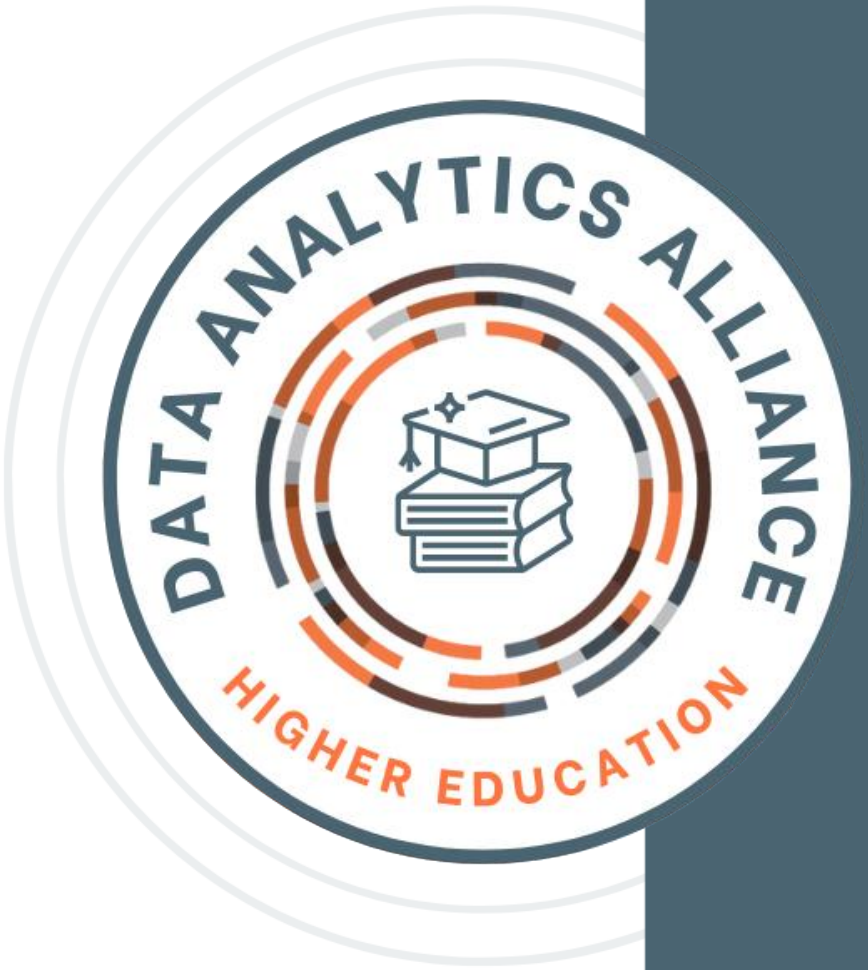
Data Analytics Alliance for Higher Education

Q2 Meeting – April 22, 2026



Meeting Agenda

- DAAHE Welcome & Overview
- Guest Speaker – **Brent Drake, PhD**, *Senior Vice President for Operations & Research* at the Gardner Institute, and *Chair* of the Board of Directors for the Association of Institutional Research
- Announcements

DAAHE Purpose

The DAAHE is a national community of higher education professional developed for the purpose of being a catalyst and resource for Higher Education organizations that are interested in using and understanding Data Analytics to solve higher education issues.

The Basics

- Peer-to-peer networking and sharing of Data Analytics best practices
- Collegial discussion of Data Analytics business and technology challenges
- Joint initiatives to address significant Data Analytics business and technology issues

Meeting Format

Meeting frequency:
Quarterly

Survey/discussions
between meetings

- Topic Selection
  - Group suggestion
  - Presented by members
  - Online collaboration
- Focus on one topic per meeting
  - Topic
  - Networking

Important Disclaimers

As a special interest group, the Data Analytics Alliance for Higher Education (DAAHE), as well as its members and participants in DAAHE meetings, must be constantly vigilant of federal and state antitrust laws, which prohibit monopolization, combinations and conspiracies in restraint of trade, collusive action, and unfair methods of competition.

To avoid the most sensitive areas, the following matters should never be discussed:

- Price, student or any elements of price or tuition policies, including costs, discounts, and similar matters
- Sensitive information – Employee or student
- Commercial liabilities, warranties, guarantees, or the terms and conditions of sales, including
- Anything dealing with “arm twisting,” trade abuses, or excluding or controlling competition.

As a DAAHE representative, my job is to prevent such conversations from occurring at this meeting. Please understand that if I suggest that a subject should not be discussed at this meeting, I will do so to ensure that our association meeting remains both risk-free and extremely productive.

Additionally, we seek active member participation, not participation by AI assistants.

Let’s start the program

The Use of Generative AI
in Institutional Research
and Effectiveness
with
Brent Drake, PhD
from the Gardner Institute

The Use of Generative AI in Institutional Research and Effectiveness

Brent Drake, Senior Vice President for Operations & Research
Chair Association of Institutional Research 2024-2026

GARDNER INSTITUTE

Agenda

- History and Current State of AI Use
- Practical Considerations
- Current Examples of Generative AI Use
- Questions

History and Current State of AI Use in IR/IE

History in IR/IE

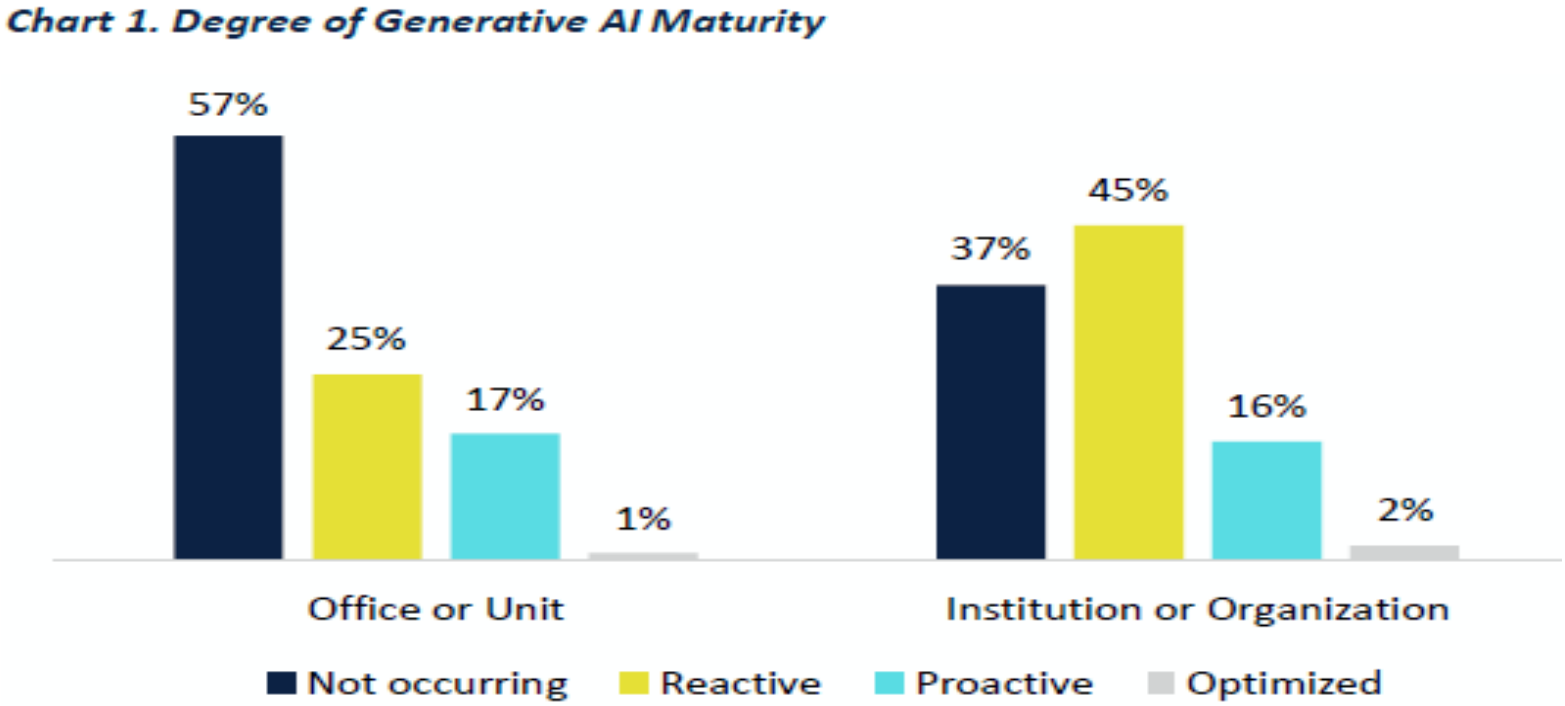
Early AI and ML

- Early iterations of AI in particular Machine Learning (ML) was a well know tool within Institutional Research
- Launch of public OpenAI Chat Generative Pre-trained Transformer (ChatGPT) 2022
- Public environment now has multiple generative tools
  - Plugin enhancements to ChatGPT
  - Generative Art Tools
  - Multimodal Large Language Models (MLLM)
  - Autonomous AI-agents

