

Statement of Work

Deloitte.

&

**Client BPC
(RADLab)**

Table of Contents

- [1. About this proposal](#)
- [2. Executive Summary](#)
 - [Current architecture:](#)
 - [Datasets high level details - Chemical Fingerprinting](#)
 - [Ecosystem](#)
 - [Proposed architecture](#)
- [3. GCP Complete Architecture Specifications](#)
- [4. Scope of work & RACI](#)
 - [RACI definitions:](#)
- [5. Deliverables](#)
- [6. Success Criteria](#)
- [7. Out of Scope activities](#)
- [8. Assumptions](#)
- [9. Roles and Responsibilities](#)
- [10. Timelines](#)
- [11. Pricing](#)
 - [Resources Cost](#)
 - [One Time Implementation project cost](#)
- [12. Stakeholders](#)
- [13. Terms & Conditions](#)
- [14. Exit terms & Conditions](#)
- [15. Acceptance](#)
- [16. Other Alternative Cloud Architectures](#)

1.About this proposal

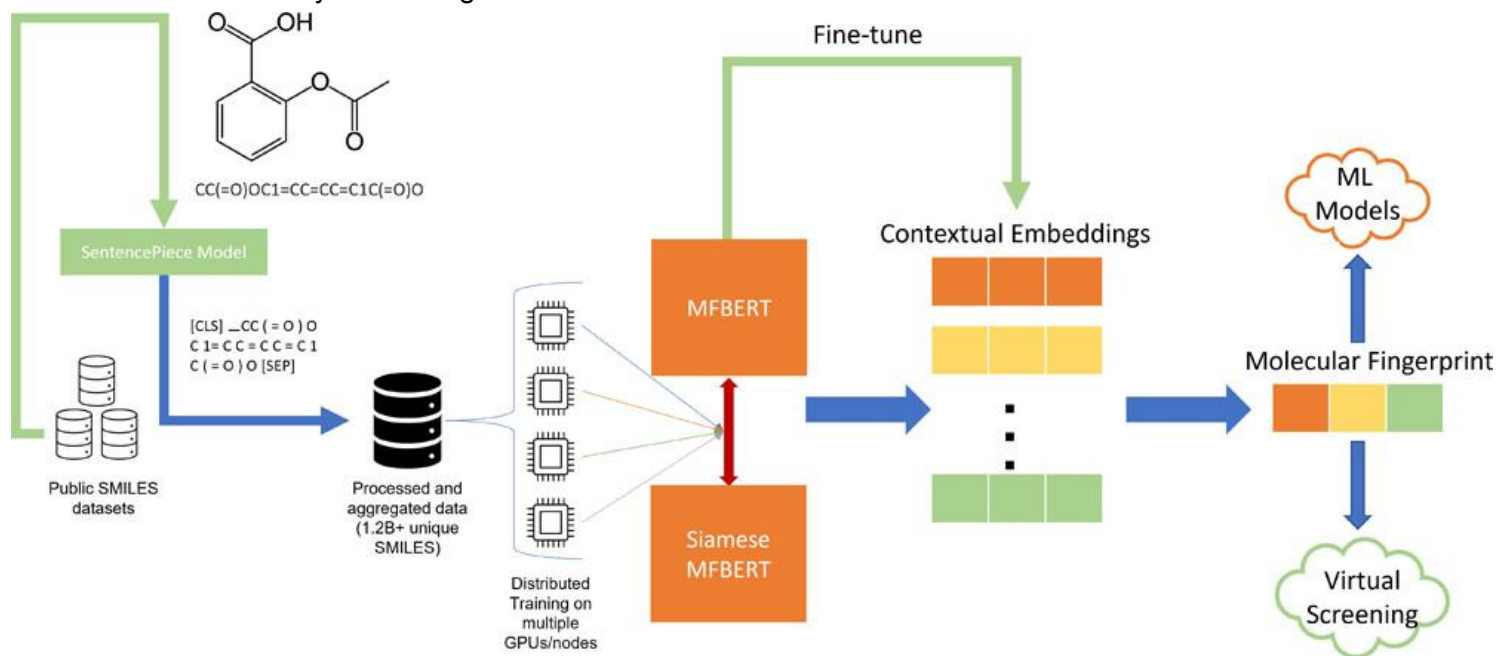
- 1) This document contains SOW between Deloitte Touche Tohmatsu Limited commonly referred to as **Deloitte** (an international professional services network headquartered in London, England.), having its registered office at ...address..... (include its parent, affiliates, sister concerns, subsidiaries and assigns) and client BPC, having it's registered office at ...address.. (include its parent, affiliates, sister concerns, subsidiaries and assigns).
- 2) This SoW is pursuant to the Master Services Agreement ("**MSA**").
- 3) This SoW applies to services and solutions described herein and is effective from (**Effective date**) and executed on("Execution date") unless terminated earlier in accordance with this SOW.
- 4) This SOW is written and negotiated in English. It is the complete and exclusive agreement between the Parties.
- 5) By Signing in the Acceptance section, the parties agree to be bound by the terms of the SOW.

