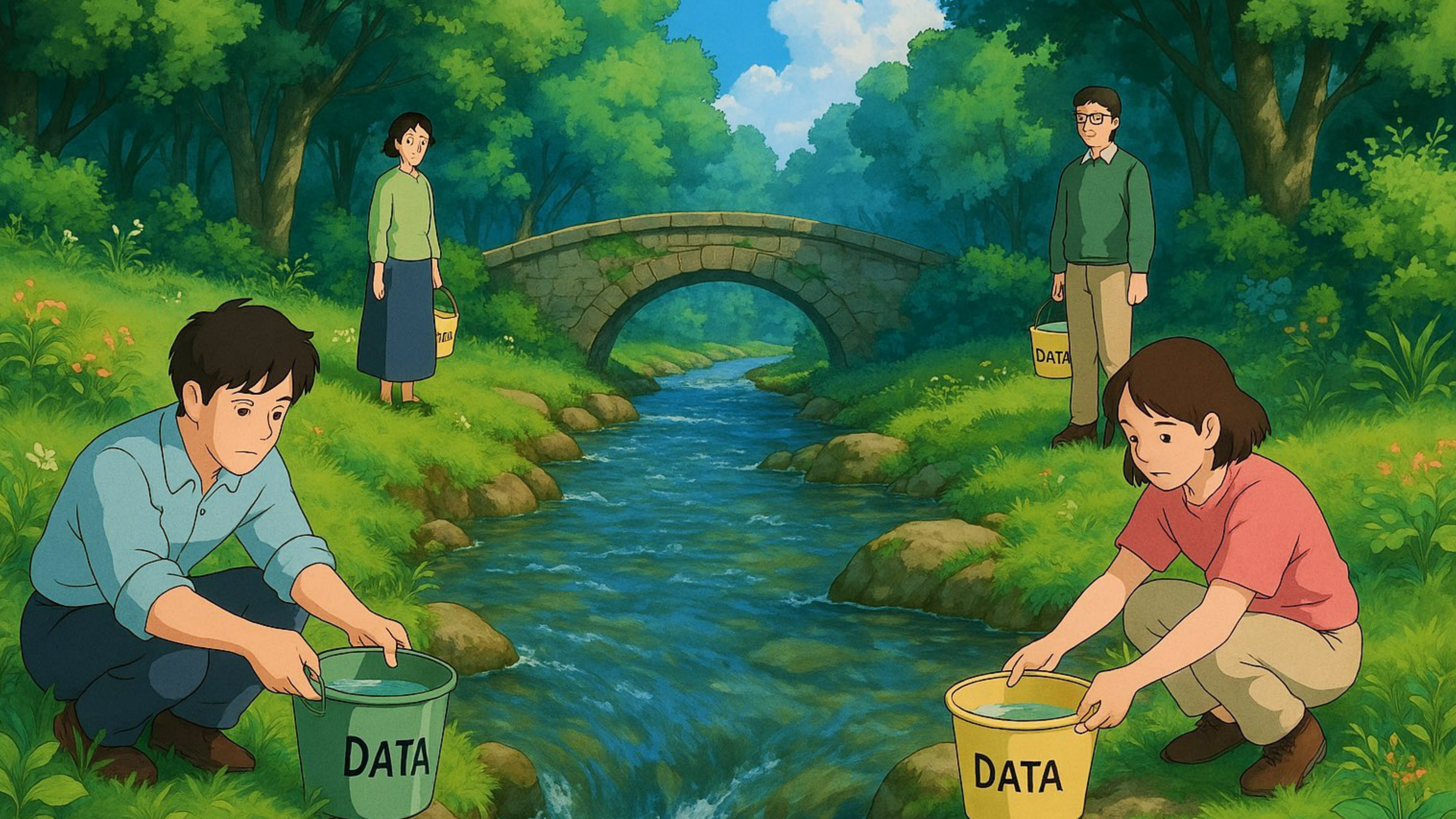


One Does Not Simply Query a Stream!

