# Iceberg Catalogs

Arsham Eslami - Greybeam
Feb 27th 2025

# What is a Data Catalog?

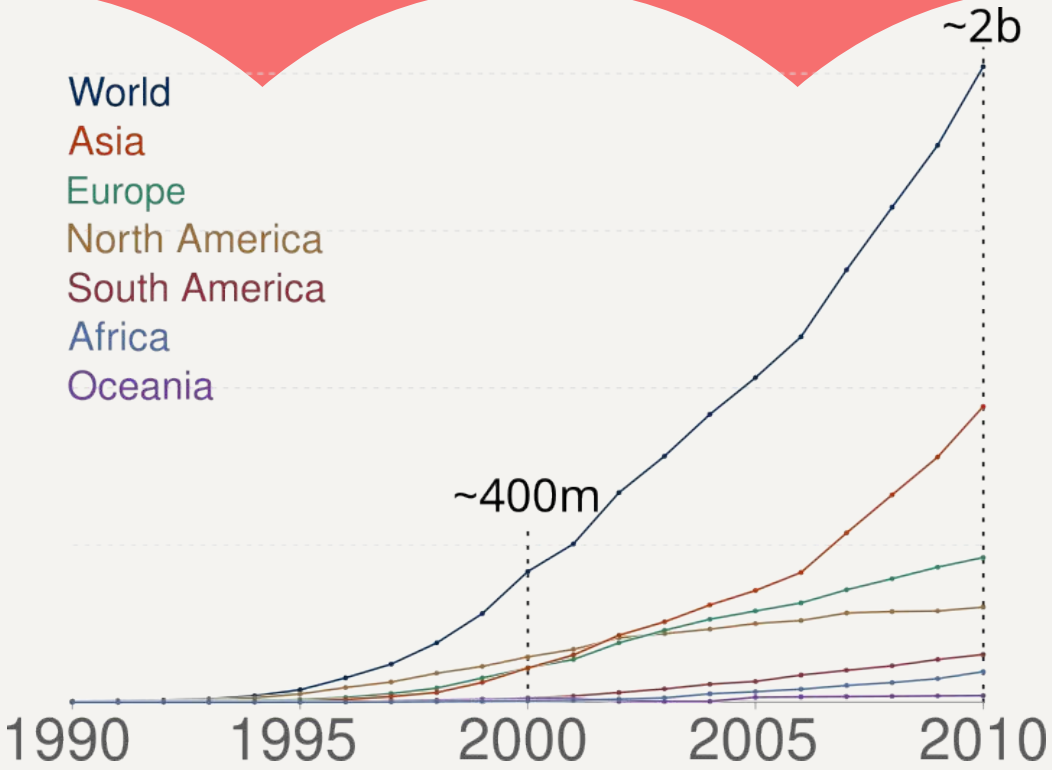
**Schema Registry**

**Metadata Store**

**Business Context**

# Why you shouldn’t listen to me!

# How did we get here?


~2b
World
Asia
Europe
North America
South America
Africa
Oceania
~400m
1990
1995
2000
2005
2010

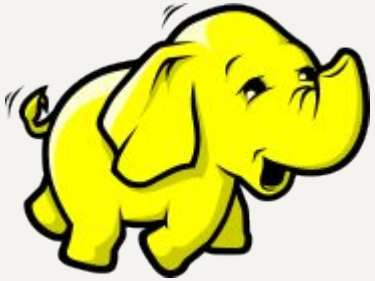
# Massive Datasets

Oracle, MySQL, SQL Server

Teradata, Netezza, Greenplum

MapReduce, GFS

# Hadoop Era

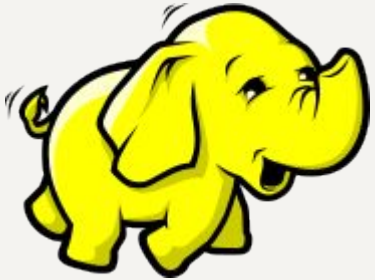


Hadoop

Hive

Avro & Parquet

# Lakehouse Era



**Object Storage**

**Open Tables Formats**

**Query Engines**

# Why do we need a catalog?

**Managing Files**

**Transaction Management**

**Improve Query Performance**

# Service Based Catalogs

Table metadata in a **relational database**

**Centralized** approach

# File Based Catalogs

Metadata is **stored as files** inside object storage

**Decentralized** approach

Metadata is **stored as files** inside object storage but…

**Centralized** via API

REST spec is open source

# The OGs

# Tabular Catalog

**Purpose**

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

**Example**

“Lorem ipsum dolor sit amet, consectetur adipiscing elit.”

# AWS Glue Catalog

**Purpose**

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

**Example**

“Lorem ipsum dolor sit amet, consectetur adipiscing elit.”

# Project Nessie

## Purpose

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

## Example

“Lorem ipsum dolor sit amet, consectetur adipiscing elit.”

# The New Kids

# Polaris

**Purpose**

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

**Example**

“Lorem ipsum dolor sit amet, consectetur adipiscing elit.”

# Gravitino

## Purpose

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

## Example

“Lorem ipsum dolor sit amet, consectetur adipiscing elit.”

# LakeKeeper

## Purpose

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

## Example

“Lorem ipsum dolor sit amet, consectetur adipiscing elit.”

# Q&A

**Interested in connecting?**

**Slides & Sources**