Data Sharing in Academic Collaborations

DRAFT REPORT

Brian Baumal
THINKLOUNGE RESEARCH | BB@THINKLOUNGE.CA

MARCH 19, 2018

TABLE OF CONTENTS

List of Abbreviated Terms

Abbreviation Used	Full Term
BCCAT	British Columbia Council on Articulation and Transfer
CDW	Central Data Warehouse
CIHR	Canadian Institute for Health Information
DSA	Data Sharing Agreement
EDI	Electronic Data Interface
FIPPA/FOIPPA	Freedom of Information and Protection of Privacy Act
FTP	File Transfer Protocol
IPC	Information and Privacy Commissioner
ISA	Information Sharing Agreement
IT	Information Technology
HEQCO	Higher Education Quality Council of Ontario
MAESD	Ministry of Advanced Education and Skills Development
MOU	Memorandum of Understanding
NSC	National Student Clearinghouse
OCAS	Ontario College Application Service
OEN	Ontario Education Number
ONCAT	Ontario Council on Articulation and Transfer
ONSIS	Ontario Student Information System
OUAC	Ontario University Application Centre
PEDAL	Public Economic Data Analysis Lab
PEN	Personal Education Number
PSE	Post Secondary Education
REB	Research Ethics Board
STP	Student Transition Project
UOIT	University of Ontario Institute of Technology
YITS	Youth in Transition Study

Purpose, Objectives & Methodology

Purpose and Objectives of the Research Study

The overall broad objective of this study is to assist institutions in Ontario with transferring student data between them. Strictly for the purposes of this report, data transfer is defined as record-level student data created by and kept at an institution and can be exchanged for the purposes of:

1) General registrar purposes, student registration or record transfer;

2) Administration of collaborative programs or co-registration programs;

3) Student redirection;

4) Administration of and research into articulation agreements and pathways between institutions; and

5) Institutional research and planning purposes.

There are two broad objectives for this study. The first is to assist institutions with the process of data sharing by providing:

- Insights into best practices for data sharing;
- A data sharing framework that outlines a broad understanding of the factors that are involved in data sharing among institutions;
- A template guide for a draft Memorandum of Understanding (MOU) that can be used by institutions when sharing data; and

The second objective is to understand the broader issues involved in data sharing in among institutions in specific and in Ontario in general and to provide suggestions for improving data sharing between institutions and through the province. The scope of this objective involved investigating best practices that can be applied from other sectors and examining how jurisdictions like BC and the United States are addressing data sharing.

There is a link between the first and second objectives in that both sets of goals for this study examined:

- The regulatory framework for data sharing at post-secondary institutions;
- The kind of data sharing that is occurring at post-secondary institutions, and for what purposes;
- Issues that may impact data sharing such as legislation, IT issues, legal issues, privacy concerns and ethics; and
- Existing data sharing agreements and arrangements from organizations like OCAS, OUAC and MAESD to learn how data sharing will evolve in Ontario as data from these large sources is combined and shared. This also includes plans to work with and share the OEN.

Research Structure, Committee & Methodology

This research was funded by the Ontario Council on Articulations and Transfers (ONCAT), and was lead by York University. Six institutions agreed to be part of the study, and each institution had representatives on the Steering Committee for the research assignment.

The research methodology consisted of three phases:

- A general literature review using Google searches to locate articles, publications and information that addressed the objectives of the research. Search terms such as "data sharing MOU", "Best practices in data sharing" and "data sharing agreements" used as terms. Also, HEQCO, ONCAT and BCCAT (British Columbia Council on Articulation and Transfer) research reports were consulted. The researcher also relied on documents provided by the client team which focused on specific agreements drafted in support of specific research activities that involved exchange of student data between them. After the literature review was conducted, a draft report was submitted to the committee to provide a draft framework and as a project check-in.

- A total of 30 one-on-one interviews were conducted with the six institutions involved in the study. Each institution was asked to provide a list of names to be interviewed at their institution. Individuals on the steering committee provided contact information to the researcher, and email and phone contact was made in order to secure interviews. At least three attempts were made to contact selected individuals before abandoning the attempt.

  A total of 30 interviews was anticipated, and was broken out by attempting to contact one individual at each of the six participating institutions in each of the following five areas:

  - Administration/Management;
  - Programming/Academics;
  - Registrars;
  - Legal/Privacy/Ethics; and
  - IT.

  Where more than one name was provided for each function, random selection was used to chose a name.

  The six institutions at which the interviews took place were:

  - York University;
  - Seneca College;
  - Trent University;
  - Durham College;
  - UOIT; and
  - Fleming College.

- Two focus groups lasting two hours each were conduced at York University. Participants were invited to attend in person or over the phone. The purpose of the groups was to serve as a check of the information that would be included in this report.

- It should be noted that Research Ethics Approval (REB) was received by each institution prior to the start of any qualitative research at the institution. Informed consent was obtained prior to the start of any interview with participants. REB's require that qualitative information in this report be kept confidential and that information in it should not be included that could identify any individual. The report has been structured in that regard.

Executive Summary

Purpose, Objective & Method

The overall objective of this study is to understand and increase willingness and capacity to engage in data sharing among institutions and throughout Ontario to improve student outcomes. As part of this mandate the research also provides a framework, best practices and guidance towards creating terms that would go into a Memorandum of Understanding (MOU) or Data Sharing Agreement (DSA) between institutions who engage in such practices. A thorough literature review using Google searches, 30 one-on-one interviews across the six participating institutions[1] in this research with representatives from five operational areas[2] across the institutions and two focus groups were conducted in support of these objectives.

This Executive Summary is presented in two parts:

- Findings that impact willingness to share data among institutions and throughout Ontario; and
- Guidance towards implementing data sharing among institutions.

Findings that Impact Willingness to Share Data among Institutions and throughout Ontario

<u>Broader Thinking is Important</u>
Overall in order to improve data sharing among institutions, broader thinking must occur between institutions and throughout the entire Ontario post-secondary system. Virtually all examples of data sharing encountered in this research are discrete and one-to-one relationships. That is, one program shares information with another program, or one institution shares information with another institution. These data exchanges occur on a fixed time schedule or period and the transfer is done via a closed system. That is, a file is created at the sending institution and exchanged via FTP or secure USB or email and then loaded into the system of the receiving institution at specific times of the year.

<u>Purpose of Data Sharing Involves Administration and Planning</u>
The purpose of data sharing is largely twofold. The first is for administration of general registration, and as a subset of that for the administration of specific programming between institutions such as collaborative programs, articulated programs, co-registration programs and student redirection. In general, this system of exchange is viewed satisfactorily and is generally functional and well-serves the institutions. Upon further inquiry though, some institutions indicated that a more open EDI (Electronic Data Interchange) system would be beneficial to make the exchange process more efficient.

The second less common use of data exchange is for strategic planning or mobility research purposes. While the frequency of exchange may not be as common as exchanges for operational and registrar function, the importance is very high. Institutions want to examine student movement in and out in order to measure success and implement planning. This need spans the entire institution. On a very broad level, it includes performance measurement, financial issues, strategic direction and even

[1] The six participating institutions were: York University, Seneca College, Trent University, Fleming College, UOIT and Durham College.

[2] The five operational areas were: Administration/Management, Programming/Academics, Registrars, Legal/Privacy/Ethics, and IT.

negotiations with various government ministries. There are even broader societal issues such as examining access for under-represented groups. In other areas of the institution strategic planning facilitated by data exchange can involve administering collaborative programs between institutions, knowing how many students are coming from the sending into the receiving institution so that faculty and space planning can occur. In other cases, institutions want to look at how to realign their programs or determine successes and outcomes of their students. Data exchanges can also be done for as few as under a dozen students in a particular program to understand pathways and the effectiveness of articulation agreements. What is common though is that student mobility and institutional planning are truly fundamental issues that institutions want and need answers for. In sharing data for these purposes, institutions may agree to generally cooperate together and exchange data. For example institutions that are in close proximity, institutions that share students or institutions that have a “transition focus” will enter into large and general data sharing agreements that can facilitate investigation of these research and planning issues. Other institutions agree to exchange very large data sets for the purposes of research projects that would involve very large scale and robust sets of data.

Open Repositories of Data Represent Potential for Improvement but are Challenging to Implement

It should also be noted that institutions can conduct research into the above issues without exchanging data with other institutions. That is institutions can use numerous internally generated data sources to obtain answers to these questions (e.g. they can look at and analyze application data, they can see where students send their transcripts). However, one of the fundamental issues with data exchanged either with other institutions or data generated internally is that it is recognized that the data is substantially incomplete in one way or another. That is, the data generated in exchanges with another institution only covers those institutions and data on student mobility that is generated from internal sources also has various issues associated with it. Within the qualitative research for this project, and within the literature reviews, when these limitations of data exchanges were discussed the solution was to create an entire open system-wide process or repository of data that would allow institutions to track student mobility between institutions. This is one of the key findings of this research – that an open repository would be of significant benefit to institutional planning and institutional operations.

Other jurisdictions, such as BC and the United States have much more advanced data sharing structures and practices regarding the issue of student transfer. The notion of a full, open exchange of data such as this exists in the United States with the National Student Clearinghouse. To a lesser extent, BC’s Student Transitions Project (STP) uses the province’s Personal Education Number (PEN) to provide research and reporting on student movement between institutions. Internationally the Groningen Declaration is encouraging movement towards electronic exchange of student transcripts and records, and Canada has signed on to it[3].

Another key finding of the research is that moving towards a full and open exchange of data between institutions in Ontario is likely very unachievable in the short term. Respondents indicate that technology is not the issue and similarly, many indicate that they already report a significant amount of data to the Ministry, so the legal and privacy issues are likely dealt with on a broad level.

[3] https://www.universityaffairs.ca/news/news-article/canada-joins-network-to-improve-the-international-exchange-of-student-data/

However, what may be one of the larger impediments to open data sharing is the limited use of the Ontario Education Number (OEN). On the one hand, the OEN is perhaps the most important development in data sharing in Ontario over the last many years. However, it is well-recognized that its use in research and analysis has been extremely limited. For example, up until recently there was legislation that expressly prohibited the collection and dissemination of the OEN, even by educational institutions. A literature review did not reveal any significant research initiatives being conducted on MAESD data using the OEN to track student movement. It should be noted though, that even in BC and the US, free and open access to student data from central repositories of information does not exist. Rather there are structures in place (i.e. the National Student Clearinghouse and the STP project) to process requests for student level data and disseminate it accordingly. It is likely that if more open access will occur to OEN-level data, it will occur through structures set-up to broker information between MAESD and requestors to ensure privacy and proper use of sensitive record-level student data. Moreover, these organizations conduct and publish research and make available papers that answer many basic questions about student transfer and mobility in their jurisdictions. The Clearinghouse in the United States can provide anonymized record-level data under certain conditions to certain requestors.

Legal, Privacy and Technology Issues do not Represent Impediments to Data Sharing between Institutions

Also, on an institution-to-institution level, legal and technological issues do not seem to be barriers to a more open exchange of data, and generally speaking implementation of data sharing agreements between institutions tends to be a fairly straightforward process, with which little difficulty is experienced.

However, Some Substantive Barriers to More Open Data Exchange Do Exist

Given that technological, legal and contractual issues do not generally pose significant hurdles to data exchange between institutions, the research examined other issues to implementing a more open exchange of data, whether it is on an institution-to-institution level or a fully open provincial exchange. What was found is that the issues inhibiting more open exchanges of data, regardless of scope of the exchange seem to be the same. That is, the research identified the following barriers to more open exchange of data in most any setting:

- Translation of records between institutions. On an institution-to-institution level data exchange is made complicated by the fact that institutions may have very different operations, procedures, grading schemes, enrollment and schedule requirements for their records. Summed-up, this is a notion of having to translate records between one institution to another. Among institutions that engage in this, they indicate that the technological solution is not the problem, but rather identifying and addressing these issues is. Moreover, it was identified that translation issues that occur between two institutions only get magnified the broader the exchange of data becomes, so creating a province-wide data exchange or open system becomes, according to many research participants, mind-boggling;

- Matching of records is also an issue encountered among some institutions that do not have common student numbers. Some institutions, however, do share common student numbers and identifiers and indicated that having this is a significant time saver for them and a very large benefit when it comes to sharing data. While the OEN may be used for this in the future, there is a recognition that at present there may be issues with using the OEN to match records because institutions are unsure of how they can use it and the fact that it does not cover everyone in, or entering into, the system; and

- Agreement on what can be said about or done with shared data. Data sharing is seen a beneficial and worthwhile among institutions in order for better operations and planning. However, there is some concern over how the data can be used, how it will be circulated and what can be said or interpreted from the data. This remains a particularly large concern in light of the translation issues identified above. That is, an institution may be concerned that another institution is not interpreting data correctly before drawing conclusions about student movement or in using it for planning purposes.

It should be noted that all three of these issues are very standard terms and conditions in any data sharing agreement, and that institutions are aware that these are issues that have to be addressed.

There are Ways to Increase Data Sharing among Institutions

What will encourage more data sharing between institutions is the relationship that exists between them. From observing the research results and interactions between institution in the study, those that view themselves as close to one another, or even as “transfer institutions” were able to think more openly about how to make data sharing happen. They also thought more strategically about key student movement issues and how to work with other partners to gather and share that information. While the research showed that one-to-one data exchanges may be based on close relationships between institutions, moving into a broader data exchange situation will require that institutions have very close relationship with each other.

Beyond having close relationships, in order to increase the amount of data sharing between institutions, it is likely that data sharing should occur at a research and planning level in consultation with senior administration. Senior administrators often have key strategic questions that they need answered, and those in the planning area are often aware of them, so institutions involved in collaborative data sharing activities will be able to reach a common ground on the objectives of the data sharing exercise. However, the reason for more data sharing, such as a full EDI to start within research and planning is that translation and matching are not as critical to this area of operations as they are to registrar and program management. That is, if data exchange is going to be used to register students, transfer grades and other information, it has to be perfectly translated over to the other institution very quickly so that the student does not experience a delay in service or any concerns about their transfer. However, within research and planning, timelines are not as immediate, and conclusions from data can be drawn without having a completely precise data set. Moreover, researchers will often review the data for translation and matching issues. Those parameters can be discovered, shared and addressed when transferring data more for registrar and operational purposes, but starting more open data sharing in the research and planning area largely removes the barrier of full and complete translation of data which is a necessity in the operational area.

Province-Wide Data Sharing is Possible, but not Likely in the Immediate Future

Looking beyond institution-to-institution data exchanges and more towards province-wide data sharing, many participants pointed to the fact that the Ministry of Advanced Education and Skills Development

(MAESD) is collecting a large amount of data and that the Ontario Education Number (OEN) has the potential to facilitate data exchange. Though there were no interviews conducted with MAESD for this research, secondary research and an examination of other jurisdictions revealed that some of the same issues involved in exchange of data from institution-to-institution will be an issue. However, MAESD will likely have a number of other privacy and legal concerns to contend with. It would seem unlikely that Ontario can move towards a full clearinghouse system that is available in the United States. However, a smaller step may be to establish an organization, group or committee that can facilitate institutional access to research with OEN data whether it is identifiable record-level data, anonymized record-level data or even summary reports of data that can shed light on student movement among institutions throughout the province, using BC's STP structure as a template. In the STP, a full research committee and an extensive data sharing agreement govern when, how and with whom PEN data can be shared. The goal of the committee structure is to ensure that researchers have fair and equitable access to such data. Also, there were reports of OCAS providing some student movement information among institutions. That is, OCAS will indicate which program an applicant chooses to enter from those chosen on their application.

Small Steps Now Can Lead to Larger Changes

In summary, participants in the research feel that the last significant barrier to a large data sharing arrangement throughout Ontario has been crossed with the creation of the OEN. However, moving towards a large open exchange of data, whether it is throughout Ontario, or whether it is a more fluid EDI exchange between two institutions is likely quite far away because of a number of existing issues that need to be addressed. However, intermediate steps can be taken. On a broader level, Ontario could consider a research project similar to the STP in BC and two institutions that have a close relationship can attempt to create an EDI exchange that would focus on addressing some of the research and planning questions that can be answered through a more open exchange. Leadership would be required in the translation, matching and agreed-upon use of the data between institutions, but the benefits of being able to more definitively tract student movement will be very worthwhile. Moreover, as data transfer issues for research and planning purposes get uncovered and solved, the learning can be applied to data transfer for registrar and operational areas of institutions.

Guidance toward Implementing Data Sharing among Institutions

Statutory Authority to Collect and Share Data

The literature search for this study indicated that each institution in Ontario is granted the statutory authority to collect personal student level data and use that data for the purposes of administering the functions of the institution. The legislation around this tends to be very broad, giving institutions significant leeway in how they can collect and use the student level data. In examining different institution's data collection and usage policies, Carleton University had a very clear description of its authority to collect and use information for research and planning purposes and how those purposes relate to FIPPA (the Freedom of Information and Privacy Protection Act). Specifically:

> *The Freedom of Information Protection of Privacy Act recognizes the legitimate need to collect personal information in order to carry out ones 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…*

> *In the case of a University, the University Act gives only general authority for Carleton University's educational program… The University's operating program is any series of functions designed to carry out all or part of its mandate and an activity is an individual action designed to assist in carrying out an operating program…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. 2.2 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. 2.3 Any written agreement or contract with an outside agent should stipulate that the collection, protection, retention and disclosure of personal information will be governed by the Act*[4].

The main message from the passage above is that institutions can collect and disseminate most of the data they possess if it serves the purpose of helping the post-secondary institution function.

FIPPA Governs Data Sharing among Ontario Post-Secondary Institutions

While post-secondary institutions can collect and use data to meet their mandate, the use, transfer and sharing of that data is governed by FIPPA. Arguably, compliance with FIPPA (Freedom of Information and Protection of Privacy Act) constitutes best practice when it comes to creating and executing data sharing agreements and MOU's between post-secondary institutions in Ontario. Some of the key aspects of FIPPA include:

- Its recognition of the fact that post-secondary institutions can use and share data for research, transfer, articulation and registrar purposes since those functions fall within the scope of the original purpose of the data collection;

- FIPPA only applies to personal information. Besides "tombstone data" such as name, address, birth date, personal information defined by the act includes data concerning the education of an individual and any identification number assigned to the individual, including a student number or the new Ontario Education Number (OEN). As such, the Information and Privacy Commissioner (IPC) of Ontario, the authority that manages FIPPA, makes a point of indicating that data that is properly de-identified is not subject to its regulations;

- The IPC interprets the spirit of FIPPA in the following way, suggesting that the purpose of information sharing needs to be carefully considered prior to engaging in a data sharing activity. Specifically, Tom Wright, the Privacy Commissioner of Ontario in 1995 mentioned:

  > *Sharing personal information between two organizations runs counter to two of the most fundamental principles of data protection — that personal information should be collected directly from the individual to whom it pertains, and should only be used for the purpose for which it was collected [with limited exceptions]. Data sharing respects neither of these principles. Data sharing involves information that has been collected indirectly, and used for a purpose which may not have been intended at*

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

> *the time of the original collection.*[5]

Finally, Regulation 460 of FIPPA indicates that data sharing agreements are required to be used when sharing personal data and provides some specifics about how they should be structured. Specifically, the regulation states:

> 1. The person shall use the information only for a research purpose set out in the agreement or for which the person has written authorization from the institution.
>
> 2. The person shall name in the agreement any other persons who will be given access to personal information in a form in which the individual to whom it relates can be identified.
>
> 3. Before disclosing personal information to other persons under paragraph 2, the person shall enter into an agreement with those persons to ensure that they will not disclose it to any other person.
>
> 4. The person shall keep the information in a physically secure location to which access is given only to the person and to the persons given access under paragraph 2.
>
> 5. The person shall destroy all individual identifiers in the information by the date specified in the agreement.
>
> 6. The person shall not contact any individual to whom personal information relates, directly or indirectly, without the prior written authority of the institution.
>
> 7. The person shall ensure that no personal information will be used or disclosed in a form in which the individual to whom it relates can be identified without the written authority of the institution.
>
> 8. The person shall notify the institution in writing immediately if the person becomes aware that any of the conditions set out in this section have been breached. R.R.O. 1990, Reg. 460, s. 10 (1).[6]

The Data Sharing Framework

The research mandate required the production of three pieces of information to assist institutions in planning data sharing agreements: 1) A Data Sharing Framework; 2) Provision of Best Practices in Data Sharing; 3) Provision of considerations for a Draft MOU or Data Sharing Agreement (DSA) between institutions. All three are provided in the report, and are generally summaries of the broader themes discussed throughout the report. The Data Sharing Framework appears on the next page and is a summary of many of the items discussed in the report. Summaries of Best Practices and an MOU/DSA draft appear in their relevant sections in the report.

