

https://dspace.org/

Institutional repositories: description of VITAL as an example of a Fedora-based digital assets management system.

Saadallah, Nabil

2012-06-14
http://hdl.handle.net/10539/11529

Downloaded from a DSpace repository. https://dspace.org/

Institutional repositories: description of VITAL as an example of a Fedora-based digital assets management system.

ICADLA-2, Johannesburg, South Africa



Nabil Saadallah
Manager, Middle East and Africa
VTLS Europe
Email: saadallahn@vtlseurope.com

Challenges

- Increase of “Born digital” content
- Increase of Digitized content
- Content is not only books or “paper”, but a variety and a collection of different media types
- Content is saved and available in several locations
- Need for an easy access mechanism to content
- Need for a preservation mechanism

‘Born digital’ Resources (1)

Increase in ‘born digital’ resources

These new formats, collections and communities of origin create new challenges for stewardship.

Examples:

- **e-Prints**
- **e-Journals**
- **Electronic Theses & Dissertations (ETDs)**
- **Research Publications**
- **Scholarly Work**
- **Electronic Classrooms**
- **Electronic Communications**

‘Born digital’ Resources (2)

Increase in ‘born digital’ resources

So what is at risk?

- **Resource discovery**
- **Long-term preservation strategies**
- **Collection maintenance**
- **Standards and policies**
- **Ownership**

'Born digital' Resources (3)

Impact – One example: Scholarship

1. Academic researchers are using newer technologies such as personal websites, blogs, wikis, and discipline-specific sites to post their scholarly work in advance of traditional publishing practices.
2. The dynamics are changing due to early and increased exposure, early response and citation from peers, and early identification of 'communities of research'.
3. Also, it is generally believed that open-access publications receive broader exposure to research communities of interest than do traditional, and more restricted, journal publication.

Proper management of these materials is rapidly becoming a necessity.

The Institutional Repository is an important solution to this issue.

Digital Preservation Issues

- Two Problems : viewers and media
- Digital content requires “viewers”
- Viewers depend on hardware / software (HW/SW)
- Obsolescence rate of HW/SW is very high
- Preservation depends on “migrations”
- Media life is unknown – may require “migrations”
- Repository services will help as Institutional Repository is “preservation friendly”.

Solution : Repository Services

Repository services are critical to developing, managing and leveraging enterprise-wide digital content and bringing greater value to institutional output. Clifford A. Lynch, Executive Director, Coalition Networked Information, defines an institutional repository as a "set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members".

- Lynch, Clifford A., "Institutional repositories: essential infrastructure for scholarship in the digital age", *ARL Bimonthly Report 226*, February 2003.

Institutional repository needs Repository services

What are Repository Services?

- Maintains storage of digital objects
- Exposes external interface for Digital Objects
  - Creation
  - Modification
  - Access
- Enforces access policies
- Provides for content type disseminations
- Offers preservation assistance

RS – is not just software

- Policies and guidelines for submission, storage, usage and dissemination need to be created by participating institutions.
- Digitization of content by author, institution or vendor
- Use cases need to be determined – so proper metadata formats can be selected
- Metadata types
  - Technical – needed for Digital Preservation
  - Descriptive – used for search, retrieval and display
  - Administrative – for rights, ownership, provenance, licensing

RS -- Goes by Many Names

- Digital Object Repository
- Digital Asset Repository
- Digital Repository
- Institutional Repository

The word Repository is always there! Each representing a slight variation in focus.

We use the terms: "Repository Services"

Acronyms in Repository Services

- **OAI-PMH** – Open Archives Initiative Protocol for Metadata Harvesting
- **SRU & SRW** – Search Retrieval URL / Search Retrieval Web
- **DOI** – Digital Object Identifier
- **RDF** – Resource Description Framework
- **METS** – Metadata Encoding Transmission Standard
- **FOXML** – Fedora Object XML
- **LDAP** - Lightweight Directory Access Protocol

Acronyms in Repository Services - Cont'd

- **XML** – eXtensible Markup
- **TEI** – Text Encoding Initiative
- **EAD** – Encoded Archival Description
- **MADS** – Metadata Authority Description Schema
- **MARC21** – MAchine-Readable Cataloging



VTLS Information Technology for Advanced Learning

What is Fedora™?

Flexible Extensible Digital Object Repository Architecture



Fedora Commons™

DURASPACE SERVICE PROVIDER

Fedora™ Advantage

- Extensible digital object model
- Repository exposed by Web services APIs
  - Management (Creation, Deletion, Maintenance, Validation)
  - Access (Search, Disseminations)
- Scalable, persistent storage for content and metadata
- Content can be local and/or remote
- Content versioning
- Open source solution

Fedora™ Digital Object Architecture

Persistent ID (PID)

Globally Unique Persistent ID

Disseminators

Public View: access methods for obtaining "disseminations" of digital object content