2. Executive Summary

1. Client BPC (we are naming it as RADLab) is a Big Pharma Company that would like to implement the algorithm below in their workflow: [ChemBERTa: Large-Scale Self-Supervised Pretraining for Molecular Property Prediction](#).
2. Their overall project goal is to have a solution for chemical fingerprinting used to represent molecules for property prediction in downstream applications.
3. Deloitte will provide services to implement a rapid move and improvement of compute, data, machine learning and related workloads from Client's data center to Cloud (GCP or AWS or Azure).
4. In the existing set up, Client BPC uses an on-premise set up, and the current architecture diagram is below.

Current architecture:

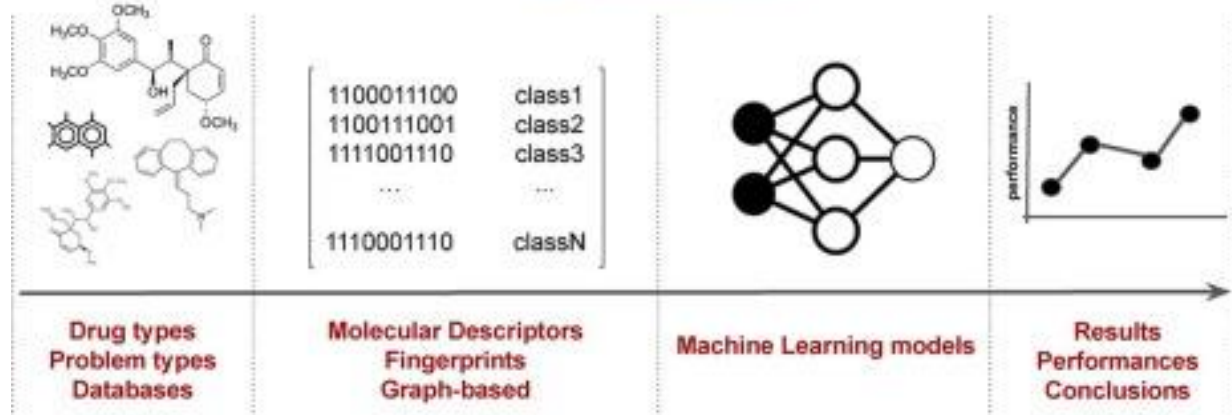
a. System Design



b.

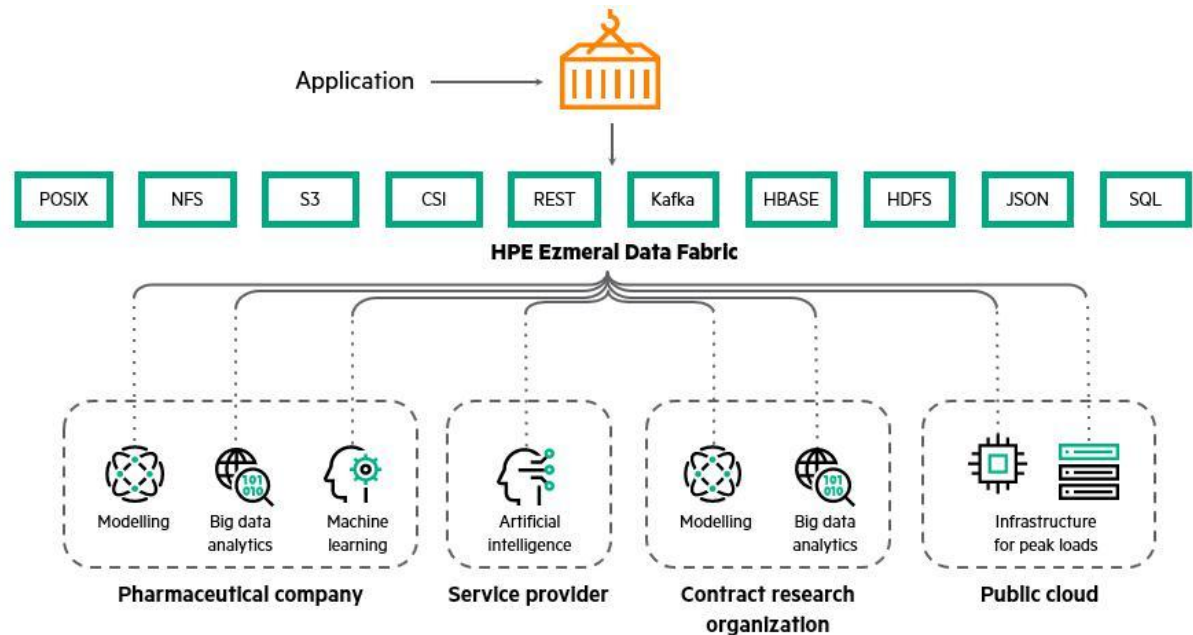
Datasets high level details - Chemical Fingerprinting

Machine Learning in Drug Discovery -- State-of-the-art Review -- (from 2016 to 2020)



c.

