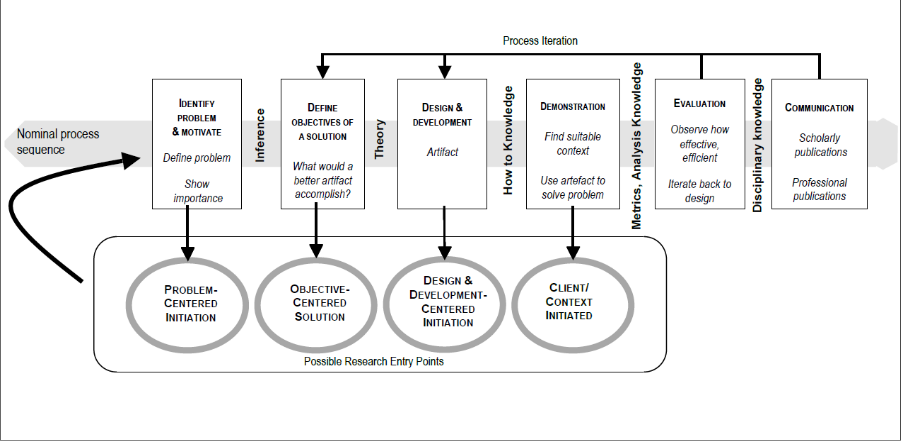
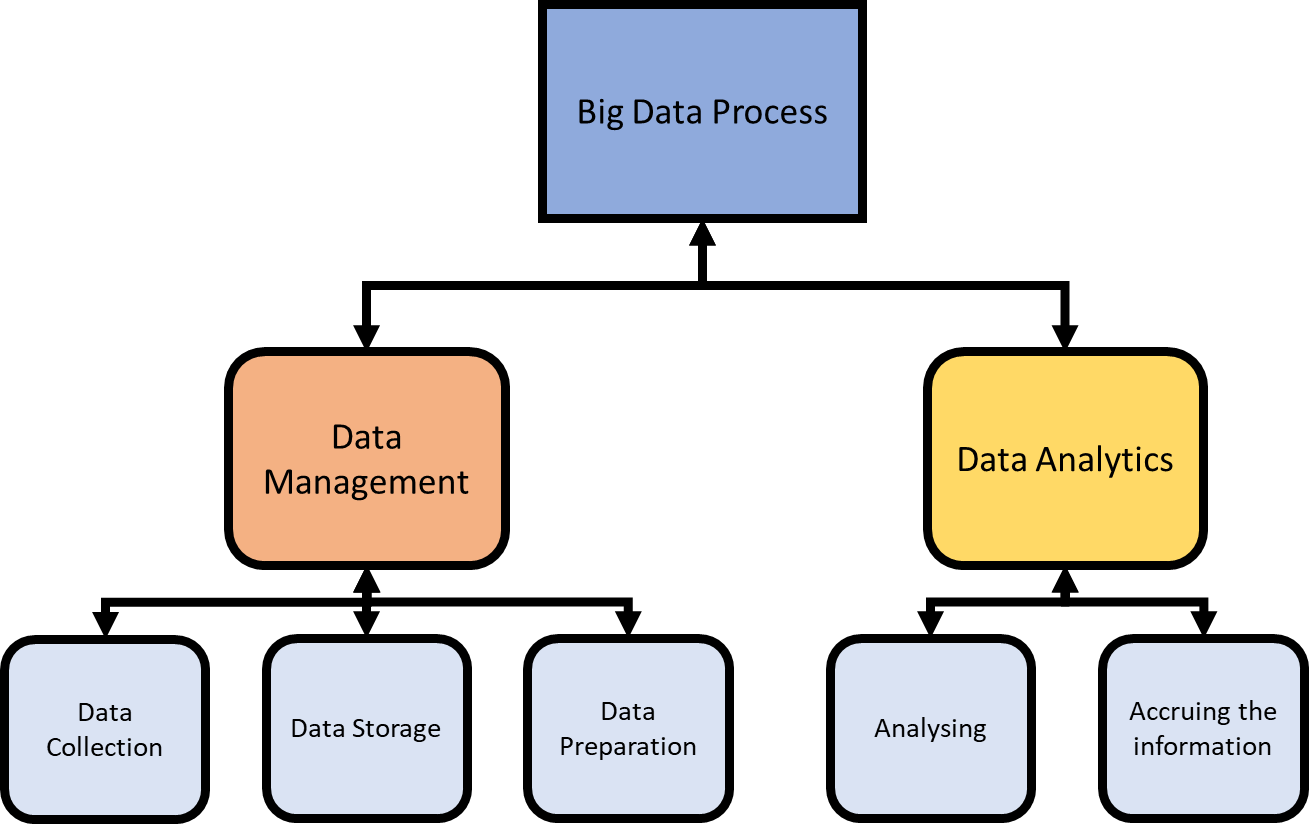
# ANNEX A – FRAMEWORKS USED IN THE RESEARCH

