Data Classifier:
Software Architecture
version 1.0

Team Flash
Computer Science Department
California Polytechnic State University
San Luis Obispo, CA USA

November 7, 2018

Contents

Credits

Name	Date	Role	Version
Joey Buelow	November 7, 2018	Document Owner	1.0
			1.0

Revision History

Name	Date	Reason for Changes	Version
Joey Buelow	November 7, 2018	Initial baseline	1.0
Kent Tran	November 7, 2018	Problem Description	1.0
Bonita Galvan	November 7, 2018	Editing and Image Formatting	1.0

1 Introduction - Joey

This document provides a detailed description of the system architecture for the Data Classifier component of the Mark Logic Data Discovery Workbench. Contained in this document is an outline for the development of the aforementioned system. This includes an overall description of the system, including diagrams, definitions for the components of the system, design considerations, testing plan, and issues.

2 Problem Description

The system is a data classifier that uses machine learning to categorize input data and outputs the data formatted into the appropriate categories. Additional features of the system include editing the generated categories, marking specific data fields as sensitive, and a web-based user interface.
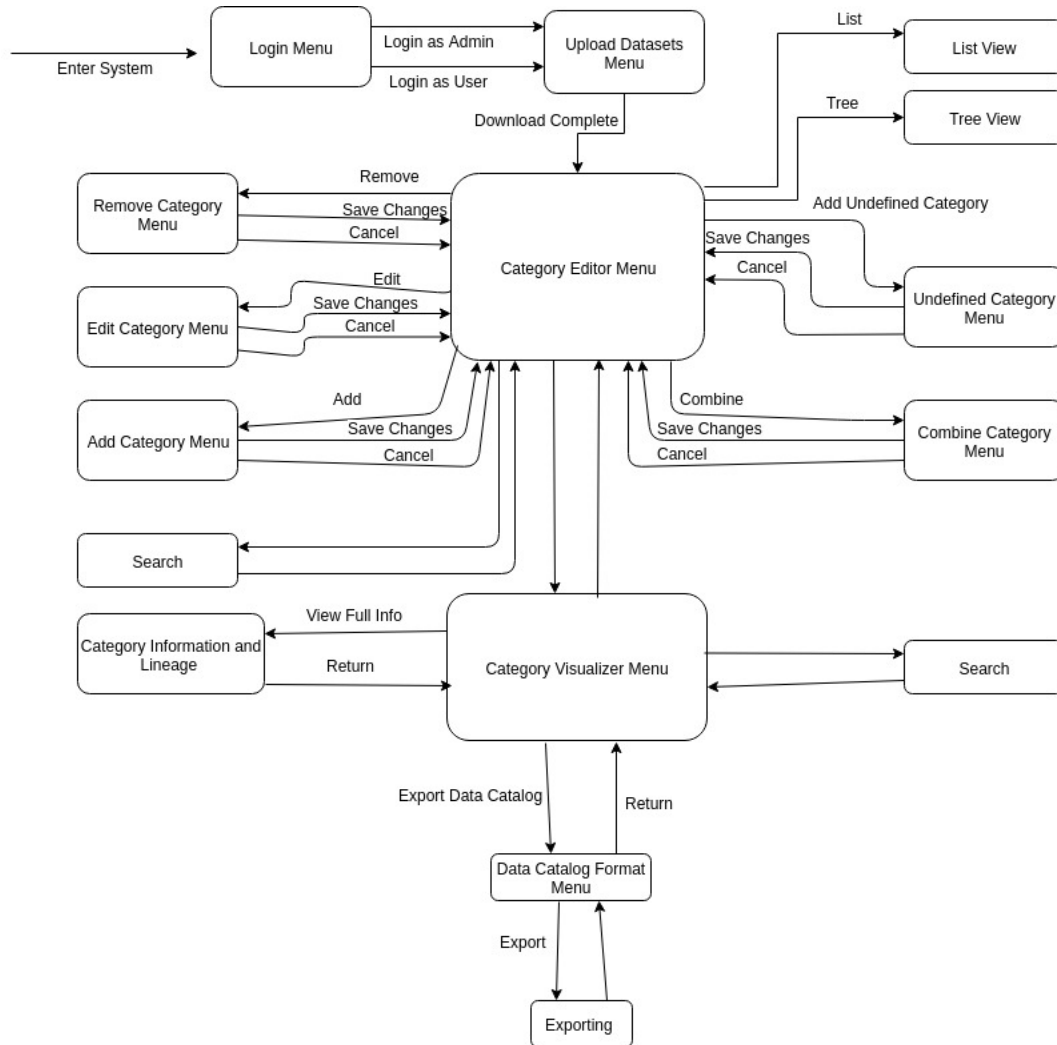
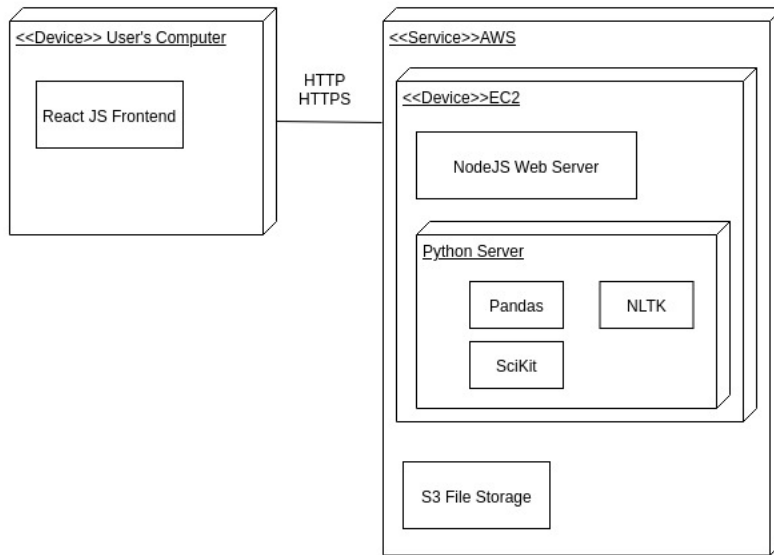# 3 Solution

