Multinational Federated ADL Search and Retrieval – II: Sharing Across NATO

James Xu
Advanced Distributed Learning Initiative
Alexandria, Virginia
james.xu.ctr@adlnet.gov

Geir Isaksen
Norwegian Defence ADL Centre National
Oslo, Norway
gisaksen@fee.mil.no

Bill Railer
Canadian Defence Academy
Kingston, Canada
railer.wg@forces.gc.ca

Giridhar Manepalli
Corporation for National Research Initiatives
Reston, Virginia
gmanepalli@cnri.reston.va.us

ABSTRACT

The Advanced Distributed Learning (ADL) Initiative of the United States Department of Defense (DoD), the Norwegian Defence Education Command (NoDEC) and the Canadian Defence Academy (CDA) are engaged in a joint effort to develop an infrastructure to allow multiple NATO nations to search, discover, view, and share technical and instructional content. This capability provides near real-time delivery of training and performance support for common interest information that is tailored to the specific requirements of a given instance.

This paper reports our progress on a project initially described at I/ITSEC 2008. This project revolutionizes access to multiple heterogeneous repositories containing learning objects distributed across different countries. By incorporating an instance of the Content Object Repository Discovery and Registration/Resolution Architecture (CORDRA), the ADL-Registry, learning objects across all registered repositories are uniquely identified so they can be discovered globally without the limitation of requiring a specific content repository system. This also leaves the total control of learning objects to local repositories. Upon their discovery, learning objects, regardless of their physical locations and countries of origin, can be easily viewed, retrieved, and constructed into training materials according to learning and training objectives and uploaded into SCORM and non-SCORM-compliant learning management systems (LMSs) from any participating nation. The main benefits of the project are to share learning assets and knowledge, provide instant access to learning objects, reduce or eliminate the cost of developing redundant content, and increase collaboration and cooperation among participating nations. The paper also recommends basic requirements for additional participation in the infrastructure.

ABOUT THE AUTHORS

James Xu supports the Advanced Distributed Learning (ADL) Initiative by bringing emerging technologies into the learning, education, and training space. Before joining ADL, he spent over 15 years in the software industry serving as Chief System Architect, VP of Software Development and CTO. Mr. Xu designed and implemented one of the largest data warehouse. During the past several years, he has concentrated on developing enterprise systems using service-oriented architecture (SOA) and other Web 2.0 technologies. Mr. Xu is currently a professor at DeVry University teaching Game and Simulation Programming. He holds a Master of Science degree in Mechanical Engineering, and is pursuing a Ph.D. in Modeling & Simulation.

LtCdr Geir Isaksen is a former submariner working at the Norwegian Defence Education Command (NoDEC) in Oslo, Norway. He is currently an ADL advisor at the Norwegian Defence (NoD) ADL Centre.

LtCdr Isaksen has worked at the Norwegian Navy Submarine School as a head instructor in the ULA-class submarine simulator in Bergen for 2 years. In recent years, he completed further education in adult pedagogy, crew resource management, project management, and learning styles. He has been responsible for the development of the NoD ADL regulations and several major e-learning projects like the development of the courses in law of armed conflict and several R&D projects. LtCdr Isaksen is the current chairman of the NATO Training Group Working Group for IT/ED ADL subgroup and a part of the newly established Norwegian ADL Partnership Lab in Oslo.

Bill Railer brings a mixed-bag of over 7 years of experience in solutions development for corporate finance and equity trading as well as several years of teaching information technology at the university level. In his current role as Director of Learning Technology for the Canadian Forces and the Canadian ADL Partnership Lab, Mr. Railer promotes the development and adoption of new learning concepts, methodologies, and solutions. He is also responsible for the implementation of global e-learning standards within the Department of National Defence (DND).

