TFS DATA PORTAL

Keenan D'spain, Khiem Nguyen, Kundan Chaudhary, Loc Pham, Sabina Khanal

MEET THE TEAM

Kundan Chaudhary

Team Lead
DevOps

Khiem Nguyen

Backend
Lead

Keenan D'spain

QA
Lead

Loc Pham

Frontend
Lead

Sabina Khanal

Database
Lead

Bingyang Wei

Advisor

TOYOTA TEAM

- Project Leadership and guidance
  - Gaurav Lall – gaurav.lall@toyota.com
- Project Co-ordination and technology support
  - Tai Trieu – tai.trieu@toyota.com
- Axon and EDC (Data Catalog and Business Glossary)
  - Aritra Das – Aritra.das@toyota.com
- Informatica Data Quality (IDQ)
  - Arpita Santra – Arpita.Santra@toyota.com

AGENDA

- Problem Statement
- Project Background
- Sprint 1
- Sprint 2
- Sprint 3
- Sprint 4
- MVP Demo
- Winter Sprint

PROBLEM STATEMENT

The problem of

- *Disjointed system to record, store, check and correct all the data in the ecosystem/ no holistic view of data*

affects

- *Employees/ Business partners of TFS*

the impact of which is

- *Unorganized data, manual process of linking physical and business data elements which is time consuming*

a successful solution would be

- *Data to be organized by business areas such as Loan Originations, Insurance, Servicing etc.*
- *Ability to do google like search for any data element which would bring up business definitions, physical attributes, data quality rules & profile and any related data associated with it*

PROJECT BACKGROUND

Create a portal where we can visualize data duly categorized that end users of TFS can use for visualization and solving their business problems

Connect the datasets spread in different systems (Axon, Informatica Data Quality and EDC)

Use CKAN as Database Management System

SPRINT #1(OCT 6 - OCT 20)

Explore CKAN

An open-source DMS (data management system)

Powers data hubs and data portals.

Makes it easy to publish, share and use data.

Launch CKAN on AWS

Understand project requirements

Created Use Cases

SPRINT #2(OCT 20 – NOV 03)