Urmeneta, 2023 Reflections on AI Transformation

Current State IR/IE

AIR Survey on Generative AI



AIR, 2023 Use of Generative AI Survey

GARDNER INSTITUTE

Current State IR/IE

AIR Survey on Generative AI

Chart 2. Using Generative AI Can Improve the Efficiency of an Office

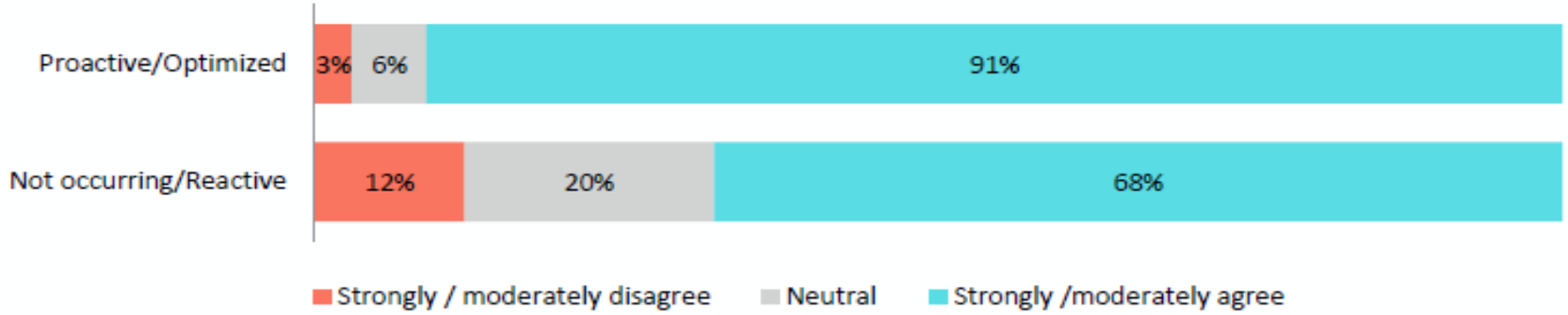
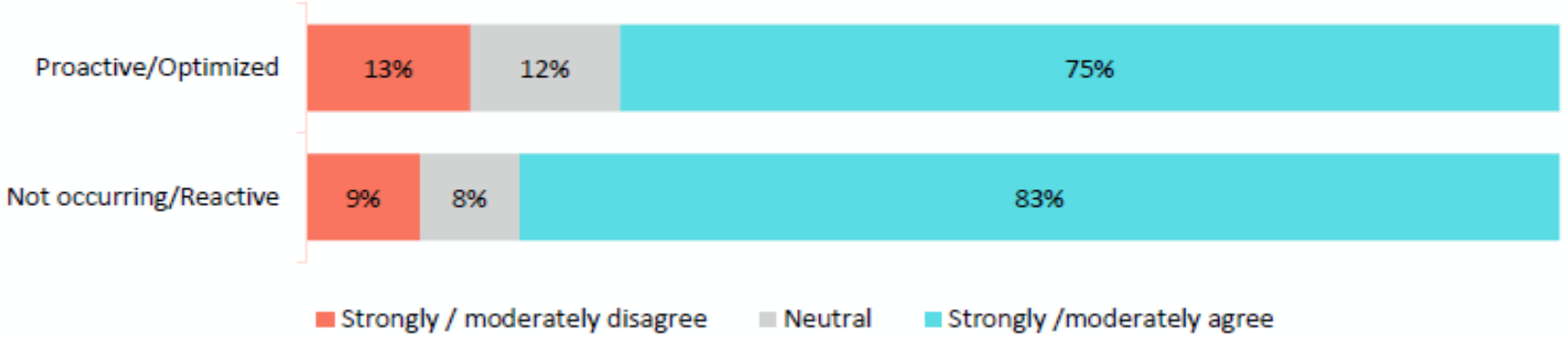


Chart 3. Using Generative AI Will Require Significant Professional Development



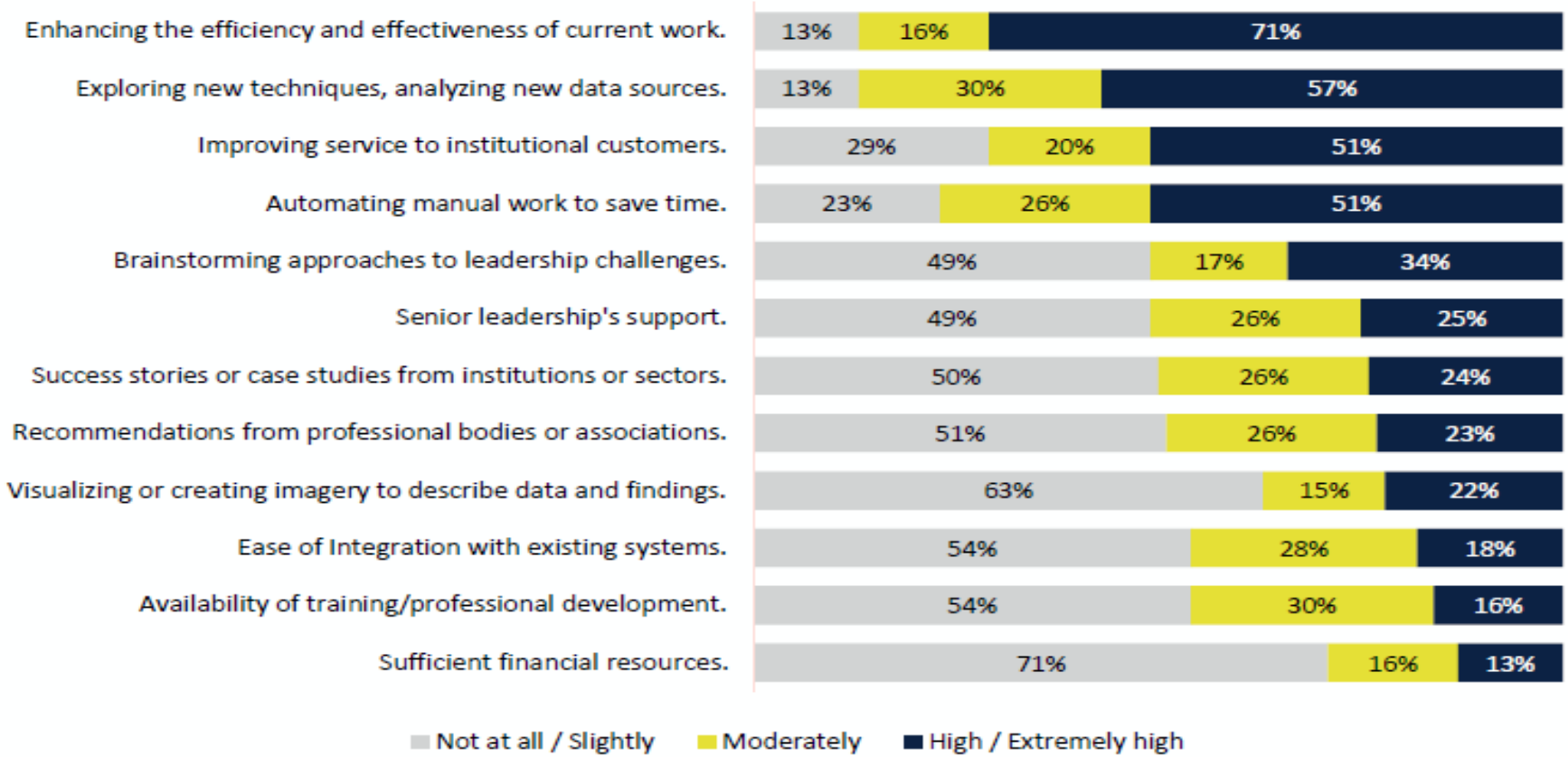
AIR, 2023 Use of Generative AI Survey

GARDNER INSTITUTE

Current State IR/IE

AIR Survey on Generative AI

Chart 5. Degree that Items are Reasons Why IR/IE Offices Use Generative AI



AIR, 2023 Use of Generative AI Survey

GARDNER INSTITUTE

Current State IR/IE

AIR Survey on Generative AI

Table 1. Generative AI Solutions

Solution	% of Total
Chat GPT	73%
Bing Chat	5%
Claude	5%
Bard	5%
Canvs AI	2%
Dall-E	2%
Tableau GPT	2%
AdAstra	2%
Miscellaneous	4%

