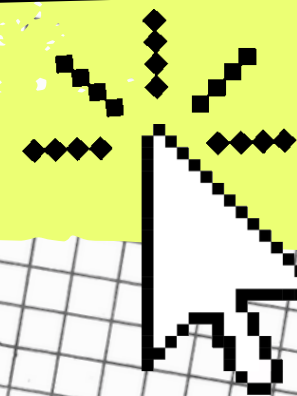


DATA LAKE



VAIBHAV KUTE


WHAT IS A DATA LAKE?

A data lake is a centralized repository designed to store, process, and secure large amounts of structured, semistructured, and unstructured data.

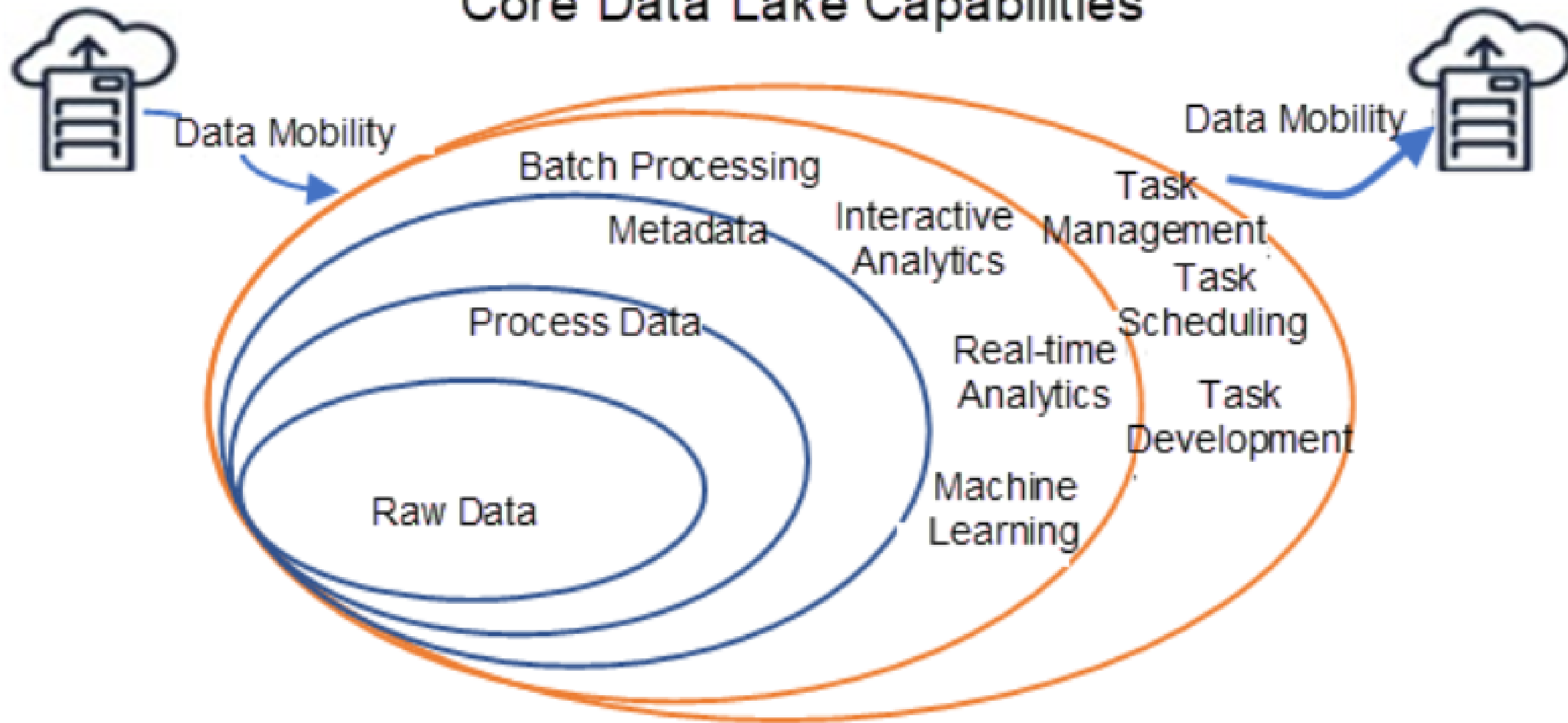
It can store data in its native format and process any variety of it, ignoring size limits.

DO YOU NEED A DATA LAKE?

Developed primarily to handle large volumes of big data, companies can typically move raw data via batch and/or stream into a data lake without transforming it.