Upload Example Data

Upload data using CKAN UI

Upload data using python script

Software Architecture

Learn Data classification

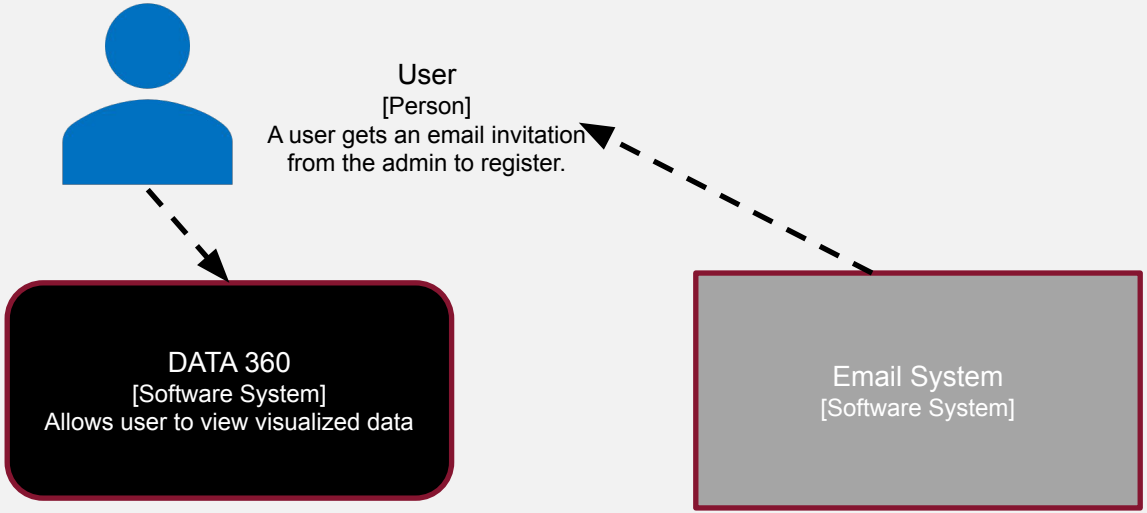
Analyze Data connectivity

Design the architecture

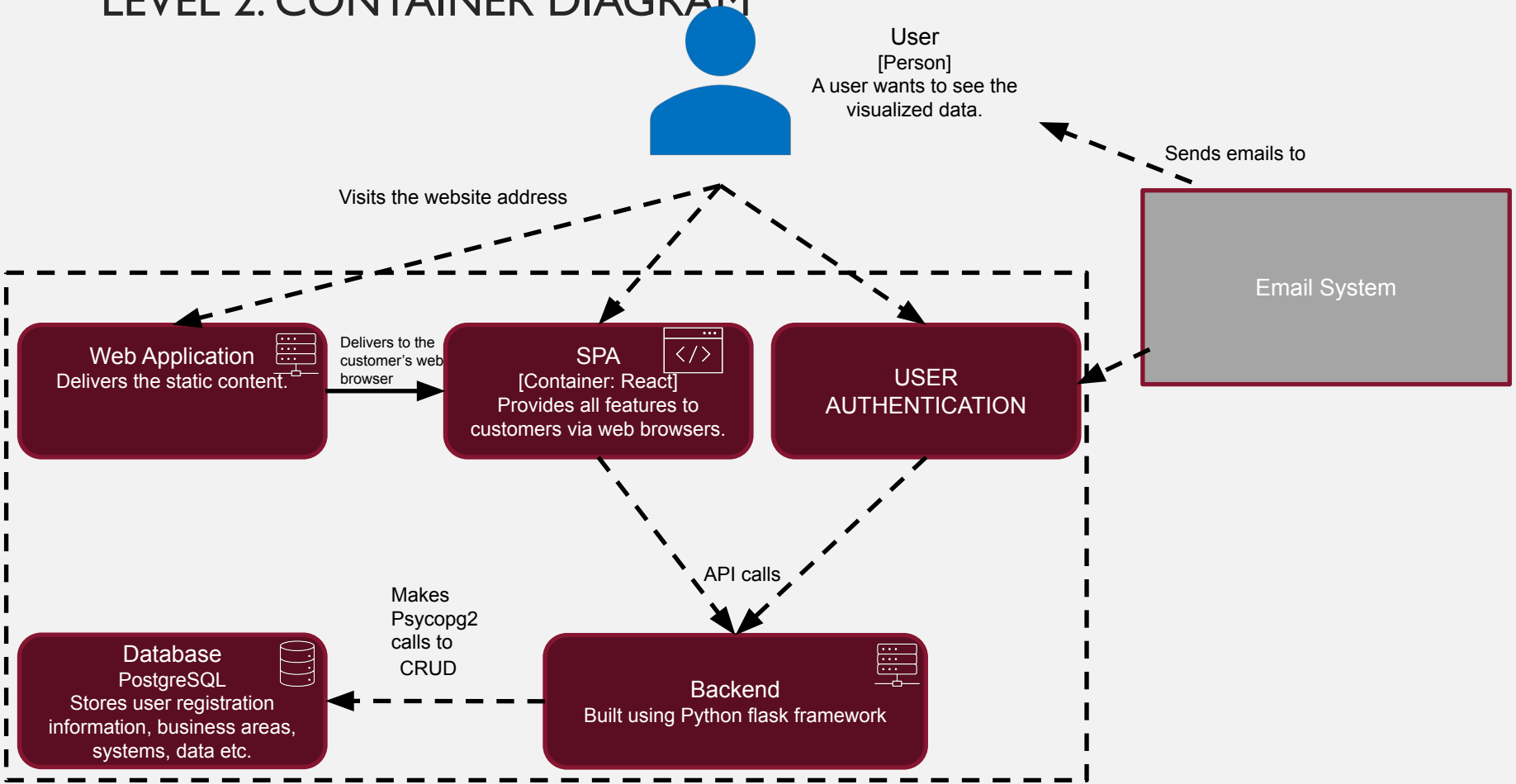
Application Prototype

Link for the prototype

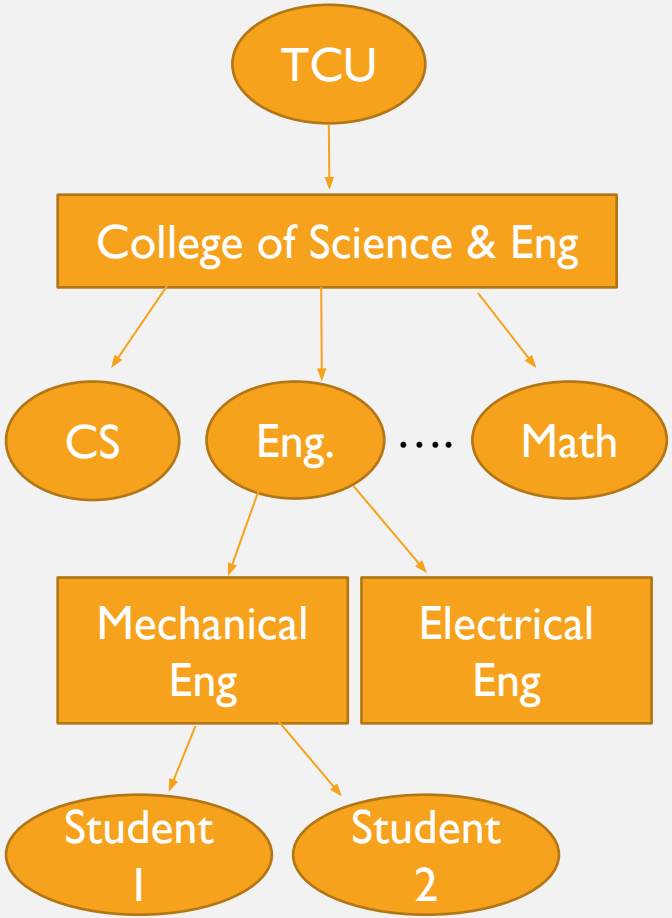
LEVEL 1: SYSTEM CONTEXT DIAGRAM

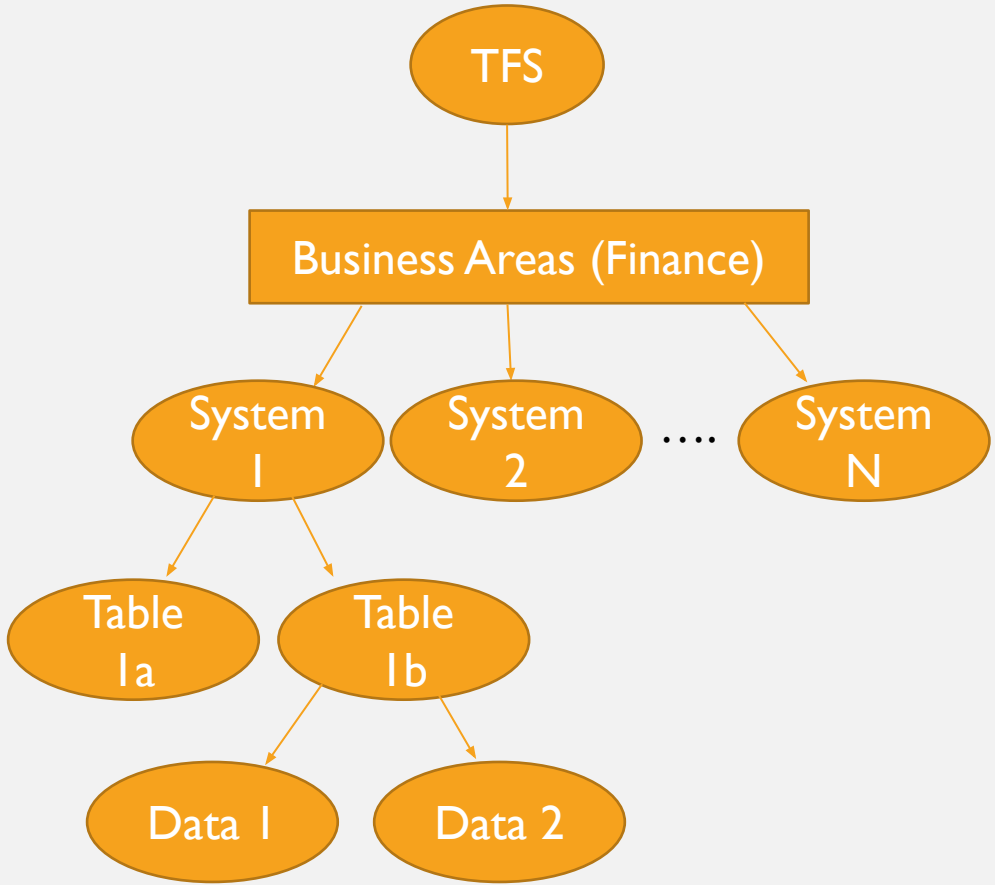


LEVEL 2: CONTAINER DIAGRAM

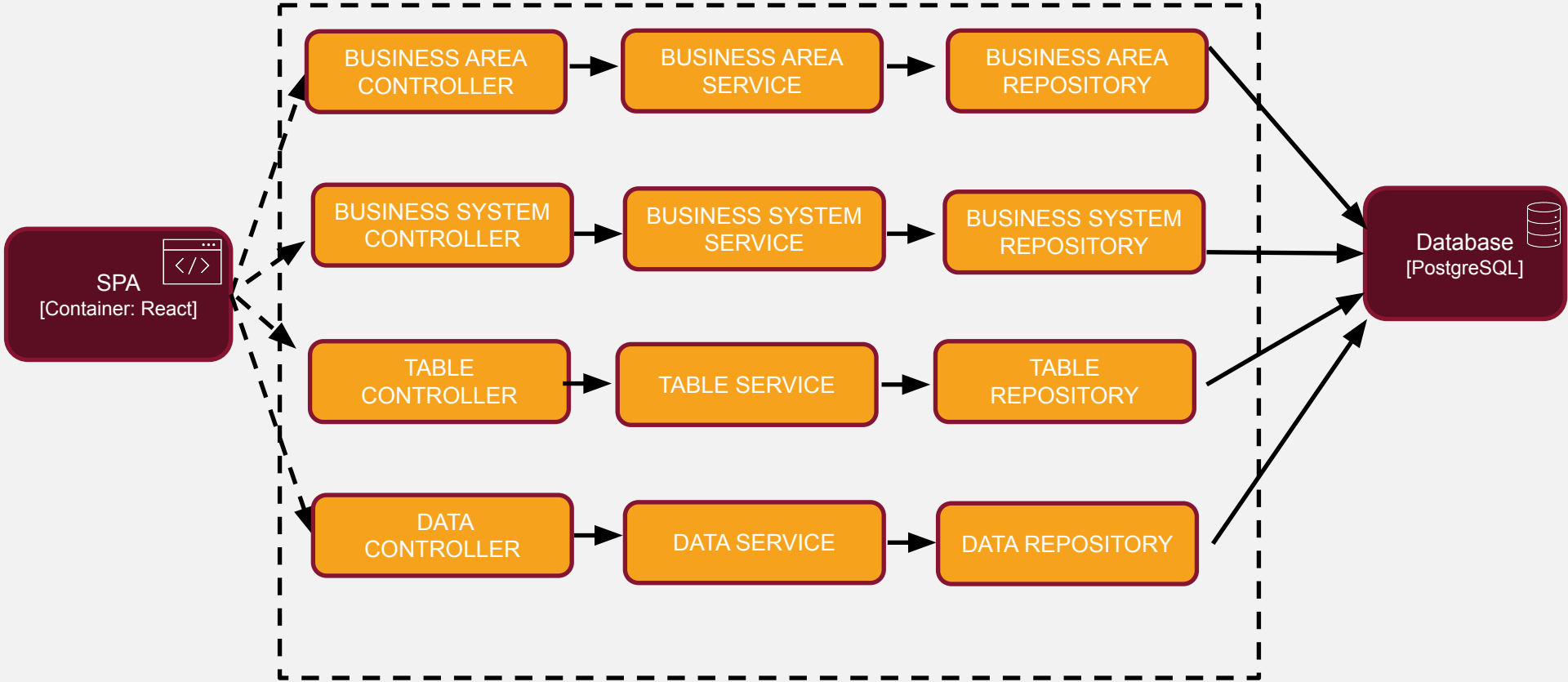


BACKEND


TCU
College of Science & Eng
CS