The Data Sharing Framework outlined on the next page is founded on whether or not the data being shared is considered personal information as defined by the Freedom of Information and Protection of

[5] http://www.ontla.on.ca/library/repository/mon/3000/10301262.pdf

[6] https://www.ontario.ca/lawdats/regulation/900460

Privacy Act (FIPPA). The Framework then encourages collaboration with other areas of the institution to enhance the data sharing activities, and then discusses how to plan and finally execute a data sharing agreement.

DATA SHARING AND MOU FRAMEWORK

Understand Legal and Privacy Issues

Data Sharing Agreements are only required if institutions plan on exchanging legally protected data that is defined as private or identifiable information. If the data to be exchanged is not private, a Data Sharing Agreement is still recommended.

Institutions must determine if the data they are going to share is private and legally protected. Any data that can identify a student (i.e. record level data or summary data that can identify a student) is protected, as is data that contains information on any of the following attributes: 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; *Any identifying number, symbol or other particular assigned to the individual*; Address, telephone number.

Institutionally, when planning to exchange data, ensure the following: 1) A business case to exchange data and breach privacy must be made; 2) Ensure that students have been informed and consented to data exchange by checking the institutional website, and application; 3) Consider having students sign consent to have their data exchanged, especially for registrar and shared programs.

Collaborate

Involve other areas of the institution for broader perspective

Other areas of the institution may have strategic or operational issues that could benefit from involvement in data being shared. Also, other areas may have input or suggestions into how to share data more effectively.

Plan

Use best practices in data sharing to properly execute the exchange.

Best practices in exchanging data include: Exchanging only relevant data; de-identifying data for research; determine how to link data; determine data access levels; plan on secure data transfer; determine storage and verification plans; plan for data breaches; determine reporting conventions and what can be said about the data.

Agree

Draft a data sharing agreement using Best Practices

After planning, create a data sharing agreement focusing on: Compliance with FIPPA Regulation 460; The legislative authority granted to share data; Stating the business case; Indicating the personal information to be shared and used; Plan for future disclosure of the data; Indicate if the data will be de-identified for research; Indicate how data will be shared and linked; Indicate accuracy and security measures; Indicate if and how the data is to be released; Indicate termination of the agreement.

1. Types Of Data Sharing Requests

This Report Considers Five Types of Data Sharing Requests

Data sharing between institutions can be categorized into two groups each having specific functions within them:

- General registrar purposes, student registration or record transfer. This may occur when students transfer into an institution and record-level data about the student may be requested or transferred and generally occurs when a student generally transfers into an institution, as opposed to transferring in to an institution for reasons outlined below. The difference with this kind of transfer is that institutions may not have agreements with other institutions for the transfer of data directly.

- Administration of collaborative programs or co-registration programs. This data sharing occurs among institutions where institutions have signed agreements to create relationships that allow for a collaborative program (i.e. one credential is awarded but where learning occurs at two institutions) or co-registration programs (i.e. where credit from courses at one institution are transferred for credit at another institution).

- Student redirection. Where students are given the option to attend another institution in order to better take advantage of courses and offerings at the receiving institution. Redirection programs may be formalized (e.g. any student who meets specific criteria should be given the option to attend the receiving institution), or they may be implemented by academic advisors or staff on an individual basis.

- Administration of and research into articulation agreements and pathways between institutions. Where institutions create credit transfer agreements or arrangements with other institutions so that students can transfer from one institution to another. Sometimes student records may follow the student under an articulation agreement for administrative purposes, and in other instances student outcomes and/or mobility may be researched in order to determine potential articulation agreements and pathways, or the results of articulation agreements (i.e. are students who are given credit from a sending institution succeeding at the receiving institution, and what factors influence outcomes).

- Institutional research and planning purposes. Data is exchanged to understand student movement between institutions for any number of purposes. Examples of data exchanged for this purpose could include:

  - Planning for space and faculty requirements (i.e. how many students from a sending institution will be attending a receiving institution);

  - Determining success metrics and overall student outcomes on an institution-wide level;

- Making strategic decisions about program planning (i.e. which programs should be offered, which programs are drawing students in and from where); and

- Understanding student movement and access throughout the system, where some institutions house research units that attempt to investigate broader student movements that occur among students throughout the province.

It should be noted that conceptually this report treats all data sharing similarly given the fact that the same data sets are exchanged for all of the above purposes. However, there may be some very practical differences in data exchange when it is undertaken for the various functions. For the purposes of this report, the biggest difference would be the fact that data for program, administration and registrar purposes cannot be "de-identified". That is, the identity of the student is fundamental to these data exchanges. However, data for research purposes can be de-identified since the identity of the student is not necessary to analyze data in aggregate for research purposes. The necessity of transferring identifiable data and de-identification of data creates some significant differences in how data sharing in general, and data sharing agreements in specific are drafted for these functions. The report highlights these differences where necessary.

Also, data transfers in this case exclude the exchange of actual transcripts and does not directly cover data transferred to MAESD and among OUAC (Ontario University Application Centre) and OCAS (Ontario College Application Service). Transfers of transcripts require student consent and have very defined purposes, so while they are protected by FIPPA legislation, they are not the subject of the research. Also data exchanges occur between institutions and MAESD, whereby institutions report data into MAESD for a number of specific purposes. Similarly data is exchanged among institutions and OUAC and OCAS, but is not covered by the draft agreements and frameworks in this report. However, attention is given to these transfers in regard to trends and influences that are occurring in data transfer throughout Ontario. For example, as MAESD collects data using the relatively new OEN, it has the potential to track student movement in a way that is very valuable for institutions, and given that institutions exchange data with and amongst themselves for this same purpose, having a sense of how MAESD is engaging in this function is worthwhile. Similarly, OCAS is providing some performance measurement data to institutions in that, for example, it is providing data on which courses students actually enroll in based on the enrollment choices available to them when they apply to Ontario Colleges. This performance measurement function is also a purpose of data exchanged between institutions, so OUAC and OCAS data transfers are discussed in a broader context in this report.

The Nature of Transfer Data Makes Privacy a Fundamental and Legal Issue

As mentioned above, personal and/or identifying information about a student will have a high likelihood of being transferred between institutions so that records from both institutions can be matched so that data can be analyzed for research and policy or used for administration. Institutions need a way of linking data between them so that the records about the student can be joined together from the sending institution at the receiving institution. In some cases that identifying information may "tombstone data" such as name, address or date of birth. In other cases, it may be a student identifier number, or even a combination of tombstone data and student numbers. Under Ontario law, a student ID number (including the newly implemented OEN) – even if it is transmitted without any other identifying information such as name or address – is considered personal or identifying information[7]. Ontario law has very specific requirements about the transfer of information that is personally identifiable, so much of the best practices discussed in this paper are to comply with the legal requirements necessary for post-secondary institutions.

Some examples of how student information is transferred for the purposes of research and planning is described below. The results come from documents provided to the researcher and from a literature review of information and reports published by HEQCO, ONCAT and BCCAT. These projects required the exchange of personal data which is protected under Ontario Law:

- Some institutions, even though they operate separately, share the same student numbering system so that the student will have the same identifier at both institutions, should they decide to transfer from one to another. The Institutional Research Offices in two such institutions signed an agreement to exchange student information identified by this common student number for the purposes of administration, research and planning.[8]

- A researcher received data from some Ontario universities containing names, addresses and other identifying information so that institutional records could be matched with Statistics Canada income data to determine labour market outcomes and measures for transfer students. The researcher was not able to chart the transfer pathways of students between institutions. Rather, one of the fields in the data records indicated the student's previous educational pathway, whether it was from another PSE, or whether it was from a post secondary institution or other source.[9] For this study the researcher followed two strict protocols with the data transfer that could represent best practice. First, Statistics Canada did the data matching between personal information and income, and then stripped the personal information from the data prior to returning it to the researcher such that the income data was anonymous when it was returned. Second, the researcher followed Statistics Canada's usage and disclosure rules to help further ensure anonymity of the data.[10] That is, the researcher respected the rules of the organization that appended data to his own.

- Institutional Research Offices at some universities and colleges exchanged student information to research student pathways and outcomes for transfer students between their institutions. For example, a study done between two institutions with no common key or identifier between them required records to be matched on tombstone information such as names, date of birth, address and other information.[11] Examination of the original REB documents and data sharing agreements showed that the researchers followed protocols to keep the data anonymous,

[7] https://www.ontario.ca/laws/statute/90f31#BK53, s2(1)(c)
[8] MOU between UOIT and Durham College dated January 31, 2011 provided to researcher by Project Authority
[9] http://www.oncat.ca/files_docs/content/pdf/en/oncat_research_reports/2016-08-Final-Report-University-of-Ottawa-How-Student-Pathways-affect-Labour-Market-Outcomes.pdf
[10] http://www.oncat.ca/files_docs/content/pdf/en/oncat_research_reports/2016-08-Final-Report-University-of-Ottawa-How-Student-Pathways-affect-Labour-Market-Outcomes.pdf, Footnotes P6
[11] http://www.heqco.ca/SiteCollectionDocuments/Transfer-Pathways-in-PSE-ENG.pdf (Smith, R., Decock, H., Lin, S., Sidhu, R., & McCloy, U. (2016). Transfer Pathways in Postsecondary Education: York University and Seneca College as a Case Study. Toronto: Higher Education Quality Council of Ontario.)P18

including signing a data sharing agreement which specified data would be stripped of all identifiers after matching and before sharing with other researchers.

- Institutions in this study transfer data for the purposes of redirection, with permission of students. In some cases the redirection is from university to college, so the data sharing saves students the cost of having to apply through OCAS.

- In talking with institutions in this study, many exchange information to administer programs between them. Data sharing agreements are in place and terms discussing transfer and use of the data are discussed. Data in files is transferred at discrete times via email, USB and FTP complying with institutional privacy policies.

- In BC, there is the Student Transition Project (or STP), which represents a model for MAESD, ONCAT and post-secondary institutions to follow to realize the potential of the OEN in tracking information. The goal of the STP is to track student data from public school into the post-secondary system and then to track movements and results through that system. The STP uses BC's Personal Education Numbers (PENs) to accomplish this. [12] One very high-level chart produced by the STP, for example, to describe student mobility at the post-secondary level can be seen here:

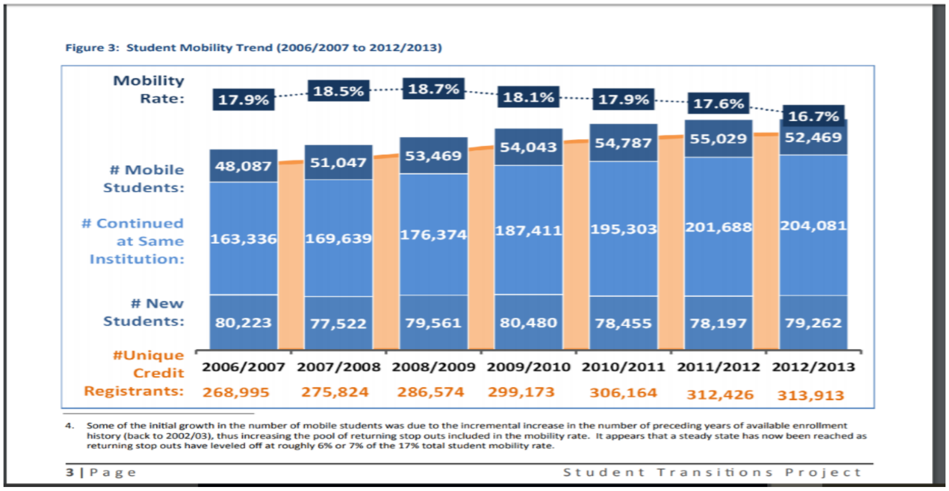


Figure 3: Student Mobility Trend (2006/2007 to 2012/2013)

4. Some of the initial growth in the number of mobile students was due to the incremental increase in the number of preceding years of available enrollment history (back to 2002/03), thus increasing the pool of returning stop outs included in the mobility rate. It appears that a steady state has now been reached as returning stop outs have leveled off at roughly 6% or 7% of the 17% total student mobility rate.

3 | Page Student Transitions Project

By contrast, an example of where the subject of a report is student transfers and where Ontario law would not apply is one where only anonymous student data was requested. "Giving Credit Where Credit Is Due" [13] was a study that examined transfer and mobility patterns among Ontario colleges, where only anonymous data was requested, no linkages between data sets occurred (i.e. no linkages between data supplied by ONCAT to survey data so hence no identifying key needed to be provided) and ONCAT itself emailed respondents to participate in the survey, such that personal information was not

[12] https://www2.gov.bc.ca/gov/content/education-training/post-secondary-education/data-research/student-transitions-project

[13] http://www.oncat.ca/files_docs/content/pdf/en/oncat_research_reports/2014-31-Final-Report-Credit-where-credit-is-due-understanding-credit-transfer-in-Ontario-Colleges.pdf P30

released by ONCAT to the researchers. Within institutions researched for this project many will exchange aggregated summary reports to track student movement. FIPPA protocols do not apply to these reports but many include statements specifying confidentiality of the summary reports and/or limiting use of the report to a specific purpose.

2. The Legal Framework In Ontario

This section is divided into two broad categories. The first addresses the basic legal authority given to institutions to collect and manage data, and the implications that has on data sharing. The second section addresses the laws which govern how those records can and should be managed. In this case the relevant Ontario law is FIPPA.

Each Institution in Ontario Has an Ontario Act that Allows It Broad Authority to Create Student Records

Prior to discussing the laws around privacy of records maintained by post secondary institutions, it is worthwhile to understand the authority by which institutions have the legal right to create and collect information about their students. It appears that each post-secondary institution in Ontario, by act of statute, can create, collect and manage student records. While each act may have slightly different wording or interpretation, the acts that govern post-secondary institutions allow Governing Councils or the institution itself to take actions that are necessary to administer the business of the institution including functions related to planning, registration and statistics creation and management. As some institutions' websites imply directly, the creation and maintenance of student records that contain a broad amount of information are considered necessary for the administration of the institution.

A full review of every Ontario institution's acts and privacy policies is beyond the scope of this project. It is assumed that all Ontario acts contain the same rules and statues necessary for an institution to create and populate a student record. The brief review highlights the partner institutions to this study, where information could be found. Also, Carleton University, which provides a particularly relevant summary of the interaction of the legal framework to data sharing agreements is discussed throughout this section.

<u>York University</u>

The Privacy section of York University's website says that it collects data under the authority of the York University Act (1965). There is a statement which indicates "By applying for admission to York University and by enrolling in a program at the University, students consent to the collection of their personal information by York University for educational, administrative and statistical purposes. The information is needed… for related recordkeeping purposes." There is also a section on the webpage highlighting indicating information may be shared with the following parties to facilitate fundamental activities:

- 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.[14]

The notation of York sharing data with other institutions with which it runs collaborative programs addresses the registrar functions of exchanging data with other institutions for the purposes of managing collaborative programs.

Trent University

Trent University's Privacy Statement indicates:

> *Students' personal information is collected, used, and disclosed by Trent University under the authority of Section 18(3)(c) of The Trent University Act, 1962-63… Trent University may collect and use personal information from prospective students to communicate with them about University programs, process applications, determine eligibility for admission and student awards, administer surveys, research enrolment issues, and maintain related statistical data… Once admitted and enrolled in an academic program, a student's information is used by the University to deliver academic and administrative programs and services. This includes but is not limited to: recording academic progress, creating the permanent student record, providing financial aid, delivering student services, conducting program reviews/appraisals, and communicating with students regarding University business. Personal information may also be used by the University, its authorized agents, approved researchers, and/or the provincial and federal government for statistical research purposes… Where students are enrolled in collaborative academic programs, Trent University may be required to transfer personal information to another post-secondary institution. Wherever possible, students will be provided with a separate notice explaining any information-sharing required to collaboratively administer their program*[15]

Seneca College

Seneca's Privacy Policy, compared to others, provides a definition of what personal information under FIPPA is. In data sharing agreements, listing the data that will passed between two parties that is legally considered personal information is considered a best practice. As such, Seneca's listing of what is personal information makes it clear to students that specific information collected by the College has legal protection. Seneca states that it collects the following personal information:

- *race, national or ethnic origin, colour, religion, age, sex, sexual orientation, or marital or family status of the individual;*

- *information relating to employment or educational history;*

[14] https://registrar.yorku.ca/privacy

[15] https://www.trentu.ca/administration/pdfs/CollectionNotice.pdf

- *information relating to the medical, psychiatric, psychological history, prognosis, condition, treatment or evaluation;*
- *any identifying number (e.g. S.I.N., student number), symbol or other particular assigned to the individual;*
- *home address and/or telephone number and;*
- *personal opinions of, or about, an individual for a research purpose with a research agreement*[16]

Seneca's privacy policy, like others also states that "FIPPA prescribes the use of students' personal information as necessary to accomplish Seneca's academic, pedagogical and operational activities"[17] meaning that the institution has the right to use personal information for research, planning and operational purposes.

