



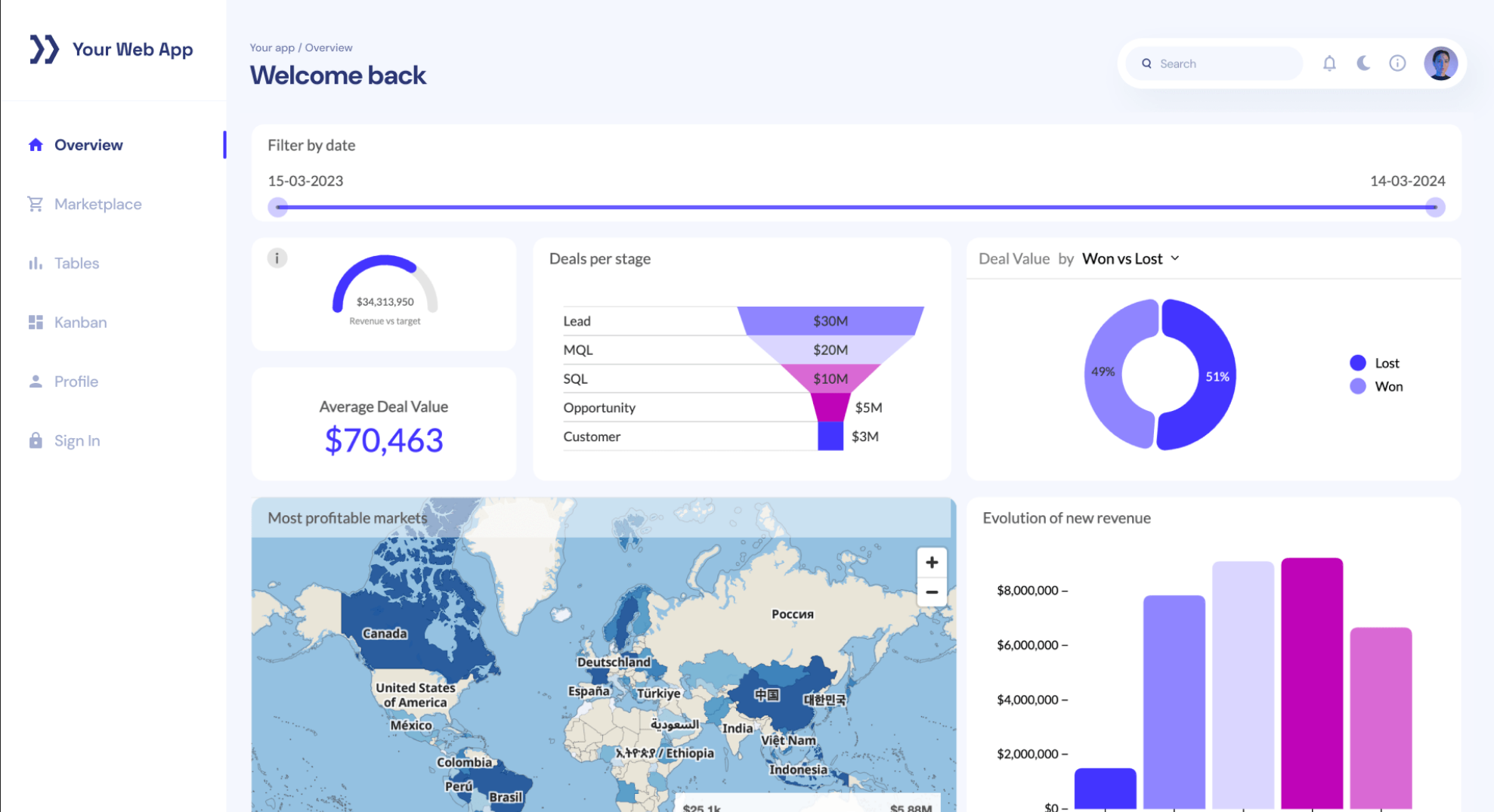
# Using ClickHouse for fun & profit

March 2025

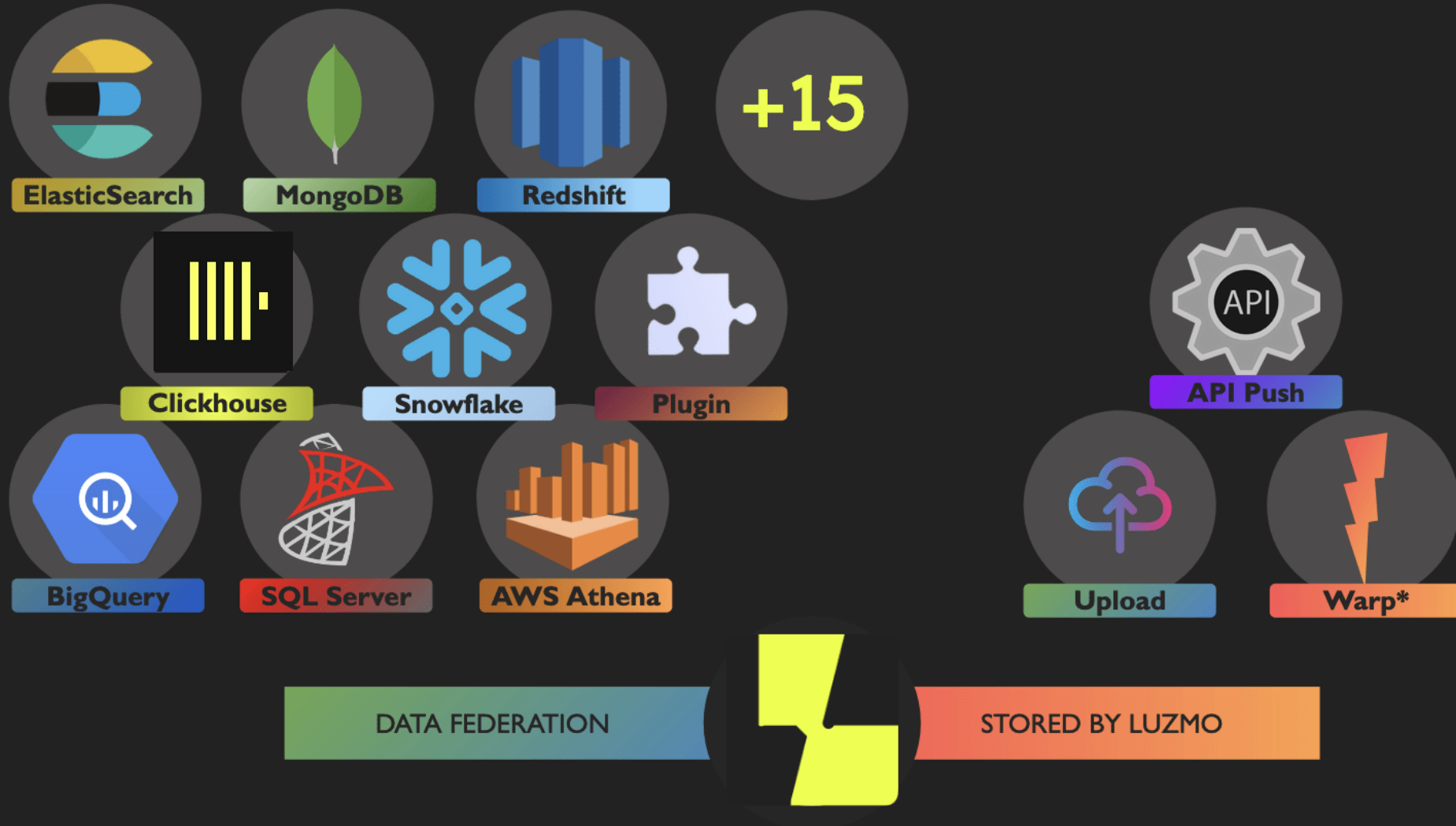


Haroen Vermylen  
Founder & CTO

# Giving your clients the right insight at the right time



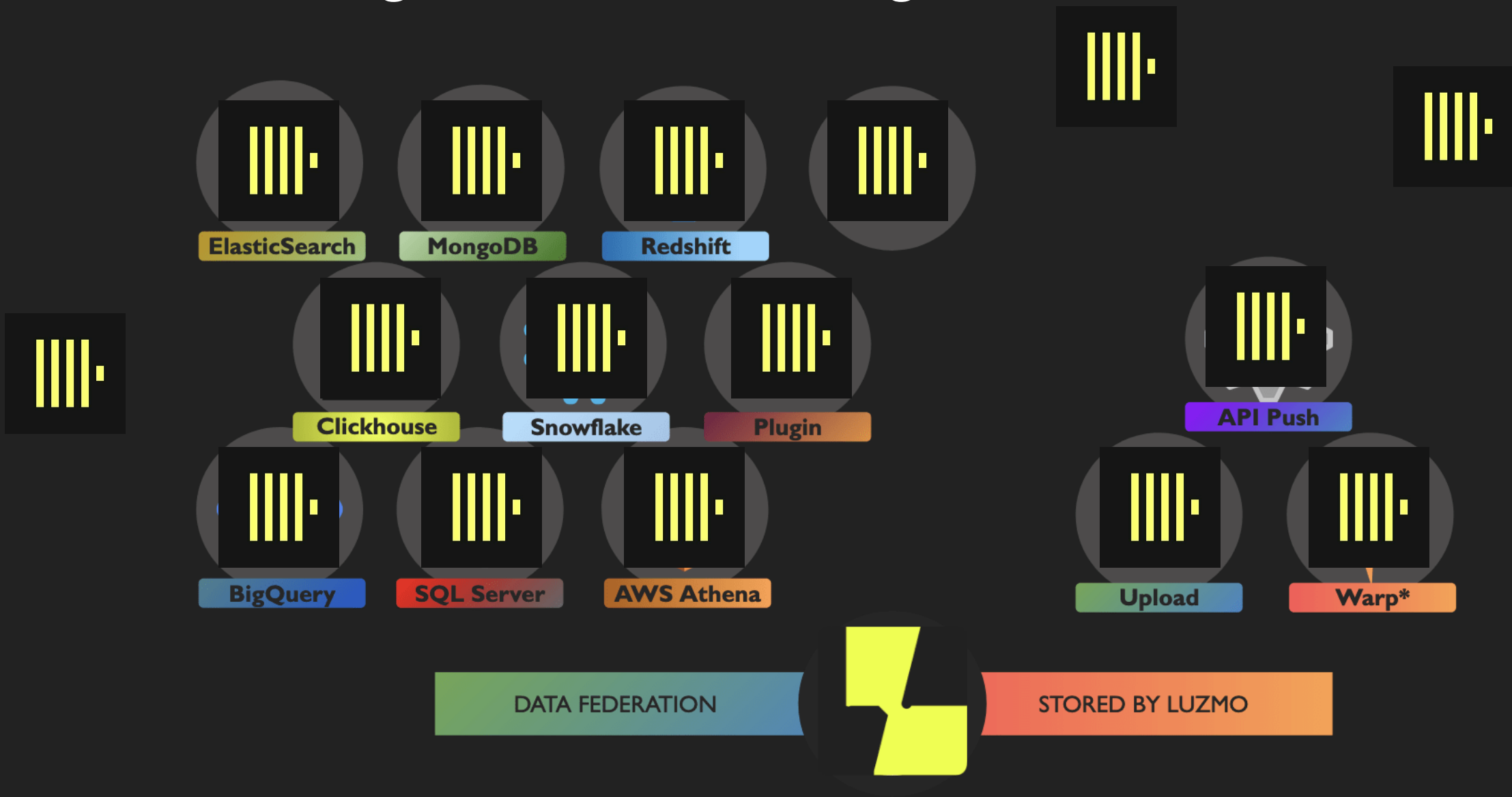
# Data source-agnostic



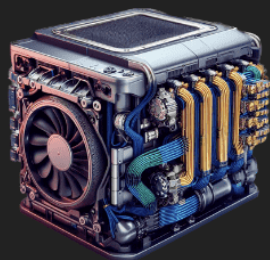
Data source-agnostic... but betting on ClickHouse since 2017



Data source-agnostic... but betting on ClickHouse since 2017



# Client data is everywhere + dashboard performance is critical



## Query Engine

- Optimized for our clients' workload  
**{ low-latency , federated , analytical }**
- Up to 70x faster dashboard load times
- Multi-tenancy with shuffle sharding

**100M rows/s**  
processed (8T / day)

**8.6M rows/s read**  
**275k rows/s write**  
on our ClickHouse