SystemMetadata

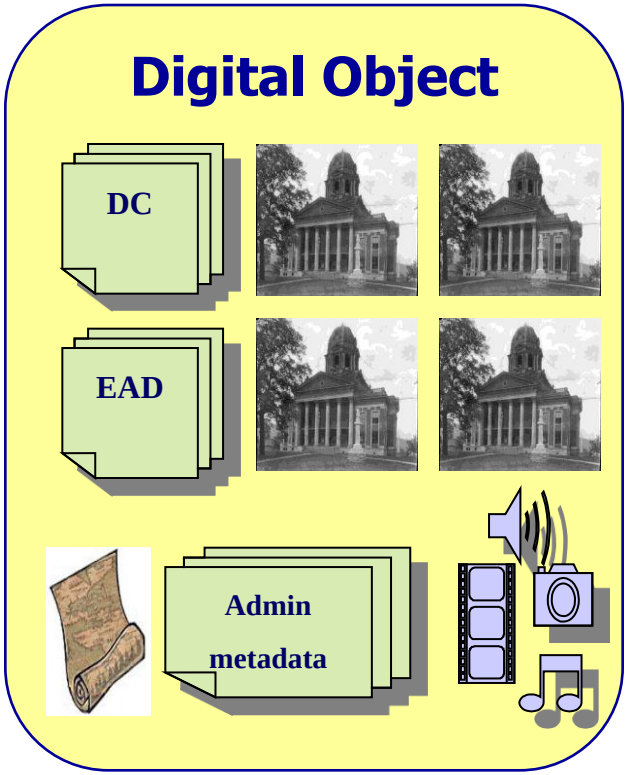
Internal View: metadata necessary to manage the object

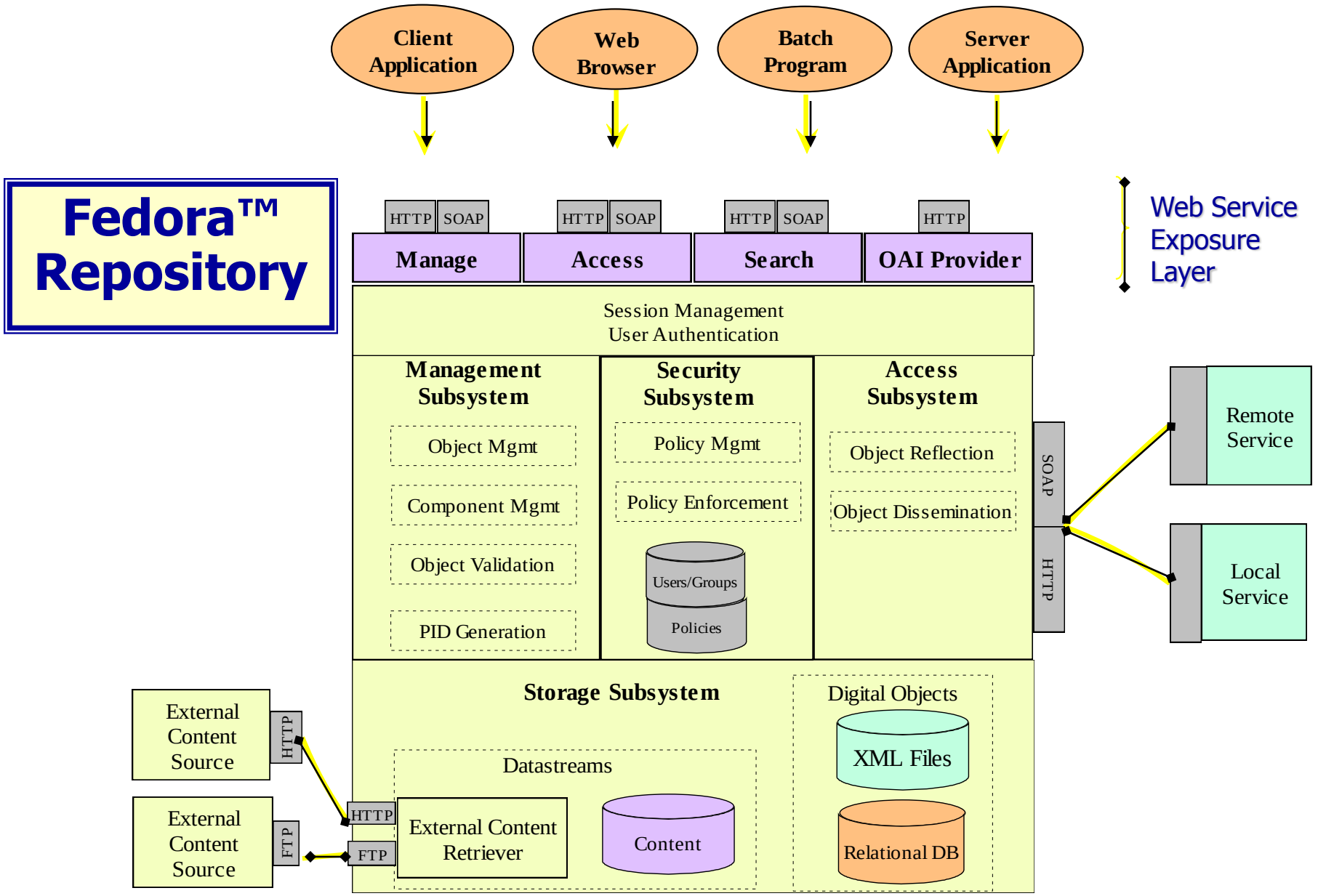
Datastreams

EAD, TEI, DC, MARC, VRA Core, MIX, etc.

Images, e-books, e-journals, music, video, etc.