Giridhar Manepalli is a Researcher and Senior Software Engineer at the Corporation for National Research Initiatives (CNRI). He is the technical lead for the Advanced Distributed Learning (ADL) Registry and leads the technical effort to design a more generic metadata registry, known as the Digital Object Registry. He is also building information management services required for the networking communities that are participating in the Global Environment for Network Innovations effort, using the Digital Object Registry. Prior to joining CNRI, Mr. Manepalli worked at the Old Dominion University Research Foundation, focusing on the Archive Ingest and Handling Test, a Digital Preservation project sponsored by the Library of Congress, to extract and translate information out of terabyte-sized archives. He also worked as a Research Consultant at the Virginia Modeling Analysis and Simulation Center, where he studied the impact of mass casualty disasters on health care systems using agent modeling and simulation techniques. He has a Masters degree in Computer Science from Old Dominion University, and Bachelors degrees in Computer Science and in Computer Engineering.

Multinational Federated ADL Search and Retrieval – II: Sharing Across NATO

James Xu
Advanced Distributed Learning Initiative
Alexandria, Virginia
james.xu.ctr@adlnet.gov

Geir Isaksen
Norwegian Defence ADL Centre National
Oslo, Norway
gisaksen@fee.mil.no

Bill Railer
Canadian Defence Academy
Kingston, Canada
railer.wg@forces.gc.ca

Giridhar Manepalli
Corporation for National Research Initiatives
Reston, Virginia
gmanepalli@cnri.reston.va.us

INTRODUCTION

Aided by the introduction of the Sharable Content Object Reference Model (SCORM) by the U.S. Department of Defense's Advanced Distributed Learning Initiative (ADL) some years ago (Advanced Distributed Learning Initiative, 2009c), the concept of developing training courses and content has been evolving. Instead of developing training courses from the ground-up every time, the focus has shifted to developing interoperable and reusable sharable content objects (SCOs) and assets that can be reused in many different training courses on any SCORM-compliant LMS (see Figure 1).

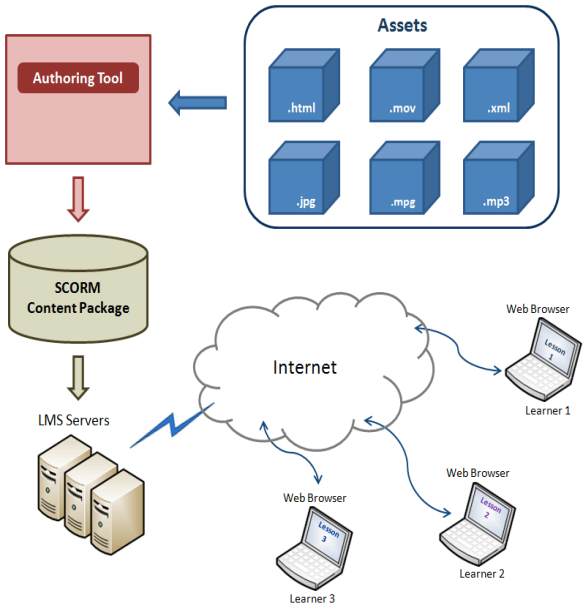

Figure 1. Course Implementation in a LMS

Some clear benefits of developing interoperable and reusable content are reduced development costs, shortened course development time, and easier deployment.

Increased cooperation and collaboration among NATO countries in general also increases the benefits of sharing course training materials and development resources across NATO countries. Imagine the impact on NATO response and reaction time when Forces from any participating NATO country could, through a common interface, search and share content developed and managed by other NATO countries without expending precious resources and time on the design and development of training materials. Forces would be better prepared and able to respond or react more quickly than they currently can.

To realize these benefits, we had to address a number of challenges and questions, including content development, storage, discovery, distribution, and construction. This approach requires an infrastructure that is flexible and efficient enough to connect all the content repositories from NATO countries to allow content sharing.

In the Multinational Federated ADL Search and Retrieval – II (M-FASR II) project, we are designing an architectural prototype that addressed some of those challenges and demonstrated the capability using a specific NATO training course. In this paper we summarize the results of the first phase of the M-FASR project, detail our work on the follow-on Phase II project, which extended the scope of M-FASR, and discuss future challenges.

BACKGROUND

Learning-centered organizations from three NATO nations, the Norwegian Defence Education Command (NoDEC), Canadian Defence Academy (CDA) and ADL, participated in the successful first phase of this project (Isaksen, Lamothe, & Railer, 2008).

