, semi-transparent blue and white circular shapes that create a sense of depth and movement.

OPEN SOURCE BUSINESS CONFERENCE

Building Your Big Data Future with Open Source

COMPUTERWORLD
OSBC
SAN FRANCISCO



Building Your Big Data Future with Open Source

Lessons Learned: Initiating and Coordinating a Tech Project Across Disparate Organizations

Patrick Schmitz

Semantic Services Architect

U.C. Berkeley



Building Your Big Data Future with Open Source

CollectionSpace is an open-source, web-based software application for the description, management, and dissemination of museum collections information – from artifacts and archival materials to exhibitions and storage.



 collectionspace



Building Your Big Data Future with Open Source

Project Partners

- Museum of the Moving Image, New York
- University of California, Berkeley, Information Services and Technology Division
- University of Cambridge, Centre for Applied Research in Educational Technologies
- OCAD University, Adaptive Technology Resource Centre, Fluid Project

MUSEUM
OF THE MOVING
IMAGE



caret.

Fluid



Building Your Big Data Future with Open Source

Funding

- The Andrew W. Mellon Foundation, Program in Scholarly Communications and Information Technology
- Institute of Museum and Library Services (IMLS)
- Collaborations with Mellon-funded projects
 - ArchivesSpace
 - OLE Project
 - ConservationSpace
 - Fluid Engage
 - Project Bamboo



Building Your Big Data Future with Open Source

Community Source

CollectionSpace development is based on the community source model:

“A hybrid model that blends elements of directed development, in the classic sense of an organization employing staff and resources to work on a project, and the openness of traditional open-source projects like Apache...the distinguishing feature of the Community Source Model is that many of the investments of developers’ time, design, and project governances come from institutional contributions... rather than from individuals. The project often establishes a software framework and baseline functionality, and then the community develops additional features as needed over time.”



Building Your Big Data Future with Open Source

Why am I here?

- Where's the Big Data in Museums & Archives?
- Why do they need Enterprise solutions?



Building Your Big Data Future with Open Source

ECM as a Platform

- Fills many needs, has wide support
- Major OSS players: Alfresco and Nuxeo
 - Alfresco has strong US presence, nice application
 - Nuxeo has strong platform focus for developers
- CMIS (OASIS) standard for interop
- Initially a hard sell in Higher Ed, seeing greater adoption with time



