

Report and Recommendations

Directory Services Infrastructure Working Group*

November 18, 2003

Contents

1	Executive Summary	3
2	Introduction and Background	5
2.1	The Group's Charge	6
3	Directories	6
3.1	What is a Directory?	6
3.1.1	What is a Directory Service?	6
3.1.2	Types of Directories	7
3.2	Roles and Identity	8
3.3	Authentication and Authorization	9
4	The Larger Directory Context	10
4.1	The Role of Directories	10
4.1.1	Directories in a Commercial Setting	11
4.1.2	Directories in an Academic Setting	12
4.2	Summary	15

*This document is intended for internal reviewers only. Please do not redistribute it. Please send comments or feedback to the document's editor, paul@umich.edu and include the version number (1.7.1.22, last edited 2003-11-18 10:07:18-05) of the document. This report reflects the substantial work and input of the team and is a jointly authored document. Errors and omissions during the writing and editing process are the editor's.

5	Directories on Campus	16
5.1	The Campus Directory Context	16
5.2	Directories on Campus	16
5.3	UMOD as a Point of Integration?	18
5.3.1	A Perspective on UMOD in a Changing Directory Environment	19
5.3.2	UMOD, ITCS, and the Future Directory Environment	21
5.4	Overall Problem Areas and Issues with Directories on Campus	21
5.5	The Direction of Change	22
5.6	Conclusion: Moving Towards an Enterprise Directory for Campus	23
6	Recommendations	24
6.1	Identity Management and Roles Data	25
6.2	Feeds	26
6.3	Governance	27
6.4	Leverage	27
6.5	Alignment	28
7	Next Steps and Moving Ahead	28
7.1	Pilots and “Low Hanging Fruit”	28
7.2	Other Recommendations	29
7.3	An Enterprise Directory Project Definition	31
7.3.1	Project Structure and Planning	31
A	Active Directory and UMOD	33
B	Roles, Eligibility and Authentication	36
B.1	Identity	37
B.2	Extensibility	38
B.3	Namespaces	39
B.4	Directories and Authentication	39
C	Feeds	40
D	Benchmarking	42

Abstract

This report summarizes the results of meetings and other activities of the Directory Services Infrastructure Working Group. This report is intended to have value for the layperson and also provide enough information and context for those who may choose to become engaged in the discussions and decisions in an effective way.

1 Executive Summary

A broad statement of this working group's conclusion is that the campus requires but lacks a comprehensive *Enterprise Directory*. The campus's goals and ambitions require that we move from the current "white pages" and e-mail forwarding functionality to a much richer and more application-oriented enterprise-wide directory service. Whether or not this new enterprise directory service will be evolve from existing services or be newly formed is an open question.

Defining what this new enterprise directory should be is not straightforward. Two "marketplaces," each with their own approaches to directory services and characterized in this document as "commercial" and "academic" are inherently different and integrating them poses a challenge. In addition, the assessment of the value present in directory services compared to the work required to provide them and maintain related services is not entirely objective. Instead, it reflects the observer's overall biases towards middleware and its impact on the computing environments he or she is most familiar with.

The lack of an enterprise directory on campus carries with it some consequences. With limited alternatives, units that resort to implementing their own directories incur extra costs and produce directories that limit interoperability with other units. Projects that require a consistent and comprehensive campus-wide directory service are not able to proceed. Factors that contribute to this include issues specific to this campus (see section 5) and quite possibly a fundamental difference between how the academic and corporate sectors view directory services and related products (see section 4).

An enterprise directory would have the following capabilities and attributes. Creating an enterprise directory on campus would require major tasks to develop and improve capacity in these areas: (see section 6):

- *Identity Management*, a sophisticated management of user accounts and consolidation of user profiles. The scope of Identity Management is large,

dealing with lifecycles of identities across the enterprise or even beyond, as it may include vendors and others with a variety of relationships to the University. Identity management includes both directory and security services. The goal of identity management systems is to allow the consistent cross-campus use of enterprise-wide identity and security data and management policies.

- *Roles Data*, or the grouping and sharing of user attributes. In combination with an identity management system, roles data would allow for finer grained and better authorization decisions.
- *Feeds*, or the management of data flows into the directory and the resolution of issues related to data precedence, updating and expiration.
- *Governance*, or the oversight and alignment of the overall operation and direction of the directory from an institutional point of view. Governance would cover both policy and operational issues, such as privacy and data organization in the directory.
- *Leveraging* the directory, or allowing units on campus to take advantage of as much centrally provisioned directory services as possible and to minimize the number of isolated unit directories.
- *Alignment* with other directory-related projects or products, or being more consistent with national and international efforts in the academic and research worlds as well as with industry directions.

The work required to address these issues would be organized in the following categories:

1. Choosing near-term tactical and operational tasks that need to be accomplished in order to provide or maintain basic levels of service in current directory and directory-related services on campus.
2. Examining farther-reaching architectural and strategic components that would define the enterprise directory. This category consists of two sub-areas that are worth noting as distinct areas.
 - (a) Architecting and selecting components of the new enterprise directory service on campus.

- (b) An open-ended strategic and scoping discussion of topics such as identity, roles and alignment. How these terms are defined will play a critical role in determining what new functionality would be provided in a campus enterprise directory.

Whether or not there are two or three overall categories of work here may be debated. It is clear work in one will inform and affect work in the other(s). A larger discussion of the cost and benefit of implementing different levels of functionality and what should be implemented as near-term projects will provide the backdrop to these other efforts.

Specific recommendations are provided in section 6. The three phases of work are described in section 7.

2 Introduction and Background

The directory services working group is one of three original working groups that examined fundamental and basic core infrastructure services on campus. Directory services are a key component of fundamental information technology infrastructure. The growing importance and centrality of directory services and the dependence that other “middleware” components have on directories demand a heightened awareness of directory services on campus. Directories provide a basic and accessible view of a potentially wide range of data. They support and inform authentication and authorization environments in fundamental ways. The usability of current technology on campus will be enhanced by a richer set of directory-based resources. More importantly, investments in future technology such as course management systems, grid-based computing, digital asset management systems, video conferencing and voice over IP will be more effectively used if an underlying and capable directory environment is present.

We assume that if directory services can be shared and leveraged more effectively across campus, more time and resources will be available to advance other projects more central to specific units’ objectives. The strong association between a campus directory and UMOD in many peoples’ minds combined with the long standing use of whitepages directories on campus both helps and hinders the debate and understanding of the role of directories on campus.

2.1 The Group's Charge

Our charge was to:

- understand what UM's current directory capabilities are, what needs we have that current systems cannot meet, and what barriers stand in the way of meeting those needs;
- investigate what other institutions and Internet2 are doing, what commercial solutions, etc. are available (benchmarking).
- determine the technical, policy, and procedure options available to us for moving ahead

We conclude this phase of our group's activity by releasing this report which we feel fulfills our requirements.

3 Directories

3.1 What is a Directory?

A directory is like a database, but tends to contain more descriptive, attribute-based information. It is not a replacement for relational databases or file systems but performs very well when supporting fast and efficient retrieval of data elements for a wide variety of applications. An entry in a directory is a collection of attributes. Each directory entry has a name, and these names are used to locate information in the directory. A familiar directory is the white pages section of a telephone book. That directory's entries are telephone numbers and street addresses that are associated with and looked up by peoples' names. The U-M Online Directory (UMOD) is probably the most familiar example of an electronic directory to the campus community.

3.1.1 What is a Directory Service?

A directory service is the information model combined with a protocol for querying and manipulating it. LDAP, the Lightweight Directory Access Protocol, is an Internet-standard protocol for accessing and managing directory information. LDAP is a well-known protocol with many implementations available. The LDAP provides answers to the following questions:

- What kind of information can be stored in the directory?
- How is the information arranged?
- How is the information referenced and accessed?
- How is the information protected from unauthorized access?

LDAP is the directory service this document focuses on as it is the recommended standard for higher education and academic enterprise directories.

3.1.2 Types of Directories

For the purposes of this document, we will consider three different types or categories of directories.

