Data Integration for Future Success

From Self-Service at Scale to Unifying Efforts and Collaboration to Transform the Office of International Education Reporting Strategies & Office of Undergraduate Education

Presenters:

Kenji Terawaki
Business Intelligence Developer
Office of Information Technology

Jimmy Kriigel
IT Engineer Lead
Office of International Education

Stefan Smith
Data Scientist
Office of Undergraduate Education

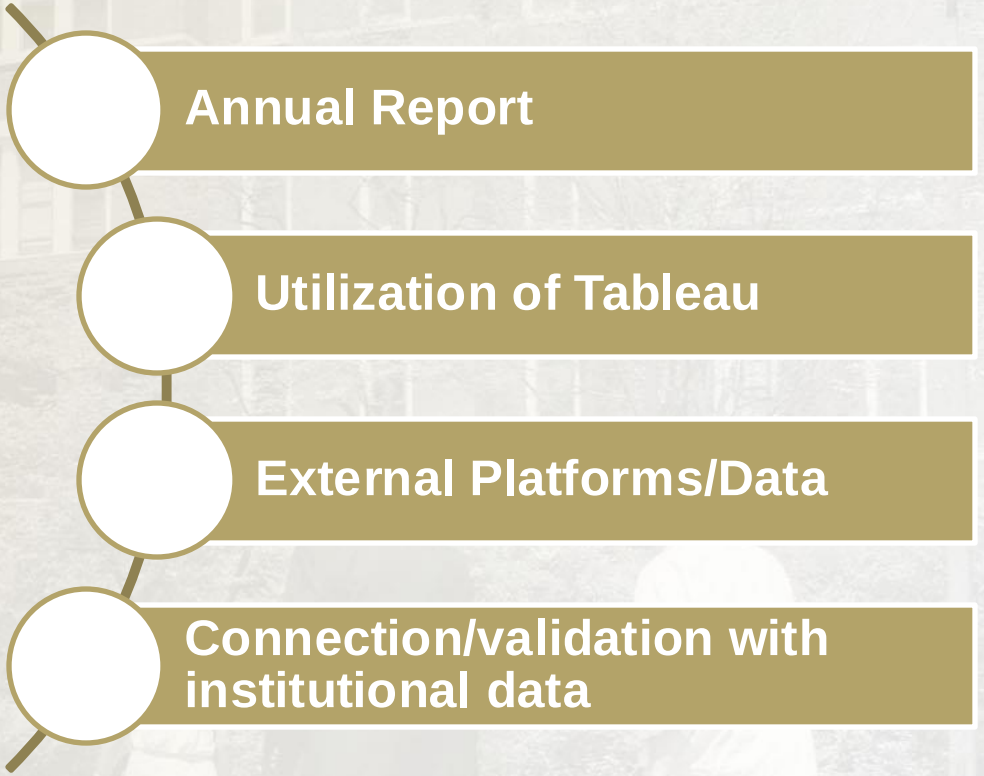
Georgia Tech

Office of International Education (OIE)

Education Abroad gives students an opportunity to improve their global competencies while meeting their academic goals and fulfilling their degree requirements.

International Student and Scholar Services helps international students and scholars with their initial adjustment to GT and Atlanta, provides information and advice about immigration regulations, cultural differences, financial matters and personal concerns all while assisting them throughout their stay through cross-cultural programming.

On Campus Internationalization offers programming that allows campus to globally engage with diverse cultures, learn about important global issues, develop global competencies, and participate in meaningful intercultural learning—all



Georgia Tech

Challenges – Office of International Education
Education Abroad

Terra Dotta is a cloud-based software management of study abroad programs, exchanges, and international experiences (Internships, Affiliated Travel Programs).

TERRADOTTA

- Platform is cloud-based and closed-source
- Each program may involve complex processes
- A one-size-fits-all student application creates data challenges for specialized programs
- Banner indicators integration
- Data does not feed into Banner
- Term cycle is Fall -> Spring -> Summer differs from Summer -> Fall -> Spring

Georgia Tech

Challenges – Office of International Education International Students and Scholars (ISSS)

- Sunapsis - Comprehensive immigration compliance software. Interfaces with SEVIS – U.S. Immigration and Customs Enforcement
- Closed Source Platform
- Hosted on-premises - Ingestion into data lake
  - ✓ Over 400 tables to shift through for relevant data sets.
  - ✓ Requires complex queries for data retrieval
  - ✓ Some historical data resides in archived tables
- Data does not feed back into Banner.