Protected View: content that makes up the "basis" of the object




Client Application
Web Browser
Batch Program
Server Application
Fedora™ Repository
HTTP
SOAP
Manage
Access
Search
OAI Provider
Web Service Exposure Layer
Session Management
User Authentication
Management Subsystem
Object Mgmt
Component Mgmt
Object Validation
PID Generation
Security Subsystem
Policy Mgmt
Policy Enforcement
Users/Groups
Policies
Access Subsystem
Object Reflection
Object Dissemination
SOAP
HTTP
Remote Service
Local Service
Storage Subsystem
Digital Objects
XML Files
Relational DB
Datastreams
External Content Retriever
Content
External Content Source
External Content Source
HTTP
FTP
HTTP
FTP

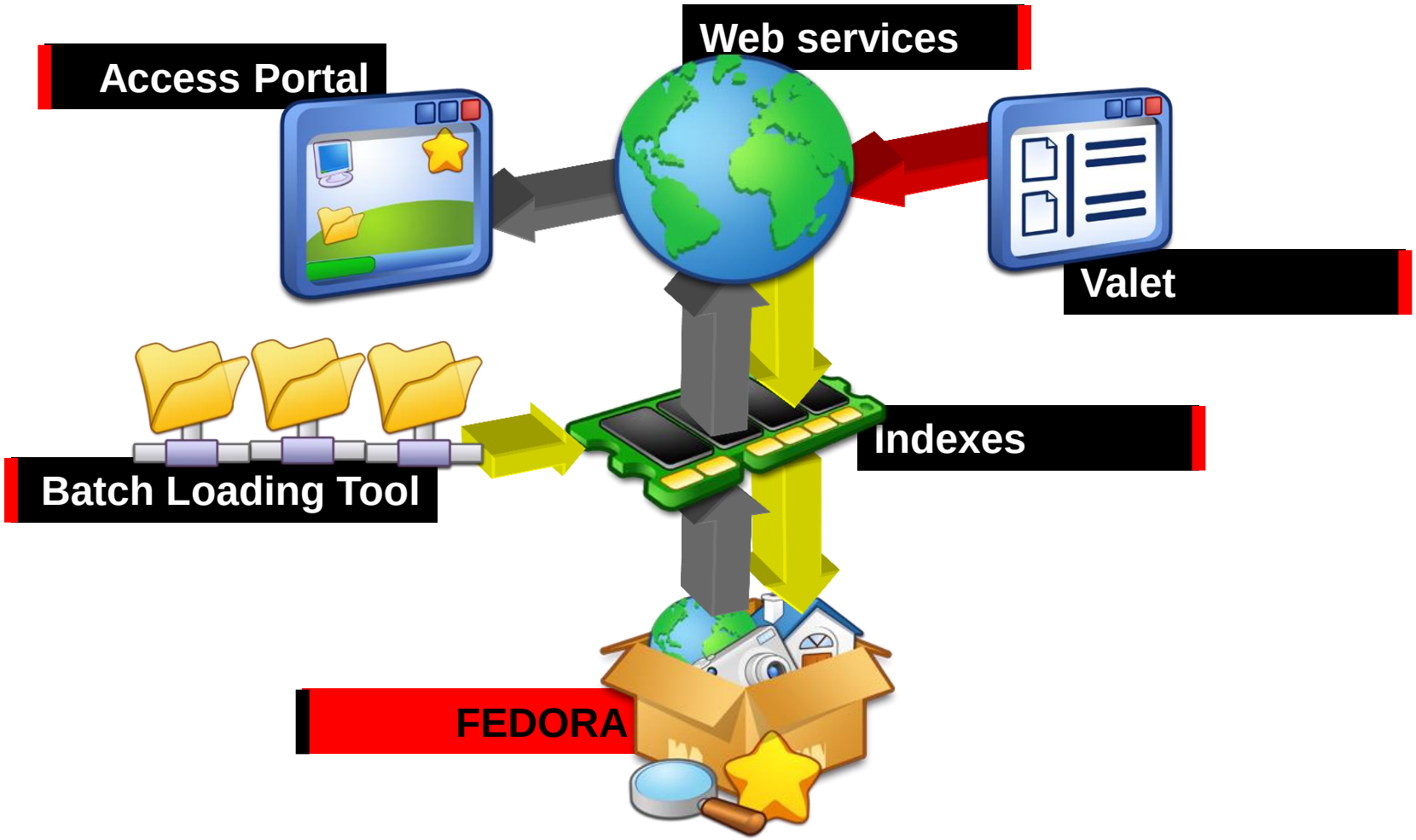
Reasons to base VITAL on Fedora™

- The “Flexible” and “Extensible” aspects:
  - **Flexible** – The development can be in any programming language because the API is based on Web services
  - **Extensible** – We can write workflow tools that facilitate different types of workflows
  - XML Submission and Storage – Digital objects are stored as XML-encoded files that conform to an extension of the METS schema
  - Focuses on Object Repository model and not how the repository will be used
  - Native OAI-PMH support
  - Open Source – VTLS is active in the open source movement

Reasons to base VITAL on Fedora™

- Extensible digital object model (architecture) allowing for complex objects
- Repository exposed by Web services APIs
  - Management (Creation, Deletion, Maintenance, Validation)
  - Access (Search, Disseminations)
- Scalable, persistent storage for content & metadata
- Content can be local and/or remote
- Content versioning
- There is a community developing other add-ons to Fedora that will add value and services for adopters
- Low cost procurement for customers

