The Core User Directory

Tony Brett
OUCS

Agenda

- Why are we doing this?
- What is the CUD?
- How did the project start ?
- Who is involved ?
- What has been done ?
- What were the findings?
- Workshop May 2008
- Audience Participation and questions

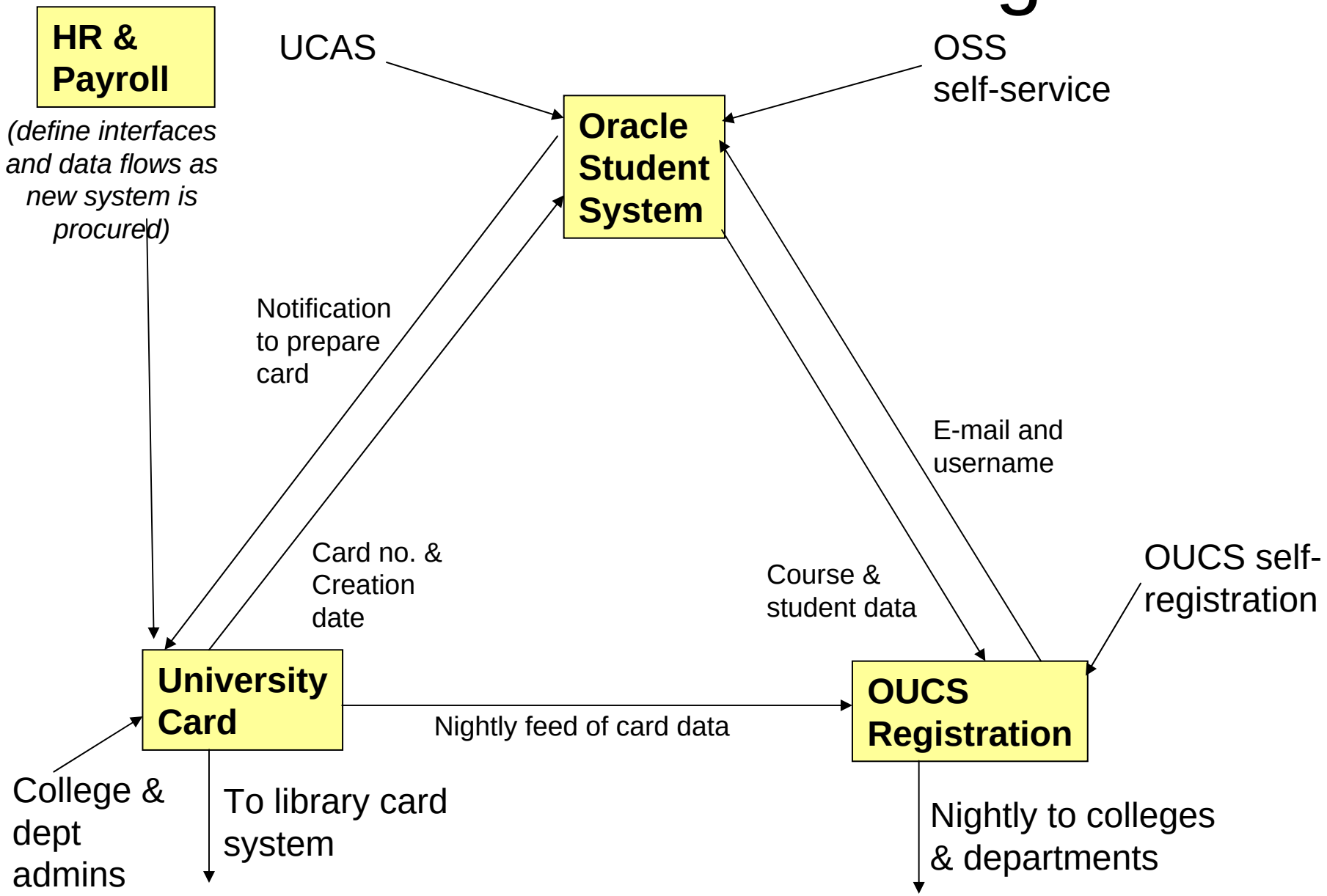
Why are we doing this?