Ecosystem

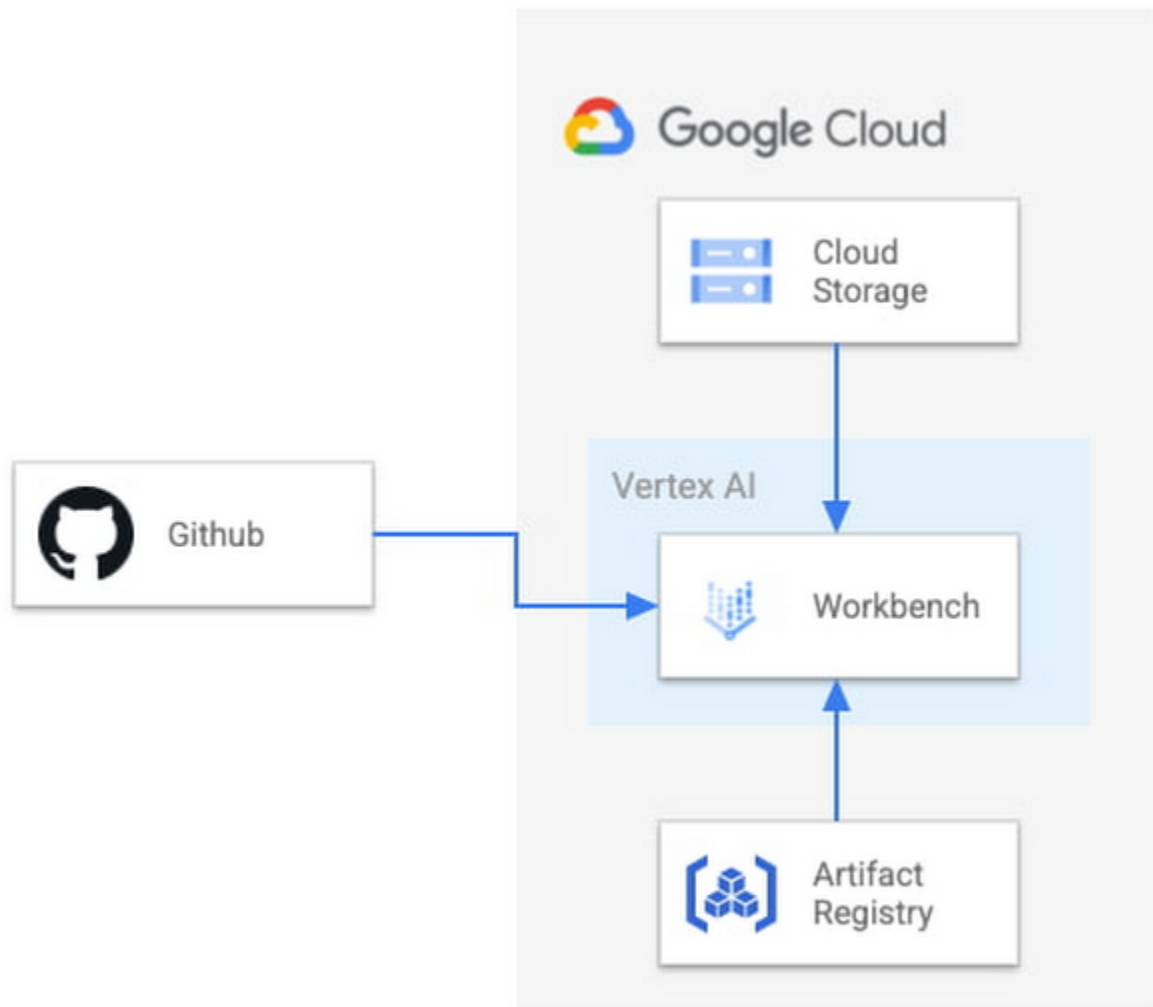


d.