VITAL Fedora Relationship

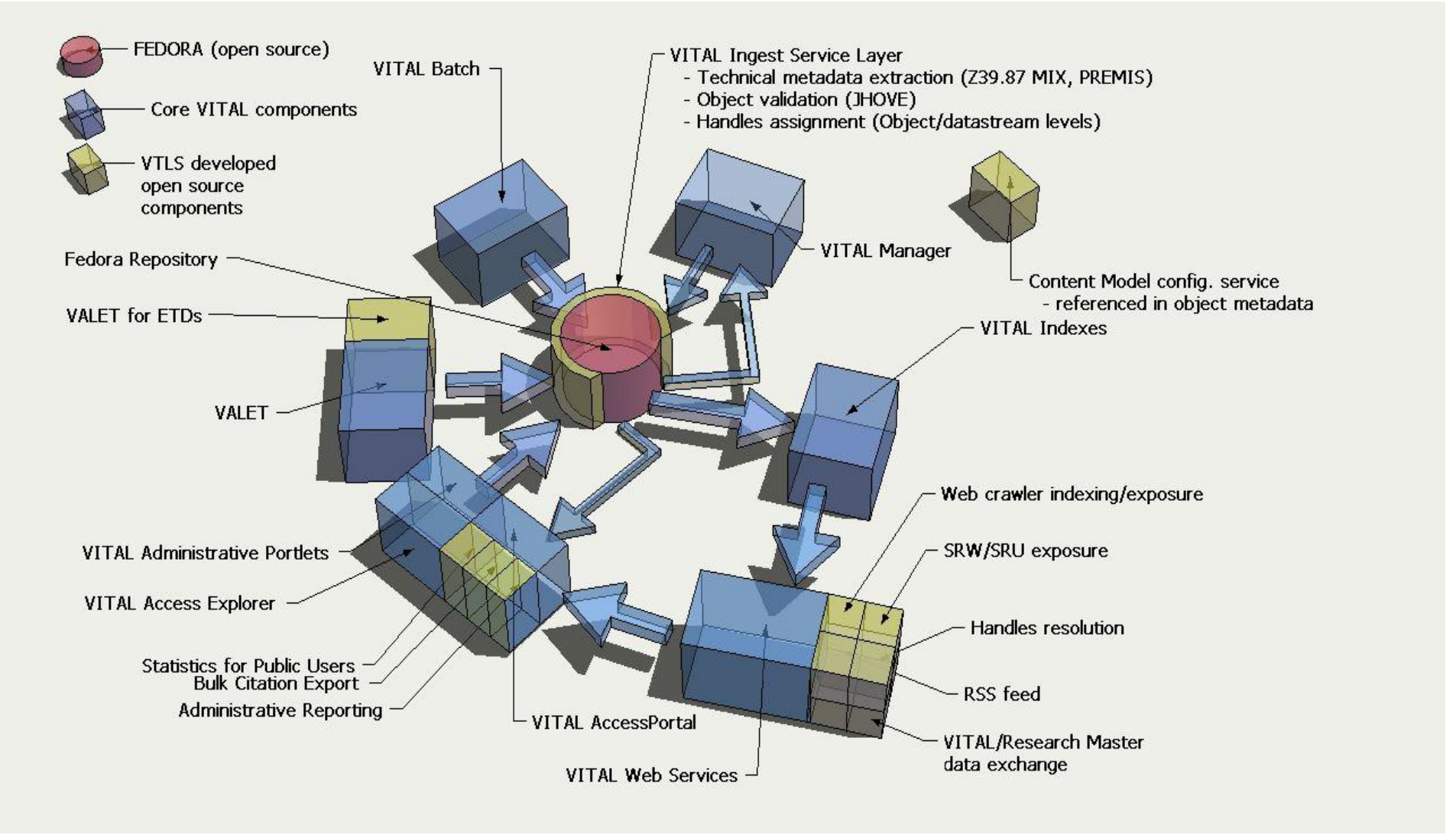


How does VITAL work with Fedora™?

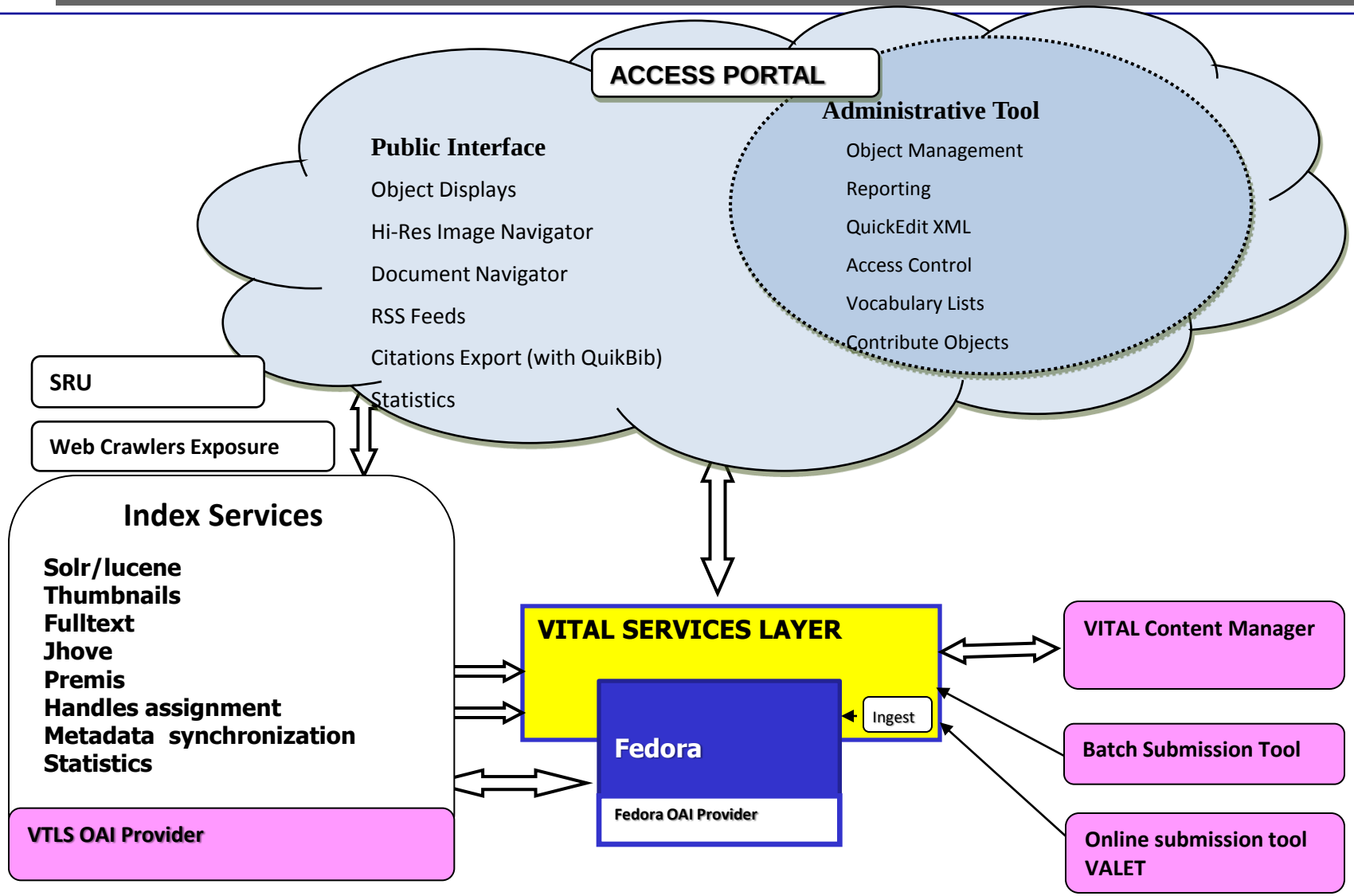
- Workflow tools that simplify the creation, modification, submission and dissemination of digital objects
- Web Service Interfaces (API's)
- Management Service (API-M)
  - Ingest – XML-encoded object submission
  - Create – interactive object creation via API request
  - Maintain – interactive object modification via API requests
  - Validate – application of integrity rules to objects
  - Identify – generate unique object identifiers
  - Security – authentication and access control
  - Preserve – automatic content versioning and audit trail
  - Export – XML-encoded object formats

