IBM Cloud & Cognitive Professional Certification Program

Study Guide Series

Exam C1000-136 IBM Cloud Pak for Data v4.x Solution Architecture

Contents

Purpose of Exam Objectives

When an exam is developed, Subject Matter Experts work together to define the role the certified individual will fill. They define the tasks and knowledge that an individual would need to successfully perform this job role doe the product or solution. This creates the foundation for the objectives and measurement criteria, which form the basis of the certification exam. Question writers then use these objectives to develop exam questions.

It is recommended that you review these objectives and ask yourself the following questions:

- Do you know how to complete the task in the objective?
- Do you know why that task needs to be done?
- Do you know what will happen if you do it incorrectly?

If you are not familiar with a task, go through the objective, perform that task in your own environment and read more information on the task. If there is an objective on a task, there is a high likelihood that you WILL see a question about it on the actual exam. Review the recommended learning designed to prepare you to take the certification exam.
After reviewing the objectives in this guide and completing your own research, take the assessment exam. While the assessment exam does not indicate which specific questions were answered incorrectly, it does indicate overall performance by section. This is a good indicator of preparedness or if further preparation is warranted.

High-level Exam Objectives

Section 1 - Cloud Pak for Data Architecture	
1.1	Understand the underlying infrastructure and installation
1.2	Understand sizing and deployment options
1.3	Understand Cloud Pak for Data reference architecture
1.4	Understand Cloud Pak for Data reliability options
1.5	Secure the solution and client data
Section 2 - Build Data Science algorithms	
2.1	Describe the differences between traditional programming and machine learning as well as no code, low code and visual programming
2.2	Understand the process and available tools in Cloud Pak for data to build, deploy and monitor ML/AI algorithms
2.3	Identify and explain how to collect, explore, and prepare data for ML/AI algorithms in Cloud Pak for Data
2.4	Identify and explain how to build and deploy ML/AI algorithms in Cloud Pak for data
2.5	Map business opportunities into a Data Science use case
Section 3 - Machine Learning Operations	
3.1	Describe the key considerations when selecting a platform for model deployment
3.2	Explain the workflow of deploying and monitoring models
3.3	Monitor deployed models inside Cloud Pak for Data
3.4	Monitor machine learning models running on an external platform
3.5	Manage risk and regulatory compliance using Open Pages
Section 4 - Analytics	
4.1	Explain the difference between Descriptive, Prescriptive, Predictive, Diagnostic, and Cognitive Analytics
4.2	Describe the capabilities of AI for financial operations in Cloud Pak for Data
4.3	Describe the capabilities of business intelligence in Cloud Pak for Data
4.4	Map business opportunities into an Analytics use case
Section 5 -Data Governance	
5.1	Describe the capabilities of a data fabric topology
5.2	Define the Governance structure
5.3	Use smart ingestion for auto cataloging
5.4	Understand how workflow is used in Cloud Pak for Data
5.5	Explain the use of Guardium in auditing and monitoring data
5.6	Leverage the platform to understand data flow and usage
5.7	Explain the concepts of Knowledge Accelerators
5.8	Map business opportunities into a data governance use case
Section 6 - Integration, Implementation, Deployment, and Scaling	

6.1	Develop an appropriate process to take a Data and AI solution from inception to production
6.2	Develop a strategy to monitor the Data and AI platform
6.3	Accelerate the solution using Industry Accelerators and External Data sets
6.4	Integrating Business Applications using Cloud Pak for Data

Detailed Exam Objectives

Section 1 - Cloud Pak for Data Architecture

1.1. Understand the underlying infrastructure and installation.

SUBTASKS:

1.1.1. Describe the features in Cloud Pak Foundational Services
1.1.2. Describe the value of running on Redhat OpenShift
1.1.3. Describe the benefits of modernization
1.1.4. Describe how CPD can Simplify lifecycle management (upgrades/fixpaks)
1.1.5. List the benefits of/use cases for multitenancy
1.1.6. Describe the differences between Cloud Pak as a Service and Cloud Pak v4.0
1.1.7. Describe advantages of using Operators
1.1.8. Understand the installation procedure for Cloud Pak

REFERENCES:

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=planning-architecture
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=planning-cloud-deployment-environments
https://www.ibm.com/products/cloud-pak-for-data
https://www.ibm.com/cloud/paks
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=overview-whats-new
https://dataplatform.cloud.ibm.com/docs/content/wsj/getting-started/compare-cpd.html
https://catalog.redhat.com/software/operators/detail/5e9874803f398525a0ceb00a

1.2. Understand sizing and deployment options.

SUBTASKS:

1.2.1. Demonstrate an understanding of the Cloud Pak deployment options.

REFERENCES:

https://www.ibm.com/products/cloud-pak-for-data/deployment-model-options

1.2.2. Explain the components of sizing Cloud Pak for Data including hardware, software and storage

REFERENCES:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=requirements-hardware
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=requirements-storage
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=requirements-software

1.3. Understand Cloud Pak for Data reference architecture.

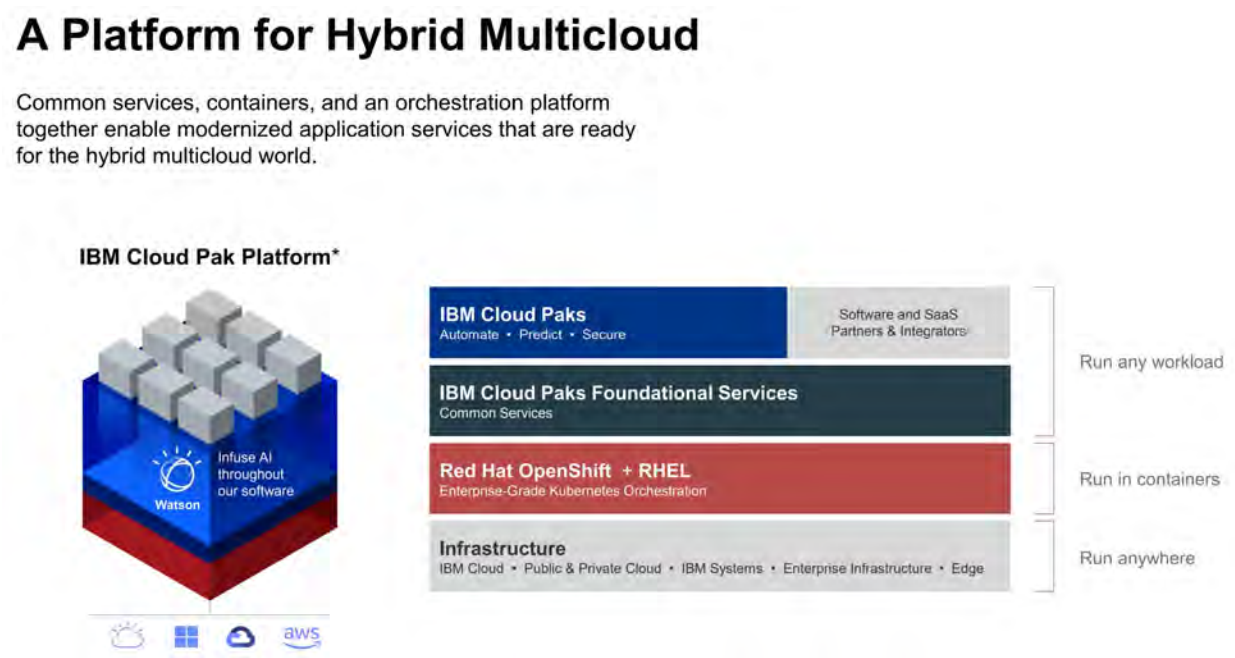
SUBTASKS:

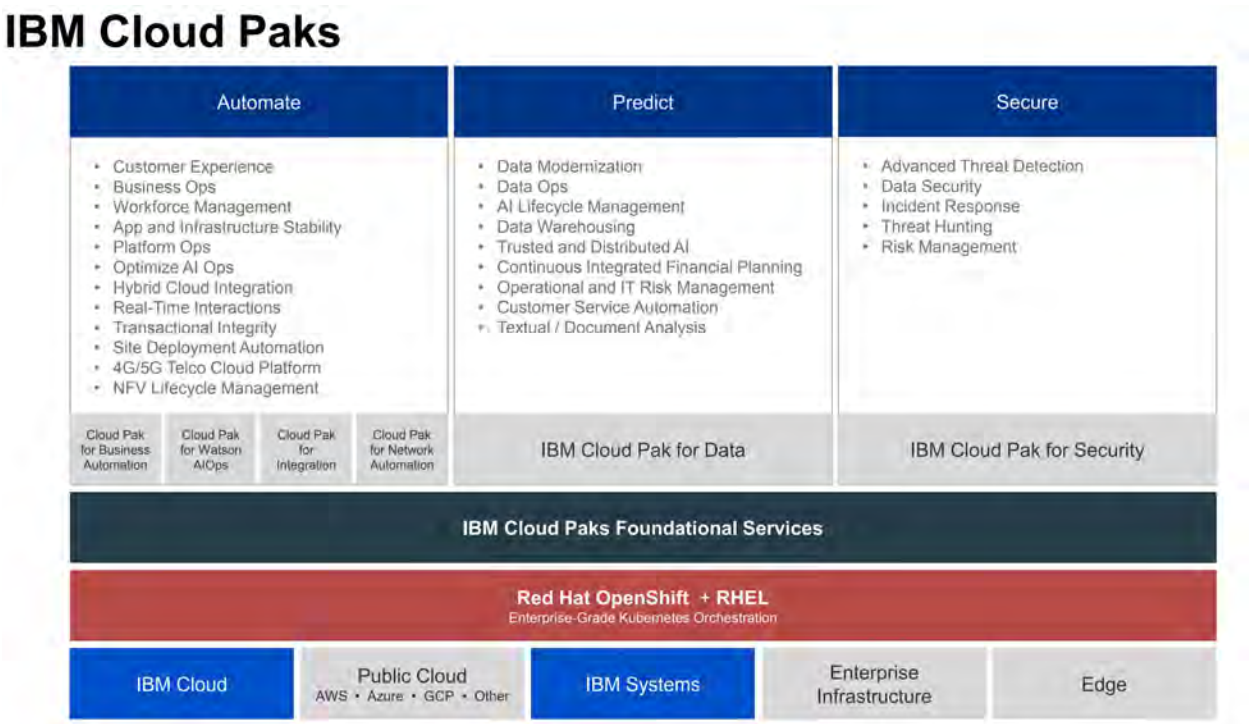
1.3.1. Demonstrate an understanding of the Cloud Pak for data architecture.

