

ANALYTICS COMMUNITY OF PRACTICE · FALL 2026 KICKOFF

Using AI & analytics to curate myUMBC's first annual report

Collier Jones · Campus Portal Architect, DoIT

Thursday, September 24, 2026 · Engineering 102 and online

TODAY

Three questions

01

What did AI make possible?

Why a report like this never existed, and how it finally got built.

02

Where did the numbers lie?

The pitfalls, and the Blackboard story.

03

What did the report show?

Everyone's slice, the whole picture, and what data can't tell us.

WHAT AI MADE POSSIBLE

Why myUMBC never had an annual report

2

databases on different engines — MySQL and PostgreSQL

475M

rows of activity logs across the two systems

78

tables, many with retired or look-alike columns

1

shared key to connect a person across systems

Every number meant hand-written SQL on a different engine, a role model that differs by system, and a spreadsheet to reconcile it all. So questions got asked one at a time — only when someone needed the answer.

WHAT CAME OUT

myUMBC Annual Report, 2025–2026

01 Access & Logins

02 Page Views

03 Tools & Resources

04 Search

05 Community

06 Group Email

07 Messaging

08 Profile Sharing

09 Guest Experience

10 Sites

43

charts and tables

+ a **methodology** with every query, so it can be rebuilt

+ a **Markdown version** written for AI agents to read

my.umbc.edu/annual-report-2025-2026.html

What changed: the cost of asking a question

Before

- One question, one hand-written query
- Export, reconcile roles, chart it by hand
- "What if" questions rarely got asked
- The method lived in my head

With AI

- Ask in plain English; queries run on live, read-only data
- Results, charts and prose in one conversation
- Follow-ups are cheap, so curiosity gets followed
- Every query captured for next year

WHAT AI MADE POSSIBLE

How it was built

Data dictionary

Standing context in a Claude Project: schemas, enums, gotchas.



Claude

Writes SQL and runs it over read-only MCP connections to both databases.



The report

One self-contained web page. Ten sections, 40+ charts.



Methodology

Every query, caveat and rule, so it can be rebuilt.



Me, throughout: steer the questions · challenge the numbers · decide what's publishable

MY SETUP

The “myUMBC Data” Claude Project

Instructions

“Do not guess at schema – look it up.”

- Which database is which, and which to ignore
- Query safety: IDs, time zones, system accounts
- How to scope the big log tables
- MySQL gotchas
- How I like answers written

Knowledge

The data dictionary, as a project file

- System overview and ID crosswalk
- Every table: columns, enums, rough counts
- Index strategy for the big logs
- Business logic and background jobs
- Retired tables and a corrections log

Tools

MCP connections to both databases

- One for MySQL, one for PostgreSQL
- Run SQL and list tables
- Reader accounts, plus a scratch area for staging
- The report conversation lives in the project too

The reusable part isn’t a prompt. It’s the context: written down once, used in every conversation.

What the schema won't tell you

Names drift

UI "Posts" are news; "Files" are group_documents.

Noise to filter

Paw values -2, -3 and -110 are legacy junk. Count only value = 1.

Who to exclude

System and guest accounts: two IDs in one system, four UUIDs in the other.

Joins that lie

Answer logs point to the answer event, not the query. The obvious join is all NULL.

Where history ends

The 328M-row activity log starts in 2017. Earlier dates return nothing, silently.

Who really did it

When staff proxy, content shows the proxied user. The real one is only in the logs.

Status isn't kind

"Retired" is a group status, not a type. Groups have exactly three kinds.

Some text is sealed

Search answers are encrypted at rest. Raw SQL returns ciphertext.

Speed traps

The biggest log has no date-only index. Lead with the action, or wait.

My job moved from writing queries to **interrogating answers.**

The AI was fast and fluent. It was also confidently wrong in exactly the places that matter.

PITFALLS

Rule one: one person, one role



People hold many roles at once — a staff member can also be an alum, an advisor and a grad student. For reporting, each person counts once, under their highest role.

“Other” is real too: **86K logins from 12,796 people** — mostly alumni, plus applicants, delegates and continuing-ed students.

PITFALLS

The join that quietly multiplies

Join memberships to roles and you get one row per role. Count rows, and someone with three roles counts three times.

Faculty hold the most roles, so they inflated the most.

Unique-person counts stayed right — so nothing *looked* broken.

