

**An  
Information Technology  
Architecture  
for  
Emory University**

**Directory Service  
Domain Architecture**

*Adopted by CIRT  
February 20, 2002*

Committee on  
Information Technology Architecture

March 6, 2002  
Version 2.4.6

**Directory Service Domain Task Force**

John Cyran (ITD and Chair)  
 Barbara Anderson (ITD)  
 Lance Basler (Emory Healthcare)  
 Parrish Brown (ITD)  
 Tim Brown (Business School)  
 Chris Camacho (Netcom)  
 Peter Day (ITD and Editor)  
 Alan Dobkin (ITD)  
 John Kyle Fenton (General Libraries)  
 David Johanning (School of Medicine)  
 Chang-Kwei Lin (Yerkes)  
 Weiming Lu (ITD)  
 Belinda Maaskant (School of Public Health)  
 Elmer Masters (Law Library)  
 John Pine (Facilities Management Division)

**DOCUMENT REVISION HISTORY**

<b>Release</b>	<b>Description</b>	<b>Date</b>
1.0	Insert summary, overview and principles	December 4, 2000
1.1	Updates based on Dec 5, 2000 meeting	December 11, 2000
1.2	Exec Summary, Overview, changes from Dec, 12 TF meeting	December 17, 2000
2.0	New template format, editing, updates, rewrite summary & overview	March 19, 2001
2.1	Benefit of standards, new configuration diagram	March 21, 2001
2.2	Overview diagram, categorize principles	March 25, 2001
2.2.1	Section numbers in numbers for figures & diagrams	March 27, 2001
2.3	ITA Feedback, Copy editing, §12 examples 2 & 7	April 2, 2001
2.4	§12: Increase clarity, correct terminology, align with current vision	April 3, 2001
2.4.1	Updated info on BigIP in §8.2	April 9, 2001
2.4.2	SSA, DD, RP feedback, x.500 terms. group memb. access control	May 14, 2001
2.4.3	DJ feedback plus formatting & copy edits especially in §§10,11,12	May 24, 2001
2.4.4	Added eDirectory, “used” status, and updated to §§7.1, 8.1	June 22, 2001
2.4.5	Changed status to “Ready for Adoption”	Sept. 19, 2001
2.4.6	Changed status to Adopted; added copyright & more bookmarks	March 6, 2002

## TABLE OF CONTENTS

1. Executive Summary.....	1-1
2. Introduction.....	2-1
3. Overview.....	3-1
4. IT Trends.....	4-1
5. Directory Service Principles.....	5-1
6. Technologies.....	6-1
7. Standards Compliance.....	7-1
8. Standard Products.....	8-1
9. Configurations.....	9-1
10. Considerations & Next Steps.....	10-1
11. Glossary.....	11-1
12. Appendix: Justification and Implications of the principles.....	12-1

## LIST OF FIGURES

Figure 2-1. Process to derive domain architectures.....	2-1
Figure 3-1. Directory service concept.....	3-3
Figure 9-1. Directory Service Implementation.....	9-2

## 1. Executive Summary

---

An Emory-wide directory service would provide a means to look up official information about Emory people, places and things that are of Emory-wide applicability and that Emory thinks authorized people or IT systems should be able to obtain at any time from anywhere on the Emory network or the Internet.

Although such information is typically available in other systems at Emory, determining which system it is in and obtaining access can be difficult. Even when access is made available, the software and method for getting to it from a desktop computer can vary with the information source. Also access might be limited to certain times or locations. Looking up the data through a single service could leverage economies of scale to provide the benefits of special features such as shared, continuous availability, Internet access to data, and very fast response without having to incur the cost of implementing these features on each of the systems that would otherwise have to be accessed to find the information.

Just as the Emory Campus Directory has sub-directories for units, faculty & staff, and students, and the public phone directory has white pages and yellow pages, the directory service would be able to provide access to information on multiple topics. Examples of the type of information that might be available via the directory service include:

- Official names and identifiers for Emory people, places, and things. Examples of things include Emory departments, schools, divisions, organizations, and courses.
- Information about Emory people (such as their status, email, phone numbers, web link to their CV, Emory-wide authentication information, student id, employee id, and other identifiers).
- Entries documenting the existence of sets of research data, the type of data, and a web link to a description of the data and who owns it.
- Information for cooperation with other universities such as to support sharing of resources.

To ensure that the directory service is used appropriately yet allow it to easily grow in size and add information about additional things, the directory's contents will need to be carefully managed. An Emory governance body that is inclusive and representative of the perspectives of all components of Emory will be needed to decide what can be included in the directory and fend off attempts by individuals and systems to use it for file storage or to store data of only local interest. For data that is allowed, personnel will be needed to identify a source that is an "authority" for that data in the sense that the source provides correct, valid answers. The preferred authority would be an existing source that already integrates and consolidates data from multiple authoritative sources, since that would reduce the number of sources and thus the complexity of obtaining the data. In the case of individually maintained information, the individual would be the source and provide the updates. Regardless of the source, each entry in the directory must have a process to get timely updates.

Although Emory already has the beginning of an Emory-wide directory service, the current system is limited in scope, capacity and availability. The directory service architecture in this document defines an infrastructure to move to a higher level of service by building an Emory-wide directory service that can grow and change quickly enough to meet Emory's needs while keeping down complexity and support costs. The document proposes that the service be based on a general purpose, standards-based directory, rather than other types of directories such as a Network Operating System-based directory or an application-specific directory. It also specifies technologies, standards, and products to achieve the needed capacity, availability, and accessibility and reduction in complexity and support costs. Finally, it lists a number of points for consideration and suggests some next steps.

## 2. Introduction

In Document 1 we agreed on a statement of what Emory wants to achieve, and how it will use an IT architecture to help reach its goals. At the end of that document, we arrived at a statement of the requirements our technical architecture should meet, e.g. “facilitate change in academic and administrative processes” and “provide a campus network that allows communication and exchange of information.”

In Document 2 we identified the design principles (called the Conceptual Architecture Principles or CAPs) we will use as we evolve our new IT environment, as well as the categories (called “architecture domains”) where we need to decide on policies, standards, etc. The following diagram illustrates the process of deriving the domain architectures.

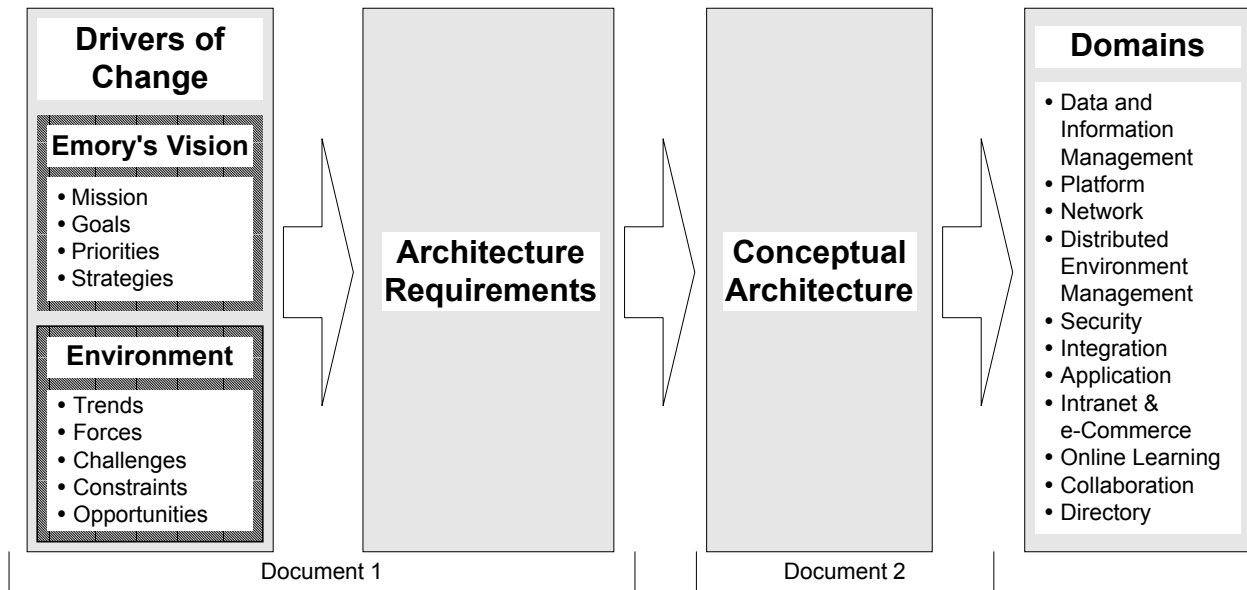


Figure 2-1. Process to derive domain architectures

This document is one of a set of documents—one for each domain—that take the next step and provides those policies, standards, etc. for each of the domains. In outline, this document:

- Provides an **Overview** describing this domain architecture;
- Points the reader to other **Related Domain Architectures**;
- Identifies the **IT Trends** relevant to this domain;
- Describes the **Design Principles** associated with this domain;
- Shows how the principles of this domain provide **Support for the Conceptual Architecture Principles**;
- Lists relevant **Configuration Principles**;
- Describes **Standards, Products and Configurations** for the domain;
- Points out **considerations** to address for this and other domains and suggests **next steps**;
- Provides a **Glossary** of terms;
- Shows in an Appendix the detailed **Justification and implications of the principles** so that those interested can understand in more detail the subtleties of meaning and the thinking behind each principle.

## 3. Overview

---

### 3.1 General

*“If you want to understand your government, ... read selected portions of the Washington telephone directory containing listings for all the organizations with titles beginning with the word `National” —George Will*

*“Knowledge is of two kinds: we know a subject ourselves, or we know where we can find information upon it.” —Samuel Johnson (1709 - 1784)*

*“If you don't find it in the index, look very carefully through the entire catalogue.”  
—Unknown, Sears, Roebuck, and Co. Consumer's Guide, 1897*

**Purpose.** The Emory-wide directory service would provide a means to look up official information about Emory-related people, places and things that are of Emory-wide applicability and that Emory thinks authorized people or IT systems should be able to obtain at any time from anywhere on the Emory network or the Internet.

Although such information is typically available in other systems at Emory, determining which system it is in and obtaining access can be difficult. Even when access is made available, the software and method for getting to it from a desktop computer can vary with the information source. Also access might be limited to certain times or locations.

**Benefits.** Having a service through which the data can be looked up has benefits such as the following:

- Storing identifiers from multiple campus systems with the people, places or things they identify enhances the ability to look them up and link together information about them.
- Providing infrastructure systems and other authorized systems with shared access to data eliminates the management cost and complexity of maintaining a copy of the information separately on each of those systems.
- Looking up data in the directory provides the following features without having to incur the cost of implementing them on each of the systems that would otherwise have to be accessed to find that data:
  - Continuous availability and Internet access to make information available for look up from anywhere on the Emory network or on the Internet at any time.
  - Very fast reads that are orders of magnitude higher performance than a typical database to provide real-time access to data in support of the operation of a wide variety of applications for a wide variety of purposes.
  - Extensibility to store any type of data, and easily change the entries, directory organization and access control while the service is running.
  - Ease of being widely distributed and replicated in multiple places.
  - Interoperation with other universities to support sharing of resources.

**Directory Data.** Just as the Emory Campus Directory has sub-directories for units, faculty & staff, and students, and the public phone directory has white pages and yellow pages, the directory service would be able to provide access to information on multiple topics. Examples of the type of information that might be available in the directory service include:

- Official names and identifiers for Emory people, places, and things. Examples of things are Emory departments, divisions, schools, organizations, classes, job titles, and security roles.

- Information about Emory people, such as their status, organizational unit, title, phone numbers, locations, preferred email address, web link to their personal home page, web link to their CV, their identifiers in various Emory enterprise systems, identifiers of courses in which they are enrolled, identifiers of organizations of which they are members, job role, and Emory-wide authentication information.
- Entries documenting the existence of sets of research data, the type of data, and web link to a description of the data and who owns it.

**Compared to other storage systems.** The wide scope of data that might be in the directory along with the ability to bring together data from disparate systems plus the possibility of providing personal data maintained by individuals raises the question of how a directory differs from other types of storage systems. To begin, a directory differs from a file system, because a file system is optimized to store data in chunks called files that can be very large, whereas a directory is optimized to store data in small pieces. A directory is different from a general-purpose database, because databases generally are intended to be able to record transactions potentially involving large volumes of data, whereas a directory is optimized for searching that can span multiple data items. A directory differs from a data warehouse, because a data warehouse uses a general-purpose database to store extracts and summaries from multiple operational systems to be used for analysis and reporting, whereas a directory is optimized to do a quick look up. Finally, a directory is different from an operational data store, because such a store is intended to provide a nearly current snapshot of data from operational systems for immediate reporting needs, whereas a directory is intended for data that does not change often.

