

Michael Magee

**Author**

Michael Magee ([magee@ucalgary.ca](mailto:magee@ucalgary.ca)) is a Ph.D student in the Graduate Department of Educational Research at the University of Calgary. He has been researching, designing and implementing software for the use of e-learning specifications for the past 4 years. Correspondence concerning this article should be addressed to: Michael Magee, Graduate Department of Educational Research, University of Calgary, 2500 University Drive, Calgary, AB, Canada, T2N 1N4.

**Abstract**

The development of a large body of e-learning specifications, such as IMS and SCORM, has led to the proposal for a new way to facilitate content workflow. This involves the movement of educational digital content and the knowledge of pedagogical communities into an online space. Several projects have looked at the theoretical structure of these specifications. They implemented a series of tools in order to examine and research the issues around the actual usage of these specifications. The CAREO, ALOHA and ALOHA 2 projects were designed to expose both individual users and whole institutions to these ideas. Initial research into the result of those interactions indicates that there is some utility in the adoption of e-learning specifications. The future success of them will depend on their ability to adapt and meet the needs of the educational community as they begin to adopt, use and evolve the way they use the specifications and the tools created around them.

**Résumé:** L'élaboration d'un grand nombre de caractéristiques sur l'apprentissage en ligne, comme le SGE (*IMS*) et le modèle de référence SCORM, ont mené à la proposition d'une nouvelle façon de faciliter le cheminement du contenu. Cela signifie le déplacement du contenu éducatif numérique ainsi que les connaissances du milieu pédagogique dans un cyberspace. De nombreux projets ont envisagé la structure théorique de ces caractéristiques. Ils ont mis en oeuvre une série d'outils afin d'effectuer un examen ainsi que des recherches sur les questions entourant l'utilisation réelle de ces caractéristiques. Les projets CAREO, ALOHA et ALOHA 2 ont été conçus afin de présenter ces idées aux utilisateurs individuels ainsi qu'aux institutions dans leur ensemble. La recherche initiale portant sur les résultats de ces interactions indique qu'il y a des avantages à l'adoption de caractéristiques de l'apprentissage en ligne. Leur succès éventuel dépendra de la possibilité de les adapter et de répondre aux besoins des intervenants de l'éducation alors qu'ils commenceront à adopter et à utiliser les caractéristiques et outils créés pour eux et alors qu'ils modifieront la façon dont ils les utilisent.

**Introduction**

An institution, a discipline of study or a geographic locale ties together the many communities of practice that exist in education. In each of these, some form of common interest and context binds them. It is within these realms that a new community has been evolving over the past few years that includes both the individual and the institution. This group is tied together by a common wish to find, share, and manage digital educational content. More than that, they are looking for effective and efficient ways to manage the knowledge that arises from a complex mixture of content, context and community. This desire crosses boundaries of institution and discipline and is held together by the belief that a knowledge economy based on easy-to-find, re-usable educational content will reward the world with more efficient, effective teaching materials.

A vision that focuses on pedagogy and sharing is not original. The new opportunities arise in the sheer scale and distributed nature of the Internet to facilitate that sharing. In addition, new technologies are rapidly changing the ability to create, manipulate, and transmit content. This paper is about the e-learning standards that are evolving as the language to describe and organize educational content, the motivations for creating this knowledge-based workflow, and a case study of some of the tools that have been developed to explore and evaluate these ideas.

### **Overview of the Specifications**

A range of e-learning standards has been evolving to facilitate simple discovery, organization and description of online digital content. These are the main ones involved with the design and implementation of this series of pilot project tools:

1. IMS metadata was designed to create a common framework for resource discovery and use of educational content (IMS Global Learning Consortium, 2001, 2).
2. IMS Content Packing was designed to organize, deploy, and transport educational content in a simple framework (IMS Global Learning Consortium, 2003, 1).
3. The Sharable Content Object Reference Model (SCORM) has the same goal as IMS content packaging in providing a common technical framework for educational content (ADL 2004, 1).
4. IMS Vocabulary Definition Exchange (VDEX) was designed to allow the description and creation of controlled vocabularies that could be easily interchanged (IMS Global Learning Consortium, 2004, 1).

All of these specifications are focused on creating an interoperable framework for the organization, management, and transport of educational content. Their goal was to limit the amount of time required to find, assess, and repurpose existing content, allowing resources to be focused on pedagogical, rather than technical issues.

With the exception of the work of ADL's SCORM, these specifications are, for a large part theoretical, based on previous research in other related fields. Many of them have not undergone the rigors of testing and iterative feedback that occurs during the adoption of new ideas and technology by a large community. As the purpose of any research is to implement and test out the underlying hypothesis of that research, it was necessary to build tools capable of placing these new specifications in the hands of users. It is only then that the flow of information and knowledge from the community into the framework surrounding educational content could be analyzed and, assumptions, tested. Under this impetus a series of tools were designed to implement these e-learning specifications.

### **Overview of the Framework**

One of the first steps in investigating the implementation of e-learning standards was to examine the current workflow that had been put in place by various stakeholders using digital content. The other major factor was to understand the motivation for those stakeholders to actually implement an interoperable framework. Both of these issues would define the kind of information architecture

necessary to meet their needs, and to evaluate the potential of e-learning standards to address them.

Initial investigations found an informal knowledge management system already existed among a number of professional post-secondary faculty using educational content (See Figure 1). They would gather and discuss the digital content they used in teaching as well as the various approaches they used in the classroom. Often the discussions would result in the transfer of a web link or an email attachment so that their colleague could use the educational content themselves. Although not consciously designed to do so, this informal system met several of the requirements of a knowledge management system, including knowledge creation, sharing, and the promotion of learning and innovation (Gupta et al., 2004, p. 3).