During this phase of the project, all three participants used the same content management system. These three test instances of the content management system formed a "cluster" which served as the framework (see Figure 2) for the project and enabled the partnership nations to search and access each other's content.

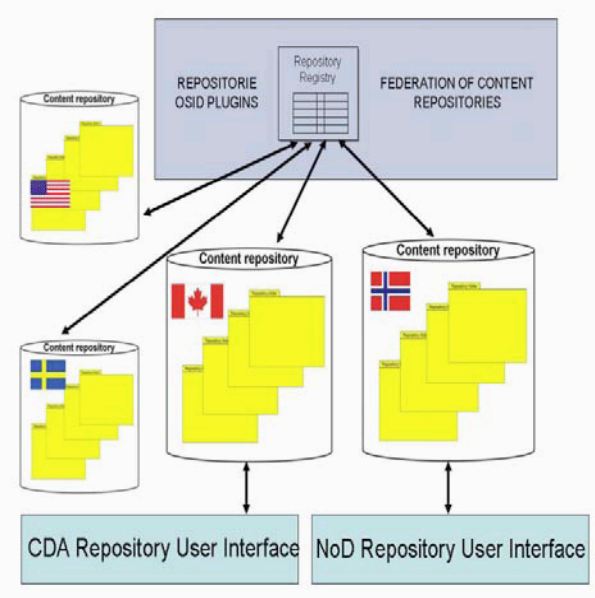

Figure 2. M-FASR Phase I Framework

Although Phase I demonstrated a technical solution to multi-national content sharing, it presented some serious challenges and limitations:

- Phase I did not address interoperability of metadata schemas and taxonomies. Different metadata structures implemented on different instances of the content management system would produce inconsistent search results.
- The Phase I framework required a single content management system. Any NATO nation that wanted to join the framework would have had to deploy the same content management system regardless of its existing repository system.
- Since the content searches across the content management system "cluster" were performed in a distributed fashion (see Figure 3), the performance would likely decrease as the size of the "cluster" increased.
- Without using a common unique identifier system across the entire framework, the same content could be duplicated on multiple repository systems and that duplication would appear in search results. This could impact performance and usability of the information.

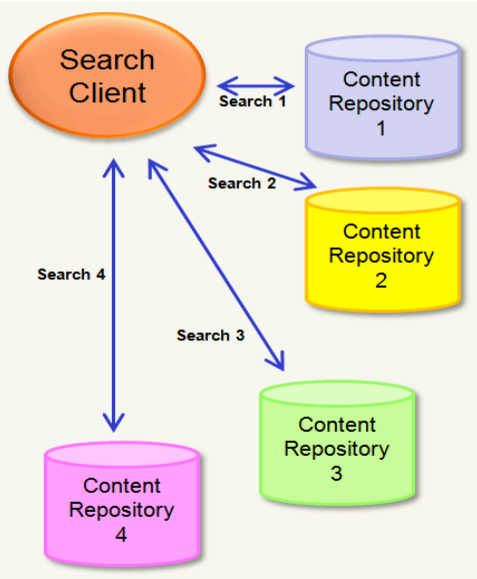

Figure 3. Aggregation Through Distributed Search

These problems are addressed by a new architecture introduced for Phase II of the project.

DESIGN

The basic design principle for Phase II of M-FASR is the federation of heterogeneous repositories through a system of registries. Done successfully, this maintains the advantages of the first phase. It provides a single starting point for search and retrieval while leaving content management in the hands of the content owners. At the same time this approach ameliorates the significant disadvantages of mandating a common repository and using only a non-scalable distributed search for aggregation. The Phase II architecture adheres to the principles of the Content Object Repository Discovery and Registration/Resolution Architecture (CORDRA) (Jesukiewicz, 2008), and is itself a step in the evolution of CORDRA.

