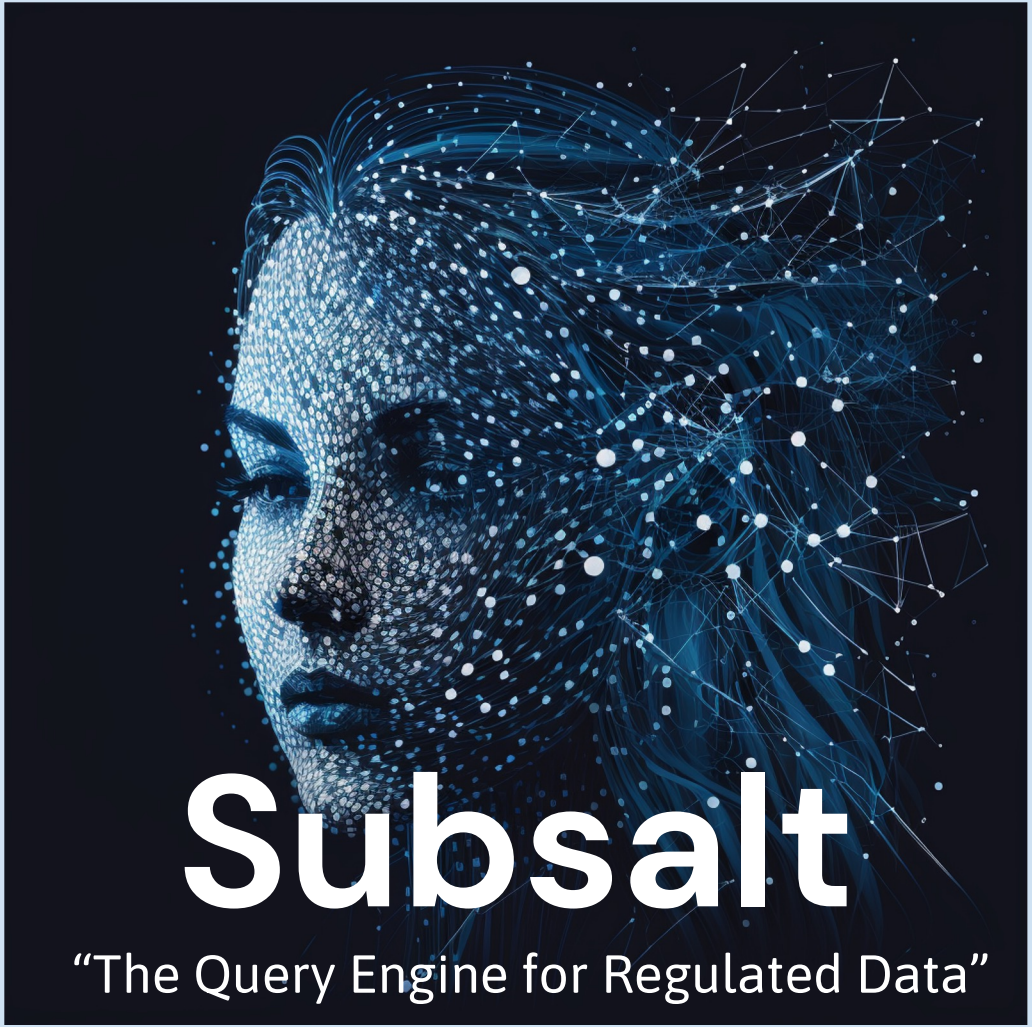

Subsalt
“The Query Engine for Regulated Data”