UOIT
UOIT states:

> *UOIT undertakes to collect only the specific personal information that is required to carry out its academic mandate and perform related administrative functions. Personal information will be collected in a manner that is consistent with the Act, and its use and disclosure will be limited to the purposes for which it was intended. In addition, a number of safeguards have been put in place to ensure these records are retained and protected from unauthorized access.*[18]

Durham College
Durham College, like Seneca also enumerates what personal information is:

- *3.2.1. Information relating to the race, national or ethnic origin, colour, religion, age, sex, sexual orientation or marital or family status of the individual;*
- *3.2.2. Information relating to the educational, medical, psychiatric, psychological, criminal or employment history of the individual or information relating to financial transactions in which the individual has been involved;*
- *3.2.3. Any identifying number, symbol or other particular assigned to the individual;*
- *3.2.4. The address, telephone number, fingerprints or blood type of the individual;*
- *3.2.5. The personal opinions or views of the individual except where they relate to another individual;*

[16] http://www.senecacollege.ca/policies/fipp.html
[17] http://www.senecacollege.ca/policies/fipp.html
[18] https://usgc.uoit.ca/policy/policy-library/policies/legal,-compliance-and-governance/access-to-information-and-the-protection-of-privacy-policy.php

- *3.2.6. 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;*

- *3.2.7. The views or opinions of another individual about the individual; and*

- *3.2.8. The individual's name where it appears with other personal information relating to the individual or where the disclosure of the name would reveal other personal information about the individual.*[19]

Like other institutions, it describes the authority to collect information and how it will be used, which includes statistical purposes:

> *Personal information is collected under the authority the Ministry of Training, Colleges and Universities and will be used for educational, administrative and statistical purposes.*
>
> *By applying for admission to Durham College and by enrolling in a program at Durham College, students consent to the collection of their personal information by Durham College for educational, administrative and statistical purposes.*[20]

Durham's privacy policies also directly discuss its data sharing and transfer policies and the requirement to create and use data sharing and confidentiality agreements between the College and anyone with whom it shares its personal data:

> *All third-party organizations and student organizations are required to sign FIPPA-compliant confidentiality agreements with an authorized officer of Durham College before obtaining access to student personal information. Student information may only be used or disclosed in accordance with the provisions of the confidentiality agreements.*
>
> *Information may be shared with the following parties to facilitate fundamental activities:*
>
> - *Other universities and colleges with which Durham College maintains a collaborative program partnership;*
>
> - *Service providers contracted by Durham College to support business processes*[21]

Carleton University

In researching different post-secondary institution's privacy policies in Ontario, Carleton has an exceptionally detailed policy. It is referred to here generally and then more specifically throughout this report as a reference to support how post-secondary institutions in Ontario can use data for research

[19] https://durhamcollege.ca/wp-content/uploads/243-access-to-student-records-and-protection-of-privacy.pdf
[20] https://durhamcollege.ca/wp-content/uploads/243-access-to-student-records-and-protection-of-privacy.pdf
[21] https://durhamcollege.ca/wp-content/uploads/243-access-to-student-records-and-protection-of-privacy.pdf

and planning purposes and the legal requirements involving data sharing agreements institutions must follow. Specifically, Carleton says:

> *"The collection, storage, utilisation, and dissemination of Personal Information concerning members of the Carleton community is only undertaken as part of ongoing efforts by the University to ensure decision making practices are based on accurate information. The university also ensures that the information gathered for one purpose is not being used inappropriately for another, and that the privacy of an individual is not compromised by disclosure of personal information to third parties without the proper approvals."*[22]

Within a document entitled "Collection of Personal Information" it highlights the authority given to it to collect personal data as follows:

> *The Freedom of Information Protection of Privacy Act recognizes the legitimate need to collect personal information in order to carry out ones 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… In the case of a University, the University Act gives only general authority for Carleton University's educational program… The University's operating program is any series of functions designed to carry out all or part of its mandate and an activity is an individual action designed to assist in carrying out an operating program…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. 2.2 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. 2.3 Any written agreement or contract with an outside agent should stipulate that the collection, protection, retention and disclosure of personal information will be governed by the Act*[23].

What is seen from this passage is the fact that the acts prescribe very broad and general powers to collect and use data for their administrative purposes, and it is perhaps the passage above that most clearly indicates that post-secondary institutions in Ontario have the right to collect and use most any kind of data they need for the purposes of research into their own programs and students for running the institution.

The Freedom of Information and Protection of Personal Privacy Act (FIPPA) Governs the Five Areas of Student Data Transfer

The preceding discussion focused on the fact that educational institutions in Ontario have the right under law to create records necessary to administer the business of their institutions. The actual statutes in each institution's governing act are broad and allow for a broad range of data to be collected and for a broad range of uses. Some of the institutions cited above indicate that they will use the

[22] https://carleton.ca/privacy/fippa-at-carleton-university/

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

collected data for research and administration purposes. However, FIPPA is the law in Ontario that governs the protection and disclosure of the records that institutions can create by statute.

Relevant parts of FIPPA are produced and discussed below:

Educational Institutions Are Subject to The Act

- Section 2 (1) – "An educational "educational institution" means an institution that is a college of applied arts and technology or a university"[24]

- Section 2(1) – "(a.1) a service provider organization within the meaning of section 17.1 of the *Ministry of Government Services Act*" [25]. The Ministry of Government Services Act indicates "1 (e) a university, college of applied arts and technology or other post-secondary institution in Ontario"[26]

Student Numbers Along with Information Related to Education are Considered Personal Information under the Act

The act defines "personal information" that must be protected and properly disclosed. The relevant sections for analysis of student transfer data are:

- Section 2 (1) b – "information relating to the education… in which the individual has been involved" [27]

- Section 2 (1) d - the address, telephone number, fingerprints or blood type of the individual, [28]

- Section 2 (1) e - the personal opinions or views of the individual except where they relate to another individual, [29]

- Section 2 (1) c - any identifying number, symbol or other particular assigned to the individual, [30]

The last bullet is of importance because it covers student numbers or other forms of identification that go beyond the name, address and date of birth of an individual. The act considers a student number or even the OEN as "personal information" that must be protected under the act. As will be discussed later, there was also legislation that up until recently prohibited collection and transfer of the OEN. However, that legislation has recently changed and begins to pave the way for future research based on the OEN.

Use and Transfer of Personal Information, Including Conducting Research about Transfer Students, Whether Done Just Within the Institution Outside Is Permitted

[24] https://www.ontario.ca/laws/statute/90f31#BK7
[25] https://www.ontario.ca/laws/statute/90f31#BK7
[26] https://www.ontario.ca/laws/statute/90m25
[27] https://www.ontario.ca/laws/statute/90f31#BK7
[28] https://www.ontario.ca/laws/statute/90f31#BK7
[29] https://www.ontario.ca/laws/statute/90f31#BK7
[30] https://www.ontario.ca/laws/statute/90f31#BK7

Carleton University best describes the relation between FIPPA and conducting research with personal records stored at a post-secondary institution such that research is a function for which personal information can be disclosed:

> *Access to personal information for research, statistical, archival or historical purposes will be allowed under conditions specified in Sections 21 (1)e,… of the Freedom of Information and Protection of Privacy Act.* [31]

The relevant sections of FIPPA referenced by the quote above is:

- 21 (1) A head shall refuse to disclose personal information to any person other than the individual to whom the information relates except,

    (e) for a research purpose if,

    > (i) the disclosure is consistent with the conditions or reasonable expectations of disclosure under which the personal information was provided, collected or obtained,
    >
    > (ii) the research purpose for which the disclosure is to be made cannot be reasonably accomplished unless the information is provided in individually identifiable form[32]

There are two other relevant sections of FIPPA that address research purposes as well. Broadly speaking, FIPPA allows an institution to use and transfer personal information for a purpose consistent for which the information was obtained. As previously mentioned an institution is given broad authority to collect information to serve its mandate and that mandate is broadly defined to include research, planning and registrar purposes. FIPPA allows use and disclosure of personal information is permitted as long as its use and disclosure is consistent with the purpose for which it is gathered. The specific sections of FIPPA that address this are:

- 41 (1) An institution shall not use personal information in its custody or under its control except,

    (b) for the purpose for which it was obtained or compiled or for a consistent purpose; [33]

- 42 (1) An institution shall not disclose personal information in its custody or under its control except,

    (c) for the purpose for which it was obtained or compiled or for a consistent purpose;

[31] https://carleton.ca/privacy/wp-content/uploads/policy_access1.pdf
[32] https://www.ontario.ca/laws/statute/90f31#BK57
[33] https://www.ontario.ca/laws/statute/90f31#BK57

> (d) where disclosure is made to an officer, employee, consultant or agent of the institution who needs the record in the performance of their duties and where disclosure is necessary and proper in the discharge of the institution's functions;[34]

FIPPA Governs How Data Can Be Transferred and the Creation of Data Sharing Agreements

In 1995, Tom Wright then Information and Privacy Commissioner (IPC) of Ontario wrote a template for a model data sharing agreement that would comply with FIPPA. In introducing the agreement template, he points out that FIPPA gives Ontario residents "The right to privacy with respect to the protection of their personal information contained in government records."[35] In fact, the document is quite stern in saying:

> *Sharing personal information between two organizations runs counter to two of the most fundamental principles of data protection — that personal information should be collected directly from the individual to whom it pertains, and should only be used for the purpose for which it was collected [with limited exceptions]. Data sharing respects neither of these principles. Data sharing involves information that has been collected indirectly, and used for a purpose which may not have been intended at the time of the original collection.*[36]

As was discussed previously, Post-Secondary Institutions in Ontario can collect and use data for the purposes of managing their activities and institutions, and this is defined very broadly. As such, data sharing for registrar, research, articulation or planning purposes is not prohibited, and many institutional privacy policies indicate that information collected from students will be used in this manner. However, it is important to consider that FIPPA makes it very clear that due respect and consideration for the law is important in these circumstances.

FIPPA Also Requires a Signed Agreement from The Researcher

One of the primary objectives of this report is to produce an MOU template that researchers and registrars can use when transferring data about students from one institution to another. According to Carleton University's interpretation of FIPPA, there must be a "written agreement of the researcher to comply with all relevant sections in Reg. 460 (10) of the Freedom of Information and Protection of Privacy Act and with the University's policies and procedures relating to the protection of personal information."[37] As such, the regulations cited below must form a key part of the MOU and they will form the minimum standards that must be set-out in the MOU.

The relevant section of FIPPA that indicates an agreement is necessary for personal information to be released is:

- 21 (1) A head shall refuse to disclose personal information to any person other than the individual to whom the information relates except,

[34] https://www.ontario.ca/laws/statute/90f31#BK57
[35] http://www.ontla.on.ca/library/repository/mon/3000/10301262.pdf
[36] http://www.ontla.on.ca/library/repository/mon/3000/10301262.pdf
[37] https://carleton.ca/privacy/wp-content/uploads/policy_access1.pdf

> (iii) the person who is to receive the record has agreed to comply with the conditions relating to security and confidentiality prescribed by the regulations;[38]

The regulations referred to above and in the Carleton University document cited above are Reg 460 (10) of FIPPA. Those regulations state that for research purposes:

- 10. (1) The following are the terms and conditions relating to security and confidentiality that a person is required to agree to before a head may disclose personal information to that person for a research purpose:

    1. The person shall use the information only for a research purpose set out in the agreement or for which the person has written authorization from the institution.

    2. The person shall name in the agreement any other persons who will be given access to personal information in a form in which the individual to whom it relates can be identified.

    3. Before disclosing personal information to other persons under paragraph 2, the person shall enter into an agreement with those persons to ensure that they will not disclose it to any other person.

    4. The person shall keep the information in a physically secure location to which access is given only to the person and to the persons given access under paragraph 2.

    5. The person shall destroy all individual identifiers in the information by the date specified in the agreement.

    6. The person shall not contact any individual to whom personal information relates, directly or indirectly, without the prior written authority of the institution.

    7. The person shall ensure that no personal information will be used or disclosed in a form in which the individual to whom it relates can be identified without the written authority of the institution.

    8. The person shall notify the institution in writing immediately if the person becomes aware that any of the conditions set out in this section have been breached. R.R.O. 1990, Reg. 460, s. 10 (1).[39]

De-Identified Records Are Not Governed By FIPPA, But De-Identifying Transfer Records Is Challenging

The legal framework described above puts an emphasis on the fact that it protects "personal information" of individuals collected by post-secondary institutions in Ontario, with FIPPA defining exactly what constitutes "personal information". This is important because records that have been properly de-identified are not subject to FIPPA regulations. This does not mean that transfer data must be de-identified under law. Rather, a prudent practice would be for researchers to attempt methods of de-identifying data prior to transfer from one institution to another. If it is not possible, then having the smallest number of trusted individuals work to match the data and then de-identify the resulting dataset

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

[39] https://www.ontario.ca/lawdats/regulation/900460

would be appropriate. Regardless though, FIPPA does allow identified data to be released if the conditions above are met, including those around confidentiality of the data.

As was discussed earlier, de-identification of transfer data remains a challenge because of the required matching process between institutions. Moreover, if data is de-identified of basic tombstone information such as name, address, date of birth or student ID number, it is still possible for transfer data to be identifiable. For example, transfer data may include information about an individual's area of study, the year of graduation, one's age at graduation and the language a person speaks. This issue becomes more problematic if data is transferred among faculty members for the purposes of planning articulation agreement within a program area, as it is possible that data may be identifiable to faculty who know certain student's personal characteristics. As such, even without the tombstone data a record may be identifiable. As such the best practice in data sharing for research and planning is to attempt de-identification, but if that is not possible, it makes data sharing agreements even more important under these circumstances. The law under FIPPA clearly states that if personalized data is to be shared, agreements must be signed.

3 The Views Of Institutions

A total of 30 one-on-one interviews were conducted with representatives of the six partner institutions to this study. Various roles were interviewed at each institution – Academic; Administration; Legal/Privacy; Information Technology (IT); Research/Planning and Registrar. This section discusses the results of those interviews by focusing on the present situation in various institutions areas and opinions towards data sharing for all roles interviewed. The discussion then focuses on how to move data sharing forward.

At Present, there are Few Legal or Technical Barriers to Data Sharing among Institutions. Data Sharing Happens Directly among Users and Works Well for Collaborative Programs and Redirects

All participants, regardless of their area were asked about the legal and technical issues surrounding data sharing. Across all roles there is a good understanding of the legal and technical frameworks regarding data sharing between institutions, to the point where participants generally did not describe these as significant barriers to data sharing in any way. Many participants indicated that their websites clearly state the circumstances under which data would be shared with other institutions, and whether permission would be required to share data. Participants did not encounter many, if any, technical issues regarding data sharing, though one caveat discussed below is the fact that proper planning needs to be done upfront to avoid any technical or translational issues in data sharing. Though, participants indicated that once the planning is done, data transfer is an easy function to implement. A few specific findings that lead to these overall observations include the following:

- IT staff were asked if they were involved with any direct record-level transfers of data. Few reported that they actually engaged in that activity, saying that the various areas of the institution have access to the data that they need to perform their functions and that permissions are granted to staff based on their institutional role. As such, institutional staff are generally free to access, download and use data to which they have access. In fact, some IT staff had never been involved in record-level student transfers at their institutions. IT staff directly indicated that transferring record-level data would not be a technical issue for them and echoed the broader view of all staff that most issues regarding data sharing occur during the planning stages of data sharing, an issue discussed in more detail below.

- There is a generally good understanding among all institutional staff interviewed about privacy legislation and transfer of personal data. Both IT staff and legal and privacy staff were asked how they managed privacy issues given the fact that virtually all institutions allow access to student level data among staff, and there were common and consistent answers provided among participants. Access levels granted to staff were cited as a key way of managing privacy. Of equal, if not greater importance, is regular privacy training seminars and mechanisms that reinforce the issue institutionally. Participants indicate that these sessions are well-attended and engaged. Finally, within the institutional setting, there is sense that privacy is a paramount issue, and that institutional staff quickly develop a “feel” for when their actions will violate privacy laws. Staff indicate that when a privacy breach does occur, they rely upon institutional procedures to correct the breach, and all are aware that they must inform the institution with whom they share data that a breach has occurred. Without providing specific information, participants indicated that breaches they are aware of tend to be relatively minor, without malicious intent associated with them.

- Perhaps one area where legal and IT staff can be more involved is the secure transfer of data between institutions. This could involve privacy measures such as encryption, password protected files, secure transfer of data (FTP or otherwise) and secure storage of the data at the receiving institution. These participants indicate that with the liberal access staff have to data and the ability to transfer it, it is up to those front-line staff to implement proper transfer security protocols. Some IT staff recognized that secure transfer of data may be beyond the scope and knowledge, and even awareness, of many institutional staff. Unsecure transfer of student data is considered a privacy breach, and was cited by a few participants as having to be addressed.

- Registrar or academic staff that are involved with collaborative programs or redirection generally understand that they can share data with the other institution for the purpose of managing the program. These individuals did not indicate any technical issues with data sharing, other than for a few instances of timing, where a student record may be needed at an institution but has not been transferred because the agreed date has not occurred. They also tend to understand the legal environment and constructs under which data get shared. There were some participants who were aware of the fact that privacy sections of the institutional website or application give notice that data will be shared. Other indicated that they were not aware of this, and a few indicated that there needed to be better communication about this issue. Most participants that were affiliated with institutional registrars indicated that they will expressly have students sign forms acknowledging the data transfer between institutions. Some recognize that signing a paper may not be required if notice is on the website, but most indicate that the prefer to obtain explicit permission to share data to ensure full informed consent and transparency. Participants were asked if they encountered any issues with students signing releases to transfer data, and only two minor issues were mentioned. There were some who indicated that some students were anxious about records, specifically grades, being transferred because of poor performance, or a concern that they would be removed from the program at the receiving institution. Registrars indicated that this issue is not frequent and that they are capable of managing the issue. Also some students, if they do not read the form, feel that the consent form is a formal "reapplication" to the collaborative program at the receiving institution.

- One interesting comment was mentioned a few times regarding data transfer and separate OUAC and OCAS application processes. Collaborative programs, where the final outcome is a degree from a university require an OCAS application to be filled-out because the student will start in a college program. Some academic and program staff indicate this can be confusing to some students, and in a worse case, it can turn the student off of the program in general, especially if they need to pay a fee to fill-out a separate OCAS application. Moreover, some redirect programs from universities to colleges will transfer student data so that a college application need not be filled out, thus saving students the expense and trouble of a separate application.

Participants who Transfer Data Indicate that Agreements are in Place

Participants who transfer data between institutions indicate that they do have data sharing agreements in place that govern all the relevant aspects of transferring identifiable student record-level data. Many indicated that they work with the IT department and the privacy/legal department to draft the data sharing agreements, and they find that the agreements tend to work well to facilitate data transfer, such that generally no, or few issues occur. Some specifics about data sharing agreements:

- Some institutions can use previous data sharing agreements that have been in place as rough templates.
- Data sharing is a primary consideration when institutions talk about collaborative programs or redirections. However, data sharing itself is not the issue, rather it is all the other issues surrounding it that tend to be critical, such as timing, purpose, translation of the data from one institution to another, equity in sharing and overall use of the data.
- As mentioned previously, participants are aware of privacy regulations that govern data sharing and have a good place from which to start to draft agreements.
- One issue that was brought-up is a more strategic one in regard to planning. Sometimes data sharing agreements, especially for research purposes, limit the scope and/or time that an institution can use the shared data for. In some cases, the data may be useful for other purposes beyond what was stipulated in the data sharing agreement, so planning needs to balance limiting use and allowing for a broader use of the data should it be necessary. To this point though, participants indicated that they always abide by the terms of the data sharing agreement.

For Collaborative Programs or Co-Registration, The Larger Issue with Data Sharing Seems to Involve Planning

While IT staff and those involved with data transfer indicate that the actual transfer is fairly straightforward (to the point where many staff transfer data on their own, without IT assistance), if there is an issue to be addressed it is the planning around data transfer. This is examined by function:

- For research and/or planning purposes, planning how to match the data is key and fundamental. Moreover, much time is spent after the data is shared and amalgamated sifting through it to determine if there are invalid records or data, and how the overall research plan, whether the change involves one of timing, objectives or scope of the research. It is recognized, however, that this is often a reality of research, and is often not something that can be accounted for until the actual data is merged and exchanged.
- For data exchanges involving collaborative or co-registration programs between institutions, there was a sense that proper planning was the most important aspect to data sharing. While sometimes data sharing can be easy for some collaborative programs if the collaborative program data share simply "completes the registration", for other collaborative programs and co-registration program this can be more difficult. This largely focused on planning issues outside the direct issue of data sharing itself, but involved, for example, such issues as timing at both institutions (e.g. intake, exam timing, timing of released grades, residency rules, scholarships, financial data, school schedules), grading scales, definition of a graduate, course descriptions/timing and how often the data should be shared in order to achieve the smooth

running of the program. There were no examples of collaborative programs that gave "open access" to another institution to their data, for where collaborative programs occurred, and this is unlikely to occur because it is too complex to implement on both sides. Moreover, there is a question about who legally owns the data when it is transferred and open data access tends to make this issue complex as well.

The openness or willingness to exchange data for these programs was often related to attitudes of the institutions towards competition or willingness to be open and share data. Some institutions have very good and open relationships with each other and recognize that their programs compliment rather than compete with each other in this regard. Registrars that have open attitudes around this issue not only exchange this data more freely, they exchange it more strategically. That is, they will look at the data exchange in order to plan strategic issues for the program and the institution as a whole. For example, they will try to look at where students are going or coming from, estimate demand for specific programs and use it to enhance the student experience. There is a more holistic view of data exchange than just viewing it transactionally in nature. One suggestion to improve this is to involve other areas of the institution in the data exchange process. That is, if both institutions agree that there can be a broader purpose or benefit to the data being exchanged, other areas of the institution may be included to better guide both the data exchange and how the data can be used more broadly in both institutions.