Metric	First draft	Verified
Voluntary join events	18,004	16,802
Faculty voluntary joins	2,990	2,039
Avg groups per faculty member	13.0	9.1
Auto-enrollment events	58,262	55,981

✔ The check: role splits must sum to the un-split total.

PITFALLS

Plausible numbers are the dangerous ones

UNIQUE SEARCHERS

~~49,446~~

19,474

Monthly uniques were summed — one person counted every month they searched.

UNDERGRAD LOGINS

~~1.38M~~

1.83M

The draft never pulled from the official login view. One re-query fixed it.

GO-LINK "TOTAL"

~~4.37M~~

4.73M

The "total" left out the Other role, so every share was off — undergrads 62% became 57%.

Every one of those errors was written by the AI — and caught by the AI, **once I asked for a line-by-line check against live data.**

Role splits sum to the total

Never sum monthly uniques

A “total” means everyone

WHEN THE HEADLINE LIES

Blackboard: the #1 search term in myUMBC

~16,500

searches for Blackboard

6×

the next most-searched term

~10%

of all search volume



The obvious conclusion: people can't find Blackboard.

WHEN THE HEADLINE LIES

Then count people, not searches

~1,300

people searched for Blackboard

~2%

of the 58,449 people who logged in

1

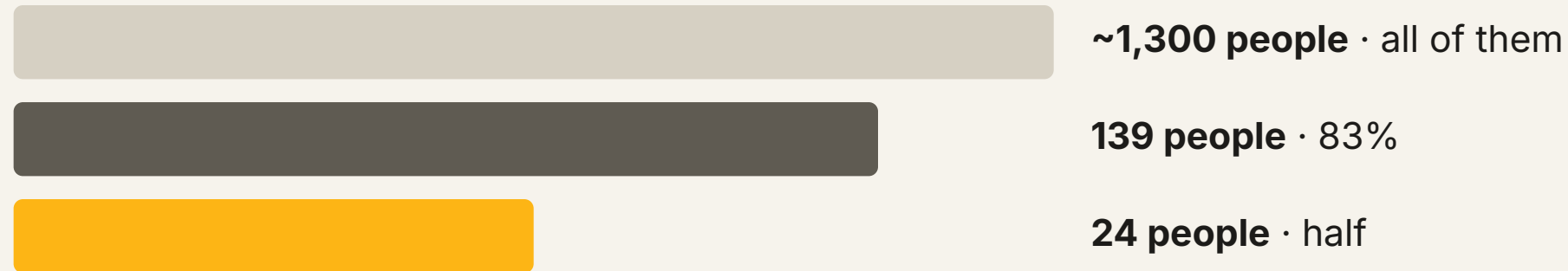
search all year for the typical one

So how do ~1,300 people, most searching once, add up to 16,500 searches?

WHEN THE HEADLINE LIES

Drilling down: who makes the number?

SHARE OF ALL BLACKBOARD SEARCHES



Those 24 searched on ~146 days each, about twice a day. Everyone else averaged two Blackboard searches all year.

A few dozen daily searchers make the number. Everyone else mixes search, links and the tray: **using the tools, not struggling.**

HOW IT WAS FOUND

One question, 16 queries, 2 databases

ME

“Could you profile the users doing the search?”

AI, A FEW MINUTES LATER

A few dozen people made most of the searches. For them, Blackboard is most of what they search, and they click the link directly too.

WHAT THE AI RAN

- 1 Staged the year's searches in a working table
- 3 Grouped spelling variants of "blackboard"
- 1 Split searches by device from browser info
- 3 Built a per-person profile, grouped by how often they searched
- 1 Added role and class year
- 7 Crossed to the other database by login to count their Blackboard link clicks
- 12 PostgreSQL + 4 MySQL · production read-only, working tables in a scratch area

I wrote no SQL. My job was noticing that the average didn't smell right, and checking what came back.

WHEN THE HEADLINE LIES

Same data, two audiences

IN THE PUBLIC REPORT

A neutral fact, with context

- One sentence and a short note
- Counted in every total
- Shown separately in the themes chart so it doesn't swamp everything else

FOR THE PRODUCT TEAM

A preference to support

- Some people would rather type than click
- Could the "blackboard" result appear instantly?
- The link works: ~48 clicks for every search
- Two-thirds of these searches come from desktop, where the link is in plain view

WHEN THE HEADLINE LIES

More headlines that mislead on their own

Headline	Tempting read	What we think is going on	What we'll look into
Staff search 3× as often as undergrads	Staff are lost	Niche needs make search the faster path; homepage resources still work for them	Do heavy staff searchers also click trending resources and links, or only search?
Staff and faculty belong to 9–13 groups each	They're the most engaged	Faculty and staff accumulate groups over years; most students focus on classes, a few are heavily involved	When were memberships created, and how are students spread from none to many?
38M guest page views	The public loves the portal	Mostly search engines landing on public posts and events	Compare guest and logged-in views of the same pages: same content, or different kinds?

WHEN QUESTIONS ARE CHEAP

Small things you only find by asking

912

undergrads searched "1098," the tuition tax form. Twice as many as searched for grades.

27

people asked search "what is my major." 183 asked "what is my GPA."

451

faculty searches for the two letters "sa." All from just 3 people.

700+

misspelled Blackboard searches. 8 were "blackboardblackboard."

Rules for a public report

No system names

Readers see myUMBC — not databases, tables or my3/my4.

No unverifiable comparisons

Year-over-year only where last year's number could be re-checked.

No guessed causes

An odd month is reported as odd, not explained with a hunch.

Prose interprets

Text says what the chart can't. It never restates the table.

Accessible by default

Gold text fails contrast on white, so roles get a color swatch, not colored type.

Keep the internal internal

Product findings go to the team that owns them, not the report.

PART THREE

So what did the report show?

Mostly what we already knew. The value is letting everyone else see it too.

WHAT THE REPORT SHOWED

Everyone sees a different myUMBC



STUDENTS

A launch pad, and a window into their info

Registration, billing, degree progress, their profile



FACULTY

A launch pad to the tools they need

Gmail, teaching schedule, Blackboard, advising



STAFF

A launch pad, and a place to publish

Timesheets, HR, finance, and more group posts than any other role

Each group sees its own slice. Few people see all of it.

WHAT THE REPORT SHOWED

The breadth most people never see

58,449

people logged in, 2.92M times

9.87M

group emails sent through myUMBC

814K

office messages to 28,979 people; about 3 in 4 opened in email

14–24K

people logged in every week of fall, lowest at Thanksgiving, highest at finals

16,802

voluntary group joins, on top of 55,981 automatic enrollments

~²/₃

of group posts written by faculty and staff

WHAT'S NEXT · A PREVIEW

Now we can see what gets clicked

7×

Recommended Resources drew 457K clicks from 16K people, seven times the promo image tiles

28%

of Pawpular clicks go to the #1 slot, 2.4× the #2 slot

11×

more clicks on a post's title than on its thumbnail

7%

of visits that scroll a quarter of the way down ever reach the bottom of the homepage

Item-level homepage click tracking since March 5, 2026 · scroll depth since July 1, 2026 · next year's report gets a full year of both

WHAT THE DATA CAN'T TELL US

Usage data is only part of the picture

Clicks aren't engagement

A promo tile builds awareness just by where it sits. Fewer clicks doesn't mean it isn't working, and no log captures that.

An incomplete picture misleads

Decisions about what to change or cut need the whole picture, not just what's easy to count.

The rest of the picture

Surveys, focus groups and conversations fill in what usage data can't.

The data shows what people do. **It can't show what they notice, value or need.**

THE ANSWER

What AI let me do that I couldn't before

Ground it first

The data dictionary mattered more than any prompt. The traps matter more than the tables.

Follow every question

Not just the ones worth a day of SQL. The Blackboard drill-down only happened because follow-ups were cheap.

Verify everything

Re-checking every number against live data became cheap enough to actually do.

Make it repeatable

The methodology was a by-product, not a second project. Every correction became a written rule.

What it didn't replace: knowing the data well enough to smell a wrong number.

NEW FOR THIS GROUP

Explore the data yourself



Views

Weekly views and users of Home, Profile, Guide and Events, by role, with semester starts and breaks marked

my.umbc.edu/stats/views



Clicks

What people actually click, by section and role. Drill into any section to see its individual links

my.umbc.edu/stats/clicks



Groups

Content statistics for each group: an overview of what every group publishes

my.umbc.edu/stats/groups

Updated nightly · Views and Clicks are open to anyone; Groups requires a myUMBC login

Questions, and your hunches

What would you want to see in next year's report?

Read the report: my.umbc.edu/annual-report-2025-2026.html

Collier Jones · Campus Portal Architect, DoIT