1. A *general purpose* directory is a large, shared directory of general interest to the membership of the broad University community. A general purpose directory's data will be useful and important across administrative domains and inter-institutionally. As a result, the process by which we determine what data will be in the general directory requires negotiation and consensus building among multiple constituencies. An "Enterprise Directory" would presumably fulfill this role.

A directory is often fed from a directory database or repository. This is because relational databases are usually the source of authoritative information about people and their roles in the institution and can easily be used to bring together multiple sources (feeds in our terminology) to a single source stream from which to populate the directory.

2. *Application directories* are intended to support a specific application. Microsoft's Active Directory is an example of a large-scale application directory. The directory that nearly every school and college on campus operates to provision local accounts creation or remember who gets door keys would be an example of a smaller scale application directory. Stand alone directories that are shipped with particular applications to support their authentication and authorization requirements are examples of application directories. A reservation or checkout system will likely include an application directory. Application directory functionality, such as provided by AD, will need to be supported and used as the applications demand, but LDAP should be

viewed as the target protocol for applications on campus. An Enterprise Directory would presumably provide application directory services to a range of applications of campus.

3. A *metadirectory* is a centralized hub that collects data from any number of other directories and data repositories (including legacy systems, unless those systems and repositories are connected using the directory database approach) and joins them together into a logical whole. The primary purpose of the metadirectory is to provide a name or ID space reconciliation, central name space management and cross indexing of identifiers. One example of reconciliation would be determining if the network ID yyonson, the e-mail address yon.yonson@umich.edu and Yon Y. Yonson in another data source all referred to the same person. A metadirectory is a more application transparent means for integration than a central application directory. A metadirectory may also be characterized as a “directory of directories” which may evoke a different approach to an integration of various directory services.

On campus, there are a lot of application directories, a small number of “enterprise” directories where the enterprise is smaller than the entire University campus (such as the eDirectory deployment by MCIT) used in some units and a set of feeds and data transformations related to those feeds that represent some metadirectory functionality. There is not a single campus-wide directory with full ED functionality.

3.2 Roles and Identity

An individual may be described in terms of groups that include her. For example, Babs Jensen may be an undergraduate in LS&A, a resident in South Quad, a temporary employee at CAEN and an applicant for admission to the MBA program in the Business School. These group membership attributes of Babs are called “roles data.” Associations made based on group memberships are of special interest in a directory service context. Role data could easily be used to grant access to data for people in the university community.

Being able to use data in a directory to answer a questions like “Who is Babs Jensen?” and “Is the person identified as ‘Barbara Jensen’ by the Psychology Department and ‘Barbara Z. Jensen’ by the payroll system the same person?” are also of special interest to those who are operating a directory service. Enabling

those questions to be answered they same way regardless of where the question is posed on campus is also of special interest. “Identity management” is also an important aspect of directory services and identity management systems provide the mechanism by which these issues are addressed. We believe that the nature and importance of identity management in higher education are significant determinants of overall directions of directory systems on campuses.

The name Babs Jensen and the roles data that describe Babs in terms of her memberships and affiliations comprise an identity. Collections of these identity associations are sometimes called name or identity spaces. Each name space is associated with a specific set of policies. Those policies indicate how the identity was created and how it will be destroyed, in which applications this ID is valid. or what kind of data can be accessed by the ID. Different name spaces may be in different administrative domains. That means there are different authoritative sources for the data associated with the names.

3.3 Authentication and Authorization

Authentication is a process by which the identity of a user (or any entity) is proven or verified. Authorization is the granting of access rights to a user (or other entity). Directories have a linkage to authentication systems because both of them are involved in the management of peoples names and identifiers and the establishment of a policy for the name space. Directories can provide these identity, authentication and authorization services directly from data stored in the directory or they may be used as a common “front end” to other sources or systems such as Kerberos for authentication or relational databases for particular data fields.

A person who successfully authenticates as Babs Jensen will be able to access applications and data depending on authorization decisions made based on the role data associated with Babs in the directory. Because these systems support names and identities over a range of time and across different contexts, they are called lifecycle management systems. Internet2’s best practices guidelines point out that “[T]he problems of making directories work are closely related to the problems of making identifiers and authentication work.” Name spaces are discussed further in appendix B.3.

4 The Larger Directory Context

The NSF Middleware Initiative (NMI) defines middleware as “software that connects two or more otherwise separate applications across the Internet or local area networks. More specifically, the term refers to an evolving layer of services that resides between the network and more traditional applications for managing security, access and information exchange.” Middleware software applications cover the areas of identifiers, authentication, authorization and resource management. Internet2 describes a directory’s role in an institution’s infrastructure as

... the operational linchpin of almost all middleware services. They can contain critical customization information for people, processes, resources and groups. By placing such information in a common storage area, diverse applications from diverse locations can access a consistent and comprehensive source for current values of key data. In future information technology environments, directories will be among the most critical services offered.

Middleware is an important and developing area of software. Directory services is a key and very prominent facet of middleware.

4.1 The Role of Directories

One’s perspective on the general question of “what is a directory?” is critical in understanding U–M’s situation related to directory services. There are many views of what a directory is and varying expectations of the directory’s functionality. These views are strongly informed by an orientation between two different perspectives of directory services. The first is one’s personal or unit’s experience with managing data to provide directory-like functionality. The second is the directory “market” or community one is in. Other influences could include beliefs about “build or buy” and “open or closed source” decisions as well. Others are expressed in terms of specific and unique solutions a unit or organization may be seeking.

As noted above, directories play a significant and critical role in the overall architecture of a modern enterprise information technology implementation. They either directly provide identity, authorization and authentication services or are closely integrated with systems that do. Directories are a key piece of the middleware software environment.

Across the campus, perspectives have been formed without a direct awareness of an environment that includes pervasive directory-enabled applications. This, combined with exposure to an information technology infrastructure that is common in higher education has played a major role in creating our expectations regarding directory-enabled services on campus. Comparing “commercial” and “academic” requirements may help illustrate this.

4.1.1 Directories in a Commercial Setting

The commercial view of directories is represented in the large scale systems world by large scale sourcing providers and consultancies such as Burton Group, Price-WaterhouseCoopers, IBM, Accenture and their clients. This market sees directories as being integral to the centralized provisioning of identities and the management of applications. The process of creating a user id and getting a password or token, resetting the password, transitioning roles and eventually deleting the user id — all parts of lifecycle identity management — has as its goal the realization of lower costs by making this a more efficient process and system across the entire enterprise. Return on investment and time to payback justify the deployments of these systems.

Some of the major tasks that a large scale commercial directory project will undertake include:

- User account provisioning
- Integrating data required by different enterprise-wide services, such as:
 - network operating environments,
 - ERP environments like Peoplesoft, and
 - applications, by providing the location for saved configuration and preference information.
- Development and support of Public Key Infrastructure (PKI) support as a robust model for security.
- Self service applications that allow individuals to change their passwords or demographic information where business process allows doing so.
- Providing a convenient and reliable source for authoritative institutional data

Centrally maintaining user IDs and associated roles data is a standard approach in a commercial setting. The same central system tends to provision and strongly control other services that the directory supports.

Not surprisingly, the benefits to be gained from such centralization are seen as lower user provisioning costs through the entire lifecycle of identity management, a central location for employee information and a lower support cost from reduced signon processes.

4.1.2 Directories in an Academic Setting

Because of the large size and wide scope of an institution like University of Michigan's business operation, a comprehensive approach similar to those seen in commercial settings management of directories could be expected.

However, the higher education and research community's requirements lead to a view of directories that is more sophisticated and nuanced than this. In the academic world, the selection of primarily open-source implementations of standards-based protocols (IMAP, POP, SMTP, LDAP, AFS, and Kerberos are prominent examples), the expectation of full interoperability with these implementations, and minimal vendor lock-in define much, if not all, of the enterprise infrastructure. The independence of campus units and the heterogeneity of their information technology environments are reflected in the choice of infrastructure design and requirements. While these implementations certainly interoperate (IMAP uses Kerberos) they are not highly integrated. The point of integration for many of these applications is at the user's desktop at the application level, and not at the server or infrastructure level. Constraints on dollars available for software purchase combined with what can be argued are cultural predispositions help create the path that has led to our current environment for academic computing on campus. These issues also place constraints on the style and pace of adoption of commercial software, sometimes to good effect and sometimes not.

