A Framework for Data Sharing in Academic Collaborations and Pathways

2017-29

Final Report

(April 2018)

Table of Contents

Introduction

Data Sharing in Academic Collaborations and Pathways was a qualitative research project which examined this practice in six Ontario postsecondary institutions. A larger more detailed report has been produced about the transfer of data among academic institutions in Ontario. However, below is a synopsis and excerpts of highlights from the larger report. As the lead institution on the project, York University followed institutional guidelines for the tendering of the research through its Procurement Services office. Thinklounge was the successful bidder and was subsequently hired to conduct the study by the project team.

Executive Summary

This research project was developed to explore data sharing practices between college and university partners in Ontario; and to provide broader guidance to PSE institutions across the province engaged in student data transfer. Three Universities and three colleges participated in the study: York University and Seneca College; Trent University and Fleming College; and the University of Ontario Institute of Technology (UOIT) and Durham College. Central to the research was an exploration of the current data sharing practices among institutions; how and for what purposes data are being shared inter-institutionally; the legal, privacy and other challenges that needed to be mitigated to facilitate data sharing; and an examination of the data sharing approaches in postsecondary jurisdictions outside of Ontario, namely in the US and British Colombia. Two objectives were central to this study. The first was the provision of a data sharing process for institutions. This was to include best practices from other jurisdictions and industries; the creation of a framework to assess the parameters for data sharing in academic collaboration; providing understanding of the broader factors involved in data sharing; and a draft data sharing agreement or MOU (Memorandum of Understanding) template that could be used by institutions. The second objective was to provide insight into the issues involved in data sharing among Ontario postsecondary institutions and suggestions for improving this practice more broadly across the province, with the goal of improving student mobility outcomes within the higher education sector.

A steering committee comprised of representatives from the participating institutions was created to provide project oversight. This group was also instrumental to/ supported the participant recruitment process. The research comprised of a literature review, thirty-one interviews and two focus groups. The participants were recruited from a variety or roles/ institutional functions which included but were not limited to administration; faculty/ academic and registration staff; IT, legal, privacy and ethics offices. The data collection began in late Fall 2017 and concluded in March 2018. Additionally, staff from the Ontario Universities Application Centre (OUAC); the Ministry of Advanced Education and Skills Development (MAESD), Ontario; and the Student Transitions Project in British Colombia was interviewed for the research.
The study found that the various University Acts gave institutions broad authority to collect and create data for educational, research/ statistical and administrative purposes. On the other hand, among Ontario Colleges, personal data is collected under the authority of MAESD

(Ministry of Advanced Education and Skills Development) for educational, administrative and statistical/ research purposes. Additionally, all educational institutions in Ontario are subject to FIPPA (Freedom of Information and Protection of Privacy Act). However, while “institutional legislation governs the creation and maintenance of student records…FIPPA is the law in Ontario that governs the protection and disclosure of records that institutions can create by statute” (Baumal, April 2018). In short, FIPPA sets out the parameters for data sharing in the Province, including the information being shared among institutions. Moreover, FIPPA governs the use of personal information and the transfer of personal information (Baumal, Data Sharing in Academic Collaboration: Final Presenation, 2018).

The research also found that the data collected by academic institutions can be broadly classified into two categories: 1) research and planning; and 2) administration. Data collected for research and planning were often used to conduct studies on student mobility and for institutional policy planning such as determining space capacity, strategic planning and for the development of agreements. On the other hand, the data collected for administrative purposes involved the sharing of data during/ for student redirection; for the administration of dual credentials and other types of collaborative programs among institutions, and to facilitate general registrarial level transfers.

Institutional staff interviewed indicated that there was strong knowledge of the FIPPA legislation and high rates of compliance within their respective organizations. Accordingly, despite the presence of privacy policies which broadly outlined the authority to collect personal data from students, many of the institutions still sought / had specific processes in place to notify students for obtaining their permission to share record level data with other institutions. However, the study found that most of the occurrences of data sharing in the province were bilateral, as Ontario does not have a centralized data repository or warehouse.

Challenges and limitations of data warehouses and the use of the OEN

FIPPA legislation provides oversight into the type of information that can be provided; how a recipient will use the information; and when and to whom information can be shared/ disseminated. Therefore, the research revealed that privacy issues were not the greatest impediment to data sharing among educational institutions in the province of Ontario. Instead, the translation and matching of records, and agreement on what can be said and done with shared data, remained the primary concerns. In many instances, institutions may have different grading schemes, operational procedures, exam schedules, enrolment and scheduling requirements which differed greatly from another potential institutional partner- resulting in data translation challenges. Additionally, very rarely did institutions have shared student numbers which made the matching of student level records difficult and onerous.