Table 2. Generative AI Tasks

Task	% of Total
Development of content (e.g., code, emails, reports, summaries, templates)	59%
Summarizing information (e.g., open-text questions, quantitative data, articles)	20%
Researching information	6%
Brainstorming/ideation	6%
Image generation/graphic design	4%
Miscellaneous	5%

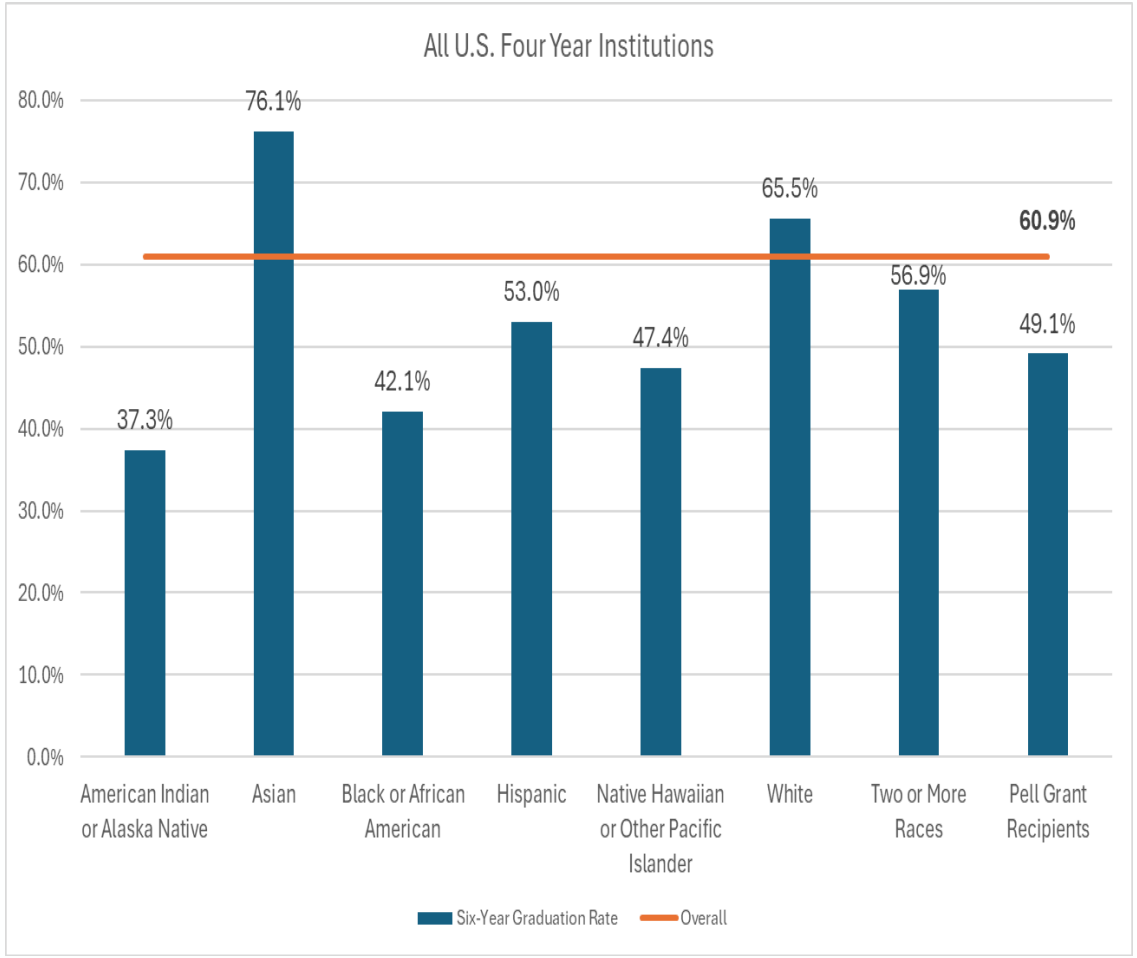
AIR, 2023 Use of Generative AI Survey

Practical Considerations

- Ethics
- Governance
- Privacy and IP
- Human in the Loop

Ethics

- **Inherent Bias** – biases in underlying data are magnified
- **Hallucinations** – tools still make guesses



GARDNER INSTITUTE

Governance

- What tools is your organization comfortable with using
- Who can access the tools
- Create an accountability chain

Privacy and I.P.

- Most tools presently being used are public sourced
  - Utilize data minimization
  - Must protect user data and PII

- Protect your institutional Intellectual Property
  - Necessitate private network RAG models

Human in the Loop

“another critical aspect that limits the use of AI in IR is the lack of inherent understanding of the cultural and socio-political contexts form which the data originates…resulting in outputs that overlook or misinterpret critical subtleties, particularly on how algorithms tend to embed values and social practices from where such algorithms are conceived” (Avila et. al. 2025, Sailing through he Challenges of AI in IR)

Some Current Examples

Coalition of OH, GA, KY Institutions

Our Solution: AI-Powered Analytics Dashboard

01

Natural Language Queries

Simply ask questions in plain English—no technical expertise required to extract powerful insights from complex datasets.

02

Instant Visualizations

AI generates charts, graphs, and dashboards on-demand, customized to your specific question and institutional context.

03

Predictive Models

Built-in forecasting analyzes historical patterns to predict retention rates, readiness gaps, and intervention opportunities.

04

Export & Share

Download insights with metric explanations to share with stakeholders and decision-makers.

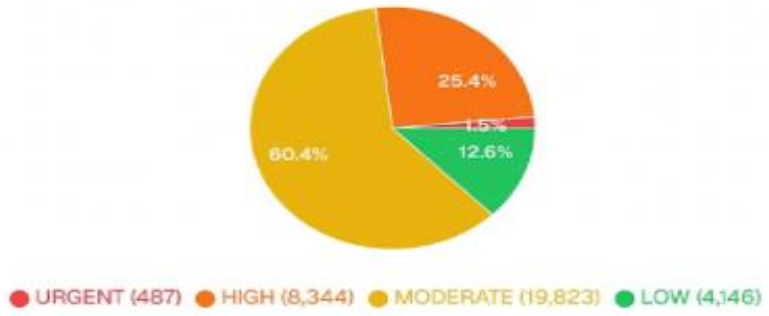
Student Success Dashboard

Export Report | SQL Query Interface

Overall Retention Rate	Avg Predicted Retention	Students at High/Critical Risk	Avg Course Completion
51.5%	51.5%	8,831	82.3%
32,800 total students	ML model prediction	Require immediate intervention	Credits earned / attempted

