The Lightweight IBM Cloud Garage Method for Data Science

Architectural Decisions Document Template

# Architectural Components Overview



IBM Data and Analytics Reference Architecture. Source: IBM Corporation

## Data Source

### Technology Choice

Please describe what technology you have defined here. Please justify below, why. In case this component is not needed justify below.

UCI ML Repository (<https://archive.ics.uci.edu/ml/index.php>)

### Justification

Please justify your technology choices here.

This data was publicly available for research and was high quality.

## Enterprise Data

### Technology Choice

Please describe what technology you have defined here. Please justify below, why. In case this component is not needed justify below.

GitHub

### Justification

Please justify your technology choices here.

Easy to upload changed via GitHub and sync across devices.

## Streaming analytics

### Technology Choice

Please describe what technology you have defined here. Please justify below, why. In case this component is not needed justify below.

Not applicable

### Justification

Please justify your technology choices here.

Not Applicable

## Data Integration

### Technology Choice

Please describe what technology you have defined here. Please justify below, why. In case this component is not needed justify below.

Not applicable

### Justification

Please justify your technology choices here.

Not applicable

## Data Repository

### Technology Choice

Please describe what technology you have defined here. Please justify below, why. In case this component is not needed justify below.

IBM Watson Studio

### Justification

Please justify your technology choices here.

Familiarity and ease-of-use.

## Discovery and Exploration

### Technology Choice

Please describe what technology you have defined here. Please justify below, why. In case this component is not needed justify below.

I used Python libraries Seaborn, Matplotlib and Pandas for data exploration and visualization.

### Justification

Please justify your technology choices here.

I did not need to use Spark as the dataset was small and did not need additional worker nodes to reduce processing time.

## Actionable Insights

### Technology Choice

Please describe what technology you have defined here. Please justify below, why. In case this component is not needed justify below.

Not applicable

### Justification

Please justify your technology choices here.

Not applicable

## Applications / Data Products

### Technology Choice

Please describe what technology you have defined here. Please justify below, why. In case this component is not needed justify below.

I used IBM Watson Studio, and Jupyter Notebooks.

### Justification

Please justify your technology choices here.

Ease-of-use and familiarity as we used this throughout the course.

## Security, Information Governance and Systems Management

### Technology Choice

Please describe what technology you have defined here. Please justify below, why. In case this component is not needed justify below.

Not applicable

### Justification

Please justify your technology choices here.

Not applicable