

Data Engineer

VS

Software Engineer

Same Same, But Different





Tomas Peluritis

Head of Data @ Mediatech

AKA Uncle Data



<https://www.linkedin.com/in/tomaspeluritis/>



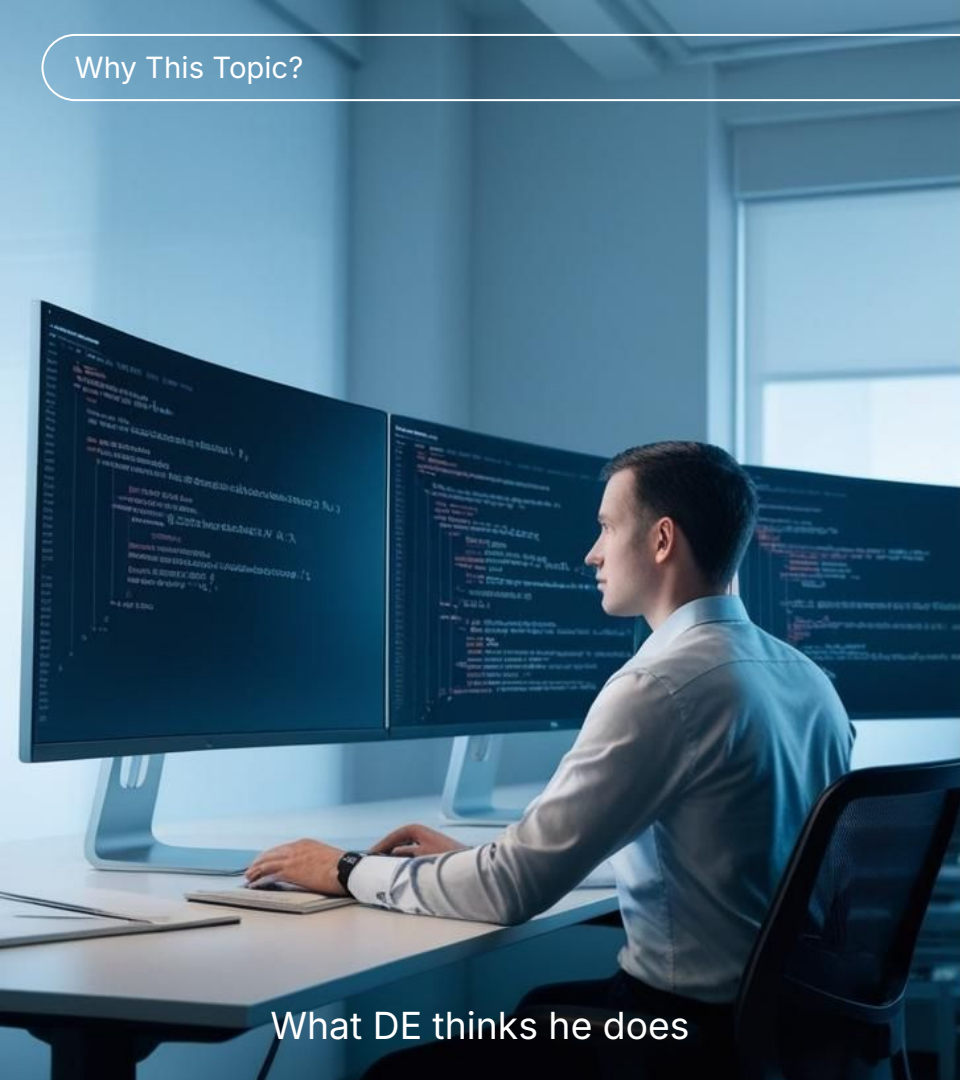
<https://podcasters.spotify.com/pod/show/duomenu-dede>