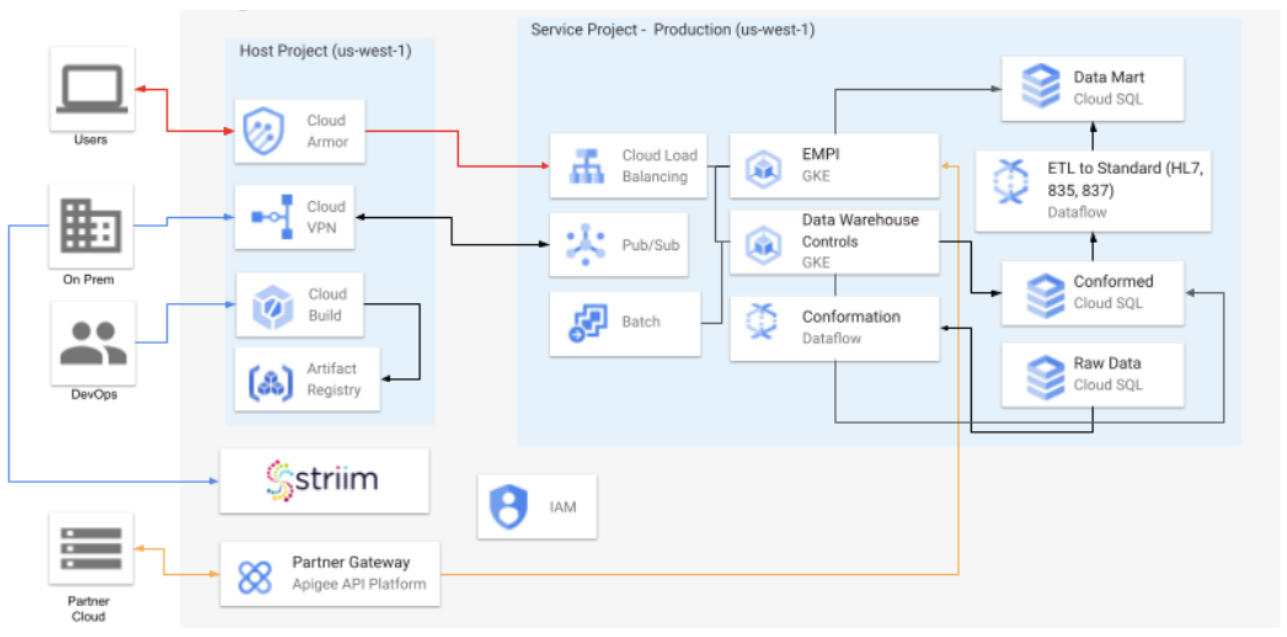
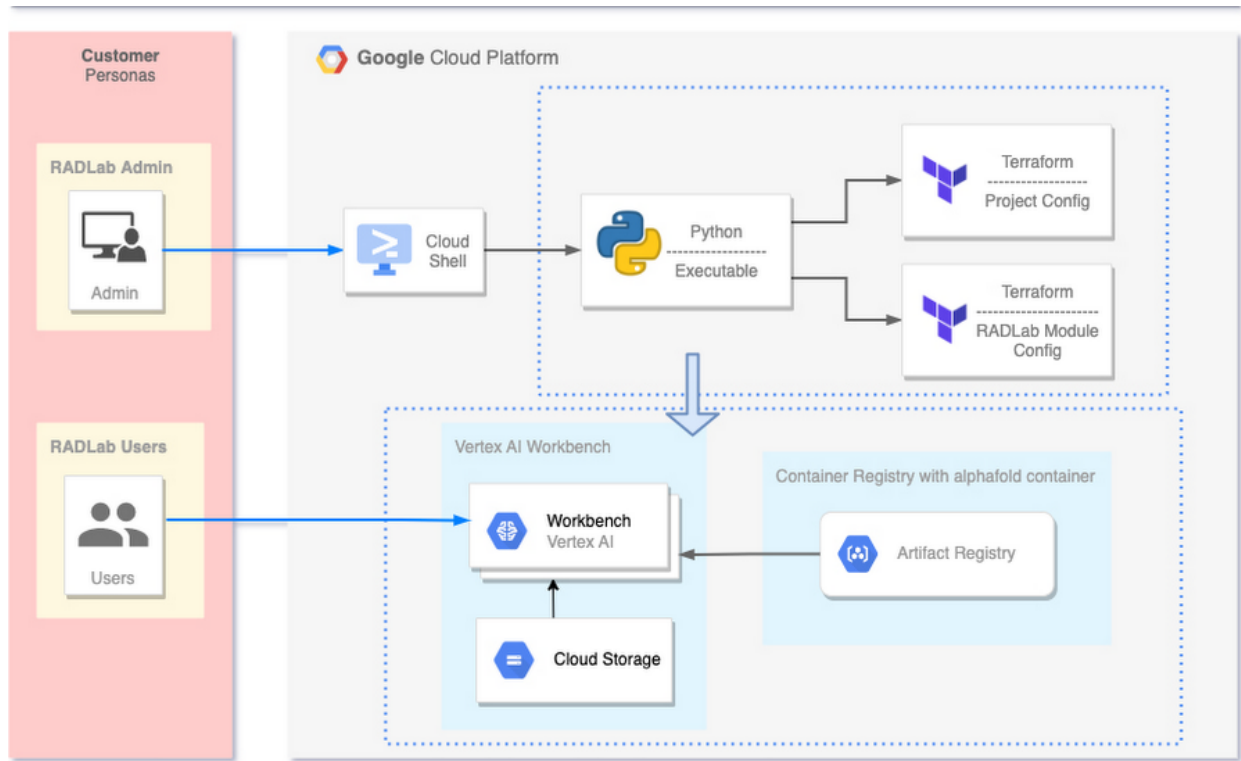
-

Deloitte will implement a solution from a similar proposed cloud architecture as part of this engagement:

Proposed architecture



•



3. GCP Complete Architecture Specifications

Component	Details
Landing Zone	<ul style="list-style-type: none"> • Network Topology <ul style="list-style-type: none"> ◦ Application Users ingress through Cloud Armor WAF ◦ Messaging bridge to on-prem through Cloud VPN • Security <ul style="list-style-type: none"> ◦ IAM policies managed through Cloud Identity ◦ All data is encrypted at rest and in flight throughout Google's network and services
API Gateway	Apigee is a platform for developing and managing APIs. By fronting services with a [roxy layer, Apigee provides an abstraction or facade for the backend service APIs and provides security, rate limiting, quotas, analytics, etc.
Application Hosting	Google Kubernetes Engine <ul style="list-style-type: none"> • Client's sensitive data and Data warehousing Controls application hosted and orchestrated by managed Kubernetes • Client's sensitive data will communicate with on-prem services using Kafka messaging bus or PubSub
Databases	Cloud SQL (for any structured data/metadata) <ul style="list-style-type: none"> • Conformed and raw Databases hosted in PostgreSQL • Striim: One-time and continuous data migration from on-prem SQL Servers to Cloud SQL using Striim on Compute Engine.
Pharma Data Security	Data Loss prevention (DLP) <ul style="list-style-type: none"> • To scan, mask and de identify Personal, sensitive and protected data (PHI/PII)
Message bus and batch jobs	PubSub for cloud based messaging as well as bridging to on-prem using the Kafka connector Batch - Scheduled container and script based jobs
CI/CD	Cloud Build <ul style="list-style-type: none"> • Automated build pipelines triggered by post-commit hooks (from code repository) or manually Artifact Registry <ul style="list-style-type: none"> • Container registry, scanning, and binary authorization
Batch jobs for data mart creation	Dataflow <ul style="list-style-type: none"> • Batch data pipeline to transform Conformed Data to Standard Forms Cloud SQL <ul style="list-style-type: none"> • Datamart Storage
Database search and Deduplication	Google Kubernetes Engine <ul style="list-style-type: none"> • To deploy the database search and clustering algorithm • It will result a JSON object containing the cluster IDs

Vertex AI	Vertex AI Workbench <ul style="list-style-type: none"> Jupyter notebook-based development environment for the entire data science workflow.
Vertex Pipeline for training the clustering model	Vertex AI Pipeline <ul style="list-style-type: none"> To setup and deploy MLOps pipeline for training clustering model. Storing trained weights on Storage buckets for future use. Set up model monitoring for monitoring model metrics Cloud Scheduler and Cloud Function for auto-triggering the model training

4. Scope of work & RACI

The list below is the detailed scope of work and responsibilities of Deloitte and Client BPC for SOW:

Sl. no	Tasks	Deloitte	Client	Week
1	Prerequisites: <ul style="list-style-type: none"> Identifying SPOC from both teams Creating Cloud Billing accounts Setting up project Client's team to share the specifications for Data attributes stored across different tables in the existing database (HL7, FHIR formats) Source code repositories RTO & RPO for Disaster Recovery planning and design 	CI	RA	Week 0
2	Phase 0 - Project	RA	CI	Week 1

	Kickoff: Presentation, Introductions, Communication channels set up			
3	Landing Zone: Discovery, Technical design document, Initial Implementation (IAM, Networking)	RA	CI	Week 1 - Week 3
4	Phase 1 - Application architecture and set up: <ul style="list-style-type: none"> • Data modeling and Cloud SQL, Containerization , Data Warehouse, AI/ML workbench • Troubleshooting • Health check and code update • Review & Sign off 	RA	RCI	Week 4 - 7
5	Phase 2: Dev Environment GKE, CI/CD, SQL, AI/ML model, PubSub, Apigee, MLOPs, Monitoring, Review & Sign off	RA	CI	Week 8 -15
6	Phase 3 - Test Environment GKE, Load balancer, SQL, AI/ML model iterate	RA	CI	Week 16-17
7	Phase 4 - Production Environment GKE, Load balancer, SQL, MLOps (Kubeflow, Pipelines), Monitoring	RA	CI	Week 18-24
8	Testing & Monitoring	RA	CI	Week 24 - 26
9	Knowledge Transfer (KT)	RA	CI	Week 27 - 29
10	Milestone: Production Complete	RA	CI	Week 29 - 34

	<ul style="list-style-type: none"> • Disaster Recovery plan • Functional testing • Load testing • Cutover plan • Initiating cutover • Any changes • Go Live • Review & Sign off 			
11	Final Testing and Validation <ul style="list-style-type: none"> • Validate as per SOW • SLAs and troubleshooting • Customer Sign off on deployment 	ACI	RACI	Week 35-40

RACI definitions:

Responsible	R	Responsible for performing the task
Accountable	A	Accountable for making the business decision or delegating specific tasks to other teams
Consulted	C	Consulted for inputs and feedback; however agreement or action on input is not required
Informed	I	Informed of the final result, task completion, and/or deliverable distribution

5. Deliverables

- GCP Architecture Diagram for the deployments done
- Data Modelling Documentation
- Cost calculator based on the resources deployed
- Technical Design Document (TDD)
- Deployment details for the environment deployed

6. Success Criteria

- a. Successfully migrate (datasets), develop and deploy applications to support the experiment on Cloud for Dev, Test and Production Environments.
- b. ML metrics - test accuracy, efficiency, training hours etc. as in consensus with the client.

