



TFS DATA PORTAL

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

MEET THE TEAM

Kundan Chaudhary



Team Lead
DevOps

Khiem Nguyen



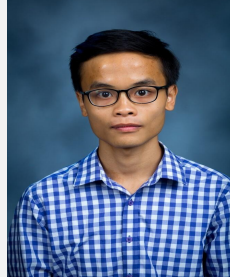
Backend
Lead

Keenan D'spain



QA
Lead

Loc Pham



Frontend
Lead

Sabina Khanal



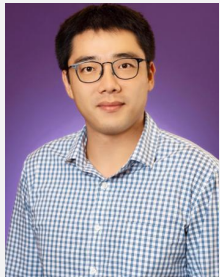
Database
Lead

Krishna Kadiyala



Advisor

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

Information Security

- Boulton Fernando

AGENDA

- Problem Statement
- Project Background
- Data Hierarchy
- EDC, Axon, and IDQ
- Software Architecture
- Database (PostgreSQL)
- Development
- Demo



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*



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



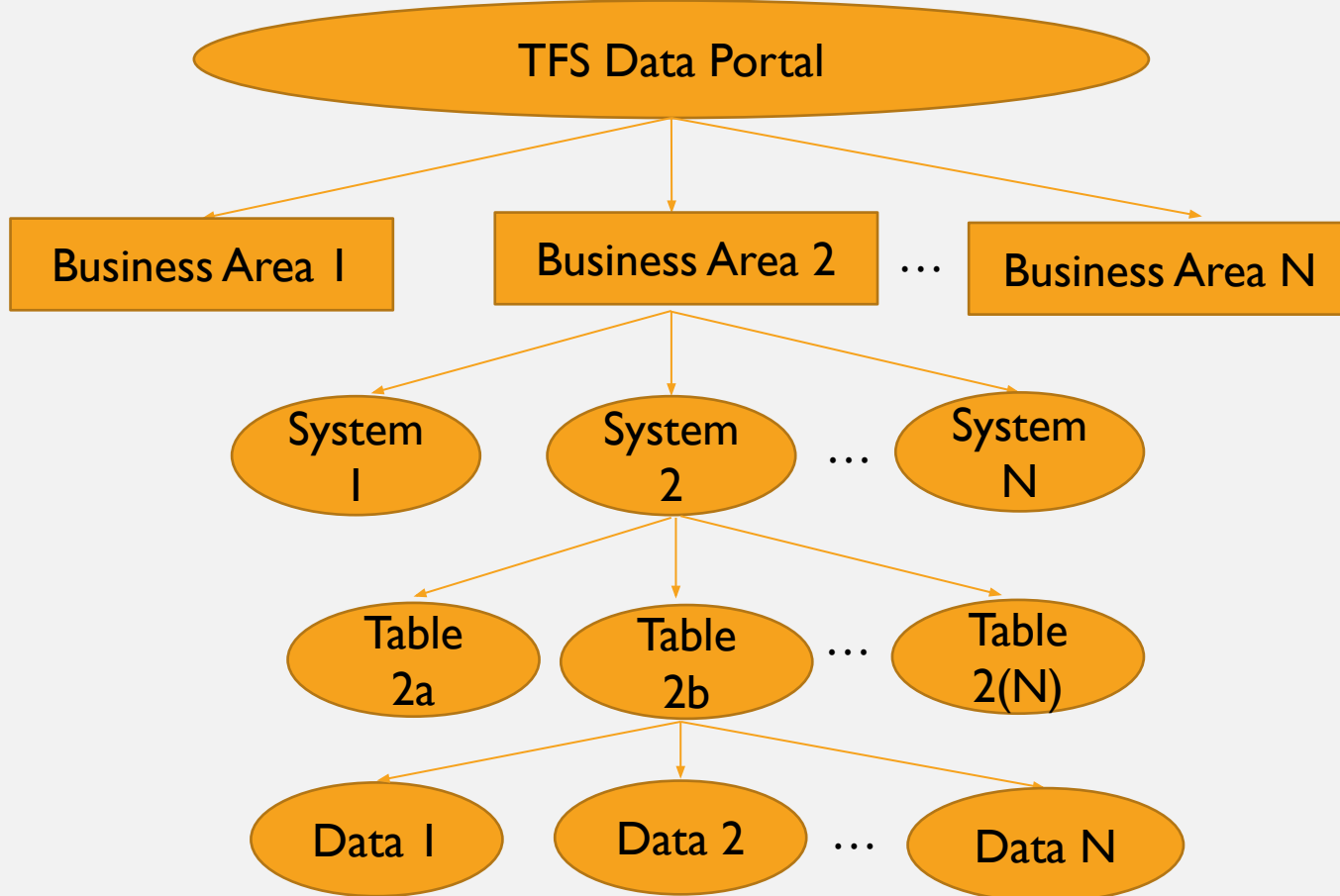
Connect the datasets spread in different systems (Axon, Informatica Data Quality and EDC)