**Directory database features.** So what makes a directory special and better suited to the envisioned use compared to other types of storage? A directory is designed for:

- **Fast reads:** A directory must provide real-time access to data in support of the operation of a wide variety of applications for a wide variety of purposes. Thus it must provide orders of magnitude higher performance to reads than a typical database, even at the expense of speed of response to write requests.
- **Extensibility:** A directory can store any type of data, and the entries, directory organization and access control can be easily changed, even while the directory service is running.
- **Distribution scale:** The data can be easily distributed to different directories, and entries in one directory can refer a request to another directory.
- **Replication scale:** Copies of a directory or parts of it can be automatically maintained in multiple places to provide higher reliability, availability and performance.
- **Standards-based access:** A directory supports an Internet standard protocol that allows for access from anywhere on the Internet at any time.

**Types of directories.** In addition to other types of storage systems, Emory also has a number of types of directories in operation. Thus the question arises as to how the Emory-wide directory service differs from them and to what extent they could be the basis for this service. To better delineate the nature of the directory envisioned by the architecture, it is useful to distinguish the following types of directories.

- **Network Operating System (NOS)-based directories** are intended to support the specific needs of the NOS, and generally are necessary to make full use of the features of the NOS.
- **Application-specific directories**, such as Lotus Notes name and address book, Microsoft Exchange Directory and Novell GroupWise Directory, come bundled with or embedded in applications.
- **Purpose-specific directories**, such as the Domain name System (DNS) directory, are intended for a narrowly defined purpose and are not designed for much extensibility.

- **General purpose, standards-based directories** are intended to serve the needs of a wide variety of applications and be accessed using widely-supported standards, such as the Internet standard Lightweight Directory Access Protocol (LDAP).

Given its wide scope of access by many types of desktop and server systems and its varied contents, the Emory-wide directory service architecture envisions using a general purpose, standards-based directory. Although Emory already has an instance of such a system in operation that might be the basis of an Emory-wide directory service, as it stands that system is limited in scope, capacity and availability.

**Content management.** To ensure that the directory service is used appropriately yet allow it to easily grow in size and add information about additional things, the directory's contents will need to be carefully managed. An Emory governance body that is inclusive and representative of the perspectives of all components of Emory will be needed to decide what can be included in the directory and fend off attempts by individuals and systems to use it for file storage or to store data of only local interest.

**Update management.** Updates would also need to be carefully managed. Data that is allowed would come by prearranged means from sources that are authorities for the data they supply. An "authority" provides data that is correct, which implies that the data is valid and up-to-date. The source could also be the ultimate authority (called the "authoritative source"), that is, the official person or system that maintains the data. An example is the telephone database for on-campus phone numbers. In the case of individually maintained information, the individual would be the authoritative source and provide the updates. Alternatively, the source could be an authority that already integrates and consolidates data from multiple authoritative sources. Using such an authority would be the preferred approach, since doing so would reduce the number of sources and thus the complexity of obtaining the data. Regardless of the source, each entry in the directory must have a process to get timely updates. The architecture for obtaining the data is an issue for the Data Management Architecture Domain.

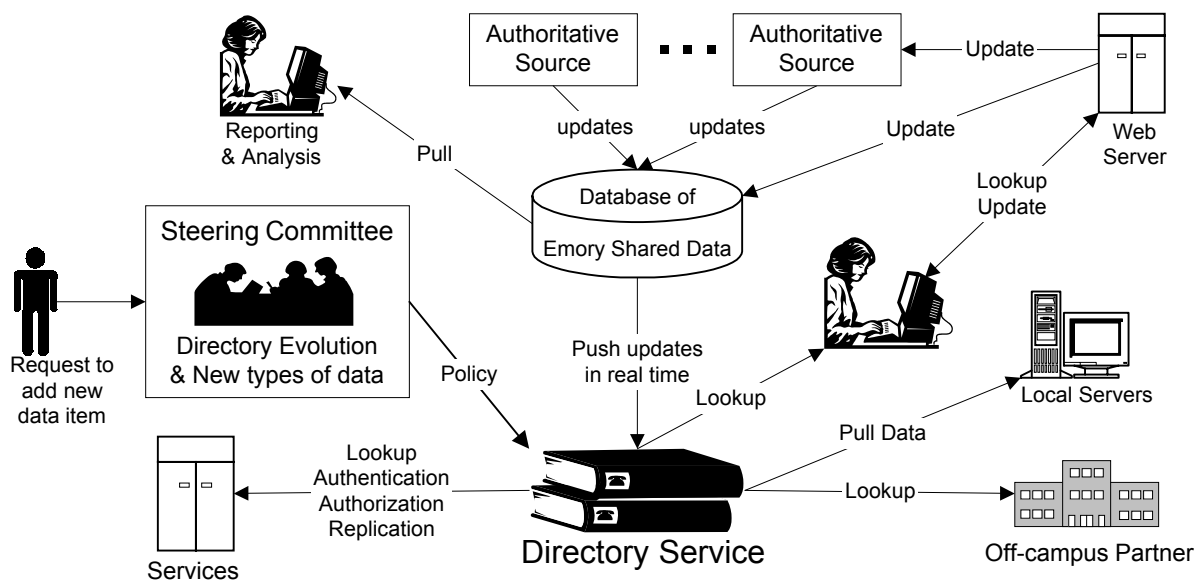


Figure 3-1. Directory service concept.

**Security.** In addition to decisions about contents and updates, decisions would also need to be made about the security classification of everything in the directory, and access controls would need to be implemented accordingly. For example, some data, such as userids and passwords would need to be tightly secured with restricted access. Other data might be available only to people or IT systems at Emory. To implement sharing of resources (such as databases) with other schools, authenticated access to some entries by specific systems at other schools might be allowed. Finally, some of the information (such as faculty names, their departments and a link to their Vitas) might be open for lookup by anyone, including the public. The directory service would thus need to allow and control access coming from many different types of systems. To do this it would likely need to be partitioned to allow placing data in the zone of trust appropriate for their classification.

## 3.2 The directory service architecture

**Purpose.** The Emory directory service architecture defines the infrastructure to create an Emory-wide directory service that meets the requirements outlined above, and that can grow and change quickly enough to continue to meet Emory's needs while keeping down complexity and support costs.

**Standards.** To accomplish this goal, the architecture specifies use of Internet standards for read and write access to the directory service as well as for authenticating to it and securely communicating with it. The architecture also specifies a standard format for feeds to load and update data plus utilities to convert common export formats to that standard. This approach only requires the directory product to support one standard format for feeds, yet accommodates other common formats.

**Replication.** In addition to the Internet standard for reading and searching the directory, the architecture envisions receiving data from the directory service via replication. Using this approach, an up-to-date copy of the directory or a portion of it is maintained elsewhere (in another directory) for use by another service. For example, Emory-wide userids and passwords, could be replicated to a directory in a password synchronization service.

Replication can also be used within the Emory-wide directory service itself to allow the directory information to be duplicated across multiple platforms in multiple locations. The replicas are accessed through multiple instances of load balancing equipment to provide:

- **Reliability:** In case one copy of the directory goes down, requests will be directed only to those that are up;
- **Availability:** Replicas can be located on different parts of the network to provide service even if part of the network is down;
- **Scalable Performance:** An increase in volume of queries can be handled by adding additional replicas.
- **Security:** Directory entries can be located on different systems in different zones of trust according to the security classifications of the entries.

The initial architecture envisions single master replication, which requires that all updates be referred to a single master that then sends the changes to replicas. Multimaster replication (to be supported later) would allow having multiple masters that could accept updates and send changes to keep all the copies synchronized.

### 3.3 Related Domain Architectures

The Directory Service domain is related to all other architecture domains. However, it strongly interacts with the following domains because of its dependence on them. In particular, principles of these domains are applicable to it. See their Domain Architectures for additional information.

Domain Name	Domain Description	Relationship to Security Arch.
Security	The security architecture defines the component processes, data feeds, and deployed hardware and software to electronically protect, preserve and control access to Emory's information technology assets. This domain covers such topics as firewalls, identification, authentication, and authorization.	Provides authentication and authorization information and other operational security policy.
Network	The Network Architecture provides the voice and data communication infrastructure for the distributed computing environment. It covers structure and topology, bandwidth management, cable plant, electronics (hubs, PBX, routers, switches), protocols (access, routing, naming, DNS), carrier services (frame relay, leased channels, ATM, WAN, SONET ring), wireless, firewalls, Internet connections.	Allows network-attached devices to communicate with the directory service.
Data and Information management	The Data and information Management Architecture defines the components and standards for accessing, exchanging, modeling, storing, converting, organizing, and distributing data and information. Product and technology categories governed by this domain include databases, data warehouses, data marts, data repositories, report writers, data modeling tools, data replication tools, data administration tools, data extract tools, data movement tools, and data cleansing tools.	Defines principles and standards for organizing, storing and retrieving data.
Platform	The Platform Architecture defines the technical computing components of the infrastructure: the client and server hardware platforms, the operating systems executing on those platforms, and the database environments and interfaces supported.	Defines principles and standards for platforms where the directory servers run.
Intranet and e-Commerce	The Intranet Architecture defines the technologies, standards, and guidelines for seamless, platform-independent Emory-wide communications and universal access to information. It includes web browser, inter/intranet servers (mail, web, news, proxy, ftp), intraware (middleware for inter/intranet), content management, web database connectivity, search engines, languages, Java development tools, web authoring tools, e-commerce, and web-based portals.	Defines principles, standards and products for Internet access.

## 4. IT Trends

---

### 4.1 Trends from Document 1

The following are the trends identified in Document 1 and their implications for the Directory Service.

#### 1. Hardware will get faster, cheaper, denser and more diverse.

- ❑ More and more platforms will need to access the directory.
- ❑ The directory will need to be able to easily support new platforms.
- ❑ There will be more use of the directory as platforms become more pervasive, which will cause the capacity of the directory service to need to be increased.

#### 2. Demand for capacity will continue to increase.

- ❑ New applications, uses and formats will drive the need for new types of data in the directory thereby increasing the complexity of the directory schema.
- ❑ There will be an increase in volume and rate of transactions to the directory.
- ❑ There will be an increase in storage requirements as additional information is added.
- ❑ More Emory departments will want to put more data in the directory due to wanting to put more information online.

#### 3. Demand for access to "anything" from "anywhere" will continue to grow.

- ❑ More and more systems will need to access the directory.
- ❑ There will be requirements to access the directory from anywhere on the Internet, so the directory will need to be accessible worldwide.
- ❑ Emory will increasingly need to be able to interoperate with external partners, such as ISPs and other universities.
- ❑ The directory might need to support a replica at the end of a slow, unreliable link for fault tolerance.

#### 4. Security will be a primary concern with increased dependence on network applications.

- ❑ Additional authentication methods continue to appear and need to be supported, so the directory service must be able to respond and add support for new ones.
- ❑ Use of the directory for authentication and authorization will require the highest level of availability and fault tolerance that Emory can provide. To make this affordable might require partitioning the directory data according to availability needs and providing the level fault tolerance appropriate to that need.
- ❑ There will be more interest in obtaining information from the directory for marketing purposes.
- ❑ Additional security for the directory service including physical security will be needed.
- ❑ The directory will become a well-known target, which might require hiding it, parts of its data, or parts of its functionality.

**5. IT expertise will continue to be scarce and its cost will continue to rise.**

- ❑ The need for knowledge specific to directory service technologies and vendors will lead to specialization, and specialists will be able to obtain certification
- ❑ Administration of the service will need to be easy so that fewer people are required, especially as the service grows.
- ❑ Cross training, documentation, and other measures will be required to allow for staff turnover.
- ❑ Documentation specialists will be needed to help ensure that the service is documented.

## **4.2 Domain-specific Trends**

The following are trends specific to the Directory Service.

1. Directory service standards and products will continue to undergo rapid change for the next few years.
2. The use of products that depend on directories will continue to increase.
3. The usage of directories will increase. Network Operating System-specific products will add the capability for directories to communicate with each other.
4. Network Operating System directories will continue to be needed in the short term (next three years).
5. There will be a staggered move toward directory interoperation among Network Operating System vendors.
6. Microsoft will continue to have proprietary directory features that impede interoperability.
7. The Extensible Markup Language (XML) will become the preferred standard for importing, exporting, and exchanging files of data.

## 5. Directory Service Principles

---

The following principles—derived from the Conceptual Architecture Principles of Document 2—apply to the entire Directory Service domain. They are to be used to guide the implementation of the Emory-wide infrastructure that creates the Directory Service. In particular, they are to be used in the evaluation, selection, design, construction, implementation and deployment of products and systems. Following these principles will not only help the implementation succeed, but will also promote logical consistency with other domains and the architecture requirements and Conceptual Architecture Principles (CAPs).

The table in Section 5 (page 5-5) shows how these principles support the CAPs. While giving just the titles and statements of the principles is sufficient to summarize them, such a list is insufficient for a full understanding of them. To see the reasoning behind a principle and the implications of following it, the reader is advised to consult the details in Section 12.

