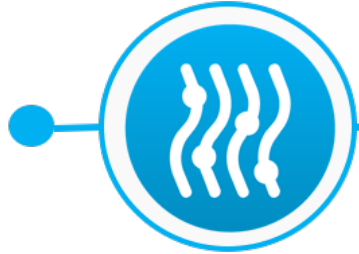


INGESTION & STORAGE



Real Time (data stream):

- Automated scalability
- Real time monitoring
- pay as you go

Batch files

(transactions, log-files):

- Management of large data quantities
- Data Governance (ingestion rules)
- Secure (access to data is controlled), scalable, redundant, long-term storage
- Monitoring

PROCESSING



Data Manipulation:

- Conversion of data into optimized formats
- Entities extraction and parsing

Data Enrichment:

- Matching of multiple data sources
- Inference and Machine Learning

ANALYSIS



Real time:

To analyse specific data sets quickly and return information in real time

Distributed:

To analyse large quantities of data in batch (leverages Hadoop)

Data exploration:

To explore large quantities of data through a number of tools, such as Hive, MySQL, Athena, Tableau

INTEGRATION



API:

To transfer information to external platforms / Apps

Data Visualization:

To analyse and visualize data through reports

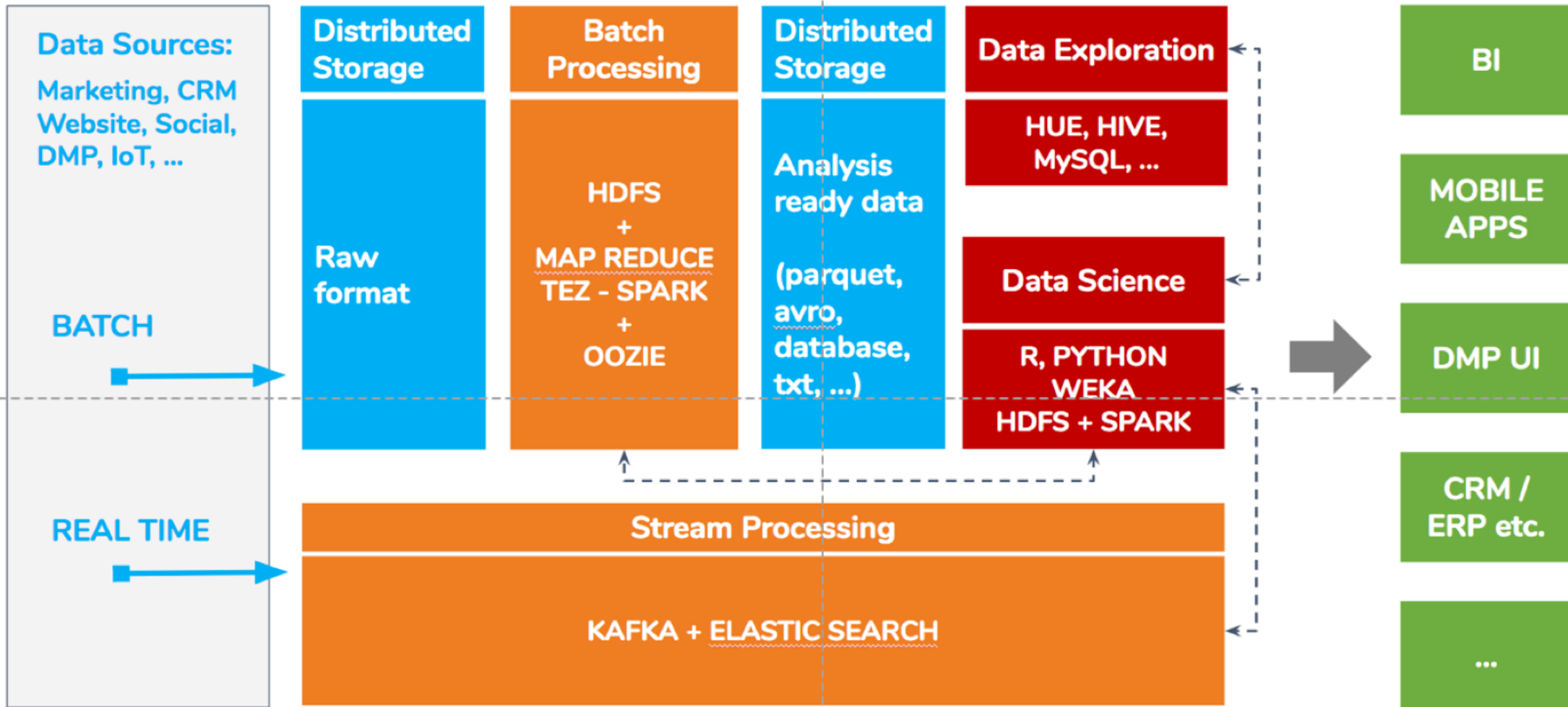
Data Extraction:

To extract users lists

Prototype Analysis:

To develop specific applications

DATA LAKE – PLATFORM ARCHITECTURE



LEGENDA

- INGESTION & STORAGE
- PROCESSING
- ANALYSIS
- INTEGRATION