TBD

3.1 Overview

TBD

3.2 Components

TBD

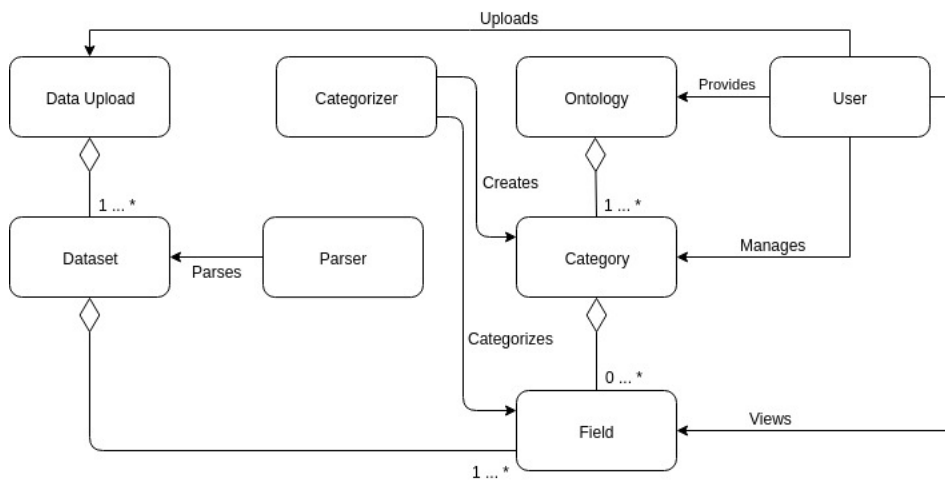## 3.3 Design

Bonita Dialog Map

Diana Deployment Diagram

Deployment

Diana Class Diagram

Concepts

