Data Authorization and Access: Empowering Faculty and Research Units

This document is the collaborative effort of staff in Computer Science, IT Services, and Graduate Studies.

In this document, we identify a large gap in processes that leverage UBC’s institutional data. We show that the major cause of this gap is that there is no authoritative, comprehensive, and unambiguous policy by which all units can understand how to gain and/or grant access to data owned by the University.

UBC must prioritize the governance of institutional data. We suggest the creation of a data stewardship advisory committee to advise the highest level of university administration. Such a group comprised of data stewards, faculty representatives, and central IT Services is necessary for effective creation, maintenance, and defence of UBC policy for access to institutional data. We understand “prioritization of governance” to mean that a senior administrator (such as a VP) take this on as one of their areas of responsibility.

1.0 Process description
Consider the following process for leveraging institutional data: a unit provides a service to a user, based on knowing who they are and if they are allowed to have the service. The former information is provided by Campus Wide Login (CWL) credentials, and the latter by cross-referencing a list. These “cross-referenced lists” contain a large set of UBC’s institutional data.

This scenario is repeated dozens of times in any given department: for administrative workflow and tracking tools, undergraduate service tools, research collaboration tools, desktop authentication, and more (see Appendix A for specific use cases.)

2.0 Process gap
(See http://en.wikipedia.org/wiki/Gap_analysis for a definition of “process gap”.)
Access to institutional data is gained via request to data owners: for example, Student Information Systems (SIS) for academic data, Human Resources (HR) for operational data, etc. Successfully gaining access to this data often depends upon:

- negotiations with individual staff, who have varying historical knowledge of previous negotiations;
- staff’s individual levels of comfort with privacy and access issues;
- long waits for responses, even to find out the correct avenue for access requests;
- unclear or ad-hoc criteria upon which decisions are based.

This process does not scale: direct requests of data owners is increasingly becoming a bottleneck and handling these requests is not part of data owners’ core missions (See Appendix B for examples of specific challenges.)
This gap is becoming increasingly apparent as IT Services builds on their role as a “data integrator”, creating services such as the Enterprise LDAP (ELDAP) to enable large-scale, convenient data authorization and access for the entire UBC community. This project will transform many units’ abilities to leverage organizational data, yet there is no written policy that enables automated distribution of any of it. Units are rushing to adopt these new tools, and yet IT Services is stymied by a lack of clarity around what data they can release to whom, and is handling each request – multiple per week – as a one-off ad-hoc thing.

This gap also has worrying implications for the university at large; a lack of clear data stewardship means there is no clear source of responsibility for which data accesses are being granted, from and to what applications, who they’re granted to, and who is responsible for communicating changes.

We note that UBC’s central service providers such as the Centre for Teaching, Learning and Technology (CTLT), Campus Security, and Student Housing are not immune to the effects of this gap. It affects all integrators.

3.0 Privacy and Security Considerations
Data owners feel heavily their legislative responsibility to protect personal information, and interpret that as needing to decide each new authorization for each new data access and business case. Allowing this decision making to reside at this level has serious limitations:

- Data owners lack a full appreciation of the business needs of the units, and their decision-making process does not scale.
- Units lack a full appreciation of the aggregate use of data across campus.
- Everybody lacks a full appreciation of the law, except the UBC lawyers.

Only a group that includes all stakeholders and UBC decision makers can devise a sane policy for the use of UBC institutional data.

4.0 Suggested solution
To address this problem, we propose that the University create an institutional data governance and stewardship advisory committee with mandatory stakeholder representation. The advice given by this committee must be acted upon by the highest level of university administration, by which we mean that a senior administrator (such as a vice president or the provost) take this on as one of their areas of responsibility.

Primary tasks would be:

- standardize and publish policy for access to UBC’s institutional data
- create business rules and scalable mechanisms for approval
- create clear processes for units to request and receive access to data systems
- include data owners and integrators as key stakeholders
- include Privacy & Security personnel and UBC lawyers
• own the database of requests, applications, and data authorization models
• be aware of privacy constraints
• make commitments to data consistency and standardization

This group would develop a high-level view of the university’s future data usage and needs, and would own the interface between data owners, data integrators like IT Services, and end users. They would act as stewards with an informed and comprehensive understanding of data needs on campus, and the legislative and privacy implications pertaining to the many potential uses of campus data (see Appendix A).