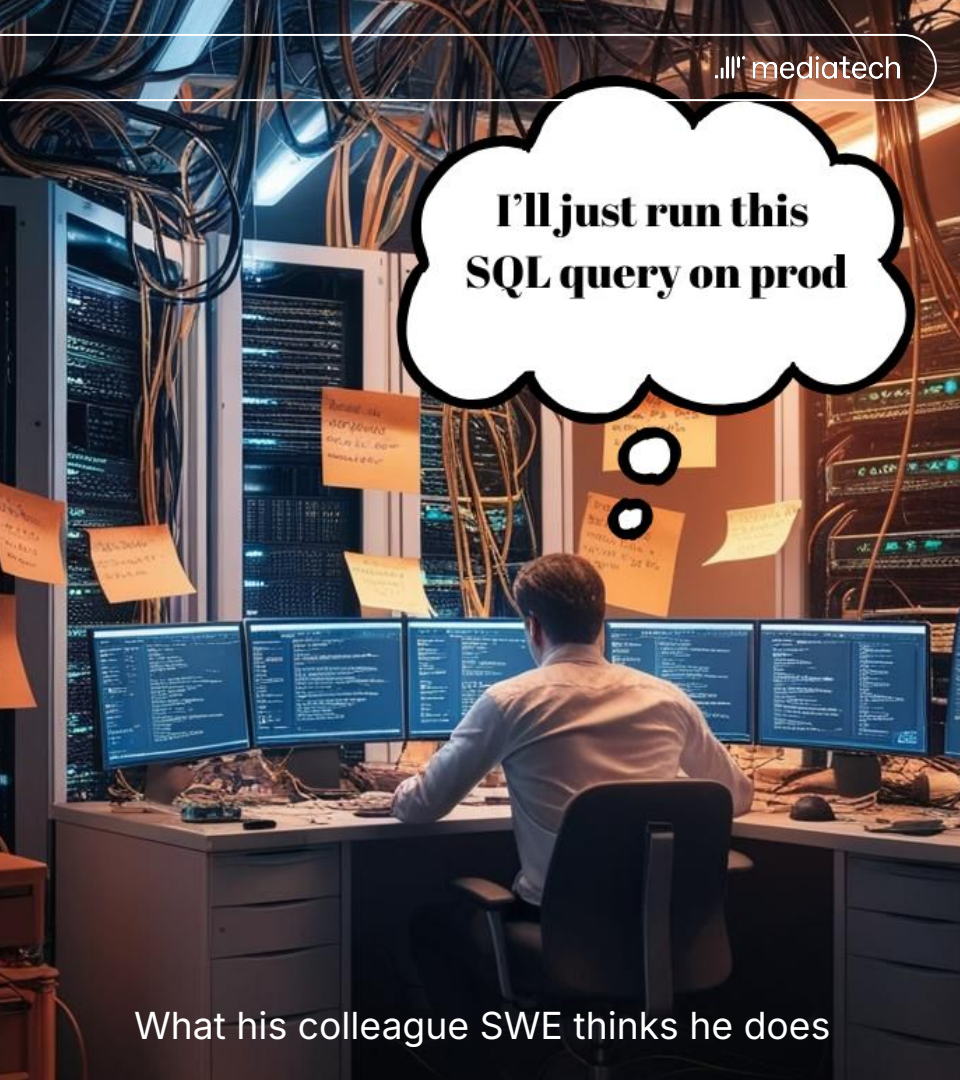
<https://uncledata.substack.com>

Why This Topic?