7. Out of Scope activities

- a. Installing software/tools and troubleshooting on the on-prem applications that client may already have.
- b. Filesystem encryption and decryption of systems that client may have.
- c. Providing test cases/metrics to calculate success for client's end.
- d. Commitment or SLA's from Cloud Provider (GCP / AWS).
- e. Deployment on any new features released by the Cloud Provider after the day of deployment.
- f. Any workload identified outside the scope mentioned in this document.
- g. Procurement of licenses or subscriptions of any kind.

8. Assumptions

- a. Client BPC will share the user accesses to the necessary systems with Deloitte.
- b. Client BPC will provide necessary information needed to complete the scope of activities as defined in this document.
- c. In case of any delays in the code due to dependencies on the client side will require project extensions with a re-evaluation on time.
- d. Additional licensing costs will be borne by the client.
- e. Client BPC will provide the transformation logic for any conformed data (like to HL7v2, 835 and 837 standards).

9. Roles and Responsibilities

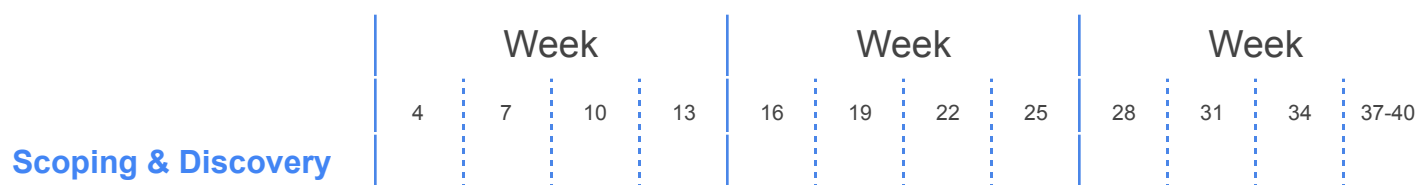
- a. Deloitte Roles and Responsibilities

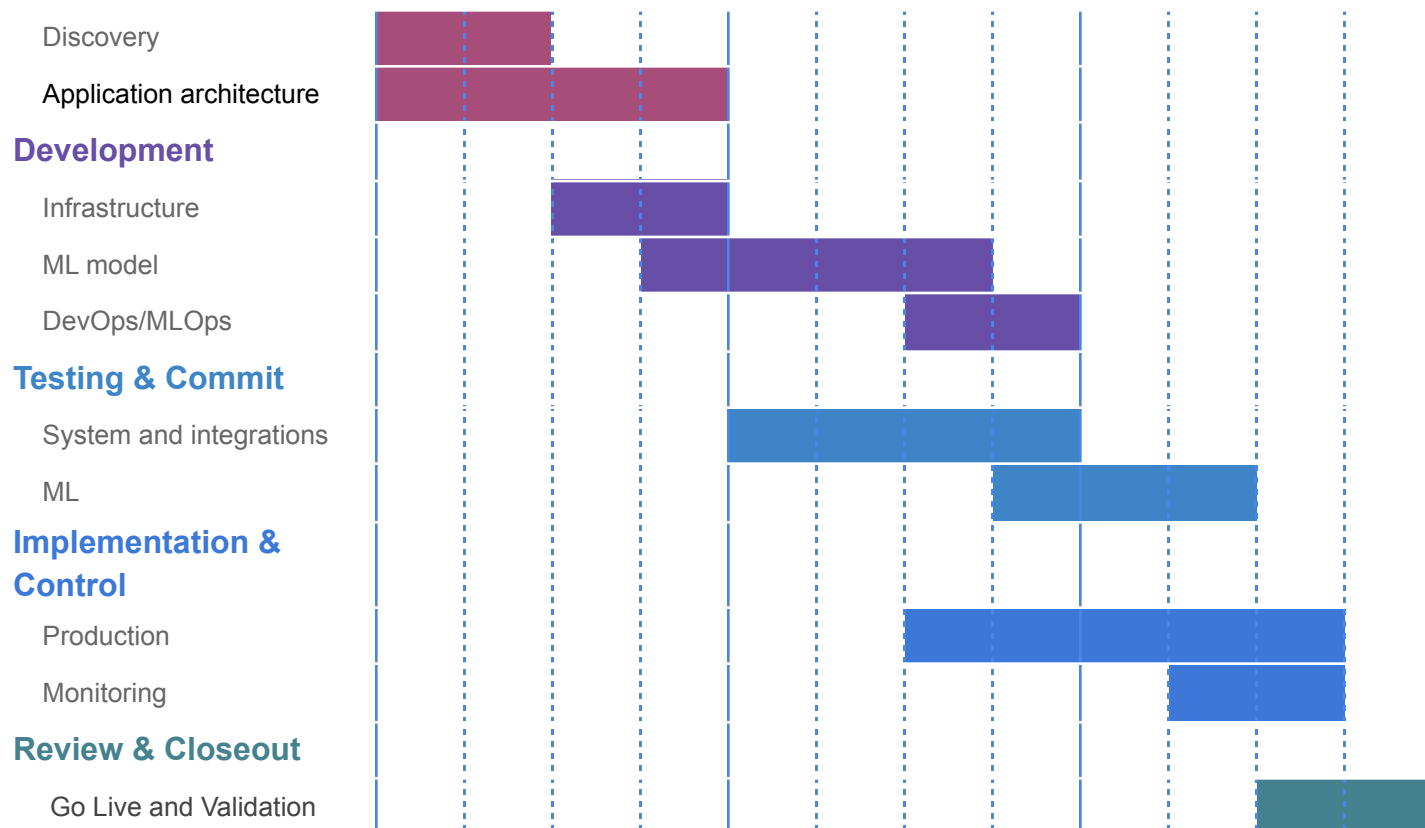
Role	Project Needs (Quantity)	Responsibilities and Skill Sets
Engineering Project Manager	1	Sets up proper execution plan, Daily/weekly updates, Conduct kickoff meeting and weekly sync up calls, escalate issues and address concerns. <u>Skills</u> : Project management concepts, topics and tools.
Cloud Engineer	2	Setting up integrations, DevOps, CI/CD, creating technical

