Digital Libraries

State-of-the-art and Prospects

John Impagliazzo

Hofstra University
New York

Digital Libraries

- Some familiarity
  - IEEE/CS Digital Library
  - ACM Digital Library
  - Others
- Content
  - Refereed and reviewed online materials
  - Interactive
  - Search facilities
  - Download materials (fee structure may apply)

What's a NSDL?

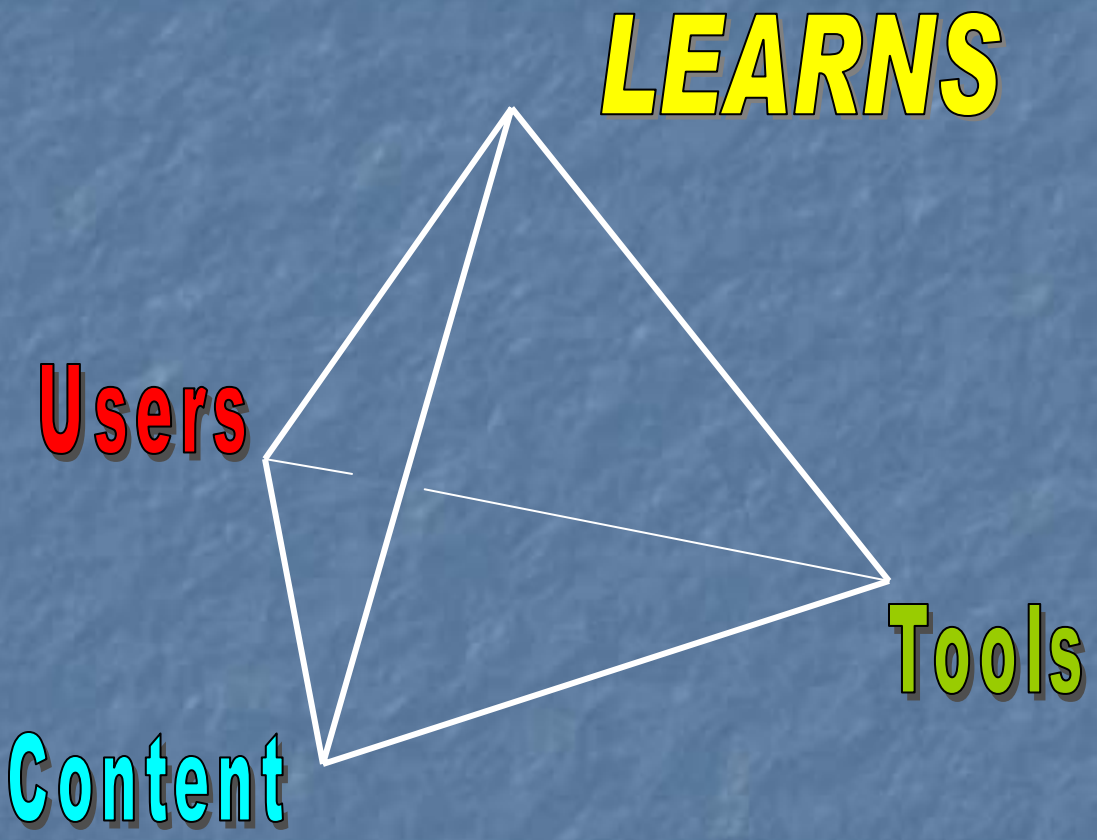
National Science Foundation **S**TEM (Science, Technology, Engineering, Mathematics) Educational **D**igital **L**ibrary

Over $100 million initiative

Transitions

- NSDL will operate as a distributed federation
- Contains separate parts for each key discipline
- Expected to be a global effort in the future
- NSDL → *LEARNS*

Learning Environment and Resources Network for STEM Education (*LEARNS*)



"The network is the library."

LEARNS Connects

- Users

  Students, educators, life-long learners

- Content

  Structured learning materials
  Large real-time or archived datasets
  Audio, images, animations
  Primary sources
  Digital learning objects (e.g., applets)
  Interactive (virtual, remote) laboratories ...

