

MICHIGAN STATE
UNIVERSITY

Project Plan Presentation

Microsoft Excel Data Extractor/Modeler

The Capstone Experience

Team Roosevelt Innovations Knowledge Science

Brendan Bushbaker

Kathryn Nagy

Het Patel

Rithwik Pulicherla

Adrian Self

Department of Computer Science and Engineering

Michigan State University

Spring 2024



*From Students...
...to Professionals*

Project Sponsor Overview

- Provides SaaS for Insurance Analysts
- Founded as a subsidiary of Delta Dental of Michigan
- Headquartered in Okemos, Michigan

Project Functional Specifications

- Define relationships between data in Excel workbook
- View data relationships with intuitive GUI
- Convert insurance rate calculation data to a more flexible format
- An easy-to-use web interface that helps save time and effort and gets the operation done in an efficient manner



Project Design Specifications

- Upload and view Excel workbooks in web application
- Add labels and measurements to workbooks
- View relationships between labels and measurements with modeler
- Export data as insurance rate calculation represented by GRACE, a domain-specific language

Screen Mockup: Upload

Excel Data Extractor

Spreadsheets		
RooseveltGraceTester.xlsx	OPEN	MODEL
DataModeler.xlsx	OPEN	MODEL
RooseveltTestingWorksheet.xlsx	OPEN	MODEL
PatientRatesJan2024.xlsx	OPEN	MODEL
RooseveltGraceClaims.xlsx	OPEN	MODEL
RooseveltWorksheetPatientX.xlsx	OPEN	MODEL

UPLOAD WORKBOOK

Screen Mockup: Spreadsheet View

			2021			2022			2023			D3 Customer	
			Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Name	Role
Michigan	Individual	PPD	3.3000339803C	0.0194713489E	5.2203202187D1	1.18902218807	7.3000332883E1	8.9	4.31784434209E	5.0	7.8		
		Promer	3.8204908833E	6.7380942289E	3.30989361707E	6.3049834418E	8.8317801722E	3.7098921247E	3.7388621247E	5.7388621247E	3.7388621247E		
		OCN	4.1602900039E1	4.0549504370E	5.1726410030E	3.4862279291E	5.3018840559E	6.4059221208E	3.43	2.89	5.1726410030E		
Spouse	PPD	7.8306881948E1	1.1123687928E	3.4843031811E	1.8682273008E	2.5372951829E4	7.3680000209E	8.46	1.1123687928E	5.4843031811E			
		Promer	2.811460888E	0.782078807E	7.8888812954E	2.012080588E	3.4868211881E	6.8632948108E	6.18	0.782078807E	7.8888812954E		
		OCN	2.3200432152E	5.7280001154E	1.8964798071E	5.002028379E	3.7298210049E	0.078871060E	2.811460888E	5.7280001154E	1.8964798071E		
Dependent	PPD	1.3817384334E	4.2880881831E	4.5744313825E	3.5887602707E	5.1671820964E	0.728715911E	2.208042076E	4.2880881831E	4.5744313825E			
		Promer	7.980217052E	4.438482942E	4.900309645E	4.012443816E	6.844816079E	1.8116274827E	1.3817384334E	4.438482942E	4.900309645E		
		OCN	3.375807811E	2.8548168743E	3.900300084E	3.058870708E	1.336240648E	3.706118647E	7.288417052E	2.8548168743E	3.900300084E		
Ohio	Individual	PPD	7.1294911143E	0.3110079102E	4.8749921136E	0.0789068172E	1.4240001239E	0.0040000000E	5.3798887801E	0.3110079102E	4.8749921136E		
		Promer	1.7870448821E	1.2718849321E	4.0706521893E	3.008948200E	4.0030403080E	0.0226048488E	7.1294911143E	1.2718849321E	4.0706521893E		
		OCN	3.8804386748E	4.8342045254E	3.8325472843E	8.1810078148E	1.2882828282E	1.400481843E	1.7870448821E	4.8342045254E	3.8325472843E		
Spouse	PPD	1.3303108970E	7.426788806E	1.8110918008E	6.4481289123E	4.1884870073E	3.0071819112E	3.8842861847E	7.426788806E	1.8110918008E			
		Promer	8.2244758878E	5.7082181163E	4.1728432031E	4.748482834E	6.8871288381E	6.5128448388E	1.3303108970E	5.7082181163E	4.1728432031E		
		OCN	7.4023818670E	4.0223279073E	1.8786146106E	2.8327477000E	6.888030301E	1.888352134E	8.4484875887E	4.0223279073E	1.8786146106E		
Dependent	PPD	3.8978573330E	2.752292825E	4.2504147170E	1.0645028174E	2.4808470157E	0.2457401733E	4.223818078E	2.752292825E	4.2504147170E			
		Promer	5.802580894E	1.139019385E	7.290488484E	2.1670589548E	4.1412868381E	0.732791488E	3.8978573330E	1.139019385E	7.290488484E		
		OCN	5.701381174E	4.8620752027E	3.1796654715E	2.0603328176E	3.8431057638E	0.721249125E	5.802580894E	4.8620752027E	3.1796654715E		
Indiana	Individual	PPD	1.4441280487E	2.7848187338E	7.1780025783E	4.1482474934E	2.9040050794E	0.3888002789E	8.79738811774E	2.7848187338E	7.1780025783E		
		Promer	4.3304164330E	7.0184888337E	5.2314880388E	3.9008807683E	3.8378807737E	6.7471942708E	1.4441280487E	7.0184888337E	5.2314880388E		
		OCN	0.291828918E	4.4878181005E	1.3082872870E	3.0111727516E	0.8046811609E	7.3278732884E	4.3304164330E	4.4878181005E	1.3082872870E		
Spouse	PPD	1.880807882E	2.041103428E	1.8128108807E	1.632729470E	4.868724888E	5.182992848E	0.818086158E	2.041103428E	1.8128108807E			
		Promer	3.474483884E	5.284818847E	2.588805948E	3.808703778E	5.888802582E	1.488788484E	1.880807882E	5.284818847E	2.588805948E		
		OCN	1.8711000304E	1.7368027201E	6.3004454200E	3.0148278025E	6.8908780400E	5.1438880330E	3.474483884E	1.7368027201E	6.3004454200E		
Dependent	PPD	0.8478837788E	0.8888808478E	3.8088808478E	3.8474818478E	4.838877788E	0.8888817788E	8.871888887E	0.8478837788E	3.8088808478E			
		Promer	7.892284891E	7.302284891E	3.847881891E	6.528887891E	2.194088089E	5.847880891E	0.8478837788E	7.302284891E	3.847881891E		
		OCN	4.3743743201E	7.082818900E	2.88138828E	1.897088034E	3.888848884E	5.888771847E	7.892284891E	7.082818900E	2.88138828E		

