Leveraging your data – a collaborative approach with IT

Richard Levin, Executive Director, Enrolment Services & Alexandra Agostino, Business Intelligence Project Leader

University of Toronto
Presentation Agenda

- About U of T
- Enrolment Services – UTBI collaboration
- Development Methodology
- Demos
- Lessons Learned
- Questions
Who is in the audience?

- Registrars
- Recruitment
- Admissions
- IT/Systems
- Financial Aid
- Data Analyst
- Other

Leveraging your data – a collaborative approach with IT, Session 1504, #SEM2018
UofT - General Facts & Figures
#21

World University Ranking
by Times Higher Education
#2 in the world for research output
$15.7B
annual impact of research on the Canadian economy
12th
in global employability
More than 557,000 alumni in more than 190 countries and territories
Enrolment

Total Enrolment 88,766

- Downtown (St. George) 60,595
  - Undergrad 43,523
  - Graduate 17,072
- UTSC 13,430
  - Undergrad 13,132
  - Graduate 298
- UTM 14,741
  - Undergrad 14,073
  - Graduate 668
- International 17,452
  - Undergrad 14,467
  - Graduate 2,985
- Domestic 71,314
  - Undergrad 56,261

Leveraging your data – a collaborative approach with IT, Session 1504, #SEM2018 #SEM18
Lexicon

Division = Major academic unit (campus or faculty or college)

Faculty = US College or School, e.g., Arts and Science, Music, Architecture, Medicine, Engineering

College = Smaller student community within Arts and Science. Home of advising and most residences.
Leveraging your data – a collaborative approach with IT, Session 1504, #SEM2018

26 Academic Divisions

Seven “first-entry” divisions:
1. Arts and Science
2. Applied Science and Engineering
4. Music
5. Architecture and Landscape Design
6. University of Toronto Mississauga
7. University of Toronto Scarborough

Second-entry professional and graduate programs
Structure

- Each Division has a registrar, e.g., Arts and Science + Colleges, Engineering, Law, Medicine, Dentistry, etc. plus Mississauga and Scarborough campuses.

- Each registrar has enrolment responsibilities, plus some recruitment, admissions and financial aid responsibility.

- Registrars report within their own divisions.

- SIS is homegrown; Ontario secondary school students apply through the Ontario Universities Application Centre.

- Incoming first-year students are admitted to program streams, select majors after first year.
Enrolment Services Facts and Figures

- Responsible for admissions, recruitment, financial aid, transcripts and registrarial policy
- Serves primarily the seven first-entry divisions
- Processes about 90,000 applications per year
- Approx. 60% accept rate
- Annual incoming first-entry enrolment of about 16,000
- About 30% of incoming students are International
How do students with a high incoming average perform?
Fall 2010 cohort
Why did this collaboration start?

- Started in approximate 2012/13 and was encouraged for a number of reasons including:
  
  - Recognition that gathering basic statistics was very labour intensive with no ability to even explore beyond the basics
  
  - Rising acknowledgement that data is valuable
  
  - Tools/techniques becoming available to store/analyze data
  
  - Next Generation of SIS system project to dedicate resources for a BI team
Value of Data

- Allocation of recruitment staff
  - 900 high schools in Ontario, significant international recruitment
  - Where to focus efforts?

- Admissions
  - Strong correlation with high school marks
  - Can we make it better?

- Financial Aid
  - What’s the yield on various scholarships
  - Scholarship combinations
Value of Data

➤ International
  • Regional performance: apps, yield, citizenship vs. country of origin
  • Effectiveness of initiatives

➤ Demand for data from many stakeholders
  • Complex data sets
  • Need for consistency
  • Reduces demand on central staff
Show of hands

- How many feel they are using an evidence-based approach to recruitment and admissions?
- How many schools are aware of whether their school has a BI team/analytics team? Does this exist at your institution?
Initial goal of our warehouse

➢ To develop a 360 degree view of the student.

➢ It started out with this view:
Overview of Warehouse Content

- 8 years to develop
- Currently working on adding Retention
How Projects Begin

- **Give Me the Answer**: Point & Click Reporting
- **Show Me the Data**: Drag & Drop Inside a Data Sandbox
- **Let Me Explore**: Drag & Drop Reporting connecting to any data
- **Statisticians**: Create their own code/queries
- **Developers**: Create their own code/queries

Users with Technical Expertise
- Project Management - JIRA
- Storage – IBM DB2 Blu
- ETL – Informatica PowerCenter
- Data Analysis – SQL queries, Tableau Dashboards, Cognos PowerCubes & Reports…and limitless other tools
Leveraging your data – a collaborative approach with IT, Session 1504, #SEM2018
Leveraging your data – a collaborative approach with IT, Session 1504, #SEM2018

- **Non-IT Team**
  - Business Lead
  - Functional User Group

- **IT Team**
  - Data Architect
  - Database Administrator
  - Analyst(s)
  - ETL Developer(s)
  - Report Developer(s)
  - Project Manager (Scrum Master)
Culture in place to adopt agile methodology for IT project management

SIS staff members attended a weeklong “agile crash course”

Encouraged to adapt the methodology to suit business intelligence

We had coaches available to lean on and in turn, we encourage others
- More collaboration…a business lead is named and they aren’t an IT person

- Less focus on documentation and more on development and collaboration

- More iterations gives users exposure to deliverables early in the process

- If things are heading in the wrong direction, this can be detected earlier in the process
### Tools

Not like this....