A federated, somewhat loosely coupled approach is usually taken, with all participants agreeing to interoperability at a fairly arm's-length manner with interfaces defined by the protocols mentioned above with multiple implementations. In the directory setting, this places a greater weight on supporting these protocols. Units tend to maintain authority over their data and tend to lack the authority to mandate particular modes of access and use. There is a lack of centrally provisioned data or data definitions. The key application on campus remains e-mail, and the directory (UMOD) supports e-mail very well. E-mail addresses and groups are self-maintained and are not "institutional" data.

In addition, privacy tends to be more highly valued in an academic setting. Information technology and provisioning tend to be much more highly distributed as well and the decision of whether or not to provide services is made unit by unit. It is important for identity and related credentials to be exported and used by a variety of units who may be authoritative for different data elements. Individuals in academia often have multiple roles and move freely between and among units which amplifies the importance of being able to transmit and use credentials and other data across campus. As a result, a significant contribution for a comprehensive directory infrastructure in an academic setting would be to provide adequate identity management so at least units are spared the costs of that aspect of providing services.

Higher education may be a niche market to much of the industry and different demands and constraints drive institutions of higher education to take a different view of commercial products than many corporate entities. The list of directories in section 5 supports this view. The view of directories in this context is better expressed by organizations like Internet2 and the National Science Foundation Middleware Initiative. Federated and voluntary associations develop standards and APIs with reference implementations. Local business process tends to be abstracted out of this view, and each participant is left to do that integration on their own.

In the national higher education and advanced research setting, one sees examples of “big” middleware, such as projects with thousands of people sharing and creating information across hundreds of units or institutions. Grid-based computing is commonly cited as an example of an important new technology that requires a capable set of middleware applications. Sharing resources across and within institutions like the University of Michigan and their various partners, vendors and constituencies requires a rich and robust set of services that we are only beginning to see on this campus. Any *web services* or XML-based distributed computing architecture will also depend heavily on middleware services.

In higher education markets the notion of a central directory for a wide range of integration activities is replaced, or at least augmented, by a range of independent network operating environments and application specific directories. These environments and directories are governed by unique policies and procedures that differ between and within institutions. In this environment, identity management becomes even more of a critical issue and is still an important objective. Yet the same forces that create the requirement for such an approach also militate against its implementation and that is the heart of the issue of an integrated campus-wide approach to directory services.

Identifiers are a strong motivator for the foundation of the campus middleware infrastructure. The ability, or lack thereof, to share identifier information across a number of independent or loosely coupled units defines the campus experience in this area. If an entity cannot be named it cannot be found. If it cannot be found, it cannot be provisioned or authenticated.

This background can be seen in this list of motivations for an enterprise directory in higher education and research.

- There are usually a lot of gray areas surrounding access policies for particular information technology resources. Rules for access may be very organic, developed informally over time and reflect many small and subtle “tweaks” as finer and finer distinctions about who gets what kind of access is decided. In some cases, there may be a lack of internal consistency of definitions and process. There is also a lack of consistency about how access is defined across units.

At the same time, authorization decisions are monolithic, in that all authenticated entities enjoy the same level of authentication to most services. Unlike the use of roles data in directories in a commercial setting, user account publishing is used to provide the basis for authorization decisions on campus. Creating a campus-wide definition of particular account types that are useful across campus is difficult. At the same time, it is becoming more difficult for separate systems to remain islands; increasingly, business processes and academic goals require the use of more than one system. These systems can be a combination of local, institutional or national or consortia-based.

- Making one or more central organizations responsible for life-cycle management of these IDs, however unlikely that may be, is likely to result in a more coherent and systematic management of the ID space. An ID provider could vouch for some aspects of the person or entity represented by the ID as part of the account creation process. If an ID space exists, then departments have the option of leveraging that ID space (both people and groups) and using it in the management of their own local systems. More broadly, unit-specific applications do not need to build their own middleware, but can leverage the enterprise directory when possible to save development time and money.

This is likely to make the departmental systems more secure by allowing the owner or steward of related data to make decisions that will be consistent

across all users of the data regarding privacy or other policy-related issues. This could include access rights or the level of authentication required to view or modify the data. If departments are leveraging a central ID space to provide user IDs and passwords on their local servers, then a central organization at the university can manage that space in times of crisis. This makes the smooth functioning of the ID space less dependent on the uninterrupted availability of particular individuals.

- If departments can leverage a central ID space, then departmental system administration time will be reduced to the amount required manage only their local IDs and passwords. Departmental administrators are presumably in the best position to populate and manage locally-required entries as well as to verify the accuracy of centrally maintained data. Allowing local administrators to become more involved with data management should lead to an improvement in the quality and accuracy of data.

The university directory design is more complicated than commercial designs because it has different motivations. Because of this, it represents what we are calling a different market. It remains to be seen whether or not one directory will serve both the “business” and “academic” sectors of a campus like the University of Michigan’s.

As opposed to the use of roles data in directories in a commercial setting, user account publishing is used to provide the basis for authorization decisions. Data integration is less likely in an academic setting. “Data” that represents the intellectual production of a unit is unlikely to be moved from the unit that created it to a central system. Instead, controlling access to the data in a very federated and loosely coupled environment is of more interest.

4.2 Summary

The polarity between corporate and academic perspectives on directories combined with the specific experiences at the University of Michigan put the campus in a set of circumstances that are very nuanced and subtle at the same time they are very clear. The combination of increasing demand for a campus enterprise directory and the ability and willingness for units to craft their own solutions have led to a situation where technical, cultural and communications issues all need to be recognized and addressed as part of a directory services solution.

5 Directories on Campus

5.1 The Campus Directory Context

There has been a long history of directory development and use on campus. Some may remember *userdirectory on MTS. The campus also participated in a project called QUIPU. QUIPU was part of the ISODE, an openly available implementation of the upper layers of OSI. QUIPU provided an X.500 Directory System Agent as well as a number of Directory User Agents. This provided the basis for the X.500 service on campus that predated LDAP. The insights and experiences that were gained by these earlier endeavors directly informed and inspired U–M’s crucial role in LDAP’s initial design.¹

As a result of this work, we enjoyed an early advantage in terms of directory-related expertise and provided leadership in the field. In some areas related to directories, we still may. In others, we have been matched or surpassed. Conversion and transition issues that we face as a site with a large installed directory service are likely more difficult than startup issues that a new installation faces in its initial deployment. This is heightened by the fact that higher education does not “own the table” as much as in years past and that much of the work in directories is being done in the private sector.

5.2 Directories on Campus

Understanding U–M’s current directory capabilities begins with a synopsis of existing directory services on campus. There are a plethora of directories on campus, both large and small. They may appear as databases, spreadsheets or flat files. The address book that supports a desktop e-mail client is a directory. Small directories are primarily serving the role of application directories. Each one of these small directories comes with its own update and access procedures and policies and contribute significantly to “stovepiping” of information technology infrastructure services and the complexity and expense of its operation.

Some larger and more significant directories on campus include:

- UMOD, the University of Michigan Online Directory. UMOD is the general purpose campus-wide LDAP directory and supports e-mail groups and e-mail forwarding for users and groups @umich.edu.

¹http://wp.netscape.com/columns/techvision/innovators_th.html is an interview with Tim Howes, credited with being the author of LDAP, where he describes the development and background of LDAP, including its roots at U–M.

- UMIAC, the University of Michigan Infrastructure for Academic Computing;
- MIRLYN, the Michigan Research Library Network;
- ABS, ITCS's Account and Billing System;
- Uniqname, the campus uniqname system;
- IDCard, the campus ID Card System;
- Novell's eDirectory which is run by both the MCIT and ITCS; and
- Microsoft's Active Directory, that support servers, desktops and users in the Windows 2000 and later environments and is run by several units on campus.

