HARVARD UNIVERSITY

Enterprise Architecture
Enterprise Technical Architecture Board

Recommendations for the Deployment and Management of Directory Services

Editor: Greg Charest	**Audience Level:** • Strategy Planning and EA Leader • Solution Architect and Program Manager
Version: 2.2 **Last Revised:** 10-Jan-2019 **Status:** Draft **Document Type:** Guidance	**Distribution Scope:** CADM
ETAB Workgroup Members: Richard Borroff Greg Charest Tim Gleason Harry Hoffman Rob Ruma Raoul Sevier Michael Thomas	**Reviewers:** Greg Charest Raoul Sevier Jane Hill Harry Hoffman

1. Preamble

Enterprise Architecture recommendations and opinions are intended to provide middle to long term guidance in key technology areas. They must be considered in terms of:

- Service to business needs
- Grounded in strategic technology goals as expressed by the senior IT leadership
- Cognizant of related costs
- Flexible based on changing technical and business environments

2. Topic Statement:

Within Harvard University, multiple directories have been deployed, many of which now have overlapping services and data, inconsistent provisioning, and varying levels of support. Most are on-premises while other IT directory resources are moving to the Cloud. This work proposes a strategy to rationalize the current state.

3. Executive Summary:

Directory services are a well-established, standardized way of providing authentication and authorization tools as well as person and resource attribute information. These services are accessed using the LDAP protocol, as implemented by several UNIX-like LDAP server solutions as well as Microsoft's Active Directory solution. Some of the traditional uses for LDAP based directory services are being superseded by more modern and manageable services. This document outlines a mid-term future-state vision, strategy and design approach for directory services across the University.

4. Key Recommendations:

- Use one master directory dataset for enterprise person/resource data.
- Continue use of both UNIX-like LDAP and Microsoft Active Directory (University AD) implementations to meet University enterprise needs but consolidate over time to an AD implementation for enterprise directory services, including HarvardKey.
- Create a HUIT Managed Directory Service offering for organization-specific data as an alternative to self-managed directories.
- Deploy local directories only where explicitly needed by schools/organizations.
- Synchronize between enterprise, managed, and local directories using automated provisioning services.
- Consolidate LDAP skills into one CA organization for critical mass and coverage.

5. Introduction:

The FY18 HUIT Top 30 Strategic Goals[1] document recognized the need to define a strategy, service approach and architecture for LDAP directories. Although this goal was subsequently put on hold, the FY18 University IT Strategic Plan[2] identified a need for directory strategies to support the CIO Council Collaboration initiative, enabling this work to continue. Other key drivers for this effort include:

- Concerns for the security of the stored data and malicious use of directory servers as an access point to other resources.
- The availability of alternative tools and services that replace the roles of some legacy directories.
- Impeded Cloud migrations due to incomplete understanding of directory roles and structures.
- The existence of several unmanaged and/or poorly supported directories.
- Duplicative and inconsistent data models that are provisioned by programmatic and manual means.
- The inability to retire legacy capabilities and data such as found in FASLDAP.

6. Discussion:

6.1. Existing Harvard Directory Services Managed by HUIT

A number of directories that are candidates for rationalization exist within HUIT. These include:

- **University AD (UNIVAD)**
  An effort is underway to provide Active Directory as a Service in order to increase security, streamline user experience and reduce the administrative burden required to operate the large number of AD instances currently used across the University. Although increased resiliency to attack and fewer AD related security incidents is a primary goal of the effort, the project includes the explicit objective of building a single University Active Directory for all schools and centers. Within the Enterprise service a School/Center will retain autonomy to manage Groups, Computers, Application access, and Group Policy for their organizational units on a delegated basis.

- **HarvardLDAP**
  The primary enterprise UNIX LDAP directory, HarvardLDAP, holds HarvardKey credentials and attributes for people (HUID holders). HarvardLDAP is one component of Harvard's Enterprise Identity and Access Management (IAM) solution. This directory has been migrated to the Cloud.

- **PublicLDAP**
  PublicLDAP holds person attributes that are accessible to public and Harvard applications.

- **FASLDAP**
  FASLDAP supports a wide variety of directory functions for the Faculty of Arts and Sciences. A number of efforts to analyze the type and volume of FASLDAP usage have been undertaken in

the past. Attempts to consolidate the data into other directories and retire FASLDAP have been stymied by lack of documentation and multiple ownership of the internal resources.

6.2. Other Harvard Unix LDAP and Microsoft AD Directories at Schools and other organizations

In addition to HUIT managed directories, organizations such as HBS, Research Computing and HMS all maintain their own LDAP directory instances with person, group and technical attributes necessary to support local needs.