Use CKAN as Database Management System (Requirement changed to Postgres later)

PROJECT BACKGROUND

Data Hierarchy



EDC, Axon and IDQ

EDC

- Stands for Enterprise Data Catalog
- AI-powered data catalog
- Provides a machine-learning- based discovery engine to scan and catalog data assets across the enterprise
- EDC extract is used to create tables and initiate the data within that table in our web portal

Axon

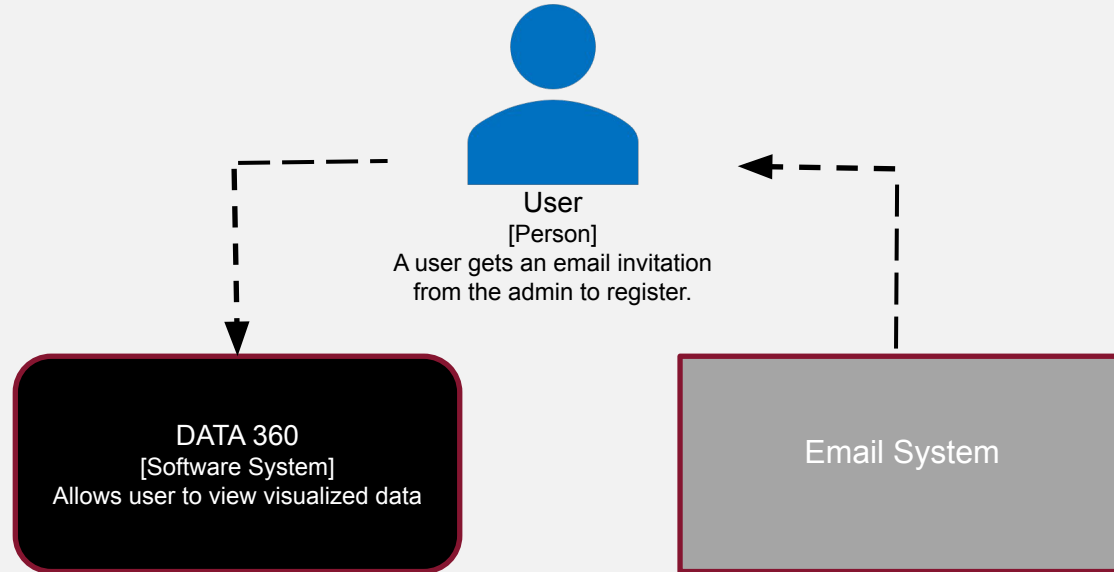
- A popular data governance tool
- Facilitates a collaborative environment to collect different data objects (metadata collection) and link them
- Axon extract provides all the relevant information about the data given the data is already initiated