Joey Architecture Diagram
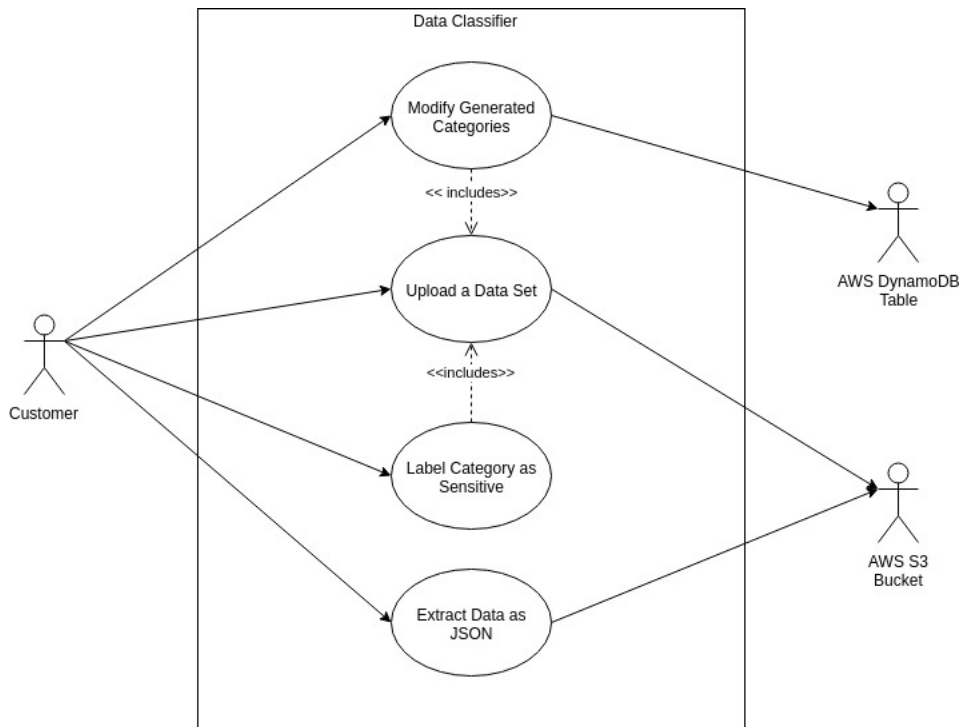
Workflow

Data ⟶ Parse ⟶ Categorize ⟶ Visualize

Ontology ⟵

GUI

Visualizer

draws from

Graphs / File System

Observes

expands

Parser