What DE thinks he does



**I'll just run this
SQL query on prod**

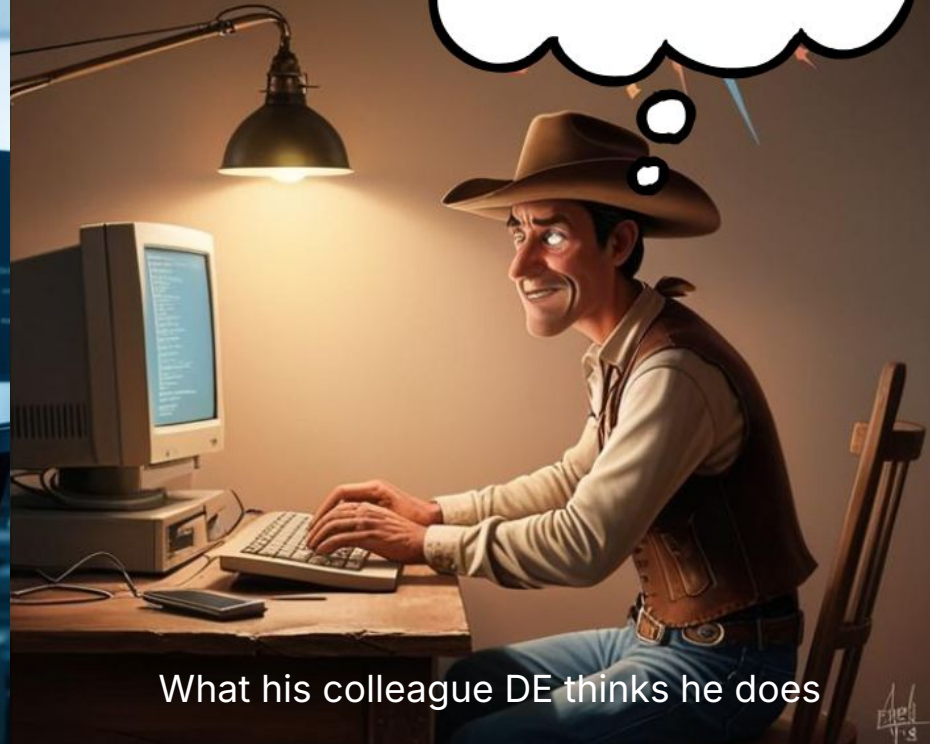
What his colleague SWE thinks he does

Why This Topic?

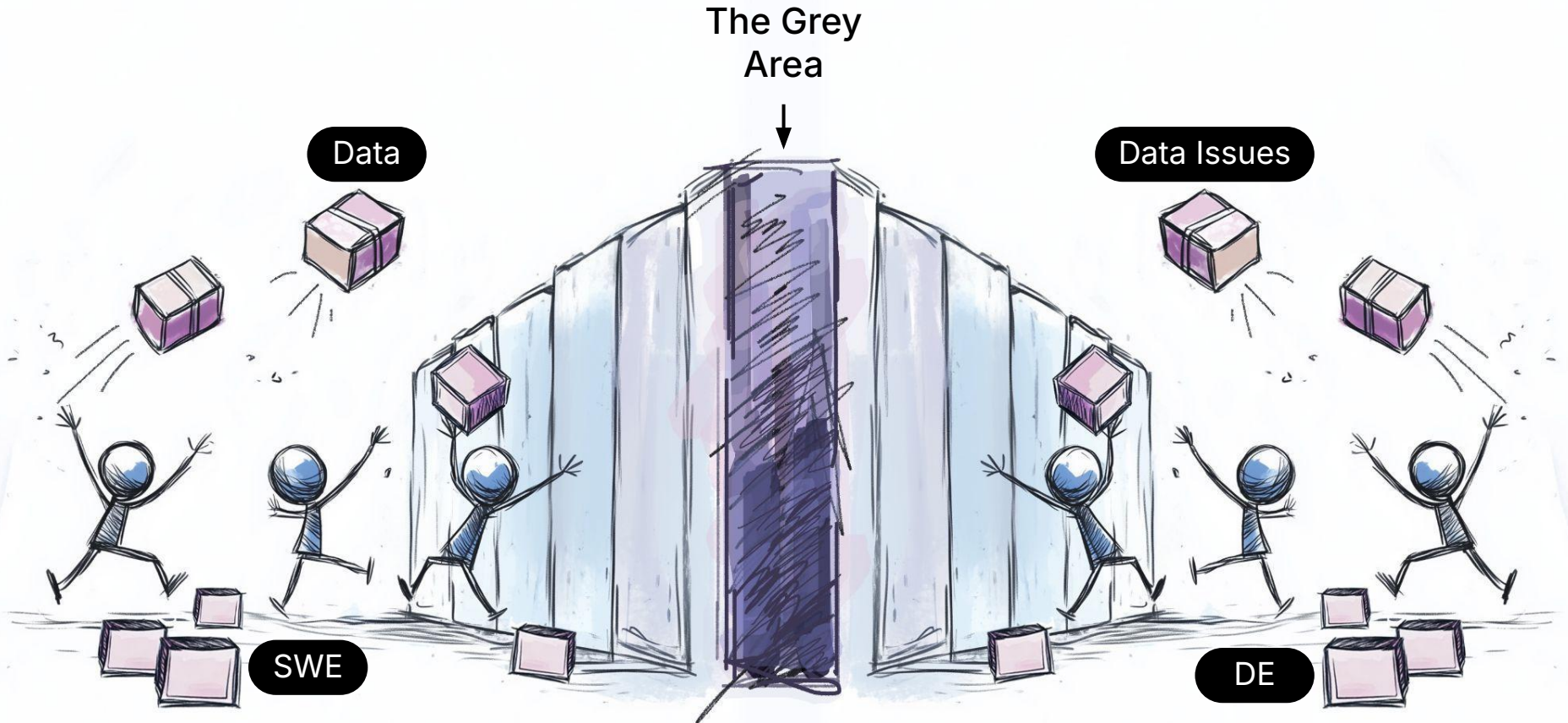
mediattech



What SWE thinks he does



What his colleague DE thinks he does



The DNA Difference



Our Engineering DNA



—>

Frontend:

How will user interact with this?