IDQ

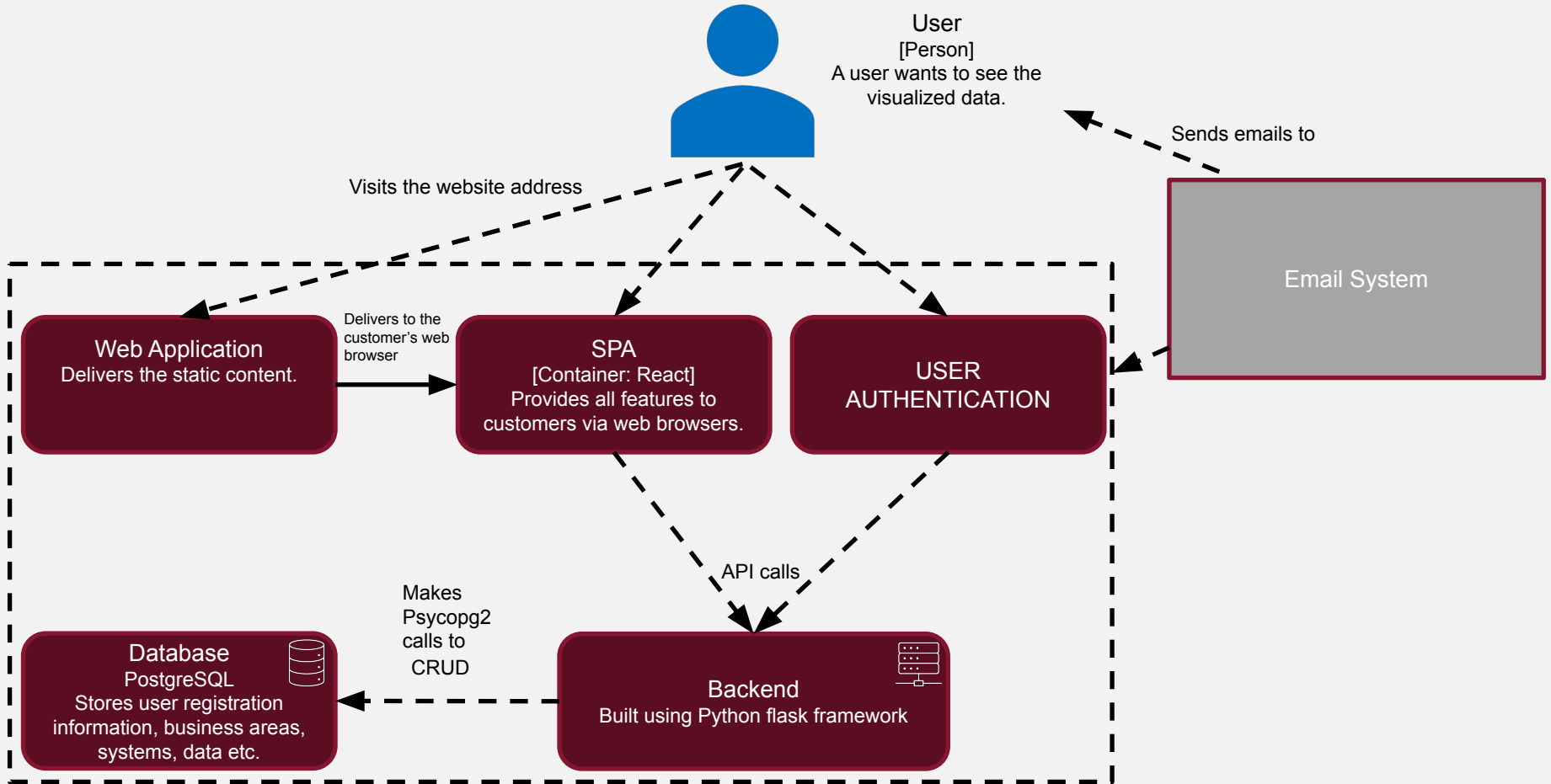
- Informatica Data Quality uses a unified platform to deliver quality data for business initiatives and applications
- IDQ extract provides a scorecard for data quality and other relevant information about their property given the data is already initiated