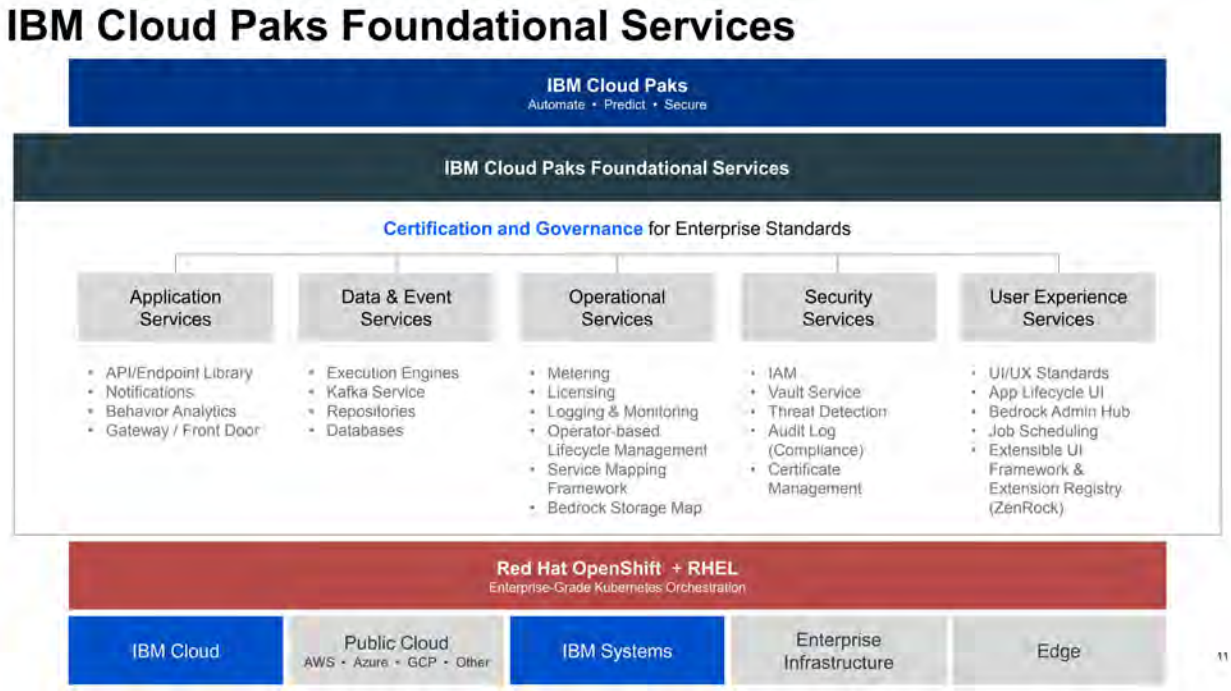
REFERENCES:
https://w3.ibm.com/w3publisher/ibm-data-and-ai-sales/what-we-offer/lead-with-offerings/cloud-pak-for-data
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=planning-architecture

1.3.2. Explain the Cloud Foundational services

REFERENCES:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=planning-system-requirements
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=tasks-installing-cloud-pak-foundational-services







1.3.3. Explain the Control Plane & Services

REFERENCES:

https://www.ibm.com/uk-en/products/cloud-pak-for-data?p1=Search&p4=43700052283971652&p5=e&gclid=CjwKCAjwsNiIBhBdEiwAJK4k

hpm7xkxt-g1wkVvB7BExiyD-X4D5j9vV75ErGy3TEs5U-w6IVFxrjhoC9rQQAvD_BwE&gclsrc=aw.ds
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=integrations-services

1.3.3.1. Services are grouped together based on their function

AI
Analytics
Dashboards
Data governance
Data sources
Developer tools
Industry solutions
Storage

1.3.3.2. Multi-tenancy

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=planning-multitenancy-support

1.3.4. Explain projects, namespaces and tethered projects

REFERENCES:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=planning-architecture
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=tasks-creating-projects-namespaces

1.4. Understand Cloud Pak for Data reliability options.
SUBTASKS:

1.4.1. Explain how Cloud Pak for Data uses OpenShift to achieve service availability.
1.4.2. Understand the backup and restore strategy.

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=cluster-adding-nodes-your

1.4.3. Describe how to scale up/downservices

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=cluster-scaling-services

1.4.4. Indentify when you would need to add Nodes

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=cluster-scaling-services

1.4.5. Demonstrate an understanding of the Cloud Pak deployment options

REFERENCE:
https://www.ibm.com/products/cloud-pak-for-data/deployment-model-options

1.4.6. Explain the components of sizing Cloud Pak for Data including hardware, software and storage

1.5. Secure the solution and client data.
SUBTASKS:

1.5.1. Explain security for users available on CPD
1.5.2. Differentiate between authentication and authorization on Cloud Pak for Data
1.5.3. Explain encryption on CPD
1.5.4. Describe Network security on CPD
1.5.5. Explain how Red Hat Openshift Container Platform provides security for CPD
1.5.6. Describe Security and Auditing considerations
1.5.7. Describe SSO, IAM and LDAP use in user management

REFERENCES:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=overview-regulatory-compliance
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=planning-security-considerations
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=considerations-basic-security-features
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=considerations-multitenancy-network-security
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=considerations-auditing-cloud-pak-data
https://www.ibm.com/support/producthub/icpdata/docs/content/SSQNUZ_latest/cpd/admin/ldap.html
https://www.ibm.com/support/producthub/icpdata/docs/content/SSQNUZ_latest/cpd/admin/users.html

Section 2 – Build Data Science algorithms

2.1. Describe the differences between traditional programming and machine learning as well as no code, low code and visual programming.
SUBTASKS:

2.1.1. Explain the differences between traditional programming and machine learning.

REFERENCES:
https://productcoalition.com/difference-between-traditional-programming-versus-machine-learning-from-a-pm-perspective-3802b02bc7f6

https://medium.com/mlearning-ai/what-is-the-difference-between-traditional-programming-and-machine-learning-f6128ed4f595

2.1.2. Describe visual programming.

REFERENCES:
https://www.outsystems.com/blog/posts/what-is-visual-programming/
https://en.wikipedia.org/wiki/Visual_programming_language

2.1.3. Describe what is meant by no code and low code and their differences

REFERENCES:
https://en.wikipedia.org/wiki/No-code_development_platform
https://en.wikipedia.org/wiki/Low-code_development_platform
https://www.bettyblocks.com/no-code-low-code-application-development
https://www.bettyblocks.com/low-code-application-development

2.1.4. Consider some of the examples of Cloud Pak for data tools that are no code, low code and programming tools
2.1.4.1. Auto AI
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=models-autoai

2.1.4.2. (SPSS) modeler Canvas

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=services-spss-modeler

2.1.4.3. Python Notebooks in JupyterLabs

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=models-jupyterlab

2.1.4.4. Decision Optimization.

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=services-decision-optimization

2.2. Understand the process and available tools in Cloud Pak for data to build, deploy and monitor ML/AI algorithms.
SUBTASKS:

2.2.1. Explain the process of Model Ops and Tools in Cloud Pak for data projects are used to develop AI/ML algorithms.

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=cases-modelops-use-case

2.2.2. Describe the available tools to choose when developing the models.

REFERENCE
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=projects-choosing-tool

2.2.2.1. Auto AI
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=models-autoai

2.2.2.2. Modeler Canvas
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=services-spss-modeler

2.2.2.3. Python notebooks and Jupyter labs
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=models-notebooks
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=models-jupyterlab

2.2.2.4. R in R studio
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=models-rstudio

2.2.2.5. Decision Optimization – CPLEX
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=services-decision-optimization

2.2.3. Explain the how Cloud Pak for data projects in Watson Studio are used to develop AI/ML algorithms.

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=analyzing-data-building-models

2.2.3.1. Explain integration of GitHub with CP4D projects

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=projects-git-integration

2.3. Identify and explain how to collect, explore, and prepare data for ML/AI algorithms in Cloud Pak for Data.
SUBTASKS:

2.3.1. Understand the Collect process for ML/AI algorithm development in Cloud Pak for Data

REFERENCES:

1. Data Virtulization
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=data-virtualizing
2. Watson Knowledge Catalog
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=services-watson-knowledge-catalog

3. Data Refinery

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=services-data-refinery

4. Data Stage

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=data-creating-transformation-job

5. Connections

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=data-connecting-sources

6. Data Management Console

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=database-using-db2-data-management-console#work-with-dmc__monitor

7. Watson Studio

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=data-adding-analytics-project

2.3.2. Explain the Explore process for ML/AI algorithm development in Cloud Pak for Data

2.3.2.1. Data Refinery

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=data-refining

2.3.2.2. Cognos Dashboards

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=services-cognos-dashboards
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=analytics-visualizing-data-cognos-dashboards

2.3.3. Explain the Prepare data process for ML/AI algorithm development in Cloud Pak for Data

REFERENCES:

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=data-refining
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=data-transforming

2.4. Identify and explain how to build and deploy ML/AI algorithms in Cloud Pak for data.

SUBTASK(S):

2.4.1. Explain in more detail some of the development tools in Cloud Pak for data for building ML/AI Algorithms?

2.4.1.1. Watson Studio

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=services-watson-studio

2.4.1.2. Integrated, Supplemental services

Decision Optimization	Find the most appropriate prescriptive solutions to your business problems by using CPLEX optimization engines to evaluate millions of possibilities.
Jupyter Notebooks with Python 3.7 with GPU	Access compute environments for Jupyter notebooks that use GPU-accelerated Python 3.6 libraries.
Execution Engine for Apache Hadoop	Integrate the Watson Studio service with your remote Apache Hadoop cluster so you can explore data and build and deploy models on your remote cluster.
Jupyter Notebooks with R 3.6	Access compute environments to create Jupyter Notebooks that use R 3.6 libraries.
RStudio Server with R 3.6	Access the RStudio IDE.
SPSS Modeler	Create flows to prepare data, develop and manage models, and visualize data. No coding required.
Watson Machine Learning	Build, train, and deploy machine learning models with a full range of tools.

2.4.1.3. Integrated Related services

Cognos Dashboard	Identify patterns in your data with sophisticated visualizations. No coding needed.
Data Virtualization	Integrate data sources across multiple types and locations into one logical data view.
Analytics Engine Powered by Apache Spark	Automatically spin up lightweight, dedicated Apache Spark clusters to run analytical and machine learning jobs.
Watson Knowledge Catalog	Create catalogs of curated assets with this secure enterprise catalog management platform that is supported by a data governance framework.

2.4.2. Describe the process and options for **deploying** models with Cloud Pak for data.

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=deploying-managing-models-functions

2.4.3. Explain the options for integrating externally developed models in Watson Studio. (Models created outside of Watson studio, some can be integrated into it – python and R Code etc, other cannot. However other models created outside of Watson Studio can be monitored in Watson Openscale.)

REFERENCE:
https://developer.ibm.com/technologies/artificial-intelligence/patterns/monitor-custom-machine-learning-engine-with-ai-openscale/

2.4.4. Consider the role of decision optimization.

REFERENCES:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=services-decision-optimization
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=models-decision-optimization

2.5. Map business opportunities into a Data Science use case.
SUBTASKS:

2.5.1. Explain how the popular use cases translate to Cloud Pak for data services and implementation.
2.5.1.1. ModelOps

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=cases-modelops-use-case

2.5.2. Selected Industry accelerator business use cases

REFERENCE:
https://community.ibm.com/accelerators/?context=analytics&type=Cloud%20Pak%20for%20Data%20industry

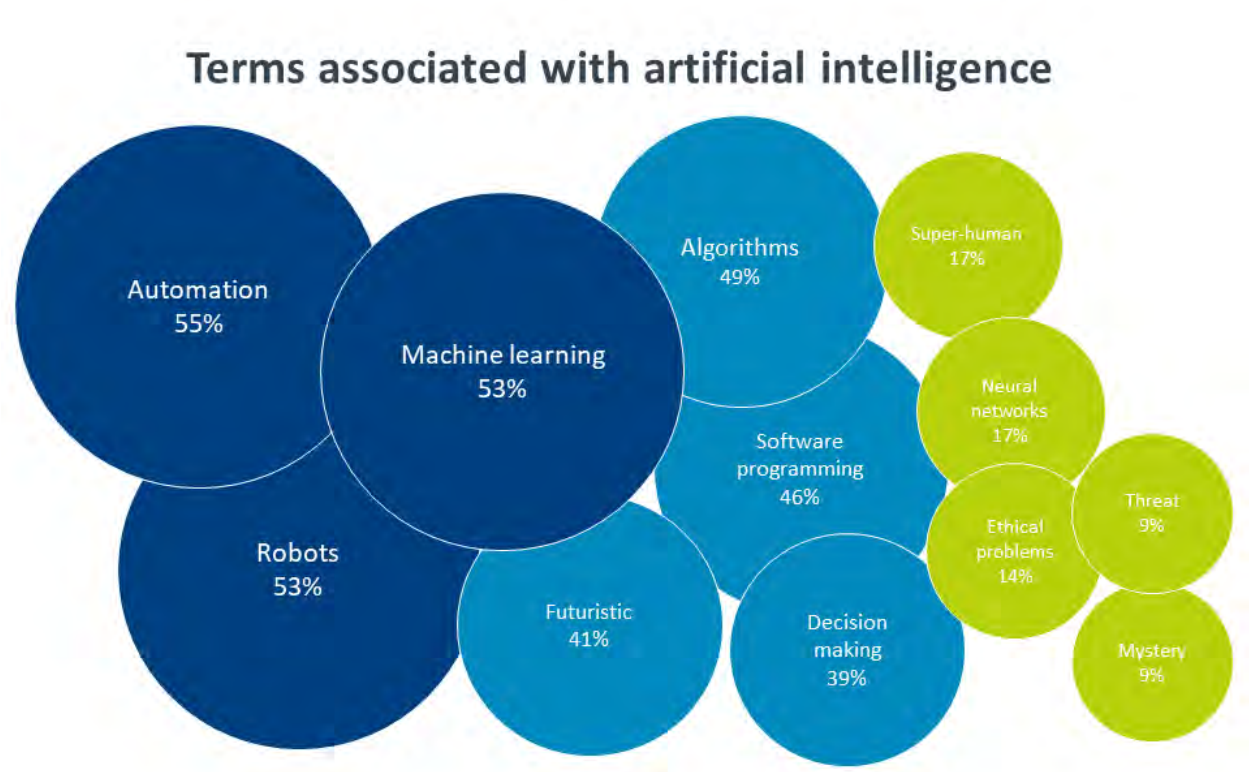
1. Insurance Loss Estimation using remote sensing data
https://community.ibm.com/accelerators/catalog/content/Insurance-Loss-Estimation-using-Remote-Sensing-Data

2. **Intelligent maintenance** - Intelligent asset management and predictive maintenance to streamline your operations
https://community.ibm.com/accelerators/catalog/content/Intelligent-Maintenance

3. **Utilities Customer Attrition prediction**
https://community.ibm.com/accelerators/catalog/content/Utilities-Customer-Attrition-Prediction

2.5.3. Identify the Emerging Business opportunities areas in AI

REFERENCES:
https://www.ibm.com/blogs/digital-transformation/in-en/blog/opportunities-and-challenges-in-ai-adoption
https://www.comptia.org/content/research/emerging-business-opportunities-in-ai



REFERENCES:

https://www.comptia.org/content/research/emerging-business-opportunities-in-ai
https://www.ibm.com/cloud/learn/devops-a-complete-guide

2.5.4. Chatbots, Virtual Assistants in service domain

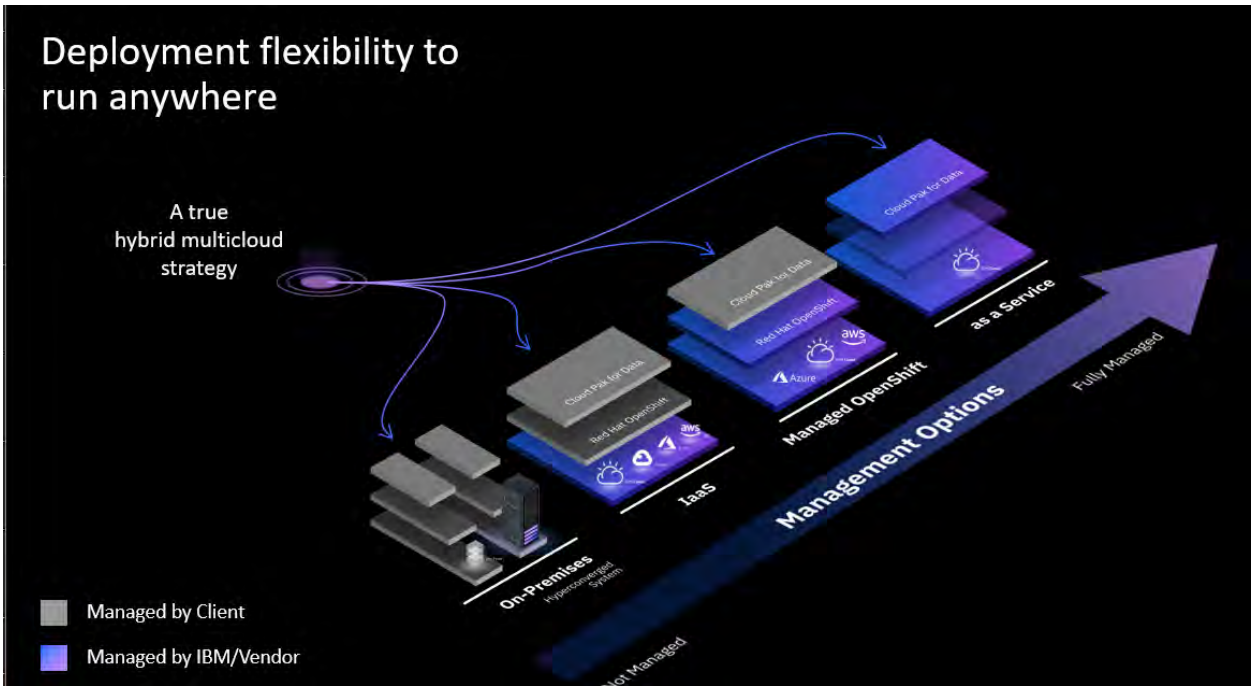
https://www.ibm.com/cloud/learn/devops-a-complete-guide
https://www.ibm.com/cloud/devops