Eng.
....
Math
Mechanical Eng
Electrical Eng
Student 1
Student 2


TFS
Business Areas (Finance)
System 1
System 2
....
System N
Table 1a
Table 1b
Data 1
Data 2

LEVEL 3: COMPONENT DIAGRAM



SPRINT #3 (NOV 03 – NOV 17)

Analyze CKAN Risks

Lack of resources

Lacks helps from developers

Not what we need

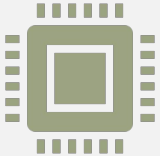
Difficult to modify

Setup Development Environment From Scratch

Setup Flask app for backend

Setup React NextJS for frontend

Setup PostgreSQL Database



Build the Business Area Domain (end-to-end)

Build Model (Database Table), Controller, Service and Repo

SPRINT #4 (NOV 17 – DEC 1)

Build the System domain (end-to-end)

- Build Model (Database Table), Controller, Service and Repo

Deployed frontend NextJS on Vercel

- Platform for frontend frameworks and static websites

Deployed backend Flask app on AWS

- Using Elastic Beanstalk EC2 with CodePipeLine

Set up PostgreSQL database on AWS RDS

Integrated CI/CD with Github

- Continuous Integration/ Continuous Delivery/Deployment

MVP DEMO

WINTER SPRINT (DEC 19 – JAN 08)