1. [Angry face](#)
2. [Dumbbell](#)
3. [Car](#)
4. [Happy face](#)

### Team

Like this!

1. [Sad face](#)
2. [Skateboard](#)
3. [Bicycle](#)
4. [Motorcycle](#)
5. [Car](#)

Source: https://agiletribe.wordpress.com/2016/02/19/incremental-development/
Selling this methodology to the organization

- We basically indicated to clients that agile was the new direction we were heading towards
- Asked for a greater commitment/participation of administrative & academic staff
- Better deliverables will come about with an increase in participation and collaboration
Project Approach

- Stakeholders nominate university community members to participate in the functional user group (FUG)
- FUG members articulate and prioritize User Stories
- User Stories address real business needs
Who should be part of your FUG?

- Is an analyst that year in and year out painstakingly prepares data for you
- Is a staff member who understands your business processes and their quirks
- Is an academic who asks good questions about the department, faculty or institution
- Is a staff member who interacts with students and wants data to support their work...to in turn support students

And is committed to working as part of a larger team!
Anatomy of a User Story

- FUG members compile and submit User Stories
- The story should be a statement in business-like language describing a need
- They have a structure of Who, What, Why
Anatomy of a User Story

Who
As a Registrar, I need to understand the year-over-year changes in applications by program to ensure we are aligning to targets.

What
An unexpected surge or drop in applications can have a significant impact in how we...

Why
What are the Prioritization Factors

- Should be between 4-7 factors
- This story is important because...
  - Our President, Provost, Registrar, etc. always asks us for this information
  - It will be used to meet government regulatory requirements
  - It provides us with data we currently don’t have or takes several people and several days/weeks to prepare this data manually
  - It helps us understand if we are fulfilling our strategic goals
### Prioritization Matrix

<table>
<thead>
<tr>
<th>User Story</th>
<th>Is important to President</th>
<th>Manual effort to prep data</th>
<th>Required for Gov't Reporting</th>
<th>Data we currently don’t have</th>
<th>Score Total</th>
<th>Possible Points</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Story 1</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>15</td>
<td>20</td>
<td>75%</td>
</tr>
<tr>
<td>User Story 2</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>13</td>
<td>20</td>
<td>65%</td>
</tr>
<tr>
<td>User Story 3</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>11</td>
<td>20</td>
<td>55%</td>
</tr>
<tr>
<td>User Story 4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>15</td>
<td>20</td>
<td>75%</td>
</tr>
<tr>
<td>User Story 5</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>14</td>
<td>20</td>
<td>70%</td>
</tr>
</tbody>
</table>
Architect a Data Repository

- Physically design data mart(s) to support data that answers the User Stories

- We typically off-load this data from our student (and other systems) as these are resource intensive processes...hence we shift analysis to a data warehouse

- Reporting from the warehouse is strategic not operational

- Data has latency...could be past term or yesterday
Star Schema Architecture

Incoming Marks DIM
- Interim Avg
- Final Avg
- Best 6 Senior Avg

Measures
- Applied
- Admitted
- Accepted
- Registered
- Total Schlp

Incoming School DIM
- School Name
- School Type
- Address

Faculty DIM
- Faculty Name
- Stream
- Program Name

Candidate DIM
- Name
- Age
- Citizenship
- Fee Type
- Address

Applicant Type DIM
- Applicant Type
- English Proficiency Requirement

Leveraging your data – a collaborative approach with IT, Session 1504, #SEM2018
Leveraging your data – a collaborative approach with IT, Session 1504, #SEM2018

**Dimensions**

- Incoming Marks DIM
  - Interim Avg
  - Final Avg
  - Best 6 Senior Avg

- Faculty DIM
  - Faculty Name
  - Stream
  - Program Name

- Incoming School DIM
  - School Name
  - School Type
  - Address

- Applicant Type DIM
  - Applicant Type
  - English Proficiency Requirement

- Candidate DIM
  - Name
  - Age
  - Citizenship
  - Fee Type
  - Address

**Measures**

- Applied
- Admitted
- Accepted
- Registered
- Total Schl’p

<table>
<thead>
<tr>
<th>Student</th>
<th>Choice</th>
<th>Faculty</th>
<th>Program</th>
<th>Incoming School</th>
<th>Applied</th>
<th>Admitted</th>
<th>Accepted</th>
<th>Registered</th>
<th>Schl’p Amt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary Smith</td>
<td>1</td>
<td>Engineering</td>
<td>Chemical Eng.</td>
<td>Lakeview HS</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>$0</td>
</tr>
<tr>
<td>Mary Smith</td>
<td>2</td>
<td>St. George</td>
<td>Life Science</td>
<td>Lakeview HS</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$0</td>
</tr>
<tr>
<td>Mary Smith</td>
<td>3</td>
<td>UTSC</td>
<td>Commerce</td>
<td>Lakeview HS</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$5,000</td>
</tr>
<tr>
<td>Mary Smith</td>
<td>N/A</td>
<td>UTM</td>
<td>Humanities</td>
<td>Lakeview HS</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>$7,000</td>
</tr>
</tbody>
</table>
Our Admissions Datamart
Why model (transform) data?