2.5.5. Modernizing customer IBM Expert lab and Managed Services

https://www.ibm.com/products/expertlabs/managed-services

2.5.6. IBM Software-as-a-service solutions

https://www.ibm.com/products/cloud-pak-for-data/deployment-model-options
https://www.ibm.com/products/cloud-pak-for-data?p1=Search&p4=43700064659101157&p5=p&gclid=Cj0KCQjw7MGJBhD-ARIsAMZ0eeuJ8BD1a8iLuciD_vYwyBjReGpAKG974jwoun-J1iuA3jrTVoXC7rgaAjeAEALw_wcB&gclsrc=aw.ds#2957466



REFERENCE:
https://www.ibm.com/cloud/garage/toolchains/

2.5.7. How IBM Business partners can build high values

REFERENCE:
https://www.ibm.com/partners/
2.5.8. Palantir – AI Democratization

REFERENCE:
https://www.ibm.com/products/palantir-cloud-pak-for-data

Section 3 – Machine Learning Operations

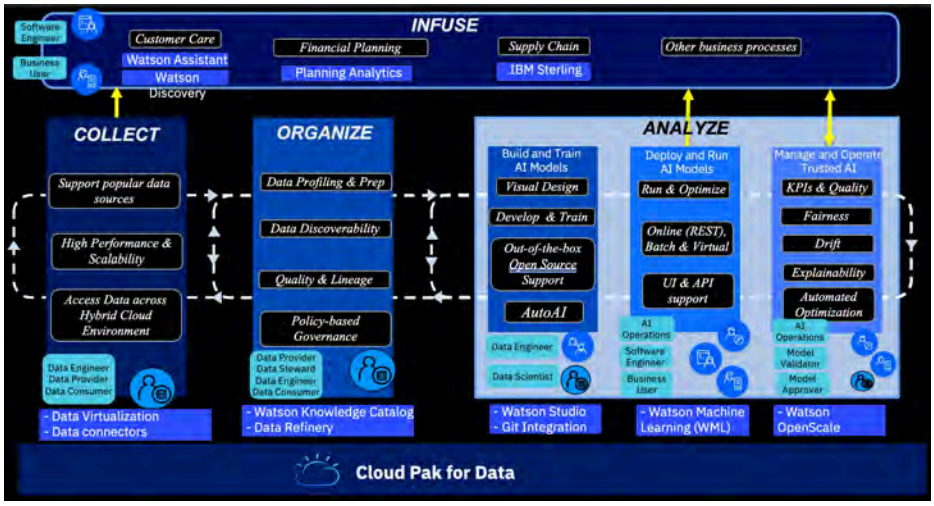
3.1. Describe the key considerations when selecting a platform for model deployment.

SUBTASK(S):

3.1.1. Model Operationalizing requirements

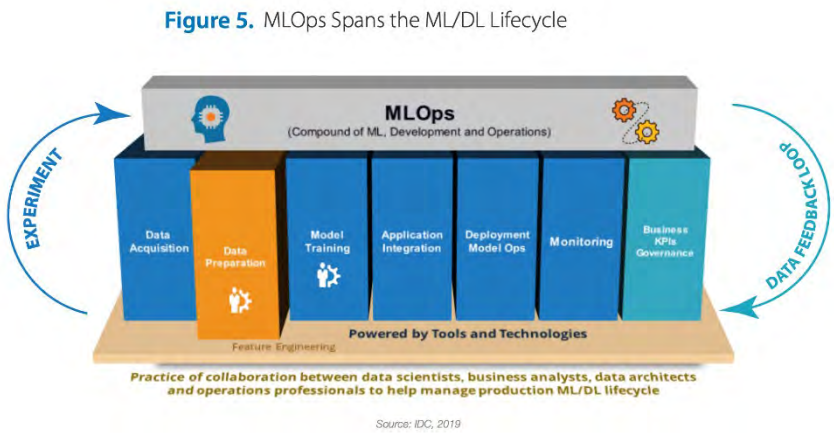
REFERENCES:
https://medium.com/@ODSC/ml-operationalization-from-what-and-why-to-how-and-who-2299dd048f36
https://www.ibm.com/cloud/architecture/architectures/machine-learning-ops-solution/
https://www.ibm.com/downloads/cas/ENDG17K3
https://www.datarobot.com/mlops-101/

https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning



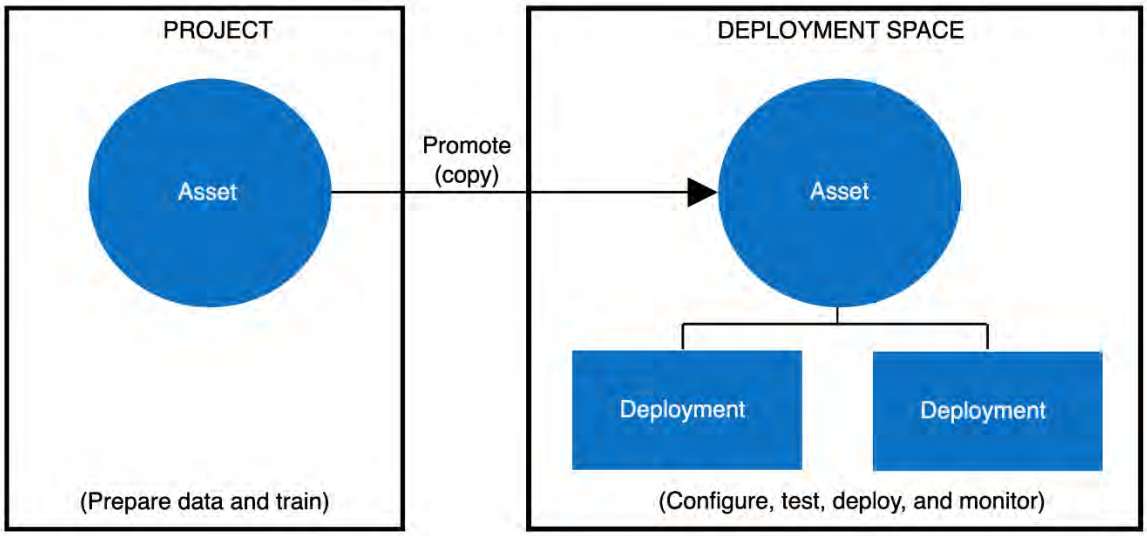
3.1.2. Describe the benefits of a machine learning pipeline

REFERENCE:
https://developer.ibm.com/blogs/kubeflow-pipelines-with-tekton-and-watson/_&mhsrc=ibmsearch_a&mhq=modelops



3.1.3. Operationalize Machine Learning Models build by data scientists

REFERENCES:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=deploying-managing-models-functions_
https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/ml-spaces_local.html#promote



3.1.4. Evaluate deployed models for bias using Watson Studio (Watson OpenScale)

REFERENCES:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=functions-validating-monitoring-ai-models-watson-openscale
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=openscale-faqs#wos-000-fairness
https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/ml-deploy-manage.html?audience=wdp
https://www.ibm.com/support/producthub/icpdata/docs/content/SSQNUZ_latest/wsj/model/wos-monitor-fairness.html
https://www.ibm.com/docs/api/v1/content/SSQNUZ_4.0/wsj/model/images/wos-explainability-controllable-features.png

3.1.5. Model Risk management

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=governance-manage-model-risk

3.1.6. Model Governance

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=governance-model

3.2. Explain the workflow of deploying and monitoring models.
SUBTASK(S):

3.2.1. Define the importance of ML lifecycle

REFERENCES
https://www.ibm.com/cloud/blog/ai-model-lifecycle-management-build-phase
https://www.ibm.com/docs/en/spss-modeler/SaaS?topic=dm-crisp-help-overview

https://towardsdatascience.com/rethinking-ai-machine-learning-model-management-8afeaa31d8f8
https://www.ibm.com/blogs/academy-of-technology/ai-model-lifecycle-management-white-paper/

3.2.2. Collaborative Platform - Watson Studio

REFERENCES:
https://www.ibm.com/cloud/blog/ai-model-lifecycle-management-build-phase
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=projects
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=integration-syncing-projects

3.2.3. Describe steps involved in preparing the models for monitoring

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=openscale-prepare-models-monitoring