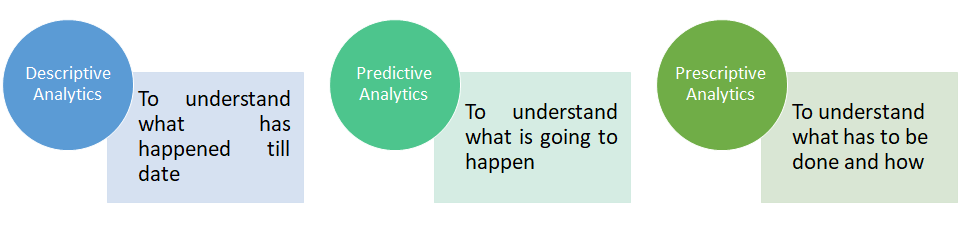
Design science research methodology (Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007, p. 54)



Data Analytics processes (Gandomi & Haider, 2015)



Types of data analytics (Sivarajah, Kamal, Irani, & Weerakkody, 2017)



Three cycle view of design science research (Hevner A. , 2007)

A close up of text on a white background

Description generated with high confidence

Reverse engineering process (Bhandari, 2014)

A close up of text on a white background

Description generated with high confidence

Black box reverse engineering (Lorenz , 2008)

A screenshot of a cell phone

Description generated with very high confidence

Reverse engineering process (Reverse-Engineering-Process-Assignment-Help, n.d.)

A screenshot of a cell phone

Description generated with very high confidence

De-compilation and Reverse engineering (Emmerik, n.d.)

A close up of a map

Description generated with high confidence

Model Driven reverse engineering (Nantes, 2011)

A screenshot of a cell phone

Description generated with very high confidence

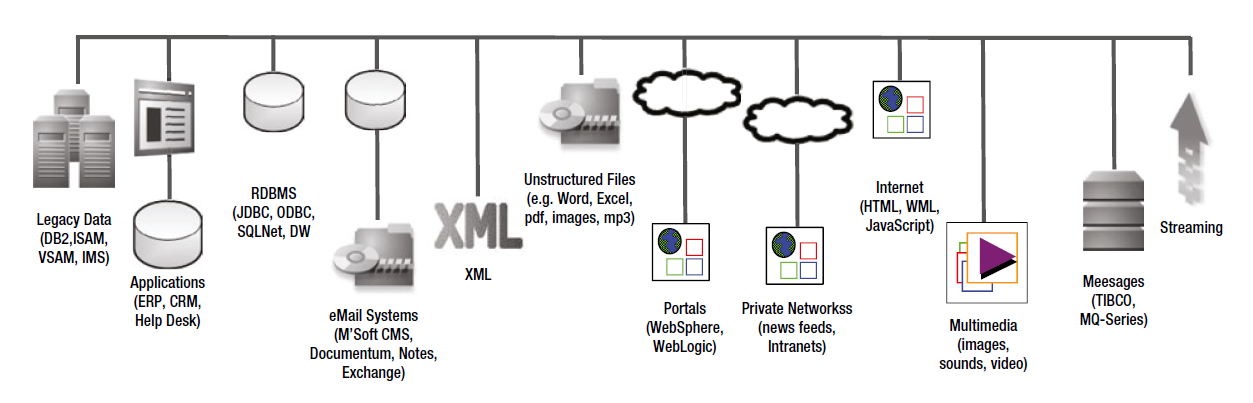
Framework for a reverse engineering process (Ajila, 2002)

A screenshot of a cell phone

Description generated with very high confidence

# ANNEX B – Artefacts used for the RE-DSR.

Multiple data types are shown below.