Viktor Gamov, Confluent
@gamussa



Amsterdam JUG
June 4, 2025

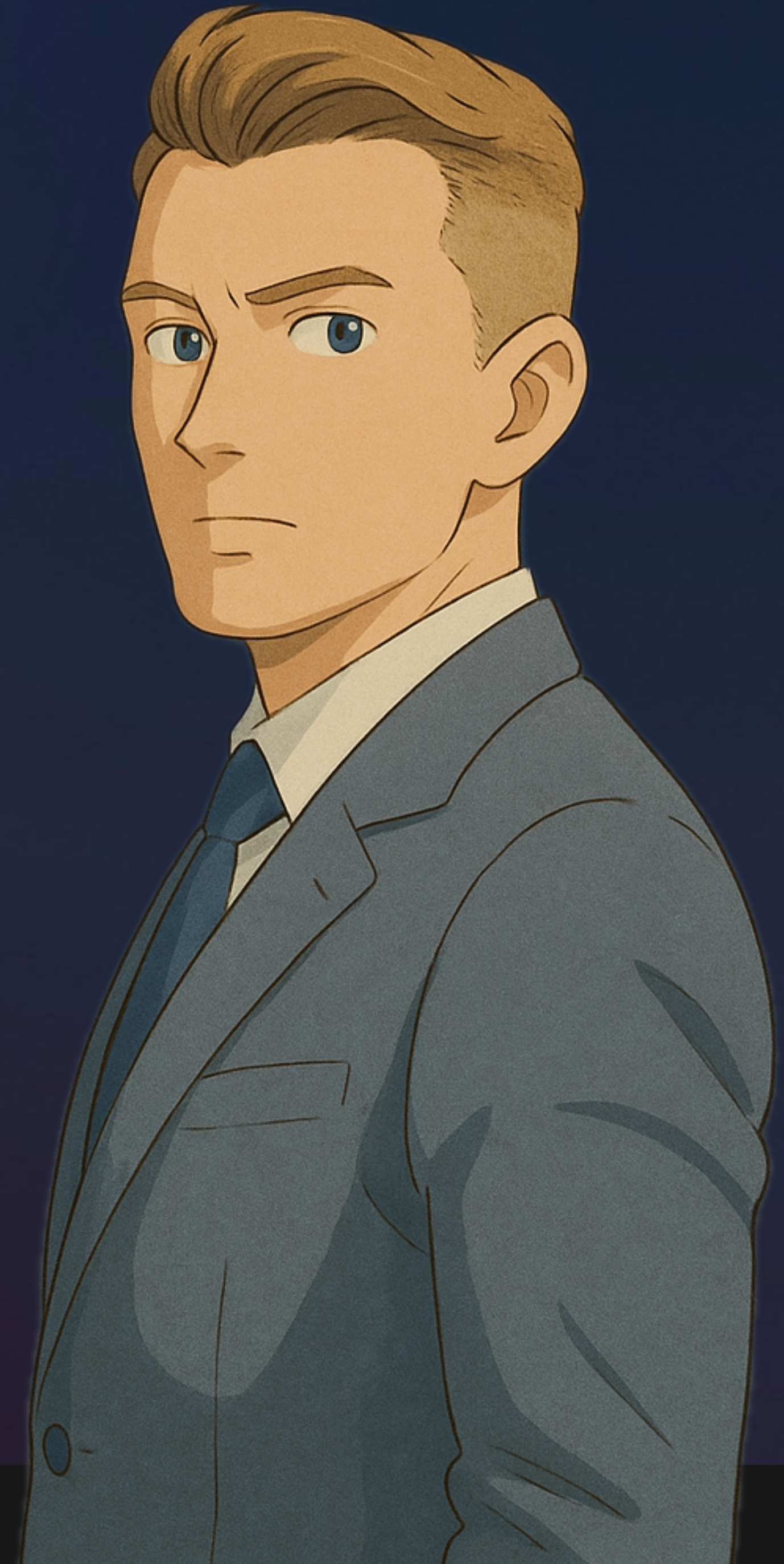


Viktor

GAMOV

Principal Developer Advocate | Confluent
Kafka in Action, Co-author

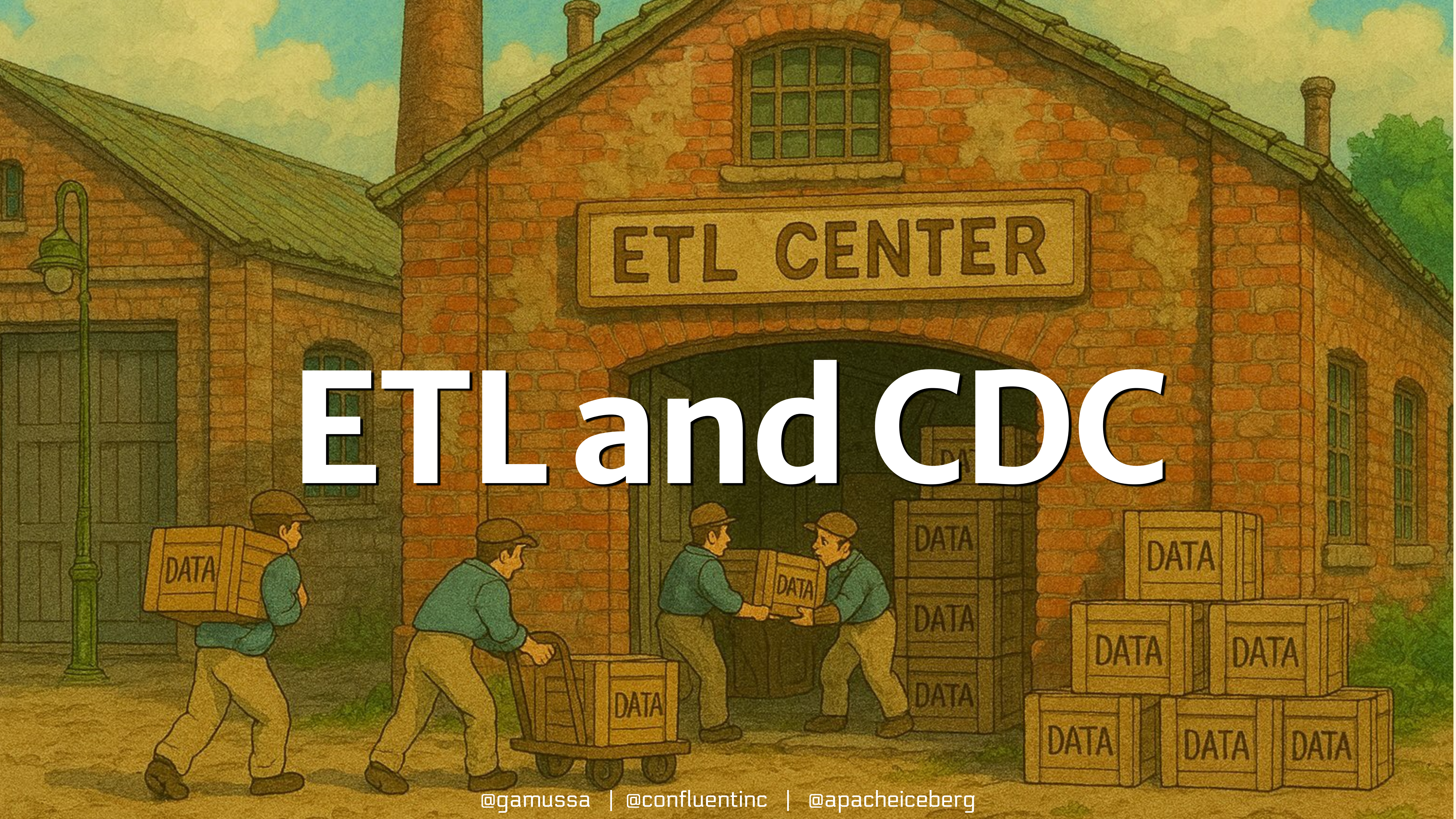
X and Bluesky: @gamussa

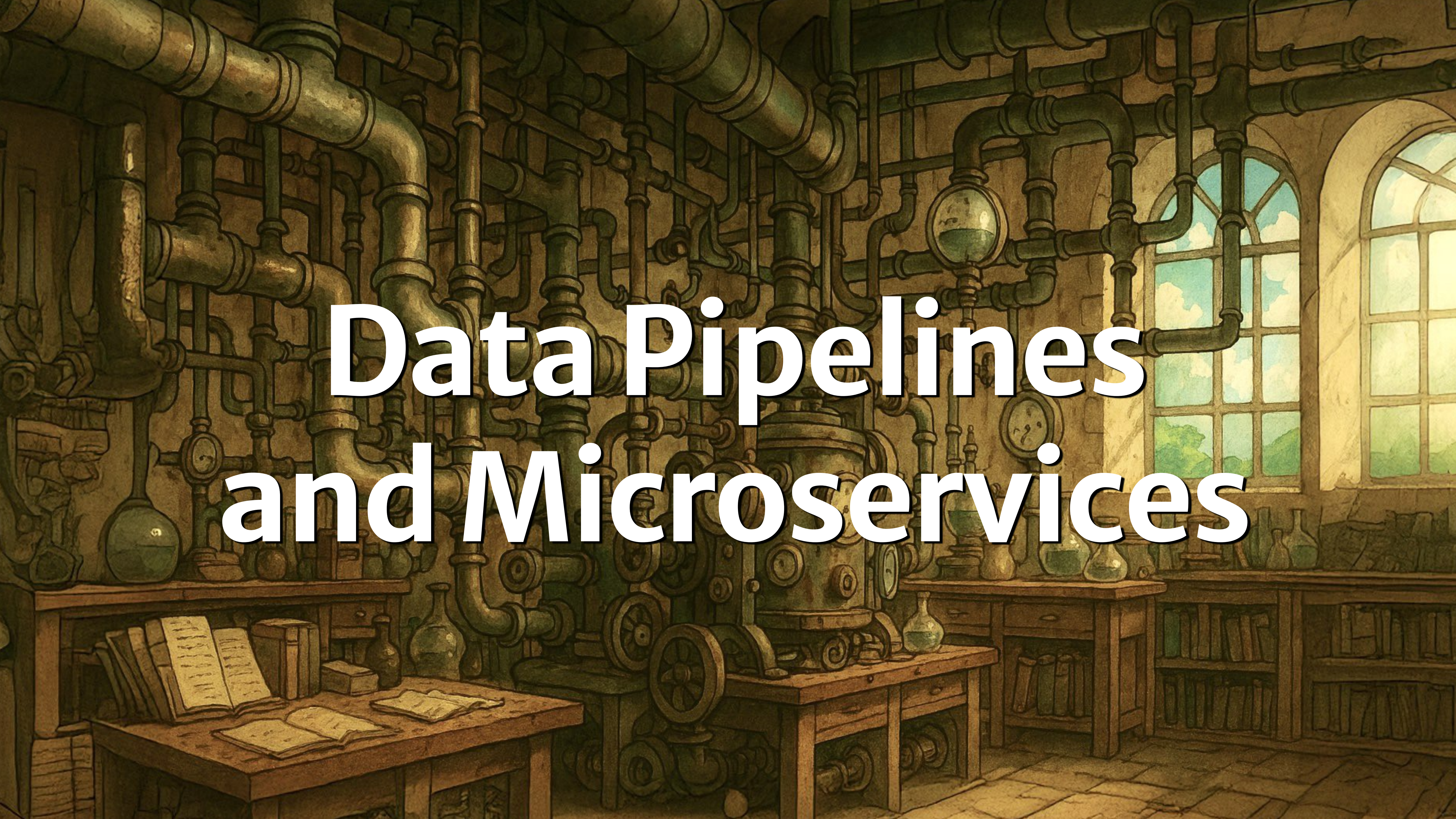


Monolith



ETL and CDC



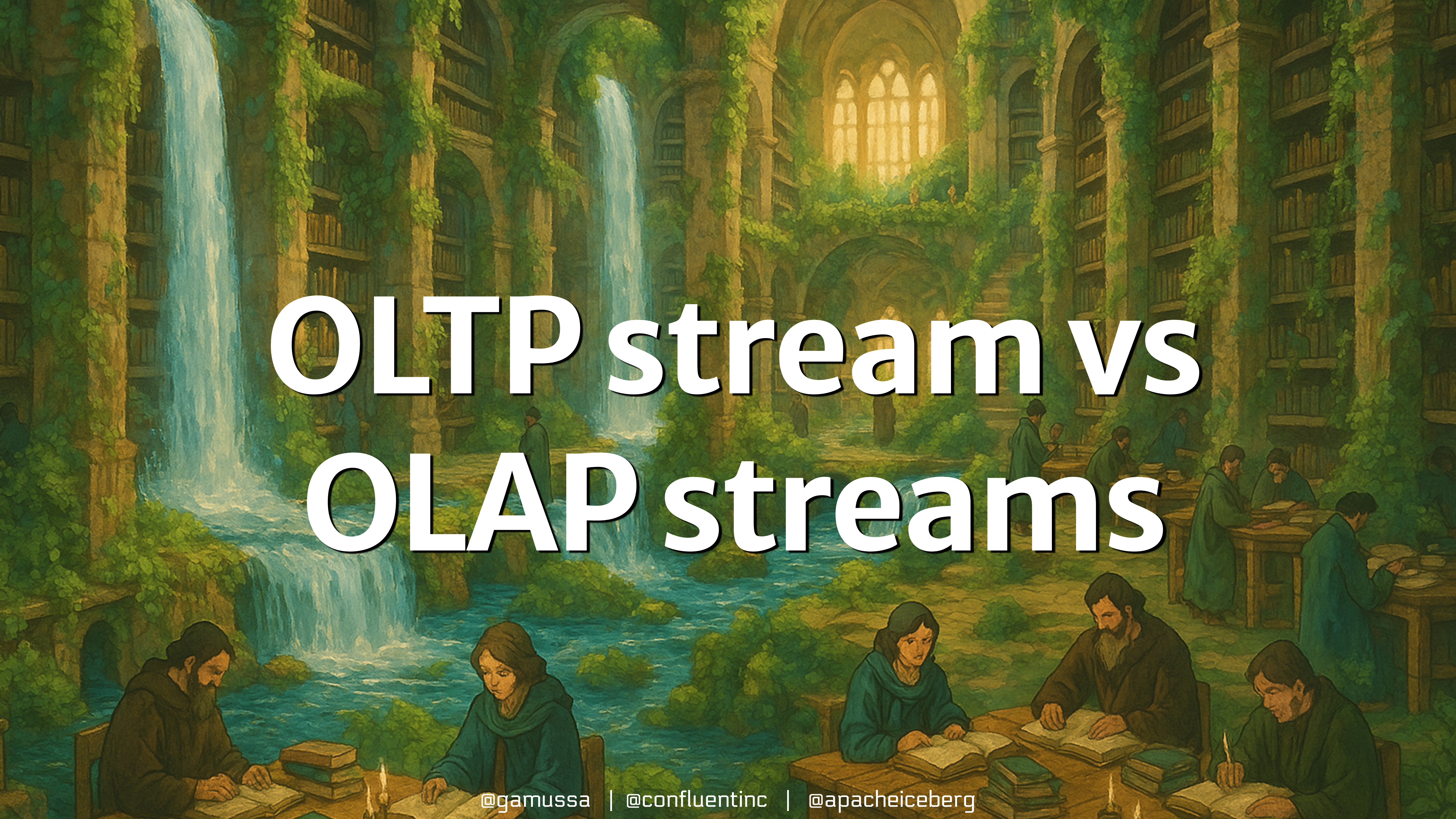


Data Pipelines and Microservices

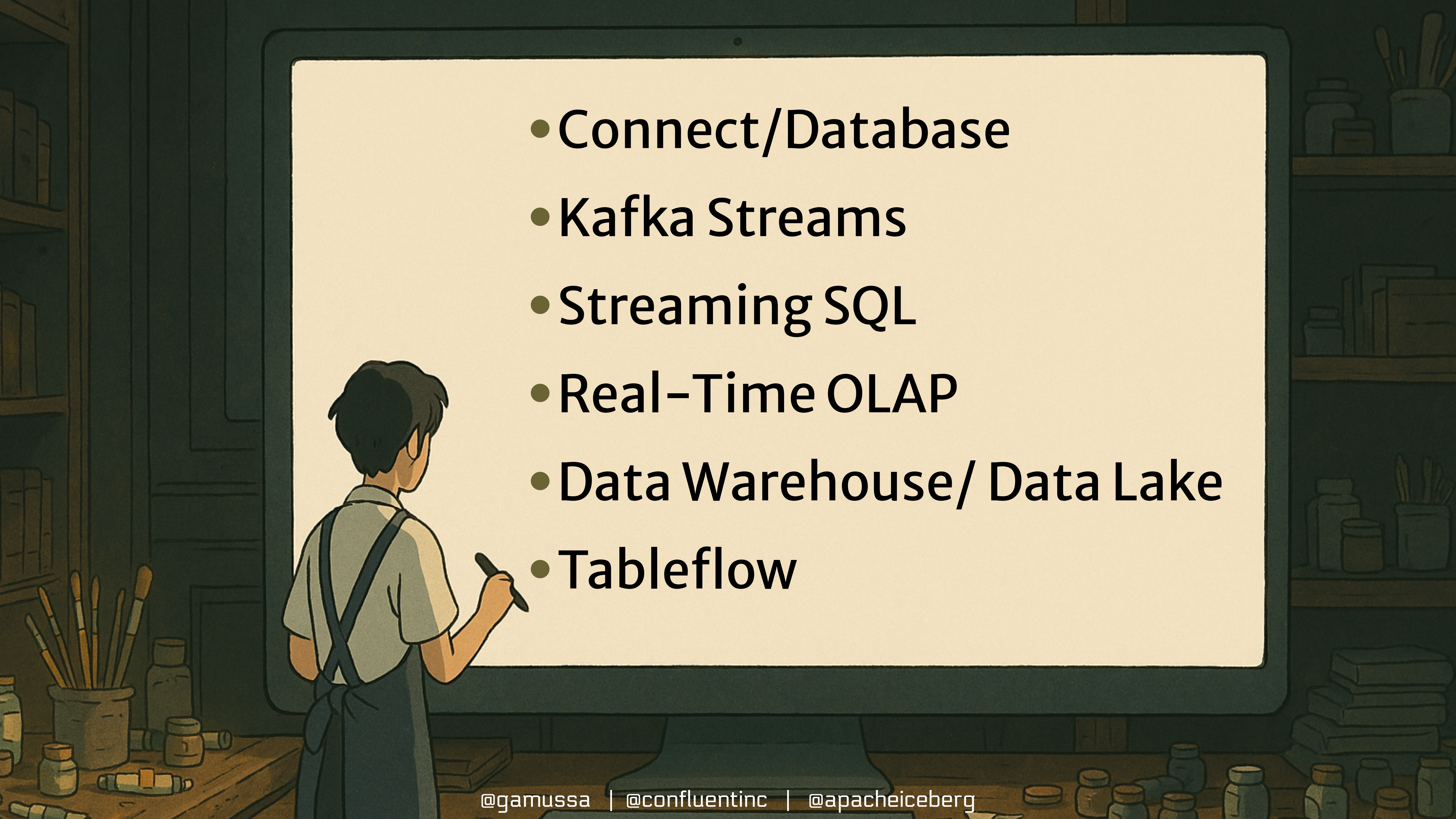


LOG

APPEND-ONLY
NO EXIT

A detailed illustration of a medieval library. Two large waterfalls flow down the sides of a grand, vaulted stone building. The interior is filled with tall, arched bookshelves packed with books. In the foreground and middle ground, several scholars in medieval attire are seated at wooden desks, focused on reading or writing. The scene is lit with warm, golden light from a large window in the background, creating a serene and scholarly atmosphere.