5.0 Conclusion
As a group of cross-campus staff including academic departments and IT Services, we believe that enabling better data authorization would provide immediate benefits to units. It would allow them to reduce duplication, save money, become smarter data analysts and users, and create more complex, adaptive applications to build operational efficiencies across dozens of different business needs.
Appendix A: Data use cases
Institutional data is used in dozens of important and appropriate ways. Data use cases listed here are related to the objectives of UBC as an educational institution:
1. Academic (Teaching)
2. Research
3. Administration

A.1 Academic data
There are many academic programs on campus, and each has workflows and processes that must be managed; typically this management is done virtually, and our tools operate in silos where access must be managed manually by departments or individuals, because the university lacks standardized mechanisms for granting and revoking access based on an individual’s academic relationship with the university.

1. Instructor teaches section of a course to students enrolled. A department (through staff) allocates resources to students taking the department's courses. These resources include, for example:
   a. Learning management system (LMS) application access.
   b. Group discussion, feedback, or progress-tracking applications.
   c. Teaching assistant (TA) management applications.
2. Faculties work together through multi-disciplinary programs that wish to share computing resources. For example:
   a. Collaboration applications (e.g. a wiki or administrative tools).
   b. Shared access to servers for course development activities.
3. Academic groups on campus control access to various physical facilities, using a very large variety of groups (groups created uniquely to application or unit).
4. The university tracks alumni and provides online services to them, which varies per department and includes, for example:
   a. Participation in outreach activities such as mentoring and event participation and signup.
   b. Participation in online communities of interest.
5. The university secures cost-effective (often free) software site licenses that require a fine-grained understanding of students eligible for the license. Software such as Microsoft Office 365.

A.2 Research data
Research groups span the institution and beyond. They require collaborative tools that must often be hosted in Canada and on university systems, depending on granting agency requirements.

- Research groups have labs that require electronic access control.
- Research labs have network ports used by individual researchers or by a group.
• Research groups have resources for the use of a group that is often wider than just the research group. For example:
  o Shared server environments for collaborative use by members of a field of interest.
  o Shared online applications for collaboration, analysis, and administration of interdisciplinary or inter-institutional relationships, or with granting organizations for transparency and data-sharing.

A.3 Administrative data
There are two broad administrative business areas at UBC: employee, and academic/research.

A.3.1 Employees
The AccessUBC system manages active employee’s access to services and applications (e.g. FASmail).

A.3.2 Academic & research units
• Program-specific workflow management applications.
• Additional student information systems (e.g. supervisor, academic progress, advising appointment management and notes).
• Student administration on many levels is decentralized with functions for subsets of students being provided by various organizational units. Organizational units require a mechanism to identify “their” constituents, e.g. a faculty needs to restrict access for all members of the faculty or a department needs to query information for all members of that department.

Appendix B: Specific challenges around data owners
Data owners have no vested interest in allowing access to their data; it represents, at best, a threat to their important legislative responsibilities to protect personal information. However, they are being increasingly asked to make decisions based on use cases they do not have the capacity to evaluate, and to appreciate the security and privacy implications of applications and systems they do not have the technical background to understand.

Each data owner is replicating the work of the others, and there is little incentive for them to develop processes that would make their data more popular and easier to access, and create more evaluation problems for them of this sort.

As a result, many units and departments have given up asking permission and prefer to ask forgiveness; they circumvent data owners if they can. For example, we have found units:
• getting data for apps in their unit by piggybacking on unrelated but already-authorized applications (sometimes from apps not even in their own units);
• building tools to manually scrape data from existing administrative system screens’ and
• getting staff to manually copy data from one system to another every term.
All of these have been easier than getting authorization to use university data through ‘formal’ channels, such as they are: they result in out-of-date or incorrect data, work duplication, and uncertainty around the future of on-campus data access.

This also suggests that large enterprise changes, such as a $100M replacement of the Student Information System (SIS), will break many systems that the university doesn’t know exist; and that, without a new mechanism for getting departments access to the new system’s data, that larger, more expensive problems will crop up because the old informal, ad-hoc access mechanisms will no longer work. Without better data stewardship, the SIS replacement will have even higher associated costs.