IOWA STATE UNIVERSITY
College of Engineering
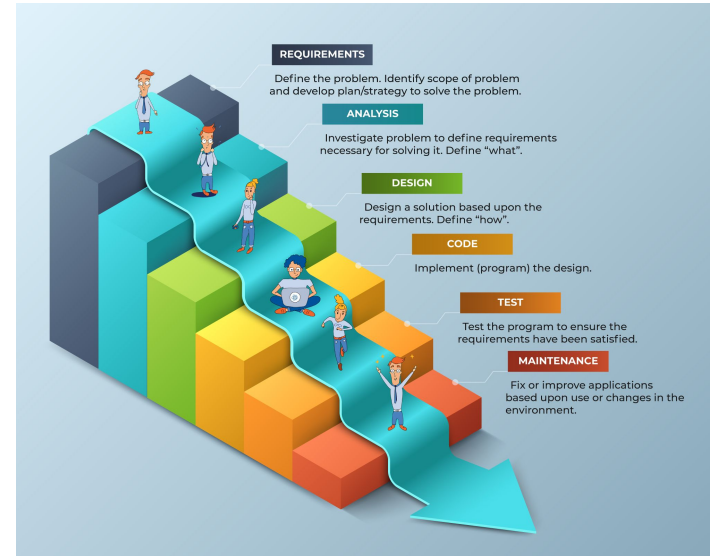
Text extractions from documents into Elasticsearch

Team Members: Bruce Bitwayiki, Jared Hayashi, Rushal Sohal, Tiffany Mayberry

Project Management Style

Task Decomposition

1. Project Planning
2. Develop Web Application User Interface (UI)
3. Develop REST API
4. Setup Elasticsearch Node
5. Develop File Text Extractor
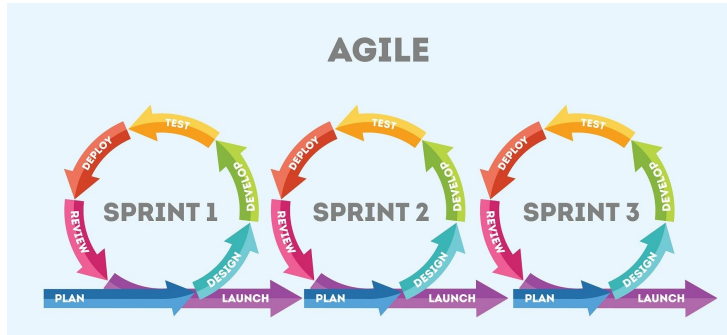6. Application Testing

Text Extraction from Documents into Elasticsearch for BuilderTrend

Tasks/Subtasks	01/25/22	1/31/22	2/7/22	2/14/22	2/21/22	2/28/22	3/7/22	3/14/22	3/21/22	3/28/22	4/4/22	4/11/22	4/18/22	4/25/22	5/2/22	5/9/22	Summer Break	8/22/22	8/29/22	9/5/22	9/12/22	9/19/22	9/26/22	10/3/22	10/10/22	10/17/22	10/24/22	10/31/22	11/7/22	11/14/22	11/21/22	11/28/22	12/5/22	12/12/22	Completed

1. In-Class Assignments — Spring Break / Finals Week / Summer Break / Fall Break / Finals Week
- Team Building
- Team Initiation Assignment
- Professionalism Assignment
- Team Website
- Requirements & Standards Planning
- Design Planning
- Testing Planning
- Project Planning

2 Develop Web Application UI
- Design core components
- Create UI Mockup
- Implement Prototype
- Implement Search Bar and Results View
- Implement Filter Bar
- Implement File Uploader
- Integrate REST API Calls
- Revise Implementation

3. Develop REST API
- Design Calls between Web App and Elasticsearch
- Design Calls between Text Extractor and Elasticsearch
- Implement Web App API Calls
- Implement Text Extractor API Calls

4. Setup Elasticsearch
- Request server to host
- Install / Setup Elasticsearch on server
- Setup Elasticsearch Node

IOWA STATE UNIVERSITY   College of Engineering

Text Extraction from Documents into Elasticsearch for BuilderTrend

Tasks/Subtasks	01/25/22	1/31/22	2/7/22	2/14/22	2/21/22	2/28/22	3/7/22	3/14/22	3/21/22	3/28/22	4/4/22	4/11/22	4/18/22	4/25/22	5/2/22	5/9/22	Summer Break	8/22/22	8/29/22	9/5/22	9/12/22	9/19/22	9/26/22	10/3/22	10/10/22	10/17/22	10/24/22	10/31/22	11/7/22	11/14/22	11/21/22	11/28/22	12/5/22	12/12/22	Completed
5. Develop File Text Extractor								Spring Break								Finals Week	Summer Break														Fall Break			Finals Week	
Install Apache Tika on server																																			
Design component interactions																																			
Implement file type detection																																			
Implement file meta data detection																																			
Implement .docx text extraction																																			
Implement .txt text extraction																																			
Implement .xlsx text extraction																																			
Implement .pptx text extraction																																			
Implement .pdf text extraction																																			
Implement .jpg text extraction																																			
Implement .png text extraction																																			
Integrate REST API Calls																																			
Revise Implementation																																			
6. Application Testing								Break								Finals Week	Summer Break														Break			Finals Week	
Create frontend unit tests																																			
Create text extractor unit tests																																			
User acceptance tests																																			

IOWA STATE UNIVERSITY

College of Engineering

Risks/Risk Mitigation

1. Project Planning (Risk Probability: 0.3)
2. Develop Web Application User Interface (Risk Probability: 0.3)
3. Develop REST API (Risk Probability: 0.4)
4. Setup Elasticsearch Node (Risk Probability: 0.2)
5. Develop File Text Extractor (Risk Probability: 0.4)
6. Application Testing (Risk Probability: 0.1)

Personal Effort Level

1. Project Planning (Est. Hours: 75)
2. Develop Web Application User Interface (Est. Hours: 50)
3. Develop REST API (Est. Hours: 100)
4. Setup Elasticsearch Node (Est. Hours: 75)
5. Develop File Text Extractor (Est. Hours: 75)
6. Application Testing (Est. Hours: 25)