		documents and deployment runbooks, provisioning and configuring API endpoints to proxy applications. <u>Skills</u> : Operational efficiency, Building automation, CI/CD
Solution Data Architect with HLS specialty	1	Landing Discovery, design solution, data modeling and designing, end to end testing with UAT, execute data and ML pipelines. Advise on products for Healthcare and Life Sciences (HLS). Also communicate with Managers and Clients on data as needed. <u>Skills</u> : Documentation, Architecture design, HLS expertise.
Lead Principal Architect with Data/AI/ML Specialty	1	Supervising best practices of relevant Cloud Services, setting up testing end to end deployment, Cost optimizations, Validating the deployed architecture technically. Ensure HIPAA and local laws compliance to the projects. Also communicate with Managers and Clients on both short and long term success as needed. <u>Skills</u> : AI/ML advanced architectures, Cloud best practices
Data Engineer	3	Cleans data, Ingests data, Creation of data pipelines, data processing jobs, data validation scripts, ETL, Scheduling jobs and acting on monitoring/logging data. <u>Skills</u> : Python, Spark, Cloud Big Data, Messaging bus, ETL
Senior Data Scientist	2	Developing ML models for the chemical fingerprinting dataset, testing and improving the model, supporting the architects/engineers in deploying the model on Cloud. <u>Skills</u> : AI/ML concepts, Molecular Data concepts, ML model, Validation, Enhancements, Cloud machine learning

10. Timelines

Estimated Timeline will be a total of 40 weeks from Scoping/Discovery to Go Live/Validate on Production.





11. Pricing

Resources Cost

Role	Rate	Hours	Total (in US \$)
Engineering Project Manager	\$200	120	$120 \times 200 = 24,000$
Cloud Engineer	\$75	240	$2 \times 240 \times 75 = 36,000$
Solution Data Architect with HLS specialty	\$250	120	$120 \times 250 = 30,000$
Lead Principal Architect with Data/AI/ML Specialty	\$300	120	$120 \times 300 = 36,000$
Data Engineer	\$150	480	$3 \times 480 \times 150 =$

			216,000
Senior Data Scientist	\$175	240	$2 * 240 * 175 = 84,000$
Total cost			\$426,000

One Time Implementation project cost

Sl. no	Component	Frequency	Rate (in US \$)
1	Project Cost: <ul style="list-style-type: none"> Resources - \$426,000 Hardware, platform, etc. - \$74,000 	One Time	\$500,000
2	Managing Services post production (if the client needs)	Monthly	\$ as per Deloitte's charge
3	Services Funds from Cloud	One-Time	10% of total cost in 1

12. Stakeholders

Sl. no.	Name	Role	Email ID
Client BPC (RadLab)			
1	Mr. A	Data Scientist	a@bpc.com
2	Miss B	CTO	b@bpc.com
Deloitte			
1	Mrs. C	Director, Cloud Consulting	c@deloitte.com
2	Mr. D	Sr. Director, AMER	d@deloitte.com

13. Terms & Conditions

To be mentioned as per Deloitte's general terms and conditions.

14. Exit terms & Conditions

To be mentioned as per Deloitte's general exit terms and conditions.