# Clients have outdated data stacks

## Warp

- One-click sync from transactional sources
- Up to **350x** faster than relational DBs

Clients ~~have outdated data stacks~~  
don't use ClickHouse yet! So we make them.

### Warp

- One-click sync from transactional sources
- Up to 350x faster than relational DBs



Lots of lessons learned & famous last words

# Lots of lessons learned & famous last words

*"Let's just run our own Zookeeper cluster. It can't be so bad..."*

ClickHouse launched ClickHouse Keeper in 2021

We migrated to ClickHouse Cloud in 2023

# Lots of lessons learned & famous last words

*"Let's create a ClickHouse table  
per uploaded / Warped dataset"*

ClickHouse hates having 50k+ tables  
This works surprisingly well though:

```
CREATE TABLE data (  
    dataset_id Uuid,  
    columns Map(Uuid, Dynamic),  
) ENGINE = MergeTree();
```

# Lots of lessons learned & famous last words

*"Let's emulate atomic large batch inserts by creating temporary tables and merging them."*

Avoid DDL queries during "normal" operations. Just build on ClickHouse's read/write performance instead:

- Insert with a version number
- Batch failed? Just ignore that versions and clean it up async

# Lots of lessons learned & famous last words

*"Nobody will store dates older than 1600, right?"*

SQL Server loves *0000-00-00*.

ClickHouse added Datetime64 support in 2020

# Technology-driven innovation

<https://instach.art>

# ClickHouse as a vector store

## 1. Create indexed embeddings for (textual, image, audio, ...) data

```
clickhouse-client --query "CREATE TABLE embeddings (...,  
    vectors Array(Float32)  
    INDEX idx_vectors TYPE vector_similarity('hsnw', 'cosineDistance')  
) ENGINE = MergeTree();"

curl http://ollama:11434/api/embeddings -d '{  
    "model": "llama3",  
    "prompt": "Get ready to sink your teeth into the most realistic and deliciously flavored rubber pizza  
you've ever had! Made from BPA-free materials, Tasty Rubber Pizza is the perfect snack for anyone."  
}'  
| clickhouse-client --query "INSERT INTO embeddings (vectors) FORMAT TabSeparated"
```

## 2. Search relevant documents

```
clickhouse-client --query  
    "SELECT id, cosineDistance(vectors, [0.1, 0.2, 0.3] ) AS relevance ORDER BY 2 DESC LIMIT 10"
```

# Introducing Luzmo IQ

<https://ayd.luzmo.com/luzmo-demo/analyst>

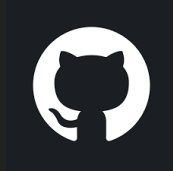


Luzmo

# Questions?



Haroen Vermylen  
Founder & CTO



@kagaherk