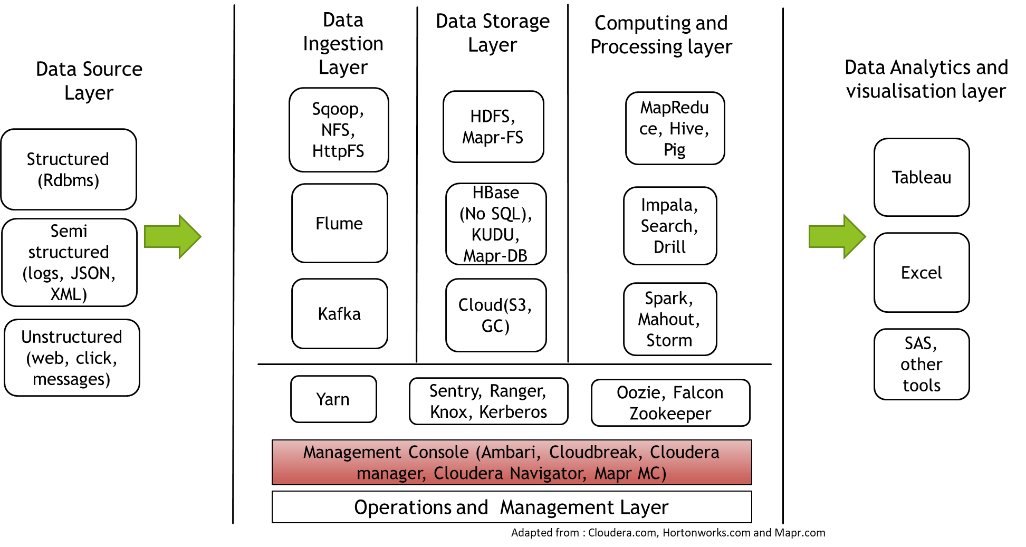
Variety of Data types (Sourced by (Sawant & Shah, 2013) )

Design Theory diagrams

Flexible Hadoop components architecture. Source: (White, 2015) (Cloudera, n.d.) (Hortonworks, n.d.) (Mapr, n.d.)

****

Reverse engineered Hadoop big data platform. Source: (Cloudera, n.d.) (Hortonworks, n.d.)



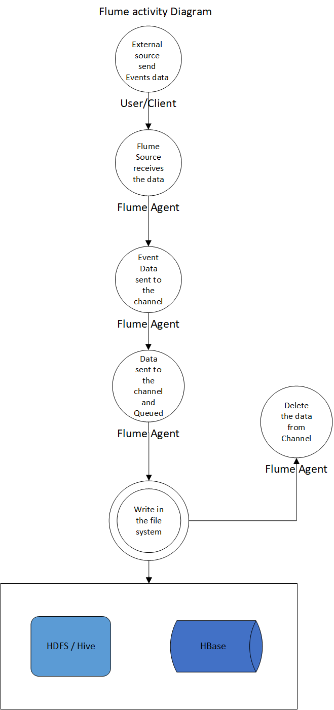
Sqoop activity and process diagram. Source: (White, 2015)

A close up of a map

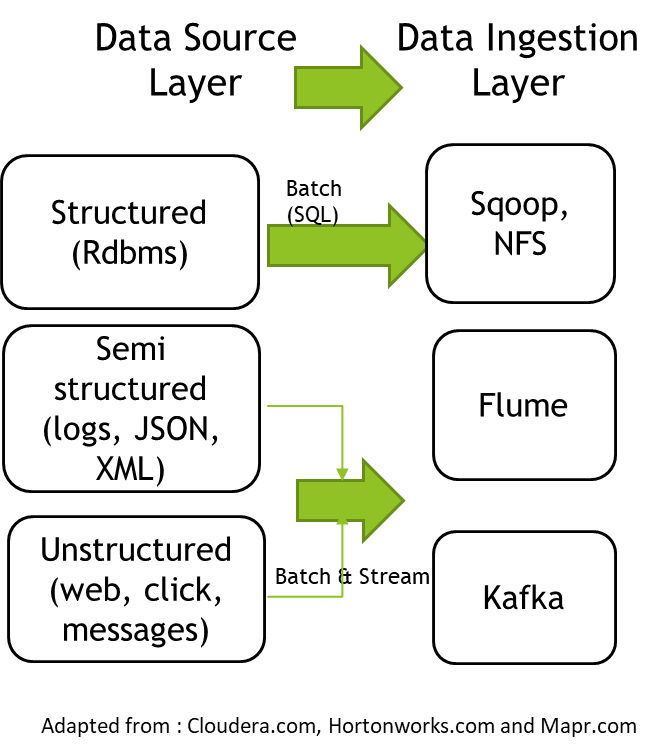
Description automatically generated*A screenshot of a cell phone