—>

Backend:

How will this system function and scale?

—>

Data:

How will data flow and transform?

When "Fast" Means Different Things

Render time



RPM



Data Volume



SQL or .py?

SQL or .py?

Yes.

SQL or .py?

Yes.

- Data folks love SQL;

SQL or .py?

Yes.

- Data folks love SQL;
SWE folks tolerate it (barely)

SQL or .py?

Yes.

- Data folks love SQL;
SWE folks tolerate it (barely)
- Both use Python—but for wildly
different things

SQL or .py?

Yes.

- Data folks love SQL;
SWE folks tolerate it (barely)
- Both use Python—but for wildly different things
- Data: Spark, dbt, Airflow
SWE: Flask, FastAPI, React

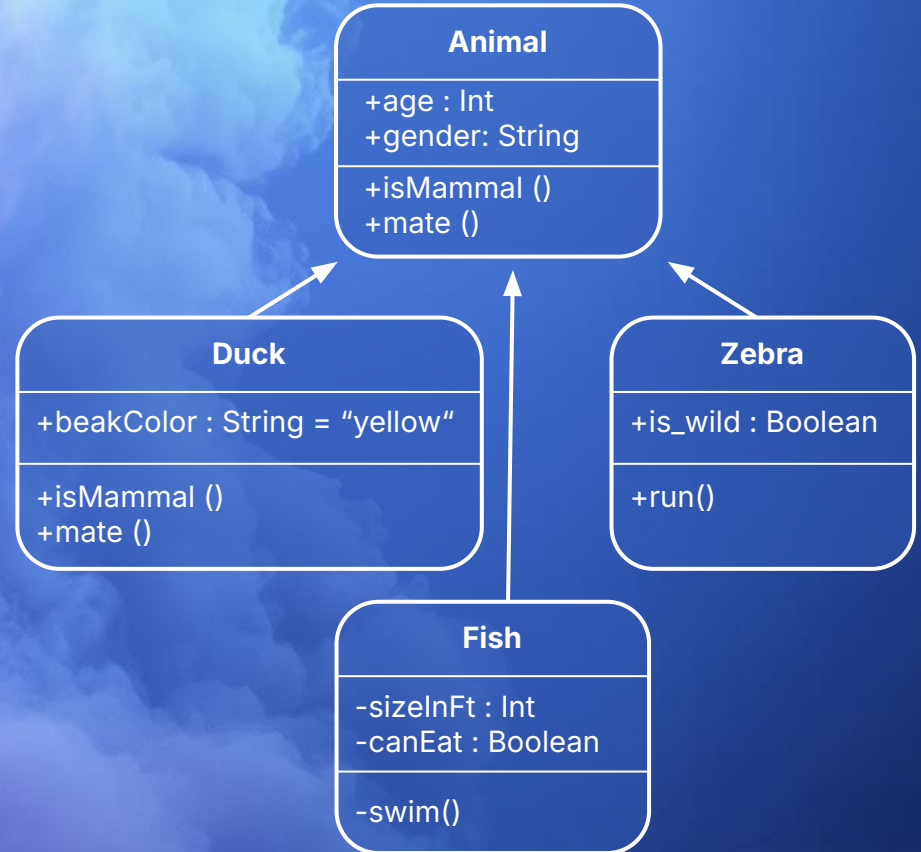
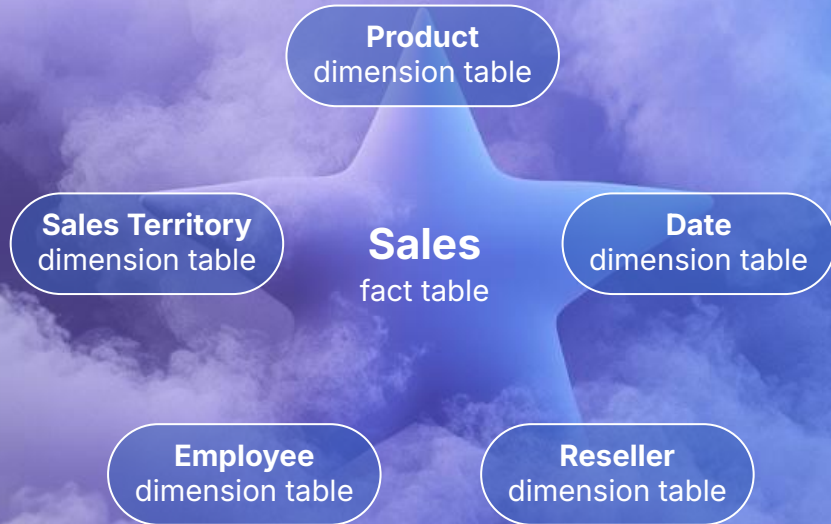
my linkedin profile