15. Acceptance

In Witness Whereof, Deloitte and Client have executed this SOW on the execution date:

Deloitte (Tax ID Number EIN)

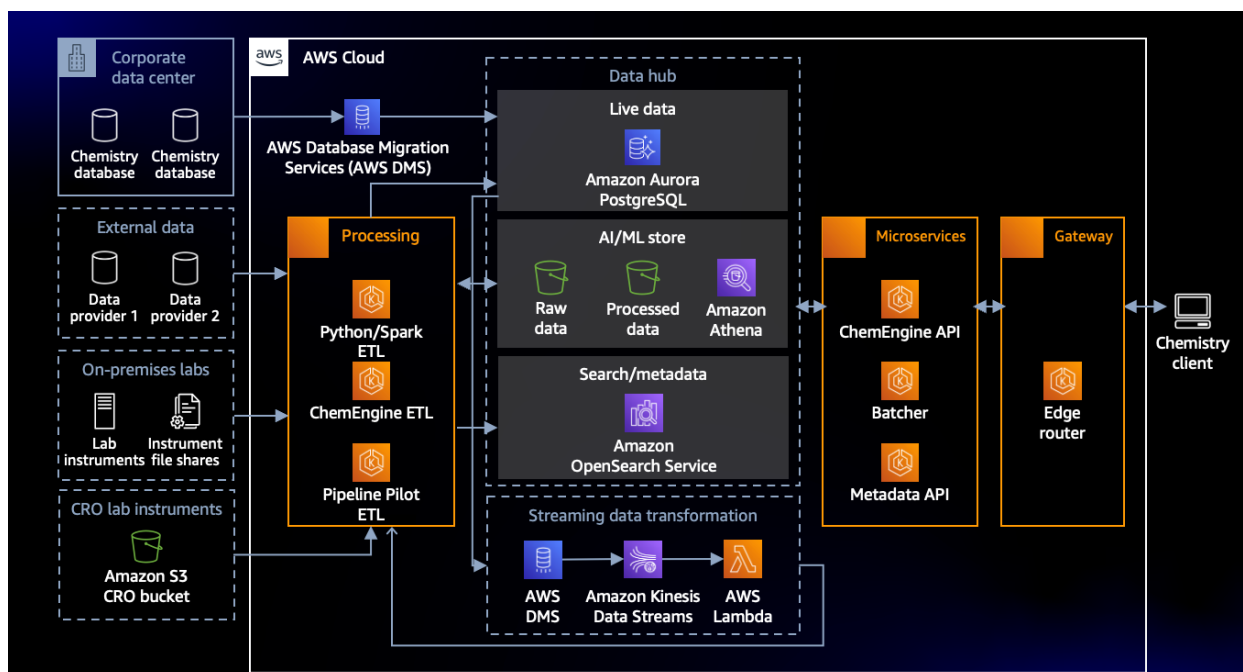
- (Signature by Director, Client Success)

Client BPC (Tax ID Number EIN)

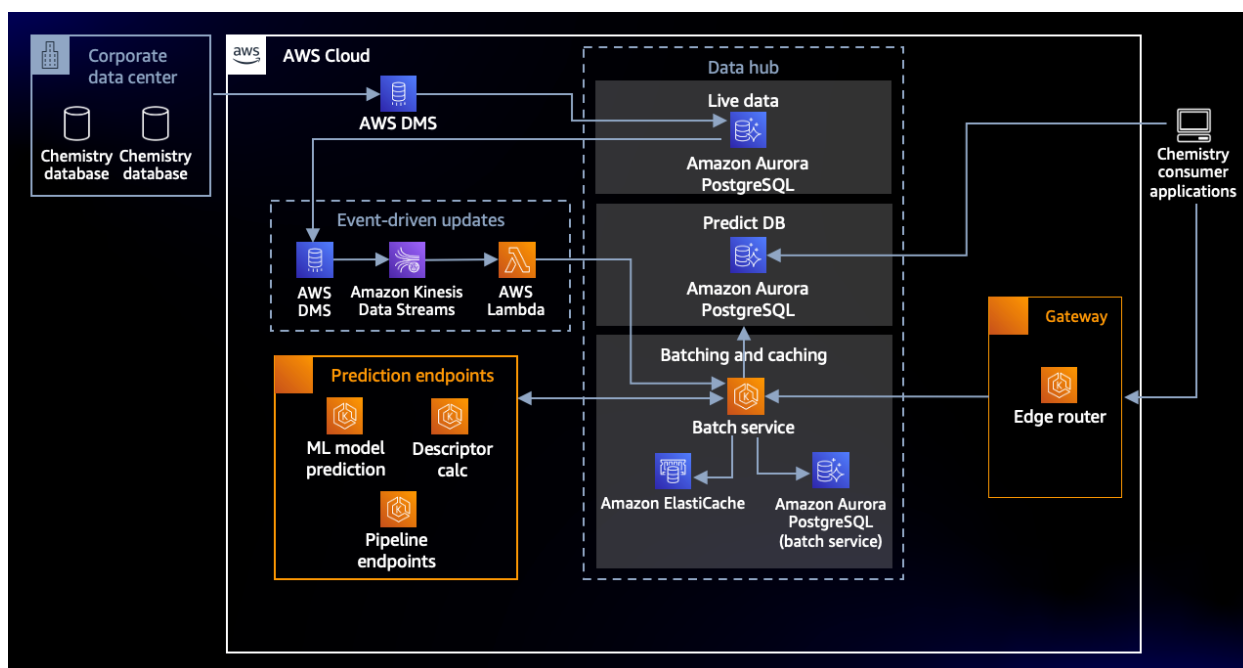
- (Signature by CTO, BPC)

16. Other Alternative Cloud Architectures

AWS:



ii.



iii. Azure:

