

Project Plan Presentation

Microsoft Excel Data Extractor/Modeler

The Capstone Experience

Team Delta Dental Knowledge Science

Ethan Bransdorfer

Morgan Mundell

Peter Ro

Xochitl Weiss

Department of Computer Science and Engineering
Michigan State University

Fall 2021



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

Functional Specifications

- Make Excel sheets easy to understand
- Adds implied context to data
- Data tethered to labels not just location
- Ability to retrieve data at a different time without Excel
- Improve Efficiency of Delta Dental Employees
- Dynamic, functional for all data models
- Remove Reliance on code developers

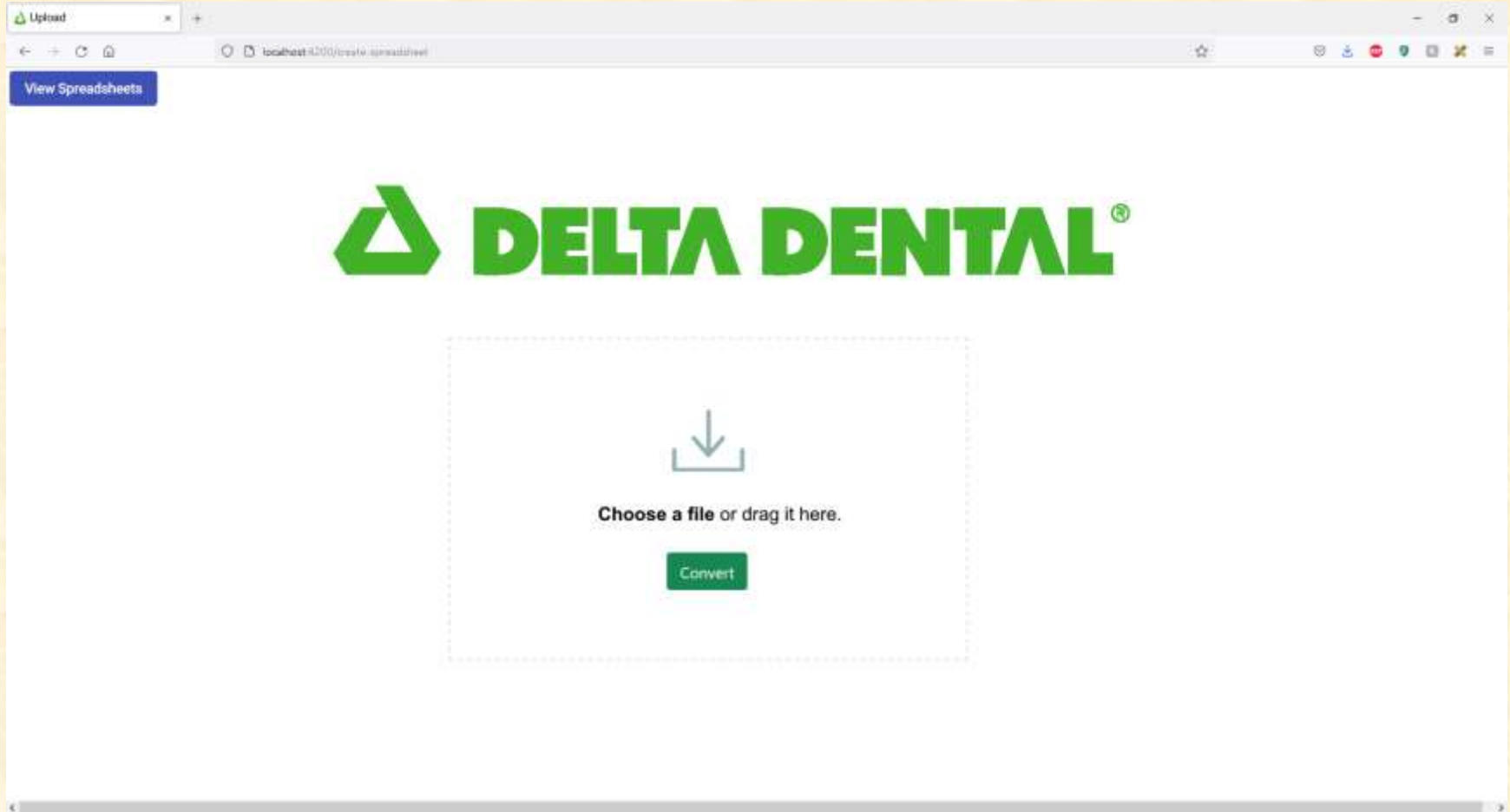


Design Specifications

- Intuitive web application
- Drag & Drop upload
- Previously uploaded spreadsheet viewer
- Browser based spreadsheet with labeling functionality
- Simple display to sort through labeled data



Screen Mockup: Upload



Screen Mockup: View

The screenshot shows a web browser window with the URL `localhost:4200/spreadsheets-list`. The page features a blue button labeled "Add Spreadsheets" and the Delta Dental logo. Below the logo is a table listing various spreadsheets.