CORDRA is an effort to enable the discovery, location, and reuse of learning content across a heterogeneous collection of repositories and content types, through the use of federation and registration (see Figure 4). The first instance of CORDRA is the ADL Registry (Jerez, Manepalli, Blanchi, & Lannom, 2006). The ADL

Registry federates content from multiple repositories through the registration of standardized metadata, based on the Learning Object Metadata (LOM) standard (IEEE Learning Technology Standards Committee, 2002), for learning content that is distributed across repositories. It leaves the control of that content with the individual repository managers/owners. Each object in the ADL Registry is assigned a unique persistent identifier, known as a handle (Corporation for National Research Initiatives, 2009). The handle allows duplicates to be easily discovered and allows static references to digital entities to persist beyond state changes such as when a digital item is moved to a new location.

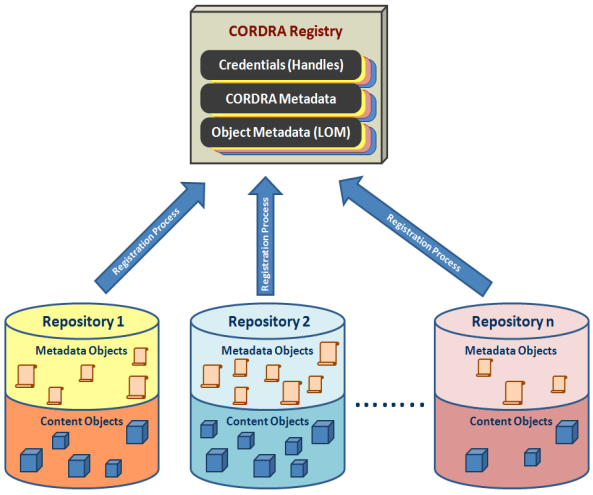

Figure 4. Federation Through Registration

M-FASR Phase II extends the ADL Registry to enable federation of learning assets beyond the confines of the U.S. DoD. The goal of Phase II is to demonstrate the building of a course from assets that are widely distributed and stored in repositories owned by multiple NATO counties. This will be done by building a higher level registry that can aggregate all of the metadata in the ADL Registry plus metadata from repositories that do not participate in the ADL Registry federation (see Figure 5).

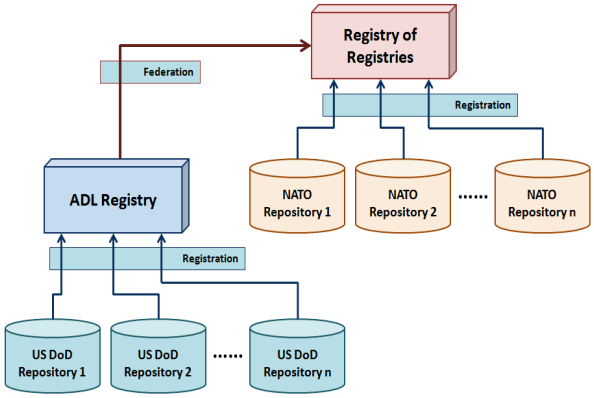

Figure 5. M-FASR II Architecture

One of the objectives of Phase II is to make the evolution and management of the participating repositories independent of the federation**.** Some benefits of federation at the repository level include:

- Being repository independent; our federation promotes interoperability across different types of repository systems.
- Adding additional repositories easily without changing the architecture.
- Searching the registry instead of individual repositories; resulting in a high-level of scalability.

Since searches are performed on the registry instead of repositories, the implementation and integration of user interfaces are simplified as well.

- Search requests can be submitted through a consistent Web interface.
- User-interfaces can be implemented and integrated easily from any Web site or application.
- Search results can be displayed in HTML format by default.
- Search results can also be returned in raw XML so they can be re-formatted and integrated in any Web site or application.

The benefits continue at the content level.

- Every content object is identified by a globally unique identifier for easy discovery, eliminating the potential for content duplication.
- The ability to attach multiple-annotations to a single content object provides support for domain and/or version specific notations; eliminating the need for saving multiple copies of the same content.

- Content granularity is defined by users/repositories; providing the flexibility to discover any type of content objects.

Federation of the assets from across various repositories and the ADL Registry results in registries and individual repositories that can be aggregated to "Registry of Registries" status (see Figure 5). This would allow:

- The community to continue to grow while providing a hierarchical structure, as needed.
- Initial searches of the "Registry of Registries" before searches are performed on individual registries; further improving performance and scalability.