### 5.1 Overall Domain Principles

#### Flexibility

- 1. Flexible searching.** The Emory-wide directory should support searching for all entries that meet specified criteria and should allow specifying what information is to be returned from the matching entries. The directory should be able to restrict what can be searched and how much information is returned according to the authorization of the requester. The search criteria should allow specifying complex Boolean expressions of tests that taken together allow information in the directory to be readily located. (Page 12-2)
- 2. Support groups with dynamic membership.** The Emory-wide directory service should be able to define groups whose members are specified using search criteria that are evaluated each time such a group is accessed. (Page 12-3)
- 3. Allow easy change to contents, structure and privileges.** The directory service should be built so that it is easy to:
  - Add new types of data;
  - Change the source of an entry;
  - Delegate administration or other privileges of a entry;
  - Import and export all or subsets of the data in a form that can be readily imported or exported by other systems; and
  - Change to use of another directory service product.
  - Change the directory schema. (Page 12-4)

#### Scalability & Extensibility

- 4. Deploy a scalable and extensible directory service.** The Directory Service should be able to expand quickly and economically in capacity, scope, availability and reliability to support an increase in data storage, speed of response to requests, uptime, read and write transaction rate, simultaneous connections, manageability, sources of information, numbers of identities that can access it and types of access they have, and the number, sizes and types of things stored. (Page 12-5)
- 5. Make it easy to interface with new technologies, devices and systems.** The Directory Service should be able to support access by new technologies, devices, systems, and types of clients as quickly as needed. (Page 12-6)

- 6. Employ extensible and widely supported standards-based interfaces.** To the extent possible, the directory service should employ widely supported standards-based interfaces through which it can interact with existing and new technologies and systems. It should support adding and extending interfaces to add new capabilities without disturbing existing portions of the infrastructure. (Page 12-7)

### Construction

- 7. Build the directory service in a modular, loosely coupled way.** The directory service should be engineered with a bias toward using highly discrete, modular, loosely coupled components. (Page 12-8)
- 8. Seek simplicity balanced with other needs.** The directory service architecture should contribute first to simplicity in overall campus infrastructure, second to simplicity of use of the directory service, and third to simplicity of its own internal structure. The directory service architecture should also balance simplicity in design, operation, and usability with requirements for functionality, responsiveness, and flexibility. (Page 12-9)
- 9. Provide defined and reusable methods and interfaces.** The directory service should have defined and reusable methods to add, update, search, bulk load, extract, and otherwise manipulate its data. Access to these methods should be provided through defined and secured interfaces. The directory should seek to minimize the number of interfaces that it must support. (Page 12-10)
- 10. Standardize for interoperability.** The Emory-wide directory should use standard interfaces, protocols, names, and values that enhance interoperability with systems and clients in common use within Emory and the higher education community. (Page 12-11)
- 11. Use industry standard solutions where feasible.** The Emory-wide directory should to the extent feasible use software and hardware that are vendor supported and widely used in Emory's peer communities (such as healthcare and higher education). The software should allow needed customizations in such a way that they are supported in future versions. (Page 12-12)

### Information flow

- 12. Data comes only from a single source of authority for that data.** Each data item in the Emory-wide directory service should come from a single source that is an authority for that data and from nowhere else. (Page 12-13)
- 13. The directory service should be event-driven.** The directory service should accept events in real time and act on them. It should send events as soon as they are ready. (Page 12-14)
- 14. Document directory contents and information flows.** The flows of information into and out of the directory, the meaning and source of each type of thing in the directory, what is replicated and where, and the flows within the directory's components should be documented. The documentation should be done for each type of thing in the directory. (Page 12-15)

### Contents & Usage

- 15. Provide unique names.** The directory service must provide each entry with a unique ("permanent") identifier that does not change. Each entry should also have an additional unique name that can change as long as it remains unique. (Page 12-16)
- 16. Standardize names and their values judiciously.** Emory resources with names in the Emory-wide directory must have standard names and values (including formats). Standardization of values whose authoritative source is personal should only be

enforced to the extent necessary to ensure that the values can be used for their intended purpose. Additional names should be defined and used when other formats or variations must be introduced. (Page 12-17)

- 17. Provide data useable campus-wide.** The directory service should contain or have links to as much data and information as are of Emory-wide interest or can be reused across Emory. (Page 12-18)
- 18. Support information sharing.** The Emory-wide directory service should provide an easy way to share information according to Emory policy. (Page 12-19)
- 19. Facilitate access to Emory-wide resources.** The directory should include information about Emory-wide resources that is sufficient to help assess a resource's potential suitability for a particular purpose and that tells how to access the resource or who to contact to request access to it. (Page 12-20)
- 20. Encourage appropriate use of the directory service.** Encourage adding data useful to Emory. Centrally bear the cost to create and maintain the directory. Make usage that is encouraged easy and no additional cost, and charge for usage that is allowed but discouraged. (Page 12-21)

## Security

- 21. Provide fine-grained access control.** The Emory-wide directory service should be able to provide fine-grained access control to its contents according to Emory policy. (Page 12-22)
- 22. Support the needs of security.** The Emory-wide directory service should support the needs of the security architecture—especially access control—according to Emory policy. Corollaries: a. The directory service should be able to provide data for Emory-wide common security. b. The directory service should be able to change as needed to provide the required level of responsiveness and availability. (Page 12-23)
- 23. Be highly secure and interoperable with security layer.** The directory service should be highly secure. It should be interoperable with the common security layer and use it for its own security. (Page 12-24)
- 24. Secure the directory data.** Everything in the Emory-wide directory must be classified for access control, be secured accordingly, and have a documented owner and source of authority. (Page 12-25)

## Management & Evolution

- 25. Do not outsource the directory service.** The Emory-wide directory service must be connected directly to the Emory network and supported by Emory staff. The support staff should be available to respond at any time day or night. Only Emory staff may do maintenance of the operating systems on platforms where portions of the directory service run. Under these conditions, maintenance and operation of the hardware and support of the directory service software may be outsourced. (Page 12-26)
- 26. Develop directory competency of Emory staff.** Develop Emory staff competency in the skills and knowledge to use, support and evolve an Emory-wide directory service. (Page 12-27)
- 27. Manage the evolution of the directory.** The evolution of the Emory-wide directory should be planned and governed across the enterprise, with at least a yearly review at a point in the budget cycle that allows its projects to seek funding. (Page 12-28)

- 28. Manage the Emory-wide directory as an Emory asset.** The Emory-wide directory should be managed as an Emory asset by one of Emory's enterprise IT organizations. In addition, it should have a steward, and to the extent consistent with Emory's security policy, it should allow for remote, highly automated management and make its documentation accessible via Emory's Intranet. (Page 12-29)
- 29. Manage the Emory-wide Directory in a unified way.** Emory's IT departments and unit leaders should have a common vision and understanding of what it means to have an Emory-wide directory, and there should be a process to implement and enforce that vision across Emory. This is needed even when IT responsibility is decentralized. (Page 12-30)



## 5.3 Configuration principles

### 5.3.1 Schema

1. Minimize the use of Emory-specific names, but distinguish them by putting “emory” in front of them. Reason: Provides interoperability and uniqueness outside Emory and with standard clients.
2. Do not use an OID as a unique name for an entry. Reason: Such use of OIDs is not conventional.
3. When a conflict of values occurs due to distinct use in different systems, store them all with distinct names. Reason: They are all correct in their own context. Having them all enhances look-ups.
4. Entries for people should be in a separate flat directory space. This means that it is not hierarchically organized. Reason: Reduces maintenance as people move and structures change. Also allows people to have multiple affiliations and makes for simpler look up.
5. Minimize the amount of business rules and edit checks that are in the directory. Reason: A general purpose, standards-based directory is typically not able to support embedded business rules and edit checks well.

### 5.3.2 Replicas

6. Use no more replicas than needed. Reason: Reduces complexity.
7. Place replicas to maximize availability and lookup performance and minimize latency between them. In particular, do not put a replica at the remote end of a slow link when the replication traffic exceeds the remote lookup traffic. Reason: Reduces complexity and improves performance.
8. Do not require clients to accept referrals. Instead allow authenticated clients to do updates. (See also #14 below). Reason: Few clients support referrals.
9. Use the same schema across all replicas. Centrally manage the enterprise directory and all its replicas and schemas. Reason: Needed to reduce complexity and maintain security.

### 5.3.3 Content

10. Limit the number of interfaces, but provide enough to simplify the whole infrastructure. Reason: This approach strikes a balance in integration complexity between one extreme in which too few interfaces are supported to allow easy integration with most of the infrastructure, and the other extreme in which more interfaces are supported than needed to integrate with most of the infrastructure.
11. When data to be included in the directory already has a standard format, the format of the directory entry should honor the existing format rather than introducing a new one. Reason: Reduces overall complexity.
12. Reduce the number of sources of input to the directory to the extent practicable. Reason: Reduces complexity.
13. Use existing authorities as the source of directory data. Reason: Leverages existing integration and consolidation; reduces complexity by reducing the number of sources of feeds.
14. Any directory data entered or updated by individuals must have an associated application (preferably web-based) for doing so that enforces edit checks and business rules. Use of

such an application would normally be the only way for an individual to enter or update data in the directory. Reason: This is necessary to minimize business rules in the directory yet maintain standards for the data.

15. When deciding on whether to put data in the directory or put a pointer to the data, prefer to put data in the directory that has a need for high availability, high rate of read access, and low rate of update. Prefer to point to data with a lower need to be available, lower rate of read access, and higher rate of update. Reason: Providing high availability and responsiveness is expensive. Also, the directory service is not tuned for a high rate of update.

#### **5.3.4 Security**

16. Use transport layer security where appropriate, such as for password synchronization, loading passwords, and otherwise transmitting or receiving sensitive information. Reason: Principles 22 and 23.
17. Partition the directory data in such a way that data supporting security can be given the strongest security available at Emory. Reason: Principles 22 and 23.
18. Consider authorization via the directory mainly for cross-application authorization rather than for use by an application to control access to its own resources. Reason: The access control models used by applications—especially transaction oriented systems—are often too complex to outsource to a directory.
19. Do not attempt to synchronize any controls of access to the directory itself with those of other directories, such as NetWare or NT directories. Reason: This is not well supported, and the consequences are not well understood. It thus creates a high security risk.
20. Keep access control lists as short and simple as practicable.
  1. Fine grained access controls should be general rules that apply to all instances of an entry, rather than to a particular instance. Reason: Reduces complexity and impact on performance.
  2. Base access control that requires authentication on membership in a group. Reason: Reduces complexity of the access control lists and reduces the amount of change to them.

## 6. Technologies

---

The following are the technologies needed to implement this domain infrastructure. Those governed by this domain are given first, followed by assumed technologies from other architecture domains. These assumptions are necessary until architectures are established in those domains.

### 6.1 Governed Technologies–Directory Service domain

Technology Resource	Definition & Notes
Directory server	For running the software to provide the directory service.
Replication	For maintaining synchronized copies of all or portions of the data.
Batch import & export	For converting, loading and dumping a file of data into and out of the directory.
Schema	For naming and organizing the data.
Anonymous Client	For anonymous read access
Authenticated Client	For update access

### 6.2 Technologies needed from other architecture domains

These technology resources governed by other architecture domains are also needed by the directory service architecture.

Technology Resource	Usage & Notes	Governed by
Programming language with interface to LDAP	Write LDAP utility clients	Application
Platform	For databases, scanners, firewalls, etc.	Platform
Load balancing	Distribute load among multiple servers	Platform
Web front end	Provide access to directory via a web browser	Intranet & e-Commerce

## 7. Standards Compliance

---

For the technologies governed by this domain, this section gives standards that are in current use at Emory or that are recommended for future use at Emory plus recommendations regarding compliance with them. See the Glossary for definitions. Standards governed by the architecture are given first, followed by assumed standards from other architecture domains. The assumed standards are necessary until standards are established in those domains.

### 7.1 Benefit of standards

There is considerable variation in data look-up at Emory that creates needless complexity and difficulty. The following comparisons give an indication of where this occurs and how the use of standards can lead to more simplicity and clarity.