Description automatically generated*

Flume activity diagram. Source: (White, 2015)

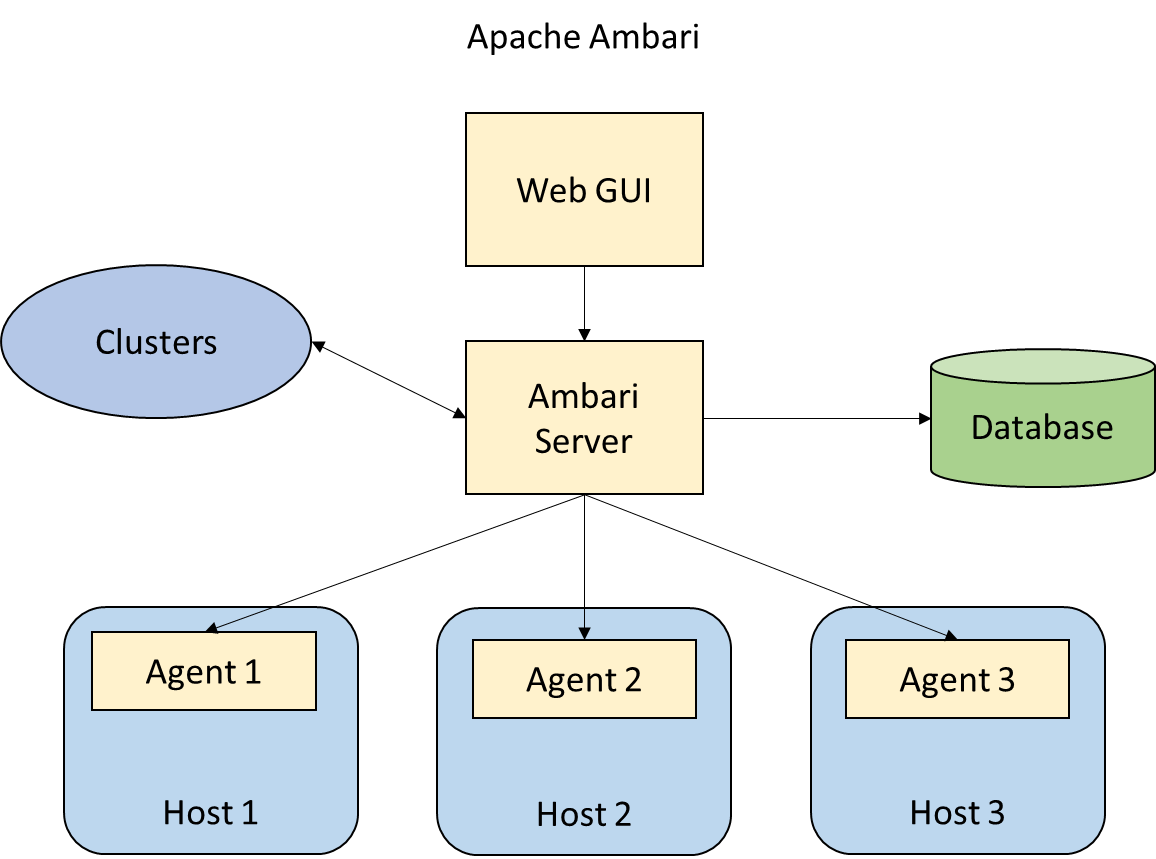
**

Hadoop ingestion components. Source: (Cloudera, n.d.) (Hortonworks, n.d.) (Mapr, n.d.)