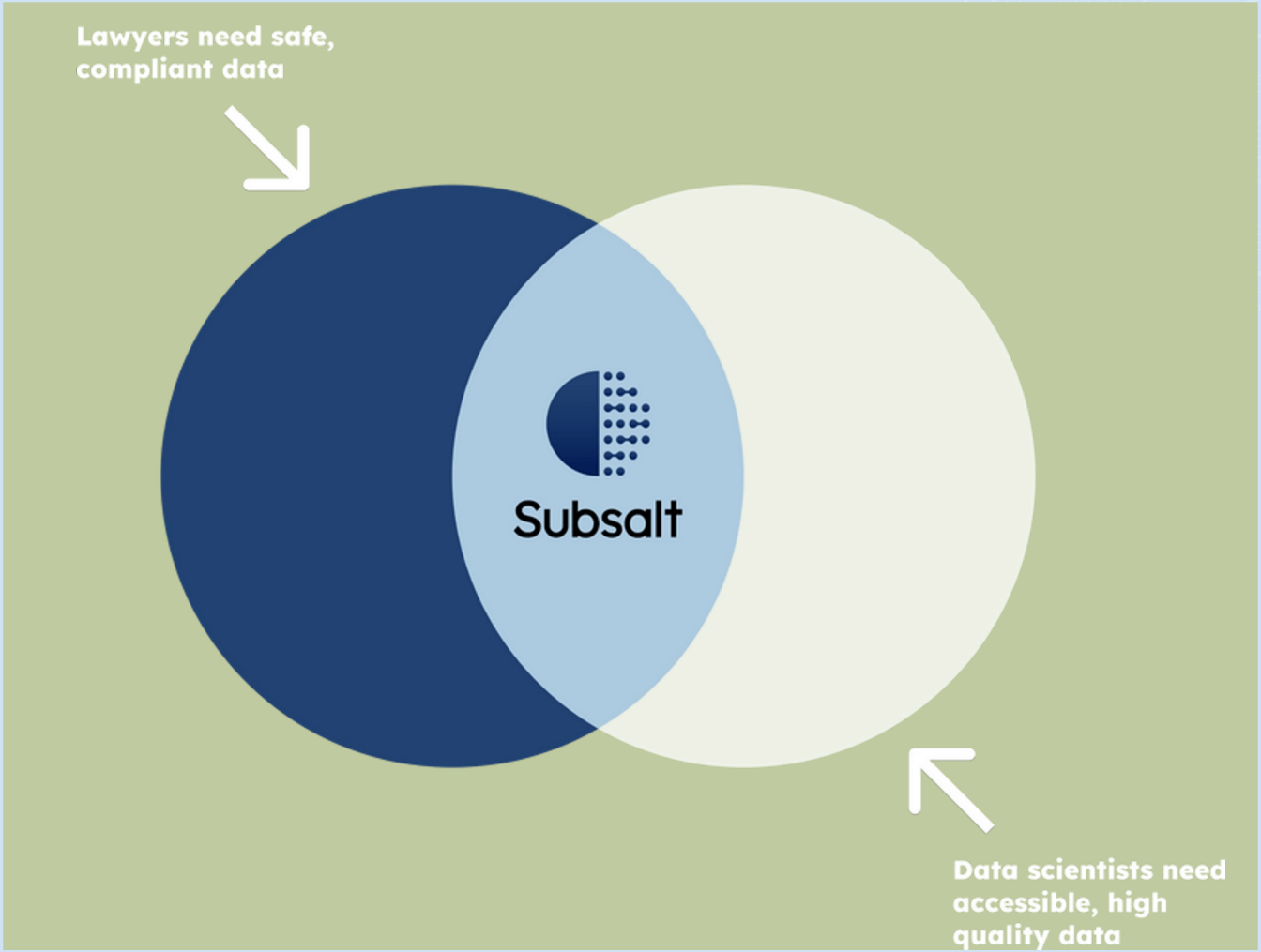
What is Subsalt?

Data is the new oil. It’s valuable, but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc to create a valuable entity that drives profitable activity; so must data be broken down, analyzed for it to have value.”