Spreadsheet ID	Date	Update
Mathematics 101 Fall Semester 2021 Maccombs	2021-09-24	Edit Delete
Computer Science 231 Intro to Python Fall Semester 2021 Smith	2021-09-18	Edit Delete
Mathematics 101 Spring Semester 2020 Maccombs	2020-05-03	Edit Delete
Computer Science 232 Intro to C++ Spring Semester 2020 Johnson	2020-05-03	Edit Delete
Mathematics 101 Fall Semester 2020 Maccombs	2020-12-23	Edit Delete
Computer Science 231 Intro to Python Fall Semester 2020 Smith	2020-12-23	Edit Delete
Mathematics 101 Spring Semester 2019 Maccombs	2019-05-03	Edit Delete
Computer Science 232 Intro to C++ Spring Semester 2019 Johnson	2019-05-03	Edit Delete
Computer Science 231 Intro to Python Fall Semester 2021 Smith	2021-09-18	Edit Delete
Mathematics 101 Spring Semester 2020 Maccombs	2020-05-03	Edit Delete
Computer Science 232 Intro to C++ Spring Semester 2020 Johnson	2020-05-03	Edit Delete
Mathematics 101 Fall Semester 2020 Maccombs	2020-12-23	Edit Delete
Computer Science 231 Intro to Python Fall Semester 2020 Smith	2020-12-23	Edit Delete
Mathematics 101 Spring Semester 2019 Maccombs	2019-05-03	Edit Delete
Computer Science 232 Intro to C++ Spring Semester 2019 Johnson	2019-05-03	Edit Delete

At the bottom of the table, there is a pagination control showing page numbers 1, 2, 3, 4, with page 1 currently selected.



Screen Mockup: Label

The screenshot displays a spreadsheet application window titled 'Data Extractor'. The spreadsheet contains data for three students: Alice, Bob, and John. The data is organized into sections: 'Assignments' (HW1, HW2, HW3, MD, FINAL), 'Exams' (Q1, Q2, Q3, Q4, Q5), and 'Homework 1', 'Homework 2', and 'Homework 3'. A 'Labeling' dialog box is open, showing 'Cell Range' as 'C4:F7' and 'Label' as 'Final'. A 'Labels' panel on the right lists 'Final' as 'E4:E7'. The spreadsheet also includes columns for 'Major' (Engineering, Business, English, Undecided) and 'Grade' (A, B, C, D, E).

	Assignments					Exams							
	HW1	HW2	HW3	MD	FINAL	Final Grade	Major						
Alice	39	47	32	64	52	54.4	Engineering						
Bob	48	53	3	96	100	60.7	Business	Q1c	2	5			Multiple Choice
John	25	98	80	75	75	61.2	English	Q1	Q1b	3	6		
Max	12	114	8	8	6	5.4	Undecided	Q1c	1	1	1	1	

Labeling Dialog:

Cell Range: C4:F7
Label: Final

Labels Panel:

Label	Range
Names	A4:A7
Homework1	B4:B7
Homework2	C4:C7
Homework3	D4:D7
Midterm	E4:E7



Screen Mockup: Browse

The screenshot shows a web browser window with a tree diagram and two data tables. The tree diagram illustrates the structure of the data, with 'Assignments' and 'Questions' as root nodes. 'Assignments' branches into 'Names' and 'Homework', while 'Questions' branches into 'Points' and 'Answers'. 'Homework' further branches into 'Homework 1', 'Homework 2', and 'Homework 3'. 'Exams' branches into 'Midterm' and 'Final'. Below the tree, two tables are displayed: 'Homework 1' and 'Homework 2'. Each table lists the names of four students (Alice, Bob, John, Max) and their corresponding scores.

View Root

```
graph TD; Assignments --- Names; Assignments --- Homework; Assignments --- Exams; Questions --- Points; Questions --- Answers; Homework --- Homework1[Homework 1]; Homework --- Homework2[Homework 2]; Homework --- Homework3[Homework 3]; Exams --- Midterm; Exams --- Final;
```