3.2.4. Describe the business scenarios for using the Monitoring options
3.2.4.1. Quality

https://www.ibm.com/support/producthub/icpdata/docs/content/SSQNUZ_latest/wsj/model/wos-quality-overview.html

3.2.4.2. Fairness

https://www.ibm.com/support/producthub/icpdata/docs/content/SSQNUZ_latest/wsj/model/wos-fairness-ovr.html

3.2.4.3. Drift detection

https://www.ibm.com/support/producthub/icpdata/docs/content/SSQNUZ_latest/wsj/model/wos-behavior-overview.html

3.2.4.4. Explainability Monitoring

https://www.ibm.com/support/producthub/icpdata/docs/content/SSQNUZ_latest/wsj/model/wos-analytics-ovr.html

3.3. Monitor deployed models inside Cloud Pak for Data.
SUBTASK(S):

3.3.1. Model Feedback Analysis

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=functions-validating-monitoring-ai-models-watson-openscale

3.3.2. Drift Detection

REFERENCES:
https://www.ibm.com/cloud/watson-studio/drift
https://dataplatform.cloud.ibm.com/docs/content/wsj/model/wos-behavior-overview.html
https://dataplatform.cloud.ibm.com/docs/content/wsj/model/wos-behavior-drift.html

3.3.3. Bias Monitoring & Mitigating bias throughout the AI lifecycle

https://dataplatform.cloud.ibm.com/docs/content/wsj/model/wos-fairness-ovr.html?audience=wdp

3.3.4. Analyzing the scoring payloads in Watson Studio

REFERENCES:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=openscale-analyzing-scoring-payload
https://dataplatform.cloud.ibm.com/docs/content/wsj/model/wos-analytics-ovr.html?audience=wdp

3.3.5. Getting Model Insights in Watson Studio

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=insights-explaining-transactions#ie-contrastive.v.lime

3.3.6. AI Fairness and Bias Mitigation

REFERENCE:
https://aif360.mybluemix.net/resources

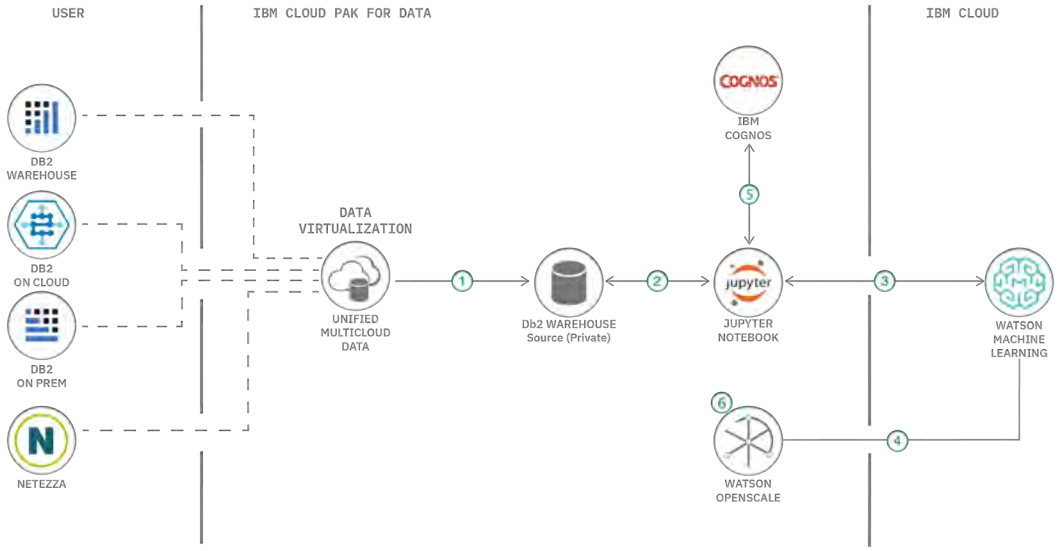
3.3.7. Responsible AI Practices

REFERENCE:
https://www.ibm.com/artificial-intelligence/ethics

3.4. Monitor machine learning models running on an external platform.
SUBTASK(S):

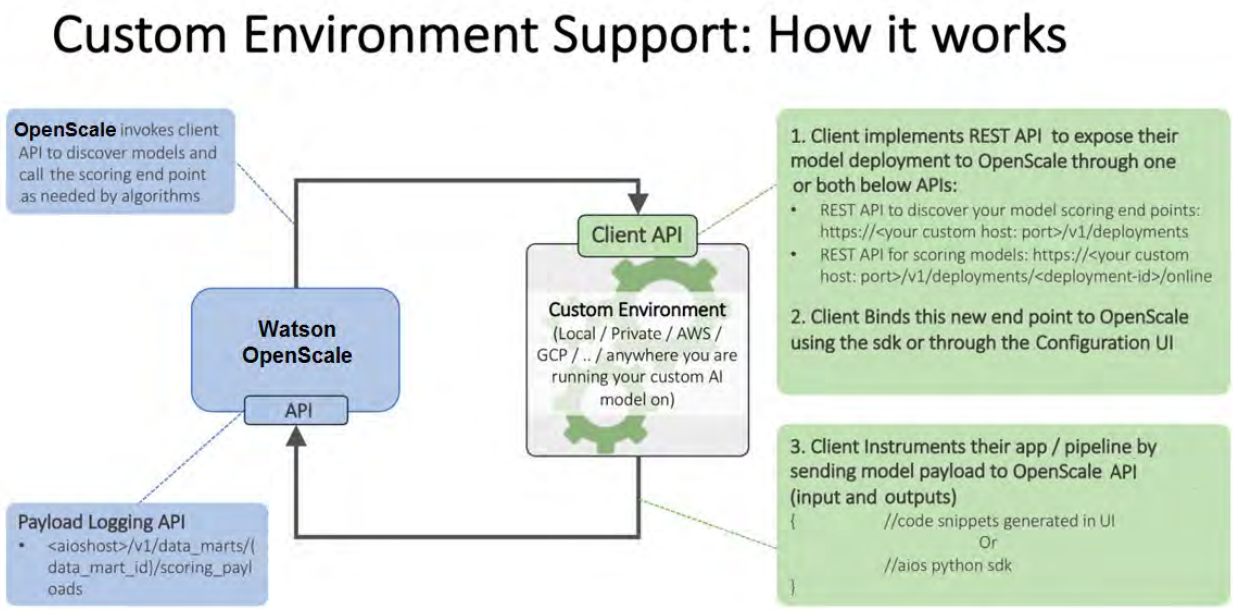
3.4.1. Identify the reasons customer would deploy ML model outside of Cloud Pak for Data.
3.4.2. Describe the AI Open scale to work with models deployed outside the Cloud Pak for Data.

a. Models deployed in WML on IBM Cloud



3.4.3. Integrating with 3rd Party ML engines

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/1.0?topic=models-integrating-3rd-party-ml-engines-watson-openscale



3.4.4. How to Integrate IBM OpenScale with Amazon SageMaker

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/1.0?topic=models-amazon-sagemaker-frameworks

3.4.5. Monitoring models Google Cloud AI platform
3.4.6. Models deployed in Azure ML Studio

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=models-microsoft-azure-ml-service-frameworks

3.4.7. Identify the opensource ML platforms
    3.4.7.1. MLflow
    3.4.7.2. Seldon

3.5. Manage risk and regulatory compliance using Open Pages.
SUBTASK(S):

https://www.ibm.com/products/openpages-with-watson
https://www.ibm.com/blogs/journey-to-ai/2021/07/citi-transforms-critical-internal-audit-with-machine-learning-nlp-and-ai
3.5.1. Build and Scale trusted AI

REFERENCE:
https://www.ibm.com/cloud/watson-studio

3.5.2. Define Governance and Regulatory Compliance (GRC)

REFERENCE:
https://www.ibm.com/blogs/journey-to-ai/2021/07/simplify-grc-management-with-30-day-trial-edition-of-ibm-openpages-regulatory-compliance-management/https://www.fdic.gov/news/financial-institution-letters/2017/fil17022a.pdf

3.5.3. Explainability and Trust in the Models

REFERENCES:
https://www.ibm.com/blogs/watson/2021/04/ibm-watson-trust-data-models-processes/
https://towardsdatascience.com/an-overview-of-model-explainability-in-modern-machine-learning-fc0f22c8c29a
https://docs.aws.amazon.com/sagemaker/latest/dg/clarify-model-explainability.html