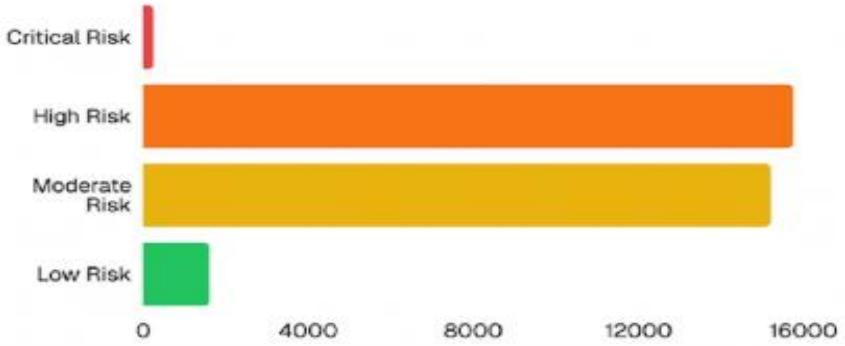
Risk Alert Distribution

32,800 total students



Retention Risk Funnel

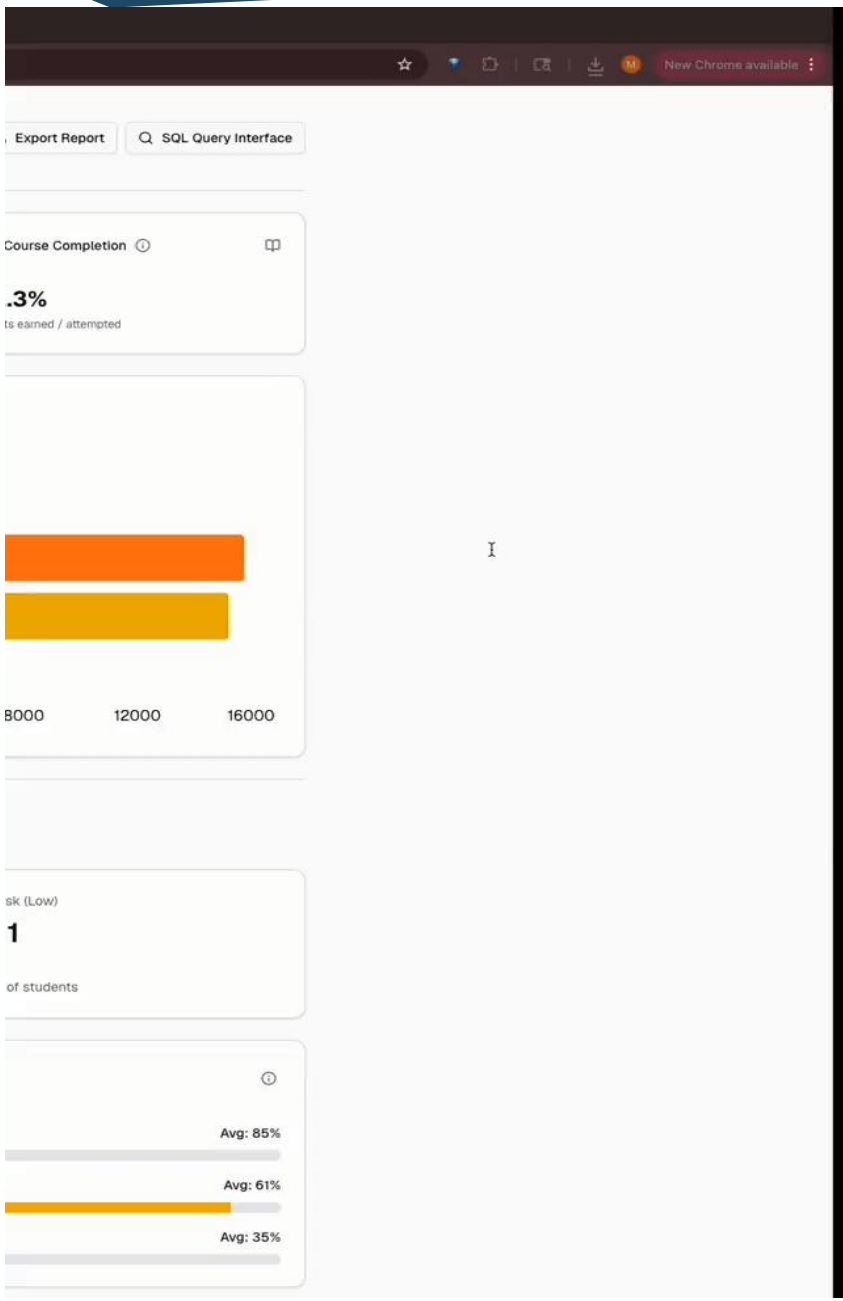
32,800 total students

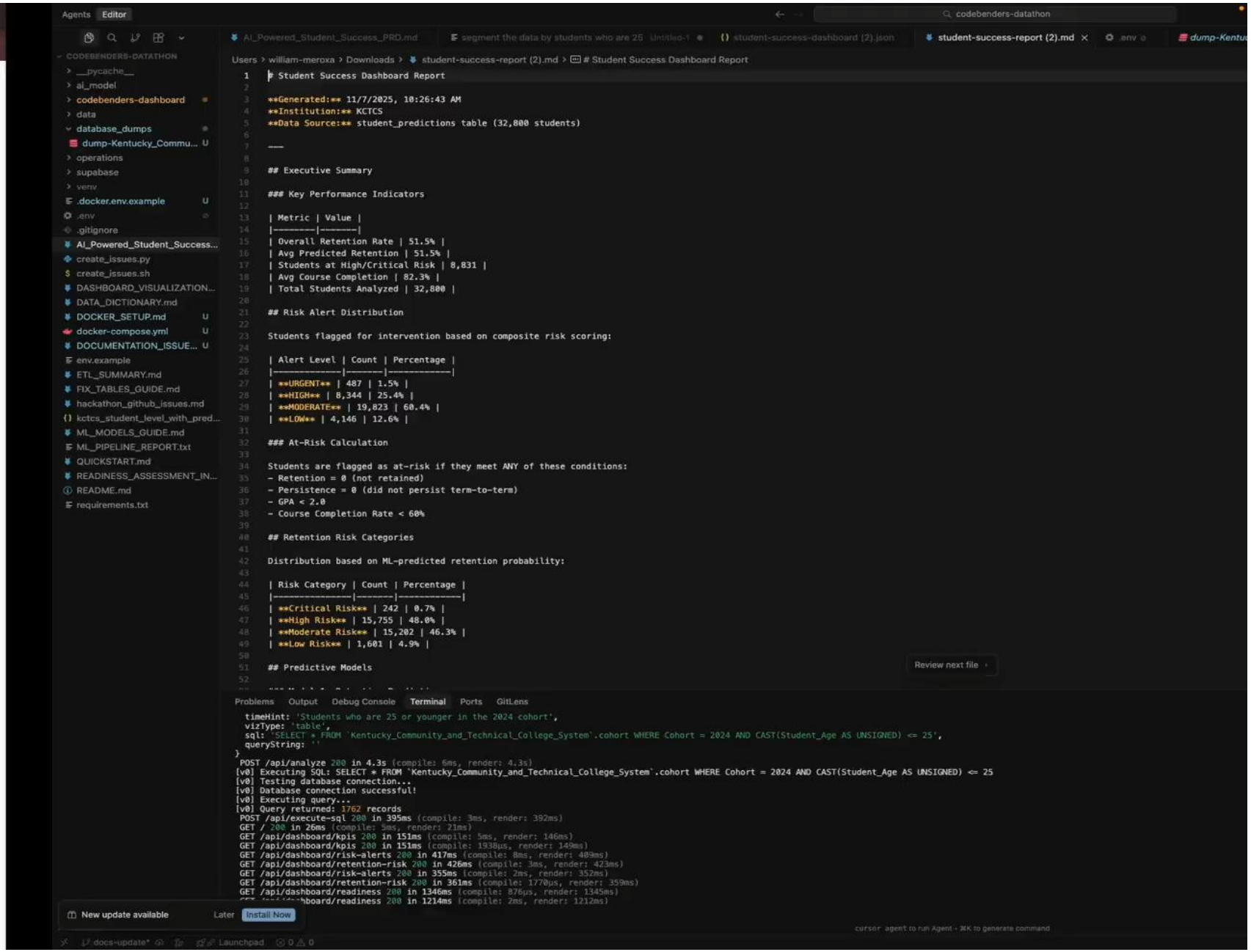


Student Readiness Assessment

AI-powered analysis identifying student preparation levels and intervention needs

Total Assessed	Average Score	High Readiness	At Risk (Low)