OLTP stream vs OLAP streams

- 
- An illustration of a person with dark hair, wearing a light-colored shirt and a dark apron, standing in a workshop. They are pointing with a marker at a large screen that displays a list of data technologies. The workshop is filled with various tools, jars, and books on shelves in the background.
- Connect/Database
 - Kafka Streams
 - Streaming SQL
 - Real-Time OLAP
 - Data Warehouse/ Data Lake
 - Tableflow



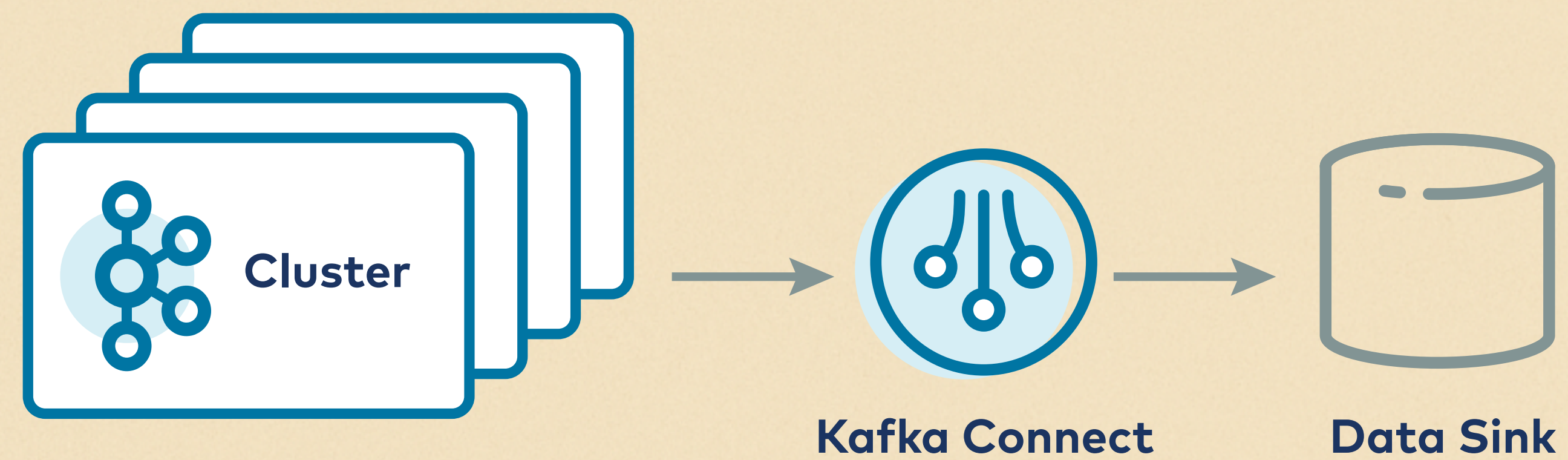
Kafka Connect

@gamussa | @confluentinc | @apacheiceberg

Connect

- Suitable for smaller data
- Transactional
- Familiar to users

Connect







Kafka Streams

Kafka Streams

(transactional)

- Ingests directly from a **topic**
- KTable
- Forms an **in-memory key/value** store suitable for querying by topic key
- Scalable across **members of a consumer group**
- Readable through **Interactive Queries**

Kafka Streams

(analytical)

- Full-featured **Java stream processing API**
- Arbitrary streaming computation
- Can emit new streams (not this talk)
- `KTables` **queryable by key**
- Every read pattern requires its own topology
- **Interactive Queries** again





Streaming SQLs

Streaming Database

- SQL for Queries
- Streaming Source is 1st class citizen
- Persistence / Storage

Streaming SQL

- ksqlDB
- Materialize
- RisingWave
- TimePlus



But Viktor, Flink
has SQL



What is Apache Flink® SQL?



ksqlDB

- «*Streaming Database*»
- Provides persistent TABLE abstraction
- Pull and Push queries
- Like **Kafka Streams**, but in **SQL**