The content of these directories may be candidates for promotion to the Enterprise or managed directories. HUIT should work with the owners of these directories to identify current requirements. Adding local organization data to the Enterprise or managed directories may be sufficient to support retirement of the local directory.

6.3. Evolution of Directory Services

Directory services have been traditionally used to:

1. Provide information about
    a. people, including unique identifiers and additional attribute data,
    b. technical resources such as printers and file shares, including their unique identifiers and additional attribute data,
    c. groups, including group identifiers and membership.
2. Determine authentication to ensure valid identity.
3. Determine authorization to access resources using attributes and groups.

Harvard continues to use LDAP[3] protocol-based services but is also introducing new services that in some cases replace LDAP functionality, and in others supplement or hide LDAP implementations.

- **HarvardKey Authentication**
  At Harvard, authentication has been abstracted into higher level service, HarvardKey. HarvardKey supports a "single identify for life" that can be used to access a variety of University applications and services. In addition, HarvardKey provides improved security features including two-factor authentication and single sign-on to over 400 applications. The implementation of HarvardKey includes an LDAP server, HarvardLDAP. This implementation is a candidate for migration to the AD-based UNIVAD directory service. This will improve security, reduce duplication of data, and simplify the infrastructure by reducing the number of supported instances.

- **Groups and group membership**
  Groups and group membership information is now provided by the Identity and Access Management[4] team through the use of Group Services. Group Services is fully integrated with HarvardKey and automatically updated based on changes to a user's status and role. It also supports the use of local groups and an associated delegation of group administration. Some group information may be provisioned to LDAP services, but on a selective basis.

- **Authorization**

The combination of HarvardKey and Group Services can be used to manage permissions to applications and websites. Individual applications may also provide this functionality directly. In general, new applications and existing application planned for significant upgrades should utilize newer tools and services such as HarvardKey and Group Services. These new capabilities, along with operational and cost considerations, are defining a new architecture for directory services.

Until they are retired or upgraded, legacy applications that only support the LDAP protocol will require the continued use of LDAP directories.

- **API Services – The Person Data Service**
  A key goal of HUIT's Data Management Services[5] (DMS) organization is a reduction in the cost, time, and effort required to share information by simplifying the way schools and central units exchange data. One of the most commonly exchanged data sets is information about people; information traditionally housed in LDAP servers.

  The Common API Platform Program (CAPP) has designed and implemented a 'Person' API that can be used to retrieve common attribute data about Harvard people. This service follows a well-defined governance process and is the preferred method to use when applications need person related information.

- **IDaaS – Identity as a Service**
  Identity as a Service are SaaS-based service offerings that allow organizations to use single sign-on authentation and access controls to provide secure access to their growing number of applications. Trends in the industry point to this approach as the long-term solution. Much of this proposed Directory Strategy prepares the ground for leveraging this capability.

7. Recommendations:

General Recommendations and Rationales

7.1. Continue to provision from one enterprise dataset for person data *- The enterprise dataset is derived from the IAM Identity registry and holds the essential person information that is used by the enterprise.*

Use of the single IAM Registry-based reference dataset avoids the following problems:

- Data silos and multiple versions of data.
- Data errors that result from manual entry and maintenance.
- Dated timeliness, including not knowing which data elements can be trusted.

At a basic level, master data management methods seek to ensure that an organization does not use multiple, potentially inconsistent versions of the same data in different parts of its operations. Master data is the consistent and uniform set of identifiers and extended attributes that describes the core entities of an enterprise. A single master dataset is necessary to provide accurate and

timely information on Harvard people and resources. Coherence in the data stored across all directories must be established, maintained and synchronized through automated means.

7.2. Two enterprise directory implementations merging into one over time – *Continue to use one UNIX-based LDAP and one Microsoft AD, which is necessary to meet the current set of enterprise directory requirements. Over time, consolidate to one AD implementation for enterprise directory services, including HarvardKey.*

The current design of two enterprise-level directory implementations, one UNIX-based and one MS-AD-based, should be consolidated over time to reduce duplication of data and simplify the infrastructure by reducing the number of supported instances. A progressive approach is recommended, starting with migrating the UNIX authentication services to AD, followed by migration of person attribute information and then the remaining non-person data. At that point HarvardLDAP should be retired.

7.3. Managed Directories - *HUIT should create a directory management service to allow schools and centers with existing UNIX-like directories the option of delegating the administration and operation of such directories to a central team with the level of technical expertise and resources required.*

The decentralized nature of Harvard often results in high levels of duplication and complexity. The shared-services approach can combine the advantages of centralization and decentralization, achieving economies of scale and scope while remaining responsive to user needs. A directory service offering would allow some parts of the organization to reduce their operational level of support for LDAP directories while meeting existing local needs.