- Tools

  Search; refer; validate; integrate; create; customize; publish; share; notify; collaborate; ...

LEARNS Supports

Learning communities

Users

(Profiles)

Application services

Tools

(Protocols)

Customizable collections

Content

(Metadata)

LEARNS Enables

Environments for

- Communication
- Collaboration
- Creation
- Validation
- Evaluation
- Recognition
- ...

and

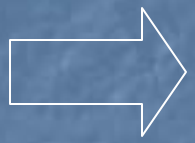
- Discovery
- Stability
- Reliability
- Reusability
- Interoperability
- Customizability
- ...

of Resources

NSDL Overall Goal

Tracks

Core Integration
Collections
Services
Targeted Research



LEARNS

(Operational in 2003)

NSDL Program Tracks

Core Integration

Coordinate a distributed alliance of resource collection and service providers; and ensure reliable and extensible access to and usability of the resulting network of learning environments and resources

Collections

Aggregate and actively manage a subset of the digital library's content within a coherent theme / specialty

Services

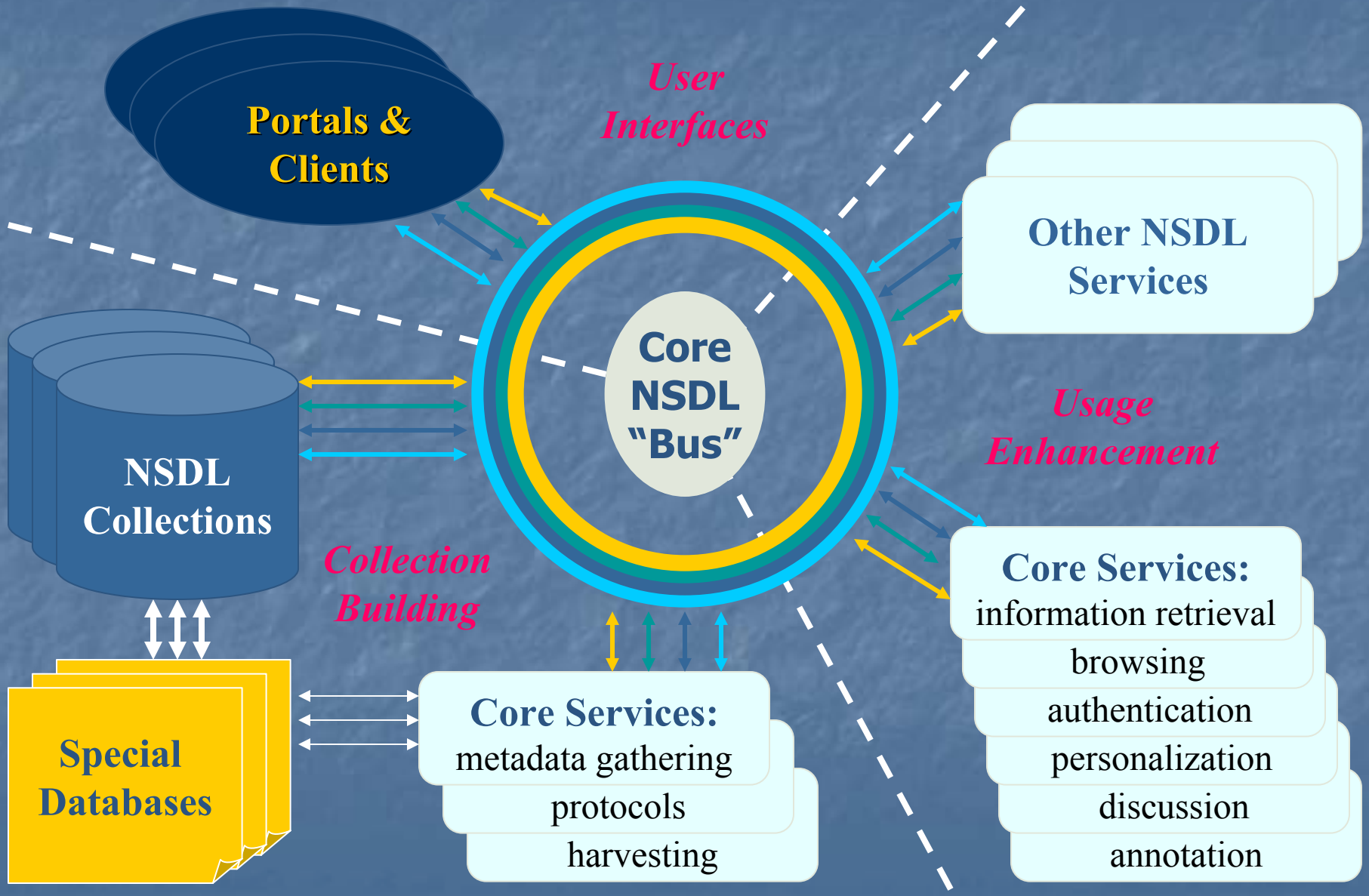
Increase the impact, reach, efficiency, and value of the digital library in its fully operational form

