Making Campus Identity Data Work For You

James Drews
College of Engineering
Agenda

• What “Data” are we talking about?
• What campus has available
• Where can I get the data from?
• Things to consider
• Example Use Case: Engineering IdM
What “Data” are we talking about
What Data?

• Centered around People

• Student Data
  • Who is a student
  • Courses being taught
  • Who’s teaching
  • Who’s enrolled
  • Degree Information

• Employment Data
  • Who’s an employee
  • HR related info (title, pay type, etc)

• “Others”
  • Anyone else affiliated with the UW
Data Sources

- Student Information System (SIS)
- Peoplesoft HR
- SpecAuth/SpecPop
- Manifest
Source: Student Information System (SIS)

- What courses are offered?
- Who’s teaching the courses?
- Who’s taking the course?
- Who are the prospective students?
- What degree programs are they in?
Source: Peoplesoft HR

- Who is an Employee
- What Department do they work for
- Who has a specific title code
- Who pays the salary (Funding)
- What pay type (one time, monthly, bi-weekly, etc)
- Tenure Track data
Special Authorizations/Populations

• SpecAuth – Older system, being replaced by SpecPop
• Information about “others”
  • Contractors
  • students enrolled at other campuses (but employed here)
• Used to get others access to campus services (netid, photo id, etc)
Manifest

- Adhoc Groups
- Data derived groups
- Netid guests
- Summer Programs (Example: SURE-REU)
Where can I get the data from?
Infoaccess

• Reporting data warehouse
• Contains Student data
  • Historical
  • Current Semester
• Contains “some” HR data
  • Contains historical data (Pre-Peoplesoft HR)
  • Directory data
  • Recruitment Data
• Contains an “ID Crosswalk” to link between systems
• Direct Oracle DB access
• Campus Query Library
• Updated Daily (not all data)
• https://it.wisc.edu/services/infoaccess/
Curricular and Academic Operational Data Store (CAOS)

- Near real time updates from SIS
- Contains Student Data
- L3 project – Noncredit Learner Data
- SOAP Web Service API
  - Programming data pulls
  - SoapUI application is your friend
- Publish/Subscribe notifications
  - Course Roster Changes
  - Academic Objective (major) changes
HRS-EPM

- Peoplesoft HR Reporting Oracle Database
- Nightly (most often) extracts from Peoplesoft HR
- Contains HR Data
- Peoplesoft EPM Reporting Web based access
- Direct Oracle Access
- https://uwservice.wisconsin.edu/administration/data-warehouse.php
Campus UDS

- Middleware Oracle Database
- Contains data from various source systems
- Has an “ID Crosswalk”
- Focused on “Directory” data
- Only way to get “Manifest” data currently (custom views)
- https://it.wisc.edu/services/iam
Things to Consider
Considerations

• Real time Query vs stored copy
• If stored, must be done securely
• What data is really needed?
• Legal Name vs Preferred Name
• FERPA
Use Case: Engineering IdM
Background

- Developed over 25+ years
- IE, it’s older than most of our students!
- Started back when getting data was “hard”
- Main driver of many of CAE’s systems
  - Engineering Active Directory
  - Linux LDAP
  - Engineering Shibboleth
  - File Storage Access
  - Self-service Group management (for AD / LDAP)
Olden Days

- CAE got a “tape” From ADP
  - Early on, was actually a reel-to-reel tape
  - Later on was a mainframe download
  - Contained mostly students taking engineering courses
  - Difficult to make changes
- Data was parsed to flat files for the two main systems to use (Unix / NetWare)
- Had to activate account (Newuser process) on multiple systems
- Potential of different login names between systems
- Had to re-activate account(s) every semester

This lead to....
Olden Days
Today

- Central IdM Store (Oracle DB)
- Pull data from campus systems daily
- Some near real time updates
  - course registrations now
  - Future – student major changes
- Highly automated
- Easy to add new exceptions/additions
Today

• Activate (newuser) account once
  • Email notifications if account is expiring
  • 30 day grace period
  • Students enrolled in Spring stay active over summer

• Self Service tools (my.cae.wisc.edu)
  • Create/manage groups
  • View rights to systems
  • Request new accounts (faculty/staff)

• Most accounts ready for activation within minutes of being added to source systems

• Expire process started once a day, weekdays

• Database entries purged 1 year after last valid semester
  • Login name becomes available
  • Unix uid/gid becomes available
Today

Oracle Database

CAOS
Direct Pull
Pub/Sub

Manuel Entry

HRS-EPM
Direct Pull
Pub/Sub

Oracle Database

SpecAuth/SpecPop
Direct Pull

Manifest

Direct Pull
Account Lifecycle

- Data Import
- Newuser Run
- Accounts Created On Systems
- Source Data Ends
- Warning State
- Account Expires
- Account Removed From Systems
- Database entry Removed
- Account Removed From Systems
Today - Issues

• Still get occasional duplicate entries
  • Primarily international students that get hired
  • Occasional SSN typos between HRS and SIS
  • Manual cleanup
    • Merge/remove duplicate in db
    • Kick off Account Self-Linking process (campus)
    • Notify student – go talk to HR Rep to fix SSN

• HR Not closing appointments
  • Often students
  • Engineering IdM will ignore Student Appointment if not a current student
Questions?