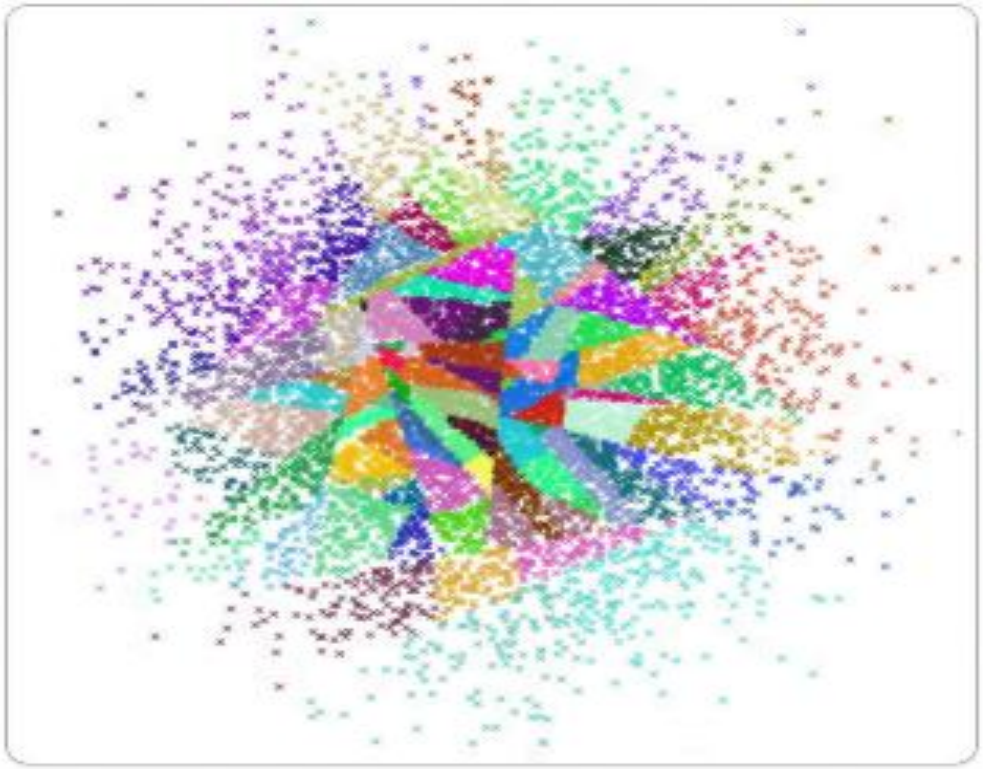

New Chrome available
Export Report
SQL Query Interface
Course Completion
.3%
ts earned / attempted
8000
12000
16000
sk (Low)
1
of students
Avg: 85%
Avg: 61%
Avg: 35%


Agents
Editor
codebenders-datathon
AI_Powered_Student_Success_PRD.md
segment the data by students who are 25
Untitled-1
student-success-dashboard (2).json
student-success-report (2).md
.env
dump-Kentu
Users > william-meroxa > Downloads > student-success-report (2).md > # Student Success Dashboard Report
CODEBENDERS-DATATHON
__pycache__
ai_model
codebenders-dashboard
data
database_dumps
dump-Kentucky_Commu...
operations
supabase
venv
.docker.env.example
.env
.gitignore
AI_Powered_Student_Success...
create_issues.py
create_issues.sh
DASHBOARD_VISUALIZATION...
DATA_DICTIONARY.md
DOCKER_SETUP.md
docker-compose.yml
DOCUMENTATION_ISSUE...
env.example
ETL_SUMMARY.md
FIX_TABLES_GUIDE.md
hackathon_github_issues.md
kctcs_student_level_with_pred...
ML_MODELS_GUIDE.md
ML_PIPELINE_REPORT.txt
QUICKSTART.md
README_ASSESSMENT_IN...
README.md
requirements.txt
Student Success Dashboard Report
Generated: 11/7/2025, 10:26:43 AM
Institution: KCTCS
Data Source: student_predictions table (32,800 students)
Executive Summary
Key Performance Indicators
Metric	Value
Overall Retention Rate	51.5%
Avg Predicted Retention	51.5%
Students at High/Critical Risk	8,831
Avg Course Completion	82.3%
Total Students Analyzed	32,800
Risk Alert Distribution
Students flagged for intervention based on composite risk scoring:
Alert Level	Count	Percentage
URGENT	487	1.5%
HIGH	8,344	25.4%
MODERATE	19,823	60.4%
LOW	4,146	12.6%
At-Risk Calculation
Students are flagged as at-risk if they meet ANY of these conditions:
- Retention = 0 (not retained)
- Persistence = 0 (did not persist term-to-term)
- GPA < 2.0
- Course Completion Rate < 60%
Retention Risk Categories
Distribution based on ML-predicted retention probability:
Risk Category	Count	Percentage
Critical Risk	242	0.7%
High Risk	15,755	48.0%
Moderate Risk	15,202	46.3%
Low Risk	1,601	4.9%
Predictive Models
Review next file
Problems
Output
Debug Console
Terminal
Ports
GitLens
timeHint: 'Students who are 25 or younger in the 2024 cohort',
vizType: 'table',
sql: 'SELECT * FROM `Kentucky_Community_and_Technical_College_System`.cohort WHERE Cohort = 2024 AND CAST(Student_Age AS UNSIGNED) <= 25',
queryString: ''
POST /api/analyze 200 in 4.3s (compile: 6ms, render: 4.3s)
[v0] Executing SQL: SELECT * FROM `Kentucky_Community_and_Technical_College_System`.cohort WHERE Cohort = 2024 AND CAST(Student_Age AS UNSIGNED) <= 25
[v0] Testing database connection...
[v0] Database connection successful!
[v0] Executing query...
[v0] Query returned: 1762 records
POST /api/execute-sql 200 in 395ms (compile: 3ms, render: 392ms)
GET / 200 in 26ms (compile: 5ms, render: 21ms)
GET /api/dashboard/kpis 200 in 151ms (compile: 5ms, render: 146ms)
GET /api/dashboard/kpis 200 in 151ms (compile: 1938µs, render: 149ms)
GET /api/dashboard/risk-alerts 200 in 417ms (compile: 8ms, render: 409ms)
GET /api/dashboard/retention-risk 200 in 426ms (compile: 3ms, render: 423ms)
GET /api/dashboard/risk-alerts 200 in 355ms (compile: 2ms, render: 352ms)
GET /api/dashboard/retention-risk 200 in 361ms (compile: 1770µs, render: 359ms)
GET /api/dashboard/readiness 200 in 1346ms (compile: 876µs, render: 1345ms)
GET /api/dashboard/readiness 200 in 1214ms (compile: 2ms, render: 1212ms)
New update available
Later
Install Now
docs-update*
Launchpad
cursor agent to run Agent • ⌘K to generate command