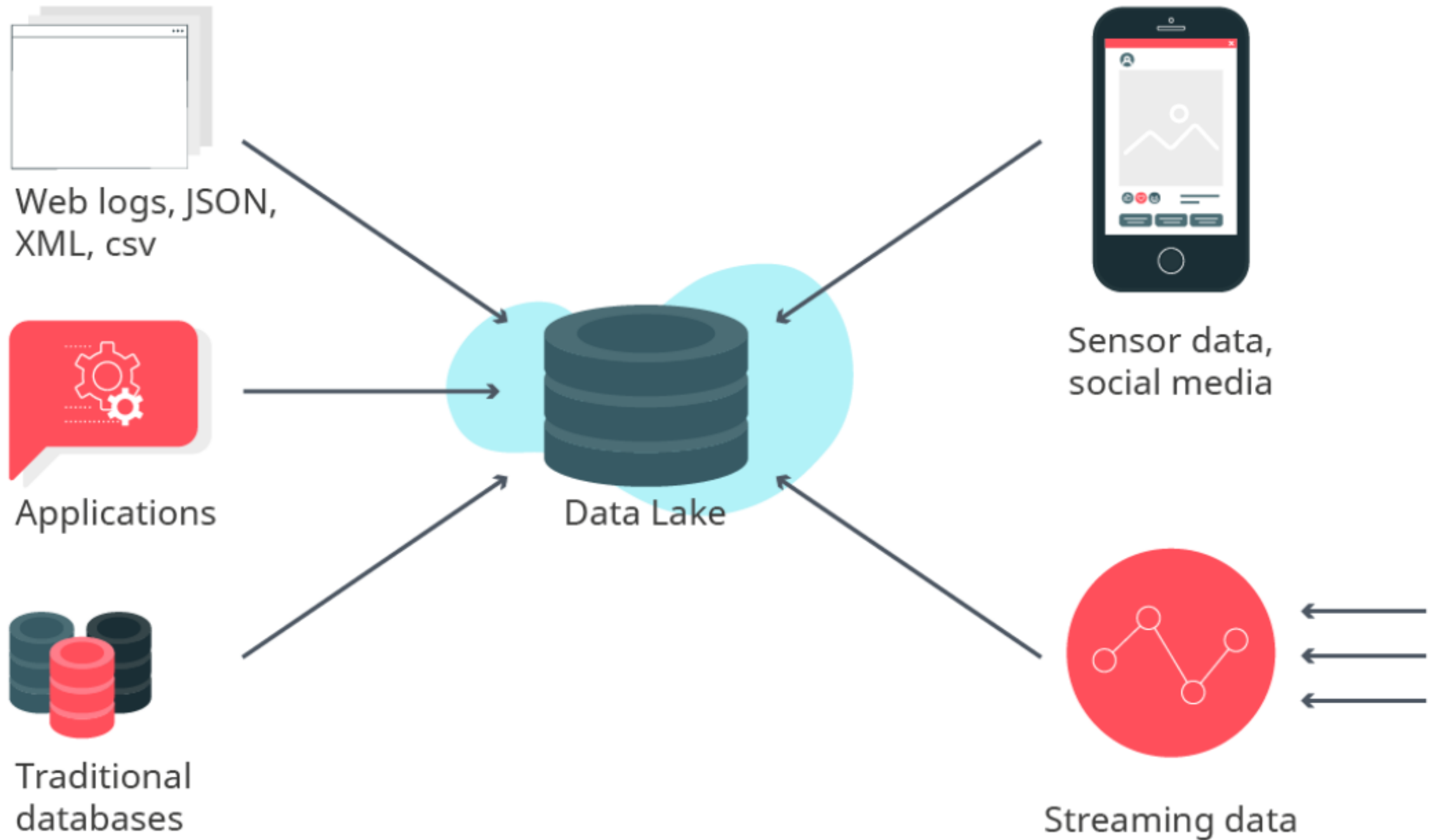


Core Data Lake Capabilities





DATA LAKE SOURCES





Machine
Learning



Analytics



On-premises
Data Movement



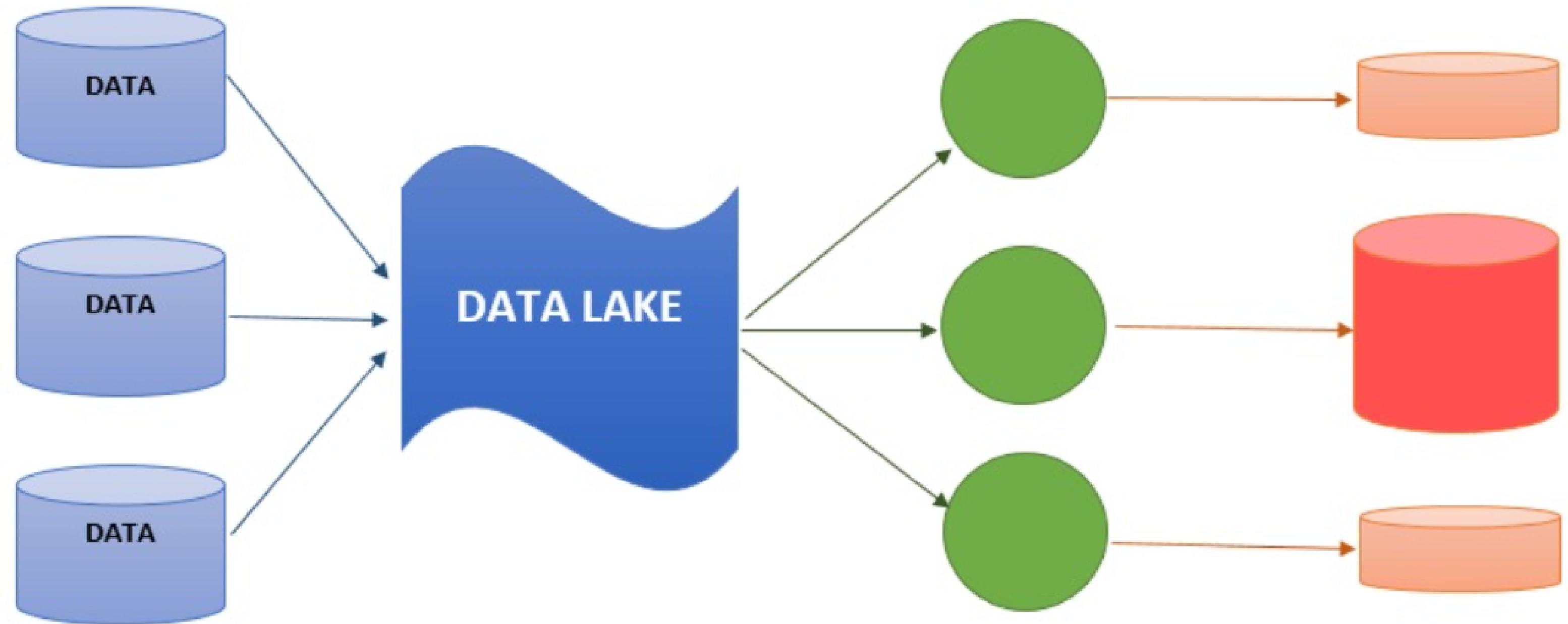
Real-time Data
Movement

Data Sources

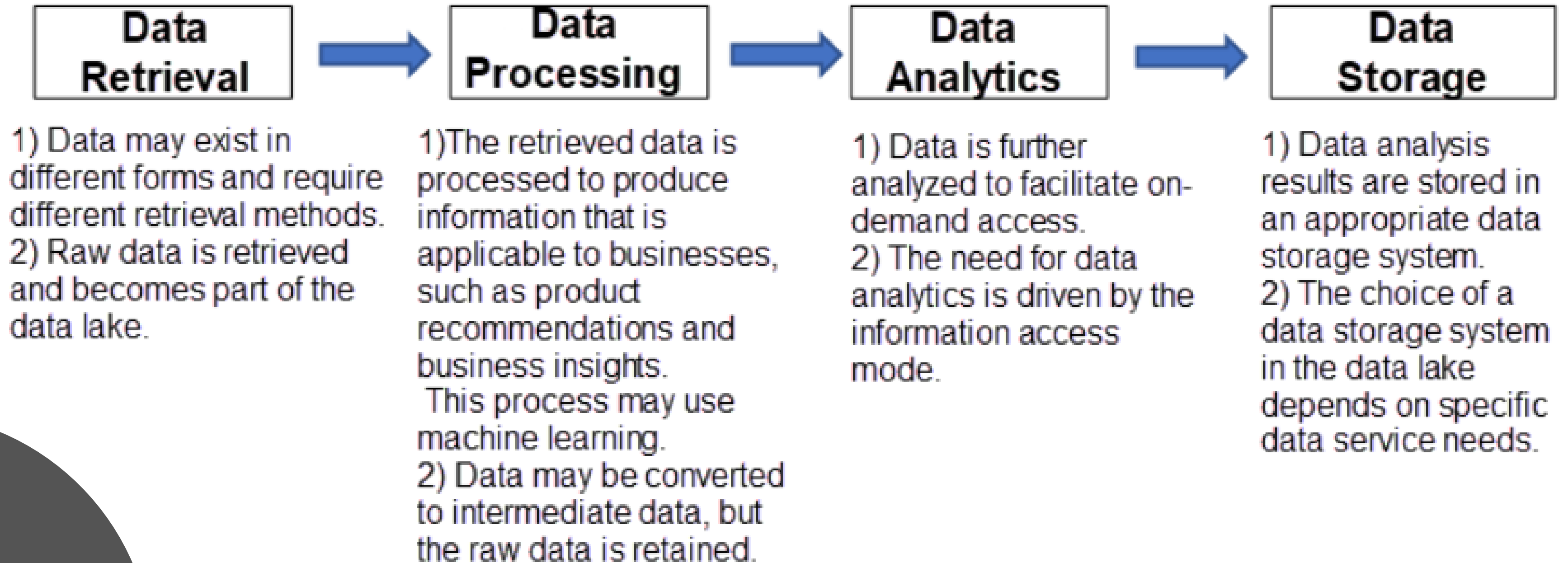
Data in raw format

Data transforms

Data ready for queries



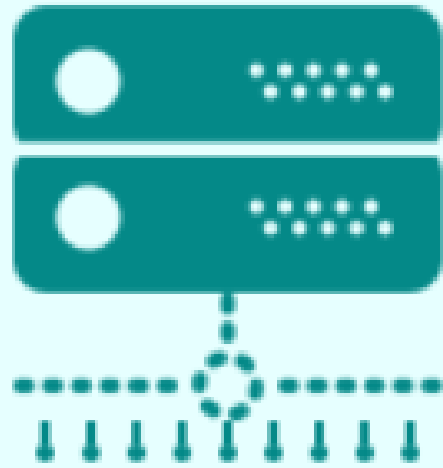
DATA LIFECYCLE IN A DATA LAKE



DATA WAREHOUSE

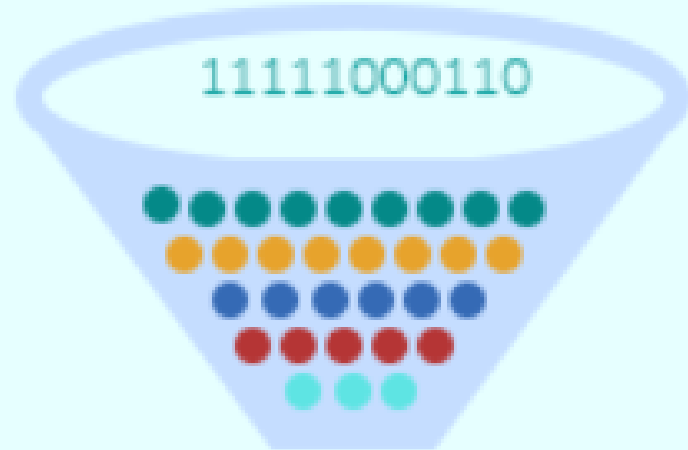
VS

DATA LAKE

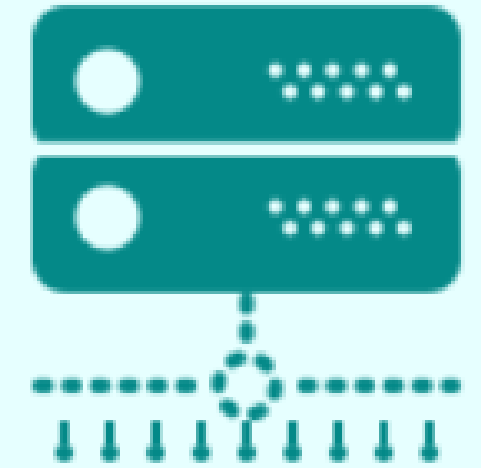


1110001101110
011011000110
11111000110

Data is processed and organized into a single schema before being put into the warehouse