Current state:

- There is a database of Emory Shared Data (ESD) whose data includes official names of departments, buildings, and rooms. It also contains information about people such as their postal addresses, preferred email address, type, and whether or not they are enrolled. Operational systems use the ESD to share data with each other and make data available for analysis and reporting. For example, the ESD is used to generate email lists dynamically.
- An LDAP-accessible general-purpose standards-based directory is in operation for use by enterprise systems. It uses the iPlanet directory server product and runs on a single NT system. It acts as the master for service-specific replicas used by Email, authentication, password synchronization with the EMORYUNIV NT domain, and roaming access. Its primary source of data is the ESD mentioned above. Only the public information in the directory is accessible via anonymous access. The web-based online directory uses this LDAP-accessible directory to look up and display a subset of the public information about people. The email relay uses it when processing incoming email to look up the system that has the recipient's mailbox. The PeopleSoft web interface uses it for authentication.
- The LDAP-accessible directory system mentioned above is limited in scope, capacity and resilience.
- Novell's NetWare Directory Service (NDS) is widely used at Emory as a NOS directory running on 61 servers in 15 trees. Although NDS has had the capability to make its data accessible via LDAP since 1998, this capability has not been used in production at Emory. eDirectory (NDS 8.5), which has native LDAP v3 support and can run on popular non-NetWare systems, is currently being tested as a candidate to implement a directory service for Healthcare. eDirectory is also being used or tested by a few other Emory units, and work is underway to update the Emory NDS tree to allow for eDirectory. However, this effort is made difficult by the variation in versions running on the 34 partitions that make up the tree. That variation also requires using two different versions to support the root.
- Domino version 5 is in use by a few units. Domino has an application specific directory that can authenticate using LDAP and can be accessed using LDAP.
- There is a wide variety of data that are unavailable for access outside the systems in which they are stored but of wider potential applicability. Even when data are available for access, determining where to look, how to access them, and what they mean can be difficult.
- It is not clear that there is a process by which the Emory community can request that additional data items be added to the existing general-purpose directory or the Emory Shared Data. The contents of both are directed to the needs of the operational systems involved and the people in charge of the data

- When the requirement is to provide general access to data or access at any time from anywhere on the Emory network or the Internet, the current practice is to put it on a web server and possibly index it in the Emory Ultraseek search facility. The result is data with a variety of formats, which makes their use by computer programs difficult. Searches are also difficult due to a full-text search returning too many hits.

Future state with standards:

- There would be a standard process by which the Emory community could request that an item of data be made available for authorized access through the directory service, and a standard process to decide whether to grant such a request. This would increase the availability of widely useful data.
- There would be a standard place to look up data at any time from anywhere on the Emory network or the Internet. This would increase the accessibility of widely useful data.
- The data would be obtainable as separately named items using standard names in a format that can be used by both people and computers. This would enhance reuse by both people and computers.
- There would be a standard way for the directory service to increase its scope, capacity, responsiveness, availability and reliability. This would help it stay current with needs.
- There would be standards for documenting the meaning of data items so the meaning could be looked up in the directory service. This would make the data more useful and through its reuse lead to more consistency of use and interpretation.
- Use of a widely-available Internet standard for access would make the data more widely available for sharing and use by authorized systems than would be the case if the systems had to interface with a relational database system using a less widely-available interface. The loose coupling provided by use of the Internet standard would make integration with other systems easier. Wider sharing of commonly needed data would reduce the cost of maintaining duplicates and lead to more consistency of results based on the data.
- The use of Internet standards for access would provide a basis to inter-operate with partners that require sharing data, such as when implementing sharing of resources with another university.

## 7.2 Descriptors for standards

- Acceptance
  - “International”: Wide use by the international technology community
  - “National”: Wide use within a country
  - “HiEd”: Wide use within Higher Education community
  - Name of a group: Wide use specific to this group (e.g. Internet2 members)
- Authority
  - Name of the organization that defines and maintains the standard
  - “De facto”: Not defined as a standard, but receives wide acceptance
  - “Info’l.”: Informational RFC. Does not have the force of an Internet standard, but is one that receives wide acceptance within the Internet technical community.
- Extent of usage of standard at Emory
  - “H” = “High” use
  - “M” = “Medium” use
  - “L” = “Low” use
  - “N” or “None” = Not used

- Emory future compliance with standard  
 Sometimes includes caveats (such as a length of time)
  - “Y” = Compliant: Full compliance with the standard
  - “L” = “Limited compliance”
  - “N” = “Not compliant”: No compliance
- Missing data:
  - “ND” = Not decided: No decision has been made
  - “NK” = Not known

### 7.3 Standards Compliance–Directory Service

Technology Resource	Standards	Authority/ Acceptance	Extent of Usage at Emory	Future Compliance	Reason
Directory server	LDAP v2 & v3	IETF / Int'l.	Medium	Y	Interoperability
	LDIF v6	IETF / Int'l.	Low	Y	No alternative
	SSL	NK / Int'l.	None	Y	Authentication
	TLS	IETF / Int'l.	None	Y	Authentication
Replication	Vendor dependent based on slurpd	Originally U of MI / NK	Low	Y	No alternative
	LDUP	IETF working group / Will be Int'l.	None	When adopted	Supports multivendors
Batch import & export	Fixed Width	Emory ESS / Local	Low	Y 1 yr	No alternative
	LDIF	IETF / Int'l.	ITD	ND	Needs study
	XML	W3C / Int'l.	None	Y	
Schema	inetOrgPerson	IETF / Int'l.	Low	Y	Interoperability
	eduPerson	EDUCAUSE & Internet2 / HiEd	Low	Y	Interoperability
	LIPS	NAC / Int'l.	Low	Y	Interoperability
	emoryPerson	Emory ITD / Emory	Low	Y	Emory needs
	OID (rfc1778)	IETF / Int'l.	Low	Y	Required
	X.500 naming	ISO / Int'l.	Low	Y	Compatibility
	X.520, X.521 objects & classes	ISO / Int'l.	Low	Y	Compatibility
Anonymous Client	LDAP v2	IETF / Int'l.	High	Y 1 yr	Few v3 clients
	LIPS	NAC / Int'l.	High	Y	Interoperability
Authenticated Client	LDAP v2	IETF / Int'l.	Medium	Y 1 yr	Few v3 clients
	SSL	NK / Int'l.	Medium	Y	Authentication
	TLS	IETF / Int'l.	None	Y	Authentication
	LIPS	NAC / Int'l.	Medium	Y	Interoperability

### 7.4 Assumed standards from other architecture domains.

Technology Resource	Standards	Authority/ Acceptance	Extent of Usage at Emory	Future Compliance	Reason
Programming language with interface to LDAP	Perl v5.05+	Larry Wall / Int'l.	High	Y	
	PerlLDAP v41.4.1 or later	Mozilla.Org / Int'l.	Low	Y	
Load balancing	Platform specific	Platform vendor / vendor	Low	Y	No choice
Platform	NT	Microsoft / Int'l.	High	Y	
	UNIX (AIX, Solaris)	Vendor / Int'l.	Low	Y	
	Hardware (IBM, Sun or Intel)	Vendor / vendor	High	Y	
	TCP/IP protocols	IETF / Int'l.	High	Y	Required
Web Front end	HTML v3	W3C / Int'l.	High	Y	
	Pure JAVA 2	Sun/ Int'l.	Medium	ND	
	"Javascript" (ECMAScript)	ECMA-262, ISO-16262 / Int'l.	High	Y	
	HTTP 1.1, rfc 2616	IETF/ Int'l.	High	Y	

## 8. Standard Products

This section gives the relevant products currently in use at Emory and their current and recommended future status relative to their use to implement the Directory Service domain architecture. Categories that do not have products in use at Emory represent gaps that need to be filled to implement the architecture. Standards specified by this domain architecture are given first, followed by assumed standards from other architecture domains.

Life-cycle status (current and future) for products and product categories:

- “Test”: Under test, not for production use
- “Pilot”: Step following test; limited production use; goes to recommended or STD if successful.
- “Rec”: use at Emory is recommended and encouraged.
- “STD”: Emory standard; use is required.
- “Maint”: In use, supported in maintenance mode only.
- “Used”: In use, not supported
- “Out”: In process of being phased out

Codes for missing data:

- “ND” = Not decided: No decision has been made
- “NK” = Not known
- “NC” = Not collected
- “NA” = Not applicable

### 8.1 Standard Products–Directory Service

Technology Product Category	Current Emory Products: name / version/ source	Consistent with position on trends / Aligns with principles / Complies with standards	Life Cycle Status	
			Current	Future
Directory server	iPlanet / 4.1+ / Sun	Y / Y / Y	STD	STD v5
	NDS eDirectory / 8.5 / Novell	Y / Y / Y	Test	ND
Replication	iPlanet / v4.1+ / Sun	Y / ND / Y	STD	v5, LDUP
	NDS eDirectory / 8.5 / Novell	Y / ND / Y	ND	ND
Batch import & export	Home grown flat to LDIF converter / NA / ITD ESS	ND / ND / Y	STD	ND
Schema	EmoryPerson extension / NK / ITD ESS	Y / ND / ND	STD	STD
	iPlanet std. / NC / Sun	ND / ND / ND	STD	STD
	Emory OID 1.3.6.1.4.1.5524 / NA / IANA http://www.alvestrand.no/objectid/1.3.6.1.4.1.5524.html	Y / Y / Y	STD	STD
Anonymous Client	Messenger / 4.x with x at least 5 / Netscape-AOL	ND / ND / ND	STD	ND
	Eudora / 4.3+ / Qualcomm	ND / ND / ND	Used	Out
	Pmail / NC / David Harris	ND / ND / ND	Used	Out
	Outlook Express / NC / Microsoft	ND / ND / ND	Used	Out
Authenticated Client	Web directory server gateway client / NC / Sun	ND / ND / ND	Rec	ND

## 8.2 Assumed products from other architecture domains.

Technology Product Category	Current Emory Products: name / version/ source	Consistent with position on trends / Aligns with principles / Complies with standards	Life Cycle Status	
			Current	Future
Programming language with interface to LDAP	Perl / v5 / NC	ND / Y / Y	STD	STD
Load balancing	Big-IP / 3.3.1 / F Five	ND / Y / Y	STD	STD
Platform	NT / 4 sp 6a / Microsoft	ND / Y / Y	STD	1 yr
Web front end	Directory Express / NC / Netscape-AOL	Y / ND / ND	ND	ND

## 9. Configurations

---

The intent of the configuration section is to provide enough detail to prescribe implementation at Emory in a way that follows the principles yet uses standards and products to reduce needless complexity and keep down support costs. At this stage of the development of this domain, the implementation details are still being worked out. Once these details are available, they would be included or linked from this section and made available according to their security classification.

### 9.1 General

The architecture reduces the complexity and the total cost of ownership through the use of:

- Identical copies of the directory and the platforms on which they run. This allows them to be more easily managed as multiple instances of the same system.
- The same schema on all replicas.
- Only one interface for general-purpose access.
- Only one format for loading data from a feed, with adapters to convert to that standard.
- As few sources of data as feasible, with aggregation of authoritative sources taking place at one of the sources rather than at the directory itself. This reduces the complexity of interfacing with systems and their staff.
- One standard—namely replication—for batch output from the directory. This is simpler to manage than converting back and forth.

Other ways that the configuration reduces complexity and total cost of ownership can be gleaned from the Configuration Principles (page 5-6).

Figure 9-1 on page 9-2 indicates generally how the standard products would be used to accomplish this. The details are yet to be decided. In many cases it is assumed that choices would be made based on standards from other architecture domains that are yet to be created. Configuration details would include:

- The products and versions or models for the software and hardware (including directory product, platforms, load balancing equipment, web access product).
- The product configuration details.
- The schemas for the data and the meta data, including their meaning, security classification, source and date of update.
- Access control lists for the entire contents of the directory.
- How the directory would be partitioned to allow data to be in the zones of trust for their security classifications.
- Location of replicas to provide needed resilience and responsiveness.
- Migration plan showing how to change the setup to migrate from the existing service to the new implementation.



## 10. Considerations & Next Steps

---

### 10.1 Considerations for this domain

1. **Userid in the distinguished name (DN).** Use of userid in the DN causes problems when the userid changes. The entry must be deleted and recreated with the new DN.
2. **Securing the Directory Service.** The information in the directory service will need to be classified and placed in sufficiently secure zones of trust. Given the varying types of information that the Directory service would likely contain, partitioning the information may be required if it cannot all be placed in or replicated to a sufficiently secure single zone.
3. **LDIF & XML.** The current de facto standard for feeding information into an LDAP directory is LDIF. However, the XML-defined Directory Services Markup Language is likely to replace LDIF as a format for exchange of data.
4. **dc naming.** The Internet2 directory group is discussing a standard for locating an institution's directory service that specifies the use of DNS SRV records (rfc2782) to locate the LDAP directory servers and "dc naming" (rfc2247 and rfc2377) to derive the root of the directory tree. Emory currently names the root of its tree 'o=emory.edu' instead of Internet2's recommended 'dc=emory,dc=edu' but supposedly Emory's directory can support multiple root names.