Materialize

- Replacement data warehouse
- Integrates with Kafka, Postgres, dbt
- The **Materialized View** is the central abstraction
- Views are **persistent and queryable**
- Postgres wire-compatible
- Positioned as an analytics solution

Rising Wave

- Distributed SQL Streaming database
- Cloud and OSS versions
- Implementation of **Flink in Rust**
- **Kafka, Pulsar, Kinesis** integrations
- Flink+persistent views
- Postgres wire-compatible

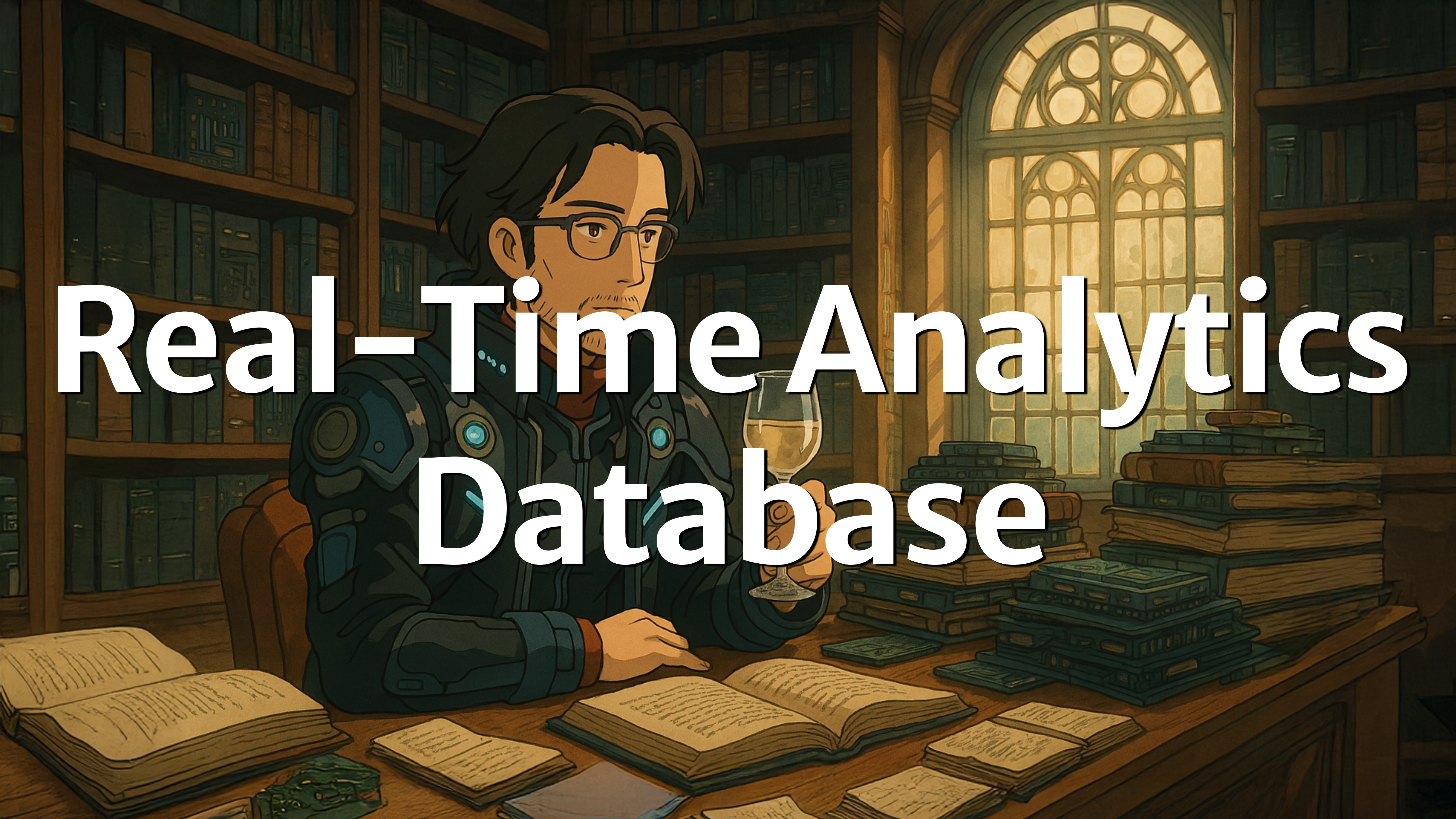


```
SELECT  
product_id.  
Av6irating)  
FROM reviews  
GROUP BY
```

```
SELECT  
room revting  
GROUP BY .....  
product_id
```

```
SELECT  
FROM orders  
WHERE status =  
'shipped'
```

```
SELECT ★  
FROM orders  
WHERE status =  
'shipped'
```

Real-Time Analytics Database

Real-Time OLAP

- Designed for **high concurrency, low latency queries**
- Ingests from **streaming and batch sources**
- Intimate integration with **Kafka**
- Conventional **tables and SQL**

Real-Time OLAP

- Analytics shaped like **real-time data**
- Analytics when **users are decision makers**