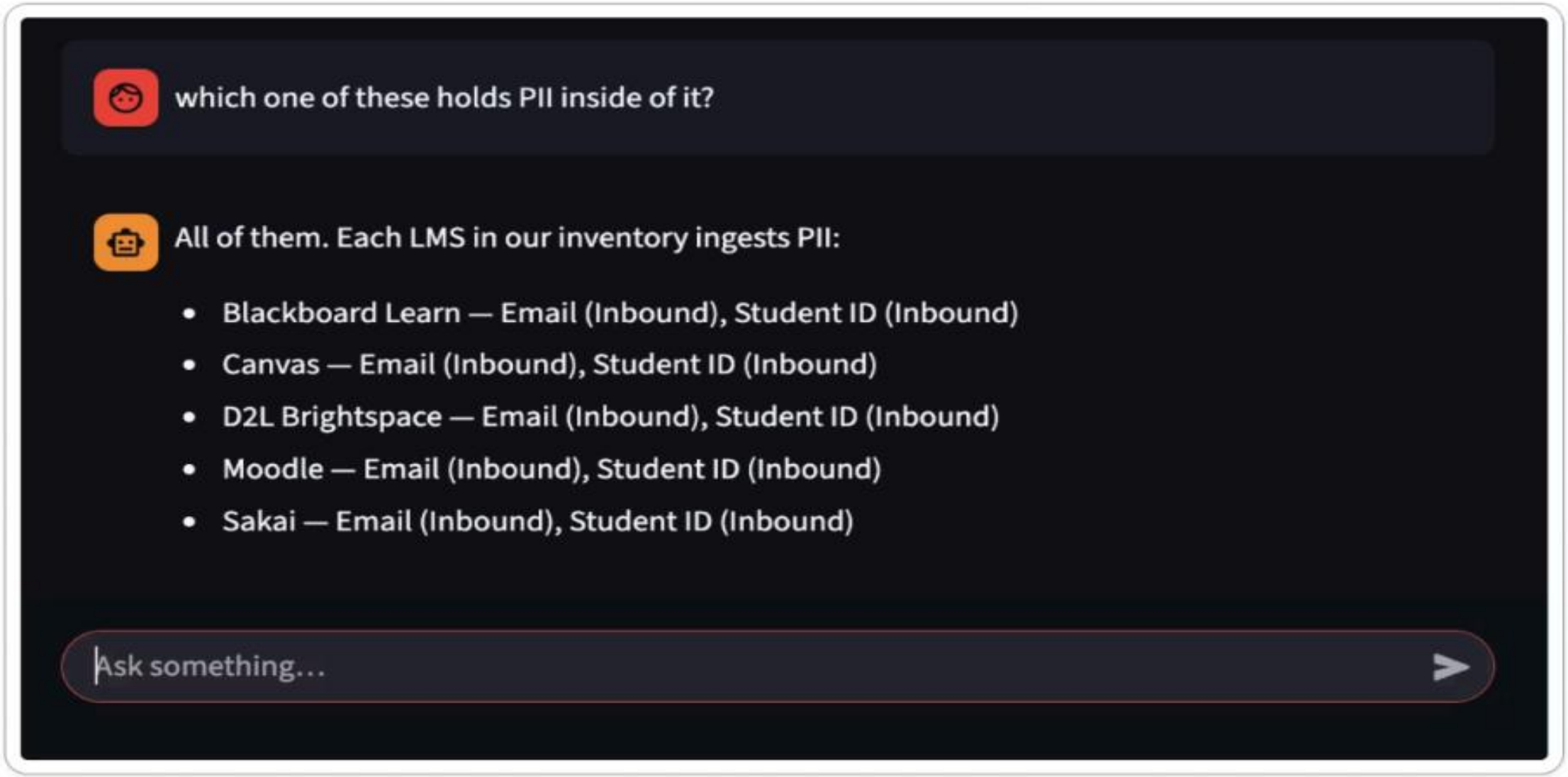
California Institution

- Large institution (30,000+ students)
- Huge technology stack across the entire institution
- Built a chat bot from LLM to be able to ask questions about vendors, data elements, tool purposes, campus owners, etc.

California Institution



Data Model: Relational + Vector Store


which one of these holds PII inside of it?
All of them. Each LMS in our inventory ingests PII:
Blackboard Learn — Email (Inbound), Student ID (Inbound)
Canvas — Email (Inbound), Student ID (Inbound)
D2L Brightspace — Email (Inbound), Student ID (Inbound)
Moodle — Email (Inbound), Student ID (Inbound)
Sakai — Email (Inbound), Student ID (Inbound)
Ask something...

Private KY Institution

Context

Who: End users e.g college advisers student success staff

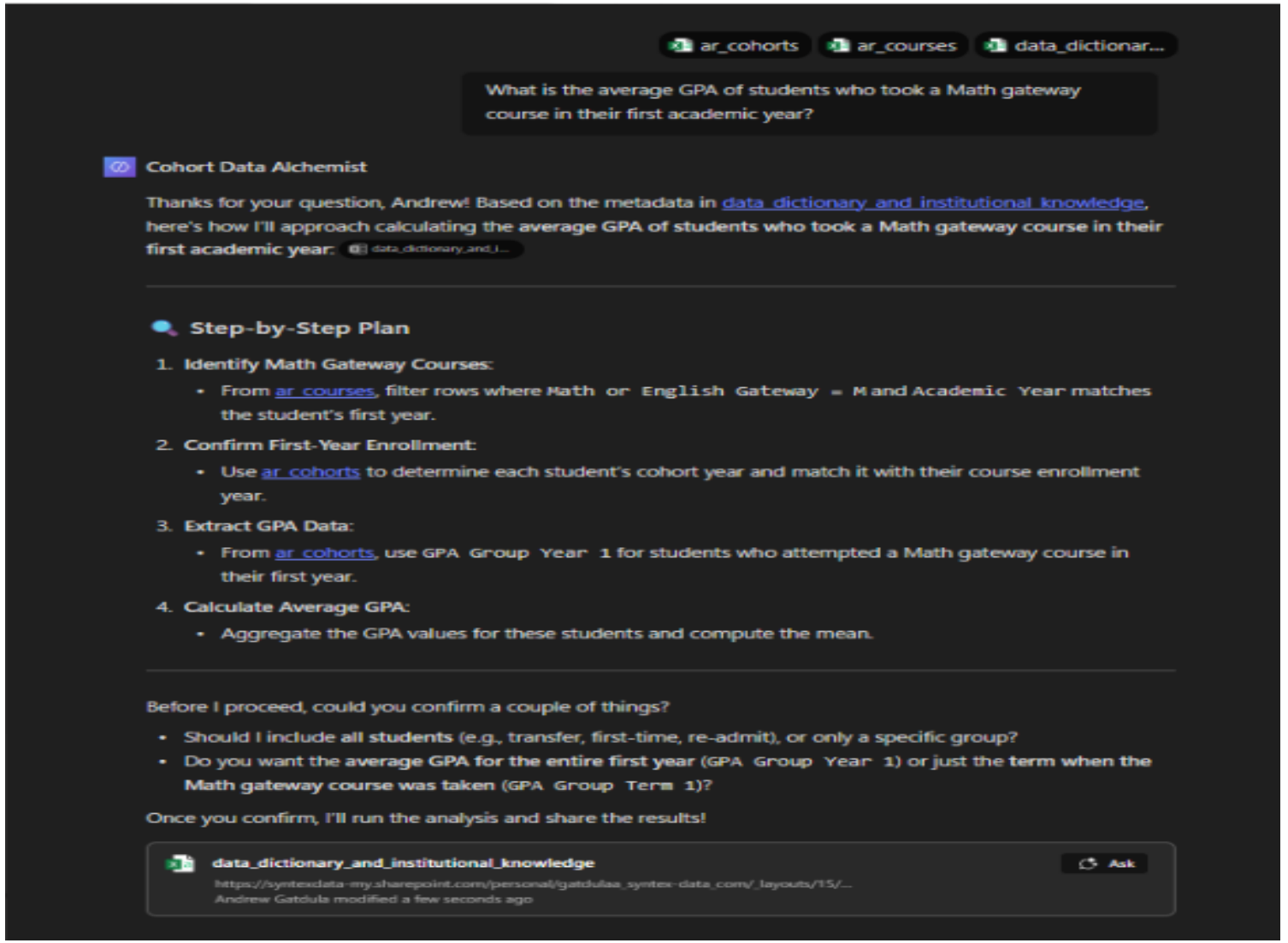
The Challenge: Have some data access but varied technical skills and context to analyze independently

Current Friction:

- Must submit requests to data analysts (creates bottlenecks)
- Clarification takes time
- Outputs might not match expectations
- Data team cycles on repetitive, one-off analyses

Impact: Lost capacity for strategic work; delayed answers for time-sensitive student needs