Targeted (Applied) Research

Have immediate impact on one or more of the other three tracks

NSDL Information Architecture

(Developed by the Technical Infrastructure Workgroup)



What is CITIDEL?

- **C**omputing and **I**nformation **T**echnology **I**nteractive **D**igital **E**ducational **L**ibrary
- Consortium
  - Virginia Tech (lead institution)
  - Hofstra University
  - Penn State University
  - The College of New Jersey
  - Villanova University
- Collection Component

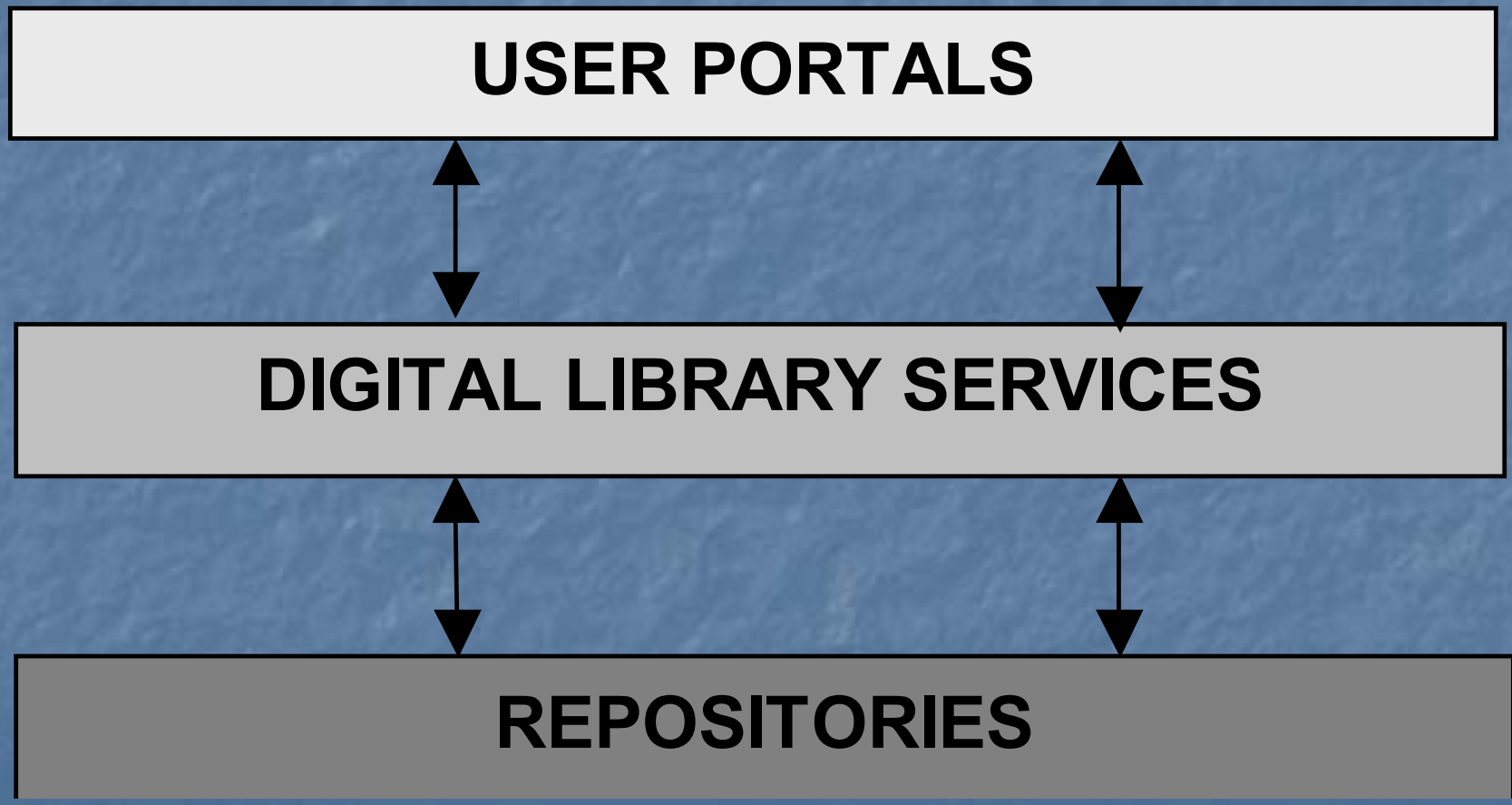
CITIDEL Content

- Builds on existing resources
  - ACM and IEEE/CS Digital Libraries
  - ResearchIndex
  - NCSTRL (Networked Computer Science Technical Reference Library)
  - JERIC and CSTC
  - Others …
- Inclusion
  - Refereed and reviewed online materials
  - Interactive
  - Search facilities
  - Download materials (possible fee structure)