SUNAPSIS®
AN INDIANA UNIVERSITY INNOVATION

Georgia Tech

Challenges – Office of International Education
Living Learning Communities / International Plan

- **IHouse, I-Lead, Global Leadership Living Learning Community & Techsplore Exchange Program**
  - Data from these communities shared the same bursar account and was segmented by Qualtrics form
- **International Plan (IP) Program Degree Designation**
  - IP requires Language and Abroad components.
  - Requires yearly milestones to achieve.
  - Banner indicators
    - Identification of student taking IP courses
    - IP Degree designation
  - Data does not feed to Banner

qualtrics XM

Georgia Tech

Complexities in Data Integration

- Three Workgroups
- All different data needs
- Data stored on different platforms
- Difficult Data Ingestion
- Manual Data export prone to integrity issues.

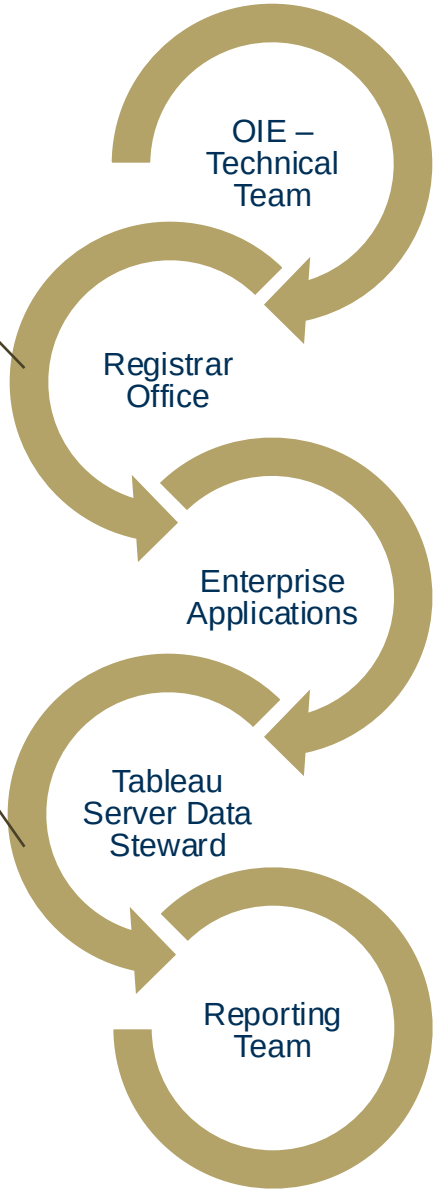
Challenges felt overwhelming

Effective Cross-Departmental Collaboration

An Opportunity for Collaboration

- Banner indicators for ilead, International Plan, and Techsplore

- Create Tableau projects for each unit
- Assigned roll data security for each folder



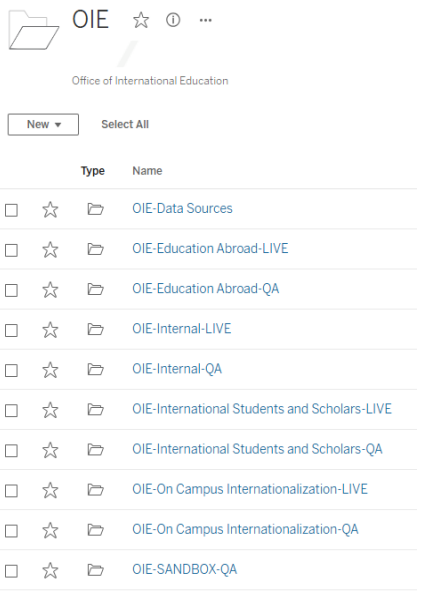
- Coordinate all three units with their data needs, reporting, data collection, Sunapsis expertise, OIE technical/data and business processes

