



# Issues in centralized identity management

TF-CSIRT Technical seminar



# Background

- University of Oslo
  - About 7 500 employees (staff and faculty)
  - About 33 000 students
  - About 1 800 PhD
  - About 6 000 “others”
  - Several research centers, hospitals etc
  - Wide research cooperation (both in Norway and abroad)
  - Yearly turnover of about 7 000 people



# Background

- Wide range of user services
  - Authentication services across the IT infrastructure at the UiO
  - E-mail
  - Home directories hosted by UiO
  - Print services
  - Controlled access to diverse administrative systems
  - Wireless internet connection, EduRoam
  - Access to the Norwegian federation (FEIDE) and cross-federated access (Kalmar, EduGain)



# Background

- Automatization and routines for user administration and identity management established early due to
  - The number of users of the IT services
  - The various and steadily growing demands for access control and integration of services into the existing infrastructure etc
- First “almost completely” centralized user administration system established in 1993-94 (the work actually started already in '88 :-))
  - Covered employees only

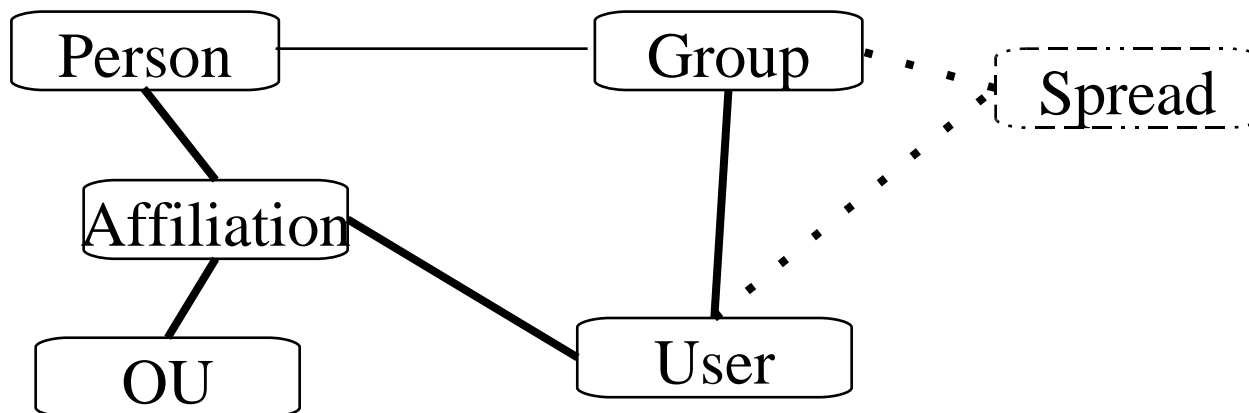


# Background

- General and complete centralized Idm solution from 1997 (UREG2000 – Y2K compliant :-))
- First Norwegian federation ideas emerging and resulting in design and implementation of Cerebrum in 2001 – 2003
  - Mainly to accomplish a generally deployable, reusable Idm solution compliant with the federation requirements

# Cerebrum

- Python and RDBMS based framework
- Covers the object types needed for Idm and user administration (people, accounts, groups, organization units etc)





# Cerebrum

- Supports automatic import and join of person data from multiple authoritative source systems into a central directory (Cerebrum db)
  - Join is a regulated act of resolving information inconsistencies in authoritative systems
- Supports automatic account administration (creation, modification, removal) across available IT services



# Cerebrum core

- Core
  - Core aims to cover the general data model for Idm
  - No dependence to the IT infrastructure of an organization introduced (and therefore easily deployed just about anywhere)
  - No business rules for an organization introduced in the core API but:
    - Extensive configuration possibilities
    - Support for implementation of business rules via python mixin classes and overrides



# Cerebrum modules

- Modules
  - Some organization specific (export of user names and e-mail addresses to the UiOs HR-system)
  - Some target system specific (E-mail module, DNS module)
  - Supporting mixin classes with possibilities for overrides of standard API functions



# A Cerebrum module - LDAP

- UiO LDAP - implementation and use
  - No authentication information is released from LDAP, only encrypted `ldap_bind()` is supported
  - LDAP-implementation at UiO consists of several trees:
    - organization (core, inetOrgPerson, eduPerson, eduOrg, norEduPerson, norEduOrg schema etc)
      - organization structure, people, groups
      - UiOs whitepages and federation backend