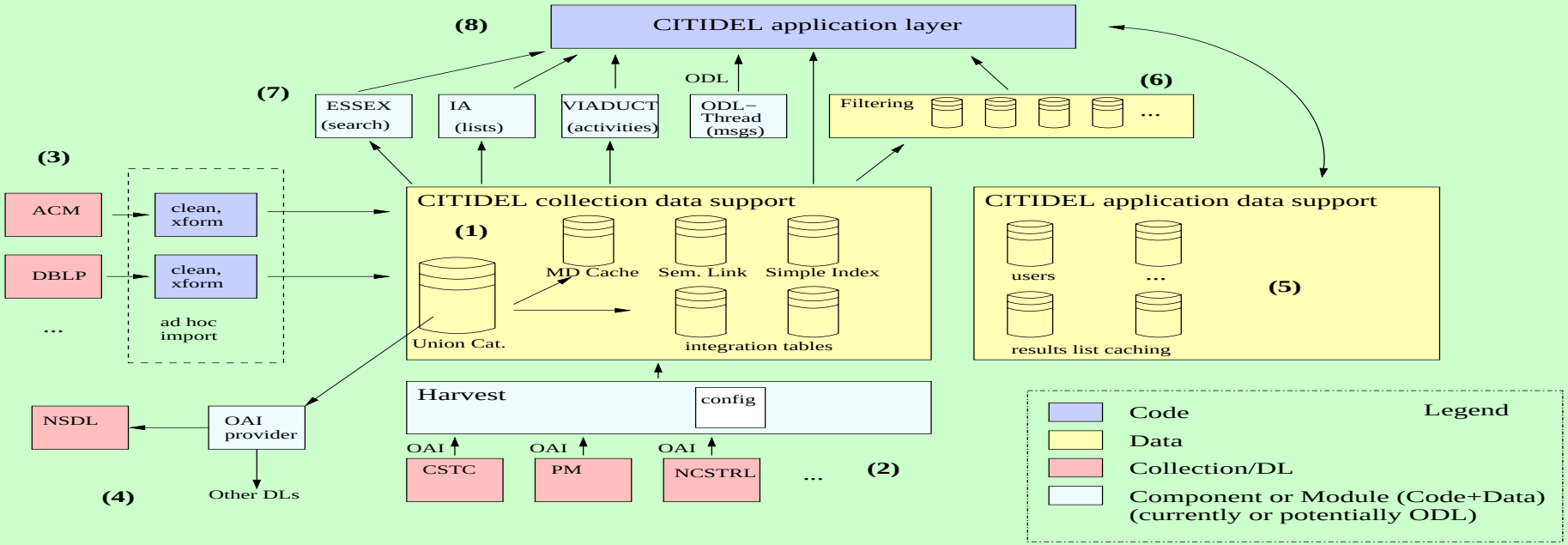
CITIDEL Players

- Led by Virginia Tech
  - Fox (Director, DL systems)
  - Lee (History)
  - Perez (User interface, Spanish support)
- Partners
  - Hofstra (Impagliazzo – Outreach, History)
  - College of New Jersey (Knox - CSTC)
  - Villanova (Cassel – Evaluation, JERIC)
  - Penn State (Giles – Selective crawling and searching)

Overview CITIDEL Architecture



CITIDEL: Current Architecture



1. The core of CITIDEL is the collection data support. This consists centrally of a union catalog, metadata cache, semantic links table, integration tables, and more.
2. The harvesting system populates the union catalog and the secondary tables from the contents of remote digital library collections, over Open Archives.
3. For collections which lack an Open Archive provider, ad hoc importing facilities must be constructed.
4. CITIDEL serves up the contents of its union catalog via an Open Archives data provider, giving other digital libraries (NSDL) access to CITIDEL's metadata.
5. The application layer data support consists of non-content-related tables and personalization tables, such as a table of users and preferences.
6. The filtering system relies on extensive database support for speed.
7. The service modules tackle the DL features of search engine, recombination into annotated and enriched lists, creation of pedagogical activities utilizing DL resources, and posting messages to DL resources.
8. The CITIDEL application ties it all together in a single user interface. Most presentation (but not all) is handled here.

What it Really Means

- Easy access to resources with known credentials
- Tools for using extra resources in preparing learning activities
- Access to others with shared interests and needs

CITIDEL-related Collections and Sizes

Summary as of 2002 October

Computing History (VT Virtual History Museum)	NCSTRL (Nat. C.S. Tech. Ref. Library)	CSTC (CS Teaching Center)	VT CS ETD (Electronic Theses and Dissertations)	Planet Math (Open Math Encyc.)	Research Index (Generated By net Crawlers)	Other (DLs and ad hoc items)
663 Items	24188 Items	76 Items	13 Items	176 Items	0 Items	16 Items
Many more expected	More expected	More expected	More expected	More expected	Over 500K expected	Over 800K expected

Approximately 25,000 items

Status as of 2004 August 11

- 877,192 collection items
- 15 collection sources
- 390,057 papers
- Over 300,000 biographical records
  - Digital Biographical Library (DBLP)
  - Contains over 500,000 entries

Technology Features

- Component architecture (Open Digital Library)
  - Re-use and compose re-deployable digital library components.
- Built Using Open Standards & Technologies
  - XSL and XML: Interface Rendering
  - Perl: Component Integration
  - ESSEX: Search Engine Functionality
- Open Archives Initiative
  - Used to collect DL Resources and DL Interoperability