### 10.2 Considerations for other domains

#### 10.2.1 For the Security Domain

1. **LDAP authentication.** The immediate role a directory service would play if any in supporting authentication of network ids is an issue. The current LDAP-accessible directory is used by a small number of trusted systems on a secure network to authenticate network ids. Untrusted services desire to do the same. However, allowing untrusted services to use the directory service to verify passwords can lead to passwords being visible on insecure networks or in insecure servers, which could result in passwords being compromised.
2. **Windows and NetWare synchronization with the directory service.** The question is to what extent to implement synchronization between data in the directory service and data in the directory of a local area Network Operating System (NOS) such as NetWare or Microsoft Windows. The issues are that much of the information is only of local interest and is under the control of the local administrator. Password synchronization would provide automatic account creation and removal in the NetWare and the NT domain, reducing the burden on local administrators and giving them authority for that information. Whether and how to delegate local administration of Emory-wide network accounts is a security policy consideration.
3. **The future of password synchronization.** The question is how to achieve the desired future state in which desktop authentication to network operating systems is done via LDAP. The issue is that desktop authentication to Microsoft Windows and NetWare servers does not support the use of LDAP and is not expected to do so any time soon. To help position the campus to achieve the desired future state, the directory service task force recommends that Emory provide a single NDS tree and NT domain that are synchronized with at least the userids and passwords in the Emory (LDAP) directory service. This is already being done for an NT domain (EMORYUNIV). Reasons: This will act as an incentive for isolated NDS and NT directories to become part of the campus tree or domain as applicable. The result is a

reduction in complexity of the campus environment. It makes reuse of the synchronized information easier and avoids investments in local scripts to extract information to NOS directories.

4. **Application Access Control.** The security architecture suggests that access control be externalized. However, the authorization model used by an application—especially a transaction based one—is often too complex to outsource to a directory.

### 10.2.2 For the Network Domain

1. **DNS and dc naming.** This plan for using DNS naming to find information in directories at other institutions (rfc2247 and rfc2377) uses DNS SRV records (rfc2782) to locate an institution's LDAP-accessible directory servers. Emory's DNS will need to support SRV records to make its directory accessible through this approach.
2. **DNS and Windows 2000 (Win2K).** Win2K includes an implementation of DNS to name and locate Win2K computers and services and allow dynamic updates to network numbers, especially when those numbers are assigned dynamically and can change dynamically. Win2K DNS can use Active Directory to store its records. Current Emory advice is to configure Win2K workstations with the address of the Internet DNS and turn off dynamic updates. Issues for the Network architecture domain, which governs DNS, include but are not limited to:
  - Would the Win2K Emory namespace be disjoint from or overlap with the existing Emory Internet DNS namespace?
  - What alternate DNS would be provided to support Win2K given that the Emory Internet DNS is not expected to support dynamic updates as a matter of policy?
  - Can DNS be configured so that Win2K's use of DNS records to identify LDAP servers does not interfere with that use by dc naming?

### 10.2.3 For the Platform Domain

1. **Active Directory (AD) and Windows 2000 (Win2K).** AD is an X.500-based directory with an LDAP interface that comes with Win2K to support Windows-based networks. Issues for the Platform domain, which governs network operating systems such as Win2K, include but are not limited to:
  - What central Win2K services would be provided?
  - What if any information would be fed into the AD global directory other than what Win2K puts there?

## 10.3 Next Steps

- Create a schema for meta data about data items.
- Classify for security the directory data and add the classifications to the asset database of the security architecture.
- Partition the directory according to the classifications.
- Further partition the directory to separate out profile information that can be written by user systems. Move these to a separate service.
- Implement load balancing.
- Choose a web-based LDAP client.
- Set up a steering committee that is inclusive and representative of all components of Emory to decide contents and its security classification.
- Create a migration plan and a funding model.

## 11. Glossary

Term	Definition
access control	Enforcement of authorization. The means by which access is explicitly enabled or restricted in some way (usually through physical and system-based controls).
attribute	A piece of information that describes some aspect of an entry. Examples for a person could be "Email" and "Last name". Some attributes can occur multiple times. Technically an attribute is a type and a set of values that include a name, OID, whether there can be more than one occurrence of the attribute in an entry, the syntax of the value, kinds of matching that can be performed on it, and other things.
authentication	Verification of a person or system's identity. Demonstrating knowledge of a password is a common approach to prove identity.
authoritative source	For an entry, especially when data is replicated to other databases: The "official" process or system of record that maintains the entry, and the place where data validity issues and creation and update rules are addressed. The ultimate authority.
authority	Provides correct, valid, timely data and information.
authorization	Granting of privileges based on authenticated identity, such as saying what a person or system can access and do.
balance	Harmonious arrangement, relation or distribution of parts or elements within a whole.
bind	The operation of establishing a connection with a directory server. It may involve proving identity or be anonymous (no identity established).
component	A part that has a clearly defined usage, function, or purpose and that can be accessed in a defined way without knowledge of its internals. A component can be a hardware part, software, a database table, a software routine, a code module, a server, a database, a system, etc.
data warehouse	A separate read-only database containing specific-purpose collections of data for analysis, reporting and decision-making. To that end, the data has been integrated, consolidated, and organized for ease of access and to remove redundancy, and has been processed to ensure correctness and consistency. It is typically historical, summarized, subject-oriented data.
directory	A specialized database designed for high read performance, for ease of extensibility, to be able to be widely distributed and replicated and to support standards for naming and access. It is best used to look up data that is highly structured, has small pieces, and does not change often.
distinguished name (DN)	Unambiguously and uniquely defines an object within the directory. In a tree-structured directory, it is an ordered list of node names that gives a path from the object to the root when read from left to right.
eduPerson	LDAP object class that provides a common list of attributes and definitions for people based on person attributes widely-used in higher education.
Emory	In the context of organizational scope, it is all sites of the Emory enterprise.
enterprise, Emory	The entire legal entity to which the University and Emory Healthcare belong.
entry	(In a directory) A set of attributes and their values. It is imagined to describe an instance of an object, such as a person, organization or device.
Fine-grain access control	Access privileges can be as specific as required about what can be accessed, who can access it, when and from where they can access it, and what they can do to it.
identification	The means by which an IT asset determines the source of a request. Often a name or userid.
inetOrgPerson	rfc2798 - LDAP object class for a person. The attributes it holds are chosen to accommodate information requirements found in typical Internet and Intranet directory service deployments.

infrastructure, IT	The foundation on which IT systems are run. The basic stuff you need to have in place before you can start to build solutions for mission-oriented problems and productivity systems. It provides storage, bandwidth and processing power. It consists of the components of a computing setup: the wiring, routers, switches, operating systems, middleware, mainframes, servers and sometimes desktop machines.
intranet	Use of Internet technologies to deliver IT services internally to an organization. It is typically an internal internet that is separated from the global Internet by at least one firewall, and may employ additional Internet technologies to increase security.
IT	Information Technology
LDAP	A specification for a client-server protocol to retrieve and manage directory information. Version 2 was published as rfc 1777 and rfc 1778. Version 3 adds new features, improves compatibility with X.500(1993) and also better specifies how LDAP can be used with non-X.500 and standalone directories.
LDAP v2	<ul style="list-style-type: none"> <li>• rfc1777: Lightweight Directory Access Protocol</li> <li>• rfc1778: The String Representation of Standard Attribute Syntaxes</li> <li>• rfc1779: A String Representation of Distinguished Names</li> <li>• rfc1959: An LDAP URL Format</li> <li>• rfc1960: A String Representation of LDAP Search Filters</li> </ul>
LDAP v3	<ul style="list-style-type: none"> <li>• rfc2251: Lightweight Directory Access Protocol (v3)</li> <li>• rfc2252: LDAPv3: Attribute Syntax Definitions</li> <li>• rfc2253: LDAPv3: UTF-8 String Representation of Distinguished Names</li> <li>• rfc2254: The String Representation of LDAP Search Filters</li> <li>• rfc2255: The LDAP URL Format</li> <li>• rfc2256: A Summary of the X.500(96) User Schema for use with LDAPv3</li> <li>• rfc2829: Authentication Methods for LDAP</li> <li>• rfc2830: LDAPv3: Extension for Transport Layer Security</li> </ul>
LDIF	LDAP Data Interchange Format (v6) -- rfc2849
LDUP	Subject of an IETF working group to standardize master-slave and multi-master LDAP v3 replication. <a href="http://www.ietf.org/ids.by.wg/ldup.html">http://www.ietf.org/ids.by.wg/ldup.html</a>
LIPS	Lightweight Internet Person Schema. Defined by NAC. <a href="http://www.netapps.org/company/lipschemafinal.htm">http://www.netapps.org/company/lipschemafinal.htm</a>
loosely coupled components	Components depend as little as possible on knowledge of the state or on the performance of other components in order to use them.
Metadata	Data about data. For example, metadata could include the nature, context, quality, condition, characteristics, owner and location of the data, as well as the name, size, and data type of fields and attributes in records and entries. (Based on <a href="http://www.foldoc.org">www.foldoc.org</a> )
NAC	Network Applications Consortium. Made up of a group of end-user companies, NAC focuses on enterprise interoperability issues. <a href="http://www.netapps.org">http://www.netapps.org</a>
NOS	Network Operating System. A NOS runs on one or more servers to provide services over a network, such as file & print services, access to shared applications, and access control (via login). A NOS typically has an associated directory that it uses to provides shared access control and information to multiple servers. The directory contains information about authorized users and network resources whose access is under NOS control.
object	A type of entry in a directory. Objects are typically associated with real world things, such as a people, organizations, devices, mailing lists, etc. An object can also be more abstract, such as a definition of a group or an attribute.
object class	A type of object. An example is "emoryPerson". Every object has an object class attribute. The object's object class determines the object's structure, including the attributes that are allowed, required, and optional. See rfc 2252.
official	Adj. Implies that the University has assigned a steward to the noun it modifies.

open	Unencumbered specifications are freely available, independent branding and certification processes exist, and multiple implementations of a single product may be created.
ODS	Operational Data Store. A database containing detailed, partially reconciled, and nearly current data used for immediate reporting needs. Users sometimes also write to it.
OID	Object Identifier. Each object in an LDAP directory has an OID to uniquely identify it independently of its name. Defined in ITU-T X.208 (ASN.1). rfc 1778 gives syntax for use in LDAP directories, and assigns internet OIDs under 1.3.6.1. Organizations can assign numbers that start with their OID. Emory's OID is 1.3.6.1.4.1.5524.
owner	The person who has final rights regarding disposition of a resource. The owner defines who the stewards are and entrusts them with responsibility for the resource.
perl	Practical Extraction and Report Language. A scripting programming language especially designed for quickly building programs to perform utilitarian tasks.
perLDAP	Perl interfaces for writing LDAP clients.
replication	A method for maintaining an exact duplicate (called a "replica") of a set of data. There are two types. (1) Master-Slave or Single-Master Replication: A replication model that assumes only one server, the master, allows write access to the replicated data. (2) Multi-Master Replication: A replication model where entries can be written and updated on any of several replica copies, without requiring communication with other masters before the write or update is performed.
RFC	Request For Comments. Documents Internet standards and other technical aspects relating to the Internet. Can be found at <a href="http://www.ietf.org/rfc/rfcxxx.txt">http://www.ietf.org/rfc/rfcxxx.txt</a> where xxx is the RFC number.
schema	A set of rules about the allowable contents and structure of the directory. It includes attribute type definitions, object class definitions, and matching rules.
security	People, process and architecture, adapting functionality to risk, creating solutions that protect yet enable.
security infrastructure	The IT infrastructure that helps to secure enterprise IT assets.
simple	Not complex or complicated or involved.
slapd	Stand-alone LDAP server process that originated at the University of Michigan.
slurpd	Standalone LDAP update replication server process that originated at the University of Michigan.
steward	Of a resource: Responsible to the owner of the resource, sets policy for it, and typically has an active, working knowledge and understanding of it and its needs. Often a senior university official with planning and policy level responsibility for a resource created or maintained within a functional area or business process of the university. May empower a custodian to manage the resource.
technology	A manner of accomplishing a task especially using technical processes, methods, or knowledge.
TLS	Transport Layer Security. rfc2830 Extension for LDAP (v3)
University, the	In the context of organizational scope, it is the enrolled students and the employed faculty and staff of Emory University no matter where they are located.
userid	A name used for identification during authentication, especially when a password is used to prove identity.
W3C	World Wide Web Consortium. Develops web standards that it calls "recommendations." See <a href="http://www.w3.org/">http://www.w3.org/</a>
X.500, X.520, X.521	X.500 is the international standard for a directory service that includes a hierarchical naming standard. X.520 specifies standard attribute types and X.521 specifies standard object classes.

## **12. Appendix: Justification and Implications of the principles**

During the period when the principles are being established, each principle has a status as follows:

Proposed	Someone has suggested the principle, and it is under consideration, but has not yet been discussed.
Under discussion	The principle needs further discussion (modification is likely).
Standards track	This principle has been discussed and is proposed as a standard.
Adopted	The principle has been reviewed by appropriate campus groups and approved by university leadership.

## 1. Flexible searching.

*Status: Adopted*

The Emory-wide directory should support searching for all entries that meet specified criteria and should allow specifying what information is to be returned from the matching entries. The directory should be able to restrict what can be searched and how much information is returned according to the authorization of the requester. The search criteria should allow specifying complex Boolean expressions of tests that taken together allow information in the directory to be readily located.

Example 1. A university's directory had attributes for its people that included name, phone number, email address, status (faculty, staff, student), unit (name of school or division), and security classification (public, private, confidential). A search for the names and email addresses of all faculty members of the college or of the graduate school could then be done by specifying the condition: status is "faculty" and either unit contains "college" or unit contains "graduate". To make it difficult for the public to use its directory to create large email lists, the university limited to 50 the number of records returned to anonymous users. It also prevented them from seeing or searching entries not classified as "public." To enable use of the directory to create larger lists for official purposes, the university defined a group whose authenticated members could view and search all records and receive all of those that matched the search criteria.