VITAL / Fedora Open Source

(through VITAL Version 4.0)



The VTLS VITAL Architecture



Key Features/Benefits to VITAL (1)

Some key benefits of VITAL include:

- Storage and management of any content format, including rich-media, due to VITAL's repository object architecture
- Integration with existing systems through open, standards-based protocols
- Search full-text content of PDF, DOC, RTF and other document formats
- Display high resolution imagery, multi-page documents and specialized data formats (MARC, EAD, TEI, etc.)
- Automatically capture preservation metadata and create long-term, citable DOIs
- Storage of content in native XML provides support/validation against any metadata schema
- Support for aggregation of like-content in collections
- Automatic metadata extraction and validation (facilitates preservation)

Key Features/Benefits to VITAL (2)

- Web-crawler indexing and exposure (Google, etc.)
- Language support and UNICODE compliance
- Support for custom index creation for resource discovery
- SRW/SRU Interface for exposure of repository content
- Tracking content changes through versioning
- Robust facilities for batch ingest of content
- Support for link resolving against an institution's OpenURL server
- Support for dynamic creation/export of content citations (EndNote, text)
- Facilities for web-based self-submission of diverse content types (VALET)
- Integration with Fedora™ repository architecture allows for extensibility

Key Features/Benefits to VITAL (3)

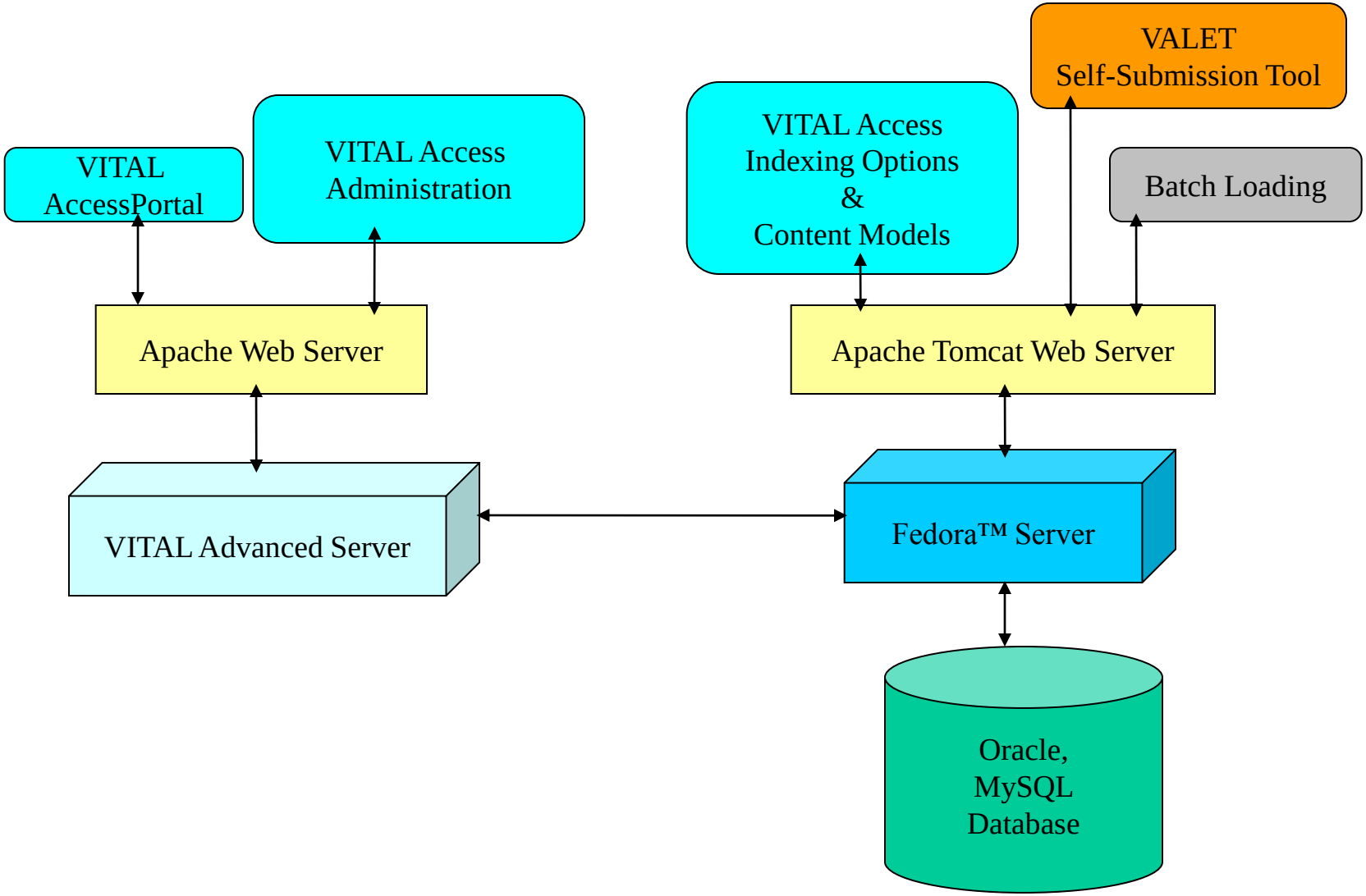
- SRU/SRW – SRW gateway connection for search from ILS or Federated Search
- Sophisticated display for Encoded Archival Description (EAD), Text Encoding Initiative (TEI), MODS, MIX, Dublin Core and MARCXML records
- Ability to edit XML content in the VITAL Access Portal
- GUI Relationship Editor between Objects
- Authority Management and Control using MADS and MARC21 Authority format
- Data exposure using faceted based search capabilities
- Consulting, Training, Documentation and Support