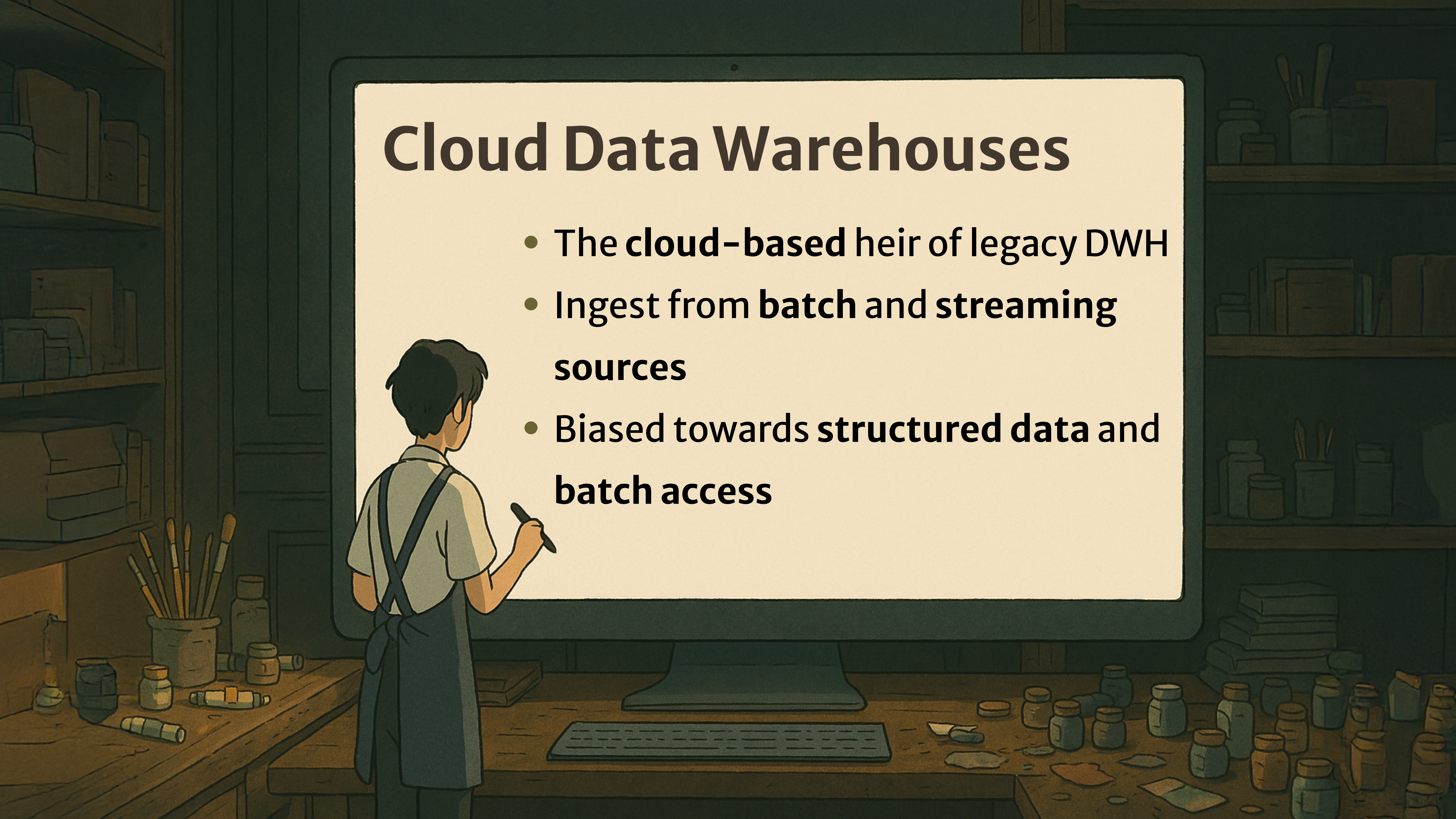
Cloud Data Warehouses

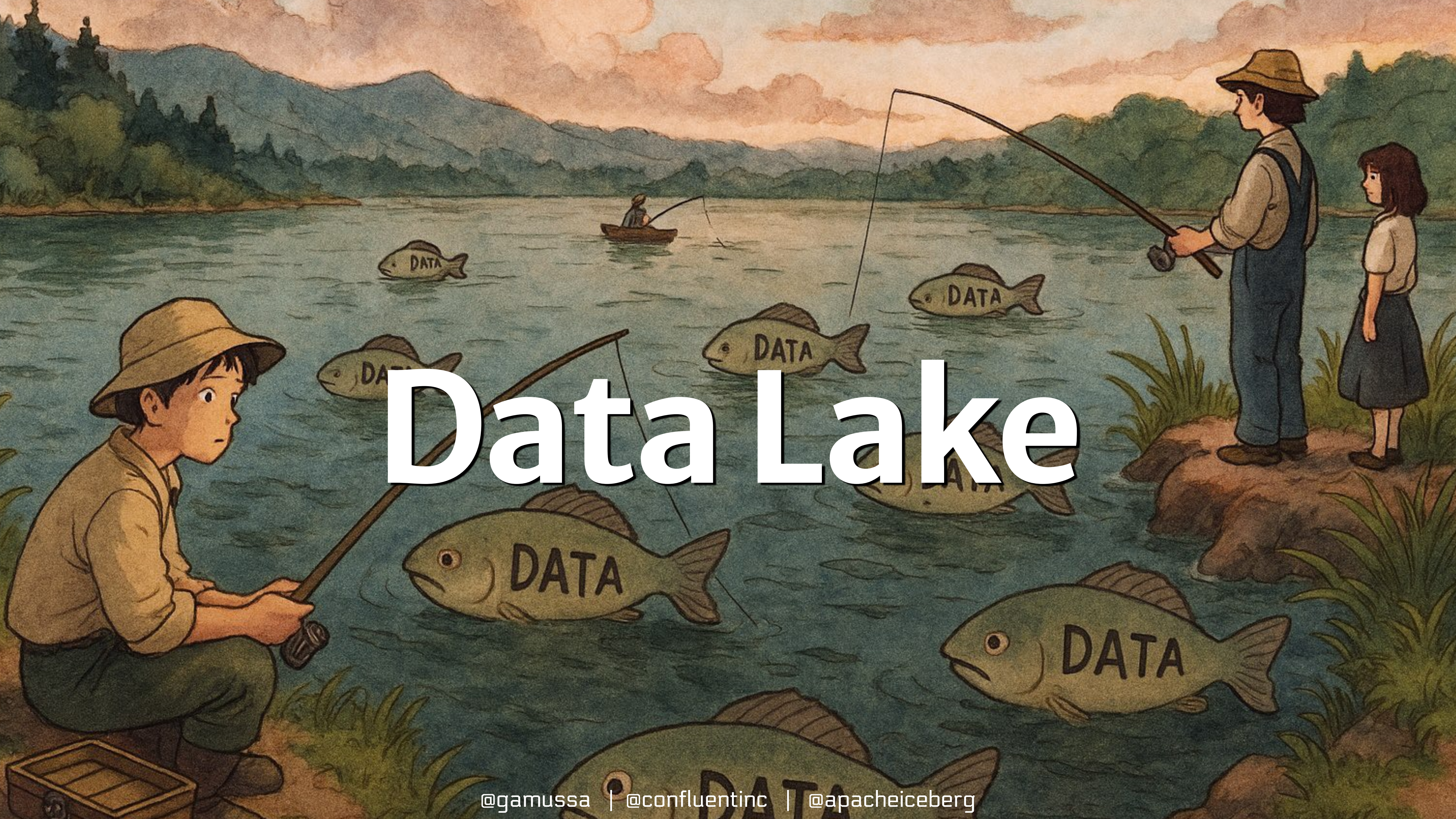
Cloud Data Warehouses



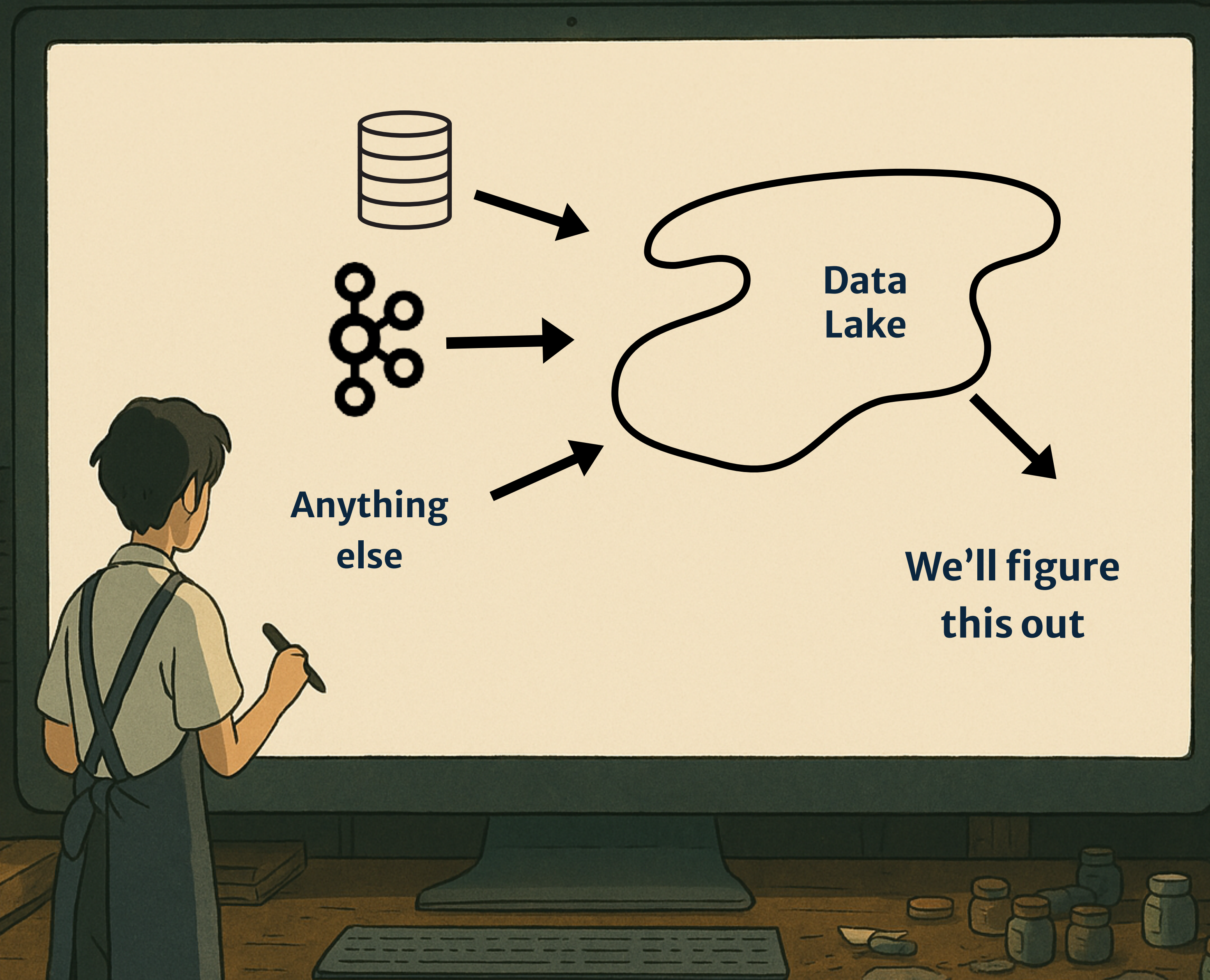
Cloud Data Warehouses

- The **cloud-based heir** of legacy DWH
- Ingest from **batch and streaming sources**
- Biased towards **structured data and batch access**



A watercolor illustration of a serene lake scene at sunset. In the foreground, a young boy in a tan shirt and hat sits on the grassy bank, holding a fishing rod. To his right, a man in a hat and overalls stands on a rock, also fishing, with a young girl in a white shirt and dark skirt standing beside him. Several fish are visible in the water, each with the word "DATA" written on its side. In the background, a small boat with another person is on the water, and distant mountains are visible under a cloudy, orange-hued sky.