### Justification

- Being able to search using complex Boolean expressions is necessary to provide flexible lookup of information.
- Being able to restrict what can be searched according to the authorization of the requester is needed to allow including information that is not public. It also allows information to be made available while preventing aggregation. An example is including home phone numbers but not allowing a search on that attribute.
- Being able to specify what information is to be returned is needed to reduce the complexity of processing the search results.
- Being able to restrict how much information is returned is needed to avoid inappropriate dumping of the contents of the directory, such as to create a mailing list of everyone at Emory.

### Implications

1. A search should be able to specify tests that include determining whether a given attribute is present, matches a given test pattern, contains a given test value as a substring, or is equal to, less than, or greater than a given test value.
2. The ability to use complex Boolean expressions implies that the tests may be placed in parentheses and connected using AND, OR and NOT.
3. The directory should allow retrieval of selected information from a single entry specified by its key.
4. A tree-structured directory should allow searches to be confined to the subtree starting at an entry identified by its key.
5. A tree-structured directory should allow searches to be confined to the entries below a base entry identified by its key.

## 2. Support groups with dynamic membership.

*Status: Adopted*

The Emory-wide directory service should be able to define groups whose members are specified using search criteria that are evaluated each time such a group is accessed.

Example 2. A university directory had an entry for each of its people that included attributes indicating their organizational status (student, faculty, staff), whether currently enrolled, and for employees their employee type (full time or part time). Using the dynamic group membership feature the university was able to define a group called "EmployeeStudent" of all people who were either staff taking a course (status is staff and currently enrolled) or student employees (status is student and employee type is not null). Any access to this group got those satisfying the criteria at the time of access.

### Justification

- Dynamic membership allows creation of groups whose members are automatically up-to-date at the time of access.
- Use of groups allows defining reusable searches without those using them having to know the criteria or having to keep up-to-date on changes.
- Managing the contents of a dynamic group is less complex than managing a group defined by explicitly listing all its members.
- Even when using a feed to update the directory contents and maintain static group contents, using a dynamic group is less complex.
- Defining membership dynamically allows the membership definition to be more quickly and easily changed.

### Implications

1. A group will need to satisfy applicable architecture principles, such as these from this Directory Service architecture:
  - The definition of a group must have a source of authority (see Principle 12).
  - Groups will need unique names (see Principle 15) that are standardized (see Principle 16).
  - Groups should be useable campus wide (see Principle 17).
  - The existence of groups will need to be documented (see Principle 14).
  - Groups will need to be classified for access control and secured accordingly (see Principle 24).
  - Appropriate use of groups will need to be encouraged (see Principle 20).
2. The ability to list all groups of which an entry is a member will be needed to satisfy the auditing requirement of security (see Principle 22, "Support the needs of security. ")

### **3. Allow easy change to contents, structure and privileges.**

*Status: Adopted*

The directory service should be built so that it is easy to:

- Add new types of data;
- Change the source of an entry;
- Delegate administration or other privileges of a entry;
- Import and export all or subsets of the data in a form that can be readily imported or exported by other systems; and
- Change to use of another directory service product.
- Change the directory schema.

#### **Justification**

- Many of these are typical changes that need to be made quickly, so they need to be easy to do.
- Getting data in and out in bulk in a form that can be imported and exported by most databases is needed to avoid being locked in to a particular directory service product. It allows easily changing to a different directory service product. It also facilitates initializing the existing system with data.
- Getting data in and out in bulk in a form that can be imported and exported by most databases is needed to allow easily changing to a different directory service product.
- Standards for cooperative use of directory services among universities, such as a standard for locating people in directory servers, are emerging. Conforming to them may require changes to the directory schema.

#### **Implications**

1. Emory will need an on-going mechanism for deciding what changes will be allowed.
2. Changes will need to be managed to minimize problems and maintain adherence to principles and standards.
3. Emory will need to track emerging standards for cooperative use of directory services among universities.

## 4. Deploy a scalable and extensible directory service.

*Status: Adopted*

The Directory Service should be able to expand quickly and economically in capacity, scope, availability and reliability to support an increase in data storage, speed of response to requests, uptime, read and write transaction rate, simultaneous connections, manageability, sources of information, numbers of identities that can access it and types of access they have, and the number, sizes and types of things stored.

### Justification

- Trend number 2 indicates that demand for capacity will continue to increase.
- The directory must be able to hold names and information about all people at Emory. It must be able to expand to hold additional information about them as well as add additional groups of people. For example, going forward, the directory must be able to accommodate adding alumni and doctors.
- Demands will include the requirement for additional storage resulting from the need to add data as well as new types of data.
- The use by operational systems will impose increasing requirements to improve availability and responsiveness.

### Implications

1. For the sake of availability, reliability, fault tolerance and performance, the directory service must be able to quickly increase its redundancy, number of masters, and number of replicas.
2. For the sake of manageability, the directory service must be able to control how it is updated, including limits on the size of fields and on what identities can modify them.
3. Local support will need to maintain awareness of local IT system and directory decisions that could impact capacity, scope, availability or reliability of the Emory-wide infrastructure or later integration with or use of the Emory-wide directory service.
4. When information about people comes from another system, then there should be a way to get temporary people into the system in a timely fashion.
5. The Emory-wide directory service should have the ability to incorporate information from varied sources and easily add additional sources.
6. Avoid changes that might adversely impact anything that reads from or writes to the directory service. Instead of changing the format, meaning or maximum size of an attribute, consider adding a new attribute.

## **5. Make it easy to interface with new technologies, devices and systems.**

*Status: Adopted*

The Directory Service should be able to support access by new technologies, devices, systems, and types of clients as quickly as needed.

*“New” means new to Emory.*

### **Justification**

- Technologies new to Emory that the directory service may need to support include PKI, digital certificates, biometric identification devices and smart cards.
- An example of an existing system that might need access to the directory service is Resource25.
- Examples of new clients that might need to access the directory service are clients of wireless networks and BlueTooth clients (which could be PDAs, doors, elevators, etc.)
- Other systems and devices that would or could need access to the directory service are DHCP, routers, and firewalls.

### **Implications**

1. Emory will need to continually test new technologies that will likely have to be supported.
2. Testing will incur cost in people's time to do the testing, in purchase of hardware and software for test systems, and in space to house the equipment.
3. The directory service will need to employ standard, extensible, and widely supported interfaces.
4. The directory service will need to be able to easily and quickly add support for new security methods and technologies.

## **6. Employ extensible and widely supported standards-based interfaces.**

*Status: Adopted*

To the extent possible, the directory service should employ widely supported standards-based interfaces through which it can interact with existing and new technologies and systems. It should support adding and extending interfaces to add new capabilities without disturbing existing portions of the infrastructure.

*“Wide support” includes not only the extent of support by other software but also support by different operating environments. “New” means new to Emory.*

### **Justification**

- Using widely supported interfaces as a standard for access to the directory service increases the probability that a new system or technology will already support one of those interfaces.
- Employing a widely used standard interface avoids the delay to implement an interface for a new technology and has the advantage of receiving wide testing through use.
- Using an extensible interface avoids the delay of having to change all the interfaces to accommodate a new capability.
- Use of the type of interface envisioned here enhances reusability of the infrastructure (and thus supports conceptual architecture principle B.1.4).

### **Implications**

1. Systems with application programming interfaces (API) will be preferred, since interfaces are typically implemented using APIs.
2. Interfaces will need to be “open” in the sense of being documented and accessible to anyone who cares to write to them.
3. Interfaces will need to be adequately secured so that they can only be used as authorized.
4. The directory service should allow new interfaces to be easily added.
5. The effect on the infrastructure of adding a new technology will need to be carefully assessed and tested to avoid introducing a vulnerability that is not apparent. Adequate testing can be expensive, but is important for any infrastructure change.

## **7. Build the directory service in a modular, loosely coupled way.**

*Status: Adopted*

The directory service should be engineered with a bias toward using highly discrete, modular, loosely coupled components.

*A “component” is a part that has a clearly defined usage, function, or purpose and that can be accessed in a defined way without knowledge of its internals. A component can be a hardware part, software, a database table, a software routine, a code module, a server, a database, a system, etc.*

### **Justification**

- This is the corresponding conceptual architecture principle applied to the directory service.
- The directory service will probably require the use of distributed components to improve performance and reliability and to provide for differing types of access. For example, replication could be used to ensure service locally when there are problems elsewhere.
- The directory service will probably use separate systems to provide interfaces to support different types of usage. Examples of usage are “one-stop shopping” access for the most accurate and timely data, and direct access by various types of operational systems for authentication and authorization decisions.

### **Implications**

1. All interfaces between components of the directory service should be message-based where possible to minimize dependencies.
2. Some communications between components, such as for replication, may have to be transaction-based to enable rollback when integrity and consistency are required.
3. Implementation will likely involve using standard parts that enable adding and removing capabilities by adding and removing the parts.
4. Access control that requires authentication should be based on group membership so that access can be changed without changing the access control list.

## 8. Seek simplicity balanced with other needs.

*Status: Adopted*

The directory service architecture should contribute first to simplicity in overall campus infrastructure, second to simplicity of use of the directory service, and third to simplicity of its own internal structure. The directory service architecture should also balance simplicity in design, operation, and usability with requirements for functionality, responsiveness, and flexibility.

Example 3. A university chose to organize its directory to place its public information in a separate partition that it then replicated to a separate “border” directory, even though doing so increased the complexity of its directory service. The border directory was placed where it would be publicly accessible, the master directory and its replicas were placed where they would receive the best protection the university could provide, and the internal systems were pointed to the master and its replicas. This approach achieved a large increase in the security of the internal directory service and thus of the entire infrastructure that depended upon it at the cost of a small overall increase in the complexity of the directory service and the infrastructure.

### Justification

- There can be tradeoffs between the simplicity of one component of the architecture and the simplicity of the whole. This is especially true of the directory service, since its role in authentication and authorization can lead to operational dependence on it by many other systems. The simplicity of the whole is of larger consequence and thus is favored over the simplicity of the service itself.
- Simplicity of use is given priority over internal simplicity to encourage wide use of the directory service for the good of the whole.
- Internal simplicity of the directory service contributes to its maintainability. Simplicity of design and operation improve reliability by making it easier to anticipate the effect of changes.
- There can be tradeoffs of simplicity in design, operation, and usability with requirements for functionality, responsiveness, and flexibility. Additional functionality and especially additional options can increase complexity. Additional replicas can increase responsiveness and provide headroom to handle unexpected demands (a kind of flexibility) at the expense of additional complexity, depending on how they are deployed and managed.

### Implications

1. Use no more replicas than needed.
2. Manage the addition of attributes. A proliferation of attributes leads to an increase in difficulty in working with the whole. The ability to add an attribute needs to be restricted to avoid abuses (such as by adding an attribute called “Contents-my-C-drive”). However, parsimony in attribute creation should not be allowed to lead to combining distinct uses into one attribute. Such “overloading” of an attribute increases complexity by requiring additional logic to distinguish the type of use.
3. engenders its own complexity through unnecessary dependencies between the entries.
4. Limit the number of interfaces, but provide enough to simplify the whole infrastructure. Architecture governance will be needed to control demand for additional interfaces.
5. Make the Emory-wide directory service easy to use. For example, make people aware of the existence of the directory service, have usage policies in place, and provide documentation and assistance.
6. Minimize the number of feeds (other things being equal). See Example 7 page 12-13.

## **9. Provide defined and reusable methods and interfaces.**

*Status: Adopted*

The directory service should have defined and reusable methods to add, update, search, bulk load, extract, and otherwise manipulate its data. Access to these methods should be provided through defined and secured interfaces. The directory should seek to minimize the number of interfaces that it must support.

### **Justification**

- Reusable methods and standard interfaces improve responsiveness and reliability.
- Fewer supported interfaces make it easier to make needed changes faster.

### **Implications**

1. When a relational database is involved, stored procedures provide a way to create reusable methods.
2. Dynamic groups can provide reusable searches. See principle 2.

## 10. Standardize for interoperability.

*Status: Adopted*

The Emory-wide directory should use standard interfaces, protocols, names, and values that enhance interoperability with systems and clients in common use within Emory and the higher education community.

Example 4. Email clients that can look up address information in a directory depend on the directory service using a standard interface and protocol, and on the use of standard attribute names to identify common directory information such as first name, last name and email address. A standard format is assumed for the email address.

### Justification

- Other systems and clients depend on standard interfaces and protocols to access the directory, and on standard attributes to identify the data.
- Emory's attributes, values and distinguished names will need to conform to standards set by standards bodies that represent communities with which Emory wishes to interact. Examples are the Internet community, the Higher Education Community, and Federal Agencies.