Alice	39
Bob	40
John	99
Max	12

Alice	42
Bob	53
John	98
Max	114

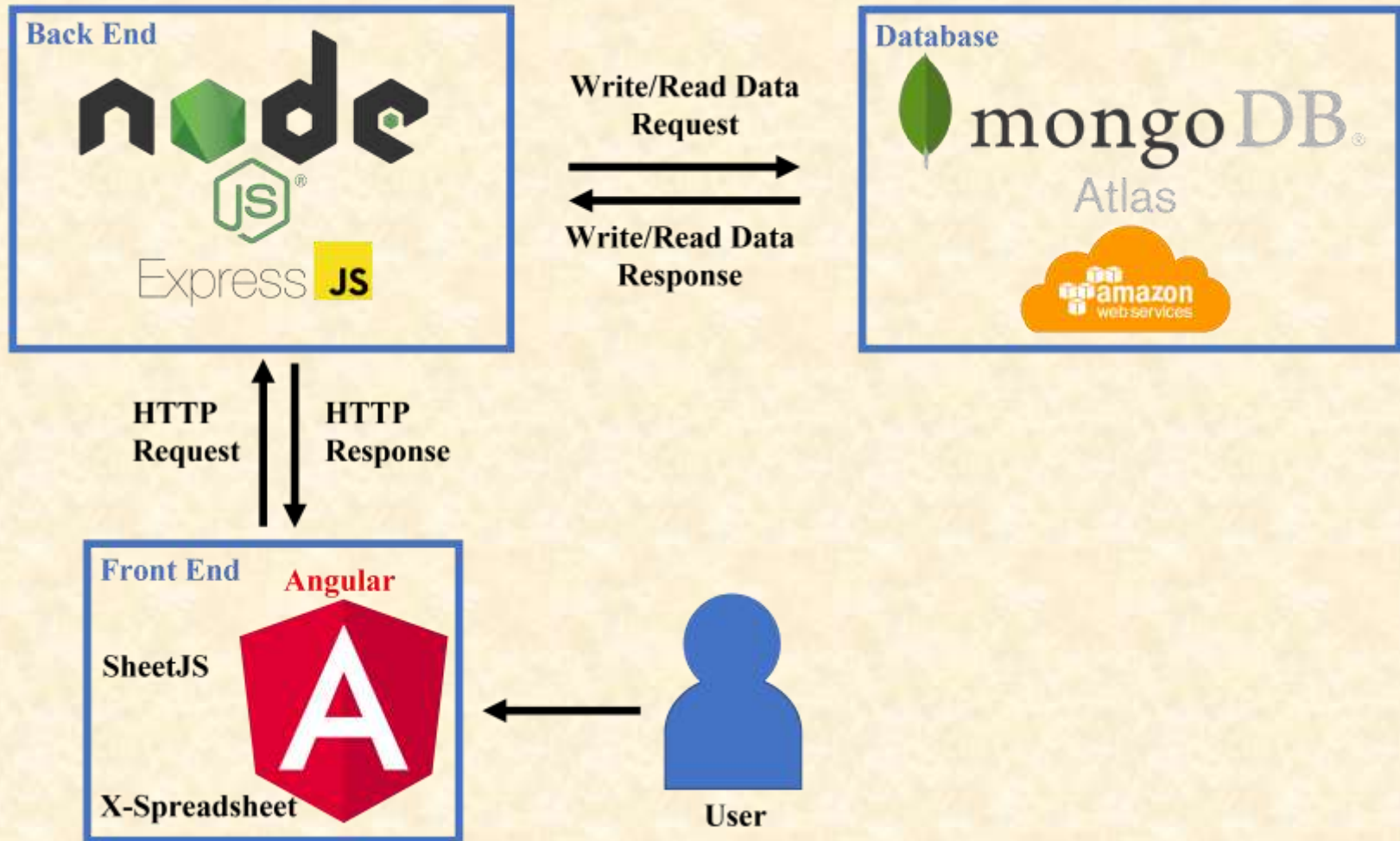


Technical Specifications

- JavaScript/TypeScript
- Microsoft Excel
- HTML/CSS
- NoSQL



System Architecture



System Components

- Software Platforms / Technologies
 - MongoDB
 - Express/Node.js
 - Angular
 - SheetJS
 - X-Spreadsheet



Risks

- UI Rendering Constraints - High
 - Rendering large datasets dire to frontend performance.
 - Virtual scrolling or progressive rendering.
- X-Spreadsheet Integration - Medium
 - Difficult integrating front-end package x-spreadsheet.
 - Investing production cycles to decide: resolve or pivot. Fallback to refactor.
- Loading Data
 - A key feature is the reproducibility of the input excel file from our backend.
 - Could save the excel spreadsheet in an open source json-based object for little additional cost.
- Multidimension Data Saved into JSON Form
 - No way to support turning an excel file into JSON unless it only has the first row as data labels and the rest as data. Need to support multi-dimensional and unique files.
 - Detailed labels and indications of dimension from user input will allow us to create a function to achieve this. We can look into SheetJS source code to aid in making this function.



Questions?

?

?

?

?

?

?

?

?

?