All of these directories recapitulate and echo the independent nature of campus units and add to problems inherent in such an environment with such a multiplicity of directories.

- They duplicate feeds and other services from authoritative data sources which in turn add cost to the development and administration of data provisioning processes. (A high-level assessment of custom interfaces to and from MAIS directory-type information revealed a list of over 50 such items. The list continues to grow.)
- There is a lack of clear differentiation in terms of services provided across these directories. Several, if not all of them, implement their own authorization schemes and provide their own roles data instead of sharing them.
- In addition to their application-specific functions, several, if not all of them, implement functionality that could and should be provided by a general purpose directory. Instead of each directory downloading its own copy of data from MAIS, for example, commonly used data could be accessed via LDAP from the UMOD. Examples of this could include class lists or the type of affiliation one has with the University.
- Institutional-level goals that have directory-related requirements, such as a campus portal, that are beyond the scope of the application requirements of any one of these services are not provided at all.

- Some are burdened by designs that reflect now bygone or at least obsolete business processes.
- They employ a range of access methods and storage mechanisms which increases overall cost and complexity to campus.
- To the degree any of them support application and authentication services, they do not provide them in a way that other directories could make use of common services.

Each of these directories, then, face the same issues any directory provider would face. Their independence and less than ideal coordination make optimal outcomes unlikely.

5.3 UMOD as a Point of Integration?

UMOD is the incumbent general purpose LDAP directory with campus-wide reach. A more effective use of UMOD in the short term could possibly stem or reverse the proliferation of local directories. Barriers to a more effective use of UMOD on campus include:

- the lack of a shared sense of how schema changes are proposed or undertaken,
- the lack of awareness of what UMOD-based options may be available to units. These could include the publication of course lists or the development of “yellow pages” or other departmental directories.
- the lack of a clear and widely disseminated statement of UMOD’s capabilities and limitations in providing the desired services.
- the lack of awareness and expertise in academic and smaller administrative units related to developing feeds to UMOD.
- the lack of clearly identified developers’ time for UMOD.
- the lack of a commonly developed and shared direction for UMOD on campus.
- the lack of a campus-wide directory oversight committee for policy and process related to UMOD.

Whatever the effect of immediate improvements that using UMOD in a more encompassing way may provide, this working group is not optimistic that UMOD and its ancillary support and infrastructure as currently implemented will in itself fulfill the role of an enterprise directory for campus. This leads the group to propose the steps it lists in section 7.

5.3.1 A Perspective on UMOD in a Changing Directory Environment

This document's editor has been asked both why he is biased against UMOD and why he is an apologist for it. Comments like that are framed by issues that go beyond the quality or use of a particular LDAP implementation. They do indicate the visibility of directory services that extend more and more beyond providing data via an LDAP directory as well. This section of the document is a reflection on UMOD in the context of possible changes in the campus's directory architecture and operation.

UMOD is a collection of processes and people and the current state of a multi-year experience of U–M's involvement in a campus-wide general purpose directory. It is presented via a particular implementation of LDAP with some of the localisms of process and history instantiated in code and data organization. Its implementation as an LDAP directory is very closely coupled with other key pieces of campus information technology. In the context of a changing campus directory environment one may reflect on UMOD's strategic and operational standing on campus.

The community associated with UMOD on campus will of necessity be involved in and affected by changes, as will its community of users. The assessments and conclusions in this report are not driven by narrow issues related to OpenLDAP and UMOD such as protocol compliance or by its ability to fulfill its current role as campus LDAP service provider. They are driven by the requirement to revisit and reassess where the campus is relative to its mission and requirements in the context of directories and to examine current and future technology choices that may fulfill the campus mission and requirements.

As it is currently implemented, UMOD has several strengths and advantages. It allows for self service operations, where people can create their own groups, modify group and personal entries, and so on. It also provides critical address book functionality for campus. An easy to use browser interface is available, as well as a more traditional command line oriented command. UMOD has strong

and well known relationships with campus user-related data and its providers. It is tightly integrated with sources and processes for maintaining user data on campus, such as uniqueness and kerberos. It also has the advantage of incumbency. People are familiar with using it and both formal and informal support networks have sprung up around it over the years.

UMOD also has some weaknesses or drawbacks. It uses some non-standard and localized schemas. Its focus on kerberos is both a strength and weakness. There is a lack of ease of use with off the shelf applications and desktop operating environments. Vendor leverage and commitment is lacking. There is an OpenLDAP community but there is no OpenLDAP vendor. As a result, the third-party market will be smaller than those of other directories. The requirement to provide staff time for its development is probably a resource constraint for UMOD developers. New work must contend with resource allocation and compatibility issues that are required to keep the current environment functional. Some of these compatibility issues involve other units on campus.

It is premature at this point to define UMOD's future role on campus or its relation to other directory-related products that may be used on campus. One scenario may be that it maintains its position as campus addressbook for mail routing by the umich.edu gateways and provides a source for user account information, like uniqueness, user and group IDs, and affiliation while new functionality is located in other systems. Another scenario is that it provides self-service functionality across a range of data, including departmental directories, network asset descriptions, and so on, with background data synchronization with other data sources on campus. Or, it may become a niche system as another directory achieves dominance on campus. It could well be that the comprehensive or "public" LDAP directory's role on campus will evolve to a given point regardless of whether that directory is based on OpenLDAP or some other product.

In order to thrive in the new directory future on campus, UMOD must emphasize its self service orientation, maintain its relationship with data providers and integration with campus process ease, and also adapt itself to the role of one of several peer directory services, and not necessarily the lead or primary directory service on campus.

Other directory services, such as Novell's and Microsoft's should both have the same access to campus data sources as does UMOD. UMOD's strategic advantage on campus will not be its unique data holdings or collection. A risk to UMOD is that other directories, such as Microsoft's Active Directory, will become the path of least resistance for those whose priorities include off the shelf compatibility with third-party applications and desktop operating environments.

Expecting vendors to always accommodate their products to a localized environment seems unlikely and requiring that compliance could well prove to be unrealistic as a vendor and product selection criterium.

In order to thrive, UMOD should exploit and maximize its advantages and address its shortcomings. It should prepare to be an effective and useful resource and tool for people in an era where data and functionality can be assumed to be available in other environments, and where “plug and play” and Web Services-based approaches become more likely.

5.3.2 UMOD, ITCS, and the Future Directory Environment

UMOD is both a campus-wide resource with strategic implications and a set of operational responsibilities and commitments for ITCS. A major portion of the planning and operational responsibilities for Novell’s and Microsoft’s directories are also located in ITCS.. As we proceed with the next steps in the campus directory environment, ITCS will also have to monitor its relationship with the directory working group and project leader. On the one hand ITCS is more strongly attuned to getting feedback and having decisions informed by the campus community. However it still has responsibility to express its own direction and leadership position to customers, announce its own strategic directions and operational directions and capacity assessments to its user community. A lack of clarity in terms of how and what resources are committed to UMOD, Active Directory, Novell, and other application directories that may exist on campus makes it difficult for constituencies on campus to allocate their efforts and resources.

5.4 Overall Problem Areas and Issues with Directories on Campus

There are some accomplishments to keep in mind with the current collection of directories on campus. After all, E-mail is being delivered, books are being checked out of libraries, computer accounts are being provisioned, M-Cards are used in some areas and day to day life on campus is proceeding. However, we should not overly congratulate ourselves on this state of affairs. We are working too hard and too inefficiently to provision what we do have and are harmed by what we lack and what should be achievable here.

The current directory structure on campus is a collection of non-orthogonal, sometimes redundant, sometimes overlapping, uncoordinated directory services

that each address some relatively narrow and focused requirement. To the degree that they interoperate, they depend on ad hoc and replicative services and operational practices. Many are inflexible and do not lend themselves to mutual improvement or leverage across the institution.

None of them, either singly nor collectively, provide the campus with what could be considered as an “enterprise directory” as that concept is commonly understood today. The most pronounced shortcoming is in the area of identity management, which is critical in terms of an academic and research directory. Questions of membership, summarized as “who’s in and who’s out?” of different groups, are difficult to deal with as a result. The meaning of one’s identity in different situations or locales and its interaction with roles data for authorization decisions is nearly impossible to integrate. The working group feels strongly that a fundamental re-examination and change in approach to directory services on campus is called for. This change should bring us closer to having an enterprise directory on campus that can address the requirements of both the corporate and administrative side of the University as well as its academic and research side.

5.5 The Direction of Change

The impetus for this change comes from several quarters. The next generation campus directory must be more extensible and serve a larger set of application needs across the campus. Directory technology and the range of supported applications have advanced since UMOD was first envisioned and we have new options available to us. Better and more effective integration with business processes and requirements on campus is needed. The next generation enterprise directory for the University of Michigan is envisioned as a central directory with rich attributes that allows for this integration of services across campus.

- More accurate and useful data from a greater number of sources will be presented via LDAP to campus. The directory will be viewed by some users as a central data store. Although a number of complex issues must be dealt with in order to develop this common data store, the presentation of the data via LDAP means that these complexities are hidden from consumers of the data. Data stewardship will be totally integrated into this process.
- Mainstream directory-enabled applications will work with a minimum of confusion or complexity presented to the desktop or browser. Directory applications will be developed by central and distributed units as needed by the

U–M community or by individual units as well as for for inter-institutional collaborations.

- Identity and roles attributes in the directory must be those that support the directory applications. These include many of the attributes contained in the current UMOD, but also include many additional attributes needed for more sophisticated uses of the enterprise directory. Some attributes will be pre-computed and stored in the directory. Other data will be available for applications that need to provide their own business logic and compute attributes “on the fly.” Schemas required to support applications of import will be provided. (This type of activity related to envisioning the next generation enterprise directory for campus is clearly related to Phase 3 activities, and indicates how blurred the boundary between phases may be.)

5.6 Conclusion: Moving Towards an Enterprise Directory for Campus

A broad statement of our conclusion is that we should move from the relatively limited, but still important “white pages” campus-wide functionality that nearly all people on campus view as the directory to a much richer and more application-oriented enterprise-wide directory service with accompanying directory-enabled applications. We feel that the technical issues surrounding the directory, such as protocol compliance, support of various authentication methods and the ability to manage the directory’s contents present less of a problem area or barrier than do issues of governance, developing a shared vision of what the enterprise directory should be and creating multi-team commitments that will be required to address shortcomings in the campus’s overall directory environment. However, even in the technology area, we need a fresh look at products and how those products are used.

The staff who work on large directories in their various incarnations on campus clearly have, in our view, the skills and insights to advance the directory environment to desired levels. It is not clear that there is a supported and comprehensive vision for directory services on campus to guide them.

There are a number of issues that need to be worked through and better understood in order to define a path for directory development on campus. The working group distilled the information and findings we developed into the following broad categories.

- Providing an *identity management* system.
- Creating and using *roles* data as a new feature in the directory and increasing the directory's capability and usefulness for authorization.
- Improving and expanding *feeds* for getting information into and out of the UMOD directory and providing a basis for better understanding of data ownership and usage rules.
- Developing *governance* frameworks on campus. As the directory is used across more administrative domains, forums for making decisions about topics such as schemas, data organization, privacy and role definition need to be created. (A schema defines what types of data can be stored in the directory and in what format or syntax.)
- Increasing the *impact* and *effectiveness* of larger directories on campus and better allocating functionality between a shared general purpose directory and large application directories. If this is successful, the number of directories and directory-type interfaces should diminish on campus. Leveraging and expanding the use of UMOD will allow application directories to be more effectively used by the units that support them.
- Achieving *coherence* and *alignment* with other projects and products. This includes multi-site efforts like Shibboleth, on-campus efforts like Course-tools.NG, and the use of commercial or free products like portals and calendaring systems. Both short term operational needs and longer term strategic requirements must be addressed. The directory clearly is a key component in the success of future initiatives on campus and must work well with the anticipated web-based technology of the future. Progress in this area needs to be driven by a "demand pull" from projects that require stronger directory services than are currently available centrally and by sponsors who prefer to avoid developing their own local solutions.

6 Recommendations

The directory services infrastructure working group has developed a set of recommendations for making progress in the directory services area on campus. We also propose a process for implementing these recommendations.

6.1 Identity Management and Roles Data

These two areas are addressed together because they are similar in their importance for future directions and very interrelated in terms of functionality. Our contention is that identity management is the most important role that an enterprise directory will fulfill in a campus setting. These are the recommendations in the areas of identity management and roles data.

1. Create one or more working group(s) to address identity management and roles data questions and issues. The structure of the group(s) should be defined to complement both the more immediate and longer term needs in this area. Identity management should be explicitly recognized as being a new area that needs attention and consideration as an important and ongoing part of enterprise directory support. Identity management should be defined in terms of current practice in industry as well as higher education. Formally recognized and centrally maintained roles and identity management data are new enough in terms of implementation and concept and have sufficient administrative issues and nuances that there needs to be (a) group(s) specifically charged with addressing them. Roles and identity management should be thought of as a core service that other applications will use, much like LDAP itself or DNS.
2. Review existing directory services on campus with the goal of incorporating their identity management and roles requirements to the extent reasonably possible in the future enterprise directory. General purpose account maintenance should be more closely integrated in the directory and moved from stand-alone applications, especially those that are totally specific and unique to this campus. Any other tools or processes that result in the creation of uniqnames and directory entries should be reviewed as well.
3. Compute and include common and easily accessible role information that will be of sufficient use to units on campus, such as common attributes (enrollment status, type of appointment, and so on). Review various definitions of common role concepts, such as faculty appointments, in various units across campus.
4. Document the life-cycle of the directory data (sources, additions, updates and purges) as opposed to relying on standard procedures to maintain data integrity.

5. Include role-oriented enrollment data in the enterprise directory to support class lists and related authorization and access decisions (course reserves, mailing lists, coursetools, and the like) both within the campus and inter-institutionally.

6.2 Feeds

These are the recommendations related to data feeds and the directory.

1. All user or people entries in the enterprise directory should be represented in a feed. Directory entries that are not associated with a feed should be purged from the directory. “Mystery” directory entries (those that are marked as “keepers” but without a clear affiliation) should be purged or checked and represented in a feed.
2. Use the feed process to help explicitly define the meaning of being associated “@umich.edu” and to understand the roles affiliated with directory members. This will help address a fundamental question of “Who’s in and who’s out” of the enterprise directory and who can add or update “external” or “sponsored” affiliate entries to the directory.
3. Allow anyone with a meaningful affiliation with the University to be represented in the directory, with appropriate roles and authorizations. For example, schools and colleges should be able to create and maintain directory entries for people affiliated with them but not elsewhere in the University. This will help reduce the number of competing name spaces on campus by providing a robust central name space.
4. Refine, more fully architect and implement our current concept of inter-directory connectors. Evaluate products that may simplify the feed process. “Bi-directional” updates that comply with data primacy rules and that do not duplicate existing functionality should be supported, so that units may more easily and efficiently correct data errors they discover or make changes and updates that they require. We should make a special effort to learn more about how this problem is dealt with at other sites.
5. The directory maintainers should provide whatever help or advice is needed to enable all units to develop and maintain feeds of their data to a central directory.

6. All sources of directory data should participate in this effort, including the Flint and Dearborn campuses.

6.3 Governance

These are the recommendations in the area of governance.

1. Convene a schema and data organization governance and oversight group with a technical and operational orientation. This would be an informed group that can begin to work through the technical and operational issues and impacts of making schema and other data organization changes in the directory and make policy recommendations when necessary. Among others, we propose this group would be able to answer a very frequently asked question, “Can we put *X* or *Y* in the directory?”