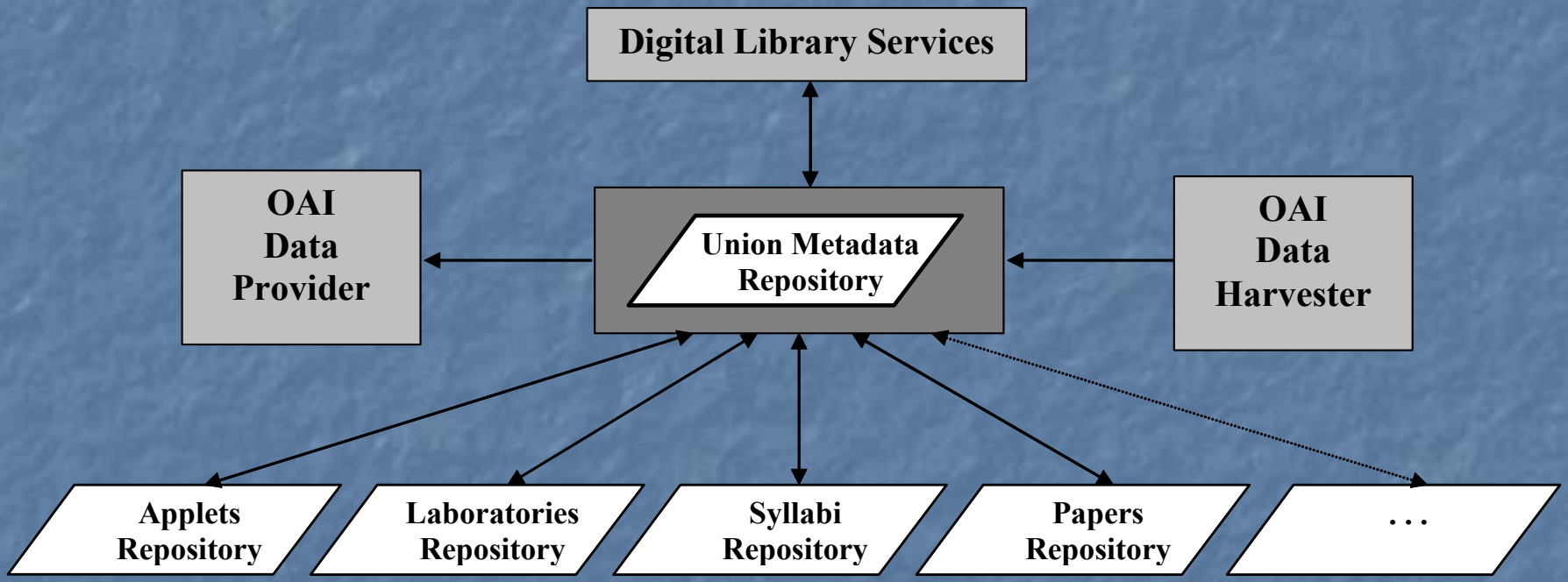
User Features

- Very large collection
- Filtered browsing and searching
  - Filters based on these user-selected sub-communities.
  - Allows customization in addition to views of all results.
- Multi-classification browser
  - Supports browsing based on curricula (familiar, professional society approved) in computing and related disciplines, as well as on classification schemes.
- Activity collection creation & tools
  - Faculty and students can extract resource references from CITIDEL search collections into learning activity templates, for sharing and interchange (with versioning).
  - VIADUCT assists in the development of a totally independent, self-generated, educational resource collection within CITIDEL. IAVT is based on Utah State's Instructional Architect.