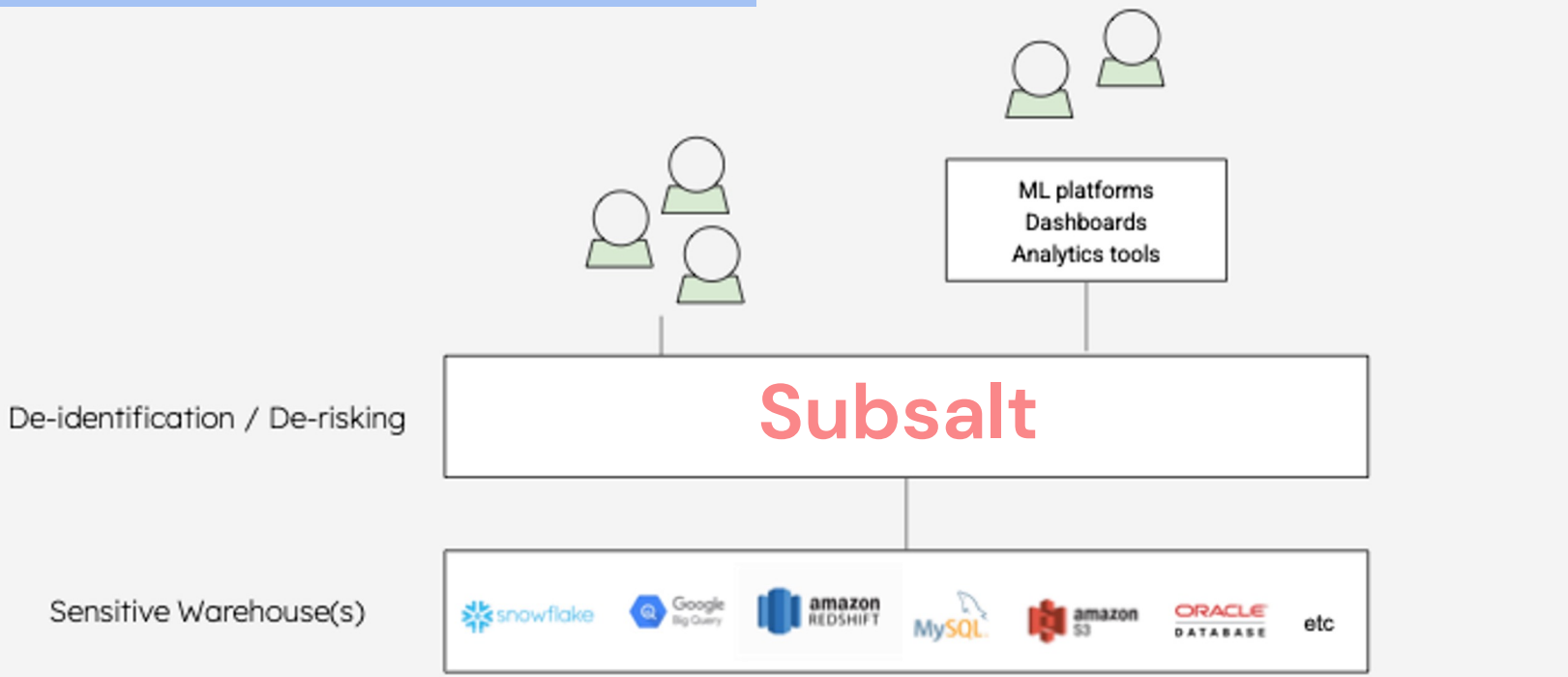
— Clive Humby, 2006

Background

- Healthcare
- Consumers
- Finance
- Defense

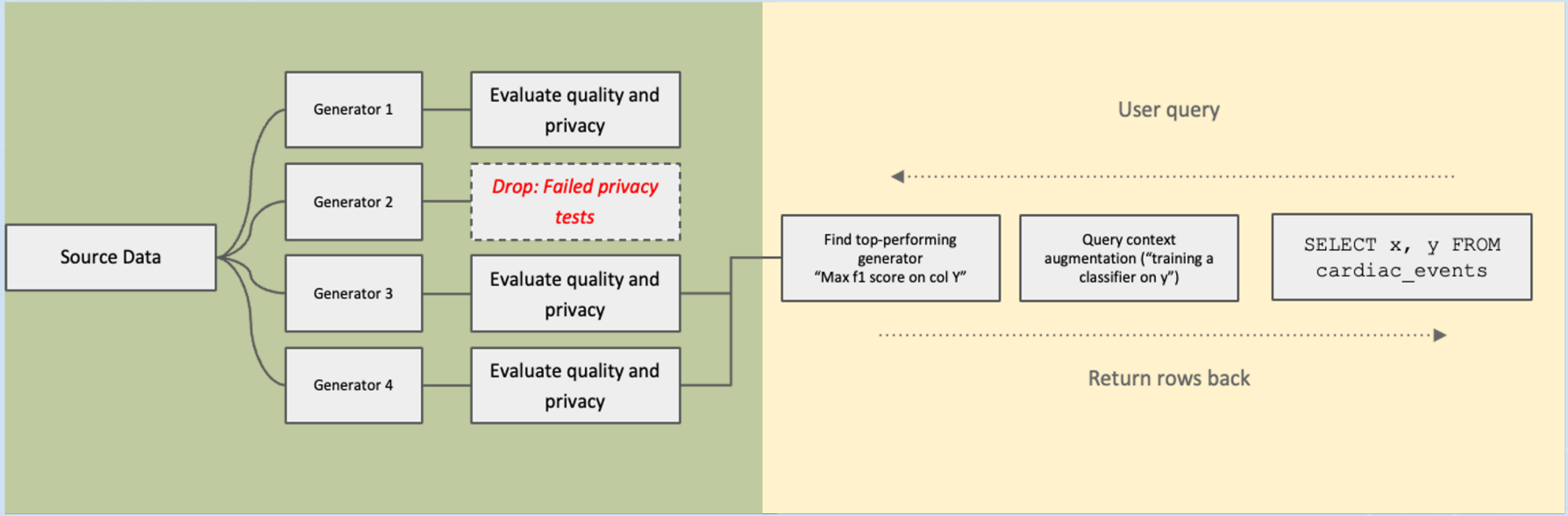


Production Process



© Subsalt, 2022

Developmental Cycle



The Problem

“High computational costs and repetitive privacy testing make large-scale data training slow and costly, reducing the product's value by hindering fast, efficient access to sensitive data.”