Finally, there were some issues cited with OEN administration but those were seen as "growing pains" in that there were some issues that needed to be worked-out with the system. A particular issue, for example, regarded accuracy and use of the OEN for students transferring in to the province or the Ontario education system outside of the public-school system. There is also a recognition that the OEN is not universal and that can create interpretation issues.

The Most Common Area of Data Sharing Occurs at the Registrar Level, Followed by the Research/Planning Area

When interviewed, registrars seemed to have a very open and experienced approach to institutional data sharing, largely because of collaborative programs, or requests made to and from their offices for various data to be shared. They tend to be involved directly in the sharing of data for collaborative programs and play a strong role in the data planning process. As described above, they have a good knowledge of the legal and privacy framework and often implement explicit agreements with students to share data when necessary. Two examples that are more broadly based beyond sharing data strictly for collaborative programs include:

- Broad data sharing between two institutions that share the same student information system whereby a student tracking mechanism was created between the institutions, where student data is transferred so that lookups at the other institution can occur to understand student mobility and to transfer registrar data. Students are informed about the data sharing agreement through various means (e.g. websites, applications).

- A registrar transfers data with other institutions for the purposes of planning for a few years out. That is, they look at which students at other institutions are in collaborative programs, or programs where they are likely to transfer into the institution so that they can plan staffing and other requirements before the students transfer in to the institution. When asked to describe how the data is shared, it was indicated that issues such as timing of the transfer and the nature of the transfer are managed beforehand, and in specific interpreting the transfer data, limits on use and how data is “translated” from one institution to another are all discussed beforehand and put in a data sharing agreement. These issues are not seen as roadblocks, but rather as issues to be managed prior to the data sharing occurring, and are grounded in ensuring that students receive the best experience possible.

Research and planning areas of institutions are also very involved in data exchange. While the data sharing in which they engage is not nearly as frequent as those encountered by the registrar, they have a significant understanding of data sharing and often a very broad institutional view that gives them particularly good insight into the benefits to and processes involved in data sharing and matching. Often research and planning areas of the institution focus on very strategic levels of operations, or areas of interest regarding access issues or particular programs that may be good fits with other institutions. These data exchanges are often “project” based, whereby a project purpose is defined and data is shared to support the research initiative. The project can be internal to the institutions, or can be more broadly based, such as conducting a project for a stakeholder organization such as ONCAT or HEQCO. Often complex matching among datasets has to occur for these projects.

Transfer Data is Generally Not Shared for Articulation Purposes.

Articulation agreements are becoming commonplace among institutions whereby credit at a receiving institution is given for courses taken at a sending institution. The process of creating articulation agreements largely involves an examination and comparison of course content between institutions that involves a judgement around course equivalency and whether the sending institution provides the pre-requisite instruction needed to succeed at the receiving institution. In some cases, where institutions and/or programs are small, instructors or academic staff may know student outcomes, or where remediation is necessary because of direct in-class exposure to students, but generally academic performance of students or other tracking of student movements is not done to create articulation agreements.

This is not to say, however, that student movement is not important in measuring outcomes under articulation agreements. Participants say outcomes for articulation agreements are quite important to consider. However, they tend not to measure outcomes using transfer data from the sending institution. Rather they tend to rely on application data, or maybe OCAS or OUAC data that can indicate a student’s sending institution. From this information, they are able to generate a number of summary reports and outcomes that provide measurement of performance of students under an articulation agreement. Moreover, it is unlikely that these reports, which are generated by the receiving institution, would be sent to the sending institution without a direct request. Sending reports or record-level data can cause concern not only around privacy but around interpretation of the results and spark issues of competition among institutions as well. Some specific concerns include:

- If record-level data was sent, would the institution analyzing it know how to properly analyze grades or other information that is known to the generating institution.

- How the data would be interpreted or used, and concern over the implications drawn from it.

While other areas of institutions that create data sharing agreements address these issues, they do not appear to be addressed in regard to articulation agreement analysis. Rather, summary reports with aggregate data are generated within the institution itself, and seem to provide all the relevant information needed to administer the articulation agreement and process. Some summary reports sent between institutions have confidentiality statements on them to address these issues.

Many felt that this kind of system of summary reports works well for measuring outcomes. However, some felt that better metrics could be used both when developing articulation agreements and measuring them. There was a phrase that described the state of articulation agreements as "articulation by anecdote" as opposed to looking at hard data. One reason for this state is because it is difficult to track where students go after they leave the institution so it is difficult to produce analysis to support creation of articulation agreements. One way some institutions have of tracking student movement through the system is by examining student requests for transcripts to be sent to another institution. That is, if an institution sees that many students are requesting that their transcripts be sent to a particular institution, it may be a signal that it may be worthwhile to investigate with that institution who is actually attending, what programs they are taking and creating a more formal tracking system and or articulation agreements.

While the process of tracking student requests for transcripts being sent to other institutions is one way of tracking student movement, the preferred method of tracking student movement is through a more open exchange of data, whether through formal agreements created between institutions, or through use of the OEN, which is recognized as having the most potential for accuracy. When asked about whether tracking student movement like this through the system would cause more competition (i.e. rather than creating articulation agreements or encouraging student movement, the data could be used to keep existing students at the institution itself), some participants indicated that the programs that they would need to create at their own institutions are simply not within their mandate, or are too challenging to create on their own. That is, there was a view among some participants that more open exchanges of data would allow institutions to play to their individual strengths, thus producing better outcomes both for students and institutions, as opposed to being a negative competitive factor.

There Are Ways to Proliferate and Improve Data Sharing in the Future

A summary of the present situation regarding data sharing among the studied academic institutions is that a moderate amount of data sharing occurs. There are two ways to describe the volume or extent of transfer between institutions. Some engage in transfer on an ad-hoc basis in a relatively siloed approach in that there is no open agreement, or even an agreement of a larger scope of data transfer among institutions. However, there are some institutions that are in relative close proximity to each other and share students back and forth, and the tend to have more extensive registrar-level data exchanges to facilitate registration and exchange of student information to make movement for students easier, and to collaboratively manage programs and even conduct analysis of data and movement between institutions. There does tend to be a certain level of competition between institutions that is present in the equally open nature of data exchange and the collaborative programs, articulation agreements and other uses that underlies the data exchange. On the whole though data sharing tends to be a series of discreet exercises, and depending on the nature of the exchange, it can be time-consuming to match data (in the case of research) or plan the underlying programs (in the case of collaborative programs or co-registration).

When asked how to improve data sharing or encourage it, some participants mentioned:

- Create a more open, or EDI (Electronic Data Interface) system between institutions. Some mentioned the OEN in particular for this, but others indicated that transcript transfers, OUAC and OCAS have networks that could possibly be use to exchange student level data and make it more open and available. This could reduce the amount of effort involved in engaging in a series of one-off data exchanges. Moreover, more open exchanges of data reduce the sense of inequity that some may feel exists because some institutions collect more data, and more accurate data, than other institutions. If there is a more open and equal exchange of data, cooperation may increase, though this needs to be carefully considered so that there is a perception of fair value provided through the exchange. For example, a few IT individuals spoke of receiving access to OCAS application data that showed which programs a student applied for (and were accepted to) at colleges and which ones they actually chose. IT individuals referred to these as "dashboards" that provided individuals within the colleges with very valuable data on their overall performance compared to other institutions.

  More open and free-flowing exchanges of data can come from three sources. First institutions can agree to individual arrangements. Second, OUAC and OCAS can facilitate this kind of transfer, according to some participants. Third, the Ministry of Advanced Education and Skills Development (MAESD) can facilitate it. In the last two cases, the dashboard reports previously sited as coming from OCAS are cited as very valuable sources of information. There was also some mention of MASED providing aggregate level data using the OEN to track student movement and/or entry into PSE in Ontario on an institution-level basis. The conclusion here is that the value of data sharing must be proven to institutions and those that hold the data. Some respondents indicated that sharing data for research into issues such as access to under-represented populations would be very worthwhile. There is also a sense that data sharing to predict demand for PSE programs in the future will be a valuable use of data sharing.

- When data transfers and exchanges occur between institutions there are roles for all areas of the institution to play in the process and it should be a best practice to involve different areas of the institution when data is shared. Specifically, as a specific data transfer exercise occurs, various roles can see if a broader exchange can occur systemically or institution-wide:

  - IT could examine if a more open exchange of data (like EDI) would be useful on a broader level and if there are other exchanges that may benefit from a more open solution than just the data exchange being planned.

  - Academic areas could see if there are program implications for other areas of the institution that could benefit.

  - Those involved in research and/or planning can see if there is a broader issue at play, such as social issues like access and student mobility. There may also be strategic issues involved. For example, if there is a regular flow of students from one institution to another institution, except for a few programs is it because the sending institution is successful at finding positive labour market outcomes for those students, or is it that the receiving institution does not have the right programs for those students? These areas may also be able to assess where the level of effort should lie in creating collaborative programs between institutions. For example, some transfers between similar programs (e.g. business to business, nursing to nursing) may be worthwhile to put a lot of effort into, there may be other, less directly connected programs that have exchanges of students between them that may go unnoticed unless there is a more fluid exchange of data. It is important to note that researchers and planners likely access the same data that is exchanged for registrar purposes and for managing collaborative or co-registration programs, so having them involved in data exchanges that occur for these purposes may be quite valuable.

  - Registrars, and those who work directly with transfer data within registrar offices likely have good knowledge of the data that is exchanged between institutions and can point-out various pitfalls or issues to address in any exchange.

  - Administrative and management staff can look at the broader partnership opportunities available to the institution.

- Among participants, institutions that are close geographically, or have a direct mandate of facilitating transfers tend to have a more open approach to exchanging data because they understand that many of their students enter and leave the institution for other programs. This belief tends to create a more open and holistic outlook towards data exchange throughout the institution. Whereas all institutions in the study seem proficient at one-off or singular exchanges of data for specific purposes, there are some institutions that by mandate think more broadly.

- There is a sense that if truly open data exchange is to occur, with the use of the OEN driving it, and to its fullest potential, that funding and/or performance metrics should be tied to data that needs to be exchanged between institutions, given that the OEN facilitate the most accurate measurement of how students succeed and where they end up transferring within the system. This could encourage articulation agreements to systematically look more closely at transfer data (whether using the OEN or other data sources until the OEN proliferates). One comment that was made is that data that may be used to measure the efficacy of articulation agreements may not be complete or entirely accurate because matching it between institutions is a challenge, and if funding or performance were tied to the success of the agreements, this may encourage more open data exchange, to increase the accuracy of the data. It may also encourage the use of the OEN, as the OEN is, in principle, the easiest and most accurate way to track student movement through the system.

4 BEST PRACTICES SURROUNDING DATA SHARING

The section below is a result of researching dozens of documents that were returned under various Google searches on terms such as "Data sharing best practices", "Data sharing agreement checklists" and "Data sharing MOUs". The results returned documents from a wide variety of sectors and activities/purposes including education, medicine, taxation, research and computer science. The researcher reviewed the documents and created with broad categories or topics that are considered best practices, and then conducted further research as to what actions constitute best practices in those areas.

This section also indicates the five areas of transfer – General registrar, program administration, student redirection, articulation agreements and research/planning – to which each best practice is applicable.

Best Practice #1 – Only Personal Information under FIPPA Needs Protection

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

As stated previously, the best practices below only apply, in a legal sense, to data that contains personal information. However, as will be shown, one of the best practices is to implement data sharing agreements for de-identified data and having all users and viewers of data sign confidentiality agreements.

Best Practice #2 - Be Clear on the Benefits to Students, Institutions and/or the Public about Data Sharing

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

Though the IPC in Ontario feels that privacy of data is very important, there is a recognition that both the law and other practical considerations allow for data to be shared, when he says "however, the right to privacy is not absolute. In certain circumstances, the right to privacy must be weighed against various public interests."[40] The document further states "organizations should prepare a detailed business case outlining why there is a need for data sharing. The business case should:

- Identify the goals or objectives of the data sharing activity and the anticipated benefits;
- Identify the potential risks or consequences of not conducting the data sharing activity;
- Clarify why personal information must be shared at this time;
- Clarify why the personal information needs to include personal identifiers;
- State the purpose(s) for which the personal information was originally collected and;
- Identify why the personal information must be collected indirectly and the advantages of sharing the data against alternative methods of achieving the same objectives

[40] http://www.ontla.on.ca/library/repository/mon/3000/10301262.pdf

In fact, most data sharing documents reviewed for this research often state upfront and very clearly that data sharing is necessary for a specified public purpose. For example, often the first few clauses of data sharing agreements directly indicate the benefit of doing so for an identified group. Note that this benefit is not a re-stating of the purpose of the research or exchange itself, but rather, an end goal that should occur because of data sharing. Some examples from MOU's are:

- *To facilitate the health of [Aboriginal and First Nations citizens*[41]*] X and Y are entering into an agreement which will allow the exchange of data and clarification of data access and utilization.*[42]

- *Improvements in information sharing, translate into many tangible benefits. Repeat diagnostic tests can be avoided. Medical errors are reduced and outcomes improved with quicker access to complete information. Time is saved by physicians, staff and patients. With less manual processing of information and fewer phone calls for results, patients can be cared for quicker. Ultimately patients will be more engaged in their care by leveraging the technology where providers and patients can securely communicate via patient portals.*[43]

- *Successful development, testing, evaluation, and deployment of these innovative [disease] management systems require expertise in measurement science and in the development of standards and partnerships with the community. The [Project] intends to take advantage of the significant capabilities that exists in these areas within the [Research Center] specifically in [Department], and the experience and knowledge of those who deliver health services in the community context, such as neighborhood health centers. Therefore, wherever possible and when mutually beneficial, the [Project] and [Health Center] seek to collaborate on research, planning, and clinical activities, and share where appropriate facilities, personnel, and scientific information to meet the recruitment, retention, and evaluation goals of the [Project]*[44]

- The Student Transitions Project (STP) in BC states:

> *A highly educated workforce is critical to British Columbia's efforts to retain its competitive position in today's global knowledge-based economy. The benefits from this Agreement range from maximizing successful completion of academic and job training programs to increased local recruitment and retention of qualified workers and investment in British Columbia through the Canada/Asia gateway… The information is used to provide educational programming, conduct research and program evaluation/improvement, track students as they progress through the K-12 and post-secondary systems, plan programs, structure institutions and allocate resources. MED, AVED, K-12 schools and the public post-secondary institutions require information about sub-groups of students because different strategies are needed to address the distinct needs*

[41] Quote modified to reflect Canadian language

[42] www.npaihb.org/images/epicenter_docs/NW-Idea/Sample%20DSA.docx

[43] https://info.clinicalconnect.ca/CC/wp-content/.../CC-Data-Sharing-Agreement.pdf

[44] https://www.yumpu.com/en/document/view/.../clinical-mou-template-2-accelerate

> *of these sub-groups. Combining K-12 data with public and private post-secondary institution data is necessary to permit evaluative and predictive research that is crucial in understanding, improving and planning for K-12/post-secondary student transitions.*[45]

A research report that documented a data sharing framework in community-academic partnerships, summarizes the balance between privacy and benefit by stating "Data, in all its forms, newly created or re-used, should be maximized for use to improve… outcomes. Without a strong partnership with good communications, clear direction for a process, and well-developed content as part of a formal agreement, there are risks to the effective use, re-use and generation of meaningful information that is of value to all partners"[46]

Finally, the IPC of Ontario notes that "Data sharing between organizations may lead to individuals' loss of control over their personal information. Therefore, where possible, sharing should not occur without exploring less privacy- intrusive means of meeting a specific objective. Before deciding to share personal data, organizations should consider all practical alternatives which are more privacy protective, and all relevant information".[47] Researchers and policy planners in academic settings should explore other options prior to sharing personal information for research purposes. While it may not seem like there are a lot of options that could avoid sharing of personal information, the last section of this report describes current and future developments in data exchange within Ontario. It appears as if Ontario is starting to move towards the US and BC models of more open student data exchange. Both those jurisdictions have incredibly large student data warehouses or clearinghouses which allow researchers access to anonymized data. For example, in Ontario OCAS now offers data analytics services to those who wish to do research into post secondary students in Ontario[48].

Best Practice #3 – Only Transfer Personal Information that is Relevant to the Purpose of the Exchange

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

One of the best practices involved in planning data transfer is to ensure that only personal information that is only necessary to the purpose at hand is released[49]. So, for example, if the subject of study for which data sharing is to occur is on transfer of students in healthcare programs between college and university, researchers would need to give special attention as to whether aboriginal status, special needs or postal code information – all of which are personally identifying information under law – is necessary for the research purposes. Similarly, if data is being exchanged for the purposes of a redirect program, consideration needs to be given to which data exchanged.

[45] https://www2.gov.bc.ca/assets/gov/education/post-secondary-education/data-research/stp/stp_isa.pdf

[46] "Data Sharing: Creating Agreements In support of Community-Academic Partnerships", Paige Backlund Jarquin, P4

[47] http://www.ontla.on.ca/library/repository/mon/3000/10301262.pdf

[48] https://www.ocas.ca/what-we-do/business-intelligence

[49] http://www.ontla.on.ca/library/repository/mon/3000/10301262.pdf

Best Practice #4 – Consider Standard Definitions of Personal Information for All Institutions in Ontario

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

While FIPPA is clear in laying-out what constitutes personal information, post secondary institutions may wish to establish directly what they feel constitutes personal information. It is important to realize that the law defines personal information to include any ID number assigned to a student, including even the OEN. Within its Information Sharing Agreements, the BC STP defines personal information in the following manner:

> *The steering committee will establish from time to time a schedule of personal information attributes which shall be posted on the Student Transitions Project secure website. MED will disclose personal information that it has in its possession to the Data Custodian. The personal information attributes established by the steering committee as of the date of this Agreement include the following:*
>
> *(a) personal education number*
> *(b) date of birth*
> *(c) gender*
> *(d) school district/school name and number*
> *(e) highest grade completed or attempted*
> *(f) date of highest grade completed or attempted*
> *(g) course information (i.e., name, grade, date, session)*
> *(h) course performance measures (i.e., school, exam and final percent; pass/fail)*
> *(i) date of graduation from K-12*
> *(j) identification of school from which the individual graduated*
> *(k) graduation credential name*
> *(l) honours flag*
> *(m) aboriginal status*
> *(n) special needs category*
> *(o) home language*
> *(p) postal code*
> *(q) passport to education data*[50]

Best Practice #5 - Create a Translation Plan for Exchanged Data

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

For registrar purposes and program administration, when data is transferred it often needs to be

[50] https://www2.gov.bc.ca/assets/gov/education/post-secondary-education/data-research/stp/stp_isa_-_nov_2016_update.pdf

translated. That is information or variables like grades, course timing, definitions of “graduation”, course completion codes and other institution-specific administrative data within the data to be translated from one institution to another. In this way, the common data shared between institutions can be used and compared for registrar, programmatic or research purposes.

The notion of data translation was described in the qualitative research as one of the most challenging parts of data transfer between institutions. The challenge occurs not on a technical level, but rather on a planning or even negotiation level between institutions. For registrars, programs and operational data exchanges there is a notion that translation of data must be perfect or all-encompassing because this directly impacts a student’s actual record when the data is transferred. From a research point of view, translation of data is also exceptionally important to ensure that data is correctly analyzed.

Those within registrars, programming and operational areas indicate that they have to account for data translation before the data exchange occurs, whereas those in research may account for it both before the exchange occurs and afterwards, once they start doing their first analysis of the data, since many times they realize that they cannot account for all the nuances that may be involved in translation until they start working with the data.