- Typically to create a life cycle around data or create pipelines

- To create snapshots of your data and key business points

- Applies business logic…if a student gets an offer to an alternate program or campus this doesn’t count as an application…but from an admitted perspective, does count
What happens during a Sprint?

- **Sprint planning** - Business Lead works with IT team to break down the user stories that will be tackled into tasks.

- **Tasks** have estimated completion times and team member names and are loaded to JIRA and posted on a whiteboard.

  - Create mapping to populate the Faculty Dimension
    - **4 days**
    - **Wk 1**

---

Leveraging your data – a collaborative approach with IT, Session 1504, #SEM2018

#SEM18
JIRA with swim lanes
Daily Aspects of the Sprint

- Daily 15 min scrum around the agile board

- Each team member answers 3 questions:
  - What did I work on yesterday
  - What do I plan on working on today
  - What obstacles are in my way

- Business Lead encouraged to attend
During the Sprint – Check-in with the FUG Group

- Progress updates

- Make decisions on data transformations…if M then Male, if F then Female, if blank then what?

- Clarify business rules such as “how is average GPA to be calculated?”

- Design dashboards/reports/analysis cubes deliverables…including prototyping where possible
The process carries on from Sprint 1 - Sprint n

- Release of Deliverable
- Determine user stories to tackle
- Sprint Planning to define tasks
- Enhancements made based on feedback
- Check-ins with FUG
- Daily Sprint Check-ins

Leveraging your data – a collaborative approach with IT, Session 1504, #SEM2018
Stakeholder’s Perspective

- This has been a learning experience for stakeholders too…”When they say Sprint, they mean it.”

- We’ve had to provide feedback quickly and be more succinct in defining needs

- We’ve found the scripting of User Stories to be more digestible and meaningful in terms of an expected deliverable

- We understand that there are more stories than time permits therefore prioritization is key
Stakeholder’s Perspective

- We have a sense of the process from concept to delivery
- We’ve become more aware of risks and bottlenecks in the development process
- We are participating in projects to a much higher degree than in the past
Demos

1. Singapore Summary – An emerging market

2. Onboarding new VP Int’l Relations – understanding latest school vs citizenship

3. Engineering Gender Analysis

4. Tableau – Recruitment Officer Analysis
Lessons Learned

- Timing is everything…are you about to switch student systems…then this isn’t the time to start a project like this

- Be aware that your non-IT staff have competing priorities…therefore will they be able to devote time to this project

- Data is never clean…processes change over time…so be suspicious of your data

- Carefully discuss “the time” aspect with users to have a clear understanding of their needs

- The more examples (sketches on napkins or copies of reports) the better…helps clarify the story
Post Project

- Consider how you’ll keep engaged post-launch. What type of training and on-going support is required?

- Maintenance – what resources will be dedicated to dealing with bugs or “small” new requirements?
Usage statistics – monitoring of hourly activity

Nov 1, 2017
Usage – ranges from 14-16,000 hits / month
Admissions Only Usage - 3-4,000 hits / month

Feb/Mar launch of new Admissions tools resulting in extensive training.
Current Project - Retention

CSRDE Definition of First Time, Full Time, 4 Year Direct Entry Undergraduate Degree Seeking

Retention – User Stories

➢ As a division/faculty, we would like to know if financial aid, or lack of resources (Grant/bursary awards/parental income), have an impact on retention/graduation.

➢ As directors of Accessibility Services, we want to know the retention and graduation rates of students registered with Accessibility services based on their primary type and time of diagnosis so that we can better understand the timespan of support.

➢ As a registrar, I would like to understand the likelihood of success for students who live first year in residence compared to students who never live in university residence. Does this differ for international vs. domestic students?
Demos

5. Retention – Early deliverables
Future Technical Direction

- **NGSIS Data Lake Pilot**: A secure storage repository holding vast amounts of raw data in its native format.
- **Divisional Data Stores**: Divisions can store their data in the cloud and lower data maintenance work.
- **Tableau expansion**: Expand capacity, licenses for developers and more training and support.

Diagram includes:
- Data Lake
- Machine Learning
- Analytics
- SQL
- Cloud
- On-premises Data Movement
- Real-time Data Movement
Questions

Contacts

richard.levin@utoronto.ca
alexandra.agostino@utoronto.ca