Functional Model (Identity Registry)

This document provides a general description of the components and functions of the identity registry component of an institutional-scale Identity and Access Management (IAM) suite. It also suggests touch points with other subsystems in such a suite. Requirements for identity registry functionality and operation can be written based on the terms and concepts presented in this model.

Overview

The function of an identity registry is to register and maintain information about entities of interest to the organization operating the registry, and to make this information available to other systems. This model is concerned with identity registries serving institutional needs: containing thousands or millions of entities, operated according to institutional policies to meet institutional goals such as accountability, compliance, security, and collaboration.

Entities, entries, identity, identifiers

An entity is a "thing" of interest to the institution, distinguishable from other entities of its type. Entities of most interest for an identity registry are typically "actors", i.e. things that initiate actions in online systems. The most common type of entity is a person, hence identity registries are often called person registries. Other common "actor" entities are processes, applications, computers, and organizations. An entity is represented in the identity registry by a record called an entry that contains structured information about the entity. Some of the data describes the entity; this is identity data. Other data, such as entry create time or access control data, is registry metadata. A data element that is designed to distinguish entities in a set is called an identifier. An entry typically contains several kinds of identifiers, as well as other data about the entity. A key goal of a registry, typically, is to ensure, as much as possible, that each entity is represented by exactly one registry entry. Each entry in a registry has a type, and each type has a schema. Different types may be handled by different registries, or a single registry may deal with several types.

Registry-managed identifiers

In addition to managing entity data sourced from various business processes, identity registries typically are source systems (i.e., are authoritative) for some data, in particular institutional identifiers. A common registry-managed identifier is a registry ID (also called unique ID, or UUID) that is an opaque non-reusable identifier serving as an institutional "key" for the entity. Another common registry-managed identifier is a network ID (also called NetID or username) that is used by end-users for login and other services such as email. Creation and management of NetIDs, and other similar identifiers such as Distinguished Names, may be integrated with credential assignment processes that also include management of authentication information such as passwords and public keys. This is a touchpoint between identity registries and authentication systems.

Registration, matching, reconciliation

Registration (also known as enrollment) is the process of creating a new identity registry entry. Identity data may come into a registry from source systems (which are typically also registries in a sense), or interactively via human entry processes. A person who engages in registering entries is called a registration agent. In support of the goal of one entry per entity, it is necessary for the registration process to determine whether a set of identity data coming into the registry refers to an existing entry, or represents a new entity, hence requiring the creation of a new entry. The process of distinguishing new from existing is called matching. The matching process may rely on many different data elements, and may involve human decision-making in addition to automated processing. The process of adding or modifying identity data in an entry based on incoming data is called reconciliation.

Merging, splitting

It may be found that due to a failure of matching in the registration process more than one registry entity exists for an entity. In this case two or more entries must be merged. Similarly, it may be found that an entry contains a mix of information from different entities. In this case the entry must be split into two or more entries. Merging and splitting are typically administrative processes; in the case of person entries the processes may involve active participation of the affected people.

Identity information distribution

Information in identity registries is made widely available to many processes and systems to support institutional-scale identity integration. This implies a touchpoint between identity registries and information distribution (or presentation) systems such as directories, web services, etc. This may imply requirements for low latency of change propagation, high fanout, etc.

Affiliations, lifecycle

Many different institutional processes bring entity information into a registry. In addition to the entity's type (person, e.g.), the registration process and the information in the entry typically reflect the nature of the process that brought the entry in. For example, the entry for a person who is a student will likely have a different input process and hold different information from that of a person who is an employee (a person may be both, of course). The different relationships that affect entry data and maintenance are called affiliations. The policies and procedures that codify how an entry is managed over time are called lifecycles of the various affiliations. Identity registries may need to support affiliation catalogs, representations of lifecycles and policies, and integration with business systems such as student and human resources systems that implement the lifecycles for key institutional affiliations. Affiliation lifecycles also often affect service access by people and other actors, implying a touchpoint of affiliation and lifecycle management with access management systems.

Contact / profile information

A common class of identity data is contact information, or more generally profile information. Contact information includes items such as phone numbers, email addresses, web URLs, etc. Profile information may include many types of information including departmental associations, interests, and other relatively unstructured information.

Identity assurance

Institutions have a wide range of relationships with people, from rich and long-term to brief and casual. Information about people in an identity registry may be well-managed and vetted by trusted business processes, or it may be self-asserted or untrustworthy, depending on circumstances. This consideration leads to notions of representing these qualities of information, a concept called identity assurance (registry information is only one piece of overall identity assurance). Identity registries may have requirements for storing and managing assurance-related information, such as vetting processes, times/locations of data checking, etc. This also implies touchpoints with other aspects of identity assurance, in particular authentication systems.

Management operations / user access

Identity registries are maintained via many processes; some of these involve interactive access by registry-associated staff to do such operations as status checking, handling merges and splits, investigating data anomalies, viewing entry history, generating reports, etc. This implies requirements for user interfaces to support these operations, including appropriate access control. Other functions may imply a need for end-user access for operations such as profile management, self-service merging, service requests, etc.

Enterprise support services

Workflow, reporting, notification, event/message services, archiving, intrusion detection/prevention, and more