# A Cerebrum module - LDAP

- system
  - Traditional NIS information (uid, gid etc)
  - user, group and netgroup trees
- mail
  - Backend for UiOs e-mail system
  - Contains information about quota limits, spam settings and other relevant e-mail data
  - Authentication information for e-mail targets



# A Cerebrum module - LDAP

- All LDAP updates are done by Cerebrum
  - Batch updates
  - Idif files distributed to the LDAP-servers at UiO



# A Cerebrum module - LDAP

- Implemented as
  - Python Factory class
  - Generic methods compliant to federation schema
  - Institution specific export rules implemented through Cerebrums general configuration
- Overrides and institution specific methods implemented as python mixin classes
  - UiO has special rewrite rules for phone numbers, addresses etc.



# Identity management

- Is about a lot of things, among others also:
  - Enhancing control over data flow at an organization
  - Enforcing policies (be they written or merely spoken :-)) through automatization and other appropriate tools
  - Providing timely and correct access level to services
  - Reducing the risk of breach in data integrity
  - Reducing registration related overhead
- But an Idm cannot substitute policy, documentation and decision making processes



# Security issues in centralized identity management

- Very obvious
  - Password control
  - Access control
- Not-so-obvious
  - Privacy issues
  - Data integrity



# Security issues in identity management

- Somewhat obscure
  - Data element interpretation and encapsulation of elements (how does a set of loosely connected elements acquire a meaning?)
  - Non-personal data dissipation control



# Password control

- Single set of authentication data for each user and all services available
- Consistent password algorithm for all IT service provided by an organization (it is actually a single password)
- Consistent password updates for the whole organization



# Password control

- Easily accessible (for legitimate users) single point of update
- Plain text password have to be kept for provisioning of services that do not understand common encryption methods and cannot use other authentication mechanisms
  - A centralized system enables you to automatically remove these when no longer needed



# Access control

- Access granting based on policies and regulations (but for the diverse shortcuts :-))
- Access revoking based on policies and regulations
- Easily accessible single point of update



# Access control

- Supposedly user friendly update methods (still quite hard to achieve)
- Role-based and other “smart” authorization systems within reach



# Privacy issues (ref. EU-95/46/EF)

- Controlled dissipation of person information
  - Easily accessible information about which systems and services require personal information (of less sensitive as well as sensitive nature)
  - Information about dissipation of data about any single person easily obtainable



# Privacy issues (ref. EU-95/46/EF)

- Routines for secure data transfer
- Implemented rules for access restriction to sensitive information (also for privileged users)
- Accountability (we know who dunnit)
- Change tracking (...and we know when and how)
- Auditing (is everything correct and proper?)



# Data integrity

- Data fetched from authoritative source systems on a regular basis
- All relevant data elements updated on a regular basis
- All inconsistencies reported to relevant offices (ie. study advisors or personnel offices)
- Identifier changes handled and dissipated throughout IT systems (if required)



# Data element interpretation

- There is a jungle out there:
  - Who are you?
  - Where are you?
  - Why are you here?
  - What are you doing?
  - Will you be doing something completely different tomorrow?
  - Really, who are you?
- Every single element addressing these questions needs to be interpreted



# Data element encapsulation

- Interpretation of single elements will take you far but often you need more
  - Who are you, where are you and why are you (here) might decide access level to a service or a set of services
  - The core data always stays the same, even if the combination of elements changes (maybe a service does not need to know why you are here)
  - The interpretation of the encapsulation needs to be consistent
  - The data provider must be “the boss” here – even if a service happens to disagree or require different interpretation of an encapsulation



# Data element encapsulation

- Does this remind you of something?
  - Establish a common data model and follow through (basic rule for fully integrated environment)



# Non-personal data quality assurance and dissipation

- This is really a side issue here, but still...
- A lot of information about internal workings of an organization is made available for identity management purposes. At a university for instance you will find:
  - Information about courses, lessons, activities etc
  - Information about employee categories, employment rules, seniority etc
  - Information about rooms (offices), telephone numbers...



# Non-personal data quality assurance and dissipation

- However all used data elements will generate more work
  - Errors are discovered and must be corrected
  - Unused data elements must be populated for new (and existing) services
- The identity management system often catalogs and controls dissipation of non-personal data but:
  - Should this be the case?
  - Should we establish other procedures for both quality assurance and dissipation?