Data Lake



Data Lakes

- Storage and compute are radically decoupled
- Structure is relatively less important
- Reads are slow
- Streaming is historically difficult

Data Lakes

- Started as the HDFS cluster
- Became S3
- That didn't help...
- ELT vs. ETL
- Iceberg/Hudi/DeltaLake

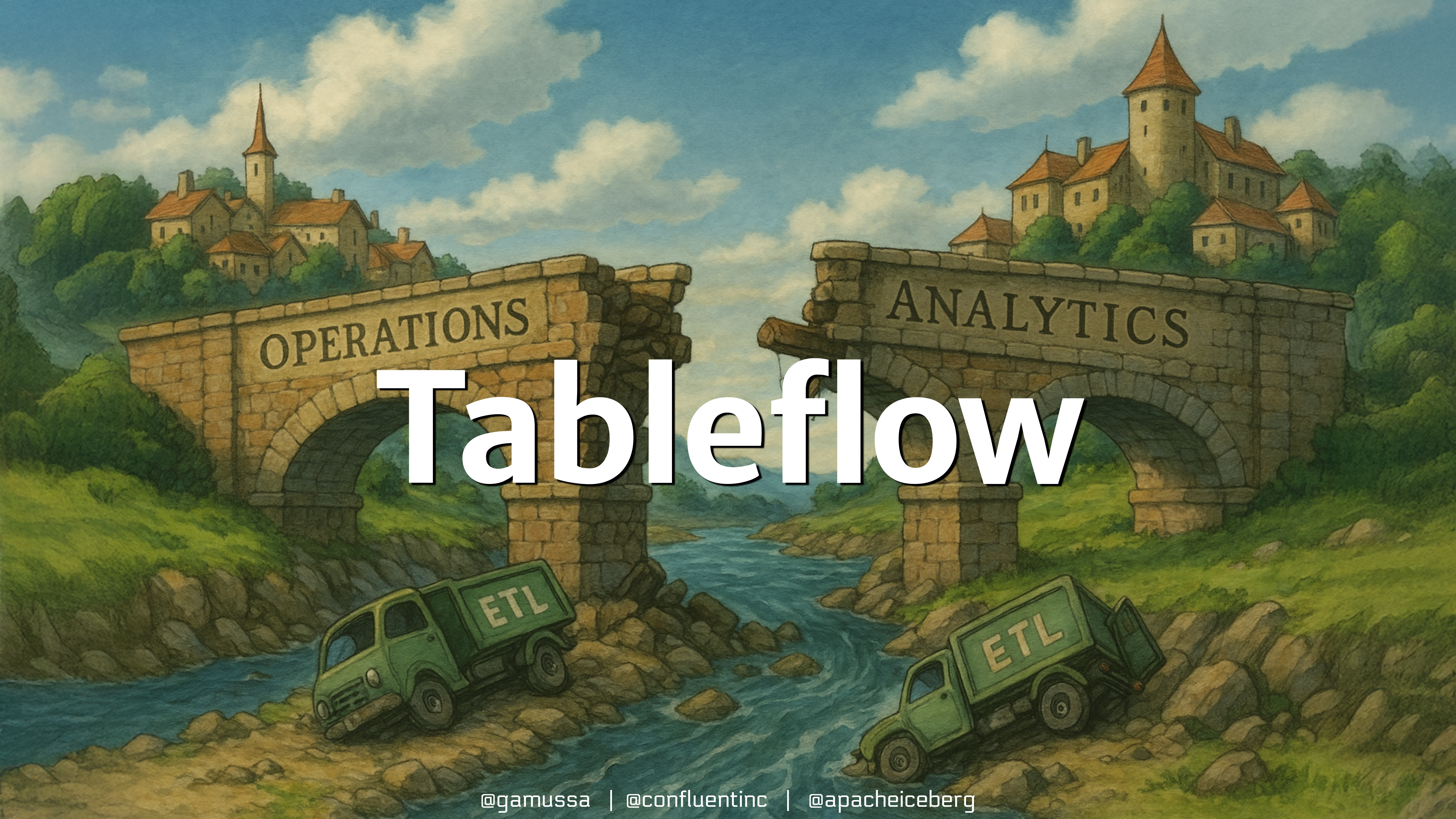
An illustration of a large blue iceberg floating in a body of water. The iceberg has a white, snow-covered peak. The word "Iceberg" is written in large white letters across the middle of the iceberg. Below the word, there are three circular icons: a clock for "TIME-TRAVEL", a network diagram for "SCHEMA", and a table grid for "TABLE". The background shows a sunset or sunrise sky with orange and blue clouds, and dark green hills in the distance.