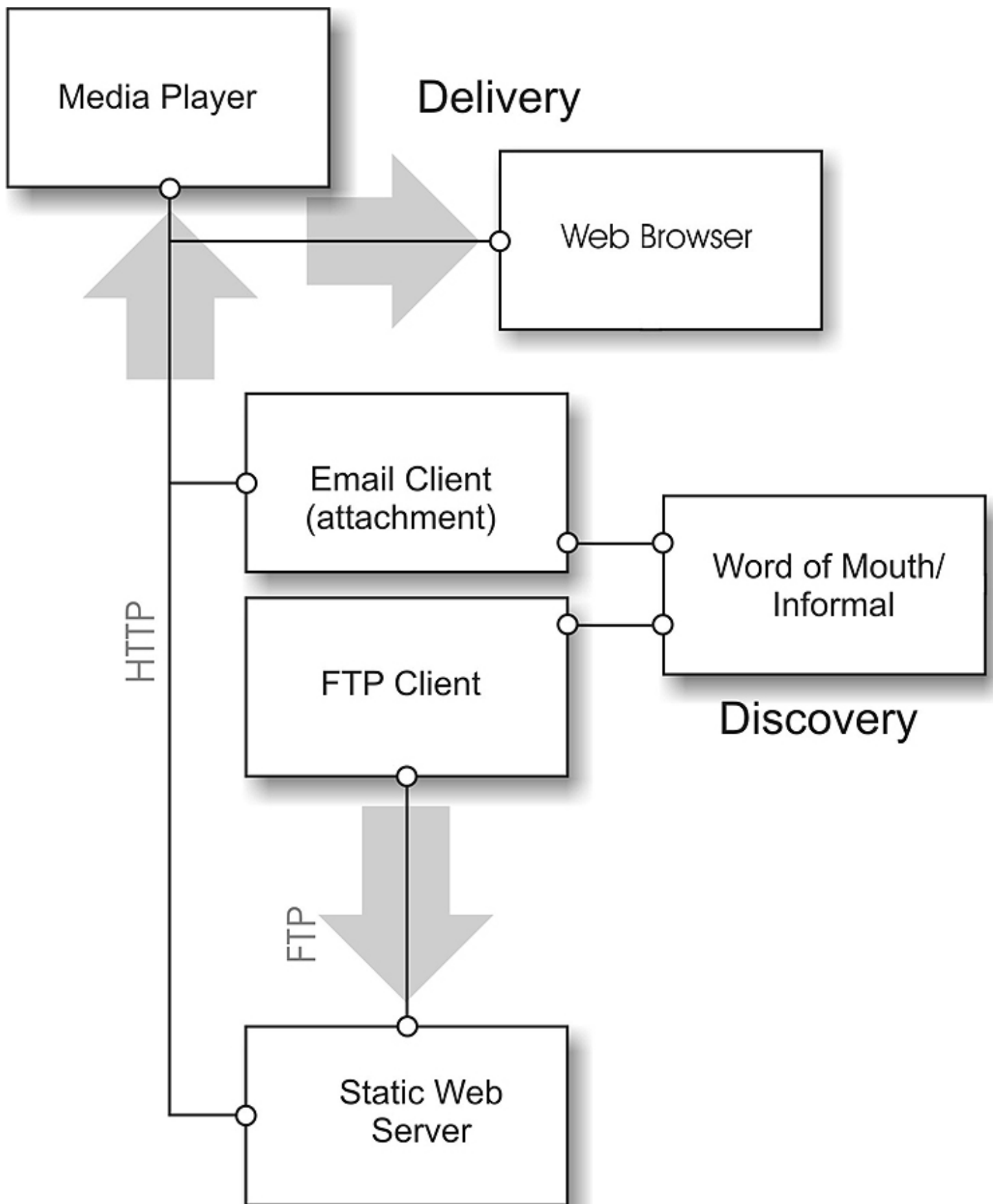


Figure 1. Informal Knowledge Management System

The shortcomings of the system became quickly apparent. As an informal network it was dependent

on personal relationships and would eventually be limited to a small group. There was also no standard way of describing the content, and so users could never be sure if the content they were receiving was even relevant. The faculty stakeholders recognized the importance of the approach and became motivated to build a more comprehensive and useful system to replace the existing workflow. This group was interested in developing a community to share their content, find relevant material, and build relationships with other members of their community. The main focus of these stakeholders included:

1. Sharing of their own content along with their ideological and pedagogical perspectives on that content.
2. Access to a large body of content that would be unattainable within their own institutions.
3. The ability to add value to the existing content and explicitly share the knowledge and wisdom they have gained through the use or re-purposing of the material.

The institutions that employed these stakeholders were also interested in creating a better framework. They had large, unordered collections of digital assets. These include documents, images, video, audio, and websites. They also recognized that the opportunity to capture and share the considerable knowledge their faculty possessed was being lost. They shared the belief, along with their faculty, that capturing relevant knowledge about teaching practices and materials would improve the experience of both the instructor and the learner. They also had other interests in building this framework in order to eliminate institutional inefficiencies and manage risk around the use of digital content. The list of institutional motivations included:

1. Building a comprehensive inventory of the content they owned and controlled.
2. Identifying gaps and deficiencies in the educational content available to the institution. This would allow them to focus their limited content budgets toward updating or licensing content to fill the gaps, or toward creating their own new content.
3. Using e-learning standards as an approach to managing the risk of investing in repositories and future-proofing the framework by providing the greatest degree of interoperability possible.
4. Defining the digital rights surrounding their content and managing access based on those attributes.

Initial investigations focused on the storing, searching, and delivery of educational content in a digital format (Jennet, Katz, Hunter, & Hawes, 2001, p. 3). It defined a fairly simple workflow that would allow users to place content online and index it with a metadata record (See Figure 2). Metadata records were simply the standardized documents that describe and organize the digital content. They were designed to allow the searching of large collections of digital content. Two tools were created to facilitate this workflow. The Campus Alberta Repository of Educational Objects (CAREO) was developed as the web interface for both the metadata created and the media stored online. The main functionality of the repository was to provide user-friendly search and retrieval of educational content. The Advanced Learning Object Hub Application (ALOHA) was designed to focus directly on creating metadata records and moving content online.