7.4. Local directories - *Local, non-managed directories should be limited to those that provide application-specific or locally valuable directory services. These typically provide directory information in narrowly-scoped environments such as individual or small sets of applications.*

The need for local directories that support specific applications and/or contain locally valuable person/resource data will continue to exist. Local data should be limited to those objects and attributes not contained in master or managed directories.

When differences exist in the business meaning of a particular attribute or set of attributes, local directories should extend the schema used in the master directory and avoid using the same attribute name(s) for locally different definitions, enabling synchronization with authoritative data at the higher level.

7.5. Synchronization of directory data - *To the greatest extent possible, all directories should be automatically provisioned/deprovisioned using a top-down approach beginning with the IAM program databases and systems.*

Consistent and timely provisioning and de-provisioning of information in directories is required for reasons that range from supporting a high-quality user experience for new employee on-boarding to rapid and secure removal/archiving of accounts for terminated employees.

Achieving this across multiple directories is almost impossible absent automated tools and standard processes. Provisioning of person information must be accomplished through a business service that automatically adds people to the directory. Manual addition of people should not be allowed. De-provisioning should be accomplished through the same business service. Provisioning and de-provisioning of technical resources should be accomplished through a similar business service.

7.6. Directory management - *Consolidation of LDAP skills into one organization for critical mass and coverage.*

Directory design, management and administration, for both Active Directory and UNIX-like LDAP directories servers, requires a unique set of skills. The University would benefit from an effort to identify and consolidate, formally or informally, staff with this experience in order to reduce duplication of effort and provide quality, sustainable services.

8. Additional Considerations

This document does not address the operational and cost considerations related to the proposed design and strategy. These would need to addressed prior to implementation of these recommendations.

Directory services should be aligned with the University's Enterprise Architecture principles, particularly the infrastructure and security principles.

9. Recommended Changes and Outcomes

Change Area	Expected Outcome	Change from Current State
Fewer directories across the enterprise	• One enterprise Directory dataset maintained in the IAM registry database • Consolidation of enterprise directories to one AD instance over time • Formal management of organizational directories on request • Guidance and standards for all directories	• Retirement of unneeded Unix LDAP and AD instances • Fewer School and organization managed directories
Simpler and more secure deployment pattern for AD	• Use implementation model developed by HUIT Security with Microsoft	• Fewer inconsistent implementations • Increased enterprise-wide directory security
Better support for managed directories	• Single point of operation and accountability	• Align with consistent and standard

		practices – shared services • One operational cost model
Improved physical and logical data model	• Enterprise-appropriate data in the Enterprise directory • Mission-specific data in the Managed and local directories	• Different collections of data stored in various directories are rationalized and streamlined • Duplication across directories is minimized • Enhanced data quality through automated synchronization
Better synchronization between different directories using provisioning services	• Single data provisioning engine provisions Enterprise and all Managed directories	• Elimination of multiple, ad-hoc mechanisms for provisioning
Consolidate LDAP skills into one organization for critical mass and coverage	• Single, critical-mass organization to manage Enterprise and all Managed directories	• Fewer organizations managing directories with better operational skills