feeds

Categorizer

informs

Ontology

contains

Category

contains

Sets

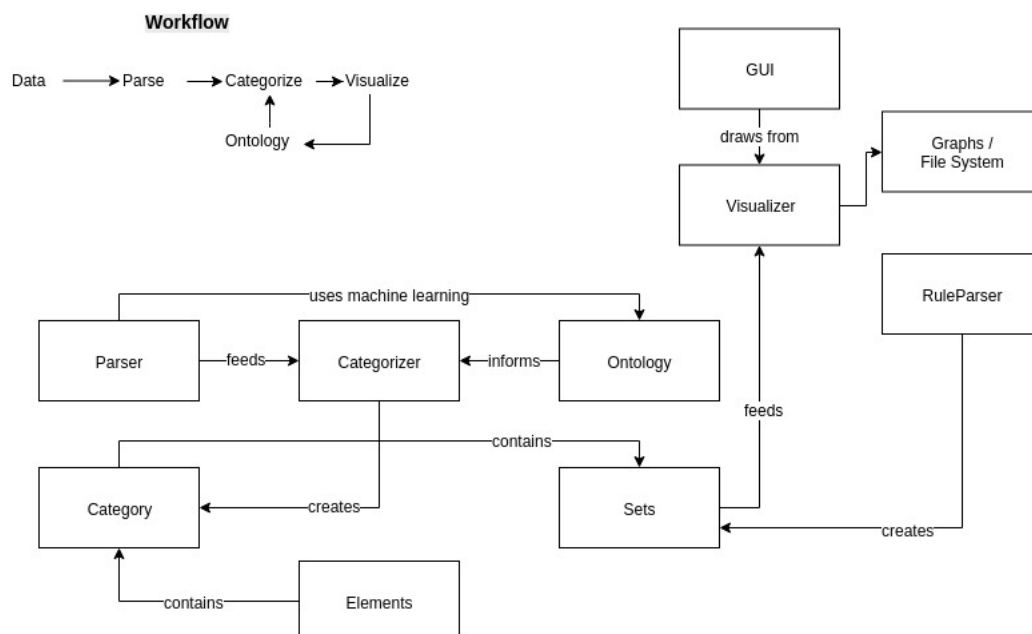contains

Elements

contains

Rule

creates
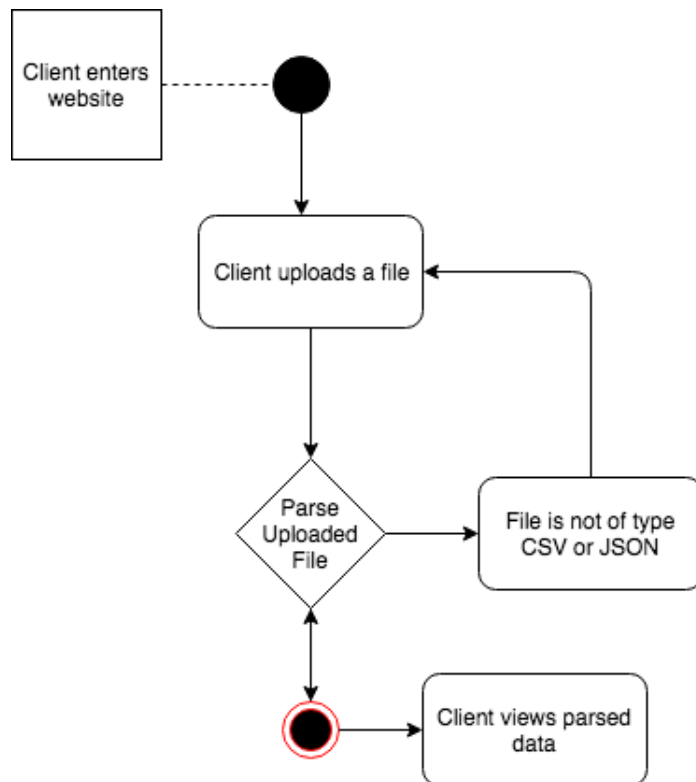