3.5.4. Model Operational Efficiency

REFERENCE:
https://www.ibm.com/downloads/cas/5N39ZO2X

3.5.5. Automate and Manage Internal Audits

REFERENCE:
https://www.ibm.com/downloads/cas/NVYXVJPA

Section 4 - Analytics

4.1. Explain the difference between Descriptive, Prescriptive, Predictive, Diagnostic, and Cognitive Analytics.

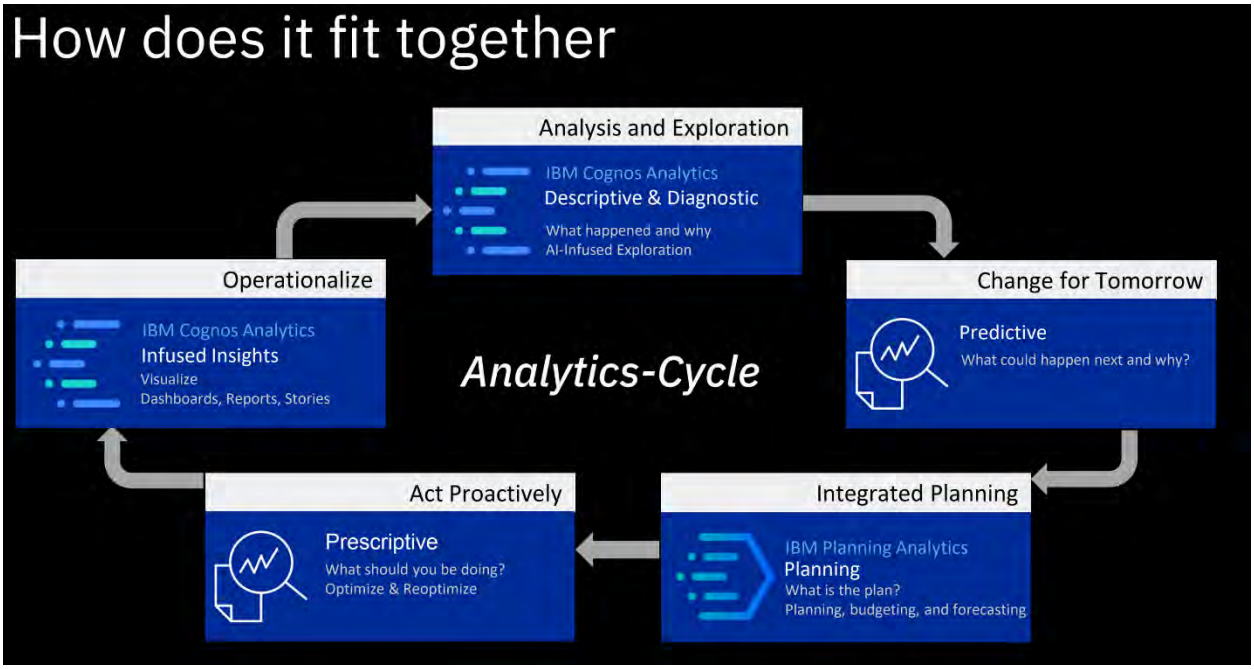
SUBTASK(S):
Predictive analytics let you forecast results and optimize outcomes. Test scenarios before implementing plans to see the potential impact of your decisions

4.1.1. Describe the services used for Descriptive Analytics
    4.1.1.1. Planning Analytics
    4.1.1.2. Cognos Analytics
4.1.2. Explain use cases for Descriptive Analytics
4.1.3. Define Descriptive Analytics
4.1.4. Describe the services used for Prescriptive Analytics
    4.1.4.1. Planning Analytics
    4.1.4.2. Decision Optimization
    4.1.4.3. Cognos Analytics
    4.1.4.4. SPSS
    4.1.4.5. Decision Optimization
4.1.5. Explain use cases for Prescriptive Analytics
4.1.6. Define Prescriptive Analytics
4.1.7. Describe the services used for Diagnostic Analytics

4.1.8. Explain use cases for Diagnostic Analytics
4.1.9. Define Diagnostic Analytics
4.1.10. Describe the services used for Cognitive Analytics
4.1.11. Explain use cases for Cognitive Analytics
4.1.12. Define Cognitive Analytics

REFERENCES:

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=business-analytics
https://www.ibm.com/analytics/data-science
https://www.ibm.com/downloads/cas/QEONJOR4



4.2. Describe the capabilities of AI for financial operations in Cloud Pak for Data.

SUBTASK(S):
4.2.1. List the capabilities of Planning Analytics (rpts, models, cubes)
4.2.2. List the benefits of running Planning Analytics in Cloud Pak for Data (single UI)
4.2.3. List the components to access data on the TM1 database
4.2.4. Identify considerations when managing users in Planning Analytics
4.2.5. Describe the use cases for using Planning Analytics

REFERENCES:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=analytics-modeling-reporting-patterns-in-data-planning
https://www.ibm.com/docs/en/cloud-paks/cp-data
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=business-analytics
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=use-cases
https://www.ibm.com/docs/en/planning-analytics/2.0.0

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=analytics-managing-users-in-planning
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=analytics-managing-users-in-planning
https://community.ibm.com/community/user/businessanalytics/home

4.3. Describe the capabilities of business intelligence in Cloud Pak for Data.
SUBTASK(S):

4.3.1. Identify how Natural language processing can be used in business reporting.
4.3.2. Describe the benefits of using Cognos Analytics for Cloud Pak for Data (access Data sources with DV/WKC that Cognos doesn't support natively, scalability, containerized)
4.3.3. Describe self-service analytics (using Cognos)
4.3.4. List options for report distribution (using Cognos)
4.3.5. Explain Data Explorer capabilities (using Cognos)
4.3.6. Contrast the Business Intelligence options in Cloud Pak for Data (dashboard / full Cognos)
4.3.7. Describe the security feature in Cognos Analytics for Cloud Pak for Data (role-based)

REFERENCES:
https://www.ibm.com/docs/en/cloud-paks/cp-data
https://www.ibm.com/products/cognos-analytics/features
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=business-analytics

https://community.ibm.com/community/user/businessanalytics/communities/community-home/all-news?communitykey=6b10df83-0b3c-4f92-8b1f-1fd80d0e7e58&tab=librarydocuments&LibraryFolderKey=&DefaultView=
https://ibm.webcasts.com/viewer/event.jsp?ei=1410095&tp_key=6ef71723aa (Cognos Modernization Deep Dive 2020-12)
https://www.ibm.com/products/cognos-analytics
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=services-cognos-analytics
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=services-cognos-dashboards

4.4. Map business opportunities into an Analytics use case.
SUBTASK(S):

4.4.1. Identify a use case for Planning Analytics (Examples: AI for Financial Operations)
4.4.2. Identify a use case for Cognos Analytics (Examples:)
4.4.3. Describe a use case for Decision Optimization (Example: Intelligent maintenance Industry accelerators)

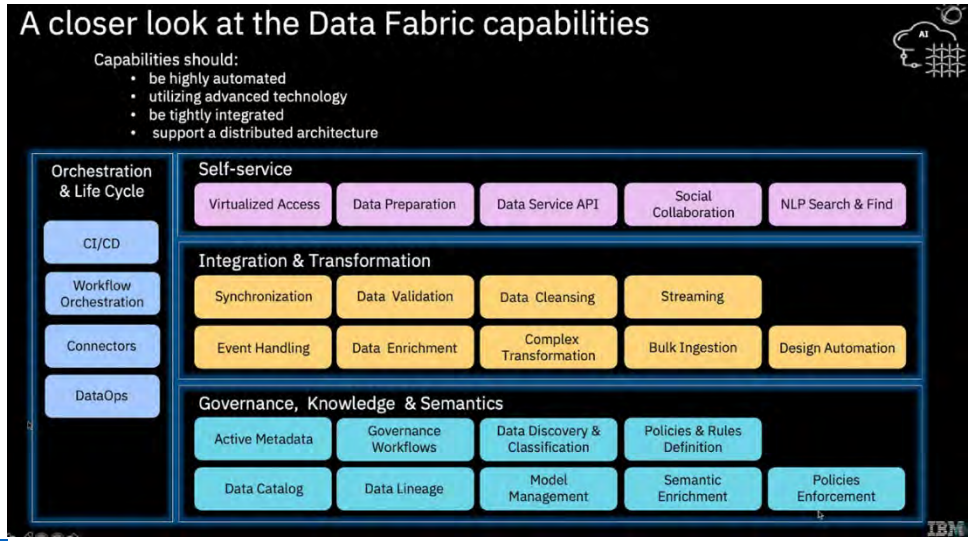
REFERENCES:
https://www.ibm.com/blogs/journey-to-ai/
https://dataplatform.cloud.ibm.com/docs/content/wsj/getting-started/faq.html
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=use-cases
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=analytics-analyzing-visualizing-data-cognos
https://www.ibm.com/products/cognos-analytics
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=overview-whats-new
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=business-analytics
https://www.ibm.com/support/producthub/icpdata/docs/content/SSQNUZ_latest/svc
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=services-decision-optimization
https://www.ibm.com/common/ssi/ShowDoc.wss?docURL=/common/ssi/rep_sm/6/649/ENUS5737-H76/index.html&lang=en&request_locale=en
https://www.ibm.com/analytics/business-analytics