6.4 Leverage

None of the directories on campus, including those listed in section 5.2, are providing as much value or being as well leveraged as they could be. Those who provide directory services on campus should facilitate their use in other units and across campus. Directories should be used as much as possible. Even if a particular project is not able to be fully served by a central directory on campus, it should minimize its own directory development and use. Units on campus should be able to take advantage of as many centrally provisioned directory services as possible and move towards minimizing the number of isolated application and other directories to the degree possible.

This is the recommendation in the area of better leveraging directories on campus.

1. More work needs to be done in the area of soliciting and canvassing for applications on campus that could better use centrally provided directory services. Directory providers need to respond to the demand on campus in a fairly cohesive, globally optimized fashion. More information is required for this to be possible. A steering committee should be formed that will help develop, vet and prioritize work related to enhancing the functionality of current directory services. This committee will solicit input from campus regarding desired directory services. At the same time, it will be informed by directory maintainers about what facilities are available or the level of effort that would be required to achieve requested facilities. A general goal of

accommodating third party applications and allowing units self-service access to directory functionality would presumably be desirable. Improvements that are to a large degree “technology neutral” — that is, would be useful for different directory implementations — would be especially valuable.

6.5 Alignment

This is the recommendation in the area of aligning directory work with other projects on campus and elsewhere.

1. Being more consonant with national and international efforts in the academic and research worlds as well as with industry directions. These efforts should be aligned and prioritized with on-campus demands.

7 Next Steps and Moving Ahead

These recommendations cover a lot of ground and imply work with different timelines and goals. Implementing the recommendations cited in the preceding section or selecting projects to drive further progress will be a key component in advancing the state of directories on campus.

The directory services environment on campus reflects our multifaceted and sometimes fractured overall environment. Improving the delivery of directory services, arriving at an overall architecture and developing a cohesive plan will all be challenges. Even developing a robust context to talk about these issues is not a given. Without a robust commitment to this activity, we will find it to be too easy to fall back on current habits when sticking points are reached.

7.1 Pilots and “Low Hanging Fruit”

Some of the recommendations call for at least conceptually straightforward actions with a fairly well defined deliverable. To the degree they require changes in internal work flow, policy or process progress made in these areas could be fairly portable across different directory technologies or implementations. They are also the same type of projects that we are pursuing now across campus. They tend to represent incremental changes.

Some specific projects or pilots may be defined and initiated in order to drive progress on the recommendations in this document. They should be non-trivial

but also not so large that they can not realistically be accomplished or developed far enough to yield useful results. Although the main effort recommended by this group is a single large project to scope and define an enterprise directory, developing and extending “shared capabilities” where possible and where they fit into unit requirements and missions should not be ignored.

An illustrative list of potential projects could include:

- Developing rules for populating current directories and adjusting existing feeds to conform to these rules. UMOD entries that are not represented in a feed should be cleaned up or purged.
- Establishing new feeds of local data from a unit in order to alleviate or remove the unit’s requirement for a local directory server, for example to maintain data for people not currently in UMOD or other campus-wide directories but whose entries are required for local business.
- Supporting departmental and faculty / researcher directories.
- Allowing units to use directories for authentication and authorization processes where appropriate and advantageous.

Work spent on projects in this category represent effort not spent on other recommendations, and the steering committee charged with defining and prioritizing these projects will need to be very aware of this.

7.2 Other Recommendations

Some recommendations envision a new environment that is desired. Implementations of these recommendations would be characterized by significant decisions about technology and vendor choices that will have a dramatic impact on the campus’s directory environment. Following these recommendations would be described as defining the campus’s next generation enterprise directory. Changes in this area would be more significant and far reaching and could change the fundamental technology base we are currently using. Hence, product evaluations and transitions could be required.

Achieving these recommendations would involve activities including:

- Determining the purpose and function of a new enterprise directory.
- Evaluating directory software for the new enterprise directory.

- Evaluating ancillary software that will interoperate with the directory software.
- Identifying the operational and other infrastructure issues (which OS to use, how to provision the supporting hardware, storage requirements, and so on) in a way that is presumably consonant with a larger direction on campus.
- Managing the transition to the new directory environment.
- Assessing the priority and role of NMI-related work relative to narrower on-campus directory-related issues.
- Forming a consensus on appropriate boundaries between “application” and “general purpose” directories and their use at a tactical level.

Other recommendations, especially in areas of identity management and roles data, relate more to fundamental design decisions and goals that reflect the institution’s requirements at abstract and strategic levels. Questions about who is in and who is out, what it means to be in various name or identity spaces, and defining the related policy questions will need to be addressed.

Achieving these recommendations would involve activities including:

- Developing definitions of the standard roles that may be calculated and stored in the directory.
- Articulating the meaning and use of “identity management” on campus.
- Identifying or defining additional roles that, once populated, could be retrieved from a general-purpose directory.
- Defining how improved feeds could accommodate the campus-wide need to collect roles information for authorization decisions.
- Forming a consensus on the location of role and identity management data between “application” and “general purpose” directories.
- Defining the correct fit between authentication and authorization services and directory services on campus.

7.3 An Enterprise Directory Project Definition

In order to accomplish the overall goals related to campus-wide directory services, we propose that the University undertake a large-scale project to implement the enterprise directory called for in this document. The project must include all the aspects of a comprehensive enterprise directory listed in this document, especially identity management and roles data and others listed in section 6. While most people will likely refer to this project as “the Enterprise Directory Project” it is important to recognize that the implementation of the directory itself is only a small part of the overall project. This project will be marked by the following directions and assumptions:

- The project must be designed to gather input from the appropriate campus constituencies. Some of this input will come through IT Commons groups such as the Stewards and the Directory Working Group, but other processes for gathering input from various campus units and business offices will also need to be developed.
- The project is a major effort for the University, and should be run using effective project management techniques, to assure that progress is appropriate, decision-making is structured and effective, and that appropriate two-way communication with campus constituencies occurs.
- University units will need to provide subject area experts on a part-time basis to support the project, but the project team itself should be made up only of staff with full-time assignments to the project. Full-time assignments will assure that priority conflicts do not put the project at risk.
- Determining whether authentication and authorization will be “in” or “out” as part of an enterprise directory implementation for campus.

No incrementalist approaches, changes based on effort “on the margins,” and localized standalone projects dedicated to satisfying narrow requirements will yield a satisfactory solution to campus-wide directory problems.

7.3.1 Project Structure and Planning

We recommend that the IT Commons Stewards develop a charge for this project. A full-time project manager should be identified for the project. The project manager will be accountable for all aspects of the project including deliverables, budgets, staffing, methodology and organizational structure.

We anticipate that the first phase of the project will likely take several months, and will consist of clearly defining the rest of the project. This will include gathering requirements from campus constituencies, defining scope, identifying resource needs and funding requirements, developing a high-level and detailed project plan and determining commitment of campus constituencies to leverage a directory in the future.

At the conclusion of this phase, there will be an evaluation and assessment of the commitment on campus to proceed. The cost and impact of the project will be considered and a decision made about going forward.

The following phase(s) of the project will likely involve several parallel activities. For example, it is likely that at one aspect of the project will be to select and implement directory software, and as this work proceeds, other teams may be working on identify management and roles work. The exact structure of these subprojects, how they are staffed, and how they will be coordinated will be the responsibility of the project manager.

In order to begin this project, two or three full-time staff members, including the project manager, will be need to be in place. Staffing requirements for the remaining aspects of the project will be identified during the first phase.

A Active Directory and UMOD

The problems of coordinating two enterprise directories as well supported and maintained on campus as UMOD and AD will also provide a sense of how difficult it will be to support any comprehensive enterprise directory as well as small directories that are not nearly as well supported on campus. The following discussion covers some Active Directory administrative issues that are currently being addressed by the AD Central Accounts pilot project. The AD Central Accounts project empowers AD departmental administrators to provision selected AD faculty and staff user accounts for Windows according to their organizational needs, while still maintaining the integrity of the AD user accounts across other places in the Windows forest environment, such as the U-M Campus Computing Sites and University Library.