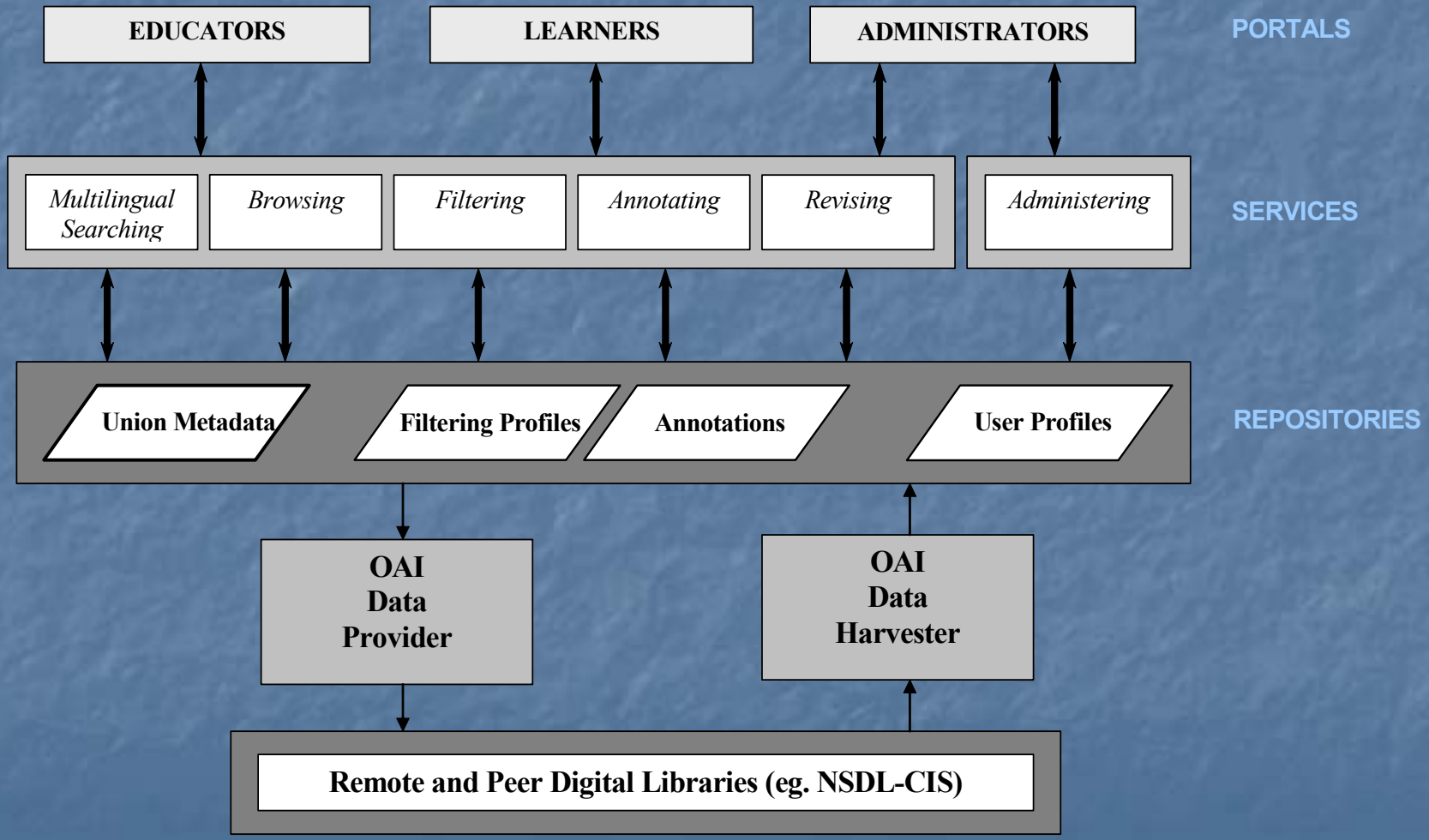
Multi-classification Browser

- Multi-classification browser allows users to browse through the CITIDEL collections based on professional society approved curricula in computing as well as classification schemes.
- As users span many disciplines related to computing, the users may browse within the scheme with which they are most familiar.
- Resources are cross-classified wherever possible through these schemes. The current schemes include the 2001 ACM/IEEE-CS Computing Curricula, the 1998 ACM Computing Classification System, the Computing Research Repository Subject Areas, and the 2000 AMS Mathematics Subject Classification.

Distributed Repository Structure



Digital Library Architecture for Local and Interoperable CITIDEL Services



VIADUCT

Virginia

Instructional

Architect for

Digital

Undergraduate

Computing

Teaching

To be able to

Take resources from a search and create lesson plans

Modify lesson plans for local use

Computer Science Teaching Center (CSTC)

- Strategy
  - Don't build large, expensive multimedia packages that become obsolete and are difficult to re-use
  - Concentrate on **small knowledge units**.
- Learners benefit from having well-crafted modules that have been **reviewed and tested**.
- Use digital libraries to build
  - A **powerful base** of support for learners
  - Include a variety of courses, self-study tutorials, & reference resources

JERIC

- **J**ournal of **E**ducational **R**esources **i**n **C**omputing
  - Accessible from <www.acm.org/>
  - ACM and SIGCSE support
- Refereed and interactive
- Part of ACM Digital Library

Contribute to CITIDEL

- How do you submit *Your* resources?
  - Create account
  - Submission form to provide resource information
  - Upload files and URLs

Want to Know More?

Check

< http://www.citidel.org/ >

< http://www.nsdl.org/ >

Спасибо

Вопросы?