Cluster management architecture – Hortonworks. Source: (Hortonworks, n.d.)

A picture containing map, text

Description automatically generated 

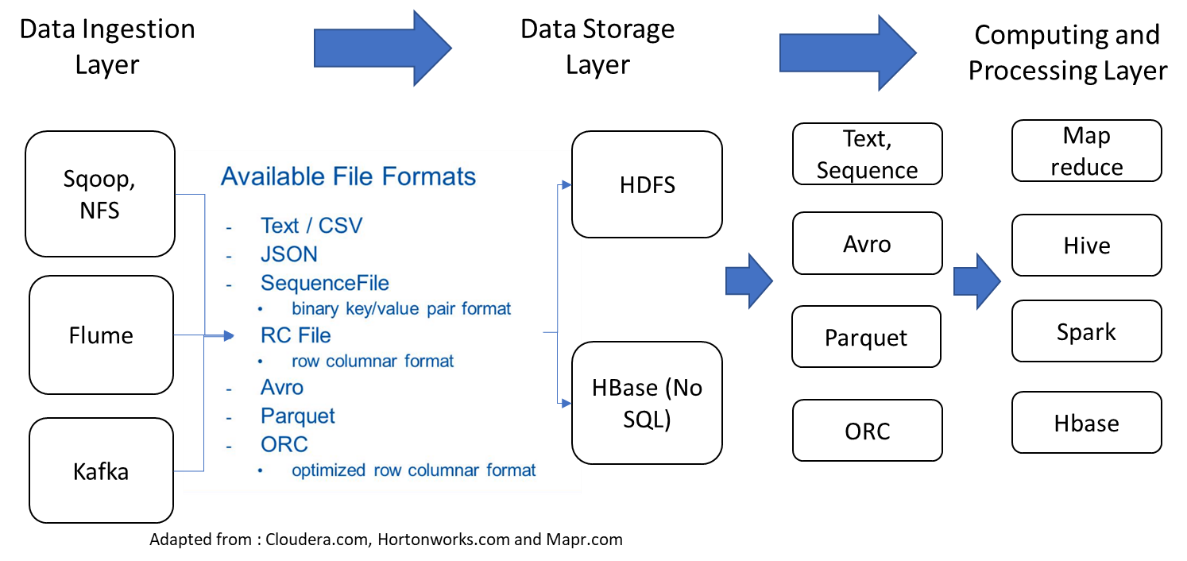
HDFS metadata management. Source: (Apache Hadoop, n.d.) (White, 2015)

A close up of a logo

Description automatically generatedA screenshot of a social media post

Description automatically generated

Hadoop architecture based on File formats. Source: (PLASE, 2017) (White, 2015) (Cloudera, n.d.) (Hortonworks, n.d.)



Hadoop replication architecture. Source: (Apache Hadoop, n.d.)

A picture containing screenshot

Description automatically generated

Hadoop sequence and MapReduce activity diagram – representing data locality. Source: (Apache Hadoop, n.d.)

A close up of a map

Description automatically generatedA screenshot of a computer

Description automatically generated

Hadoop sequence diagram – Distributed and Parallel Processing. Source: (Apache Hadoop, n.d.)