To reflect the flat structure of the U-M Online Directory, and to facilitate easy updates from the U-M Directory to AD, all users in the U-M Directory that possess a unique name are channeled into a single organizational unit (OU) in the UMROOT domain, called the "People OU." The People OU contains over 200,000 user entries, and those entries represent just about every U-M affiliated individual that might want to use U-M computing facilities. These user entries are automatically created and continuously updated, based upon corresponding entries in the U-M Online Directory. Users in the People OU of the UMROOT domain are able to authenticate with their U-M unique name and passwords, using Kerberos pass-through authentication.

AD Group Policy and OUs

In AD, Group Policy is applied at the OU level to administer Windows 2000 resources such as users, computers and printers. A W2K administrator might typically want to set a user's home directory, set a user's profile path, set a user's Windows password, and also enforce Group Policies to customize the user's environment. Unfortunately, for several reasons, none of this is possible for users located in the AD People OU. First, since all of the users are contained in a single OU, it would be imprudent to assign multiple administrators control over the same OU. Second, moving AD user objects to department OUs and assigning administrative rights to one department creates the potential for conflicts of interest.

When a Windows user logs onto a computer residing in a Windows domain, a complex set of actions take place to configure the user's Windows environment. If we give a delegated OU administrator too much control over a user's account, the

user may encounter unexpected or inconsistent behavior the various locations they may be using, such as the U–M Computing Sites or departmental labs. Arriving at a solution that is satisfactory for all potential logon locations is achieved by selectively restricting certain user object attributes for users that will be moved from the generic People OU to a specific “delegated” OU. In other words, the delegated OU administrator trades a few administrative capabilities for the right to administer a user of interest to their organization.

Many users function in multiple roles, so a primary administrator would need to be assigned, and the potential for turf conflicts and confusion over user roles could be daunting given the large user population.

Here is an example of potential conflict of interest among several Windows administrators, each in a separate organization, but wishing to administer a common Windows account associated with a particular individual, the ubiquitous Babs Jensen.

Perhaps Babs is a graduate student in Astronomy, but is also working as a staff member in the Chemistry department. Her academic and work roles require that she uses computers in three different locations. Those locations are Chemistry (work role), Astronomy (student role), U–M Computing Sites (student role). The administrator of the Chemistry department decides that for Babs to perform her duties, her Windows account must be granted disk space, a profile path, and other Windows specific attributes that are associated with permanent physical resources. Babs’s Windows user account is moved to the Chemistry OU, so the Chemistry administrator is able to set the appropriate attributes of Babs’s Windows account. Later in the week, the Astronomy department decides that all Astronomy students should be granted customized Windows resources, including disk space and a profile path. Since Babs’s Windows account (bjensen) is already a member of the Chemistry OU, she can’t simultaneously be a member of the Astronomy OU. The administrators of the Astronomy OU will either have to convince Chemistry to give up their primary administrative rights to the bjensen account, find a workaround solution or forget about provisioning the bjensen account with Astronomy specific resources. Meanwhile, Babs has dropped by the computing labs at Angel Hall to work on a term paper. She logs on to a Windows workstation, but receives quirky and not-so-funny error messages. The word processing program that she normally uses at the U–M Computing Sites is not available, but she observes that some of the programs that normally would appear when logging on at Chemistry are available.

The question of accurate role data needed to make authorization decisions is a key issue in this scenario. This example assumes that information regarding

Babs's roles in the various departments and areas of interest to her is available. In an ideal Windows administrative world, each Windows user account would be automatically assigned to a primary administrative organization. If the UMOD possesses accurate role data, the process of creating and associating resources with a user's Windows account could be automated. In lieu of a central provisioning mechanism for Windows accounts, and of accurate role data that could automate the provisioning of resources by other assigned organizations, the procedure for Windows account provisioning assumes a manual intervention component. If the numbers of user's requesting special Windows account provisioning remain small, the manual process may work for some time, but as requests for provisioning increase, accurate role data will enable more efficient and streamlined handling of Windows account provisioning tasks at the central campus level.

This fictitious example illustrates the need for strict adherence to commonly agreed upon conventions for the administration of Windows user accounts. The ITCS LNGS group believes that the potential pitfalls of Windows user administration illustrated in the preceding example can be avoided by following a set of administrative procedures that place a set of well-defined restrictions on Windows accounts, but still allow the user to access customized resources specific to a particular logon environment.

Impact on Campus

For reasons including those mentioned above, some organizations on campus have felt that they had no alternative but to create a separate W2K domain within the U-M forest. A W2K domain is security boundary, with both a separate Kerberos realm and a separate user principal namespace. In a Windows 2000 forest, all domains are connected via a two-way transitive trust relationship, so security interoperability between W2k domains is less of an issue than the separate user namespace issue. Proliferation of W2k domains means that a user can be assigned separate unqiename-based accounts, one per domain, each having a separate Windows password. Additionally, each additional domain creates more AD replication traffic, and increases the potential security exposure of the entire W2k forest. Trust relationships in the campus-wide AD environment may expose some units to security or operational lapses that occur in other units.

One other pitfall to W2k domain proliferation is that users in down-level domains cannot take advantage of Kerberos pass-through authentication using their down-level user account. The pass-through authentication process always maps the user to their UMROOT domain account. Thus requests for single-signon be-

havior at down-level domains cannot be fulfilled under the current operating conditions, at least without resorting to custom GINA software. Modifying the Windows GINA logon is tricky, and difficult to administer in all but tightly controlled environments. The College of Engineering is currently using a customized Windows GINA.

In summary, our difficulties to date in solving the user administration problem has curtailed administrative capabilities for a number of W2K OU administrators, and has contributed to the proliferation of domains in the U–M Windows 2000 forest, with the attendant complexities of multiple namespaces. However, the ITCS LNCS group is currently working on an “AD Central Accounts” pilot project which is intended to allow the movement of certain AD users from the central “People OU” to individually managed administrative OUs without breaking existing functionality within AD. If this project is successful (and preliminary results are very promising), both administrators and end users will benefit and the problems described above will be significantly ameliorated. As with many other projects with a directory focus, the availability of timely and accurate user role data will be key to automating AD Central Accounts procedures.

B Roles, Eligibility and Authentication

Who can check books out of the library? Who is eligible for a CAEN account? Is Babs Jensen a student? Or a faculty member? Who can use dial in or print services from the basic computing package? The answers to these questions depend on the role of the person related to the service or data for which they are requesting access. From a given client’s perspective, should roles be computed or retrieved? Should the data that defines roles be stored centrally or locally? How are roles like being a member of the faculty defined? Who is vouching for the quality or the authoritative nature of the data? There are answers to many of these kinds of questions, but they are spread among many locations and in systems which require a variety of access mechanisms.

Roles provide a grouping mechanism to represent what may be temporary and dynamic organizational structures. Roles are designed to be more efficient and easier to use for applications than group lists. For example, applications can locate the roles of an entry, rather than select a group and browse the members list. Roles provide information to help answer the basic question of “who can do what with which resource?”

A roles database is a mechanism used to assign a user access to data or appli-

cations. Permissions groups and roles related to centrally maintained data need to have centrally-administered controls because they define organizational security policy. Yet, associating users with groups and roles should be easier to do from the field until since local administrators are familiar with employees and their functions.

Roles data for an enterprise should be hosted centrally, and made available to remote applications as needed.

Roles are a set of criteria used as a basis for authorization decisions. Roles data can also be used in other ways, such as reporting and extracting data. At this time, there are effectively no roles-based authorization decisions made based on UMOD data. They are made based on other application directories, or services that support authorization based on a role attribute, such as AFS's PTS and ABS's determination of who is eligible for a basic computing services package.

B.1 Identity

Network Identity is the context-sensitive identity, attributes, rights, and entitlements all maintained within a policy-based trusted network. Managing Network Identity describes the technology infrastructure and business processes for managing the lifecycle and usage of an identity, including those attributes, rights and entitlements.