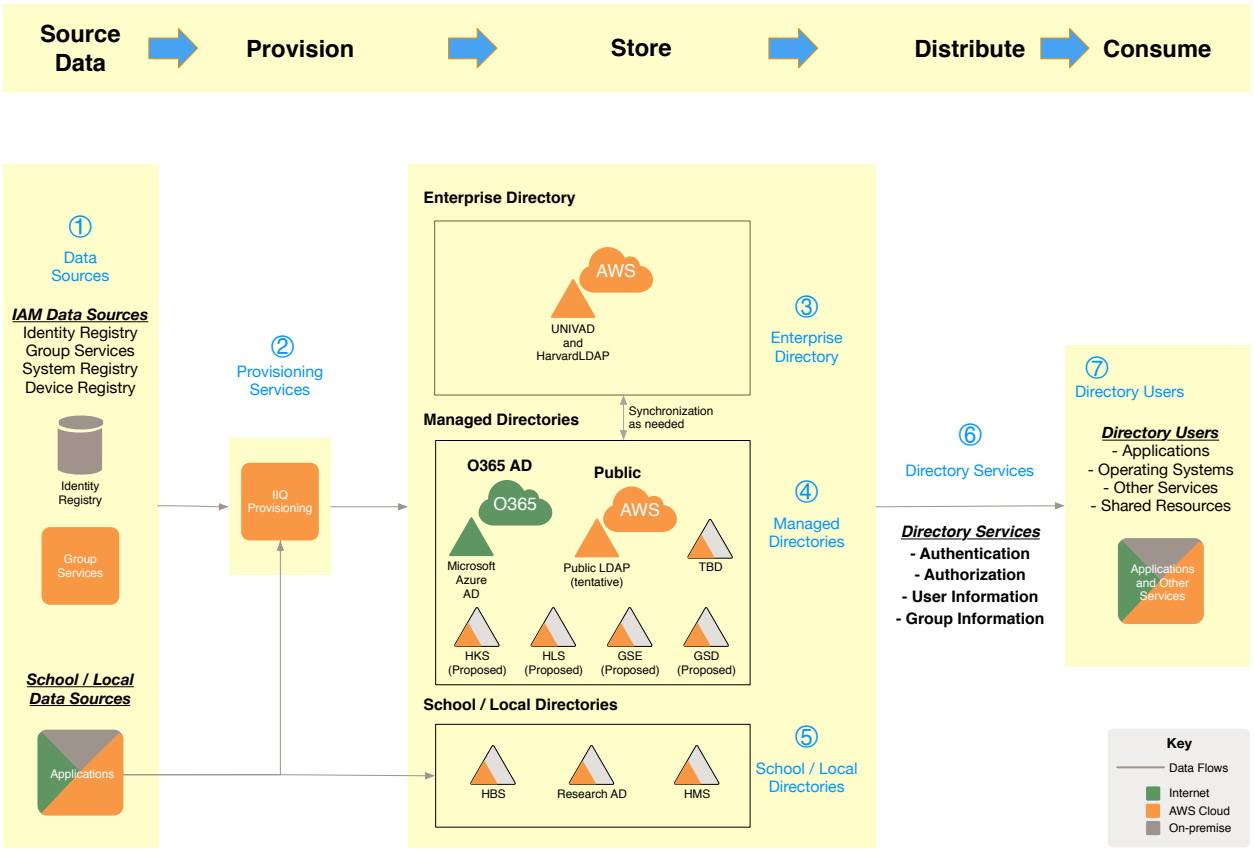
10. References

1) FY18 Top 30 Goals - https://huit.harvard.edu/files/huit/files/fy18_top30_goals-handout-v2.pdf
2) University Information Technology Strategic Plan, 2018 https://huit.harvard.edu/files/huit/files/final_2018_itstrategicplan.pdf?m=1527778654&utm_source=SilverpopMailing&utm_medium=email&utm_campaign=2018_IT_Strategic_Plan_rollout%20(1)
3) Lightweight Directory Access Protocol - https://en.wikipedia.org/wiki/Lightweight_Directory_Access_Protocol
4) HUIT Identity and Access Management – https://iam.harvard.edu/
5) HUIT Data management Services - https://huit.harvard.edu/data-management-services

11. Appendix A

The following diagram provides a general overview of a future state directory environment that includes Enterprise, managed and local directories and the associated data lifecycle from creation to consumption.

Harvard Directory Services – Future State Conceptual Lifecycle



Narrative

Harvard's Directory Services are distributed across enterprise and local organizations. Enterprise directories are in the service to all Harvard organizations. School / local directories are limited in their use, and may interact with enterprise directory services. Directories are accessed by many applications for authentication, authorization and directory information look-up purposes.

Key Features

1- Data Sources
Data for Directory Services originates from both enterprise and school / local sources. Most enterprise directory data come from IAM's Identity Registry capabilities which span all identity types across Harvard, including information about people, groups, and systems.
Enterprise data is sometimes stored in separate managed directories which are needed because of application design or security requirements.
School and local applications provide most of the data to school and local directories, under the management of the local organizations.

2- **Provisioning Services**
The IAM organizations operates provisioning services, using the IIQ product, that manages the data in the Enterprise directories, as well as more focused directories that are centrally managed.

3 - Enterprise Directories
Data from enterprise and other sources is provisioned into two primary Directory Service repositories; Active Directory (AD) and UNIX Lightweight Directory Access Protocol (LDAP). These are used to supply authentication, authorization, and directory data to Windows and non-Windows clients and applications. The primary instances of these repositories can be extended by the addition of specialized data structures or instances. Over time, these will be consolidated into one AD instance.

4- Managed Directories
Some directories are needed for specialized purposes, but are inherently enterprise in scope. These are provisioned and managed centrally. Others are needed by Schools or other organizations but are managed by a centralized Directory Service organization.

5- School / Local Directories
Schools continue to operate local directory services which are provisioned from local applications, but can also be provisioned by the central provisioning engine, and integrated with the enterprise Directory Services through standard trust and synchronization means.

6- Directory Services
Directories implement standard protocols and services that support 1) authentication, 2) authorization, and 3) retrieval of user, group, and other data for use by applications.

7- Directory Users
The vast majority of Directory Services interactions are read-only, by applications that need authentication, authorization, or directory data lookup services.