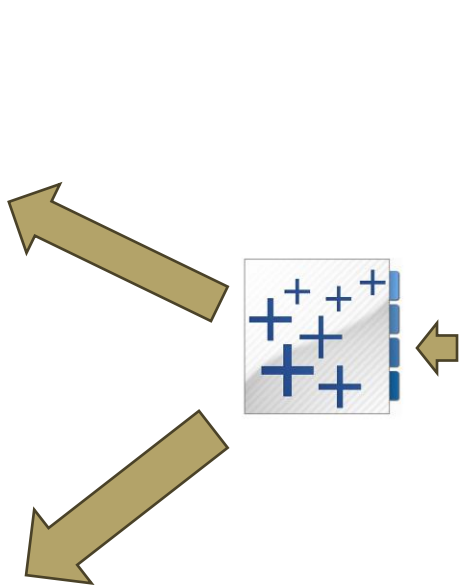
- Add IP Graduated indicator to data source
- Add Techsplore and Ilead indicators to data sources

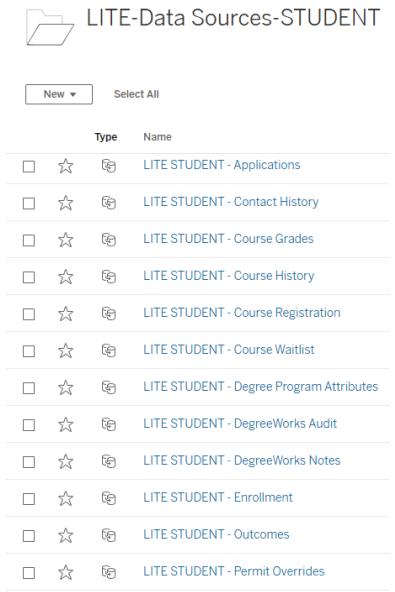
- Access to institutional data
- Facilitate Tableau reports development
- Guidance on data blending, data connection and data sharing

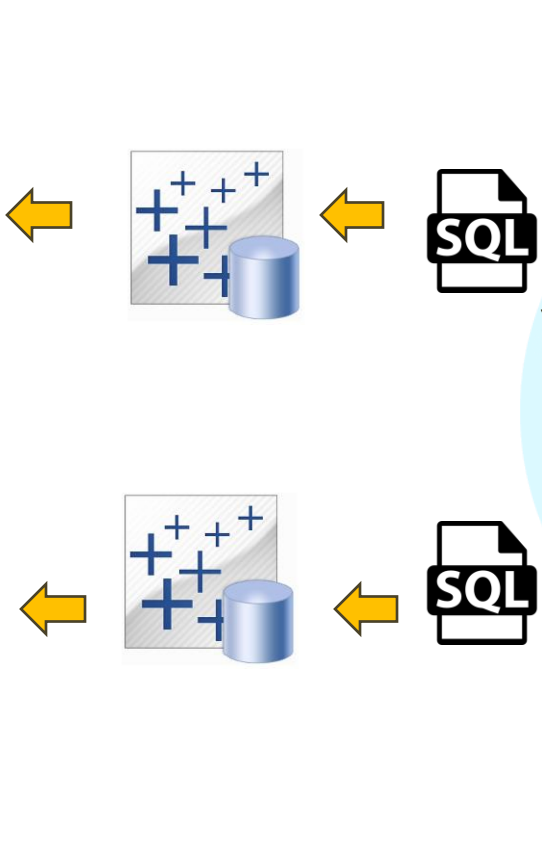
Georgia Tech

OIE Data Cycle









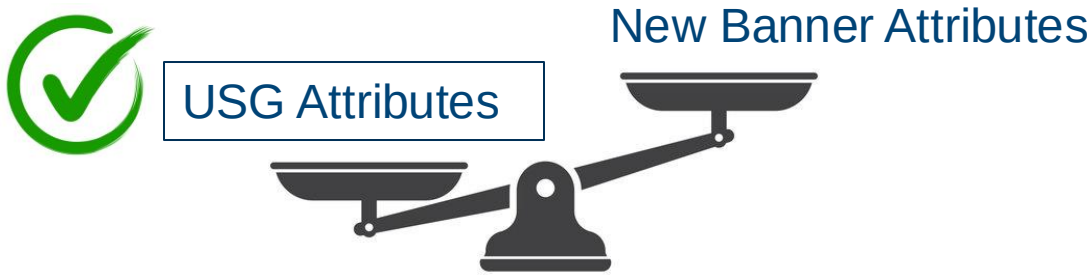
LITE.Gatech.edu

Path Forward

Georgia Tech

Education Abroad - Solutions

- Education Abroad Indicators (Banner)
- There are multiple scenarios in where students can register their studies abroad
- How to flag students in non-Georgia Tech programs?