Operational Decisions

To prototype and demonstrate the goals of this project while still keeping the project within a manageable complexity for the given time frame, we made the following assumptions and decisions:

- We selected the current ADL Registry (version 1.7) for the lower-level registry function (Advanced Distributed Learning Initiative, 2009a).
- We establied a "Registry of Registries" above the ADL Registry to prototype and demonstrate the federation of registries.
- We created two repository instances directly under the "Registry of Registries" to keep the structure simple while still demonstrating the desired functionality. (Note that Figure 5 shows the abstract architecture and not the details of the demonstration).
- We used LOM for all learning content objects on all participating repositories from all locations for this phase of the project.
- We decided that all searches are performed on the federated Registry of Registries. No distributed searches are issued to any participating repository.
- We determined that all content objects are registered in participating registries using a defined M-FASR Handle prefix. All participating content items are registered, and accessible without security and policy limitations.

Given the time constraints for Phase II, the registration process may be manual for this phase of the project. Eventually, the process will be automated.

Search and Retrieval

One search of the "Registry of Registries" will discover all matching content objects on all repositories across all participating NATO countries, including all content objects registered in either the lower-level registry or directly in the "Registry of Registries", and all entries in the lower-level registry aggregated up to the "Registry of Registries". The "Registry of Registries" uses a simple interface (see Figure 6) to facilitate the quick and efficient yet complex searches it performs. Matching content objects are displayed on the search result page with URLs pointing to the current locations of the content objects (see Figure 7).

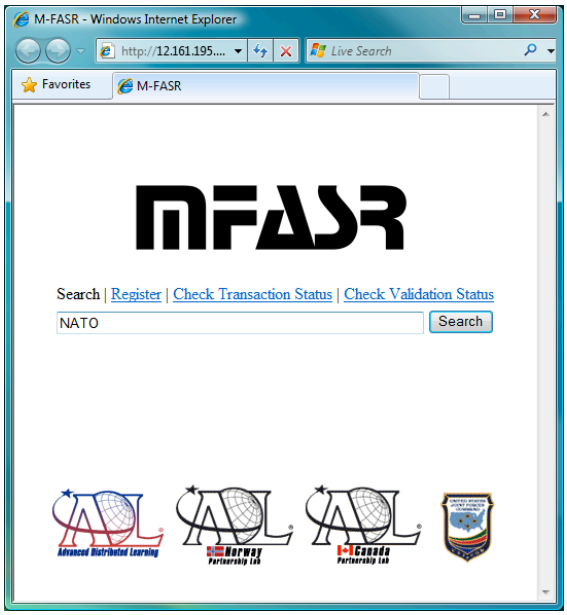

Figure 6. M-FASR II Search User Interface

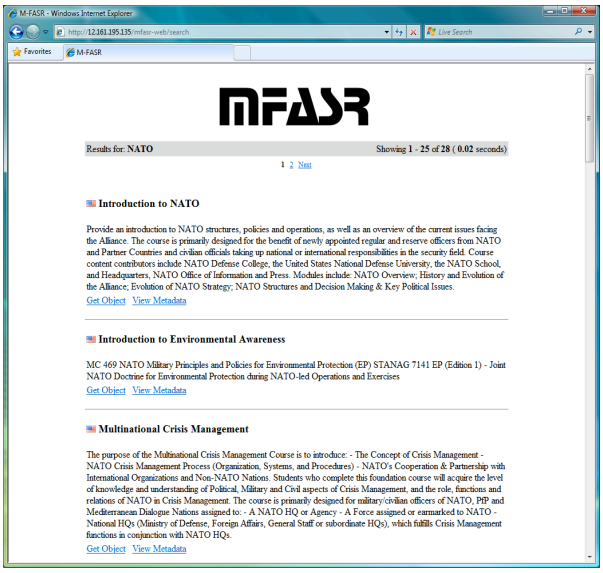

Figure 7. M-FASR II Search Results

After discovering content objects, instructional designers must evaluate the objects and determine which objects meet their requirements. Any content object that meets the requirements can be retrieved and reused in the new training course (see Figure 8). This capability should save instructional designers and content developers valuable time since they do not have to re-create content already developed by another NATO country.

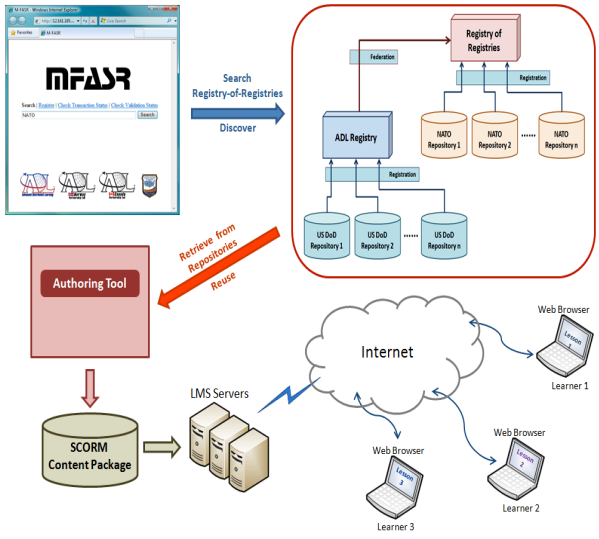

Figure 8. Search and Retrieval of Content Objects

REPOSITORY REQUIREMENTS

The project was designed to impose the fewest possible requirements on participating repositories. They must be able to register their content objects in registries and be able to distribute content objects within a designated domain. For Phase II, the designated domain is "public."

When selecting a repository or a content management system for future participation in M-FASR, a NATO country should consider:

- Metadata. Supported metadata elements must meet the LOM requirements, as specified in ADL-Registry version 1.7 documentation (Advanced Distributed Learning Initiative, 2009b).
- Versioning. A repository should be capable of tracking changes to content objects.
- Complex content object types. Users and repository managers must determine the correct type and size of objects for the design requirement, since registries do not dictate granularity or object type for registered objects.
- Web interface. A repository system should have a Web interface for content access and distribution.
- User roles and privileges. User roles and privileges are key to managing content objects within a repository because control remains at the repository level.
- Automated registration. Consistent metadata registration is more convenient when automated by a trigger following the deposit of content in a repository.

TEST CASE

Many repository systems and content management systems, including commercial off-the-shelf (COTS), government off-the-shelf (GOTS), and open-source, provide, or can be configured to provide, these capabilities. To demonstrate this, ADL deployed a COTS repository system, a GOTS repository system, an open-source repository system, and a home-grown, generic file system-based repository system in its Learning Technology Lab.

We used the "Public Information Operations (PIO)" course designed for "NATO Working Group Training and Education Peace Support Operations (WG TEPSO)" to demonstrate content sharing and reuse across multiple NATO countries for this test case (see Figure 9).

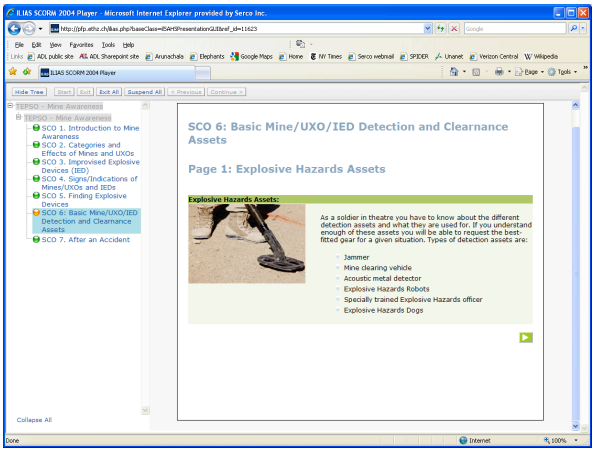

Figure 9. NATO's "Public Information Operations (PIO)" course

The SCOs and assets used to construct this course are located across participating NATO countries, i.e., United States, Norway, and Canada, in various types of repositories. After content objects are discovered by searching the registry, they are retrieved and reused to construct a new NATO training course using an authoring tool.

This phase of the project demonstrated many benefits and also revealed some challenging considerations for future action.