A screenshot of a computer

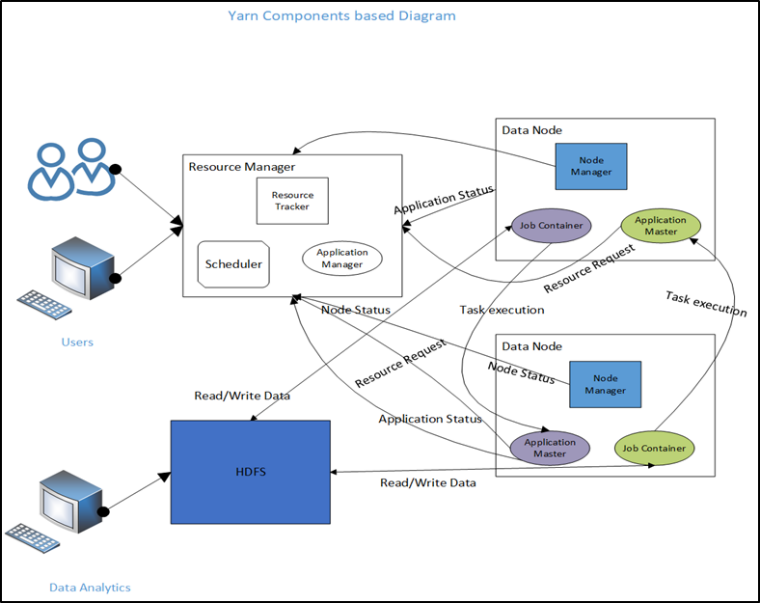
Description automatically generated

HDFS sequence diagram - Data stored in a distributed environment. Source: (White, 2015)

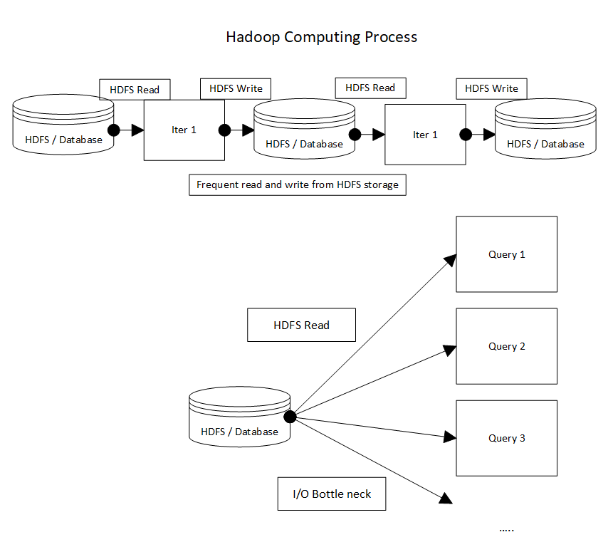
A screenshot of a social media post

Description automatically generated

Yarn components diagram – Multiple tools(applications) execution in parallel. (Apache Hadoop, n.d.)



Disk-based processing – Hadoop. Source: (Zhang, Chen, Ooi, Tan, & Zhang, 2015)

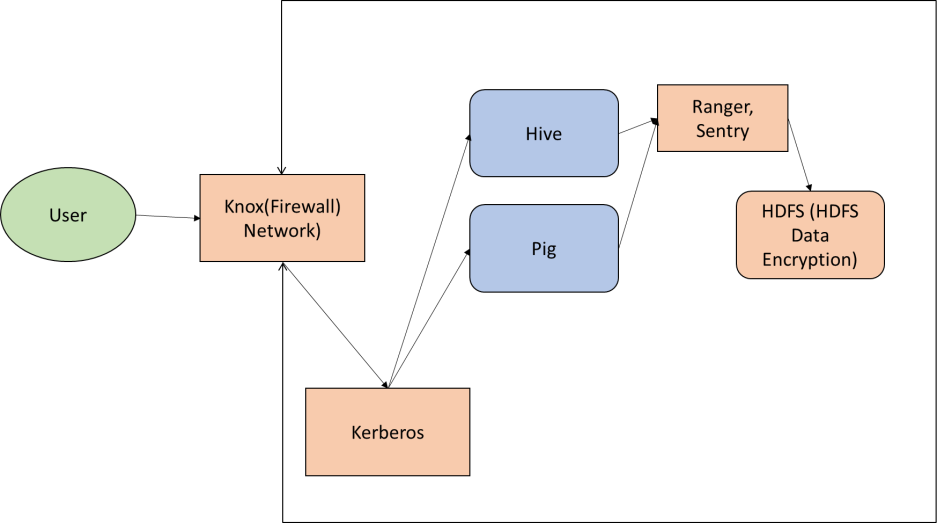


In-memory computation – Spark. Source: (Zhang, Chen, Ooi, Tan, & Zhang, 2015) (White, 2015)

A screenshot of a cell phone

Description automatically generated

Security architecture of Hadoop. Source: (Hortonworks, n.d.)



Cloudera platform management architecture. Source: (Cloudera, n.d.)