What's happening elsewhere? Examples from British Colombia (BC) and the United States (US)

Other jurisdictions such as British Colombia (BC) and the United State (US) have data repositories/ warehouses to facilitate data sharing among many institutions. That data are stored mostly as aggregate level information. While there was an awareness, particularly of the BC model in Ontario, where the Ministry of Advanced Education's Central Data Warehouse, through the Student Transition's Project in that province has granted access to aggregate level data, the same could not be said for Ontario. The widespread use of the Personal Information Number (PEN) in British Colombia, which is equivalent to Ontario's Education Number (OEN), has facilitated the access to provincial level data among BC's postsecondary institutions.

Despite the challenges to data sharing in Ontario, there remained general ongoing interest in improving and increasing data sharing among educational institutions. Particularly among institutions with formalized partnerships and those who viewed themselves as 'transfer institutions,' there was a greater willingness/ openness to increased bilateral collaboration and data sharing. Additionally, according to the research findings, senior administrators were increasingly interested in answers to more 'strategic questions' including but not limited to enrolment and retention patterns and saw the value of province-wide data sharing exercises.

One proposal for facilitating multilateral data sharing emerging from the research was the establishment of a data sharing repository/ warehouse, like BC's Student Transitions Project (STP). The latter project is a partnership between BC's Ministry of Advanced Education, Ministry of Education, the University of British Colombia. Simon Fraser University, the University of Victoria and the University of Northern British Colombia to share student level data and track student mobility throughout that province's education system. The STP governance structure includes a steering committee which has its own terms of reference; a breach of privacy protocol; a data access protocol; a data linkage policy; and a reporting protocol. This approach to data transfer offers a valuable framework and a potentially replicable model for the province of Ontario.

There were also several best practices emerging from the report. They included the following:

1. Only personal information needs protection – records that have been properly de-identified are not subject to FIPPA regulations. Therefore, where possible, institutions should consider de-identifying data before it's transferred/ shared.
2. Data transfer should be secure and include the use of Electronic Data Interface (EDI) systems / portals.
3. Access levels and data storage should be determined and written into data sharing agreements.

4. Breach protocols, audit trails and reporting conventions should be agreed to and documented in the data sharing agreement.
5. Schedules should also be determined for data exchange and a term limit should be set for the receiving institution's use of the data; as well protocols should be determined for the interpretation of data for publication.
6. Transfer only information that is relevant to the purpose.
7. Have translation plans, as often, for example, institutions use different GPA systems or interpretations of transferred data. Therefore, consideration should be given to standardizing data definitions across institutions.
8. Involve many levels of the institution in the creation of data sharing agreements/ arrangements.

Data Sharing Template & Framework

As previously mentioned, there were several deliverables for this research project. These included an MOU / Data Sharing Agreement template and Framework to guide data sharing practices between and among academic institutions. In creating data sharing agreements, there are several considerations that must be contemplated by institutions.

- At minimum data sharing agreements must comply with FIPPA regulations.
- Institutions must make a business case for data exchange of personal student records.
- Agreements must also indicate:
  - What personal information will be shared;
  - How the information will be used;
  - Whether data will be de-identified;
  - How the data will be shared and linked;
  - The protocols that will be in place to ensure accuracy and security;
  - How the data will be released; and
  - The term of the agreement/ termination date.

Finally, a framework for data sharing must include a good understanding of legal and privacy issues impacting 'protected data.' Under FIPPA, protected data is defined as private or identifiable information. Institutions must therefore clearly outline in their data sharing agreement how personal data will be legally protected. Any data that can identify a student is protected. This includes data on: "race, national or ethnic origin, colour, religion, age, sex, sexual orientation or marital or family status of the individual; educational information or financial transactions in which the individual has been involved; [As well] any identifying number, symbol or other particular information assigned to the individual, [such as] address and telephone number'' (Baumal 2018, p.13).

List of Partner Institutions

- **York University**
- **Seneca College**
- **Trent University**
- **Fleming College**
- **University of Ontario Institute of Technology**
- **Durham College**

Research Ethics

Prior to conducting the research, individual ethics applications were completed and submitted to the appropriate research ethics board/ offices at each of the participating institutions. The Ontario College strike in the Fall 2017 did however cause some delays in the receipt of response and approval, due to staffing limitations. Although Ontario Colleges have a multi-college ethics process, and all the participating colleges in the study were signatories to this agreement, it proved to be more efficient for this project to submit to institutions, individually. As well, support from members of the Data Sharing Steering Committee, proved invaluable in securing successful ethics approvals, particularly in submitting the required documentation to the Boards, initiating the connections and following up when there were delays.