R, python, javascript, shiny, dplyr, purrr, ditto, ggplot, d3, canvas, spark, sawk, pyspark, sparklyR, lodash, lazy, bootstrap, jupyter, vulpix, git, flask, numpy, pandas, feebas, scikit, pgm, bayes, h2o.ai, sparkling-water, tensorflow, keras, onyx, ekans, hadoop, scala, unity, metapod, gc, c#/c++, krebases, neo4j, hadoop.

I typically ask recruiters to point out which of these are pokemon.



Data Models vs. Mental Models



Why Are You Storing a Whole JSON Object in There?!



Why Are You Storing a Whole JSON Object in There?!

- Data engineers: this breaks everything



Why Are You Storing a Whole JSON Object in There?!

- Data engineers: this breaks everything
 - Schemas



Why Are You Storing a Whole JSON Object in There?!

- Data engineers: this breaks everything
 - Schemas
 - Queries



Why Are You Storing a Whole JSON Object in There?!

- Data engineers: this breaks everything
 - Schemas
 - Queries
 - Dreams



Why Are You Storing a Whole JSON Object in There?!

- Data engineers: this breaks everything
 - Schemas
 - Queries
 - Dreams
- Software engineers: it works and it ships



Why Are You Storing a Whole JSON Object in There?!

- Data engineers: this breaks everything
 - Schemas
 - Queries
 - Dreams
- Software engineers: it works and it ships
- True story: had to parse a nested array with 5+ levels