Key Features/Benefits to VITAL (4)

Ingesting using VALET

- Web-based form for self service or assisted submission
- Can be customised for your look and feel
- Currently 7 basic defined content models and can be extended to include more models
- Handle submission of any file format and allows contributors and other content creators to enter metadata
- User specified review processes
  - Number of review stages is fully configurable
  - Can specify different reviewers for different stages e.g. metadata, copyright

VITAL Components



Packages Available

The VITAL Package

- Fedora™ open source software (free)
- VTLS VITAL software
  - VITAL Access Portal
  - VALET Submission Tool
  - Batch Loading Utility
  - Content Special Viewers
- VTLS installation, training, support, integration and documentation



Packages Available

VITAL Hosted Solution

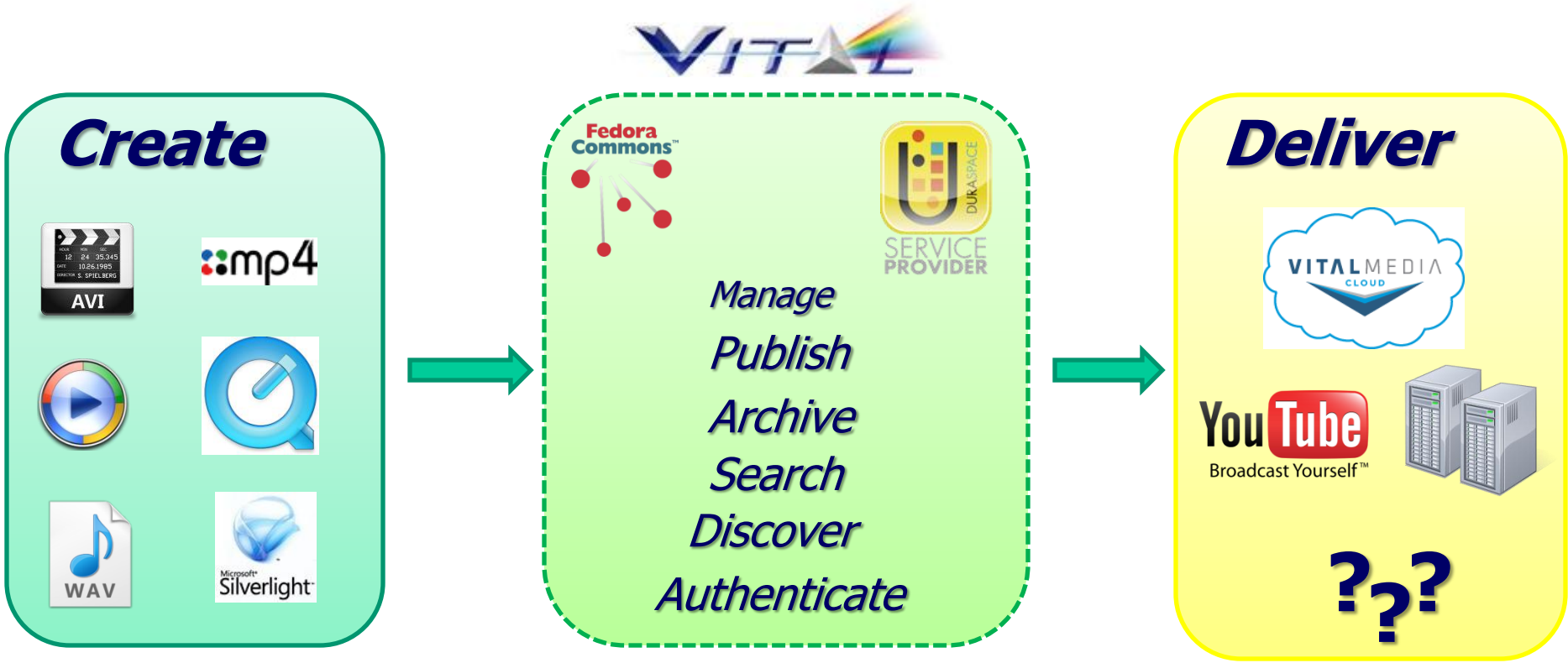
- VTLS provides ASP services for your repository services.