Research Methodology

A total of thirty-one interviews were completed, with a minimum of at least four participants from each of the six institutions. The interviews began in December 2017 and concluded in March 2018. Each of the interviews were approximately one hour in length. With permission obtained from participants, interviews were digitally recorded and then uploaded to a computer, adhering to research privacy and confidentiality protocols. Following the interviews, two focus groups were completed at York University. There was an average of 7-10 participants at each of the focus groups. These participants had the option of attending via teleconference or in-person.

Interview and focus group participants were selected from among a range of occupational functions within the various institutions. They included staff from the registrars, privacy and ethics offices; program and academic staff involved in articulation development; IT support and senior administrators. Staff from these areas were targeted for the research. Therefore, purposive sampling was the initial participant recruitment approach. If they were unavailable to participate, they were then asked to recommend others within the organization, which made snowball sampling technique the other approach used to solicit participants for the study. The steering committee for the project was also instrumental in identifying potential research participants and doing the initial introduction of the project and researcher at their respective institutions.

The interview questions addressed participants familiarity with and involvement in data sharing practices within their respective institutions. The were asked to comment on the following areas during the interviews: the student consent process/ practice within their institution; privacy, accessibility, use and frequency of use of data sharing, to name a few. Also discussed were institutional policies; technological limitations; and the legal requirements for data sharing.

The focus groups on the other hand were used to clarify and confirm information emerging from the interviews. Additionally, questions around the motivating factors for data sharing within respective institutions; the types of data sharing occurrences/ (purposes for data sharing) within institutions; the types and benefits of data sharing; the challenges such as legal and IT hurdles, as well as the process for obtaining student permission; the use of the OEN; competition vs cooperation factors implicated in data sharing; and the use and request for data from OCAS (Ontario College Application Service) and OUAC (Ontario University Application Centre). From these, themes were derived to guide the development of the report.

Information was also collected from existing documents submitted to the researcher by participating institutions during interviews and focus groups.

Literature Review

Generally, data sharing requests were grouped into five broad categories, with specific purposes or functions. They were as follows: 1) general registrial/ student registration or student record transfer; 2) for the administration of collaborative[1] programs; 3) for student redirection; 4)

[1] *The use of the term “Collaborative Programs” in this report is a general term that encompasses all programs across institutions that formally facilitate student mobility. This term encompasses* ***Co-registration programs*** *allow students to enroll in courses at another post-secondary institution for credit towards their degree program at the primary institution[1];“***A joint degree*** *program is a program of study offered by two or more universities or by a university and a college or institute, including an Institute of Technology and Advanced Learning, in which successful completion of the requirements is confirmed by a single degree document”[1]; “A* ***dual credential*** *program is a program of study offered by two or more universities or by a university and a college or institute, including Institutes of Technology and Advanced Learning, in which successful completion of the requirements is confirmed by a separate and different degree/diploma document being awarded by at least two of the participating institutions”[1]; and* ***A collaborative program*** *is a graduate program that provides an additional multidisciplinary experience for students enrolled in and completing the degree requirements for one of a number of approved programs. Students meet the admission requirements of and register in the participating (or “home”) program but*

for research into articulation agreements/ pathways between institutions; and 5) for institutional research and planning.

The research found that data collected for the administration of programs and for registrarial purposes often could not be de-identified. This was because the student's identity was fundamental to those processes. On the other hand, data collected and shared for research purposes could often be de-identified, since the student's identity would not be required for the analysis of this type of data. Therefore, in the formulation of data sharing agreements understanding the purposes or functions of the data being collected and shared was critical. Data were often transferred between institutions so that records could be matched for research, policy development or administrative purposes. Therefore, institutions needed ways to link the data so that student records of the sending institution could be matched with that of the receiving institution, particularly when there were elevated levels of student mobility between partner institutions. Data in these instances were often matched using 'tombstone data,' which included name, date of birth and address, for example. In some cases, this type of data along with a student identification number was transferred to facilitate the data linkage. In Ontario, any personal information, including an identification number is considered 'identifying information' and is therefore subject to FIPPA regulations. Legally, educational institutions can and do collect 'tombstone' level data from students and generate student identification numbers. However, FIPPA is the legislation that provides broad guidelines on how this may be shared between and among institutions.