### Implications

1. Emory will need to track and conform to relevant standards, such as those of the Internet Engineering Task Force, EDUCAUSE, and Internet2.
2. Emory should maintain metadata about the objects in the directory. It should adopt a standard for the metadata attributes and their values that is compatible with the plans of relevant standards bodies.
3. Emory will need an on-going mechanism for deciding attribute choices, what resources can be represented in the directory, what names and fields can be in the directory, what standards will be enforced on the values of the fields, and other aspects of the structure of the directory.

## **11. Use industry standard solutions where feasible.**

*Status: Adopted*

The Emory-wide directory should to the extent feasible use software and hardware that are vendor supported and widely used in Emory's peer communities (such as healthcare and higher education). The software should allow needed customizations in such a way that they are supported in future versions.

### **Justification**

- Vendor support is desired to provide accountability and a contractual obligation.
- Commonality with peer communities is needed to support the goal of collaboration within those communities.
- The ability to customize a product is often a selling point. Future versions of the product must support needed customizations.

### **Implications**

1. The vendor will need to provide some scheme for customizations that will be supported in future versions and that reduces the need to adjust Emory customizations as the result of changes made by the vendor.
2. Customizations might be supported as configuration options, leading to a tradeoff between configuration complexity and the amount of customization.
3. The directory service should be able to support relevant Internet standard RFCs, and EDUCAUSE and Internet2 standards.

## 12. Data comes only from a single source of authority for that data.

*Status: Adopted*

Each data item in the Emory-wide directory service should come from a single source that is an authority for that data and from nowhere else.

Example 5. A university had separate databases for data about students and data about employees that were separately maintained without reference to each other. As a result, a person's name could be different in the two databases. The university's directory service temporarily resolved the multiple person names by providing attributes for an HR\_Name, a Student\_Name, and a Preferred\_Name. The HR database was the authoritative source for the HR\_Name, the Student database was the authoritative source for the Student\_Name, and the person was the authoritative source for the Preferred\_Name. Having all these in the directory made finding people easier, since any of the names could be used.

Example 6. A university allowed its students, faculty and staff to list their home phone numbers in the campus directory if they chose to do so. This information came from a database that nightly refreshed the information in the online campus directory. A student discovered she can login to the online directory and update her phone number there. However, her change was overwritten that night by the refresh. Although the student was the authoritative source of her phone number, once she changed the number in the directory, the database was no longer an authority for that information (its value was wrong), although it kept acting as a source for the information.

Example 7. A university created a web-based application so that people could change their preferred email address. Since email addresses were in the online directory for forwarding of incoming email and in a database for dynamically creating mailing lists, an update needed to take place in both systems as soon as possible. The university considered three options to accomplish this and rejected two of them, because they increased the complexity of the infrastructure unnecessarily. It rejected changing only the directory, because doing so would require implementing a database notification capability on the directory system and would increase the number of sources feeding the directory. The university rejected having the application update both the directory and the database, because doing that would increase the number of sources of feeds to the directory. It would also make the application more complex by requiring it to interact with multiple systems and ensure that either both updates occurred or neither occurred. Instead, the university designed the application to update the database and put a trigger in the database to send any update to the directory immediately when it occurs. Since the database already fed the directory, the number of sources feeding the directory did not increase.

### Justification

- Use of an authority is required to maintain consistency and correctness.
- Using an authority rather than an authoritative source allows reuse of existing authorities and the results of resolving data conflicts.
- Using an authority rather than an authoritative source can also reduce the number of sources of feeds and therefore the integration complexity of the infrastructure by allowing use of an existing authority that already integrates and consolidates data from multiple authoritative sources.
- Having a single, logical source is less complex than having multiple sources even if the sources are all authorities and thus provide the same values.

### Implications

1. Existing databases, such as the Emory shared data, can be used as the main source of data for many of the data items in the directory.
2. Resolving conflicting data may be required when the data is maintained in multiple databases each of which acts as though it is the authority for that data.
3. People, places and things can have multiple kinds of names. Thus it might be necessary to have multiple kinds of names in the directory, and each name might come from a different authoritative source.

### **13. The directory service should be event-driven.**

*Status: Adopted*

The directory service should accept events in real time and act on them. It should send events as soon as they are ready.

#### **Justification**

- The directory service is intended to be an authority for the information it contains. Thus it needs to be as up to date as practicable.
- Some uses of the information in the directory, such as status for security purposes, need to reflect changes as soon as possible.

#### **Implications**

1. The directory service will need to have the ability to accept transactions and be updated and changed without taking down production access to the service.
2. Infrastructure will be needed to support exchange of events between services in a standard way. The Integration architecture domain covers this topic.

## 14. Document directory contents and information flows.

*Status: Adopted*

The flows of information into and out of the directory, the meaning and source of each type of thing in the directory, what is replicated and where, and the flows within the directory's components should be documented. The documentation should be done for each type of thing in the directory.

### **Justification**

- Documenting flows of information is needed for maintainability and manageability as the number of things and the number of flows increases.
- Documentation must be done for each type of thing that has a value, since each thing could have a different source of authority from which it comes.

### **Implications**

1. Flow documentation should be kept up-to-date and changed when the flow changes.
2. No thing should be added to the directory until all the metadata for it is in hand.
3. Access control lists (ACL) are included. It is necessary to know the source of ACLs as well as who has access to change what.
4. Things to be documented include groups, schema, object classes, objects, and attributes.

## 15. Provide unique names.

*Status: Adopted*

The directory service must provide each entry with a unique (“permanent”) identifier that does not change. Each entry should also have an additional unique name that can change as long as it remains unique.

Example 8. A university assigned unique permanent identifiers to the entries in its directory using sequential numbers. It also included in each entry for a person the unique identifiers of that person from official university databases (such as HR, student, Alumni, donor) that contained records about that person. In addition, it associated with many of the people their email addresses and the userids they were authorized to use to login to certain systems. While the other identifiers would not change, email addresses and userids could change as long as they remained unique.

### Justification

- A unique name that does not change is needed to allow an entry to be uniquely and consistently located and identified at any time.
- A unique name that can change provides the capability for a user-friendly way to uniquely identify an object, such as a delivery mailbox for email.

### Implications

1. In an X.500-based directory, an entry has two names that identify it. One called the “Relative Name (RN)” is part of the entry and distinguishes the entry from others with the same parent node (called the “superior” node). The other called the “Distinguished Name (DN)” is the list of relative names of the nodes on the path from the root to the entry starting with the root and ending with the relative name of the entry itself.
2. A DN that includes a name that can change (such as the userid) should be avoided, because a change in that name changes the DN. A DN change typically requires deleting the entry with the old DN and adding the entry again with the new DN, which is more complex than simply changing the value of an attribute in an entry.
3. When the unique name is the path through a tree to the entry, the same relative name can occur in different subtrees. This can cause problems if the entry must be moved.
4. When a tree-structured directory is used, and an entry can move to a different subtree, then the path to it through the tree cannot be used as the permanent name, since that path would change if the entry moves.
5. A directory of Emory people should not be organized in a hierarchical scheme that represents the way Emory is organized, because of the added complexity of moving people’s entries when they change jobs or putting their entries in multiple subtrees when they have multiple affiliations. Instead, put entries for people in a separate flat space and indicate affiliations using a multi-valued attribute, because it is easier to change an attribute than to move an entry.
6. Do not use an OID to create unique numbers for entries. Standard usage is only to identify objects (types of things).
7. A large set of numbers could be used to uniquely identify entries.

## 16. Standardize names and their values judiciously.

*Status: Adopted*

Emory resources with names in the Emory-wide directory must have standard names and values (including formats). Standardization of values whose authoritative source is personal should only be enforced to the extent necessary to ensure that the values can be used for their intended purpose. Additional names should be defined and used when other formats or variations must be introduced.

Example 9. A university avoided the confusion of introducing new standard names in its enterprise directory by using existing standard building names from its facilities management system, and department and division names from its financial system.

Example 10. In a directory of people, fields for which a person was an authoritative source included a “plan” field and a personal home page link. The only limitation on the “plan” field was the length and the requirement that the characters all had to be printable. The web link to the personal page was required to have valid syntax and not result in an error when accessed. Requiring that entry and update only be done through a web-enabled application enforced these edit checks.

Example 11. When a university realized that its directory needed to have multiple addresses for people for multiple uses (home address to send grades, local address to send announcements, address to send check stub, etc.), it created a separate attribute for each.

### Justification

- Values and names for Emory resources must be standardized to enable the directory service to provide standard names for designated Emory objects as well as standard information about them.
- Data provided by a person must be validated to enforce a valid format. Examples needing a valid format are email and web page addresses.
- Introducing additional formats to an existing attribute increases complexity and can have ripple effects on systems that are already using the attributes. Introducing new attributes to handle alternative formats allows for peaceful coexistence of variations when each has a separate and distinct use.

### Implications

1. Emory will need an on-going mechanism for deciding what data can be in the directory, the choice of names, the standards that are to be enforced on the data, and what changes can be made to the structure of the directory.
2. When data to be included in the directory already has a standard format, the format of the directory entry should honor the existing format if feasible rather than introduce a new one.
3. Clients used for updates to the directory should enforce standards using edit checks and drop-down menus of standard choices.

## **17. Provide data useable campus-wide.**

*Status: Adopted.*

The directory service should contain or have links to as much data and information as are of Emory-wide interest or can be reused across Emory.

### **Justification**

- This promotes the reuse of data.

### **Implications**

1. A process will be needed to allow people to request additions to the directory.
2. An addition could be a new attribute, object or object class.
3. A process will be needed to decide what will be allowed in the directory and what will not be allowed in the directory.
4. A way will be needed to determine the source of authority of additions.

## 18. Support information sharing.

*Status: Adopted*

The Emory-wide directory service should provide an easy way to share information according to Emory policy.

Example 12. A university's directory of people had attributes for each faculty member that included preferred email address, department, university phone number, link to personal web home page, link to a Curriculum Vitae (CV), and standard areas of interest. Making it easy to search on area of interest and locate the CV helped both on campus and off-campus faculty find and contact potential collaborators.

### Justification

- This principle supports Conceptual Architecture Principle (CAP) B.6 of Document 2.
- Sharing is valuable for the reasons stated in the justification of CAP B.6 (summarized below):
  - Increases the possibilities for information reuse.
  - Increases use of information, which increases the value of the information.
  - Fosters faster and more effective decision-making.
  - Leads to wider review, which increases information accuracy.
  - Increases the effectiveness of external partner relationships.
  - Allows access needs to be met immediately when the needed information is already shared.

### Implications

1. The directory service should be able to store a wide variety of different types of data, including binary data (for pictures, for example), pointers to other directories and data bases, and data describing data.
2. There will need to be data that defines the meaning of the entries in the directory and gives their source.
3. The directory should support anonymous access.
4. The directory service will need to be able to restrict searching according to the authorization of the requestor as mentioned in Principle 1 (page 12-2). See also Example 13 on page 12-22.

## **19. Facilitate access to Emory-wide resources.**

*Status: Adopted*

The directory should include information about Emory-wide resources that is sufficient to help assess a resource's potential suitability for a particular purpose and that tells how to access the resource or who to contact to request access to it.

### **Justification**

- People need to be able to see what is available, decide whether they want to access it, and find out how to access it or request access to it.

### **Implications**

1. A resource can be anything that can be given a standard name.
2. Resources include both IT resources (such as devices and data sets) and non-IT resources (such as buildings and rooms).
3. The information will need to be organized in various ways according to standard attributes that group it into various categories or subjects.
4. To be useful, the information about each resource documented in the directory will need to be kept up-to-date. For that to happen, maintenance of the information will need to be related to the work of a local unit and be part of someone's job.

## **20. Encourage appropriate use of the directory service.**

*Status: Adopted*

Encourage adding data useful to Emory. Centrally bear the cost to create and maintain the directory. Make usage that is encouraged easy and no additional cost, and charge for usage that is allowed but discouraged.

### **Justification**

- Adding data useful to Emory increases the value to Emory.
- Encouragement to add data is needed to overcome inertia and reluctance to commit time and attention to sharing data.
- Covering the cost to create and manage the directory and to encourage appropriate use is needed for the common good, because any perceived cost is a large barrier.

### **Implications**

1. There will need to be an easy process to get a review of what data goes into the directory and whether the requester must pay or not to get it in there. The review also needs to be responsive. The difficulty of the process and length of time to get a decision is part of the perceived cost.
2. To make access to the directory easy, provide awareness of the directory and its policy, make encouraged access no additional cost, and provide supported site-licensed access software. Provide administrators and developers with sample code for implementing their own access.
3. Usage of the Emory-wide directory service needs to be monitored, because the directory service is an infrastructure service, and such services should define and track metrics over time as specified in CAP C.1.b.
4. Monitoring of usage will be needed to identify stagnant (unused or little used) data for possible deletion, and to identify the schools, departments, divisions and other groups getting benefit from the directory by using it.
5. Information that is deemed local will need to be maintained by the local unit in its own directory. NOS directories (such as those from Microsoft and Novell) will have to access the directory service through the standard interface(s) for general access (see section 9). The local administrators would be responsible for implementing that interaction.