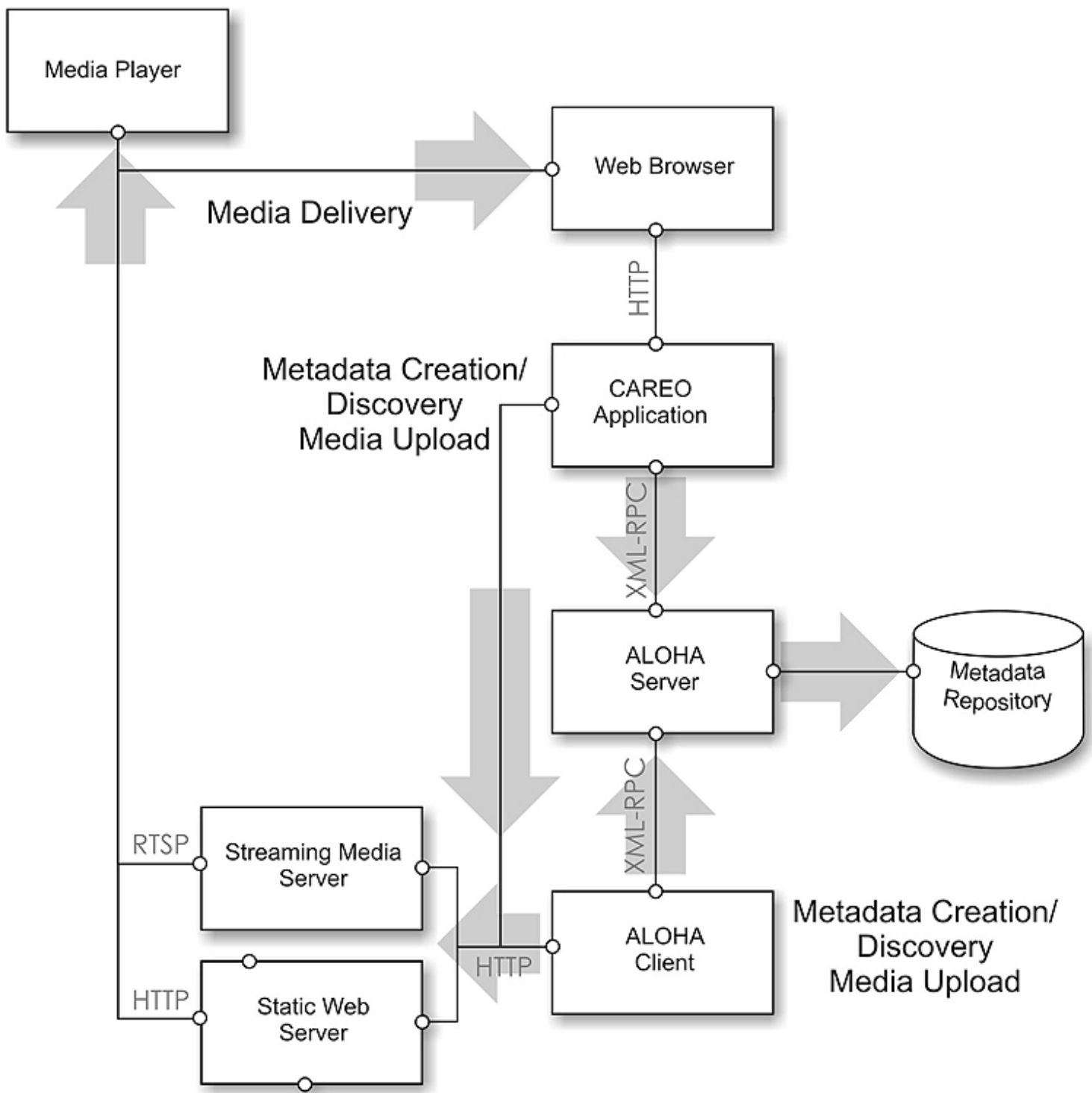


Figure 2. Information Architecture and Workflow of CAREO and ALOHA specifications

A detailed, logical diagram of the system was created that would facilitate all of the features that both the faculty and institutions had requested. This information architecture was compared against their feature list. When the logical design had satisfied the essential features for the storage, searching, and delivery of educational content development of the software began. The building of the tools to create and store metadata was not a large technical challenge. It quickly became apparent, though, that the use and adoption of this software was going to require a more detailed understanding of its users. Now that the workflow had become facilitated it was necessary to examine how users were actually using the tool to create metadata. Several questions arose over whether or not it was realistic to expect faculty to actually author metadata records. There was

also concern over whether or not they understood the specifications such as IMS metadata sufficiently to create records that were actually good enough to meet the needs of searching and retrieval. It was necessary to look at whether or not these tools were actually being used and, if so, how they were being used.

### **The New Indexers**

Suddenly there was a large group of users who wanted to start describing their content in a meaningful way and in such a way that others could find it and take advantage of it. This new group of cataloguers was trying to understand and create IMS metadata much like the library science world had been creating catalogue records for years. This group was different than librarians, however; they had years of experience in their discipline and education but none in the world of indexing and cataloguing. It resulted in a new community that was previously unknown to the indexing world of the library scientist—the naïve user. Most librarians spend years learning the skills necessary to index materials properly; the thought of having an untrained indexer creating records would be unheard of due to the problems that would be created in consistency and completeness of the resulting records.

In identifying and creating a profile of these naïve users, it became apparent that the level of knowledge these users had of metadata and indexing was very limited. The educational community tried to address this with groups like IMS and SCORM that created standardized indexing approaches. These early specifications quickly grew to become very complex and based in new technologies such as XML. The early tools created to build these XML-based e-learning standards metadata records were daunting. Exposure to complex metadata XML schemas and documents as well as numerous metadata elements with arcane names intimidated most users and limited adoption, even among those visionaries who could see the eventual point of it. It became apparent that, in order to enable this new framework of educational object repositories, tools would need to be created that non-experts can use.

One significant issue for the new community was the complex specifications being developed by standards groups. These were considered too esoteric and complex for most professionals within the educational world. As a result these new indexers were not interested in investing a significant amount of time trying to decipher and use these specifications. The most successful groups to organize their institutional knowledge were those who hired assistants for these professionals. The assistants were tasked with learning the appropriate indexing tools and specifications. Comprised of mostly students, this group was usually paid to index materials at their institution as part of a larger strategy toward knowledge management. The professional educators created very few metadata records on a voluntary basis. They simply couldn't see any kind of explicit benefit or reward that would justify the time spent away from their academic duties. Although altruistic behaviour was possible from these professionals, it was not a priority among their tasks.

### **The Tools**

There were a number of tools created during the course of the investigation into educational object repositories and metadata creation. Each one would fill a specific need, but as often occurs, would also emphasize another weakness in the workflow. This iterative process resulted in the evolution of a series of tools rather than a single application that would solve all of the issues with educational object repositories and e-learning specifications.

The driving philosophy behind this development was how to create a tool that will allow institutions to create inventories of their objects, will allow individual users to add value to the content, will

conform to emerging e-learning standards and most of all, make it easy and simple enough that people will actually use it. The whole purpose of all these endeavours is to make something useful that will improve people's lives, not create a rigid set of rules that become so unwieldy that no one will ever be able to make sense of them, let alone use them.