The analysis is done on the cleansed data in the warehouse



1110001101110
011011000110
11111000110

Raw and unstructured data goes into a data lake



Data is selected and organized as and when needed



Data warehouse vs. data lake

	DATA WAREHOUSE	DATA LAKE
DATA TYPES	Structured, processed data from operational databases, applications and transactional systems	Structured, semistructured and unstructured data from sensors, apps, websites, etc.
PURPOSE	Predefined purpose for business intelligence, batch reporting and data visualization	May not have a predefined purpose; typically used for machine learning, deep analysis and discovery
USERS	Data engineers, business analysts, data analysts	Data engineers, data scientists

COMMERCIALY AVAILABLE DATA LAKE SOLUTIONS

- 1 AWS's Data Lake Solution
- 2 Huawei's Data Lake Solution
- 3 Alibaba Cloud's Data Lake Solution
- 4 Azure's Data Lake Solution

The background is a solid light pink color. It features several abstract geometric elements: a large, semi-transparent light pink circle on the left side; a smaller, semi-transparent light pink circle on the right side; a thick pink diagonal line segment in the top right corner; and another thick pink diagonal line segment in the bottom left corner. The text 'THANK YOU!!' is centered in a bold, black, sans-serif font.

THANK YOU!!

Made with ♥ by Vaibhav