Current Pitfalls

- Timings for training and privacy
- Computational costs
- Privacy tests still running post failure
- No filtering process before training
- Have to train repetitively to check privacy

PRIVACY LAWS

Privacy Testing

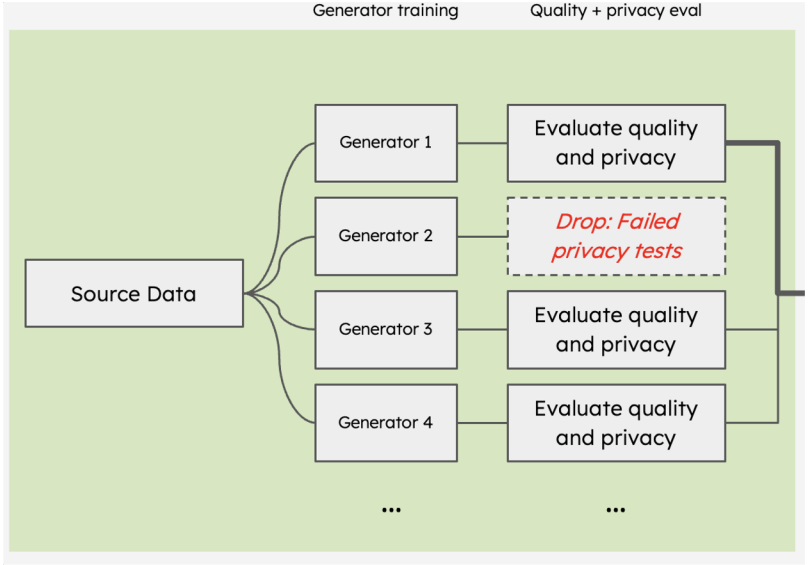
Risk Assessment for Attribute Inference

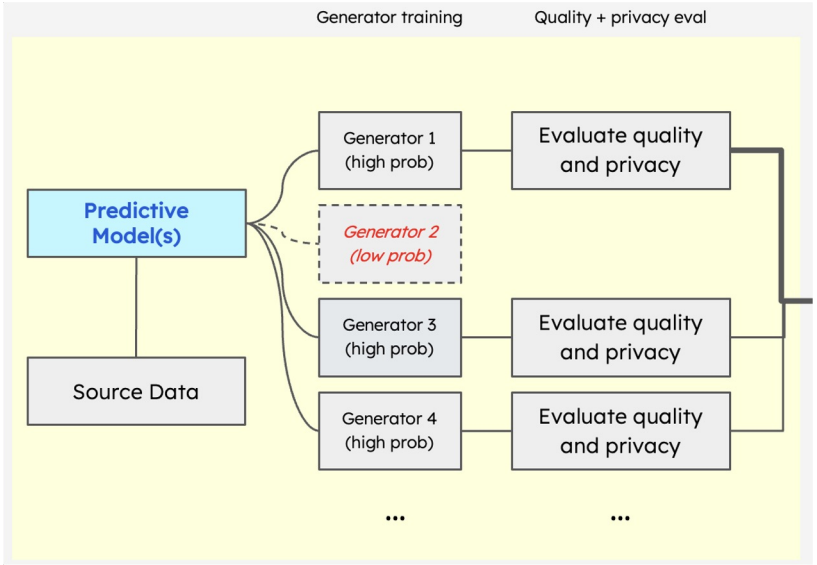
- Attacks that target the features in the data as a group in order to gain insights