- ICT Strategic Plan: "*Establish a University-wide Identity Management system which provides authentication and authorisation, and enables interoperability with national and international infrastructure.*"
- Quite a tall order!
- Many databases in many parts of Collegiate University
- No common key
  - people are duplicated and data goes out of sync
  - Significant risk of violation of DPA
  - Significant resource required to re-key data across many databases
  - Some databases e.g. University Card being abused as data sources rather than what they are there for (recording cards in this case)

Drivers & Urgency

- Our people-focussed administration processes are complex, costly, potentially error-prone and do not enhance the experience of someone joining the University community
- It is difficult, cumbersome, sometimes impossible, to join together identity attributes held in heterogeneous databases in order to provide key ICT services, whether centrally or within departments and colleges
- We’re way behind other Universities!

The “Bermuda Triangle”



What is the CUD?

- University-wide store of information for and about people who are connected with the collegiate University
- Many things to different people
- Not an ID Management solution in itself. It is required for IdM but is not IdM itself
- Simplest is a tiny set of attributes and provenances with a unique identifier
- Most complex is something that holds any attributes that are used by more than one service provider
- Both have pros and cons

Two extremes...

- Purist view is that CUD's sole purpose is to uniquely identify individuals
  - This gives the minimum set of data but doesn't solve data consistency problems
- Other extreme is that anything that is held in more than one place should move to CUD as the master
  - CUD becomes arbiter for changes and authoritative source
  - Requires decisions on business rules and attribute ownership
- Need to agree where to pitch the CUD between these two extremes
- Note that CUD "owns" nothing

How did the project start?

- Project profile written Feb 2008
- Project sponsor Director of ICT (Paul J)
- Working party formed in early 2008
  - Members from across many constituencies of the Collegiate University
  - Terms of reference tightly define the work of the group
- 2 phase project
  - Requirements analysis
  - Pilot implementation

Who is involved?

- As well as steering group there are 2x0.5FTE
- Jonathan Ward (BSP)
- Tony Brett (OUCS)
- Jonathan a contractor and Tony backfilled (Jane did much conference organising)
- Jonathan and Tony are currently just doing the work for the user requirements gathering phase. More staff will be needed in the second phase

Project Working Party

- Project Head: Paul Jeffreys
- Chair: Michael Fraser
- Secretary to the Board: Miranda Turner
- OULS: Dave Price (Neil Jefferies attends on behalf of Dave)
- Infrastructure, BSP: Tom Payne
- PRAS/Central Administration: Ilana Veitch
- BSP, HR Project : Chris Cattermole
- BSP, Student Systems: Mirjam Siderius
- Student Information Systems: Emma Potts
- University Card Office Manager: Maureen McNaboe
- Registration, OUCS: Beth Crutch
- NSMS, OUCS: Adrian Parks (until May 2008)
- ICT Support Team: Lyn Waddington
- Systems Development and Support: Ray Miller
- Department of Materials: Alana Davies
- Medical Sciences Division: Anne Bowtell
- Conference of Colleges: Peter Bushnell
- Continuing Education: Jim Davies
- OULS: Michael Heaney
- BSP: Heather Skevington

What has been done?

- Interviews with many data providers and/or consumers
  - Including: BSP, OSS, OUCS User registration, OULS, IMSU, Careers Service, Development Office, 2 Colleges
- Facilitated Workshop in May
- Use cases prepared and validated

Interviews

- Common format
  - Database schemas
  - How is data added/changed/deleted?
  - What data is provider authority for?
  - How are exceptions handled?
  - What import/export interfaces exist?
  - What do you need other providers to access?
  - What from other providers do you need to access?

Interview findings (1)

- Providers are keen to work together
  - exception may be with alumni data
- Data often moves through several systems before use
  - Should use original sources
- First-time population of CUD (pump priming) will be difficult and require manual de-duping etc.
- Many units use OUCS User Reg. at the moment but it was not designed for this purpose
- Multiple affiliation and status a big issue!