- Build the Table domain (end-to-end)
  - Build Model (Database table), Controller (JSON), Service and Repo
- Pre-process Data
  - Python script to analyze and connect data
  - Store the processed data in the database

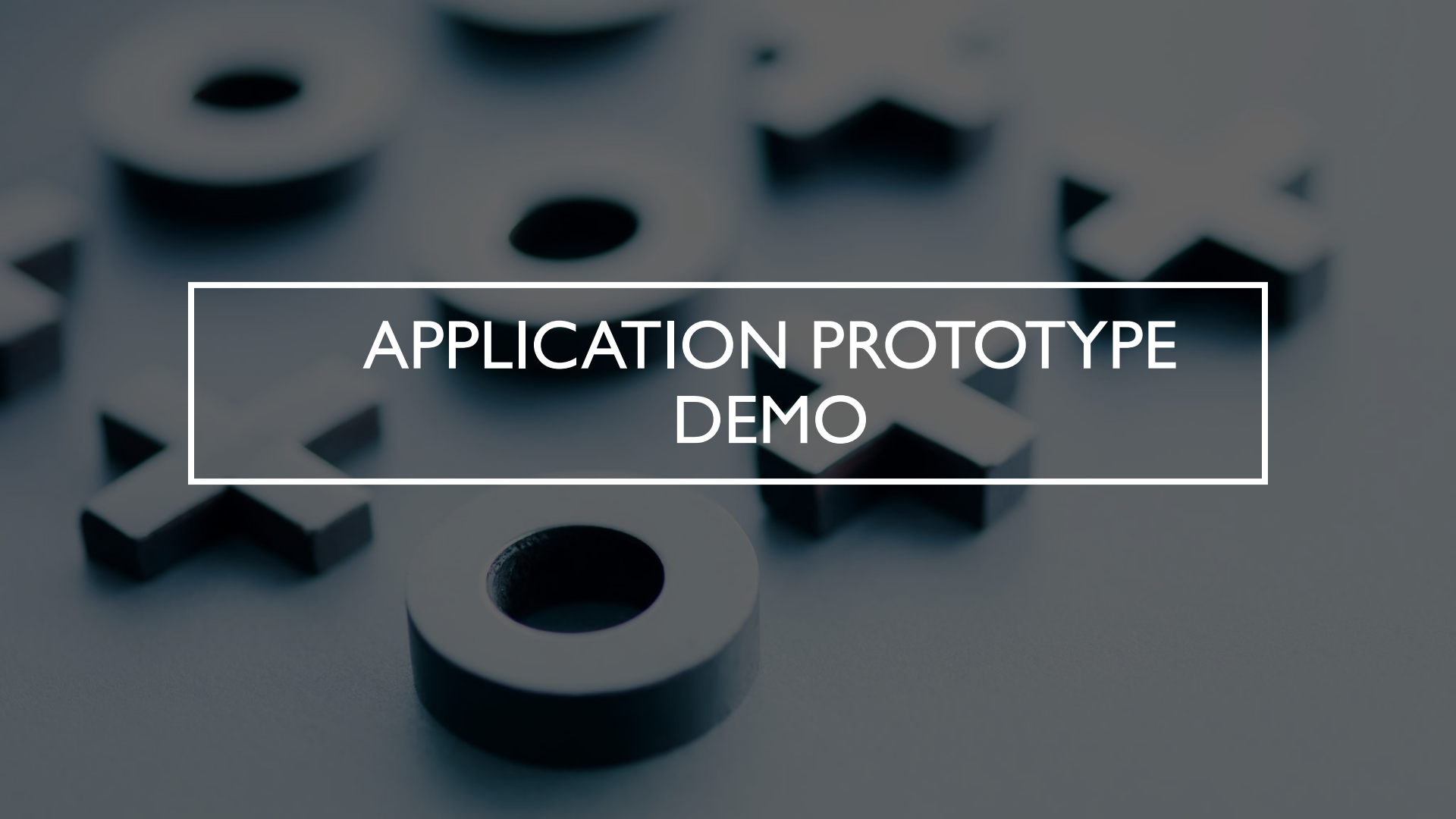
PLAN FOR REMAINING SPRINTS

Build Data domain

User Management feature

- Create Identity Access Management for admin and user roles

Implement Google-like intelligent search



APPLICATION PROTOTYPE DEMO

Thank you!

Questions?