## 21. Provide fine-grained access control.

*Status: Adopted*

The Emory-wide directory service should be able to provide fine-grained access control to its contents according to Emory policy.

*"Fine-grain access control" means that access privileges can be as specific as required about what can be accessed, who can access it, and what they can do to it.*

Example 13. A university was able to add home phone numbers and digitized pictures of faculty members to its directory by restricting access to those attributes. Although it allowed authorized university users to search on email addresses and display university phone numbers, it prevented display of pictures and home phone numbers by anonymous users and prevented search on home phone numbers by everyone.

### Justification

- Access control is needed to prevent undesired changes.
- The capability for the access control to be fine-grained is needed to avoid unnecessary inflexibility.
- Even information of Emory-wide applicability can require fine grained access controls as illustrated in Example 13 above.

### Implications

1. The directory will need to be able to restrict access to directory data down to the attribute level.
2. The directory should be able to base access control on IP address and or domain name in addition to user identity.
3. A way will be needed to authenticate systems that supply updates to the directory.
4. Care in the use of fine-grained access control will be needed to avoid creating a directory security environment whose management is unnecessarily complex.
5. Creating many access control exceptions can lead to access control lists that are so complex that they lead to configuration errors and so long that they impact performance. To avoid this, access controls will need to be general rules rather than being specific to a particular entry.
6. Distributed administration of data in the enterprise directory is distinct from a locally controlled directory and is a matter of access control administration that is outside the scope of the directory service architecture.
7. Access control that requires authentication should be based on group membership. Doing so simplifies the access control lists and reduces the amount of change to them.

## **22. Support the needs of security.**

*Status: Adopted*

The Emory-wide directory service should support the needs of the security architecture—especially access control—according to Emory policy.

### **Corollaries**

- a. The directory service should be able to provide data for Emory-wide common security.
- b. The directory service should be able to change as needed to provide the required level of responsiveness and availability.

### **Justification**

- The directory service can be a place to which people and systems refer when making access control decisions.
- The directory will need to be able to scale as more systems make use of it.

### **Implications**

1. The directory service might need to use multiple replicas to provide adequate availability and responsiveness to support direct access by operational systems, such as for authentication and authorization.
2. Providing proper security at the right level at the right time implies having the right number of replicas at the right locations to provide uniform availability and responsiveness.
3. The directory service will need to be partitioned in such a way that the entries supporting security can be given the strongest security available at Emory.
4. Distributed administration of data in the enterprise directory is distinct from a locally controlled directory and is a matter of access control administration that is outside the scope of the directory service architecture.
5. The enterprise directory and all its replicas will need to be centrally managed and all use the same schema, which will also need to be centrally managed.
6. The ability to list by user the access privileges that are granted by the directory's access control lists is needed to satisfy the auditing requirement of security.

## **23. Be highly secure and interoperable with security layer.**

*Status: Adopted*

The directory service should be highly secure. It should be interoperable with the common security layer and use it for its own security.

### **Justification**

- This architecture envisions the directory service as part of the Emory IT infrastructure whose purpose is to help secure other IT assets.

### **Implications**

1. The directory will need to support password policy management, such as keeping password histories and controlling minimum and maximum password lengths.
2. The directory service will need to be able to communicate securely where appropriate, such as by use of an authenticated, encrypted connection. Example uses are for loading passwords, replication, password synchronization, authentication, accepting requests, sending results, and otherwise transmitting or receiving sensitive information.
3. The directory service will need to be partitioned in such a way that the entries supporting security can be given the strongest security available at Emory.
4. The underlying platform and cross-platform communication will need to be secure.
5. Good procedures will be needed for changing accounts and privileges.
6. The directory service will need to be able to easily and quickly add support for new security methods and technologies.

## 24. Secure the directory data.

*Status: Adopted*

Everything in the Emory-wide directory must be classified for access control, be secured accordingly, and have a documented owner and source of authority.

Example 14. A university secured its directory data by first classifying everything in the directory based on policy, privacy and application requirements. Access control that required authentication was based on membership in a group to reduce the complexity and the frequency of changes to the access control lists. Students were able to opt out of being visible in the directory by the presence of a “Buckley” attribute in their entry. To allow hidden entries to still receive services, each application authenticated to the directory using its own identifying distinguished name. These application names were members of a group that was allowed to look at almost every attribute in almost every entry.

### Justification

- The directory service provides information in support of the security infrastructure, and thus must be highly secure (security architecture principle C-4).
- Having an owner is required by security architecture principle A-4.
- Documenting the source of allowed updates helps prevent unauthorized updates. The requirement to have a unique source of authority is directory service principle 12. Documenting that source also contributes to documenting the information flow (directory service principle 14).
- Entries, objects, attributes, object classes, schemas, etc. are all included, because both their instances and their definitions are all subject to change through the various interfaces to the directory.

### Implications

1. Everything in the Emory-wide directory will need to be classified using the Emory standard security classifications.
2. A way will be needed to enforce classification and capturing of needed information about anything added to the Emory-wide directory.
3. Access control will need to be defined for everything in the directory and implemented accordingly.
4. The directory will need to be able to base access control that requires authentication on membership in a group, since group membership will likely be the basis for access control in general.
5. The directory service should be able to restrict use of specific directory operations, including read, write, search, and compare applied to a particular entry, individual attributes within an entry, or in the case of a tree-structured directory, to a subtree of an entry.
6. Entries that must have different security treatment would need to be distinguished using an attribute.
7. The directory will need to be partitioned in such a way that everything in the directory that is classified receives the protection associated with its classification, such as location in the appropriate zone of trust.
8. Distributed administration of data in the enterprise directory is distinct from a locally controlled directory and is a matter of access control administration that is outside the scope of the directory service architecture.
9. The enterprise directory and all its replicas will need to be centrally managed and all use the same schema, which will also need to be centrally managed.

## **25. Do not outsource the directory service.**

*Status: Adopted*

The Emory-wide directory service must be connected directly to the Emory network and supported by Emory staff. The support staff should be available to respond at any time day or night. Only Emory staff may do maintenance of the operating systems on platforms where portions of the directory service run. Under these conditions, maintenance and operation of the hardware and support of the directory service software may be outsourced.

### **Justification**

- The Emory-wide directory service must rapidly adapt to support the needs of applications that affect Emory's distinctiveness and to address changing needs critical to Emory's future. Thus it should not be outsourced (CAP C.3.c).
- Indeed, the directory service is expected to be key to security and to contain sensitive data. In addition, the management of its contents requires knowledge specific to Emory.
- The need for the directory service to be directly connected to Emory's network and have staff onsite during normal business hours is due to the service being particular to Emory; the sensitivity of the data; and the extent to which other operational services will depend on the directory service.
- Certain maintenance can be outsourced, because hardware maintenance is already outsourced, and the directory service software would be vendor-supported (principle 11).
- Maintenance of the operating systems used by the directory service may not be outsourced, because of the reliance of many directory server software packages on the operating system (for security functions, for example).

### **Implications**

1. The equipment could be located in a data center that is run by non-Emory personnel, provided it is connected to the Emory network and Emory staff maintain the equipment's operating systems.
2. Emory must develop core competency in directory services (see principle 26).
3. Some support staff may be required to be on site if the security classification and protection of the directory service equipment does not allow remote administration.

## **26. Develop directory competency of Emory staff.**

*Status: Adopted*

Develop Emory staff competency in the skills and knowledge to use, support and evolve an Emory-wide directory service.

### **Justification**

- Staff competency should be developed in areas of strategic importance (CAP C.2). By externalizing data about the infrastructure, the Emory-wide directory service becomes of strategic importance as explained in the following bullets.
- The Emory-wide directory documents and parameterizes certain aspects of the infrastructure, thereby making those aspects easier to understand and to change quickly.
- The directory will be used by most components of the infrastructure to avoid separately maintaining private stores of information of wide interest, such as for authentication and authorization.
- Information in the directory also promotes reuse by making it easy to find and use resources.

### **Implications**

1. Training is needed in skills related to use and support of the directory service for those:
  - supporting the Emory-wide directory service,
  - supporting local directories that use the Emory-wide directory service,
  - supporting systems that use the Emory-wide directory,
  - supporting others in their use of the directory service,
  - entering, changing, and looking up information in the Emory-wide directory.
2. Staff will need to stay abreast of standards, technologies and product updates related to the directory service.

## **27. Manage the evolution of the directory.**

*Status: Adopted*

The evolution of the Emory-wide directory should be planned and governed across the enterprise, with at least a yearly review at a point in the budget cycle that allows its projects to seek funding.

### **Justification**

- Campus needs and available technologies change continually, so planning needs to be ongoing and the directory needs a regular review.
- The directory review needs to happen to allow for funding of enhancements.
- There must be a way to make changes to the directory and to its support.
- Establishing an Emory-wide directory takes time and involves considerable coordination and collaboration.
- Changes to the directory must be well thought out. Good change requires collaboration and collective planning.

### **Implications**

1. There will need to be a process by which the directory is developed.
2. Processes will be needed to establish policies, procedures, priorities, principles, product standards, and configurations, and to manage the directory organization and its contents, including approving changes.
3. Cooperative oversight that is inclusive and representative of the diverse perspectives of all components of Emory will be needed to manage the evolution and direction of the Emory-wide directory service. Managing the direction involves assessing whether the directory service meets business needs and is strategic.
4. Managing day-to-day policy issues is considered to be part of operating the service and as such is distinct from managing direction.
5. To keep the cost of evolution and change under control, the components of the directory must have affordable exit costs. In particular, to avoid the cost of lock-in, it must be possible to load information and business rules into and dump them from components as applicable.
6. The directory architecture will need a steward to be the champion for ensuring that strategic and operational needs of the directory are met.
7. The enterprise directory and all its replicas will need to be centrally managed and all use the same schema, which will also need to be centrally managed.

## **28. Manage the Emory-wide directory as an Emory asset.**

*Status: Adopted*

The Emory-wide directory should be managed as an Emory asset by one of Emory's enterprise IT organizations. In addition, it should have a steward, and to the extent consistent with Emory's security policy, it should allow for remote, highly automated management and make its documentation accessible via Emory's Intranet.

### **Justification**

- Emory's operations, strategic initiatives and programs will increasingly depend on its Emory-wide directory.
- The Emory-wide directory is an Enterprise-wide IT resource that needs to use extremely reliable products and methods, and be managed, maintained and upgraded for the common good.
- A steward is needed to ensure that the Emory-wide directory receives the attention it needs to be successful.
- The directory requires active management. Being able to manage it remotely and automate the activity to the extent allowed by security is needed to control costs.
- Emory-wide directory documentation is a resource of Emory-wide interest, and is needed for the directory's effective management.

### **Implications**

1. Entries will generally have varying availability needs. For some uses, such as authentication, authorization, and email address lookup, the Emory-wide directory service will need to provide continuous availability (24 x 7).
2. To provide continuous availability, the Emory-wide directory will need to have features such as hot swapping of hardware components, load sharing and balancing, and redundancy. It will also need to have service level agreements for continuous availability from those infrastructure services (such as the network) on which it depends.
3. Availability for update might not have to be as high as availability for read depending on how up-to-date the data must be.
4. Availability requirements can affect whether to put an entry in the directory or just a pointer to it.
5. To enable cost control, the entries in the directory service will need to be classified according to their availability requirements. Then appropriate strategies such as stratified availability and replicas can be used.
6. The Emory-wide directory will need to provide data and error reports, and permit active probing and management.

## **29. Manage the Emory-wide Directory in a unified way.**

*Status: Adopted*

Emory's IT departments and unit leaders should have a common vision and understanding of what it means to have an Emory-wide directory, and there should be a process to implement and enforce that vision across Emory. This is needed even when IT responsibility is decentralized.

### **Justification**

- A common vision and understanding simplifies governance. It makes it easier to make decisions for the good of the whole organization, especially when such decisions are not optimal for a particular unit.
- The customer base is the entire Emory community. A unified approach, as opposed to just a centralized approach, helps to accommodate the diversity of perspectives present in a diverse organization. Broad representation enables decisions to take into account the full context and implications of the decisions.
- An Emory-wide view helps to identify situations where an Emory-wide directory could potentially add value. Examples include situations where process and data sharing occur across Emory.
- A unified approach also facilitates a systemic, consistent, Emory-wide view of the data in the directory and of the meaning of those data items.

### **Implications**

1. A unified approach will require an organizational structure that enables cooperative decision-making and cooperative enforcement.
2. Organizationally, unification will have to represent the perspectives and needs of many constituencies informally or formally, including Emory-wide, unit, and central IT.
3. A governing body that is inclusive and representative of all the components of Emory will be needed.
4. There will need to be a process and organizational structures to establish policies, priorities, principles, product standards, and configurations, and to manage the directory organization and its contents, including approving changes.