One potential area of collaboration between all areas of institutions in all data exchanges it to involve different areas of the institution in creating a translation plan. That is, it is possible that some in research may have addressed translation issues during previous data exchanges, and similarly, those in registrar, program and operational areas may have also addressed translation issues not accounted for by researchers when planning data exchanges. Such collaboration presents an opportunity to save time by having different areas of the institution contribute previous knowledge and experience the translation function that is necessary for most data exchanges.

Best Practice #6 – Involve Relevant Areas of the Institution

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

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 engage with all relevant partners all the time for every data exchange, and there were very few instances in the interviews where participants indicated that this would be a requirement for all data exchanges occurring. For example, IT participants indicated that individual users have the ability to exchange data with other institutions without their involvement. Similarly, legal and privacy departments also indicated that it was not necessary to consult with them all the time a data sharing agreement is created. Those involved in ethics indicated that consultation with them is mandatory prior to research taking place, but that individuals within institutions are aware of receiving research ethics approval before engaging in such assignments. Some indicated that various data sharing agreement templates or suggestions for contracts exist within the institution to serve as guides to assist with planning. However, as with other respondents, use of previous contracts or templates is not required and is not seen as something that should be mandatory. Rather those documents are meant to serve as guides.

Perhaps a more fruitful area of collaboration within institutions could occur between programming, administration, operations, research and planning. That is a data exchange may occur to facilitate movement and registration of students for a collaborative or co-op registration program, but researchers and planners may be able to use the transfer data to answer some questions that they have, or may be able to provide additional input into program development based on an analysis of the transfer data. It would be important to determine if data transferred for operational use can be directly used for research purposes, but this collaboration could be a valuable source of data for all areas of the institution.

Also, as mentioned in the previous best practice, the most fruitful area of collaboration may be between areas of the institution that have had to translate transfer data previously to facilitate data transfer.

Best Practice #7 - Consider De-Identifying Data & Creating A Data Sharing Agreement That Governs De-Identified Data

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
				✓

As stated previously, de-identified, or non-personal data, is not subject to FIPPA laws and requirements. The IPC for Ontario's website states:

> *As the demand for government-held data increases, institutions require effective processes and techniques for removing personal information. An important tool in this regard is de-identification. "De-identification" is the general term for the process of removing personal information from a record or data set. De-identification protects the privacy of individuals because once de-identified, a data set is considered to no longer contain personal information. If a data set does not contain personal information, its use or disclosure cannot violate the privacy of individuals. Accordingly, the privacy protection provisions of the Freedom of Information and Protection of Privacy Act… would not apply to de-identified information.*[51]

Even more specific to the subject of post-secondary institutions sharing data for research and planning purposes, the IPC of Ontario recognizes de-identification of data as a best practice. Specifically, a guide on the IPC website states that de-identification of data is considered a best practice when sharing information between institutions for the purposes of planning and research as follows:

> *There is also a growing desire in government services for institutions to break down their "silos" and share more information within—and among—themselves. This may happen for a number of reasons. For example… information from one institution or program area may be relevant to the planning of a program or service in another institution or area…*
>
> *Data sets that contain personal information may be shared within and among institutions only if the disclosure is permitted under section 42(1) of FIPPA … If the disclosure is not permitted and the institutions still wish to share data sets, then (similar*

[51] https://www.ipc.on.ca/privacy/de-identification-centre/

> *to an access to information request or open data release) any personal information must be removed.*
>
> *However, even if disclosure is permitted under FIPPA or MFIPPA, there may still be important privacy issues to consider. While information sharing among institutions can play an important role in providing better, more efficient services, the practice may also have the unintended consequence of undermining the privacy of individuals by diminishing the amount of control individuals have over their personal information.* ***Therefore, as a best practice****, institutions should always consider de-identifying data sets before sharing them*[52]*.*

The issue, however, with transfer data between post-secondary institutions in Ontario is that to share and link data between institutions, it is highly likely that some personal information, even if it is just student numbers or the new OEN, must initially be exchanged to link the data together. The best practice in this regard is to link the data together, de-identify it and then only work with the de-identified data. However, this does not circumvent the fact that for a short while identified data is passed between institutions, thus requiring a data sharing agreement, even if the intent is to work with a de-identified set once the matching occurs. Moreover – and of critical importance – de-identification includes accounting for the possibility that an individual may be identified through data that is unique to the individual even if it is depersonalized. For example, a person could be identified if their area of study, previous country of study and grade upon exit is passed in the data set, and if the program area identified in the data only contains a small number of students.

The IPC addresses this in two ways. First it provides a comprehensive guide to de-identification of data, and second, it says that data sharing agreements should be in place even when working with de-identified data. In the case of post-secondary institutions sharing data, a data sharing agreement would be required for two reasons – 1) the initial sharing of data will likely require the sharing of personal information to create linkages; and 2) it is best practice to create a data sharing agreement even when working with de-identified data.

Addressing the de-identification process, the IPC provides the following guidelines for de-identifying data. They are provided in summary form, as the full topic of de-identification is beyond the scope of this research, other than to identify it as a best practice in data sharing and highlighting some of the factors that go into the process.

De-Identification Step	Specific Actions to Take
Determine the Data Set Release Model	This refers not to a final report release, but releasing the combined data set that provides primary data for the report. Publicly released data sets must be de-identified, but where there is only to be a private release, or sharing of data, which is likely the case with post secondary research, the guide states that there must be a data-sharing agreement in place.[53]
Classify Variables	For de-identification purposes variables are either direct identifiers (e.g. name), quasi-identifiers (student ID number or a combination of personal data) or non-identifiers

[52] https://www.ipc.on.ca/wp-content/uploads/2016/08/Deidentification-Guidelines-for-Structured-Data.pdf, P6
[53] https://www.ipc.on.ca/wp-content/uploads/2016/08/Deidentification-Guidelines-for-Structured-Data.pdf, P8

Determine Acceptable Re-Identification Risk	Researchers must evaluate how likely it is an individual can be re-identified by the data they are exchanging
Measure Data Risk	A researcher must look at the data set itself to see how many records can possibly identify an individual, even if data in the set groups people together to avoid identification of an individual when reporting.
Measure Context Risk	Evaluate the risk of re-identification based on the release model. Even if there is a private release model, researchers need to consider the technical resources, motivations and knowledge of those who may have access to personal identified information prior to it being de-identified
De-Identify Data	By masking identifiers, "Generalizing" identifiable variables (e.g. using age ranges instead of actual ages) eliminating (with notations) individual records or data that could be used to identify an individual even without identifiers
Assess Data Utility	The researcher should evaluate the utility of the data set once de-identification is complete
Document Process	Full documentation of the process should be made

[54]

Addressing the subject of working with data that has been de-identified, the IPC recommends creating a data sharing agreement that governs researchers in the following ways:

> 1) Prohibit[s] the use of de-identified information, either alone or with other information, to identify an individual;
>
> 2) Place[s] restrictions on any other use or subsequent disclosure of the de-identified information;
>
> 3) Ensure[s] that those who have access to the de-identified information are properly trained and understand their obligations in respect of such information;
>
> 4) Require[s] the recipient to notify the organization of any breach of the agreement, and
>
> 5) Set[s] out the consequences of such a breach.[55]

Another guide from the IPC says data sharing agreements for de-identified data should also cover:

- Protecting against attribute disclosure. Attribute disclosure occurs when a group, as opposed to an individual, is identified negatively from a data set or a resulting report. For example, identifying children of parents with a religious affiliation and their vaccination histories, or lack thereof, could result in stigmatization. Specifically, "while the privacy provisions in FIPPA relate to the personal information of individuals only and do not include measures to address potential harms affecting groups of individuals a best practice is to de-identify data that could be stigmatizing before releasing the data set. An ethics review of the data set may be needed to

[54] https://www.ipc.on.ca/wp-content/uploads/2016/08/Deidentification-Guidelines-for-Structured-Data.pdf, P7-20

[55] https://www.ipc.on.ca/wp-content/uploads/Resources/pbd-de-identifcation_essential.pdf, P3-4

achieve this."[56]

- Ongoing and regular re-identification risk assessments
- Auditing data recipients to ensure that they are complying with the conditions of the data sharing agreement
- Examining the disclosures of overlapping data sets to ensure that the re-identification risk is not increasing with new data releases, or that potential collusion among data recipients does not increase the re-identification risk
- Maintaining transparency around the de-identification practices of the institution
- Assigning responsibility and accountability for de-identification
- Maintaining oversight of changes in relevant regulations and legislation as well as court cases
- Developing a response process in case there has been a re-identification attack
- Ensuring that individuals performing de-identification have adequate and up-to-date training[57]

Another best practice in de-identifying data comes from the CIHR (Canadian Institutes of Health Research), where it is possible to remove identifiers from a dataset, but have a key that links the de-identified data to identified data in case there is a need to look at identifiable variables again. Th CIHR recommends that data can be:

- coded to allow a trace-back to individuals, by means of:
  - single-coding (the researcher has the key to the code to link the research data back to direct identifiers, which are held separately); or
  - double-coding (an increased level of confidentiality protection over single coding because the data holder does not give the researcher the key to re-identify individuals); or
- created without a code, if the capacity to trace the research data or results back to individuals is not required for the research purpose.[58]

Best Practice #8 – Linking Data

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

[56] https://www.ipc.on.ca/wp-content/uploads/2016/08/Deidentification-Guidelines-for-Structured-Data.pdf P21

[57] https://www.ipc.on.ca/wp-content/uploads/2016/08/Deidentification-Guidelines-for-Structured-Data.pdf PP20-21

[58] http://www.cihr-irsc.gc.ca/e/29072.html

The basic principle of any transfer data is that some sort of personally identifying information from one institution must be shared with another to create linkages between two data sets, or so that student information from the sending institution can properly be assigned to the same student at the receiving institution.

The discussion below outlines procedures for research, but equal consideration should be given to linking data for registrars, programs and articulation agreements. That is, who is going to do the linkage and how. However, there are more considerations for linkage for research purposes and the CIHR provides some guidelines as to best practices in data linkage. Specifically:

> *The most secure way of conducting data linkages requested by external researchers is for the data holder to conduct the linkage and provide linked datasets to the researcher without identifiers, and at the minimum level of identifiability required for the research purpose.*[77] *If that is not practicable, a trusted third party may conduct the linkage or the researcher may conduct the linkage on the data holder's site. As a last option, a researcher may be permitted to conduct the linkage at a secure site but under strict controls, as specified in a data-sharing agreement.*[78]

Who should conduct the linkage	***Comments***
A) Data holder (Preferred)	*The data holder performs the linkage(s) and subsequently removes all direct identifiers, or replaces direct identifiers with a code, prior to releasing the linked data set to the external researcher.*
B) A trusted third party (e.g. a statistical agency) or *C) The researcher conducts the linkage on the data holder's site*	*When the original data holder does not have the technical capacity or resources to perform linkages in-house:* • *a trusted third party acting as an information manager may conduct the linkage off site; or* • *the researcher as a "deemed employee" (e.g. the Statistics Canada model) may conduct the linkage on the data holder's site.* *The third party and the researchers should be bound by equivalent conditions of confidentiality and security as apply to the data holder and the data holder's employees.*
D) The researcher conducts the linkage off site	*If Options A, B or C are demonstrably impracticable, the researcher may conduct the linkage in compliance with a data-sharing/confidentiality agreement with the data holder, setting out their respective and shared obligations, including restrictions on use and disclosure and appropriate security requirements (see 8.3 below). In this situation, any direct identifiers or other personal data*

	not required to answer the research question should be destroyed or returned to the original data holder as soon as is practicable, and in compliance with the terms of the data-sharing agreement.

> *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.*
>
> *For example, direct identifiers (e.g. name or personal health number) or potentially identifying elements when combined (e.g. a full date of birth or full postal code) may be needed for data linkage but may not be needed to answer the research questions. In such cases, these identifiers should be destroyed as soon as is reasonably practicable or returned to the data holder, as per the terms of the data-sharing agreement.*
>
> *Universities may have specified retention periods for research data. Researchers should either destroy the new linked dataset at the end of the specified period, or use enhanced security measures to store it as per the terms of the data-sharing agreement. Within some research or statistical agencies, it may not be practicable to unlink datasets after each use. However, these institutions should document a process to ensure that the linked datasets are used only for authorized purposes (e.g. for REB-approved projects).*[59]

For programming, registrar and articulation purposes, effort should be taken to ensure that accurate data linking can occur as well. Some institutions have a common student identifier number and others can use personal student information such as name, address and birthday to link records.

Best Practice #9 – Determine Access Levels & Identify Users of the Data

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

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. Careful consideration must be given to this facet in academic settings and situations, particularly if someone with knowledge of students may be examining personal information, even if it is stripped of direct identifiers. For example, a faculty wishing to examine articulation data may be able to identify a student if the area of study is small, and if the student's age and data of graduation are listed in the data. For registrars, program administration, redirection and articulation agreements, it is likely that institutions will already have established access levels for each employee based on their function, but a best practice would be to identify which functions at the receiving institution have access to the transferred data.

[59] http://www.cihr-irsc.gc.ca/e/29072.html

According to the BC STP data sharing agreements, researchers may wish to establish various levels of access for different individuals that could include:

- **Data Steward:** The Data Steward has read/write access to identifiable data. Only a limited number of individuals requiring access to the data to fulfill their duties and responsibilities will have access to the data. It is important to state the direct reasons why this individual has access to such data, and encourage signing of a confidentiality agreement.

- **Authorized User Level 1:** Users in this group have ongoing read-only access to all identifiable data and can do so to carry out data quality assurance checks. It is important to state the direct reasons why this individual has access to such data, and encourage signing of a confidentiality agreement. It is important to state the direct reasons why this individual has access to such data, and encourage signing of a confidentiality agreement.

- **Authorized User Level 2:** Users in this group have ongoing, read-only access to anonymized data. Authorized Users Level 2 may also access identifiable data for their own institution, for conducting research to support program evaluation and accountability. This group also includes employees in institutional research/registrar/data and reporting departments of the organizations. Contractors working on behalf of an institution or consortium may be granted access for a time-limited period. It is important that these individuals sign a confidentiality agreement

- **Academic Researcher**: Academic researchers are members of academic institutions that have ethics committees or boards that approve research projects. Academic researchers may be granted read-only access to data through research agreements which allows disclosure for research or statistical purposes provided that the researcher has obtained an approved ethics certificate and that the proposed research is consistent with the original purpose of the collected data. Academic researchers must submit a written proposal that articulates the research questions that will be answered using the data.

It should be noted that while the guidelines above only suggest confidentiality agreements for the first three levels of users, best practice would suggest that confidentiality agreements be signed by anyone who views or uses the data. Finally, the Canadian Institutes for Health Research (CIHR) recommends dismissal or punitive actions be spelled-out for individuals who breach confidentiality of data.

For programmatic, operational, administrative and articulation purposes consideration should be given to who within the institution has access to the transfer data and if those individuals, either by name or function should be identified.

Best Practice #10 – Implement the Data Transfer and Arrange Schedules

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

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 can be handled directly by the staff involved.

Best practices of data transfer largely involve securing and encrypting it, and then ensuring that it is stored on a properly secured device at the receiving institution. These parameters do not have a standard definition either from the literature, or from individuals interviewed. The only standard cited is that emailing a non-password protected file is not considered secure and is also considered a privacy breach. Many individuals recognize that USB transfer and/or encryption and/or password protection is necessary. Some institutions discussed the use of secure methods of transfer (e.g. FTP). Some did indeed report using FTP and in fact require using FTP to transfer files. However, it was recognized that FTP use may be difficult to implement because it is not too familiar to users and would require special training.

The other best practice involved in transferring data is to determine how often updates, if any, will occur. Related to this is the fact that there is a possibility of open data exchange, or EDI, where data is transferred automatically from one institution to another. In the qualitative interviews participants indicated that they did not EDI, though some were aware of EDI, especially if they had been involved in other industry sectors prior to working in education. While EDI may represent a best practice in that it allows for instantaneous updates from one institution to another, it is difficult to implement and would likely not be implemented for data exchanges that do not involve a large amount of data that will be exchanged over time on a regular basis between institutions.

Best Practice #11 - Storage & Verification of Accuracy of Data

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

Best practices in data sharing require consideration of storage and verification of accuracy of the data. It is important to note that many sample data sharing agreements do not list the actual specifics of the storage methodology, there are many institutional factors that will directly impact data safety and storage, some best practices could include:

- controlling access to rooms and buildings where data, computers or media are held
- logging the removal of, and access to, media or hardcopy material in store rooms
- not storing confidential data such as those containing personal information on servers or computers connected to an external network, particularly servers that host internet services
- firewall protection and security-related upgrades and patches to operating systems to avoid viruses and malicious code
- Security of computer systems and files may include:
    - locking computer systems with a password and
    - installing a firewall system
    - protecting servers by power surge protection systems through line-interactive uninterruptible power supply
    - implementing password protection of, and controlled access to, data files, e.g. no access, read only, read and write or administrator-only permission

  - controlling access to restricted materials with encryption
  - imposing non-disclosure agreements for managers or users of confidential data
  - destroying data in a consistent manner when needed[60]

The CIHR indicates some of the following safeguards:

- Encryption, scrambling of data and other methods of reducing the identifiability of data should be used to eliminate unique profiles of potentially identifying information.
- Direct identifiers should be removed or destroyed at the earliest possible opportunity.
- If direct identifiers must be retained, they should be isolated on a separate dedicated server/network without external access.
- Authentication measures (such as computer password protection, unique log-on identification, etc.) should be implemented to ensure only authorized personnel can access data.
- Special protection for remote electronic access to data should be installed.
- Virus-checking programs and disaster recovery safeguards such as regular back-ups should be implemented.
- Where possible, a detailed audit trail monitoring system should be instituted to document the person, time, and nature of data access, with flags for aberrant use and "abort" algorithms to end questionable or inappropriate access.
- Computers and files that hold personal information should be housed in secure settings in rooms protected by such methods as combination lock doors or smart card door entry, with paper files stored in locked storage cabinets.
- The number of locations in which personal information is stored should be minimized.
- Architectural space should be designed to preclude public access to areas where sensitive data are held.
- Routine surveillance should be conducted.
- Physical security measures should be in place to protect data from hazards such as floods or fire.[61]

Another factor to consider that is in keeping with data storage is verification of accuracy of the data. Best practices recommend that institutions periodically review that the data they are working with contains the same and accurate information as was in the original file. This is an especially important function when data is de-identified, say by creating a variable set that groups ages of individuals together, rather than reporting individual ages to avoid identification of any one individual. This step

[60] Van den Eynden, Veerle, et al "Managing and Sharing Data", Best Practices Guide For UK Researchers, UK Data Archive P29

[61] http://www.cihr-irsc.gc.ca/e/29072.html

would require that one or both institutions maintain an original copy of the data and frequency counts of all records and variables in the file to confirm continued accuracy.

Best Practice #12 - Reporting Breaches & Audit Trails

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

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. There are many best practices that can be implemented in this regard:

- Authorized users will record and monitor access to the data, to establish a chain of responsibility
- All attempts to access servers, directories and files are to be logged. This provides an audit trail to assess unauthorized access attempts
- Privacy audits from either party can occur
- In the event of a breach, or possible breach such as unauthorized access, use modification or disposal a breach protocol should be implemented. [62]

If a breach, as outlined, has occurred, the following protocol should be implemented:

- **IDENTIFY:** Identify the scope of the alleged breach and take initial steps to contain the damage (this may involve determining whether the privacy breach would allow unauthorized access to an electronic information system).
- **REPORT:** Ensure that appropriate staff is immediately notified of the breach. The report should indicate whose personal information was disclosed, to whom, when it was disclosed, how it was disclosed/accessed, and what steps have been taken in response to the disclosure.
- **RETRIEVE:** Any documents that have been disclosed to, or taken by, an unauthorized recipient should immediately be retrieved or destroyed (especially for fax or electronic mail)
- **INFORM:** In cases where the breach may result in consequences that would directly affect the person whose information has been disclosed, that person should be informed of the details of the breach. They should also be informed of the Party's efforts to retrieve this information and prevent a similar breach from reoccurring.
- **INVESTIGATE:** For determining and recording all the relevant facts concerning the breach and making recommendations. The objectives of this investigation should include: a review of the

[62] Adapted from https://www2.gov.bc.ca/assets/gov/education/post-secondary-education/data-research/stp/stp_isa_-_nov_2016_update.pdf

circumstances surrounding the event as well as the adequacy of existing policies and procedures in protecting personal information.[63]

Best Practice #13 – Determine Reporting & Access Conventions

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
				✓

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. Finally, for planning purposes, it is important to consider whether faculty or staff with direct knowledge of students (e.g. instructors who teach classes may be the ones analyzing data; instructors familiar with published work of students at other institutions) at the sending and/or receiving institution will have access to the data and how confidentiality should be maintained given that some staff may be able to look at a data set and identify students in it even when the data is de-identified.