## Careo

The CAREO repository (CAREO, 2004) was browser-based and this limited the kind of technology that was available for metadata creation and media upload. CAREO utilized a basic forms-based interface for creating metadata and a small utility for uploaded media (Figure 3). It provided the basic features necessary for metadata and media workflow in the architecture that had been defined. Although the system worked well, it was not the strength of the repository that had been mainly designed to be a search and retrieval system of educational content. Although many new metadata records were created in the tool, the forms-based interface was not successful in encouraging large-scale creation. It was the lessons learned from this initial development that lead to the development of the more sophisticated and robust ALOHA tool.

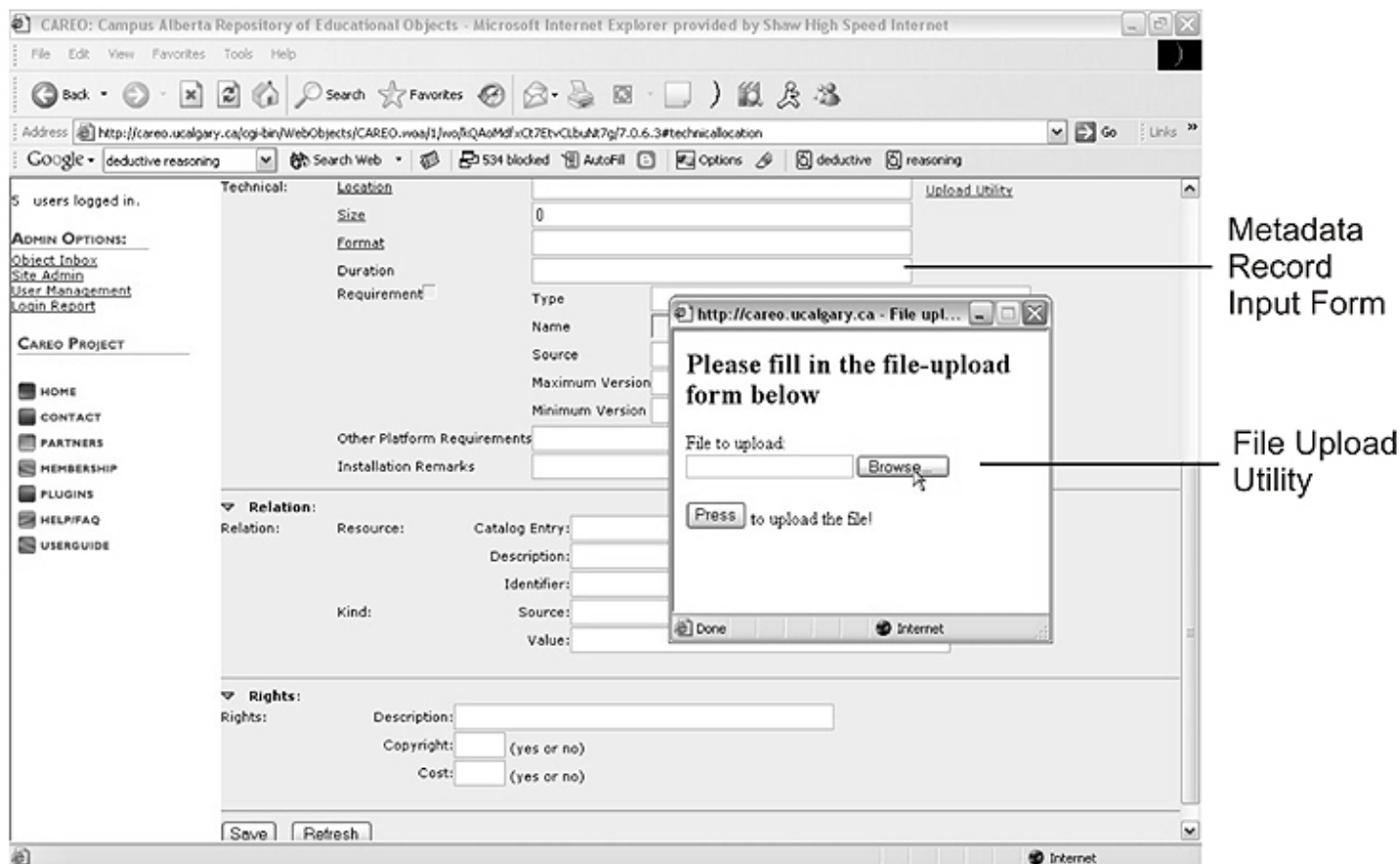


Figure 3. CAREO metadata creation and media upload screen

## Aloha

ALOHA (2004) was a Java-based client that had a primary function of creating metadata records for indexing educational content (Figure 4). Simply dragging files off a hard drive, onto ALOHA, and dropping them there could get digital assets into ALOHA. Once this was done ALOHA was able to automatically parse the media and fill in the metadata for several file types. Any file that has an identified MIME type is capable of having metadata automatically extracted. It was also capable of



identifying the type of media and using that information to send it to the most appropriate media server from an available list. This meant the QuickTime video files would go to a QuickTime video server and Real media files would go to a Real media server. It would generate the URL and automatically place it back into metadata record. It made the creation of IMS or other forms of standardized metadata much easier.

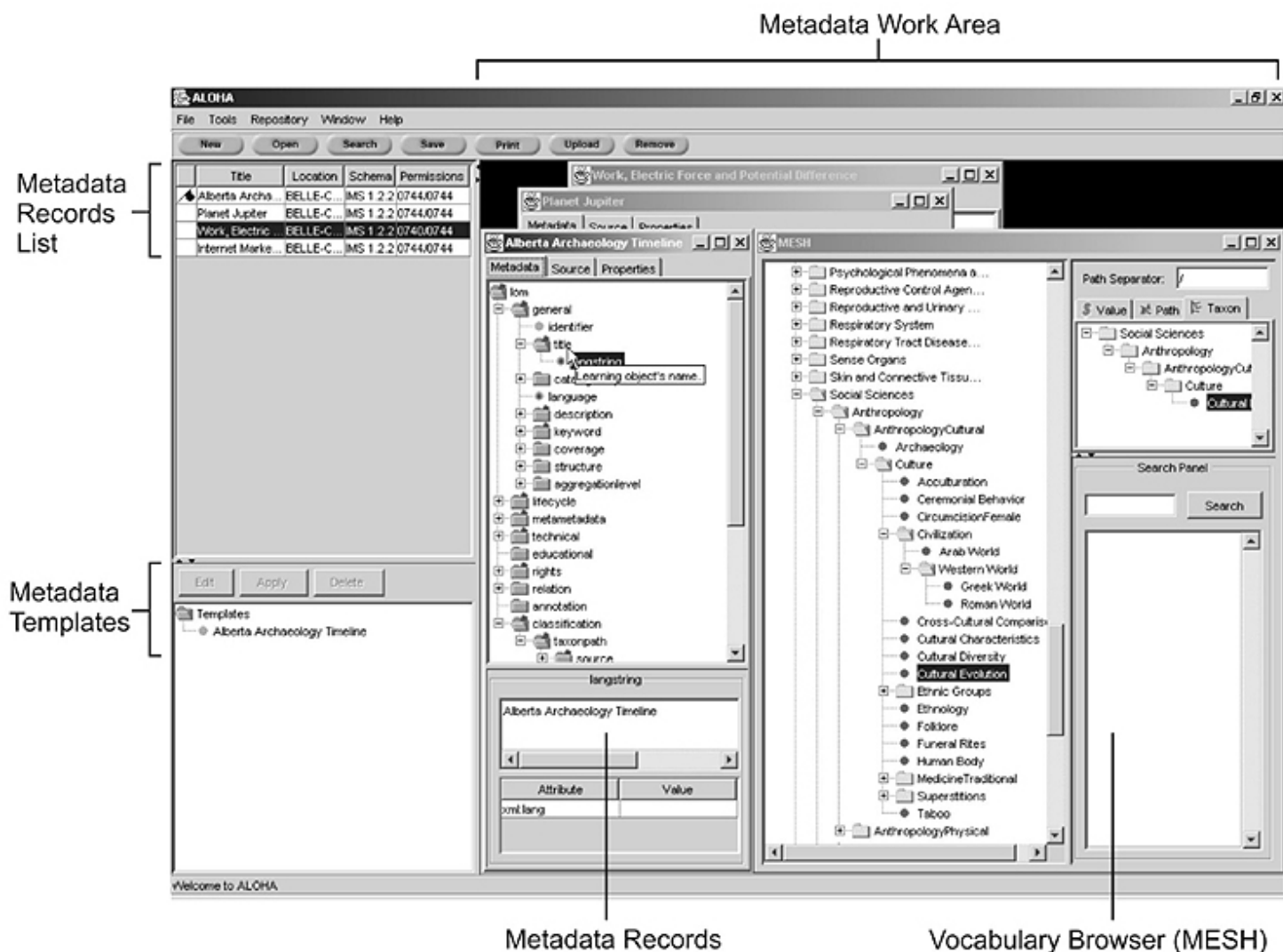


Figure 4. ALOHA Interface Screen.

It did this in two ways, (a) automating as much of the data entry as possible, and (b) automatically checking to make sure the metadata record followed all the rules of the metadata specification. Administration tools were built to manage workflow issues surrounding digital rights management. This meant that multiple indexing groups could be set up to accommodate librarians, educators, media developers.

Other tools that provided useful capabilities to indexers included the "Templates" feature that allows indexers to create metadata for a single record, save it as a template, and then automatically fill in those same fields in other records by dragging them into the appropriate template. The vocabulary browser proved to be one of the most powerful features incorporated into ALOHA. It allowed controlled vocabularies to be added to the tool. Indexers could simply drag large, complex hierarchies of vocabulary onto a metadata record and have them automatically added. This feature became very popular because the communities using the tool often defined themselves and their content using a vocabulary that defined the semantics of their discipline. If the e-learning

specification was lacking in its ability to describe their content this is often where the issue was tackled.

ALOHA proved to be extremely successful. The CAREO repository quickly went from hundreds of metadata records to thousands in a few short months. The large collection of metadata represented one of the largest samples of e-learning metadata created by a naïve indexing community and provided a unique opportunity to look at what patterns were emerging.

### **Analysis of Careo Metadata**

Table 1 represents the top 21 elements used by indexers from the educational community. They represent 3145 IMS 1.21 metadata records that each having a maximum of 86 elements. There are a number of trends that become obvious in viewing the data. Figure 5 shows the most frequently used element, `lifecycle.contribute.centity`, has a very high count and falls above the trend line on the plot of the element count. The count of the elements drops and then gradually decreases. Within the top 21 elements there are five main groups of interest. These relate to (a) general information, (b) ownership information (c) pedagogical context (d) subject classification, and (e) technical information.

The first one is the General group of elements. This includes `general.title`, `general.keyword` and `general.description`. From discussions with the indexers they identified these as the most easily understood and these provided the least difficulty in use. It was also the most likely metadata that would be voluntarily filled out by the indexers.

The second focus was the elements that defined ownership or object lifecycle information. This included `lifecycle.contribute.centity`, `lifecycle.contribute.date` and `lifecycle.contribute.role`. As most of the educational objects came from post-secondary academic institutions, they came from an environment where acknowledgement of authorship was very important. The actual identity of the contributors to an object became the most commonly used element within the IMS group of elements.

Path	Count
lifecycle.contribute.centity	5113
educational.context.value	3333
lifecycle.contribute.date.datetime	3224
general.title	3145
technical.location	3129
technical.format	3076
general.keyword	2935
lifecycle.contribute.role.value	2860
educational.context.source	2846
technical.size	2600
general.description	2026
educational.learningresourcetype.value	1982
classification.keyword	1575
classification.purpose.value	1512
classification.taxonpath.taxon	1479
general.language	918
educational.language	866
classification.purpose	683
classification.taxonpath.source	593
educational.learningresourcetype.source	490
classification.purpose.source	485

### *Metadata Element Frequency.*

The third focus was the pedagogical information, which was also an important component to the metadata. The educational.context element was used to define the situation where the object was most pedagogically appropriate. As most of the content came from educational institutions, it is not surprising that a major focus of their indexing was placing the material in an educationally relevant context. The vocabulary used to define this was the Learning Object Model (LOM) list provided by the Best Practices Guide (IMS Global Learning Consortium, 2001).

The fourth focus was the subject classification of the object. The classification.taxonpath.taxon element allowed complex hierarchical data to be placed in the metadata record. Large controlled vocabularies with several layers of data from general to specific were used to populate this element. The two most common vocabularies were MESH (Medical Subject Headings) and the Alberta Ministry of Education's Learning Outcomes vocabulary. This powerful metadata element allowed the use of several terms to describe the object, but also maintained the relationship between these terms.

The fifth group of elements was comprised of the technical details of the media. The ALOHA tool automatically extracts as much technical metadata as possible from the object as it enters, when it is dragged or loaded into the tool. Although most file formats contain a large amount of metadata, most of the data is proprietary and will not map into the IMS metadata set. This limits the metadata to the common elements of technical.location, technical.format, and technical.size. All of these are extracted by ALOHA and automatically placed into the metadata record.