Section 5 - Data Governance

5.1. Describe the capabilities of a data fabric topology.
SUBTASK(S):

(Data Connections)
5.1.1. Describe a Data Fabric

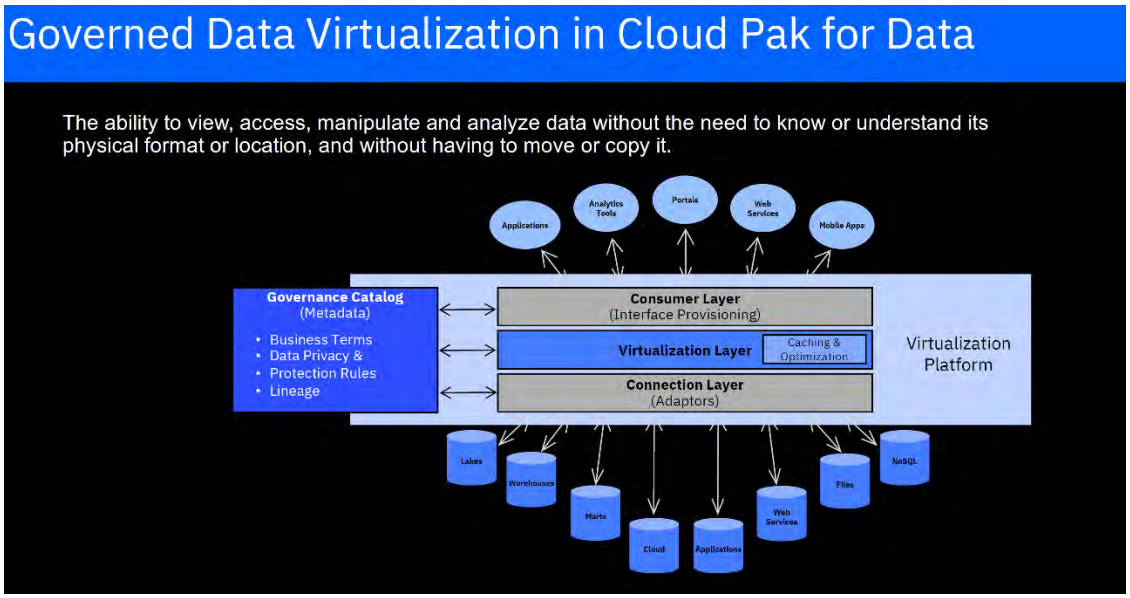
https://www.ibm.com/cloud/blog/announcements/automate-data-discovery-governance-and-consumption-with-ibms-data-fabric
https://www.ibm.com/analytics/data-fabric



5.1.2. Identify the primary considerations for created a platform connection

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=data-supported-sources

5.1.3. Explain the benefits of data virtualization



Installing connectors on remote data sources (Data Virtualization) - IBM Documentation

5.1.4. Identify the types of data sources that can be virtualized
- 5.1.4.1. Single Source
- 5.1.4.2. Multiple Sources
- 5.1.4.3. CSV Files

Creating a virtualized table in Data Virtualization | IBM Cloud Pak for Data
Creating a virtualized grouped table from multiple data sources in Data Virtualization | IBM Cloud Pak for Data
Creating virtualized tables from files in Data Virtualization | IBM Cloud Pak for Data
Installing connectors on remote data sources (Data Virtualization for IBM Cloud Pak for Data) | IBM Cloud Pak for Data

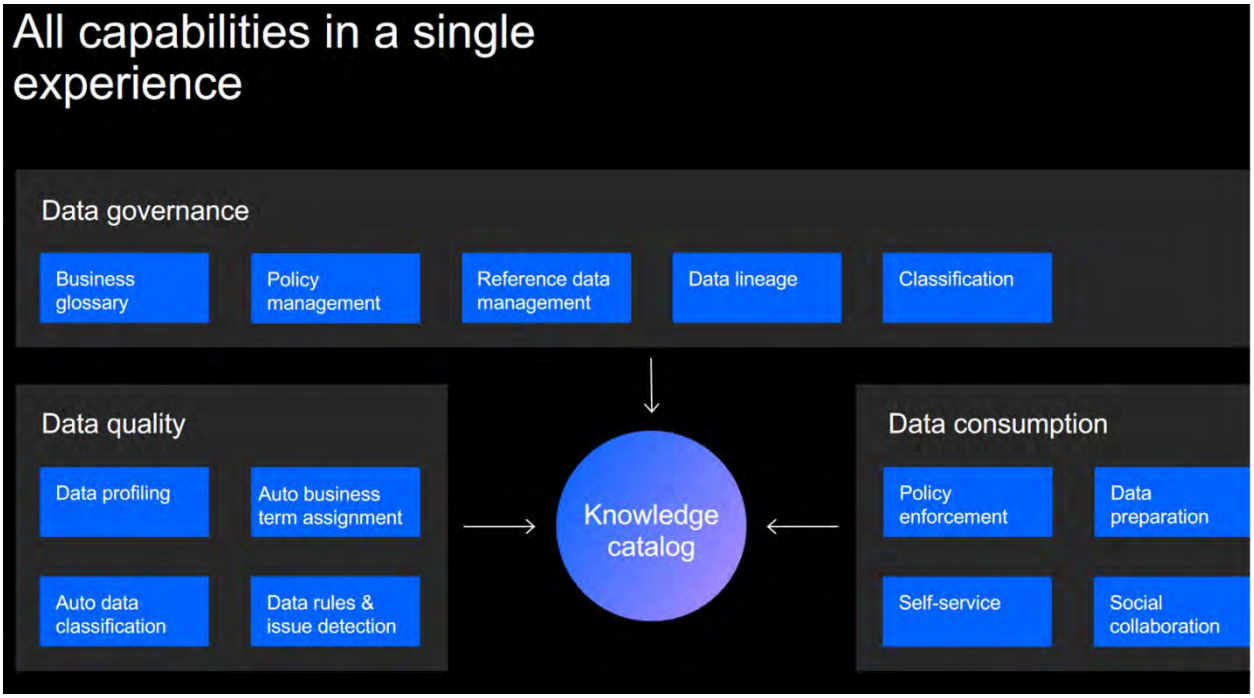
5.1.5. Explain how AutoSQL enables access across both structured and unstructured data

Manage all of your data – any type, on any cloud, from any vendor – with a high-performance universal query engine and data fabric - Journey to AI Blog (ibm.com)

5.1.6. Describe capabilities of joining virtual objects
Joining multiple tables in Data Virtualization | IBM Cloud Pak for Data
(Data Ingestion- profiling, auto catalog, data discovery)

5.1.7. Identify the data source types available within the Cloud Pak for Data
    5.1.7.1. IBM Services
    5.1.7.2. Third Party Services
    5.1.7.3. User-defined

5.2. Define the Governance structure.
SUBTASK(S):



5.2.1. Describe how data catalogs are created, maintained, and synchronized with information assets.
    5.2.1.1. Default Catalog
    5.2.1.2. Multiple Catalogs
    5.2.1.3. Catalog Permissions
    5.2.1.4. Asset Properties and Relationships
    5.2.1.5. Assets Types
Administering a catalog (Watson Knowledge Catalog) - IBM Documentation
Catalog assets (Watson Knowledge Catalog) - IBM Documentation

5.2.2. Distinguish between data policy vs data rules
Governance artifacts (Watson Knowledge Catalog) - IBM Documentation
Policies (Watson Knowledge Catalog) - IBM Documentation
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=artifacts-governance-rul

5.2.3. Summarize how data protection rules and masking are implemented to address industry privacy standards:
5.2.3.1. PHI
5.2.3.2. PII
5.2.3.3. GDPR
5.2.3.4. CCPA

Data protection rules (Watson Knowledge Catalog) - IBM Documentation
Masking data with data protection rules (Watson Knowledge Catalog) - IBM Documentation

5.2.4. Summarize the capabilities of OpenPages to assist with monitoring compliance initiatives.

Managing risk and regulatory challenges with OpenPages - IBM Documentation

5.2.5. Identify/List commonly used predefined data classes. Explain the different between predefined data classes and custom data-class artifacts

Data classes (Watson Knowledge Catalog) - IBM Documentation
Predefined data classes (Watson Knowledge Catalog) - IBM Documentation
Governing and curating data (Watson Knowledge Catalog) - IBM Documentation
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=classes-predefined-data-details
https://www.ibm.com/analytics/info/cloud-pak-for-data-modernize-upgrade

5.3. Use smart ingestion for auto cataloging.
SUBTASK(S):

5.3.1. Demonstrate how data virtualized tables are associated to business terms and published to the catalog
Governing virtual data in Data Virtualization | IBM Cloud Pak for Data

5.3.2. Explain the methods to discover assets into WKC

Discovering assets (Watson Knowledge Catalog) - IBM Documentation

5.3.3. Demonstrate the understanding of Data Quality capabilities within Cloud Pak
5.3.3.1. Column Analysis
5.3.3.2. Auto Discovery results
5.3.3.3. Project Integration

Running a column analysis (Watson Knowledge Catalog) - IBM Documentation
Data quality project settings (Watson Knowledge Catalog) - IBM Documentation

5.4. Understand how workflow is used in Cloud Pak for Data.

SUBTASK(S):

5.4.1. Explain how workflows are implemented within Watson Knowledge Catalog
- 5.4.1.1. Workflow Tasks
- 5.4.1.2. Custom workflows
- 5.4.1.3. Workflow Notification

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=started-identifying-tasks-that-need-be-completed
Monitoring workflow tasks - IBM Documentation
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=governance-workflows

5.5. Explain the use of Guardium in auditing and monitoring data
SUBTASK(S):

5.5.1. Describe the purpose of IBM Guardium
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=integrations-auditing-your-sensitive-data-guardium

5.5.2. Describe the type of data that can be protected using IBM Guardium
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=integrations-auditing-your-sensitive-data-guardium

5.5.3. Explain the high-level process that you must follow to integrate IBM Cloud Pak for Data with IBM Guardium
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=integrations-auditing-your-sensitive-data-guardium

5.5.4. List and describe the steps needed to prepare for the integration of IBM Guardium
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=guardium-preparing-your-system

5.5.5. Connect to IBM Guardium appliances
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=guardium-connecting-appliances

5.5.6. Describe how you can audit and monitor your sensitive data with IBM Guardium
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=guardium-connecting-appliances
https://www.ibm.com/docs/en/guardium/10.1?topic=monitor-audit