Given the above, a few best practices have emerged to facilitate the sharing of identifiable information between institutions. Some institutional partners have created a common student

complete, in addition to the degree requirements of that program, the additional requirements specified by the collaborative program. The degree conferred is that of the home program, and the completion of the collaborative program is indicated by a transcript notation indicating the additional specialization that has been attained."[1] *Sources: http://secretariat-policies.info.yorku.ca/policies/undergraduate-co-registration-options-with-ontario-post-secondary-institutions-policy-and-guidelines-on/)*
http://gradstudies.yorku.ca/current-students/regulations/degree-types/#inter:
http://gradstudies.yorku.ca/current-students/regulations/degree-types/#inter:
http://gradstudies.yorku.ca/current-students/regulations/degree-types/#inter:

numbering system to facilitate data transfer; while others have transferred identifiable data, following strict protocols on access as well as data encryption processes. This information was then matched by the receiving institution. Then de-identified information was sent back to the institution who originally requested the information. Another best practice that has been utilized was the creation of 'data key/ unique identifier and stripping tombstone data information after matching the records, before it's sent back.

The literature review found that by 'acts of statute,' postsecondary institutions in Ontario can create and collect information for the development of student records to "administer the business of the institution" (Baumal, April 2018, p. 19). Accordingly, all institutions in the study had published privacy statements in their online platforms indicating the types of data that were being collected and their compliance with FIPPA regulations. However, the colleges appeared to provide more detail in defining the types of personal data that they were collecting from students. Some of the colleges collected information on: 'race, national or ethnic origin, colour, religion, age, sex, sexual orientation or marital or family status of the individual; information relating to employment, criminal or educational history; finger prints, blood type; information relating to the medical, psychiatric, psychological history, prognosis, condition, treatment or evaluation; personal opinions; home address and/or telephone number and; any identifying number (e.g. S.I.N., student number), symbol or other particulars assigned to the individual; correspondence sent to the college by the individual that is implicitly or explicitly of a private or confidential nature, and replies to that correspondence that would reveal the contents of the original correspondence; to name a few. These were vital to the accomplishment of the pedagogical and operational activities of the college (Baumal, April 2018, p.20-22).

On the other hand, some of the university partners in the study indicated specifically which third parties, with whom student information can/ will be shared. For example, at York University, this included: " Other universities and colleges to verify any information provided as part of an application for admission; other universities and colleges to share incidences of falsified documents or credentials, or share information regarding fraudulent applications for admission; government offices to verify information regarding an application for admission and to support processes for government financial aid; other universities and colleges with which York University maintains a collaborative program partnership; service providers contracted by York University to support business processes.[2]" (Baumal, April 2018, p. 19-20). The University Act of each of the participating universities in the study provided the legal parameters for collecting student level data, while the Ministry of Advanced Education and Skills Development (MAESD) granted this authority to the colleges (Baumal, April 2018).

Although Carleton University was not one of the partner institutions participating in the study, its privacy statement provided some good insight for the literature review. The institution noted that;

[2] *https://registrar.yorku.ca/privacy*

> the Freedom of Information Protection of Privacy Act recognizes the legitimate need to collect personal information in to carry out one's mandate and to provide services but restricts that collection to a defined set of circumstances. The circumstances are the collection of information is expressly authorized by or under an Act… the information relates directly to and is necessary for the University's operating programs or activities… The Carleton University Act does not specify what personal data elements can be collected. However, personal information must be relevant to the purpose for which it is being collected… The University may do its own collection or may authorize an outside agent to carry out the collection on its behalf, either under contract or through an agreement or arrangement in writing with the other agency[3] (Baumal, April 2018).

In the above, Carleton University was explicit about the authorization it has to collect and use data under FIPPA and its 'University Act.' Also introduced here was the use of 'data sharing agreements' for access by third parties to student information.

The Legal Framework – FIPPA (Freedom of Information and Protection of Privacy Act)

Under FIPPA Regulation 41(1) b, c, institutions can use and transfer personal data if it is "for the purpose for which it was obtained or compiled or for a consistent purpose...where the discloser made is necessary and proper in the discharge of the institutions functions [4]" (Baumal, April 2018, p.26). There are several conditions in FIPPA related to the confidentiality and security of data. They are as follows:

1) If information is to be disclosed to a third party for research purposes, written authorization must be obtained from the institution.
2) There must be a written agreement where the persons having access to the data shall be named.
3) Before personal information is disclosed, the persons who will have access to the data shall enter into an agreement, confirming that they will not be disclosing the information to another party.
4) The information shall be kept physically secure and access only given to authorized persons.
5) The receivers of the information shall destroy all individual identifiers in the data as specified in the agreement.
6) The receivers of the information shall not contact persons whose personal information they have obtained without prior written authority from the sending institution.
7) Receivers shall ensure that no personal information shall be disclosed in a form which the individual to whom it relates can be identified without the written authority of the institution.

[3] https://carleton.ca/privacy/wp-content/uploads/policy_collection1.pdf, P8-9

[4] https://www.ontario.ca/laws/statute/90f31#BK57

8) If there is a breach, the receiver shall notify the institution in writing immediately.

FIPPA regulations contained three fundamental pillars. That is, prior to disclosing or agreeing to disclose information, what must be known is 1) the type of information that is expected to be provided/ exchanged; 2) how the recipient will use the information (it purpose); and 3) when and with whom the recipient may share information. These three pillars must also be enshrined in every privacy policy and potential data sharing agreement among institutions. FIPPA also defines personal information and stipulates that it must be protected and properly disclosed (Baumal, April 2018, p. 24).

The OEN (Ontario Education Number)

In the past there have been limitations placed on the use and transfer of the OEN between institutions. While the use of the OEN can potentially limit the need to transfer 'tombstone data' between institutions, there are limitations. For example, students transferring into Ontario PSE institutions from outside of the province won't have assigned OEN's. Thus, the accuracy and use of the OEN for students transferring in from outside of Ontario was problematic. A similar challenge emerges when there is a need to access more historical student level data, as the assignment of OENs is relatively new in the province of Ontario. Despite these challenges, in practice, the OEN is being used to carry out registrarial functions, important for the verification of the student record across institutions. Additionally, interviewees have confirmed that research has also been completed using the OEN. However, provincial-wide access to use the OEN has not yet been granted.

Environmental Scan of Emerging Trends & Key Issues

Data sharing has and can be used by institutions to understand broader student mobility patterns within the Ontario system. At present though, based on the study results, most of the data exchange, particularly for research purposes, among he six institutions studied was bilateral or in some cases trilateral in scope. Some institutions also indicated, however, that they had conducted research for stakeholder provincial organizations such as HEQCO (Higher Education Quality Council of Ontario) and ONCAT (Ontario Council on Articulation and Transfer) (Baumal, April 2018, p. 33). Despite its limitations, the use of the OEN remains central to the expansion of data sharing practices in the province of Ontario.

Analysis & Limitations of the Data

In the study, many participants indicated that their institutional websites clearly stated the circumstances under which personal data would be shared with a third party. They further mentioned that there were few technical barriers to data sharing among institutions. However, issues around timeliness, (that is, regarding when information is received by the requesting institution) given the differing exam schedules and grading schemes for example, created some challenges, particularly for data transfer for registrarial purposes.

Despite the information on the institutional websites regarding the collection and transfer of 'personal information,' some institutions also informed students at the time of application or had them sign forms, giving their expressed permission to release data. In rare instances, the signing of this attestation was sometimes a source of confusion or anxiety for students. Students in collaborative programs for example, were sometimes concerned that these forms were part of a reapplication process or were concerned about their grade transfer, particularly if they were under-performing academically. They were consequently anxious about their ability to continue in their chosen program at the university.

In many of the institutions where data sharing was occurring there were agreements in place. While the sharing of data was not particularly an issue, the timing, translation, equity in sharing and overall use of the data were some of the hindrances to this process. Additionally, concerns were raised about the narrowness of scope often stipulated in data sharing agreements. In some instances, shared data was useful beyond the purpose for which it was stipulated in the agreement. As such, balance was needed to manage the timing and the more extended use of data in partnership agreements.

Data transfer however was occurring through the various levels and role functions across the institutions studied. The most common data sharing purpose was registrarial level data followed by data sharing for research and planning purposes. The former involved data exchange for program and course registration including for collaborative, dual credential and / or co-registration programs. Data quality and accuracy were also a concern and, thus, some institutional pairs opted to use OUAC and OCAS application data to track the sending institution's students. Specific concerns were also raised about the interpretation, analysis and the implications that may be drawn from student level data transferred between institutions.

Another way that student mobility was being tracked was by the number of requests for transcripts sent to institutions and programs. Among study participants there was a growing interest in more open data exchange through formal agreements and/or using the OEN. Questions were raised though on whether more unrestricted access to this type of data would create more competition between institutions. In response, study participants agreed that the benefits outweighed the costs, noting that "more open exchanges of data would allow institutions to play to their individual strengths, thus producing better outcomes for both students and institutions" (Baumal, April 2018, p. 35).

There was however, very limited use of data transfer for the development of articulation agreements. The study found that this may have been because many institutional pairs had very close working relationships and were very familiar with the program offerings and mobility patterns of students who moved between the respective institutions. Additionally, the success of articulation agreements was not often measured by transfer outcomes or the academic performance of students in specific programs.

Data sharing was therefore not confined to the Information Technology departments of the various institutions studied. Instead it was embedded throughout various levels and functions within the institution. Additionally, when asked how data sharing could be improved and / or increased, some participants recommended the creation of a more open Electronic Data Interface (EDI) system. Others mentioned the expanded use of the OEN; greater access to OUAC and/ or OCAS data, or even data collected by the Ministry of Advanced Education and Skills Development, in Ontario.

> More open and free-flowing exchanges of data can come from [different] sources. First institutions can agree to individual arrangements. Second, OUAC and OCAS can facilitate this kind of transfer, according to some participants (Baumal, April 2018, p. 35).

The challenge with these information sources is that its mostly aggregate level data that is transferred and stored.

Limitations of the Study

The study excluded analysis of transcript level data and data transferred by institutions to OUAC or OCAS. By and large, the transfer of student transcripts requires the latter's consent and have a very defined purpose, which is often connected to the program admissions process. However, all the above data types are subject to FIPPA regulations.

Promising Practices & Challenges

The research uncovered several promising practices for data sharing in academic collaborations.

- **Only personal information is subject to FIPPA regulations**. Therefore, where possible, institutions should transfer and use de-identified information. However, data transferred across institutions for administrative and research purposes were often identifiable, particularly when there was a need to create linkages between other data sets, before it could be de-identified. It was therefore sometimes necessary to transfer identifiable data across institutions to create these linkages.
- **Be clear on the benefits to students, institutions and the public.** In recognition of the law and other practical considerations, the right to privacy must be weighed against other public interest. Institutions must therefore create a business case for why 'personal data' must/ should be shared. The business case should include:
    a. the goals and objectives of the data sharing activity and the anticipated benefits;
    b. the potential risk or consequences for not conducting data sharing;
    c. clarification of why personal information must be shared at this time;
    d. clarification on why the data need to include personal identifiable information;

e. a statement of purpose for which personal information was originally collected; and

f. identification of why personal information must be collected indirectly and the advantages of data sharing, when compared with alternative methods of achieving the same objectives.

One of the remaining concerns identified by the Information and Privacy Commissioner of Ontario was that data transfer may lead to a loss of control by individuals over their data. Therefore, where possible, data sharing should not occur without exploring privacy and/ or less intrusive means of achieving the same objectives.

- **Transfer data that are relevant to the purpose of the assignment/ for which they were granted.** Institutions are encouraged to only transfer / release information that is necessary for the purpose at hand.
- **Standardizing the definition of personal information**. As previously mentioned, the privacy statements of the colleges in the study were very explicit about what constituted private information. Additionally, FIPPA legislation clearly defined personal identifiable information. As a best practice, all postsecondary institutions should clearly define its meaning of 'personal information'.
- **Create a translation plan for exchanged data.** Information related to grading schemes or graduation for example, may differ from institution to institution. Data must be clearly understood by sending and receiving institutions for data analysis in the case of research, and to ensure the accuracy of the student record.
- **Involve relevant areas of the institution. "**Data exchange involves a number of areas of institutions including programming, registrars, IT, legal, privacy and ethics. For the most part, participants indicated that they were comfortable with the level of consultation in which they engaged with other partners. That is not to say that they always engaged with all relevant partners all the time for every data exchange" (Baumal, April 2018, p. 42).
- **Consider de-identifying data & creating a data sharing agreement that governs de-identified data.** As mentioned above, de-identified or non-personal data was not subject to FIPPA legislation and therefore constitutes a best practice. However, as indicated previously, it may not always be possible to de-identify data before its transferred, thus setting in motion the requirement for a 'business case'. Additionally, even though data may be de-identified, in smaller academic programs with limited enrolment, it may still be possible to identify certain individual students because of their unique characteristics. Therefore, as a best practice, where possible, aggregate small cell level data and create a data sharing and confidentiality agreements between researchers at the relevant institutions who are engaged in transfer.

The Information and Privacy Commissioner in Ontario has listed several steps for achieving data de-identification. They are: 1) determine the best data set release model; 2) classify variables; 3) determine an acceptable de-identification risk; 4) measure data risk; 5) measure content risk; 6) de-identify data; 7) assess data utility; and 8) document the process. The Commissioner also listed specific components for data sharing agreements which included, but was not limited to, the prohibition of the use of de-identified information with other information to identify and individual; placing restrictions on the use and subsequent disclosure of information; ensuring that those who have access to de-identified information is properly trained and understand their obligations with respect to information privacy; requires individuals to notify organizations in cases of breach; and sets out the consequences for a breach. Recipients of data are also responsible for protecting groups from ‘attributed disclosure.’ This may occur when a group is identified negatively by the data – thus stigmatizing them (Baumal, April 2018, p. 44-45).

- **Linking data.** Data linkages were sometimes necessary to ensure data/ student level record accuracy. This was particularly important for data exchanged for research purposes. One best practice that is often recommended is that the data holder does the data linkage before sending the de-identified data back to the requestor. Additionally, “following the linkage of datasets, the person doing the data linkage should reduce datasets to the lowest level of identifiability needed to accomplish the research objectives” (Baumal, April 2018, p. 48).
- **Determine access levels and uses for/ of the data.** “Institutions involved in data sharing may wish to consider assigning various access levels to different researchers and team members, especially as it relates to identifiable data” (Baumal, April 2018, p. 48).
- **Implement the Data Transfer and Arrange Schedules. “**Institutions indicated that the actual transfer of data between institutions is not tightly controlled, but rather those involved in the actual transfer of data are relied upon to implement transfer in a secure and responsible manner in accordance with institutional policies. For the most part, participants indicated that they were aware of their institutional policies regarding data transfer and implemented them. In fact, IT individuals interviewed indicated that their involvement in data transfer was generally rare because the function could be handled directly by the staff involved. Best practices of data transfer largely involved securing and encrypting it, and then ensuring that it is stored on a properly secured device at the receiving institution” (Baumal, April 2018, p. 49-50).
- **Storage and verification of the accuracy of the data.** Although FIPPA legislation does not specify data storage requirements, there are a few best practices that can be followed. They include: controlling access to rooms, buildings and computers where data is stored; logging the removal of and access to media or hard copy material; not storing confidential information such as personal information on servers or computers connected to external networks; ensuring firewall protection and security upgrades to avoid viruses and malicious software intrusion; and securing computer systems and files using passwords, firewalls, restricting access; using encryption etc. (Baumal, April 2018, p. 50-51).

- **Reporting breaches and audit trails. "**A significant amount of agreements contained provisions concerning procedures to address data breaches that would occur on the other side of the agreement. The best practice is for the Data Steward on the breached side to report the breach 'immediately' to the Data Steward on the other side" (Baumal, April 2018, p. 52). If a breach occurs the following steps should be implemented: 1) identify the scope of the breach; report the breach immediately to the appropriate staff; retrieve any documents that may have been disclosed to or taken by an unauthorized recipient; inform persons whose information may have been directly disclosed; and investigate the facts of the breach and make recommendations (Baumal, April 2018, p. 52).
- **Determining and reporting access conventions. "**Once the data set is in its final form, whether it contains identifiers or not, researchers and policy makers should consider how the data will be reported and accessed to account for privacy concerns, including grouping data so that individuals cannot be identified, and reporting conventions will avoid identifying or stigmatizing any individual or group… Researchers should [also] address levels of release for the data and the report" (Baumal, April 2018, p. 52). There are generally three kinds of data release: i) private; ii) semi-private; and iii) public. As a best practice, the 'public' release of information must be de-identified (Baumal, April 2018, p. 53). "The [data sharing] agreement should address how both the data and the report should be presented to avoid identification and stigmatization of an individual or group. This may mean grouping and/or suppressing some variables and results more broadly in both the data and report so as not to identify or negatively impact any person or identified group. The researchers should consult with institutional policies regarding privacy and confidentiality and any restrictions placed on data that may be included in the data set from third parties. In general, cell sizes that have a count of ten or less should not be released and data should be grouped to avoid results that identify" (Baumal, April 2018, p. 53) individuals within the data.
- **Determining additional usage options. "**A fundamental aspect of data sharing agreements and data usage is that the exchanged data can only be used for the purpose for which it was exchanged and/or for a certain period. Some participants in this study indicated that they were constrained from using data for additional research purposes because the data sharing agreement limited the use of the exchanged data to only a particular use and/or timeframe. Those planning on exchanging data should think about future uses of the exchanged data and consult with legal departments to determine if future uses can be permitted and how those uses should be incorporated into data sharing agreements. This may avoid situations where exchanged data cannot be used because the initial data sharing agreement is too limiting" (Baumal, April 2018, p. 54).

Conclusion

The aim of the research was to explore what student level information was being shared among Ontario PSE institutions; and to what extent current privacy legislation placed limitations and potentially inhibited or encouraged collaboration related to student mobility and research across institutions. The goal then, was the develop a framework or structure that institutions could use to guide their data sharing practices and create a MOU template for data sharing. With the increasing numbers of articulation agreements and formalized partnerships between academic institutions, particularly universities and colleges, there is a growing interest among PSEs for transfer related data to support for academic planning, program development, partnership and research, more broadly. The widespread use of the Ontario Education Number was being championed as one approach towards understanding system level mobility patterns among Ontario institutions.

The research found that provincial level institutions such as the Higher Education Quality Council of Ontario (HEQCO) was among the more vocal with regards to use of the OEN. HEQCO notes that

> the OEN informs policy files at the centre of provincial priorities: mobility, equity of access, student success, and institutional differentiation...The pendulum on protection of privacy is swinging from 20 years of “play safe: don’t share anything” to a balanced approach that protects individuals while promoting evidence-based policy and program design. The Ontario Ministry of Advanced Education and Skills Development is signaling a willingness to share OEN data, appropriately protecting privacy, with the broader community and is taking steps to do so (Baumal, April 2018, p. 66).

Centralizing and controlling access to ‘personal information,’ remains a pivotal issue in Ontario. At present, the Ministry of Advanced Education and Skills in Ontario collects aggregate level data on student mobility within the system using its ‘Open Sims’ system. Additionally, the interviews with research participants revealed that MAESD and HEQCO have been collaborating on advancing research on student mobility using the OEN (Baumal, April 2018, p. 67-68). There were several concerns raised, however, with regards to access and control. There was the concern that those who gain access to this data may act as ‘gatekeeper,’ limiting access to the data, thus creating data release inequities. Caution was also encouraged with regards to the creation of ‘private entities’ for warehousing/ storing student level data/ privatization of data storage in Ontario. Of concern was the notion that the motives or interests of for-profit entities may not be consistent with privacy concerns or the legislation. System level student data is needed. Perhaps, as the research suggested, a model like BC’s Student

Transition Project with representation from postsecondary institutions and government ministries, might be one approach that would mitigate the above concerns. Despite the above concerns, most of the research participants favoured a centralized data sharing approach. Furthermore, if in the pursuit of higher levels of accountability within education and a move towards more data informed decision-making and policy development, reliable and consistent high-quality province wide data source(s) would be needed. Additionally, at present, OUAC (Ontario University Application Centre) and OCAS (Ontario College Application Service) collect significant amounts of data. Participants in the study noted that OCAS was now offering data analytics on the student application data it collects for a fee. This has become a 'valuable source of business intelligence for many colleges' (Baumal, April 2018, p. 71).

To ensure compliance with FIPPA regulations the research found that the transfer of students' personal data, required data sharing agreements. At minimum the components of a data sharing agreement or MOU must include the following:

- Compliance with FIPPA Regulations;
- An indication of the institution's legislative authority to collect and disseminate data;
- A business case for data sharing, articulating the benefits of the research when weighed against privacy concerns;
- Clarity on the personal information to be shared;
- An indication of how the personal information will be used;
- An indication on whether there will be future disclosure of the data;
- A de-identification process;
- Clarity on how data will be shared and / or linked;
- How data accuracy would be maintained;
- Security processes and breach protocols;
- A release model for the reporting of data; and
- A termination date for the sharing agreement.

In situations where data can be de-identified, FIPPA provides some guidelines on how this would be achieved, and this process must also be included in data sharing and confidentiality agreements. If the data is to be de-identified, the MOU must state:

- The variables that will be de-identified and how they will occur;
- When the de-identification will take place;
- Who will do the de-identification and confirm that a confidentiality agreement has been signed and that they have proper and traceable access to the identifiable data;
- How long the identified records will be maintained;
- If there will be a link or key between the de-identified and identified data;

- The methods in place for ensuring that individual records are not identifiable and how variables may need to be grouped together to help in the de-identification;
- The methods in place for ensuring accuracy of the de-identified data; and
- If written notice will be provided upon successful de-identification of the data (Baumal, April 2018, p. 61).

Data sharing of system level educational data is greatly needed in Ontario. As well, there was a general acceptance across the system that student mobility is not restricted to provincial jurisdictions or simply within countries such as Canada or the United States. It is a global phenomenon and as such there is growing interest in student data portability. Sharing data across jurisdictions has become a necessity, not only to facilitate international student mobility, but also to seek greater understanding of world-wide and or regional student movement patterns. It is this recognition that has prompted the Association of Registrars of the Universities and Colleges of Canada (ARUCC), to sign the Groningen Declaration on May 6, 2015. Canada is therefore a signatory to the **Groningen Declaration,** which attempts to create and promote a 'digital student data ecosystem' to make 'digital student data portability happen'[5](Baumal, April 2018, p. 80). The implications of being signatory to the above declaration will need further exploration, particularly as institutions and the province decide on the next steps towards achieving broader student data transfer in education.

[5] *http://www.groningendeclaration.org/*

References

Baumal, B. (2018, March 26). Data Sharing in Academic Collaboration: Final Presenation. Toronto, Ontario: ThinkLounge Research.

Baumal, B. (April 2018). Data Sharing in Academic Collaborations. Toronto: ThinkLounge Research.