# IMS Metadata Element Usage

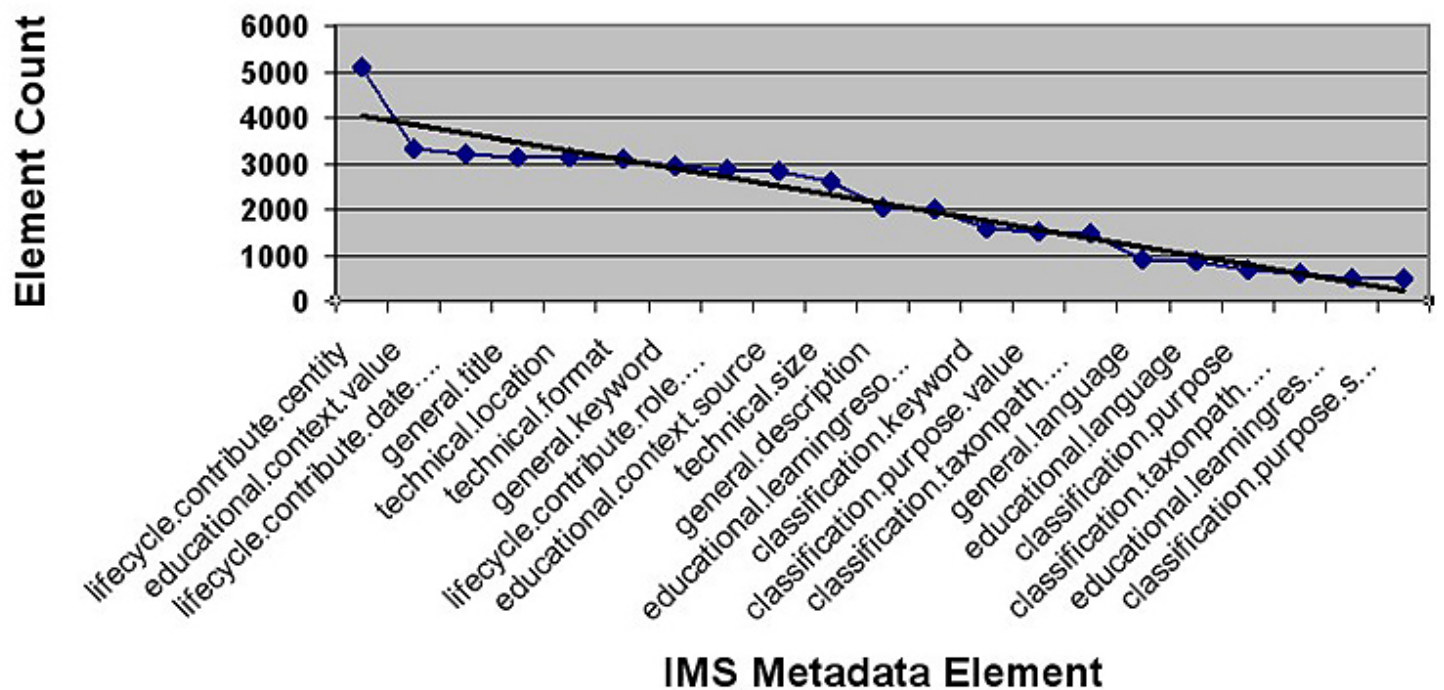


Figure 5. Metadata Element Usage

## Initial Conclusions

Attributing ownership and identifying a relevant context became a major focus for indexers during the course of the project. The ownership of the objects fits well with the importance placed on authorship in the academic realm and provides a potential mechanism for academic credit within the various institutions involved in the project. Placing the educational content into a relevant context is the other priority, although this one will require much more research into its usage. The two main communities identified in this study were pedagogical and institutional/sectoral.

An institutional context was focused on the placement of content into a relevant framework. This often meant alignment to a curriculum or course structure. The educational context element was most commonly used to define this institutional context. Most institutions became familiar with the IMS controlled vocabulary but quickly identified its limited nature. This has been serving as a catalyst for creating and publishing controlled vocabularies that are more finely tuned to meet their institutional needs.

Pedagogical context often overlapped with the way an institution viewed content. Despite this many instructors wanted to provide a more meaningful description than just a course classification. The education context element was quickly identified as being too limited so many communities gravitated towards the classification.taxonpath.taxon element. This element allowed a nested hierarchy of vocabulary terms to be used that could be used to describe subject matter and learning outcomes. This element was also used by instructors to define the semantics of their discipline. One example was the MESH (Medical Subject Headings) vocabulary that was used to describe objects being used in teaching anatomy. This controlled vocabulary worked well in providing a general down to specific list of terms to describe the object. The context became more

focused on the discipline of anatomy than on a specific educational context, and provided another example of a community-defined context.

Ownership and identity were the least complicated to understand. There was only a single place in the metadata record to place their information. The lifecycle contribute element allowed the addition of vcard-formatted descriptions that could provide a wide range of descriptive terms for identifying the owner of a particular object.

The high level of technical information elements in the metadata records was likely due to the automated metadata extraction performed by the ALOHA indexing tool. This feature did not require the indexer to create the information independently and therefore did not present any additional effort during the indexing workflow.

## Aloha 2

Many of the lessons learned in this first round of development have been incorporated into ALOHA 2 (<http://aloha2.netera.ca>). The project was part of the eduSource project (eduSource, 2004) and represented an international collaboration between the University of Calgary, LearnAlberta and Reload, a European metadata tool. Although the tool provided an increased level of functionality the focus of the project wasn't just to create a more feature-rich metadata tool. The software needed to be much more user-friendly so users could spend more time on filling in valuable knowledge and less time just trying to understand the educational standards. This message had been delivered every time the previous projects had moved to an adoption phase.