Private KY Institution



GARDNER INSTITUTE

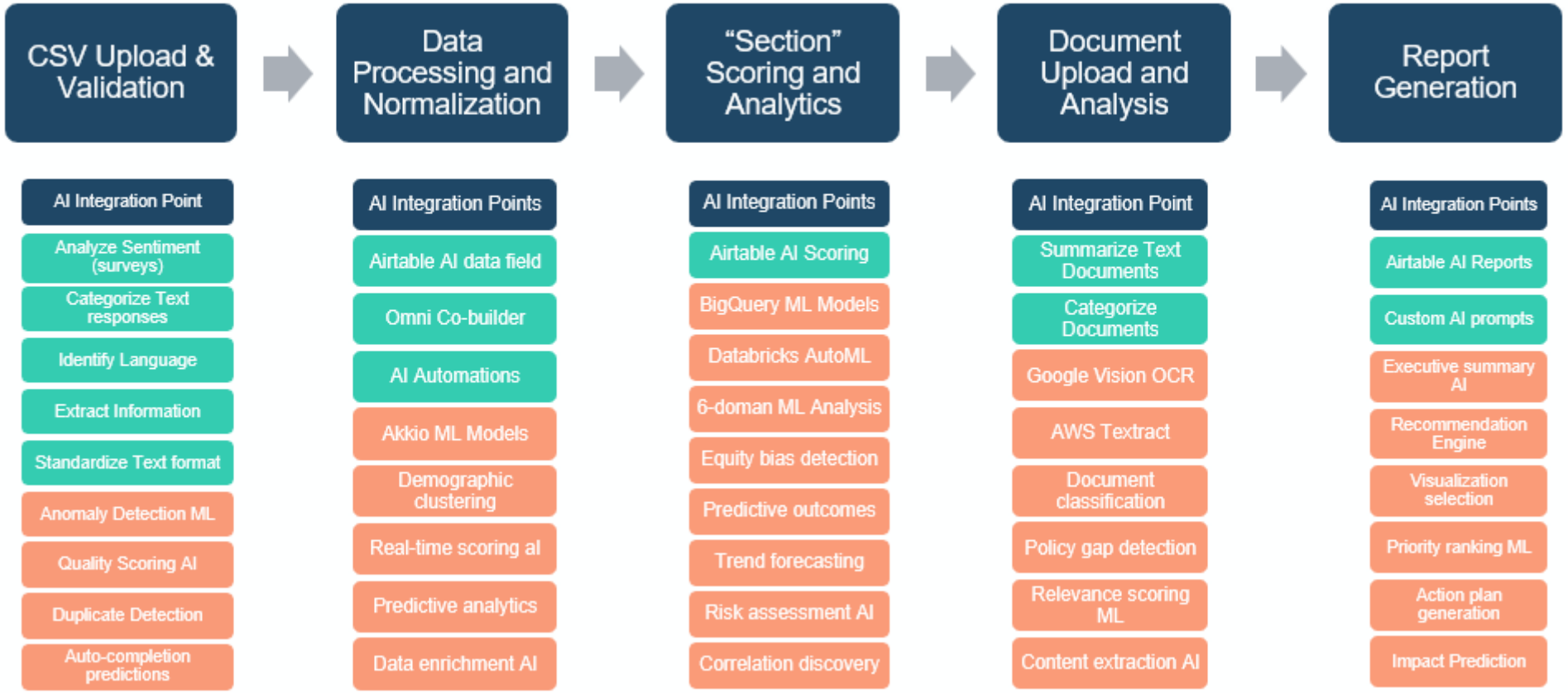
Gardner Institute

- Sensemaking
  - Score two surveys
  - Combine with institutional student data
  - Combine with institutional strategic documents
  - Prepare in a report that use for guided discussion with campus

- Process typically takes 2 staff members about 2 weeks to complete

Gardner Institute

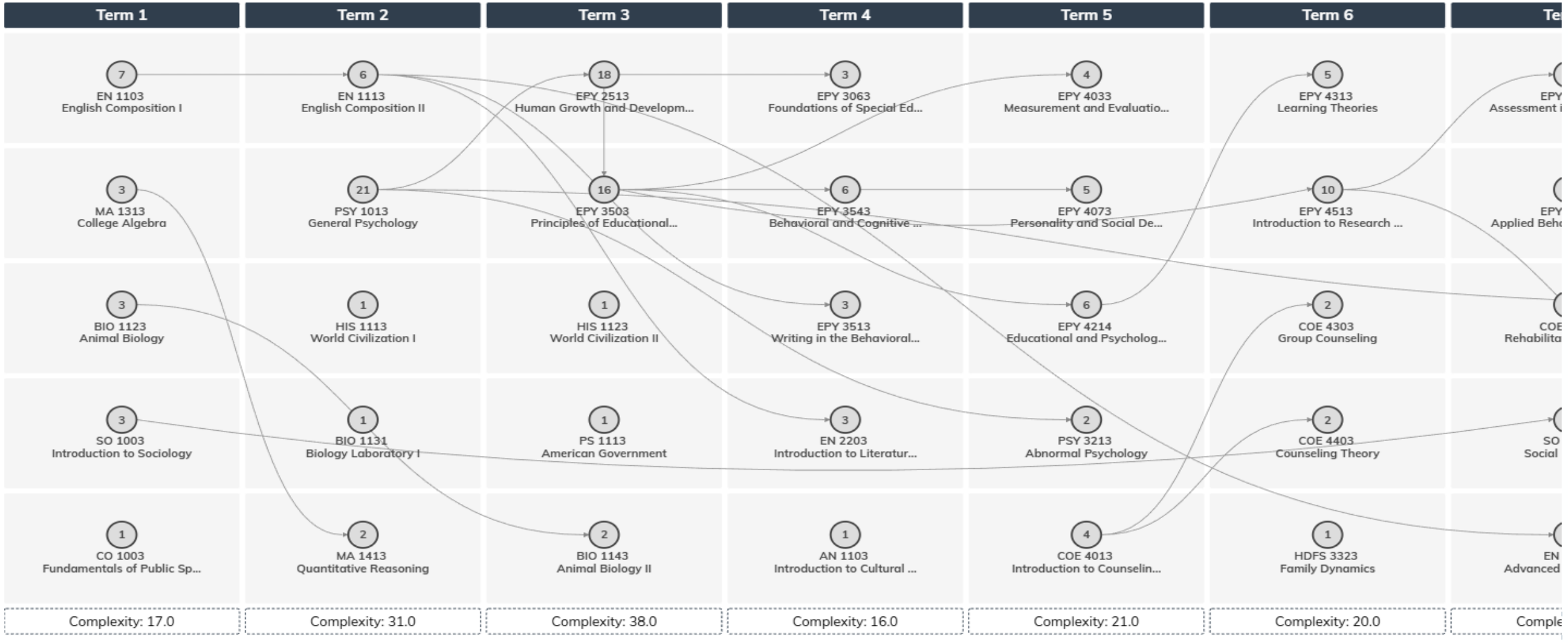
Directions & Scalability: Agentic Survey Workflow Example



GARDNER INSTITUTE

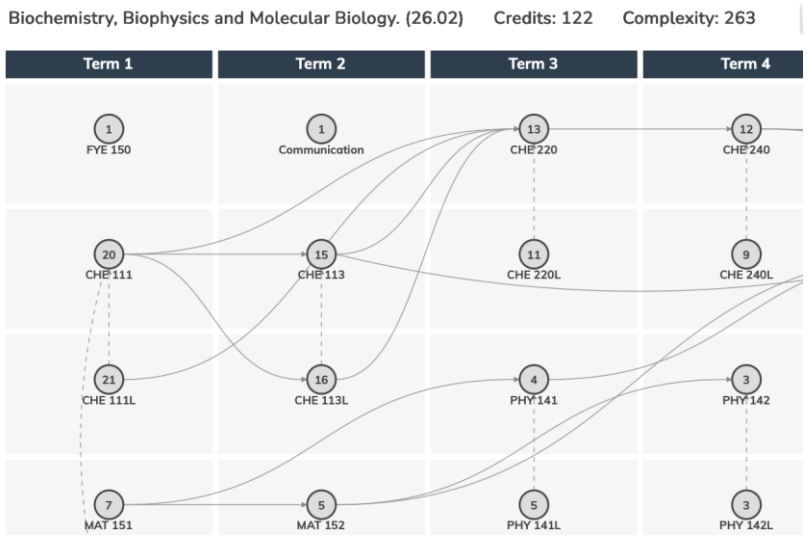
Program Course Maps

Educational Psychology with Counseling/Clinical Applications Concentration Credits: 119 Complexity: 178 Edit Curriculum Create a Copy View in list