Iceberg

TIME-TRAVEL

SCHEMA

TABLE



Tableflow

Kafka



TABLE



user@host ~/streaming-github-events

```
SELECT * FROM github.default.event_stream LIMIT 5;
```

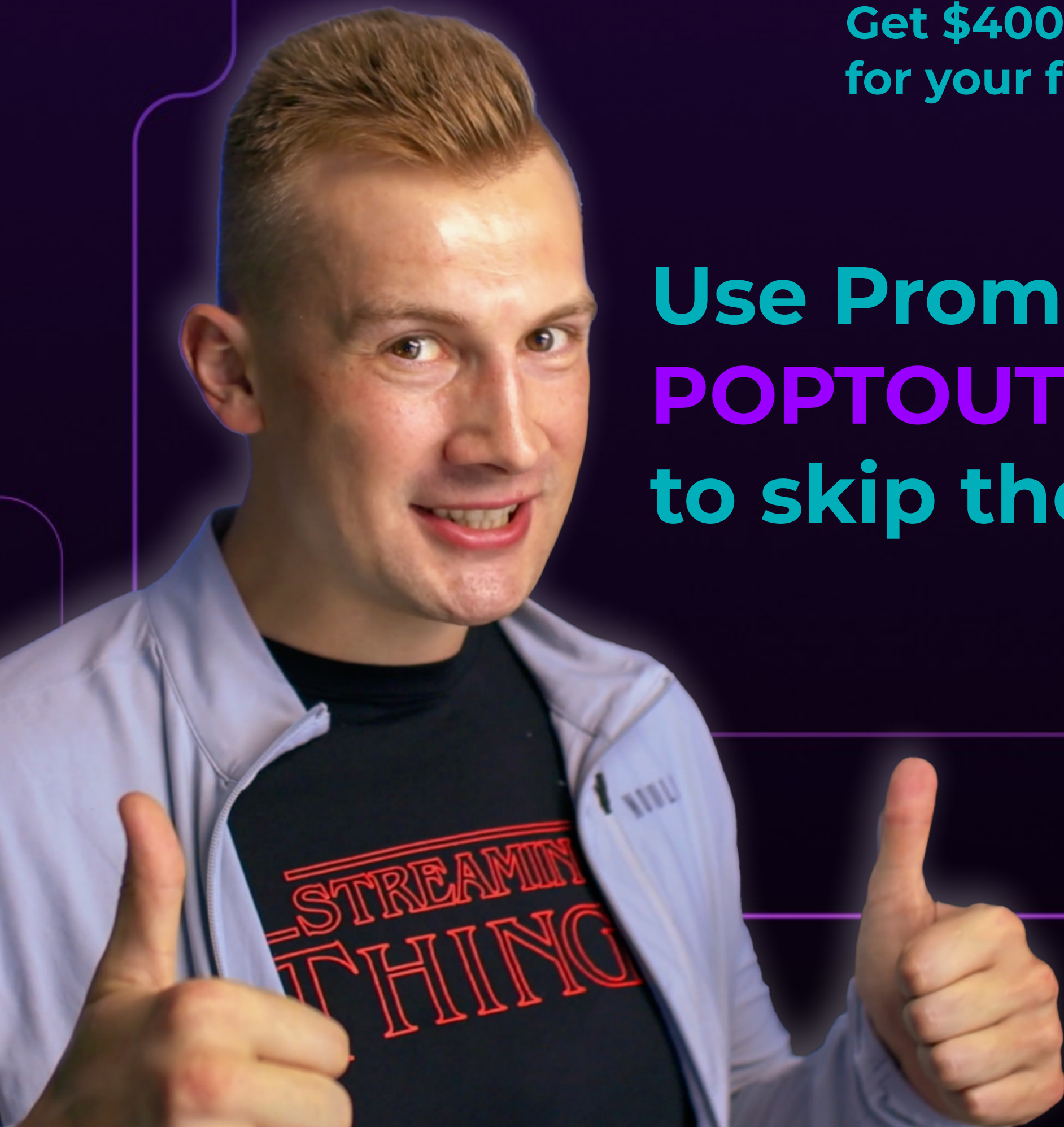
repository_name	eventtype	creationstamp
apache/pinot	WatchEvent	2024-04-24 15:31:54
open-telemetry/opentelemetry-		2024-04-24 15:31:54
python-contrib	python-contrib	
cornerstone.js/cornerstone	WatchEvent	2024-04-24 15:31:54
aws-gates/ec2-instance-connect-cli		2024-04-24 15:31:54
rancher/rancher	WatchEvent	2024-04-24 15:31:54

Skip Paywall

Sign Up for Confluent Cloud

Get \$400 worth free credits
for your first 30 Days

Use Promo Code -
POPTOUT000MZG62
to skip the paywall!



A detailed steampunk illustration serves as the background. It depicts a man with dark, spiky hair and round glasses, wearing a light-colored lab coat over a grey shirt and dark trousers. He stands in the center, looking directly at the viewer with a neutral expression. The room is filled with an intricate network of brass pipes, valves, and mechanical components. On the left, a large brass sphere with multiple gauges and a large wooden wheel are visible. To the right, another brass sphere is mounted on a wooden stand. The floor is made of stone tiles, and the overall lighting is dim, creating a mysterious and industrial atmosphere.

No Solutions only Trade Offs



Sometimes you go
with what you know

A cartoon illustration of a man with a beard and a grey cap, smiling as he reads a children's book. He is sitting in a wooden armchair in a room filled with bookshelves. The book he is holding shows a sun, a rabbit, and two children. The text "This is not bad!" is overlaid in large white letters. The scene is warmly lit by a window with yellow curtains in the background. On the desk in front of him are some papers and a small container.

This is not bad!



Performance

A detailed illustration in a medieval or Renaissance style. A man with a dark beard and a blue robe stands behind a large, complex mechanical device, possibly an early clock or astronomical instrument. He has his right hand raised in a gesture of explanation or awe. A large crowd of people, dressed in period clothing, surrounds him, looking up at him with expressions of wonder and curiosity. The setting is a grand hall with high arched ceilings, stone columns, and large windows. The lighting is warm and golden, suggesting an interior space with light coming from the windows. The word "Community" is overlaid in a large, white, sans-serif font across the center of the image.

Community

Differentiated Application Code

Area of
Exploration

Kafka



This is not bad!