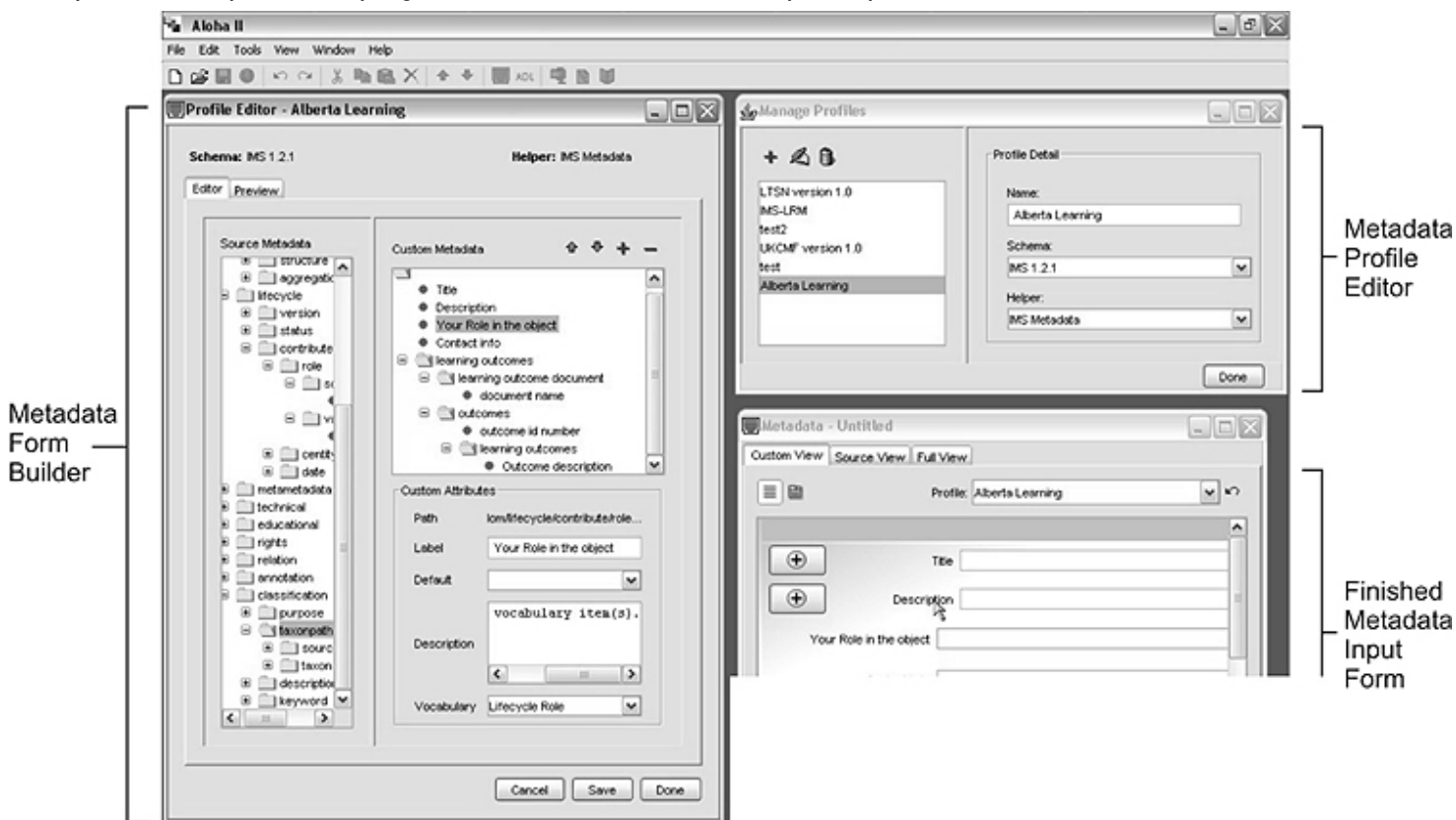


Figure 6. ALOHA 2 interface screen

In terms of new features ALOHA 2 added IMS Content Packaging, IMS VDEX (vocabulary definition and exchange) and SCORM 1.2 specifications. To improve usability it provided a new form builder.

This allowed the interface for creating metadata to be easily tailored to meet the needs of different groups and more importantly increase their willingness to author metadata. The form builder included the ability to create custom labels for the metadata elements, assign a default value to the element, provide a community specific description and assign controlled vocabularies. The creation of a more simplified interface for dealing with standards was seen as crucial to adoption. The level of adoption and any changes in the way people use e-learning specifications remains to be seen as the project is in its final stages. Initial feedback is excellent but it will take some time to determine if the tool succeeds in all of its promises.

### **Recommendations for the Future of E-Learning Specifications**

The knowledge gained from these initial research pilots must be capitalized on in order for e-learning standards to have greater adoption in the education community. We are in a new era where information technologies are offering us new opportunities and potential. Despite this we are still trying to understand how to use them effectively. We will need to figure out how to reorganize existing production and improve the workflow of placing educational content to the web. Then we will begin to see some of the promised benefits of using these new paradigms in digital content and e-learning specifications.

One of the critical steps is to make the entry of information as painless as possible. This involves streamlining the workflow for getting content online and limiting the amount of new skills required to do so. Exposing users to complex specifications and unusable software tools will quickly doom any attempt by an institution to implement strategies involving repositories and e-learning specifications. In both the CAREO and ALOHA projects users needed a very simple starting point. Once they had gained mastery of the basic concepts, they quickly began to demand more sophistication and capacity from the systems. At that point there will be less resistance towards investing in more robust content management paradigms. It is critical that the new software is not perceived as yet another technology that is being pushed for the sake of technology, rather than for productivity. To this end knowledge transfer must be recognized as important, and there is a need to do a continued return on investment on the metadata to determine which elements are most meaningful to the different user communities.

It is becoming apparent that there are different needs for different communities. Each of these communities will need to educate themselves, define the semantics of their group and identify which software is most relevant to them. They are going to need choice and flexibility in order for them to recognize a worthwhile investment in e-learning specifications. They will take on the responsibility for managing the knowledge created in their own community, and most importantly they will continue to evolve. The way they describe their online content and the knowledge they add to that content will be refined and updated with each cycle of reflection and iteration. This will result in the same content being defined in many different ways (See Figure 7).