5.6. Leverage the platform to understand data flow and usage
SUBTASK(S):

5.6.1. Describe the synchronization process between information assets and default catalog assets
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=catalogs-information-assets-view

5.6.2. Identify the assets that are synchronized
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=catalogs-information-assets-view

5.6.3. List the information asset linage report types
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=view-information-asset-lineage-reports

5.6.4. Compare the data linage report with the business lineage report
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=view-information-asset-lineage-reports#report-types

5.7. Explain the concepts of Knowledge Accelerators
SUBTASK(S):

5.7.1. Explain the purpose of the knowledge accelerators
https://www.ibm.com/cloud/knowledge-accelerators
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=artifacts-business-terms

5.7.2. Describe the assets that are included in the knowledge accelerators
https://www.ibm.com/docs/en/knowledge-accelerators/1.0.0?topic=product-overview
https://www.ibm.com/docs/en/knowledge-accelerators/1.0.0?topic=SSWNJS_1.0.0/com.ibm.ima.product_overview/pro_ove/pro_ove.html

5.7.3. Explain the process for using a knowledge accelerator
https://www.ibm.com/docs/en/knowledge-accelerators/1.0.0?topic=started-importing-into-watson-knowledge-catalog

5.8. Map business opportunities into a data governance use case
SUBTASK(S):

5.8.1. Explain how enterprise governance can ensure that your company's resources are used responsibly
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=enterprise-governance

5.8.2. Describe the data governance use cases and how the various roles and personas would manage them

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=governance-watson-knowledge-catalog

Section 6 - Integration, Implementation, Deployment, and Scaling

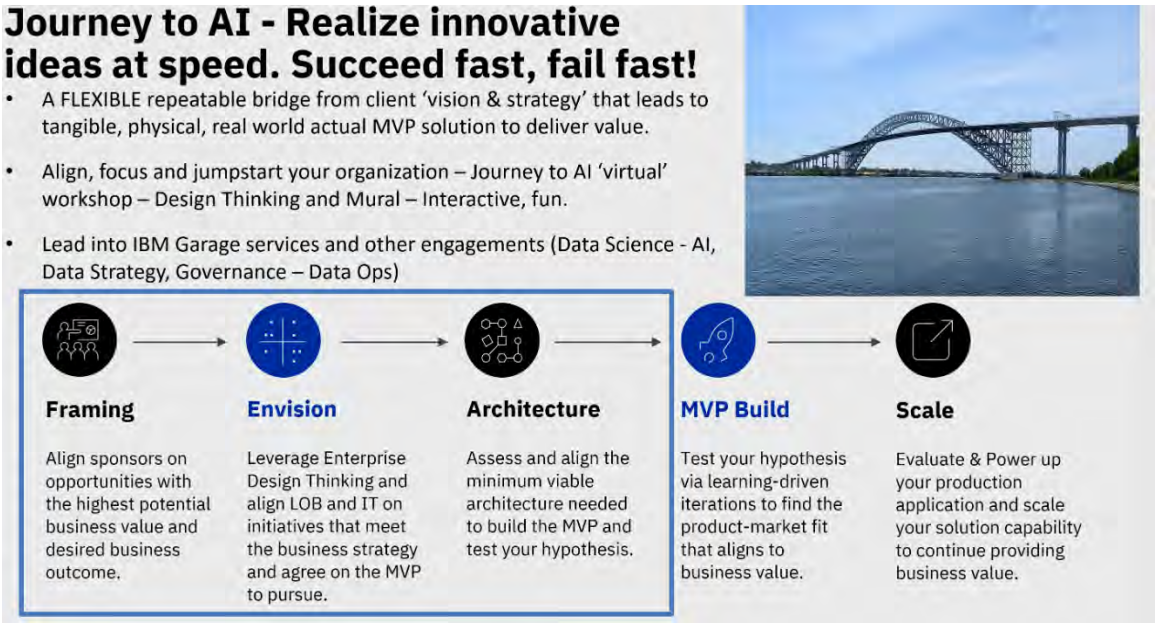
6.1. Develop an appropriate process to take a Data and AI solution from inception to production.

SUBTASK(S):

6.1.1. Describe the characteristics of a viable pilot project
  6.1.1.1. Identify key goals to reach
  6.1.1.2. Identify the right opportunity and achievable success criteria

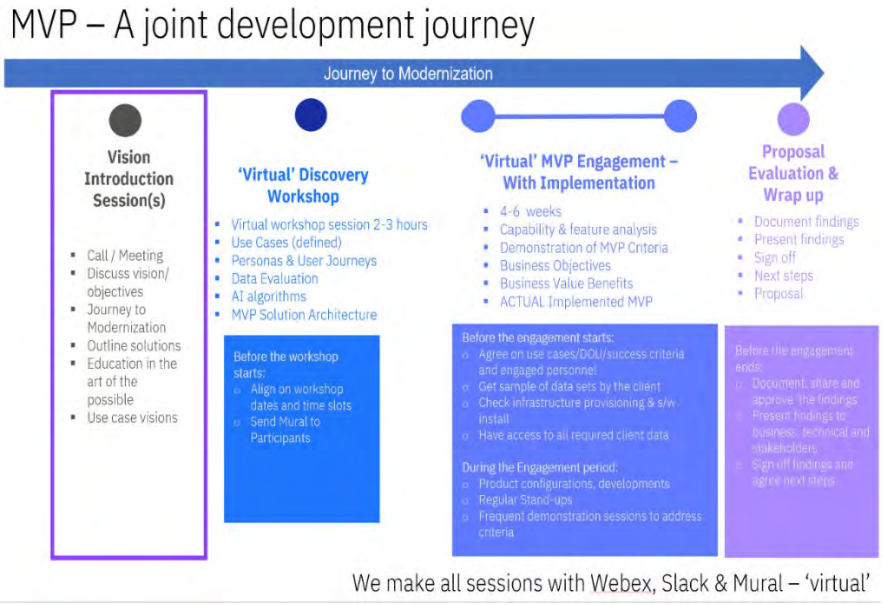
IBM Garage and Business Framing sessions

REFERENCES:
https://www.ibm.com/garage/method/
https://www.ibm.com/cloud/architecture/content/course/garage-method-for-cloud-advocate



The MVP Lifecycle

REFERENCE:
https://www.ibm.com/garage/method/practices/think/practice_minimum_viable_product/



6.1.1.3. Identify some accelerators which can get a pilot to value quicker

REFERENCES:
http://www.eiminstitute.org/library/eimi-archives/volumn-4-issue-4-september-2010/building-valuable-pilot-projects
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=integrations-external-data-sets
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=integrations-industry-accelerators

6.1.2. Identify ways to isolate development/test and production service instances

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=planning-multitenancy-support

6.1.3. Identify situations where you will desire fewer compute dense nodes vs less dense nodes
6.1.3.1. Example Jupyter notebook usage, is it more than compute on each node.
6.1.3.2. Dedicated database nodes may need special RAM configurations

6.1.4. Identify considerations needed when scaling a solution
6.1.4.1. Describe entitlement capacity
6.1.4.2. Describe when to add hardware

REFERENCES:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=cluster-scaling-services
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=assets-scaling-deployment
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=requirements-hardware

6.2. Develop a strategy to monitor the Data and AI platform.

SUBTASK(S):

6.2.1. Describe the ability to provide metrics regarding compute and memory usage and trends.
6.2.2. Describe the ability to alert at usage thresholds.
6.2.3. Identify the types of alert notification protocols supported
6.2.4. Identify how you can extend existing dashboard monitoring

REFERENCES:

https://www.ibm.com/support/producthub/icpdata/docs/content/SSQNUZ_latest/cpd/admin/alerting_and_monitoring.html
https://www.ibm.com/support/producthub/icpdata/docs/content/SSQNUZ_latest/cpd/admin/platform-management.html
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=resources-home-page-custom-cards-apis

6.3. Accelerate the solution using Industry Accelerators and External Data sets.

SUBTASK(S):

6.3.1. Describe how Industry Accelerators can be used to jump start development
6.3.2. Understand how to find and install accelerators
    6.3.2.1. Explain where to download accelerators
    6.3.2.2. Understand how to import and deploy accelerator projects and where to find instructions
6.3.3. Summarize the use cases of the Industry Accelerators
6.3.4. Identify the external data sets available with Cloud Pak for Data

REFERENCES:

https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=integrations-industry-accelerators
https://community.ibm.com/accelerators/?context=analytics&type=Cloud%20Pak%20for%20Data%20industry
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=integrations-external-data-sets

6.4. Integrating Business Applications using Cloud Pak for Data.
SUBTASK(S):

6.4.1. Identify options of exposing a Machine learning model API
6.4.2. Describe best practices when exposing APIs for external applications usage
6.4.3. Describe use cases when a business application would integrate with Cloud Pak for Data

REFERENCE:
https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=assets-getting-deployment-endpoint-url

Next Steps

1. Take the https://www.ibm.com/certify/exam?id=C1000-136 Assessment test.

2. If you pass the assessment exam, visit pearsonvue.com/ibm to schedule your testing sessions.

3. If you failed the assessment exam, review how you did by section. Focus attention on the sections where you need improvement. Keep in mind that you can take the assessment exam as many times as you would like ($30 per exam), however, you will still receive the same questions only in a different order.