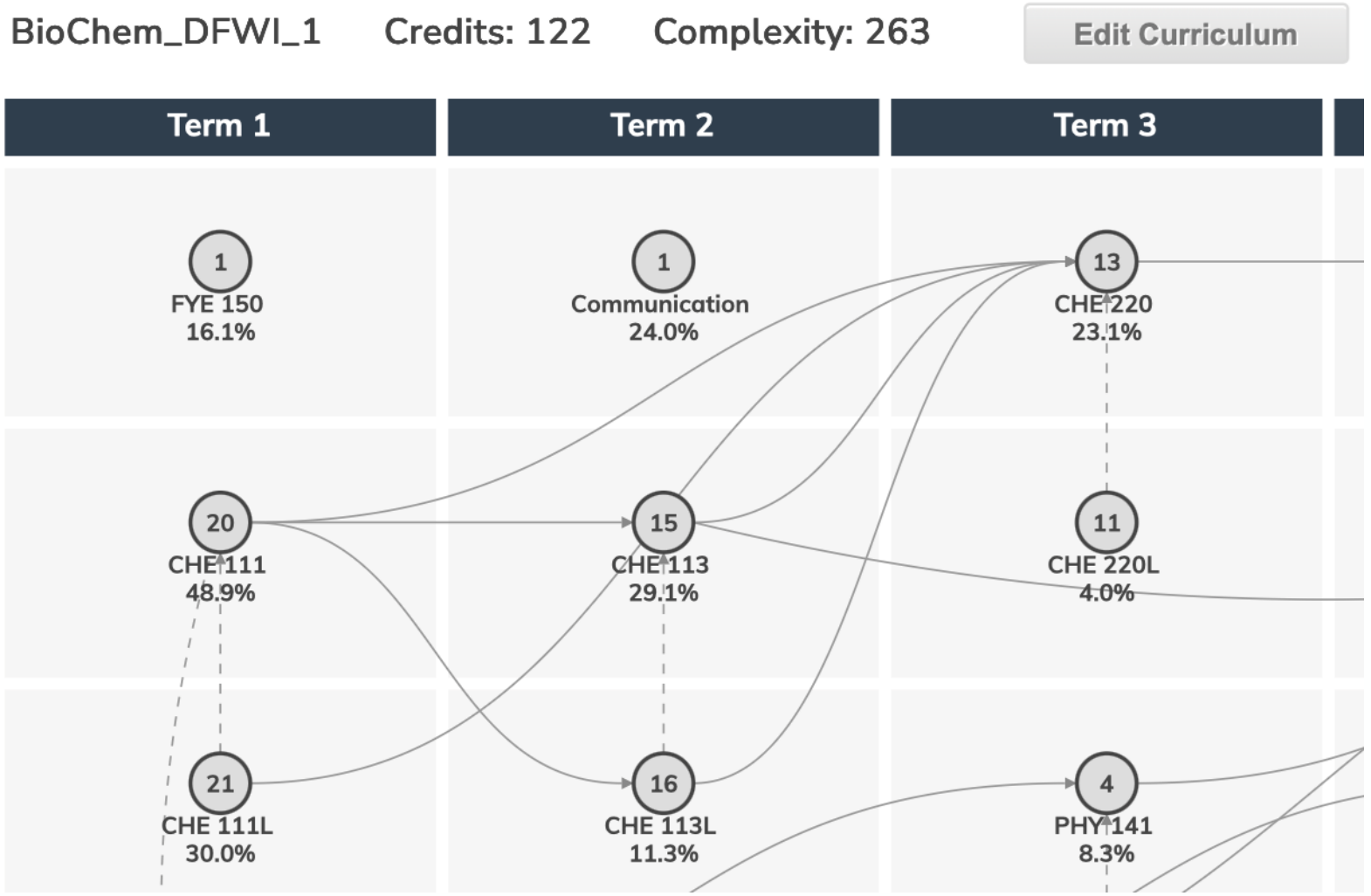
Course Maps + AI

Course	Title	Terms	Sections	Enrollment	D,F,W,I Rate
BIO 330	Biochemistry	1	1	7	00.0 %
BIO 330L	Biochemistry Laboratory	1	1	7	00.0 %
CHE 101	Introduction to Chemistry	1	1	23	56.5 %
CHE 103	Basic Chemistry I	1	3	66	45.5 %
CHE 103L	Basic Chemistry I Laboratory	1	5	59	18.6 %
CHE 104	Basic Chemistry II	1	2	38	18.4 %
CHE 104L	Basic Chemistry II Laboratory	1	3	38	07.9 %
CHE 111	General Chemistry I	1	3	88	48.9 %
CHE 111L	General Chemistry I Laboratory	1	5	80	30.0 %
CHE 113	General Chemistry II	1	2	55	29.1 %
CHE 113L	General Chemistry II Laboratory	1	4	53	11.3 %
CHE 220	Organic Chemistry I	1	1	26	23.1 %
CHE 220L	Organic Chemistry I Laboratory	1	2	25	04.0 %



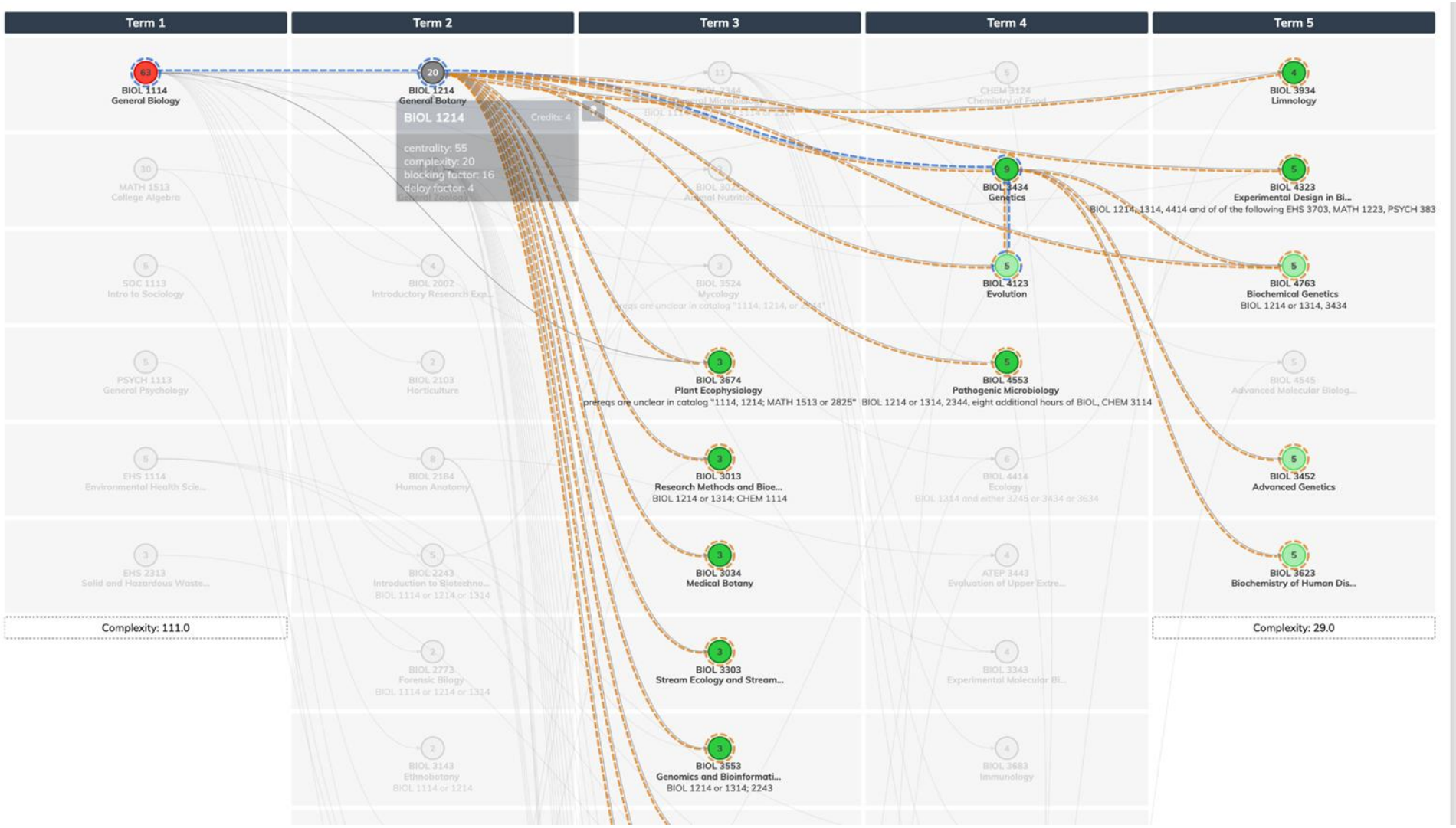
GARDNER INSTITUTE

Program Course Maps + AI



GARDNER INSTITUTE

Gateway Course Maps



GARDNER INSTITUTE

Discussion & Questions

GARDNER INSTITUTE

Contact Information

- **Brent Drake**

drake@gardnerinstitute.org

GARDNER INSTITUTE

Next Meeting

- 3rd Quarter: August 26 - Could this be YOU?

Announcements

- Datatelligent will be at: (come by and say 'hi')
  - Higher Ed Data World – April 26-29
  - AIR Forum – May 26-29
- We will soon announce an in-person DAAHE meeting in the Chicagoland area this summer. Be on the lookout for details.
- New Podcasts every other Wednesday
- Follow us on LinkedIn
  - Datatelligent & Deborah A. Phelps

Thank you for attending!

If you have feedback or suggestions,
please let us know or email dphelps@datatelligent.ai

Follow-ups

- Everyone will receive a link to the recording, the presentation, and contact information for attendees.
- If you do not want your contact information shared, please let us know.
- Look for your monthly DAAHE newsletter in your inbox.