Software Architecture

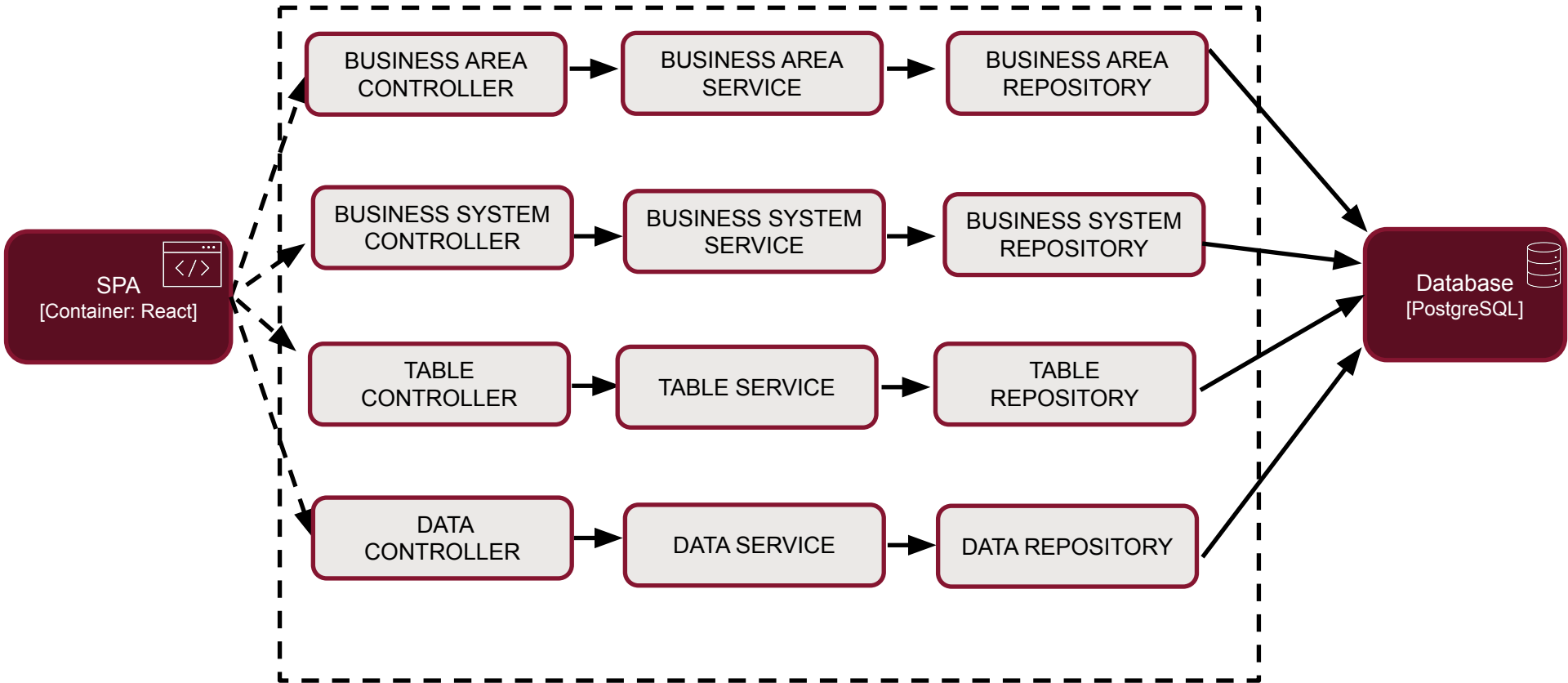
LEVEL 1: SYSTEM CONTEXT DIAGRAM



LEVEL 2: CONTAINER DIAGRAM



LEVEL 3: COMPONENT DIAGRAM



Database (PostgreSQL)

- Highly stable database management system
- Backed by more than 20 years of community development which has contributed to its high levels of resilience, integrity, and correctness
- Used as the primary data store or data warehouse for many web, mobile, geospatial, and analytics applications.
- Supported by AWS

PostgreSQL



DEVELOPMENT

Build the Business, System, Table and Data 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



DEMO

Experiences and Takeaways

- Making the software architecture flexible
- Expect changes from the client
- Sometimes expect the client to change
- Reasonable Negotiations
- Team Collaboration



Thank you!

Questions?