The Authoritative Directory serves as an integrated repository for storing and managing identity profiles, as well as application and network resource information. The directory stores user profile information including user credentials, such as public key certificates, passwords, or pin numbers. Applications and services retrieve identity information from the authoritative directory. Meta directories are an implementation of authoritative directories. Meta directories can synchronize and normalize identity information distributed across multiple enterprise applications such as customer databases, human resources applications, network operating systems, etc.

Network Identity (NI) is the context-sensitive identity, attributes, rights, and entitlements, all maintained within a policy-based trusted network framework. Managing Network Identity describes the software infrastructure and business processes for managing the life cycle and usage of an identity, including those attributes, rights, and entitlements.

Excluding mechanisms provided by some application directories, the most common authorization method on campus is based solely on successful authentication! This has forced groups on campus that require a more finely-grained au-

thorization to run their own distinct authentication infrastructures so that they in turn may make their authorization decisions based on authentication against (for example) their own Kerberos realm. Neither of these methods leverages roles-based data from a central directory.

Although developing and publishing roles is important, reasonable and informed boundaries must be applied to roles that are computed and stored in the central directory. There are many data elements used on campus, and we will be quickly overwhelmed if our goal is to represent “all” roles data and their relationships in the central directory. The focus should be on providing the most commonly used and required roles.

The location of identity and roles data as well as the retrieval mechanism required to access it will impact the usefulness of the data. Availability of data and ease of selecting or merging the data with unit-specific processing should be a goal of roles development. A separate issue is ensuring appropriate use of roles attributes when applied or used in combination with other information that may not be stored in the general purpose directory or with business logic that is specific to the service provider.

B.2 Extensibility

Extensibility is an important concept in engaging units to use the centrally managed general purpose directories. Units want to be able to add people to a directory and represent their unit-specific association or affiliation. Authorization decisions are then made based on this information. When a unit’s or service provider’s clientele are not already affiliated with the University and therefore in the directory, units choose to run their own directory services. They are confronted with the need to support a full directory in order to provision the small additional group of people not represented in the general purpose directory, but more importantly the local directory provides the roles information that the unit requires to make authorization decisions.

Individual units need to be able to add users to a central directory, manage the entries, and administer services. However, it is important for both the individual unit and the campus overall to ensure that respective boundaries for local and central administrators are well defined and enforced.

B.3 Namespaces

As mentioned earlier in this document, authentication and authorization are effectively synonymous as deployed on campus today. Under this scenario, role-based authorization is accomplished by creating different namespaces. One's membership in a directory rather than the attributes of one's directory entry defines access. For example, the College of Engineering's permission and access granting scheme and account provisioning process generally depends on one's membership in its own Kerberos realm and AFS cell when deciding to grant access to resources, as opposed to using the UMICH.EDU realm and cell and role information from a central campus directory.

The most basic definition of a name space is a definable bounded area or location in which a collection of uniquely-assigned and managed identifiers can be successfully resolved to an address. Specific attributes may be bound to or associated with those identifiers via the address they resolve to. For example, in the e-mail name space, paul@umich.edu resolves to an address that locates mailbox information. In a Kerberos realm, paul@UMICH.EDU resolves to an address that locates some identity and authentication information.

In a directory namespace, names may resolve to addresses that provide identity, authentication and authorization services. Identifiers in the directory namespace may be a unique name or a distinguished name. (A distinguished name or *DN* is a globally unique name in the directory system.)

Integrating or sharing namespaces is only the first, albeit difficult step in name space consolidation. First addressing and then reversing the tendency towards multiple name spaces used as crude roles data will allow other significant work to be done with rules. An example of this would be merging engin.umich.edu into umich.edu. Altering applications to do automatic provisioning and perhaps more importantly de-provisioning is an ever growing task, as it needs to be repeated in each application as long as central role information is not available. The business of name space consolidation is now referred to in the industry by the term "Identity Management." Consultants in industry are doing analysis and gaining field experience using Identity Management techniques to build both a return on investment and a better user experience.

B.4 Directories and Authentication

The directory relates to authentication decisions in some fairly important and direct ways. The directory can support authentication by allowing one to bind to

their entry. There are several protocols that can be used to bind. The directory can use various back ends, such as Kerberos, via plugins in order to access authentication information. The directory may also store credentials to use in authentication decisions, such as x.509 certificates and thereby play a central role in PKI efforts on campus.

Although the use of SASL and Kerberos provide a strong and useful authentication method, most third-party vendors who provide “ldap-enabled” applications will use a “simple bind” approach. In that case, SSL is, or should be used to provide a defense against sniffing plaintext passwords as they are transmitted over the network.

C Feeds

A feed is the organization and delivery of raw data to a directory. In the future, we anticipate seeing this as being an XML-based technology. At this time, it is accomplished in a more ad-hoc and informal manner. Although roles and feeds are addressed separately in this report, one must remember that they are very closely related concepts. The role/feed/maintenance loop is a tight one.

Data in the UMOD is provided by several groups or organizations who are responsible for the authoritative data source. The UMOD itself is the authoritative source for very few person-related data elements, one of which is campus e-mail addressing information.

Data feeds present the opportunity to better define and formalize the relationship between groups and areas who have the authoritative view of data and stewardship over its use in the directory. One key aspect of this is to define roles. For example, one data source may provide a list of people who are alumni, with attributes they have collected or maintain. Another source may provide a list of people who are students, temporary employees, and so on. In other words, membership in a list is itself data that should be represented as a role in the directory.

Properly defined and standardized feeds will facilitate:

- developing rules about primacy or data precedence when an individual is represented in multiple feeds. Since data is fed to a common directory from multiple sources, data primacy should be well defined and negotiated based on University business rules. It is not acceptable to “flip-flop” data with subsequent feeds.
- purging of old data, assuming all data must be associated with a feed. Each

feed would have well-defined add and purge policies and rules as well as a known authorized source or steward.

- identification of a commonly recognized point of contact and authoritative source for each field in the directory, including privacy-related rules.
- maintaining quality of data in the directory.
- consistent improvement of operational processes, so that improvements applied to load processes will be universal for all feeds and that ad hoc idiosyncratic processes that may be error prone or difficult to maintain can be retired. A standard approach to feeds now will facilitate eventual use of “web based services” or other approaches based on XML and related technology.
- campus-wide adherence to sometimes difficult to understand policies or laws, such as FERPA, by allowing data feeds to be more consistently defined and more widely reviewed.

There are currently data feeds in place that provide data for UMOD to incorporate and make available via LDAP. These are unidirectional — updates from UMOD are not passed back upstream to the original data provider. Procedures not related to the feed process such as end-users submitting paper forms or using web-based forms are used to correct central, authoritative data. Data is incorporated into the directory via the application of locally developed scripts designed for that purpose. These scripts may check and modify the syntax of supplied data and apply data primacy rules.

This paragraph needs fixing and clarification. More thought and work needs to be done in the area of developing and operating a system that allows us to interactively transform a replication stream from our LDAP directory to fit the data model of application directories, such as Active Directory. The current feed model may prove incapable of effectively handling schema mismatches and “dirty” data. These issues, as well as other requirements such as documentation and design, may require more effort to solve using the current approach than is realistically available on campus. One perspective on this approach characterizes it as more likely being more useful for sharing data than for transition activities themselves.

If we view and are able to use metadirectory products as flexible transformation engines, rather than all encompassing repositories of data, they may prove to be attractive in our U–M environment. DirXML is being tested in a UMOD

to eDirectory transform process. This product is also being considered as a replacement for the existing UMOD to AD transform. DirXML (and other similar products) could also provide well-documented development environments, cross-platform use, support for multiple sources and targets and a single, deterministic, transformation rule set for both asynchronous and batch updates. Use of this, or similar products, could help define a transition path for use in moving local customizations to an XML paradigm.

In any event, XML-based technology will almost certainly play a dominant role in future data exchanges and an assessment of “build or buy” will need to be made in this area.

D Benchmarking

To be provided.

Document Revision History

version 1.4 distributed to the Directory Services Infrastructure Working Group and IT Commons Stewards for review.

version 1.4.1.4 edited based on feedback received so far (Feb 3, 2003). Infrastructure Working Group.