Added Labels

- 2021 (D3:F29)
- 2022 (G3:I29)
- 2023 (J3:L29)
- Michigan (D3:L11)
- Ohio (D12:L20)

Screen Mockup: Labeling

The image shows a screenshot of an Excel spreadsheet with columns for years 2021, 2022, and 2023, and rows for different states: Michigan, Ohio, and Indiana. A grey overlay box is positioned over the data, containing the text "Selected Cells: D21:L29" at the top, "Label Name" on the left, "Indiana" in a white box in the center, and an orange "Add Label" button at the bottom.

Added Labels

- 2021 (D3:F29)
- 2022 (G3:I29)
- 2023 (I3:L29)
- Michigan (D3:L11)
- Ohio (D12:L20)

Screen Mockup: Data Modeler

Roosevelt
simple. seamless. smart.

EDIT SAVE

Premier Rate ↓
Severity ↓
Age
Coverage
Insurance
Frequency ↓
Age
Area
Limits
Risks
Associates
IEE Rate
Pro SP Rate
PPO CH Rate
OON SP Rate
OON CH Rate

Legend
Label Dependency
Label From Sheet
Label
Worksheet

```
graph TD; Sheet1[Worksheet] --> PremierRate[Label]; PremierRate --> Severity[Label]; PremierRate --> Frequency[Label]; Severity --> Age1[Label]; Severity --> Coverage1[Label]; Severity --> Insurance1[Label]; Frequency --> Age2[Label]; Frequency --> Area[Label]; Frequency --> Limit[Label]; Frequency --> Associates[Label]; Frequency --> Risks[Label]; Age1 -.-> Location[Label]; Coverage1 -.-> Location; Insurance1 -.-> Location; Age2 -.-> Status[Label]; Area -.-> Status; Limit -.-> Status; Location --> Sheet2[Worksheet]; Status --> Sheet2; Sheet2 --> Health[Label]; Health --> MedicalHistory[Label]; Health --> Appointments[Label]; Health --> Genetic[Label];
```