Building Your Big Data Future with Open Source

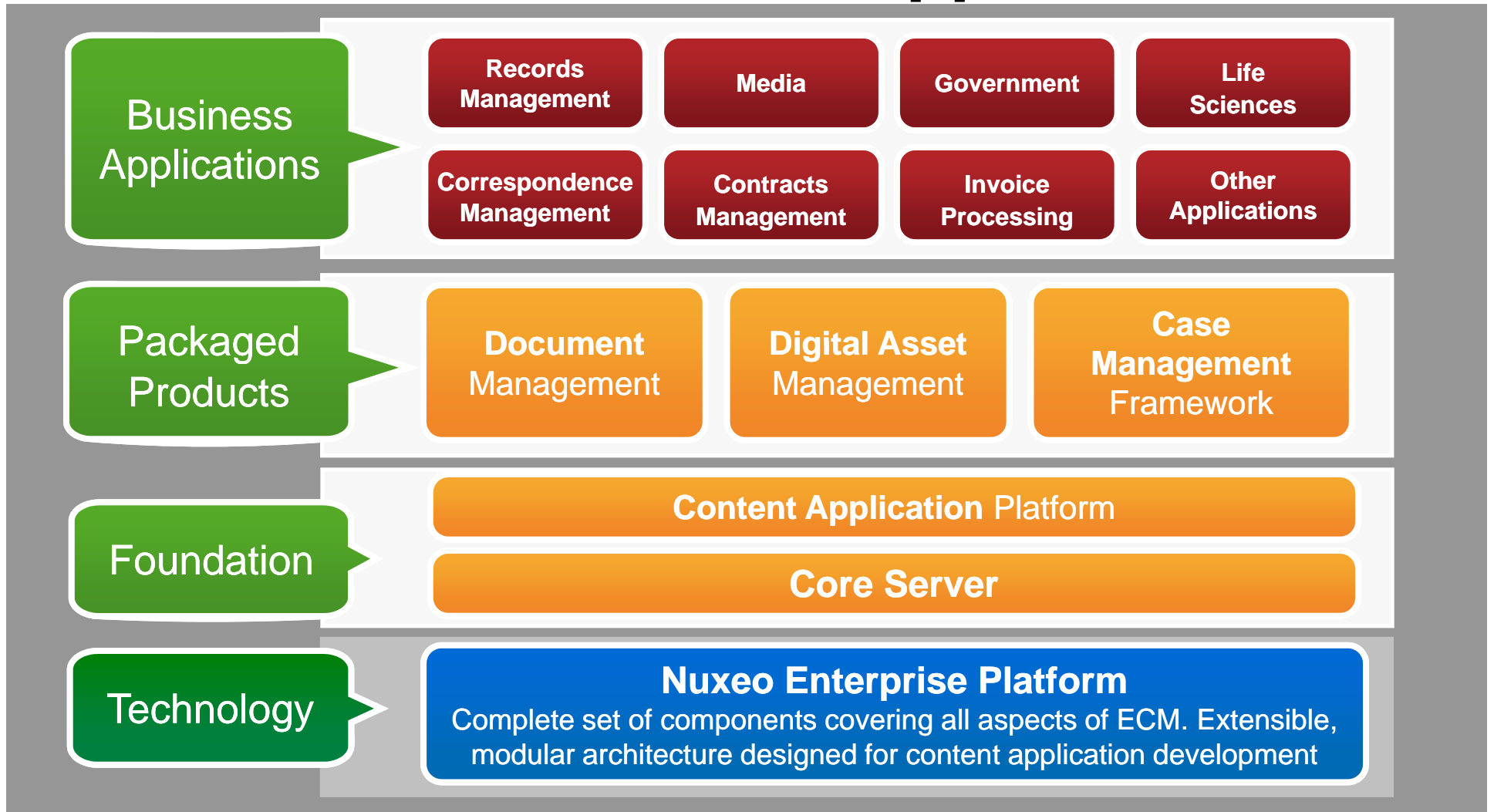
Nuxeo Background

- Focused on Enterprise Content Management, for 10 years
 - 500+ customers, 5000+ installations
- A platform approach for content applications
- Open platform, open company, embracing OSS
 - Sells support & maintenance subscriptions, with added value software and customization tooling, as services



Building Your Big Data Future with Open Source

Nuxeo Platform Approach





Building Your Big Data Future with Open Source

Community Source Project - Lessons

- Engaging your User Community
- Considering Partners for Development
- Project Roles
- Funding & Sustainability