Precision in Membership Inference

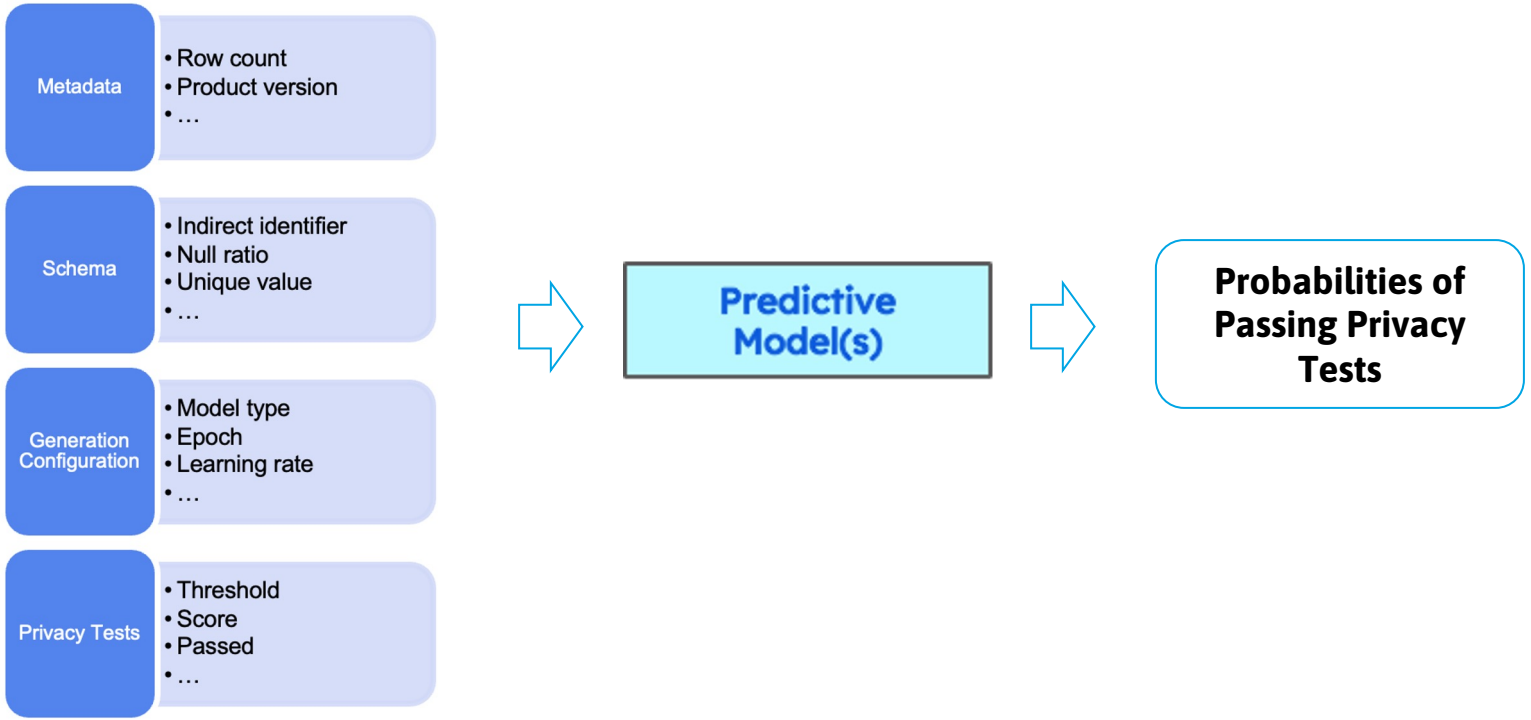
- Target specific data points

Proposed Solutions

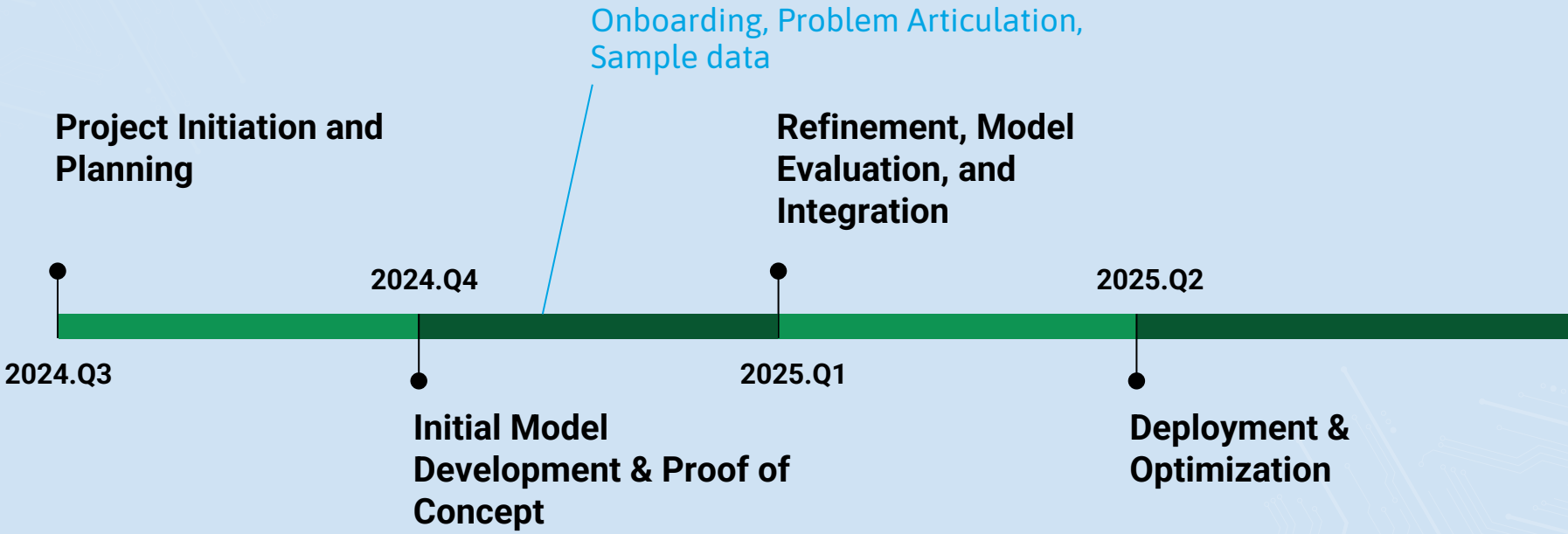




Proposed Solutions

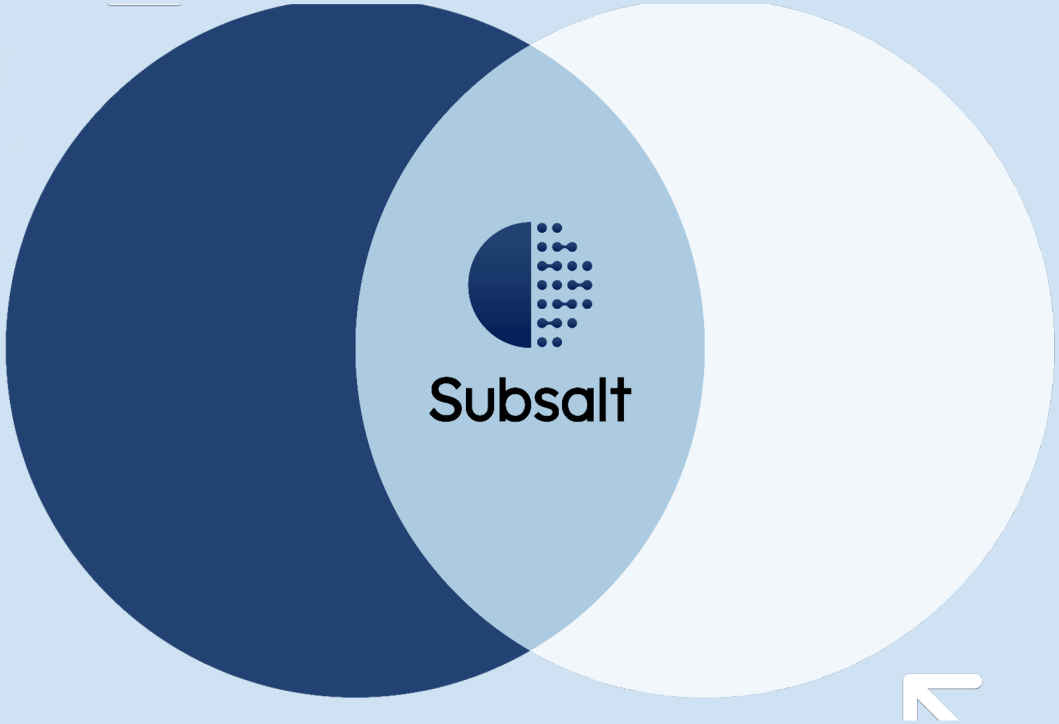


Project Timeline



Potential Risk

01	Usage of Public Datasets	• Use business data • Use synthetic data generated from business data • Use historical report
02	Gap of Documentation vs. Sample Data	• Dual diligence • Agile
03	**Schedule Management**	• High Availability • Flexibility



THANK YOU FOR LISTENING. PLEASE DONT GRILL US OR WE WILL COME AFTER YOU!