Researchers should address levels of release for the data and the report. There is a movement towards releasing the raw data that backs-up a research study or an academic plan to readers of those documents. There are three kinds of releases in general – private, semi-private or public. In this case:

- Private means just among the identified individuals in the MOU;
- Semi-private means to those in the circle of influence of the researchers and;
- Public means to anyone not in the private or semi-private spheres.

A public release of data must be de-identified, and it is best practice for private and semi-private to be de-identified. Note that the release models for the data and reporting can be different from each other. The 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 supressing 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.

A final reporting release convention is to determine whether interested or identified parties should be allowed to comment of the data. Since data shared between institutions necessarily involves at least two organizations, it is important to consider providing a provision that allows the other institution to comment on the release of the research report, especially if the other institution is identified or if comments about the other institution are made. An agreement should include provisions about how consensus about the released report should be obtained.

[63] Adapted from https://www2.gov.bc.ca/assets/gov/education/post-secondary-education/data-research/stp/stp_isa_-_nov_2016_update.pdf

Best Practice #14 – Determine Additional Usage Options

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

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 of time. Some participants in the qualitative research for 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.

5. Suggested MOU Elements

This section provides guidance on elements that should be included in MOU's or data sharing agreements between institutions. It is important that readers customize each of these elements and seek institutional approval as necessary, especially legal approval, prior implementing such agreements. Also, the best practices outlined in the previous section should be consulted in drafting agreements.

There were three references used of this section. The first is Regulation 460 of FIPPA which sets out minimal requirements and is considered Ontario law for what should be included in an agreement. The second takes the 1995 "Model Data Sharing Agreement" from Tom Wright, Information and Privacy Commissioner of Ontario at the time and uses it as a general guideline. The third is an analysis of dozens of data sharing agreements across a wide variety of sectors and incorporating best practices from those agreements into these best practices.

As with the previous section, this sections also provides guidance on the relevance of each suggestion based on functions within the institute.

<u>Element #1 – Compliance with Regulation 460 of FIPPA</u>

As discussed earlier, FIPPA provides some regulations regarding the terms and conditions relating to security and confidentiality that must be agreed to before disclosure of personal information from an institution can occur. Tom Wright's document that drafts a sample MOU for data sharing indicates "If personal information is shared for research purposes, the organization should consider the terms and conditions relating to security and confidentiality, as outlined in section 10 of Regulation 460/section 10 of Regulation 823 under the Act."[64]

The following table provides the relevant sections of Regulation 460 and some comments on how to comply and/or include them in data sharing agreements:

MOU Element		Area of Institute				
Regulation 460 Requirement	**Comments**	**General Registrar**	**Program Administration**	**Student Redirection**	**Articulation Agreements**	**Research Planning**
The person shall use the information only for a research purpose set out in the agreement or for which the person has written authorization from the institution.	An MOU must list the exact purposes of the research	✓	✓	✓		✓
The person shall name in the agreement any other persons who will be given access to personal information in a form in which the individual to whom it relates can be identified	The previous section discusses the fact that besides a data sharing MOU, personal confidentiality agreements should be signed by those who have access to personal data. It should also be noted that best practice is to assign	✓	✓	✓	✓	✓

[64] http://www.ontla.on.ca/library/repository/mon/3000/10301262.pdf

MOU Element		Area of Institute				
Regulation 460 Requirement	**Comments**	**General Registrar**	**Program Administration**	**Student Redirection**	**Articulation Agreements**	**Research Planning**
-and- Before disclosing personal information to other persons… the person shall enter into an agreement with those persons to ensure that they will not disclose it to any other person.	different roles, responsibilities and levels of access to different individuals who access the data, including who the data steward is and how communications between individuals will occur.					
The person shall keep the information in a physically secure location to which access is given only to the person and to the persons given access to the data	Institutional guidelines and other best practices will govern this action, but MOU's can and should list the precautions they have in place to keep data secure.	✓	✓	✓	✓	✓
The person shall destroy all individual identifiers in the information by the date specified in the agreement	Agreements should list the date when personal identifiers will be destroyed, how destruction will occur and that notice will be given.					✓
The person shall not contact any individual to whom personal information relates, directly or indirectly, without the prior written authority of the institution	This section can be put into both the MOU itself and individual confidentiality agreements as necessary.	✓	✓	✓	✓	✓
The person shall ensure that no personal information will be used or disclosed in a form in which the individual to whom it relates can be identified without the written authority of the institution	Best practice is to identify rules and regulations regarding cell sizes and levels of detail in reporting that could possibly identify an individual and supress those. If the research may publish small cell sizes in a way that could identify an individual, consent from the other institution must be obtained.	✓	✓	✓	✓	✓
The person shall notify the institution in writing immediately if the person becomes aware that any of the conditions set out in this section have been breached	The previous section discussed audit trails and some best practices regarding investigations of data breaches. Some of those include the fact that if a data breach has occurred that the individuals affected must be notified in person.	✓	✓	✓	✓	✓

Element #2 - Indicate the Legislative Authority to Collect and Disclose Personal Data

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

It is important that agreements recognize the legislative authority that gives institutions permission to collect and use personal data for the purposes of assisting in articulation research and the transfer process. As discussed in Section One, under the acts that regulate the institutions themselves, they are given very broad authority to collect and use records for the purposes of administering the institution itself. Best Practice would be to develop an MOU that outlines the exact areas in the legislation and/or the institutional privacy policies that allow for data collection and use for these purposes.

Element #3 – State the Purpose and Business Case for Data Sharing

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

FIPPA considers privacy of personal information to be paramount, but that there are times where privacy rights of individuals can be relaxed or compromised for activities that would be considered for the greater good of society or individuals. All transfers of data fall within this definition, but as the Privacy Commissioner states:

> *organizations should prepare a detailed business case outlining why there is a need for data sharing. The business case should:*
>
> - *Identify the goals or objectives of the data sharing activity and the anticipated benefits.*
> - *Identify the potential risks or consequences of not conducting the data sharing activity.*
> - *Clarify why personal information must be shared at this time.*
> - *Clarify why the personal information needs to include personal identifiers.*
> - *State the purpose(s) for which the personal information was originally collected.*
> - *Identify why the personal information must be collected indirectly and the advantages of sharing the data against alternative methods of achieving the same objectives.*[65]

[65] http://www.ontla.on.ca/library/repository/mon/3000/10301262.pdf

An example from the Canadian Institutes for Health Research illustrates some of the above principles:

An Example of CIHR Study Objectives & Justification for Use/Collection/Transfer of Personal Information

Study objectives: To examine and compare the health status, health care, and social involvement of distinct ethnic groups living in [region X of province Y], to inform policy development by community organizations and governments.

Research questions: (examples) What is the association between health status, experience of health care and ethnicity? What are the impacts of personal support networks and activity level on health status and perceived well-being?

Personal data needed and justification:
Initials: To assist in checking for duplicate records, using a combination of initials and demographic data.

Demographics (date of birth, gender, ethnicity...): Needed to make between-group comparisons on health variables by ethnicity, and between- and within-group comparisons by other demographics.

Physical health and sense of well-being/Use of health services: Needed to investigate health status and perceived health status by health care-related knowledge, behaviours, attitudes and use.

Meaning of health and of aging: Needed to explore the meanings of health and illness and the cultural context of aging in the ethnic community.

Family and friends/Social activities: Needed to investigate the impact of family structure and interaction and environmental factors on measures of health and well-being.

Element #4 - Indicate the Personal Information to be Shared

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

According to the Privacy Commissioner, the disclosing party will provide to the collecting party a list all elements of personal information that will be disclosed and collected for the research purposes previously identified. The Commissioner notes that personal information that is only necessary to the purpose of the transfer should be released. As such, transfers need to plan the information to be disclosed based on their anticipated research needs. Specifically, the parties should:

- *Identify whether personal information is about one individual or a group of individuals;*
- *Estimate the number of records to be shared and the current storage format or medium and;*
- *Identify how the personal information will be disclosed and the frequency of data sharing. The exact nature of the personal information to be shared must be identified in detail. The parties should use the definition of personal information in section 2(1) of the Act as the basis for*

its description.[66]

Section 2(1) of FIPPA appears below:

- information relating to the race, national or ethnic origin, colour, religion, age, sex, sexual orientation or marital or family status of the individual;
- information relating to the education or the medical, psychiatric, psychological, criminal or employment history of the individual or information relating to financial transactions in which the individual has been involved;
- any identifying number, symbol or other particular assigned to the individual;
- the address, telephone number, fingerprints or blood type of the individual;
- the personal opinions or views of the individual except where they relate to another individual;
- the views or opinions of another individual about the individual; and
- the individual's name where it appears with other personal information relating to the individual or where the disclosure of the name would reveal other personal information about the individual; ("renseignements personnels").

Element #5 – Indicate How the Personal Information Will be Used

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

This best practice requires users to list how personal information will be used during the time they are using it. A Data Sharing Agreement should indicate how the proposed use of the data complies with FIPPA – specifically that research purposes are allowed under the act. The example of the CIHR agreement above shows how personal information is justified in its agreement. Moreover, there should be an explicit statement about how the data will not be used for any other purpose other than those set out in the Data Sharing Agreement.

Element #6 – Indicate If There Will be Future Disclosure of The Data, and How That Will Comply with FIPPA

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

The data sharing agreement should indicate if the resultant dataset from the data sharing or transfer will be released again, and if so under what conditions. Best practice would be to limit the data created to the purposes of the transfer only, but as the qualitative research indicated for this study, other purposes for the data may occur. It is important for those who have access to the identified data set understand what the purposes of the data sharing are and if any additional use or disclosure can occur.

[66] http://www.ontla.on.ca/library/repository/mon/3000/10301262.pdf

It should be noted for research purposes that use and transmission of de-identified data is not protected under law, but could be specified in an agreement.

Element #7 – Indicate If the Data Will Be De-Identified

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
				✓

If the data is to be de-identified, the MOU should 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
- If written notice will be provided upon successful de-identification of the data

Element #8 – Identify the Method of Sharing Data

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

An agreement should indicate how the data is to be transferred in terms of method, security precautions and the frequency of data exchange if it is to occur more than once. The Privacy Commissioner also notes that institutions who sign an MOU should:

- *Identify any technical problems involved with the data sharing and the strategy which has been developed to minimize these problems. (e.g., physical loss of data during transfer.)*[67]

Element #9 – Indicate If and How Data Linking Will Occur

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

Since one of the main reasons to include personal information in datasets is to perform linking, the data sharing agreement should indicate the steps that individuals will take to link the data, including factors such as individuals who will perform the linkage, timing, and whether the personal identifiers will be destroyed after linkage occurs.

[67] http://www.ontla.on.ca/library/repository/mon/3000/10301262.pdf

Element #10 – Accuracy and Security of The Personal Information

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

The previous section described some of the best practices involved in accuracy and security of the personal information identified under FIPPA. These policies and abilities will also be influenced by institutional guidelines and technological resources. Best practice is to:

- Describe what steps will be taken to verify the accuracy and completeness of the personal information before it is used;
- Identify the steps that will be taken to ensure that the personal information is up-to-date;
- Describe the measures that will be taken to ensure that the personal information will be protected against unauthorized access and that only authorized persons will have access to it;
- The organization should identify the measures that will be used to ensure that personal information shared through this Data Sharing Agreement is protected against loss and unauthorized access during transfer, as well as unauthorized access, use and disclosure after transfer;
- For any personal information stored on a computer:
    - identify the controls in place to ensure the security and completeness of transmission (encryption);
    - identify the controls in place to ensure that only the required personal information will be transferred; and
    - describe the types of audit trails and/or management reports produced to ensure that personal information will be processed in a complete and accurate manner.

Element #11 – Indicate Release Model for Report & Data

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
				✓

If the data is going to be released, where data could include identified raw data, de-identified raw data and/or any summary report that uses the data, the agreement should specify the release model and what can be said about the data and by whom. The agreement should also state if approval of the other party to the agreement is required prior to the release of any data.

MOU Best Practice #12 – Indicate Termination of The Data Sharing Agreement

General Registrar	Program Administration	Student Redirection	Articulation Agreements	Research/Planning
✓	✓	✓	✓	✓

The agreement should state an end-date for the sharing activities. Consideration should be given to whether data could be used for reasons beyond the purpose of the existing exchange and built that into the agreement.

6 Present & Future Developments

This section examines factors that influence the current and future state of sharing student data in Ontario for the purposes of analyzing transfers and helping in articulation.

The Role of The OEN In the Future of Data Sharing Among Institutions

The OEN is a unique student identifier that is assigned to all students who are enrolled in a public or post-secondary institution in Ontario. As a common identifier, it has significant potential to aid in student tracking between institutions, though at present participants in the qualitative research indicated that their use of the OEN at present is limited because it is not universal among all students at this point. However, perhaps the most significant factor in the use and development of the OEN is that it is considered a piece of personally identifiable information under section 2.1 of FIPPA. Even if the OEN is the only piece of data released in two datasets such that linking between them can occur, the OEN still requires the privacy protocols mandated by FIPPA. Moreover, a report by Kelly Gallagher-Mackay on "Data Infrastructure for Studying Equity of Access to Post Secondary Education in Ontario" states that legislation has been very protective and cautious in allowing institutional collection, use and release of the OEN. Specifically, she indicates:

> *Under the Ontario Education Act (governing K-12) and the Ministry of Training, Colleges, and Universities Act, which provide for Ontario Education Numbers to be assigned, there is a general requirement that the OEN be treated as private information: "No person shall collect, use, disclose or require the production of another person's Ontario Education Number."*[68]

Her report indicates that there has been some slight movement in the legislative requirements regarding the OEN. However, from her description of changes to the legislation, it is quite clear that there is only very cautious use of the OEN in Ontario for any sort of student tracking purposes. This was also confirmed in the qualitative research in that institutions indicate they are not using it yet for tracking student movement. The repot indicates:

> *There have been a number of exceptions, notably, for purposes relating to the provision of educational services and for applications for student financial assistance. For greater clarity, the Education Act was amended in 2010 to add a new exception in s. 266 (3): "The Minister and a prescribed person may collect, use or disclose or require the production of Ontario education numbers for purposes related to education administration, funding, planning or research." In 2014, the Ministry of Training, Colleges and Universities Act, which has parallel provisions to the Education Act, was also amended to provide the minister or a college, university or other postsecondary educational institution with the same exception to share data for research and planning.* [69]

[68] http://www.heqco.ca/SiteCollectionDocuments/FINAL%20Data%20Infrastructure.pdf P19
[69] http://www.heqco.ca/SiteCollectionDocuments/FINAL%20Data%20Infrastructure.pdf P19

However, the research and planning that is referred to above is not inter-institutional data sharing but rather data sharing with larger, more established databases, as the report mentions:

> *This change was a key development in permitting the data to be shared with, for example, Statistics Canada for PSIS, or the Ministry of Finance. But even with new legislative authority, there continues to be very limited data sharing and the data sharing that does occur is limited specifically to educational purposes rather than broader public policy purposes such as research projects involving student well-being, or measuring the effectiveness of anti-poverty initiatives.*

The development of the OEN as a common identifier linking information together along with the insights provided in the Gallagher-Mackay report provides some implications regarding data sharing among institutions. Perhaps the most important is that caution is likely still required when sharing information that has the OEN attached to it, even between institutions for research purposes. While the report does not site an example of inter-institutional data sharing, it does refer to OEN data being shared between the Ministry and U of T researchers, where she indicates:

> *On a few occasions, researchers working on the evaluation of government-supported projects have obtained access to anonymized versions of these data for their work (see, for example, Ford & Oreopoulos, 2016). The process has involved protracted negotiation on a variable- by-variable basis and highly customized data-sharing agreements. This treasure trove of data is largely unavailable to the broader research community and even to agencies with mandates to inform government.”*[70]

However, from a privacy, accuracy and research point of view sharing the OEN has the potential to improve privacy when conducting student transfer research. Consider, for example, the study done on student transfer data over a twelve-year period between York and Seneca and the methodology used to match data:

> *Developing the analytical sample. The analytical sample included students who have entered either institution between 2000 and 2012 but previously attended the other at any time from as early as the 1980s.* ***Because there was no common unique identifier****, 1.2 million valid Seneca records were compared against 407,000 valid York records. The match was made using combinations of surname, first initial, gender, date of birth, permanent and secondary telephone statistics.* [71]

On the one-hand, the use of the OEN would have eliminated the need to transfer such personal information as name and address and would have represented an increase in the privacy and efficiency of the research. However, the potential harm that could be caused by a security breach would be significant. If a malicious individual obtained OEN numbers for the above amounts of students and was able to re-identify and link personal information contained in the OEN record back to the student level data, there would be a very large privacy breach.

[70] http://www.heqco.ca/SiteCollectionDocuments/FINAL%20Data%20Infrastructure.pdf P19

[71] Smith, R., Decock, H., Lin, S., Sidhu, R., & McCloy, U. (2016). Transfer Pathways in Postsecondary Education: York University and Seneca College as a Case Study. Toronto: Higher Education Quality Council of Ontario, p18

The second implication is that the OEN paves the way for the development of large data warehouses that are available in jurisdictions like BC and the United States. Gallagher-Mackay references the fact that OEN-linked data is shared with Statistics Canada and even the Ministry of Finance, and this may be a stepping stone to linking and sharing data in a way that is useful for measuring student mobility and transfers.

It is important to note that HEQCO has made some significant comments on their hopes for use of ministry-gathered OEN data:

> *The OEN informs policy files at the centre of provincial priorities: mobility, equity of access, student success, and institutional differentiation. The data has been rigorously collected on a census basis; the sample is everyone and the risk of getting it wrong is extremely low. 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 signalling a willingness to share OEN data, appropriately protecting privacy, with the broader community, and is taking steps to do so. At HEQCO, we are looking forward to doing better research with OEN tagged data*[72]

A similar comment follows:

> *It's the Ontario Education Number. Please, let's use it. Let the research community look at participation rates, demographic trends, and the implications of policy and program changes. The more open we are with our data, the better a conversation we will be able to have about what works and what doesn't in supporting our students.*
>
> *We know there are privacy concerns. They are legitimate but they can be accommodated. The governments of British Columbia and Alberta have managed to find a way. These provinces both use a unique identifier that tracks learners' educational progress so that governments and policy makers can make decisions that serve students best. Those would be evidence-based decisions, the basis of all effective public policy.*
>
> *We all want more equitable access for Ontario's youth. But how on earth will we know how far away we are from reaching that goal and whether anything we are doing is affecting that course unless we know more about Ontario students.*
>
> *Right now we are just yelling policy and program ideas into a black hole of assumptions and good intentions. Let's actually shed some light on the process. It would really be remarkably easy.*[73]

The comment above indicates that "British Columbia… has managed to find a way" and a discussion of BC's STP project occurs below. However, briefly here, the STP has a Steering Committee structure that governs the project and guards overall privacy of the data very strictly. Moreover, the project creates

[72] http://blog-en.heqco.ca/2017/08/martin-hicks-data-done-right/
[73] http://blog-en.heqco.ca/2017/04/fiona-deller-and-martin-hicks-spoiler-alert-its-the-ontario-education-number/

"levels of users" that have different access to the data for the project. In one document, access to the actual PEN itself is described as "very limited", as follows –

> *The STP Data Access Policy specifies that only a handful of "data custodians" (e.g. information technology employees in the Ministry of Education who support the STP project) and "user level 1" researchers may see the PENs used by STP. A few more employees in other provincial government ministries and agencies may see encrypted PENs."*[74]

The implication is that projects that use provincial education numbers must have extremely limited exposure of the number and use of the number requires a very strict protocol to be in place for the project.

In a report entitled "Unlocking Student Potential Through Data" Donna Quan indicates that the ministry is willing to share data – and perhaps more importantly – she indicates that data sharing agreements are critical to that process. The very limited access and strict control over the PEN in BC for the STP shows the kind of structure that is needed to make a good data sharing agreement. While the scope of this repot is limited to creating a data sharing template between institutions, when approaching the Ministry to work with OEN data, it would be very important to indicate that institutions are willing to create data sharing agreements like the ones that Quan sites below and the one detailed in the description of the BC STP:

> *Data-sharing is a key piece in enhancing knowledge mobilization and bolstering accountability. Data-sharing initiatives hold the potential to produce analyses that can directly target areas of inequity as well as inform policies and actions geared to circumvent future recurrences. In addition,* ***solid data-sharing agreements can lead to producing information that will be useful in a variety of capacities and to various organizations.*** *While government organizations stand to gain from the culmination of cross-sectional analyses, extending the collection of data and instituting data-sharing agreements can also provide academics, school boards, advocacy groups, and others with useful analytics that can help inform community-, school-, and district-level improvements. Once data-sharing mechanisms have been established, the Ministry will have to develop ways that community organizations can also leverage critical information through providing relevant analyses.*[75]

Finally, some of the one-on-one interviews indicated that OEN data is starting to be analyzed at the provincial level (i.e. not on an institutional level). Some participants cited the fact that it is possible to see aggregate-level data of student movement between institutions and the proportion of Ontario high school students that went to college and university. Also, the PEDAL (Public Economic Data Analysis Lab) lab at McMaster has had experience analyzing record-level OEN, OCAS and OUAC data, and it is an indication that under the right conditions, such data analysis can occur. Some strategic questions that could assist with planning the use and analysis of OEN data outside of the Ministry itself could include:

[74] https://www2.gov.bc.ca/assets/gov/education/post-secondary-education/data-research/stp/datalinkagepolicy.pdf

[75] http://news.yorku.ca/files/Feasibility-Study-Unlocking-Student-Potential-through-Data-FINAL-REPORT-Feb-2017.pdf P83

- What are the acceptable levels of security requirements at the facility to which the data will be released?

- Who will do any data linkages?
- What are the minimally acceptable terms of data use for such factors as storage, use, purpose of research and release of the data?

It would appear that Ontario is at a stage where student level data is available for analysis, and that there are a number of valid uses for that data. The next step would be for MAESD to determine acceptable standards under which data can be used and accessed. The BC Student Transitions Program, discussed in below offers Ontario some guidance on how to structure this important step in more wide and secure use of the OEN for research purposes.

Regardless of the Implementation of the OEN, the Current State of Data Sharing in Ontario Is Cited as Needing Improvement

While institutions will be better able to match their records with the OEN, and while there is vast potential for the Ministry of Advanced Education and Skills Development to release OEN-tagged data to researchers, the present state of data sharing in Ontario for the purposes of research and articulation is somewhat lacking. This situation is likely to continue for the foreseeable future even if the potential of the OEN takes full effect. Specifically, since the OEN is new, many older records may not possess it so existing research methods and limitations may still apply.

In reading reports that either involved data sharing to analyze student mobility in Ontario, or in reading reports that described the state of data sharing, the following passages indicate some issues with the systemic quality of transfer data in Ontario. Specifically, from Ross Finnie's study on income levels of transfer students, he notes:

> *It is important to highlight data quality issues underlying these findings. Since the applicant type variable had difficulty identifying the application types of all the graduates in the data, we could not examine potential heterogeneities among non-direct entry graduates. Thus, while this project may have demonstrated fruitful approaches by which PSE-tax linked data can be used to examine how PSE pathways are related to both pre- and post-schooling outcomes, more thorough analysis requires higher-quality data on PSE pathways, ideally full PSIS-type data for an entire jurisdiction so that specific pathways can be identified by the researcher by tracking students as they move through the entire PSE system.”*[76]

Another report that focused on Indigenous Program Pathways noted the following about data available for the study:

> *Data was also a common challenge raised. Some institutions shared that they currently do not have the capacity to track pathway learners. Other institutions do have the capacity to track pathway learners by characteristics including discipline, gender, and geographic location. They explained that once the infrastructure is in place, tracking*

[76] How Student Pathways Affect Labour Market Outcomes: Evidence from Tax-Linked Administrative Data Executive Summary March 31, 2017

> *pathway learners is not an onerous process. When students come in through the admissions program they are flagged as a transfer student via a specific code, and students can be sorted by that code. In this instance, as indicated in one follow-up phone conversation, the challenge for institutions is understanding how to best use the data that is available."* [77]

Referring to the Gallagher-Mackay report, she comments on the current state of data infrastructure by indicating that one of the most important sources regarding student transfer data, YITS (Youth in Transition Study), is no longer a priority in the shadow of the OEN:

> *Despite the importance of YITS and the fact that most comparable countries have similar longitudinal cohort studies of youth transitions, none of my informants viewed renewing the survey as a priority of data infrastructure. Instead, most researchers and institutional stakeholders emphasized the untapped potential of administrative data and, in particular, the Ontario Education Number (OEN), which theoretically allows depersonalized linkage of longitudinal data about programs and progress for students throughout K-12 and PSE. Unfortunately, while data are collected using the OEN, there is limited linkage of the data between K-12 and PSE, between institutions, and between student outcomes and programs or resources. Indeed, the Ontario Student Information System, OnSIS, systematically strips data that would allow student-level analysis. Access to depersonalized, OEN-linked data is limited and highly discretionary. Different educational agencies with key information have quite different patterns of response to data requests.*[78]

Another important finding cited in a few key informant interviews is the discretionary nature of release of the OEN data and other record-level data. That is, there is some concern that data holders become very strong gatekeepers as to who is able to see record-level data, and who is not. In theory, according to some participants in the one-on-one interviews, this puts a significant concentration of discretionary authority into the hands of data owners, such that it may impede the perception of open and transparent access to large banks of student level data among those who request it. This speaks to the importance of considering having an arms-length committee that is comprised of the data holders themselves as well as other stakeholders (e.g. individuals from educational institutions and other organizations that play a role in Ontario higher education) that can make more open and transparent decisions about who is able to access record-level data from large information banks.

Development of Large Data Warehouses by Arms-Length Organizations and in Other Jurisdictions

The discussion of the OEN above has touched upon the notion that a common identifier (i.e. the OEN) would allow linkage of several separate databases together to create a large data warehouse of student information in Ontario, which among many other things could include data that could be analyzed to provide insight into student transfers and articulations. While the idea of such data warehouses is still in

[77] INDIGENOUS PROGRAM PATHWAYS INVENTORY PROJECT PHASE ONE Prepared By: Lana Ray, PhD, Minowewe Consulting, Research Lead p30
[78] http://www.heqco.ca/SiteCollectionDocuments/FINAL%20Data%20Infrastructure.pdf P3

very early stages in Ontario, there are several practical benefits and issues that can be addressed and discussed:

- Data combined from many sources gives much more information to analyze. Other such data warehouses include entire educational profiles from k-12, financial information, employment and outcome information and detailed minority status information;

- Privacy, access, reporting, data management, confidentiality and usage agreements are all governed by one source thus standardizing these important issues and reducing chances of breaches. If there is more centralized and secure control, there may be more liberal access to the data;

- Sometimes depersonalized data is made available publicly, or more quickly. Such data may provide immediate answers to some research questions without having to ask for personalized information and;

- Some services offer request forms right from websites making access somewhat easy.

The disadvantage of such warehouses is:

- Concern over data accuracy and completeness;

- If access is denied, what other data sources are available? That is, would there be too much concentration of data such that freedom of access may be limited;

- Their interpretation of privacy policies may be more stringent than other sources thus reducing the utility of data offered;

- They may not replace custom requests for data between institutions themselves and;

- They may charge fees for access.

Also within the qualitative research many participants indicated the fact that translation of data between institutions is a fairly significant logistical issue that needs to be managed when sharing data just between two institutions. This problem may be magnified for larger data sets. The other issue cited in the interviews is interpretation and use of the data. Institutions want to be sure that their reputations are protected and that accurate interpretations and statements are made about results that involve them and their programs

Within Ontario, there would be a few potential sources for such large data warehouses:

- Ministry data itself could be made accessible to researchers. As the Gallagher-Mackay paper indicated, and as indicated by the HEQCO opinion pieces, this route is extremely limited at present. However, should the relevant Ministries adopt a model used in BC, there could be more access granted. It must be noted that all Ontario post-secondary institutions are required to submit and report data to the relevant Ministries during regular intervals, so the ministry does have a significant amount of data. All post-secondary privacy policies disclose this reporting to the ministry. The introduction of an OEN will make it easier for the Ministry itself to link data.

- OUAC and OCAS, considered arms-length organizations also collect a significant amount of transfer data related to post-secondary institutions in Ontario. OCAS, for example is now offering data analytics services on the student application data that it gathers[79]. This was cited by some participants in the one-on-one interviews as a valuable source of business intelligence for colleges.

- Institutions could strike agreements and create custom where they could share and pool data regularly. For example, Durham College and UOIT struck such an agreement to share student data based on their common Banner number.

For an example of how data sharing is working between MAESD and OUAC/OCAC, Donna Quan in her report entitled "Unlocking Student Potential Through Data" indicates:

> *The Ministry has also recognized the importance of linking its data with data collected in the post-secondary sector to examine student outcomes after secondary school. The Ministry currently has a data sharing agreement in place with the Ontario Universities Application Centre (OUAC) and the Ontario College Application Service (OCAS) to acquire student-level application and registration data. Currently, ESAB facilitates matching OUAC and OCAS data with data collected through OnSIS to create combined data sets that are used for internal Ministry analysis and made available to school boards*[80]

Finally, the United States serves as an example of how a well-developed warehouse of data can work and service many of the research needs of institutions who wish to examine transfer patterns among their students and transfer patterns throughout the system. The National Student Clearinghouse claims that it has data representing 97%, or 19.8 million, of currently enrolled postsecondary students (98% of all public and private institutions, nearly 94% of all degrees awarded in the U.S. and over 250 million historical student records. The organization further makes the claim that it enables higher education to save over $750 million each year.[81]

The Data Clearinghouse has a service called "Student Tracker" which allows for "enrollment and degree information on your institution's current students, former students, and admission applicants."[82] It further provides "student unit level data that you can combine with your own data to analyze educational trends and patterns by any variable you choose… [and] unlimited individual student look-ups via the Web."[83] Data and findings that can be created include:

- Identify enrollment trends & patterns
- Track transfer student enrollment nationwide
- Improve your ability to target, recruit & retain students

[79] https://www.ocas.ca/what-we-do/business-intelligence

[80] http://news.yorku.ca/files/Feasibility-Study-Unlocking-Student-Potential-through-Data-FINAL-REPORT-Feb-2017.pdf

[81] http://studentclearinghouse.info/onestop/wp-content/uploads/NSCFactSheet.pdf

[82] http://www.studentclearinghouse.org/studenttracker/

[83] http://www.studentclearinghouse.org/colleges/studenttracker/

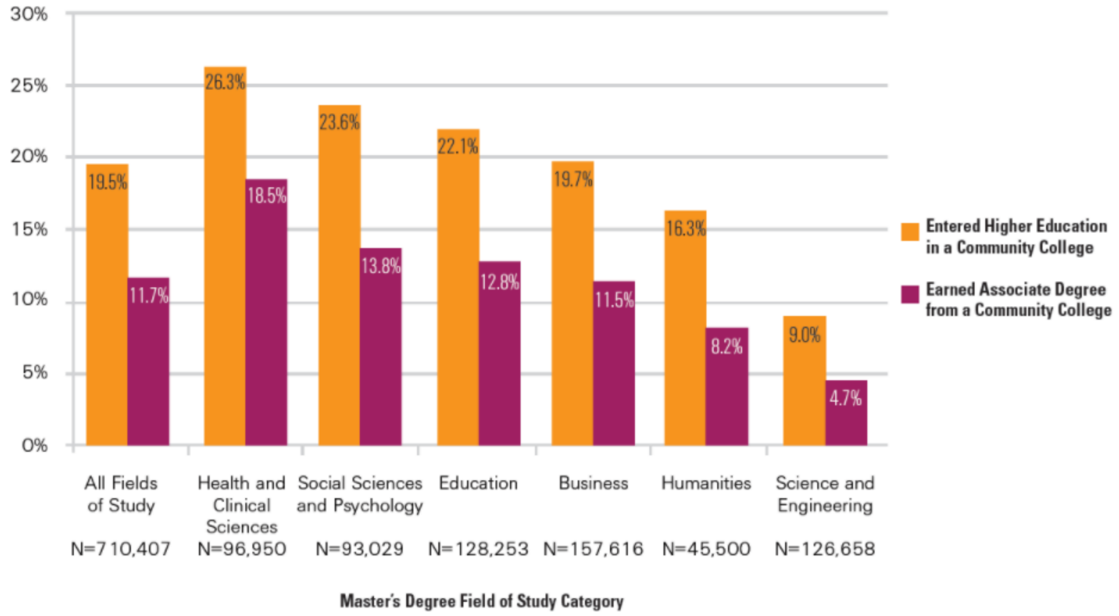
- Fulfill federal reporting requirements
- Verify & correct cohort default rates
- Determine a student's financial need level
- Plan curriculum modifications & institutional alliances
- Conduct academic assessments
- Identify students who are concurrently enrolled
- Perform accurate longitudinal & other outcome analyses[84]

On November 1, 2017 the organization released a report called “From Community College to Graduate Professional Degrees” with the preface that:

> *community colleges provide an important entry point on the pathway to graduate and professional degree completion. Nearly 20 percent of 2016-17 master’s degree earners originally entered higher education in a community college, and nearly 12 percent earned an associate degree from a community college.*[85]

And part of the report provides the following information

Figure 1. Percentage of 2016-17 Master’s Degree Earners Who Entered Higher Education in Community College (by Master’s Degree Field of Study Category)



Baccalaureate-granting institutions that predominantly award associate degrees (at least 55 percent of undergraduate degrees awarded are associate degrees) are classified as community colleges. A mapping of Classification of Instructional Program (CIP) family codes to broad field of study categories is included at end of this report. A student is considered to have entered higher education in a community college if their earliest postsecondary enrollment at age 18 or above occurred at a community college.

[84] http://www.studentclearinghouse.org/colleges/studenttracker/

[85] https://nscresearchcenter.org/snapshotreport-from-community-college-to-graduate-and-professional-degrees30/

A report entitled “The National Student Clearinghouse as an Integral Part of the National Post-Secondary Data Infrastructure” discusses some of the broad issues around participation in the NSC and how it manages data and agreements. Specifically, it mentions:

> *Participation by institutions is voluntary and each institution submitting data to NSC retains ownership of their own data. NSC acts as steward of the data and agent of the institutions in their use. NSC’s use of the data is allowed only in accordance with existing agreements that it maintains with each The National Student Clearinghouse as an Integral Part of the National Postsecondary Data Infrastructure submitting institution. These agreements specify the allowed uses and govern the ownership of the data and the terms of the agency relationship. They are designed to comply with FERPA (Family Educational Rights and Privacy Act) and to conform to data privacy best practices. A key implication of the agreements is that institution-level data, including outcomes or results derived from the data that identify the institution itself, can be made public only with the institution’s consent. If NSC were to do so without consent, it is safe to assume that some institutions would simply stop providing data. This is the current reality for policymakers or others wishing to make use of NSC data. Extracting or collecting the data from NSC, although operationally and technologically far easier, is legally no different from extracting or collecting it directly from the institutions.*[86]
>
> *Who can access the data? Outside of the relevant offices (e.g., institutional research, financial aid, etc.) located at participating institutions, access to student-level data is permitted only to researchers affiliated with an organization or institution. Their research purpose must comply with the allowable research exceptions listed under FERPA, and they must begin with their own data on a group of students that they wish to augment with the postsecondary enrollment and degree data NSC holds. This arrangement aligns with FERPA rules, it is within the scope of the contractual rights granted to NSC by the institutions, and it does not require institutions to grant permission to individual researchers. The StudentTracker® service is the primary mechanism for access to student level data. Through StudentTracker, researchers must submit a list of students, with individual identifiers, to query. NSC does not produce or provide such lists, nor does it verify or correct the individual identifiers submitted. The researcher must certify that the purpose of the request meets one of the allowable exceptions to the release of student-level educational records under FERPA. NSC then matches the submitted list to the appropriate student enrollment and outcomes data and returns the data to the requester… Access to NSC data is also provided through special requests by organizations for custom analytic reports, produced outside of StudentTracker. These reports show results at aggregate levels that prevent both students and institutions from being individually identified.*[87]

The Central Data Warehouse (CDW) and Student Transitions Project (STP) in British Columbia – An Example of Large Data Sharing with The Province to Research Student Transfer

[86] https://nscresearchcenter.org/wp-content/uploads/NSC-as-an-Integral-Part-of-the-National-Postsecondary-Data-Infrastructure.pdf p6

[87] https://nscresearchcenter.org/wp-content/uploads/NSC-as-an-Integral-Part-of-the-National-Postsecondary-Data-Infrastructure.pdf P7

The Environment in British Columbia

British Columbia has long been recognized as a leader in facilitating transfer of students between post-secondary institutions, and having the data infrastructure in place to support such transfers. BCCAT houses a document that synthesises all the research conducted on student transfers in BC. Its author, Bob Corwin, indicates ""British Columbia is distinctive in the North American post-secondary context because of its highly developed system for students not only to move among institutions but also to transfer credits."[88] He then goes on to describe the transfer-student data sharing architecture in the province by saying:

> *Over the past five to seven years [NB – the report was written in 2012], new student-by-student databases such as the Ministry of Advanced Education, Innovation and Technology's Central Data Warehouse (CDW) and the collaborative Student Transitions Project (STP) have allowed knowledge of transfer students to be viewed in the larger context of the flow of "mobile" students to and from the full set of public institutions, and increasingly over an extended period."*[89]

Use of the CDW

The CDW in BC is like data collected by Ontario post-secondary institutions which is then aggregated and reported to the provinces and the federal government. However, the BC government shows a willingness to share the data in the warehouse publicly as follows:

> *The Post-Secondary Central Data Warehouse contains standardized data relating to student demographics, programs, credentials, courses, session registration and campuses for 21 public post-secondary institutions in B.C… Data is updated in May and October… The submission process is centralized – the Ministry of Advanced Education provides data on behalf of contributing institutions to the federal government's Post-Secondary Student Information System. Data is submitted annually after the May data collection cycle…. **Privacy**: Individual students cannot be identified – student names are not included and any other student-level data is encrypted.*[90]

From this, the same website makes available the following reports:

- Headcount Totals (PDF)
- Aboriginal Identity (PDF)
- Gender (PDF)
- Program Area (PDF)
- Age (PDF)
- Credentials Awarded (PDF)

[88] http://www.bccat.ca/pubs/synthesisofresearch.pdf, p8

[89] Student Transfer, Success, and Mobility in BC Post-Secondary Institutions, Bob Corwin, P9

[90] https://www2.gov.bc.ca/gov/content/education-training/post-secondary-education/data-research/post-secondary-central-data-warehouse

It should be noted that the data here are very general and would likely not suffice for the research purposes needed by post secondary institutions, but it indicates a willingness to make data in the CDW available on an aggregate level to the public.

General Description of the STP Project and Relevance to Ontario

Perhaps the most relevant initiative in BC is the Student Transitions Project (STP). The STP:

> *uses personal education numbers (PENs) to track B.C. student data across both K-12 and public post-secondary education systems. This information guides program planning and management to help students transition successfully to post-secondary education and graduate…. Strict procedures ensure that privacy is protected – data used for the project cannot be used to make decisions about individual students.*[91]

More specifically, the STP is a partnership between BC's Ministry of Advanced Education, Ministry of Education, The University of British Columbia, Simon Fraser University, University of Victoria and University of Northern British Columbia to share personal information about students to track their movement through BC's entire education system. The STP provides an example of project governance that would likely be required in Ontario to have access to similar data from MAESD should it decide to utilize its data in a similar fashion. The two defining characteristics of the project are the fact that a Steering Committee manages the project and that there is significant documentation supporting the project. Specifically, there are at least five separate documents within the data sharing agreement system, with many containing sub-documents. Specifically, the STP has:

- A Steering Committee Terms of Reference Document;
- In Information Sharing Agreement which contains the ISA itself, and many other documents such as a "Breach of Privacy Protocol", "Confidentiality Agreements" for those who access the data, a "Research Agreement" that requests details of any research project for academic institutions who wish to study the data and a "Compliance Certificate" that affirms researchers have deleted STP data;
- A "Data Access Protocol" that defines different user levels of the data and their access rights to it (e.g. identifiable data, aggregated data, anonymized data, ability to manipulate/maintain data);
- A "Data Linkage Policy" that describes how other data can be linked to the STP dataset and;
- A "Reporting Protocol" that describes how STP data should be reported based on FOIPPA requirements.

These five elements are described below in more detail.

[91] https://www2.gov.bc.ca/gov/content/education-training/post-secondary-education/data-research/student-transitions-project

The STP Steering Committee

The Steering Committee has strict control over the data created by the project, and it is tasked with the following based on a detailed Terms of Reference:

- Determine policy-related research questions to guide analysis of data exchanged under the Student Transitions Project Information Sharing Agreement
- Determine how parties to the agreement can use the data and who can use it
- Assess third-party requests for access to the aggregate data that does not contain personal information
- Establish timelines, methods and procedures for the exchange of data
- Review data analysis reports
- Ensure that any exchange of personal information meets the requirements of the Freedom of Information and Protection of Privacy Act[92]

The STP Information Sharing Agreement

The following summarizes the Information Sharing Agreement signed by the parties involved to give a sense of how the data for this project is managed. It will give insight into the nature of a complex data sharing agreement, and the considerations a similar project would require in Ontario:

- The agreement states a concise purpose and benefit to the project – "A highly educated workforce is critical to British Columbia's efforts to retain its competitive position in today's global knowledge-based economy. The benefits from this Agreement range from maximizing successful completion of academic and job training programs to increased local recruitment and retention of qualified workers and investment in British Columbia through the Canada/Asia gateway."

- There are many roles, functions and definitions that are provided in the agreement:

    - Different types of data that will be used and created such as aggregate data (combined and grouped data without identifiers), anonymized data (data that contains encrypted personal identifiers), identifiable data (data that is "collected, including personal information with personal identifiers that have not been encrypted. This data exists for a time-limited period to match records.");

    - Data Custodian – "responsible for receiving and aggregating the personal information disclosed by the Parties for the purposes of the Student Transitions Project" and;

    - Data Protection Plan – "specifies how the user who is permitted access to the anonymized data will protect that data from unauthorized access, collection, use, disclosure or disposal and which specifies the physical security measures implemented to protect the storage media upon which the data resides".

[92] https://www2.gov.bc.ca/gov/content/education-training/post-secondary-education/data-research/student-transitions-project

- There is a listing of all the information that is personally identifiable that will be included in the data from institutions;

- Descriptions of how the collection and analysis of the data is consistent with the purposes for which it was collected under the Acts that regulate the post-secondary institutions contributing to the project;

- Ownership of the data and the subsequent reports among the contributing parties is stated, and in this case, they all share ownership rights;

- There are descriptions of who can access different levels of identified information and for what purpose. For example:

    - Only masked data will be available to the public, with masking defined by various BC FOIPPA definitions and;

    - Only authorized users will have access to unmasked data for specific purposes.

- There are specific instructions on how personal information can be used for the purposes of research and analysis and that only those purposes are allowed for use of personal data. Such purposes include linking data together, and conducting some specific analyses. The Steering Committee must approve all analyses involving personal data

- There are responsibilities provided to the Data Custodian regarding security and anonymity of the data. Such stipulations include that the data be stored in Canada, that the Ministry's data storage infrastructure shall be used to secure the data, that only certain individuals can access the identifiable data and the following specific clauses about security:

    - Any access to the identifiable or anonymized data will be under controlled circumstances, with full security measures that meet the highest government standards;

    - All authorized users agree to provide reasonable physical security measures for the data that is the subject of this Agreement in their custody or under their control, commensurate with the sensitivity of the information. Authorized users shall make persons with access to the data aware of their protection of privacy responsibilities under FOIPPA and;

    - All authorized users who will have access to anonymized or unmasked aggregate data must sign a confidentiality agreement attached as Appendix 2 and forming part of this Agreement.

- Creation of compliance, monitoring and auditing standards, including logging all attempts at data access, auditing of data access and the implementation of a "Breach of Privacy Protocol" should there be a breach in the data;

Should an external researcher wish to use the data, they must complete a "Researcher Agreement" that includes the following information:

- Identification of users of data;
- Purpose of the project;
- Identification of the specific data they wish to access, with a provision that general researchers can only access aggregate data but that academic researchers may access individual data if there is Ethics Board approval and;

- Terms and conditions of use of the data which include the following:
    - The fact that use must be consistent for the purpose with which the data is collected, which is administration of post-secondary functions;
    - That no attempt will be made to identify or contact individuals included in the dataset
    - Data will be stored in Canada;
    - Reporting cannot identify any individual;
    - A cope of the report containing STP data must be provided to the Steering Committee
    - Storage of data must meet specific "government security standards" and that audits of storage can occur and;
    - Deletion of the data must occur by a specific date.

STP Data Access Policy

The STP also creates definitions of data users and classifies their access to the data based on the user type. The classifications are contained in a document called the "STP Data Access Policy" that groups users in the following way:

- Data Steward & Authorized User Level 1 – only a few individuals who have complete access to all data to maintain it;

- Authorized Level 2 – Can access anonymized data across all contributors and identifiable information from their own institutions for the purposes of carrying out analysis;

- Academic Researchers – Can have restricted access to identifiable data, anonymized data and aggregated data with a REB certificate and Steering Committee approval of their project. Of note, the Steering Committee will not consider releasing data if a research project is considered a duplicate of a previous project an;

- The Public – can only see reported data that meets the STP reporting requirements.

STP Data Linkage Policy

Sometimes administrative data outside of the STP database could be linked to data inside of it to create a new dataset for analysis. This document describes the procedures for data linkage. In short data linkage is allowed, but only by unanimous approval of the STP Steering Committee. It is also indicated that the only way to link data from an external source to the STP dataset is through the provincial PEN number, and then the policy states that there is very strict control over the PEN. Other considerations about linking data to the STP dataset include:

- Whether the researcher is affiliated with a post-secondary institution and if the research has been approved by an Ethics Board;

- If the research has a clear public benefit and;
- How the linkage is to occur via PEN, whether it is to be done by the ministry, by STP staff, by a third-party contractor or by the researcher themselves and how the linking data is to be deleted.

Reporting Policy
The STP reporting policy indicates how STP data should be reported. Specifically:

- The policy sets different reporting requirements for project partners (the ministry and participating institutions) and third parties. Third parties are granted to the minimum amount of data needed to perform the research;
- For project partners, institutions can internally distribute information about their own institution freely. However, if other institutional partners are discussed in a public release, the institution's research director must be given an opportunity to comment;
- For third parties, both the STP Steering Committee and any identified institutions must be given the opportunity for comment;
- Data cannot be used to make decisions about individual students;
- Small cells must be supressed according to legislative guidelines/policies and;
- The report must be consistent with the purpose of the data collected.

The Groningen Declaration Illustrates Global Movement and Best Practices in Data Sharing, Especially in Relation to Data Repositories

While Ontario, like many other jurisdictions, continues to evolve and implement new and innovative initiatives regarding the sharing of student data, the Groningen Declaration discusses the issue on a more global scale and provides insight into specific movements in data sharing that have been discussed throughout this report. The declaration itself "seeks common ground in best serving the academic and professional mobility needs of citizens world wide"[93] and makes "digital student data portability happen. Citizens of the world should be able to consult and share their authentic educational data with whomever they want, whenever they want, wherever they are."[94] The mandate of the Declaration seeks not only to provide portability in student transcripts, but also addresses the function of sharing record-level student data among institutions to improve mobility of students and educational outcomes. To this end, the Declaration's website prominently posts a keynote speech given by Thomas C. Black, currently the registrar at Stanford University[95]. Excerpts and analysis of his keynote speech highlight some global trends, best practices and applicability to data sharing among academic institutions in Ontario follow:

- While Ontario itself is experiencing and managing increased student mobility, the trend is also occurring on a global scale whereby he states "the needs of our students keep changing, as they are more cosmopolitan and mobile… We must think of worldwide exchanges."[96] He grounds this in the fact that there is a culture that now "celebrates the capacities of the individual"[97] and that this is being aided with internet-based post-secondary education.

[93] http://www.groningendeclaration.org/
[94] http://www.groningendeclaration.org/
[95] https://news.stanford.edu/news/2007/july25/black-072507.html
[96] http://www.groningendeclaration.org/article/fifth-way-plea-inter-connected-central-student-data-depositories
[97] http://www.groningendeclaration.org/article/fifth-way-plea-inter-connected-central-student-data-depositories

- This particular study placed an emphasis on the transfer of record-level student data, as opposed to transcript data. The keynote speech provides an excellent notion of why such exchanges are so important for Ontario and student outcomes and mobility. Specifically, there is mention of the fact “we are NOT [sic] keeping the full transcript of the educational experience… students are involved in faculty supervised internships, research activities and community service. Most of which is not hitting the transcript… the record is a long way away from being the full record of what students have learned, or what they can do.”[98]

- The keynote speech spends a fair bit of time discussing the history and mechanics of electronic transcript and record transfer, focusing on PDF technology, EDI exchange which is a standard for exchanging data between two parties through to integrating the entire transfer of all student records combined together (i.e. the combined exchange of an application, a transcript and all other relevant electronic educational information in one format). For this last point, which Black calls integration between all student records (e.g. applications, transcripts, records) ,he states one potential benefit for student applications, where a student is transferring from one institution to another to enable the receiving institution, to which the applicant is applying to receive a student transcript from the sending institution before the applicant has even finished applying.[99]

However, and in concluding the paper, Black says that the future of student record transfer does not lie in the above factors (i.e. better use of PDF’s, EDI’s or integration of student records), because all of these innovations, while exceptionally helpful and beneficial, only facilitate data exchange between two institutions in a discrete, direct and closed manner. In order to support the spirit of the Groningen Declaration, Black refers to student data clearinghouses as a best practice solution. He states that data clearinghouses, specifically the National Student Clearinghouse in the US “can now help registrars solve a big problem… the emerging urgent challenge of… record portability… my colleagues realize that on their own, keeping up … will be expensive and time consuming.”[100] Black cites a number of benefits for a combined student clearinghouse of data. Specifically, he mentions:

- Nearly two-thirds of the students in the United States have more than one school record, and forty percent have more than two, making the need for a central source fairly important;

- A repository will help institutions know more about their students compared to other institutions; and

- As online courses go a central clearinghouse can help verify student identity.

In conclusion of Black’s paper, and to tie it into the Ontario situation as discussed in this paper, Black indicates that “adoption [of new innovation] occurs in stages… in stage 1… people acquire knowledge about innovations that are in accordance with their interests, needs and existing attitudes, and seldom expose themselves if they don’t perceive the need.” Based on the results of the one-on-one interviews and focus groups conducted for this paper:

- The institutions interviewed appear to be somewhat supportive in their attitudes towards seeing the need for full data exchange by way of a student clearinghouse. Institutions that are more open in nature are able to create fairly functional and open data sharing arrangements

[98] http://www.groningendeclaration.org/article/fifth-way-plea-inter-connected-central-student-data-depositories
[99] http://www.groningendeclaration.org/article/fifth-way-plea-inter-connected-central-student-data-depositories
[100] http://www.groningendeclaration.org/article/fifth-way-plea-inter-connected-central-student-data-depositories

with institutions and/or programs of interest. Others are able to obtain basic levels of student information just by looking at existing data without data exchange;

- Virtually all participants in the research, and as the literature review indicates, institutions see the OEN as a huge opportunity to create a more open system of data sharing for a multitude of purposes;
- However – and in relation to Black's point on adoption of change and technology – requisite interests, attitudes and needs must be in place before such an effort can take place in Ontario. Many individuals interviewed see the benefits to the entire system in sharing data in a very open nature. Others can see the benefit to their own institutions in sharing data in a similar fashion, whether or not they see the benefit on a more systemic level. Those that see benefit to both institutions and systemically indicate that if they remain focused on their mission, and/or core area of operation, they will not be in competition with other institutions. Others within institutions are asking broad and strategic questions about student movement throughout the system that will be of benefit primarily to the institution only. However, these strategic questions can only be answered through data that is available on a more open level, and as such are willing to share it with others in order to get their own answers. There is a recognition now that institutions do not have the right data to fully analyze student movement through the system.

  To this point, one potential opportunity for further research lies in looking directly at how some large data sharing efforts, such as the OCAS initiative that provides institutions with data on where applicants to their programs actually wind-up accepting offers has worked-out in terms of the intersection of the benefit of the information provided to institutions versus the impact on competition between institutions that may prevent some from wanting a more open sharing of data between them.

When the above study occurs, then Ontario will be in a better position to consider implementing a larger data sharing protocol between institutions grounded in the OEN. The first step towards this, or a stepping stone towards it, may be to consider implementing a project like the STP in BC. That study releases limited amounts of both student level data and aggregate data on student movement through the BC system, and has a significant governance and data sharing structure to it. If an initiative like this using the OEN can be implemented in Ontario, it can be examined and used as a way to understand and grow an initiative that could develop into a full-on student data clearinghouse discussed by Black.

References

(2017, October). Retrieved from CIHR: http://www.cihr-irsc.gc.ca/e/29072.html

(2017, October). Retrieved from StudentClearingHouse.org.

Access to Student Records and Privacy. (2017, October). Retrieved from Durham College: https://durhamcollege.ca/wp-content/uploads/243-access-to-student-records-and-protection-of-privacy.pdf

Administration PDFs. (2017, October). Retrieved from Trent University: https://www.trentu.ca/administration/pdfs/CollectionNotice.pdf

Black, T. C. (2018, February). *Article - Fifth Way Plea Interconnected Central Student Data Depositories*. Retrieved from Groningen Declaration: http://www.groningendeclaration.org/article/fifth-way-plea-inter-connected-central-student-data-depositories

Corwin, B. (2013). *Student Transfer, Success, and Mobility in BC Post-Secondary Institutions.* Vancouver: BCCAT.

Data Sharing Agreement. (2017, October). Retrieved from Clinical Connect: https://info.clinicalconnect.ca/CC/wp-content/.../CC-Data-Sharing-Agreement.pdf

Durham College. (2016). *Giving Credit Where Credit is Due.* ONCAT.

Education and Training, Post Secondary Education, Data Research, Student Transitions Project. (n.d.). Retrieved from Government of British Columbia: https://www2.gov.bc.ca/gov/content/education-training/post-secondary-education/data-research/student-transitions-project

Emam, K. a. (2014). *De-Identification Protocols: Essential for Protectiing Privacy.* Information and Privacy Commissioner of Ontario.

Finnie, D. M. (2016, August). *Publications and Reports.* Retrieved from ONCAT: http://www.oncat.ca/files_docs/content/pdf/en/oncat_research_reports/2016-08-Final-Report-University-of-Ottawa-How-Student-Pathways-affect-Labour-Market-Outcomes.pdf

FIPPA at Carleton University. (2017, October). Retrieved from Carleton University: https://carleton.ca/privacy/fippa-at-carleton-university/

Hicks, D. a. (2017, October). *It's Not Academic Blog*. Retrieved from HEQCO: http://blog-en.heqco.ca/2017/04/fiona-deller-and-martin-hicks-spoiler-alert-its-the-ontario-education-number/

Information and Privacy Commissioner of Ontario. (2016). *De-Identificaiton Guidelines for Structured Dats.* Toronto.

Jarquin, P. B. (n.d.). *Data Sharing: Creating Agreements In support of Community-Academic Partnerships.*

Kelly Gallagher-Mackay. (2017). *Data Infrastructure for Studying Equity of Access to Postsecondary Education in Ontario.* Toronto: HEQCO.

Lana Ray, P. M. (n.d.). *INDIGENOUS PROGRAM PATHWAYS INVENTORY PROJECT PHASE ONE.*

Laws and Regulations Section. (n.d.). Retrieved from Ontario Government: https://www.ontario.ca/lawdats/regulation/900460

Laws and Statutes. (n.d.). Retrieved from Ontario Government: https://www.ontario.ca/laws/statute/90f31#BK53

Laws and Statutes. (2017, October). Retrieved from Ontario Government: https://www.ontario.ca/laws/statute/90f31#BK7

Library. (n.d.). Retrieved from Legislative Assembly of Ontario: http://www.ontla.on.ca/library/repository/mon/3000/10301262.pdf

Memorandum of Understanding between Durham Colege and UOIT. (2011, January 31).

News: University Affairs. (n.d.). Retrieved from University Affairs: https://www.universityaffairs.ca/news/news-article/canada-joins-network-to-improve-the-international-exchange-of-student-data/

Policies. (2017, October). Retrieved from Seneca College: http://www.senecacollege.ca/policies/fipp.html

Policies, Policy Libary, Legal Compliance Governance, Access to Information and the Protection of Pribacy Policy. (2017, October). Retrieved from UOIT: https://usgc.uoit.ca/policy/policy-library/policies/legal,-compliance-and-governance/access-to-information-and-the-protection-of-privacy-policy.php

Privacy and De-Identification Centre. (2017, October). Retrieved from Information and Privacy Commissioner: https://www.ipc.on.ca/privacy/de-identification-centre/

Privay Policy. (n.d.). Retrieved from Carleton University: https://carleton.ca/privacy/wp-content/uploads/policy_collection1.pdf,

Quan, D. (2017). *Unlocking Student Potential through Data.* Toronto.

Registrar - Privacy . (2017, October). Retrieved from York University Website: https://registrar.yorku.ca/privacy

Smith, R. D. (n.d.). *Transfer Pathways in Postsecondary Education: York University and Seneca College as a Case Study.* Higher Education Quality Council of Ontario.

Snapshot from Community College to Graduate and Professional Degrees. (2017, October). Retrieved from NSC Research Centre (National Student Clearinghouse Research Centre): https://nscresearchcenter.org/snapshotreport-from-community-college-to-graduate-and-professional-degrees30/

Student Transititons Project. (2017, October). Retrieved from Government of British Columbia: https://www2.gov.bc.ca/assets/gov/education/post-secondary-education/data-research/stp/stp_isa_-_nov_2016_update.pdf

Van den Eynden, V. e. (n.d.). *“Managing and Sharing Data”, Best Practices Guide For UK Researchers.* Uk Data Archieve.

What We Do, Business Intelligence. (2017, October). Retrieved from OCAS: https://www.ocas.ca/what-we-do/business-intelligence