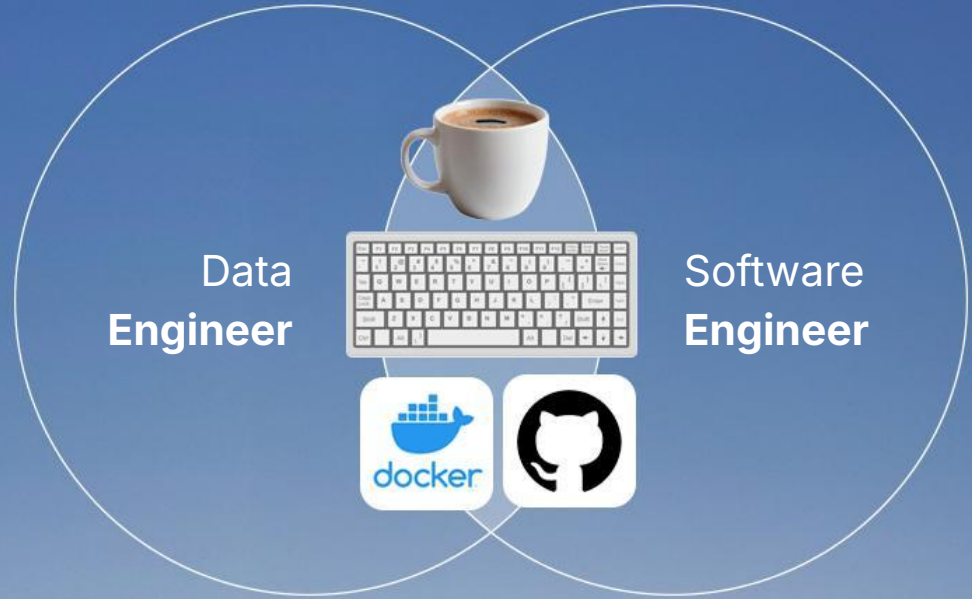
When Different Worlds Meet



Core Similarities



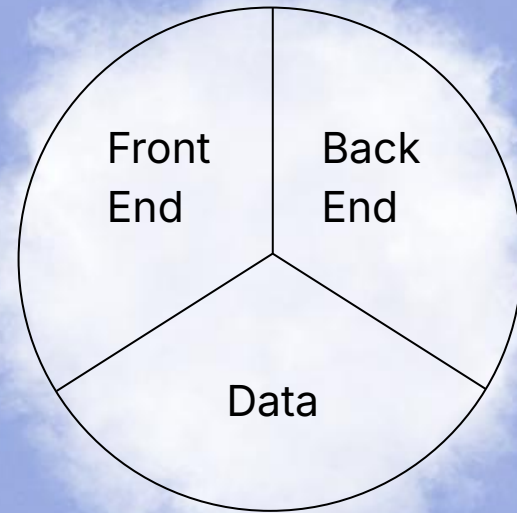
Engineers Be Engineering



Data Quality



Who's Responsible for Data Quality?




Data Quality in the Age of AI



Conclusion




How to Not Hate Each Other



Remember:
you're not supposed
to know each other's
tools

How to Not Hate Each Other

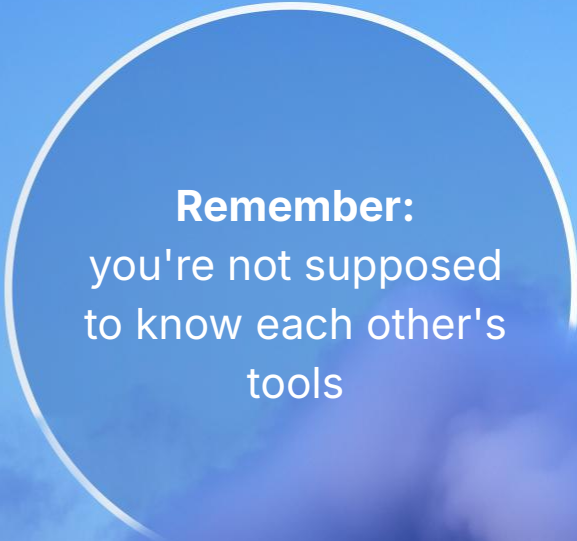
- Define interfaces early—data contracts, API specs



Remember:
you're not supposed
to know each other's
tools

How to Not Hate Each Other

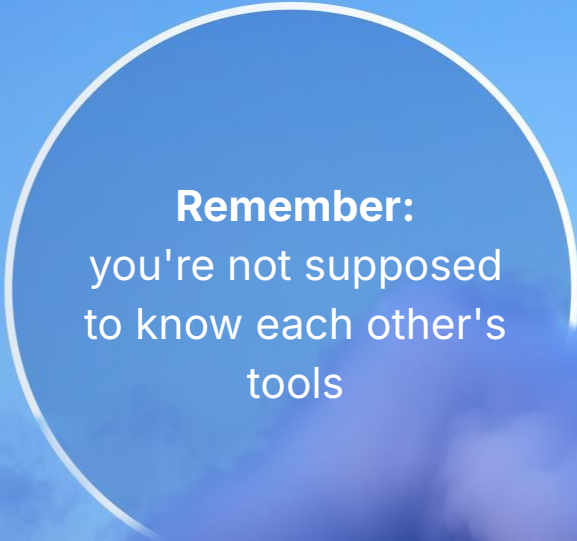
- Define interfaces early—data contracts, API specs
- Communicate assumptions—latency, freshness, format



Remember:
you're not supposed
to know each other's
tools

How to Not Hate Each Other

- Define interfaces early—data contracts, API specs
- Communicate assumptions—latency, freshness, format
- Pair on cross-cutting concerns



Remember:
you're not supposed
to know each other's
tools

Same Same, But Different



Same Same, But Different

- We solve different problems, with the same intent



Same Same, But Different

- We solve different problems, with the same intent
- Collaboration > turf wars



Same Same, But Different

- We solve different problems, with the same intent
- Collaboration > turf wars
- Curiosity, communication, and empathy = better teams



Same Same, But Different

- We solve different problems, with the same intent
- Collaboration > turf wars
- Curiosity, communication, and empathy = better teams
- And please—stop storing JSON blobs in your DB



Questions?



The End