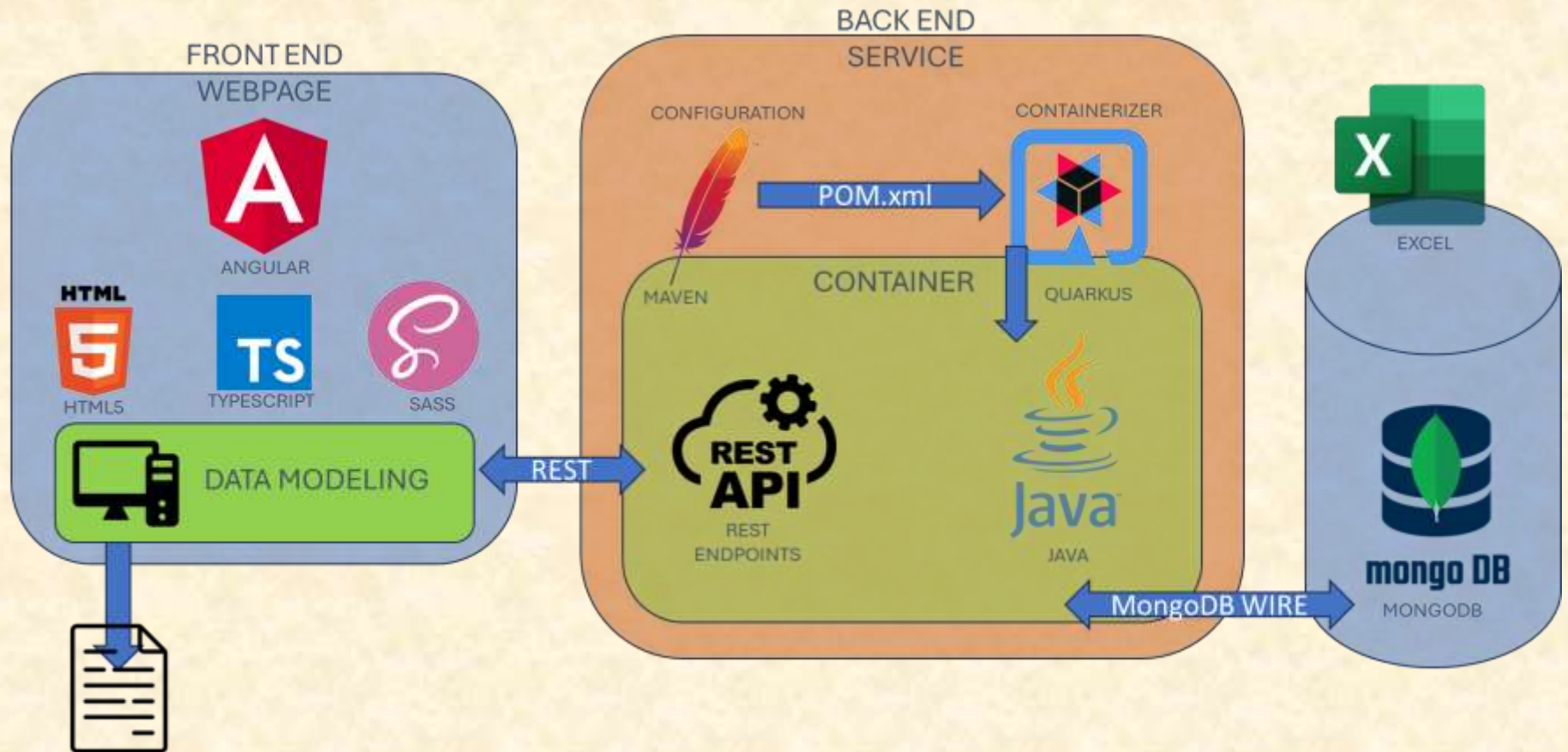

Project Technical Specifications

- Web application built using three separate entities.
- Frontend – Angular framework that...
 - Uploads/Downloads excel
 - Parses excel
 - Translates to GRACE
 - Models data
- Backend – Java in a Quarkus container that...
 - Runs computation and translation
 - Communicates with database
- Database – MongoDB cluster that...
 - Stores excel files, label and measurement information and data models.



Project System Architecture

MICROSOFT EXCEL DATA EXTRACTOR/MODELER



GRACE FILE
OUTPUT

Project System Components

- Hardware Platforms
 - Server running backend and database at Roosevelt Innovations Knowledge Science
 - Desktops viewing software through web app
- Software Platforms / Technologies
 - MongoDB
 - Angular
 - Java
 - Quarkus

Project Risks

- DSL Conversion
 - Project produces GRACE programs, an internal domain-specific language
 - Access to GRACE documentation and compiler to check validity, use of existing conversion code
- Database Server
 - We must ensure that the program is functional using only our own database server, as we will not have access to the client's database
 - Communicate with client to keep as many factors similar as possible between our internal testing server and their server
- Parsing Excel Workbooks
 - Data must be properly stored and handled from Excel workbooks, including those with multiple spreadsheets
 - Use provided backend code as a starting point, identify limitations
- Library for Displaying Excel Workbooks
 - Identify library for displaying workbooks that allows custom context actions without modifying library source code, if possible
 - X-data-spreadsheet was initially recommended, identify its limitations and possibly other libraries to work with it



Questions?

?

?

?

?

?

?

?

?

?