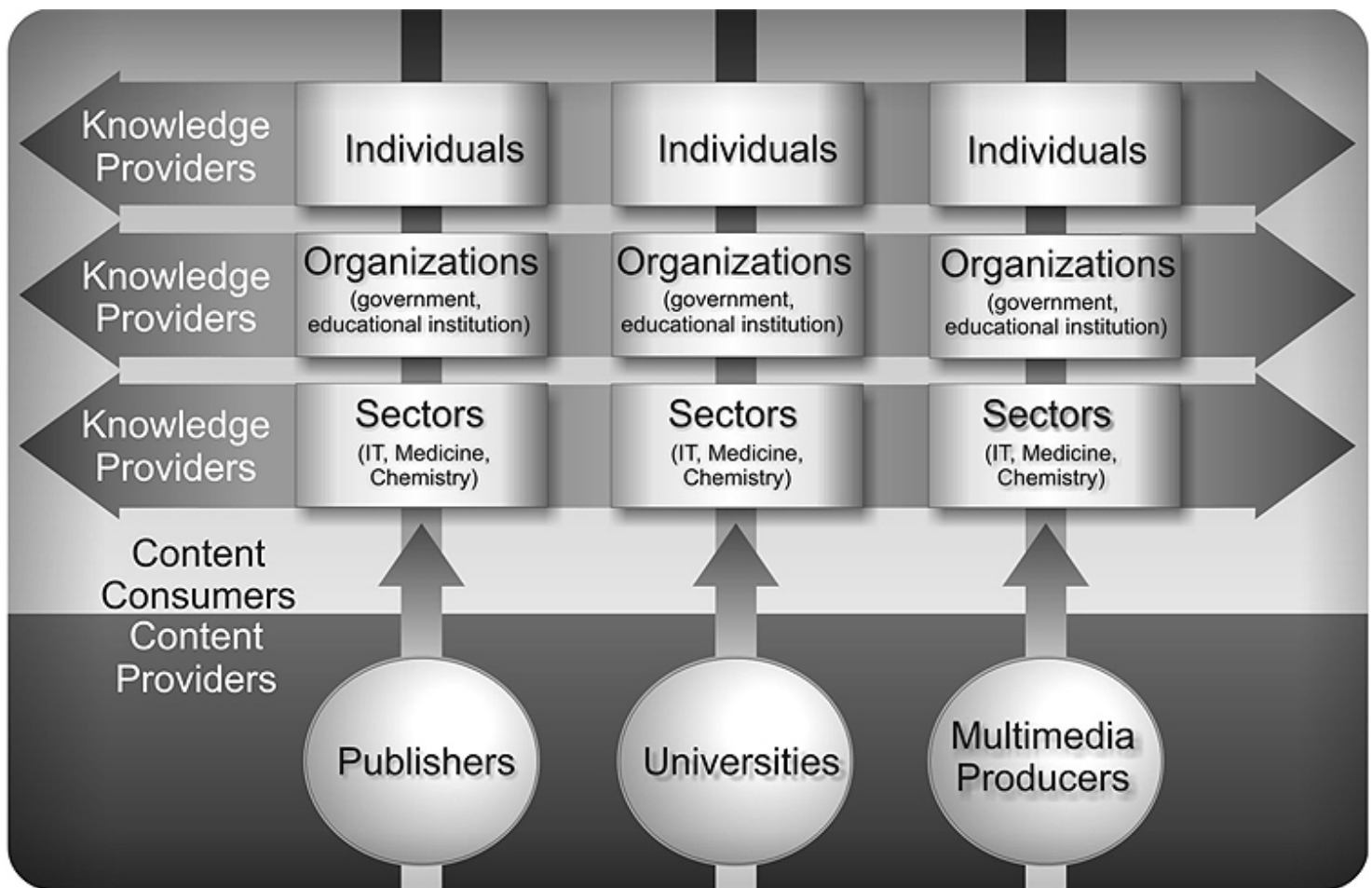


Figure 7. Knowledge Creation by Different Communities

Facilitating this kind of information flow will encourage innovation, creativity, and further development of content. It is apparent that if e-learning specifications are going to achieve adoption and become more than just a tool for organizing information, they need to be able to give each community of practice the power to place information into a meaningful context. At the point the specifications allow meaning and insight about digital media to be made explicit they will become part of the knowledge structure of that community. In order to achieve this they will need to have an active role in defining the specifications they use. These communities will not adopt a structure that is dictated through a top-down approach and not tailored for their use. They will quickly subvert the standard or ask the more fundamental question about why they should even care? They need to be able to push from the bottom in order for the specification to work for them. In order to facilitate this they will need expertise that will provide the guidance they need to make their own decisions.

The actual outcome of these new communities is still unknown. What boundaries will demarcate them and the kind of content they ingest and produce will continue to evolve as the level of awareness increases and the sophistication of the tools increase. For this reason the process will never really be finished because knowledge is never a static entity.

## References

ADL. (2004). *SCORM overview*. Retrieved May 14, 2004, from <http://www.adlnet.org/index.cfm?fuseaction=scormabt>

ALOHA (2004). *ALOHA metadata tagging tool*. Retrieved September 21, 2004, from <http://aloha.netera.ca>

ALOHA2 (2004). *ALOHA metadata tagging tool*. Retrieved September 21, 2004, from <http://aloha2.netera.ca>

CAREO. (2004). *Campus Alberta repository of learning objects*. Retrieved September 21, 2004, from <http://careo.netera.ca>

EduSOURCE. (2004). *Canadian network of learning object repositories*. Retrieved September 1, 2004, from <http://www.edusource.ca/>

Gupta, J. N. D., Sharma, S. K., & Hsu, J. (2004). An overview of knowledge management. In J. N. D. Gupta and S. K. Sharma (Eds.), *Creating knowledge based organizations*, (pp. 1-28). UK: Idea Group Publishing.

IMS Global Learning Consortium. (2001). *IMS learning resource meta-data best practice and implementation guide, Version 1.2.1, Final specification*. Retrieved May 14, 2004, from [http://www.imsglobal.org/metadata/imsmdv1p2p1/imsmd\\_bestv1p2p1.html](http://www.imsglobal.org/metadata/imsmdv1p2p1/imsmd_bestv1p2p1.html)

IMS Global Learning Consortium. (2003). *IMS content packaging best practice guide, Version 1.1.3, Final specification*. Retrieved May 14, 2004, from [http://www.imsglobal.org/content/packaging/cpv1p1p3/imscp\\_bestv1p1p3.html](http://www.imsglobal.org/content/packaging/cpv1p1p3/imscp_bestv1p1p3.html)

IMS Global Learning Consortium. (2004). *IMS vocabulary definition exchange best practice and implementation guide, Version 1.0, Final specification*. Retrieved May 14, 2004, from [http://www.imsglobal.org/vdex/vdexv1p0/imsvdex\\_bestv1p0.html](http://www.imsglobal.org/vdex/vdexv1p0/imsvdex_bestv1p0.html)

Jennet, P., Katz, L.; Hunter, B., & Hawes, D. (2001). *The multimedia database initiative: The health/education cluster project – year 2 final report*. Retrieved May 14, 2004, from <http://www.kin.ucalgary.ca/cluster/docs/final-report-2001.PDF>

---

© Canadian Journal of Learning and Technology

**ISSN: 1499-6685**