BENEFITS AND FUTURE CONSIDERATIONS

The project described in this paper implements architecture for the federation of repositories and registries and demonstrates the advantages of content sharing and reuse in a heterogeneous environment. Its further deployment would enable NATO countries to search, discover, access, and retrieve each other's content while still maintaining control over the management of those content objects nationally. Specifically, the benefits of this architecture are:

- Every content object is tagged with a globally unique identifier, eliminating content duplication and enabling persistent references to objects on a dynamic network, resulting in NATO-wide content identification regardless of national boundaries.
- Content objects can be discovered across various types of repository systems, promoting repository interoperability and protecting each country's existing investments in content management systems, thereby lowering the threshold for joining M-FASR.
- Managers maintain total control over the content objects in their own repository systems, allowing content developers and content owners to maintain content currency and accuracy.
- All types of content objects are discoverable, as the granularity of content objects is defined by each repository manager. There are no limitations on the type of content can be shared.
- The architecture is scalable, so additional repositories can be introduced into M-FASR without changing the architecture, an important consideration as additional NATO countries join.
- The federation of registries is scalable as the number of participating repositories increases, allowing easier organizational expansion and change within NATO to ensure the long term success.
- Searches for content objects are highly scalable and efficient because searches are executed on the federated registry only. Participating repositories will not carry additional load for searches so they can concentrate on their main task – managing the content.
- The user interface is easy to use and integrate with other applications. Each participating NATO country can integrate it into its own applications as part of their capabilities.

Some challenging considerations for future action:

- Most repository systems do not have the ability to automatically register a content

object in a registry when the object is put into the repository system. Although this process can be performed manually, it is time consuming and subject to error.

- Repository managers must register content with correct, meaningful, and consistent metadata to produce usable search results.
- Without automated registration, any change to any already registered content object requires a manual update in the registry.

CONCLUSION

The M-FASR project has demonstrated a technical solution that enables content sharing and reuse across heterogeneous repository systems among multiple NATO countries. The tested architecture provides a flexible, scalable, and efficient platform that participating nations can use to discover, access, retrieve, and reuse each other's content. It protects the investments of NATO countries by allowing content sharing across heterogeneous content management systems. It identifies all available content materials across NATO without the limitation of country borders. It leaves control with each country so content can be kept up to date and accurate. This architecture can serve as the foundation for developing a content sharing and reuse infrastructure throughout NATO.

Content sharing and reuse not only reduces and eliminates resources wasted on developing redundant content, but also improve communication, collaboration and cooperation among NATO countries. ADL, NoDEC and CDA, together with other NATO countries, will continue efforts to improve and expand the capabilities of this solution to store, manage, search, retrieve, and reuse content across NATO and partners.

REFERENCES

Advanced Distributed Learning Initiative, (2009a). *ADL Registry version 1.7.* Retrieved June 27, 2009, from http://www.adlnet.gov/Technologies/adlr/default.aspx

Advanced Distributed Learning Initiative, (2009b). *ADL-Registry version 1.7 LOM Cardinality.* Retrieved June 27, 2009, from http://www.adlnet.gov/Technologies/adlr/ADLRDocuments/Cardinality%20Documents/ADL-REGT-1.7-LOM-Cardinality.doc

Advanced Distributed Learning Initiative, (2009c). *SCORM® 2004 4th Edition, Overview*, Alexandria, VA: ADL Initiative.

Corporation for National Research Initiatives, (2009). *The Handle System.* Retrieved June 27, 2009, from http://www.handle.net/

IEEE Learning Technology Standards Committee, (2002). *IEEE Standard for Learning Object Metadata.* Retrieved June 27, 2009, from http://ltsc.ieee.org/wg12/

Isaksen, G., Lamothe, P., & Railer, B. (2008). New Way Of Accessing and Reusing E-Learning Between Countries. *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*

Jerez, H., Manepalli, G., Blanchi, C., & Lannom, L. (2006). ADL-R: The First Instance of a CORDRA Registry. *D-Lib Magazine, Fabruary 2006*.

Jesukiewicz, P. (2008). CORDRA Content Object Repository Discovery and Registration Architecture. *ITEC*.