VTLS Professional Digital Imaging Services

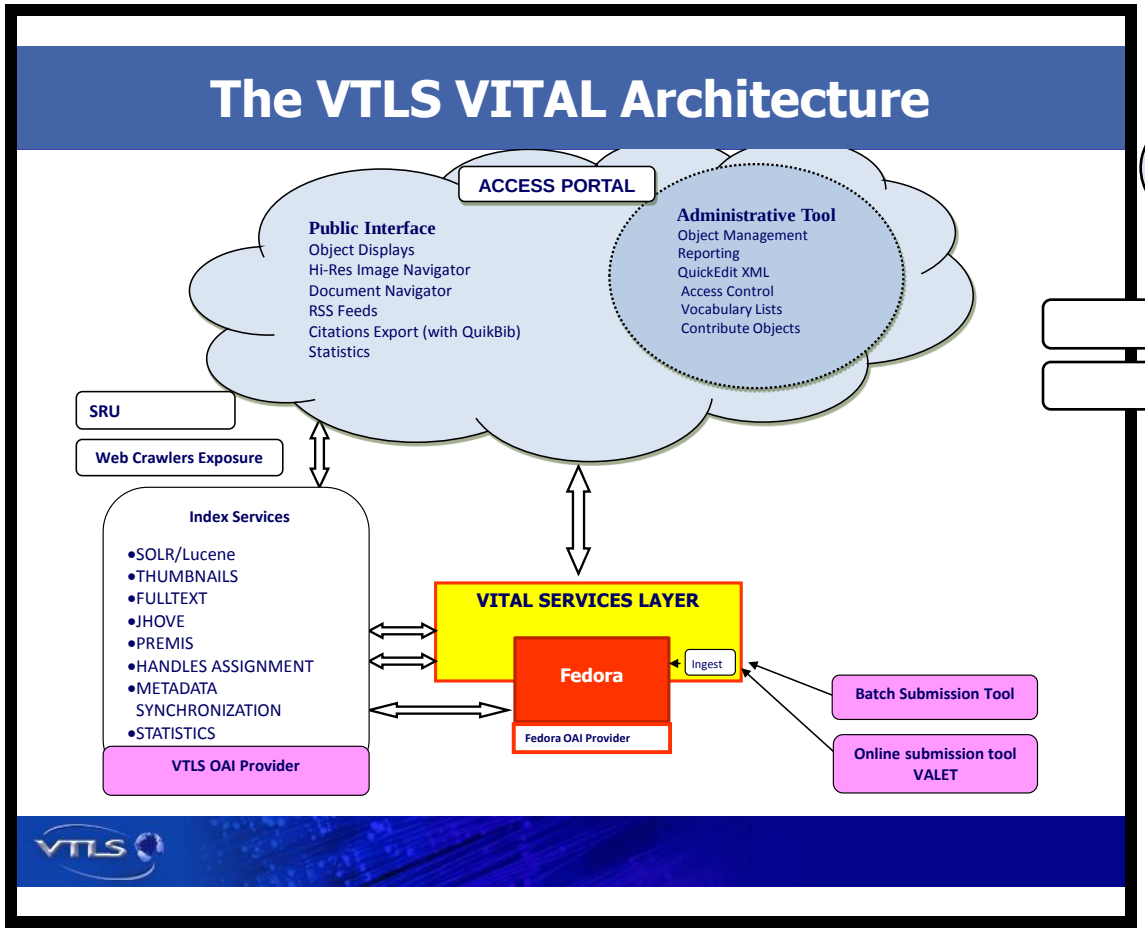
- Imaging services and project consulting can be combined with any of the above packages to provide a solution tailored to your needs