TERRADOTTA

- Review application/registration form
  - Improved control over data collection through customized registration form fields
- Ingest just what's needed!
  - Location + Term

Georgia Tech

International Students and Scholars (ISSS) - Solutions

- Sunapsis successfully ingested into data lake
- Cross team collaboration finding relevant data out of 400+ tables

- Work done: SQL - TDS
- Academic Training J1
- Scholar and Interns J1
- OPT F1

```
SELECT
    -- SEVIS Data
    sevis.sevisid AS "SEVIS ID",
    sevis.status AS "SEVIS Status",
    statusCodes.description AS "SEVIS Status Description",
    categoryCodes.description AS "Category Description",
    subjectCodes.description AS "Subject Field Description",
    -- From iuieBio table
    bio.idnumber AS "ID Number",
    bio.PRSN_UNIV_ID AS "University ID",
    -- From iuieTerm table
    term.ACAD_CAREER_CD AS "Academic Career Code",
    term.INST_CD AS "Institution Code",
    term.ACAD_TERM_CD AS "Academic Term Code",
    term.CIP_CODE AS "CIP Code",
    -- From iuieVisaCitizenship table with country information
    visa.PRSN_UNIV_ID AS "Visa University ID",
    visa.PRSN_VPRMT_TYP_CD AS "Visa Permit Type Code",
    visa.PRSN_CTZN_CNTRY_CD AS "Citizenship Country Code",
    visa.PRSN_CTZN_STATUS_CD AS "Citizenship Status Code",
    visa.DATESTAMP AS "Visa Datestamp",
    country.code AS "Country Code",
    country.description AS "Country Description",
    country.region AS "Country Region",
    country.subregion AS "Country Subregion",
    country.continent AS "Country Continent"
FROM
   sunapsis.sevisDS2019Program sevis
LEFT JOIN
    sunapsis.codeSEVISStatus statusCodes ON sevis.status =
    statusCodes.code
LEFT JOIN
    sunapsis.codeSEVISJ1Position positionCodes ON
    sevis.positionCode = positionCodes.code
LEFT JOIN
    sunapsis.codeSEVISJ1Category categoryCodes ON
    sevis.categoryCode = categoryCodes.code
LEFT JOIN
    sunapsis.CodeSEVISProgSubject subjectCodes ON
    sevis.subjectFieldCode = subjectCodes.code
LEFT JOIN
    sunapsis.iuieBio bio ON sevis.idnumber = bio.idnumber
LEFT JOIN
    sunapsis.iuieTerm term ON sevis.idnumber = term.idnumber
LEFT JOIN
    sunapsis.iuieVisaCitizenship visa ON sevis.idnumber =
    visa.idnumber
LEFT JOIN
    sunapsis.codeCountry country ON visa.PRSN_CTZN_CNTRY_CD =
    country.code
```

SQL INTENSIFIES

Georgia Tech

Living Learning Communities / International Plan - Solutions

- **IHouse, I-Lead:** A new indicator was assigned to I-Lead to track students in Banner
- Reconsideration on using Qualtrics as an intake form
- **Techsplore:** A new indicator was assigned to Techsplore to track students in Banner
- **International Plan (IP) Program**
  - Added a Student International Plan Ever flag
  - Added an IP Department Honor Indicator to Outcome data source

OUE – Office of Undergraduate Education

My Job:

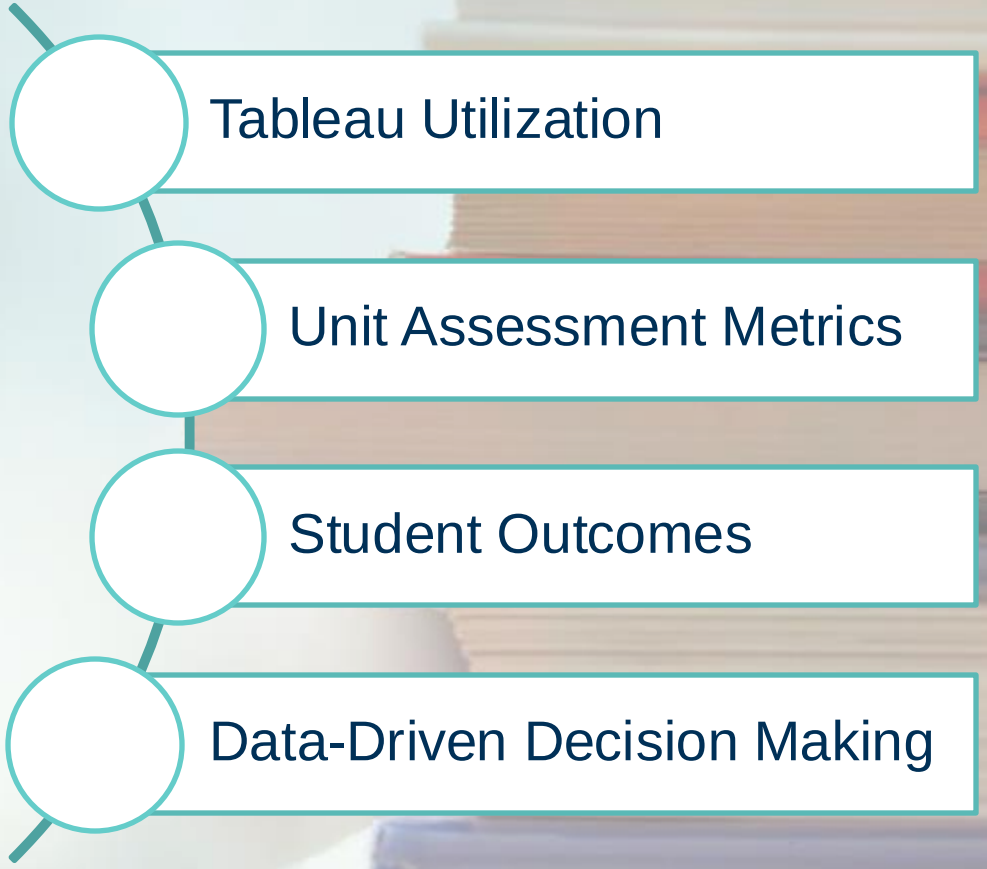
- Data Scientist, Analytics & Planning Unit
  - onboarded to LITE Tableau June '23

My scope at work

- Support various unit initiatives
  - Data Deliverables (LITE Data Extracts, and more)
  - Dashboards (LITE-based, blended)
  - Data Analytics Projects
- Improve Data Stewardship

My Challenges

- Differing Cross-Campus Data Definitions
- Blending Data Sources
  - LITE (Tableau)
  - Various Platforms (non-ingested)
    - TracCloud, AdvisorLink, Symplicity, Engage



Effective Collaboration

An Opportunity for Consolidate – Collaborate – Communicate

Data Scope

- Advising
- Tutoring
- Events
- Research
- Internships/Co-Ops

Analytics

- Participation #'s
- Demographic Breakdowns
- Regression Analysis
- Predictive Modeling

Communication

- Results
- Follow-up Initiatives
- Derivative Questions
- Able to collaborate with other departments

Georgia Tech

Effective Collaboration

- Alignment of Uniform Data Definitions (primarily external systems)
  - How attributes are derived. How this may not align with GT's standard definitions.
  - Changes in definitions over time as needs change.
  - Gaining cross-campus buy-in for consistency.
- Ingest new & existing platform data to Datalake and make it accessible to all developers
  - New: EAB Navigate
  - Existing: TracCloud, AdvisorLink, Engage, Symplicity

Georgia Tech

What
To Take Away

Integration Team is working on ingesting as many systems into data lake upon client request and priority needs.

- Closed Source & Cloud Based Applications are hard to ingest.
- On-premise Applications are easy to ingest but hard to query.
- Progress not perfection
- Shifted from solo development to cross-campus collaboration with fellow developers
- Data ingestion is the solution; data understanding is the real challenge.
- Moved from simply requesting/matching numbers to seeking deeper analysis and insights.
- Focused on uncovering the true numbers, not just the ones we believe to be true.
- Data security (Presentation: **A Programmatic Approach to Scalable Tableau Server Administration by Justin Boone) Wednesday 8:00 am**

Questions?

Jimmy Kriigel
IT Engineer Lead
Office of International Education
jimmy.kriigel@oie.gatech.edu

Kenji Terawaki
Business Intelligence Developer
Office of Information Technology
kterawaki@gatech.edu

Stefan Smith
Data Scientist
Office of Undergraduate Education
stefan.smith@gatech.edu

Georgia Tech