Interview Findings (2)

- Need ability for units to get basic data about people apparently not linked to that unit
- Source (authority?) for data changes with status
- University Card DB is much-misused (often via OUCS) and this sometimes causes problems e.g. Clinical medics
- Sometimes read-only access to CUD is sufficient
- Clear symbiosis opportunities e.g. Careers, DARS, Colleges
- Expectations and vision of CUD wide and varied so scope needs tight definition

Use Cases

- Some were written by project team
- Interesting ones provided by data providers/consumers
- e.g. Jesus College
  - Graduate students vs. Teaching Staff
- e.g. Language Centre
  - Self-registration & provisioning
- Extremely useful in helping project team to understand people's visions of how the CUD would work for them

The issue of multiple status

- University Card can’t hold more than one status
  - Only one college and/or dept
- No current way to record two roles as in Jesus College use case
- Quick win for the CUD?
- Card DB should be a consumer, NOT a provider as it currently is

Workshop May 2008

- Reviewed and Refined Requirements gathering findings
- Tried to understand which use cases the CUD would address in early stages
- Considered steps required to move into pilot implementation
- Facilitated by Dave Nesbitt, Identity Architect at Oxford Computer Group

Agreed at workshop…

- Operation of the CUD comprises two distinct activities: the 'cloud' activity involving data cleaning, reconciliation, etc; and the interfaces for data feeds in and out of the CUD. It was agreed at the workshop that for the pilot, at least, it would be sensible to release an LDAP interface
- Implementation phase to be a period of experimentation – that commencing on a solution was more important than a continued refinement of the CUD definition
- The next step should be to proceed with pilot implementation as basis for further discussion and evaluation
- The OUCS work already done on the Registration Database should be exploited for the pilot phase

Next Steps…

- Establish first set of core attributes for CUD
- Investigate solution for multiple status problem
- Establish sources for initial set of attributes
- Ensure that the pilot tests the reconciliation and provision of data from a representative; sample of data sources/users

Next steps…

- build on, and extend, the cloud activity undertaken by OUCS (Registration)
- build on, and extend, the provision of an LDAP service by OUCS (Oak)
- Establish Core User ID
- Move into Pilot Phase

Pilot options: Build on OAK

- The OAK Directory already exists and contains some of the information that would be required
- Uses as standard LDAP implementation, supports LDAP queries and will have Web Services infrastructure
- There is technical expertise within OUCS who could develop the CUD pilot
- Data Protection policies are already in place, although these would need to be examined further
- OAK could quite simply be enabled to model multiple status and affiliation

BUT

- Non-standard package and would require maintenance of in-house skills for support. It is, however, based on standard directory software.

Pilot options: Card DB Copy

- Card database already exists and contains some of the information that would be required, including photographs of card holders

BUT

- Proprietary package and possible additional licensing costs would be involved
- USB dongle required which might make large-scale rollout prohibitive
- There are restrictions on changes that can be made to the application. Typically columns may be added and the size of existing columns may be changed.
- Not designed as a CUD and contains functionality that would not be required.
- The Card Office is not placed to administer a Core User Directory.
- Card database cannot model multiple status/affiliation.

Pilot options: Commercial Package

- Product with the functionality could be purchased ‘off the shelf’
- Skills could be bought in from companies with experience of implementing identity management solutions

BUT

- The cost might be high for a pilot.
- It is unlikely that there would be in-house skills available.
- The complexity of the organisational structure of the University makes it unlikely that any package would be a good fit.

Decision!

- Building on the work of the OAK project was agreed by the Working Party in May meeting as a basis for pilot
- Workplans approved by PRAC Budget Subcommittee
- The next steps will be discussed further by the working party next week

Audience Participation

- Get together in College/Dept/Central Dept groups if you can (5 or 6 people)
- Think of a use case
  - Why do we do this?
  - What happens now?
  - How would a CUD help?
    - What would it need to do?
  - What's the risk of not doing it?
- Report back after 10 minutes

Questions/Comments?

Resources

- http://www.ict.ox.ac.uk/odit/projects/coreuser/
- Kurt Bittner, Ian Spence (2002). *Use Case Modeling*. Addison Wesley Professional, 2-3. ISBN 0-201-70913-9.