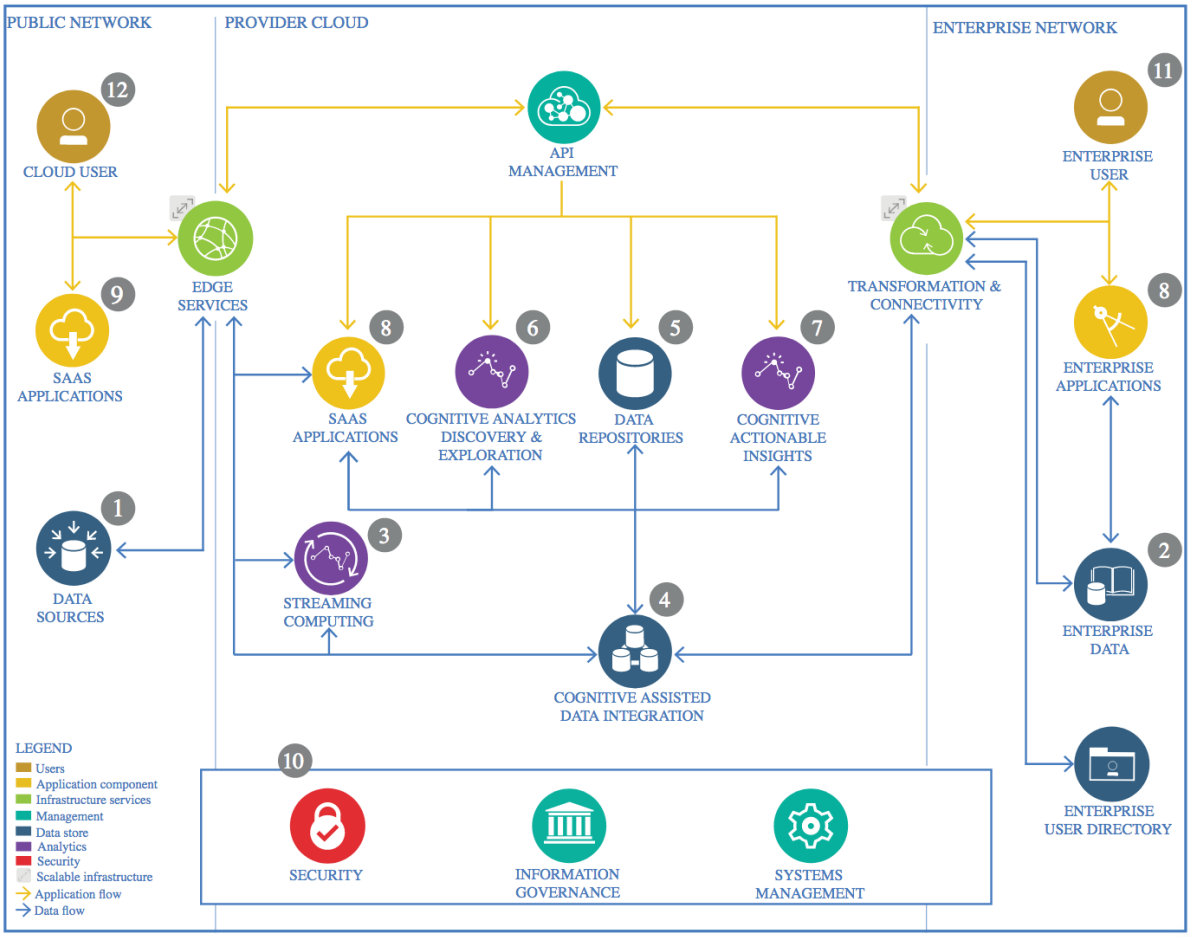
Architectural Decisions Document – Project MCA-IBM-Capstone

Architectural Decisions Document Template

1 Architectural Components Overview



IBM Data and Analytics Reference Architecture. Source: IBM Corporation

1.1 Data Source

1.1.1 Technology Choice

We used a CSV file provided by the IBM use case for heart failure prediction

1.1.2 Justification

The CSV file is an easily implementable format directly usable for data transformation and application to machine learning.

1.2 Enterprise Data

1.2.1 Technology Choice

There is no Enterprise data for this project

1.2.2 Justification

Please justify your technology choices here.

1.3 Streaming analytics

1.3.1 Technology Choice

There is no streaming process in our application.

1.3.2 Justification

Please justify your technology choices here.

1.4 Data Integration

1.4.1 Technology Choice

We integrate the data by importing and reading the data file in our Jupyter notebook

1.4.2 Justification

This process provides an easy implementation of the data and the project

1.5 Data Repository

1.5.1 Technology Choice

The data will be stored locally
We developed another version accessible via a web application to do your own prediction.

1.5.2 Justification

The local implication will provide an easy to use implementation

1.6 Discovery and Exploration

1.6.1 Technology Choice

After importing the data we look at the structure (columns, information) and data type to be able to modify them for data application
We also look at the correlation between the different data

1.6.2 Justification

The CSV file nature of the initial data makes it easy to integrate as a dataframe and modify accordingly

1.7 Actionable Insights

1.7.1 Technology Choice

We use feature engineering to modify the data and make them appropriate for machine learning

1.7.2 Justification

Feature engineering will improve the results of our machine learning process.

1.8 Applications / Data Products

1.8.1 Technology Choice

Using the dataframe structure and python libraries for machine learning, plotting and scoring we'll analyze the data and the models created.

1.8.2 Justification

Python provides very useful and powerful tools to allow us to do the machine learning and get the best results

1.9 Security, Information Governance and Systems Management

1.9.1 Technology Choice

NA

1.9.2 Justification

NA