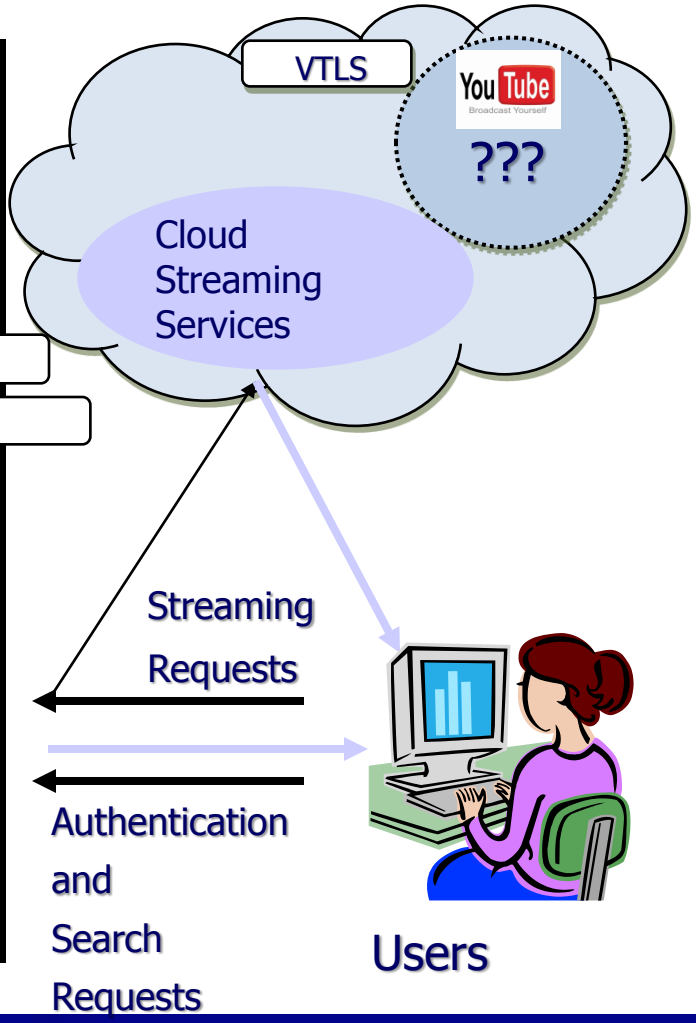


VITAL Media - Architecture



Streaming from VITAL Media Cloud





Example Lecture

VCOMTv
Edward Via College of Osteopathic Medicine Online Network

SIGN OUT
year3Student@OMS-III

Physicians to Care for those Most in Need

VCOM Students

Change Site View
GlobalView
Submit

Repository
Show All 88
Show Quick Collection 0

Search Advanced Search
Search
Show Public Objects

Browse
Communities & Collections
By Title
By Creator

Home ▷ Pediatrics ▷ Videos ▷ Newborn Exam

Breann Bailey, MD, FAAP. Newborn Exam. 12/07/2009.

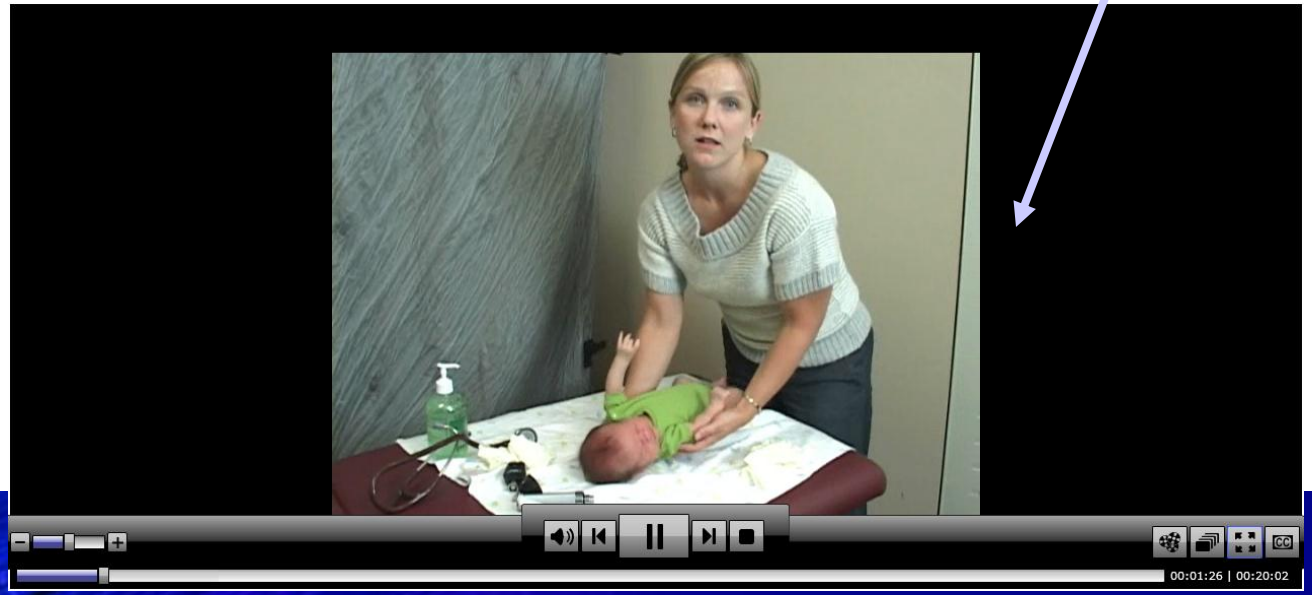
Description	Size	Format	
Silverlight Video	3 KB	HTML Document	View Video

Title	**Newborn Exam**
Creator	Breann Bailey, MD, FAAP
Date	12/07/2009
Collection	Videos
Identifier	vcomtv:1176

Streaming Requests

Cloud Streaming Services





VTLS

Who is using VITAL?

- Athens Archaeological Society
- University Of Geneva
- National Library of Ireland
- Slovak National Library
- Tomsk State University
- Columbia University
- Virginia Tech
- Duke University
- Pakistan Medical Research Council
- Pakistan Ministry of Health
- Kuwait Institute of Scientific Research (KISR)
- La Trobe University
- Macquarie University
- Monash University
- Murdoch University
- Queensland University of Technology
- Swinburne University of Technology
- University of Ballarat
- University of Newcastle
- University of Western Sydney

VTLS HQ in Blacksburg, VA, USA



More About VTLS Inc.



- **VTLS World Headquarters are located in Blacksburg**
- **Barcelona, Spain (EMEA)**
- **New Delhi, India**
- **Rio de Janeiro, Brazil**
- **Kuala Lumpur, Malaysia**
- **Melbourne, Australia**

VTLS has four major products

Virtua: Integrated Library System

VITAL: Institutional Repository Software

VISUALIZER: Discovery tool for distributed content (Now Chivas)

VTRAX: RFID based tracking & security systems for libraries



Nabil Saadallah

Manager, Middle East and Africa
VTLS Europe

saadallahn@vtlseurope.com