Building Your Big Data Future with Open Source

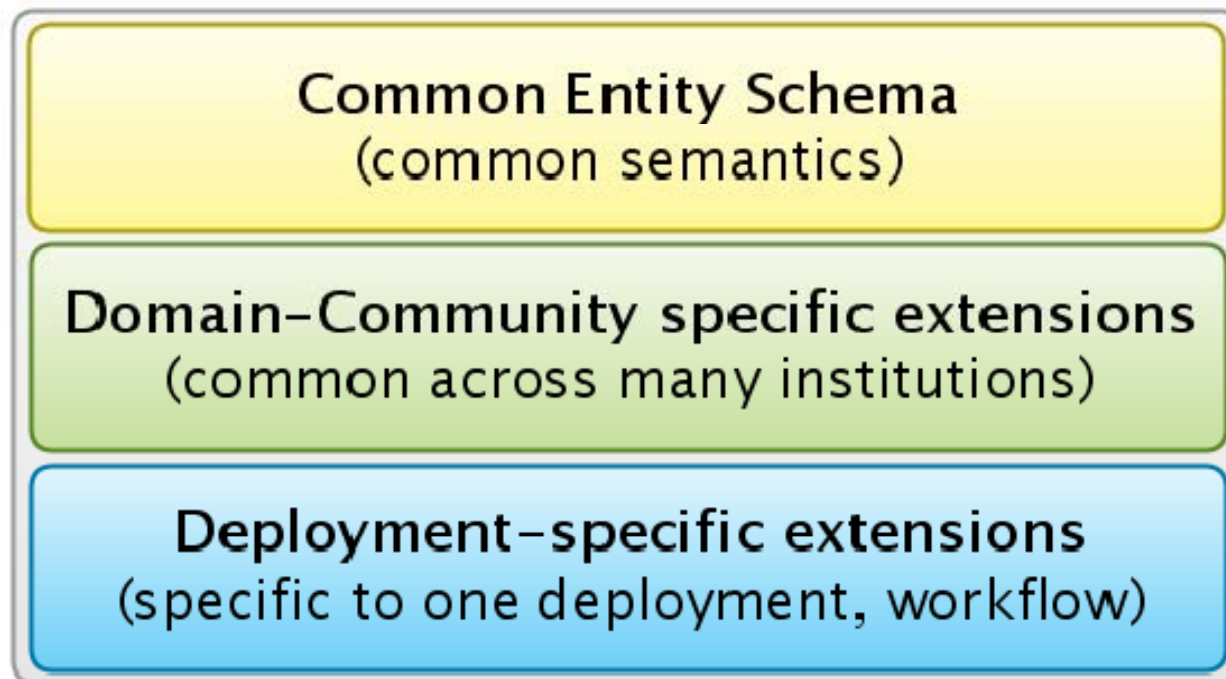
Engaging Your User Community

- Community design workshops
- Open communications
- Keep asking them questions
- Think early about adoption workflows
- Find and use natural domain communities



Building Your Big Data Future with Open Source

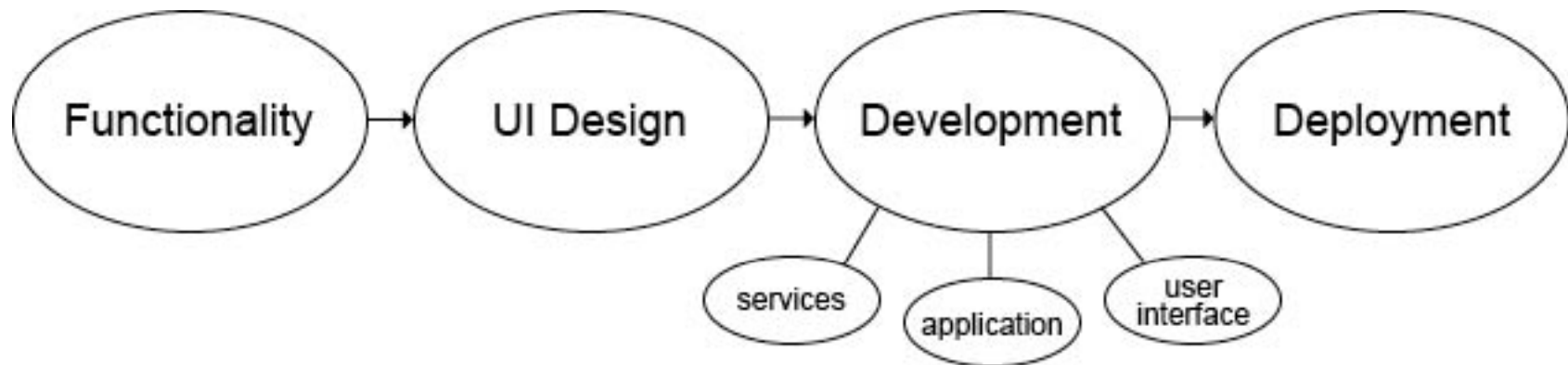
Data Architecture as Social Engineering



Schema model for a customized service deployment

Partners for Development

- Ensure your partners have skin in the game
- Localized vs. Distributed expertise
- UI/UX/UCD matter!





Building Your Big Data Future with Open Source

Project Roles

- More PM work with consortial projects
- Functional lead is essential, challenging
- Technical leadership crucial but delicate
 - Responsibility with limited Authority
 - Politeness can sink your project
- Cost/benefit of Test and Build support
- Supporting additional contributors



Building Your Big Data Future with Open Source

Funding and Sustainability

- Development vs. Deployment vs. Production
 - Different sources, different goals
 - Partner vs. Consortial
- Planning for sustainability
 - Chicken and Egg problem for adoption
 - Will your community pay?
 - Foundations, foundations.



Building Your Big Data Future with Open Source

Communications lessons

- Try various media (email, wiki, IRC, Skype, Connect) and see what works for what/whom.
- Don't fight time zones
- Invest in webinars
- Marketing (internal and external) *matters*.



Building Your Big Data Future with Open Source

Development lessons

- The many ways of Agile/XP...
- Get a stack up early!
- Use CI servers and automated testing.
- Hold periodic F2F dev camps
 - Trust underlies effective distance collaboration
 - Guided walkthroughs and shared tasks
 - Look costly on paper (for travel), but worth it.



Building Your Big Data Future with Open Source

Community Source and Governance

- Funding and internal governance is standard negotiation, etc.
- Technical governance is more of a challenge because of local constraints
- Services (SOA/ROA) governance combines local and peer project constraints.
- Traditional models need adaptation



Building Your Big Data Future with Open Source

Thanks! Questions?

- Email the team:
talk@lists.collectionspace.org
- Visit the website: www.collectionspace.org
- Browse the wiki: wiki.collectionspace.org