RuleParser
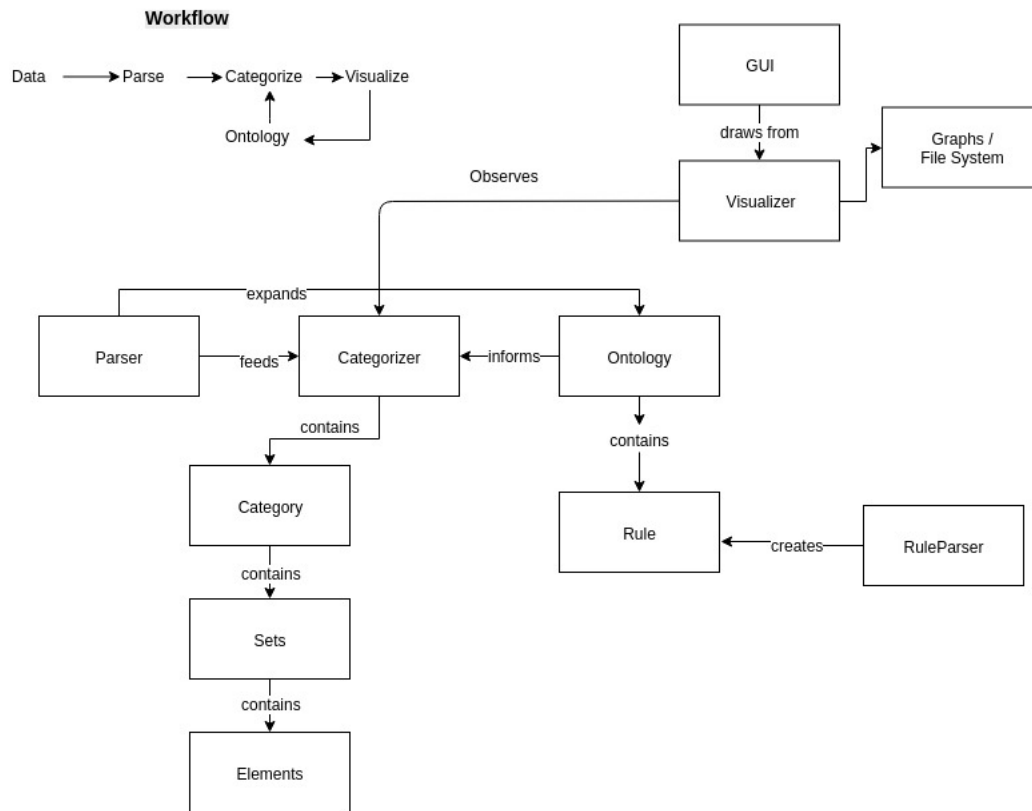
Joey Activity Diagram
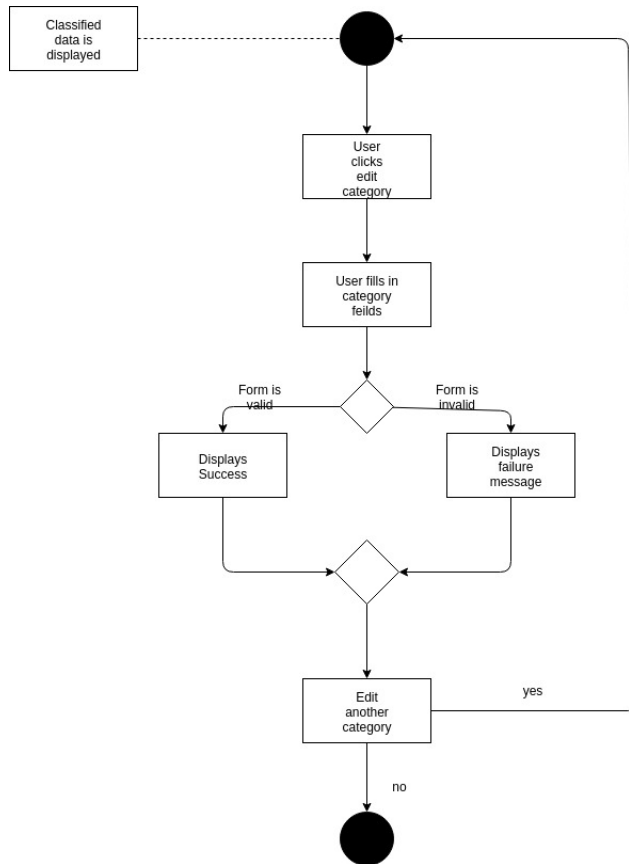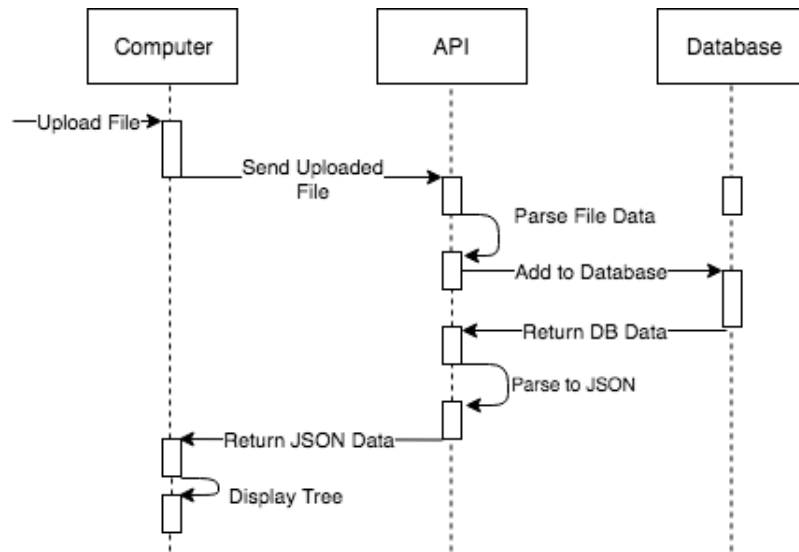
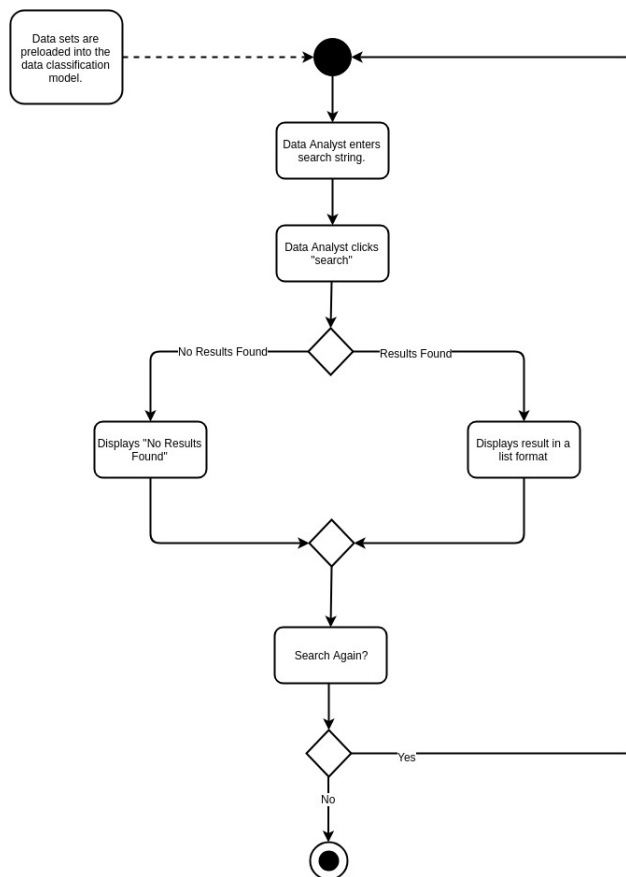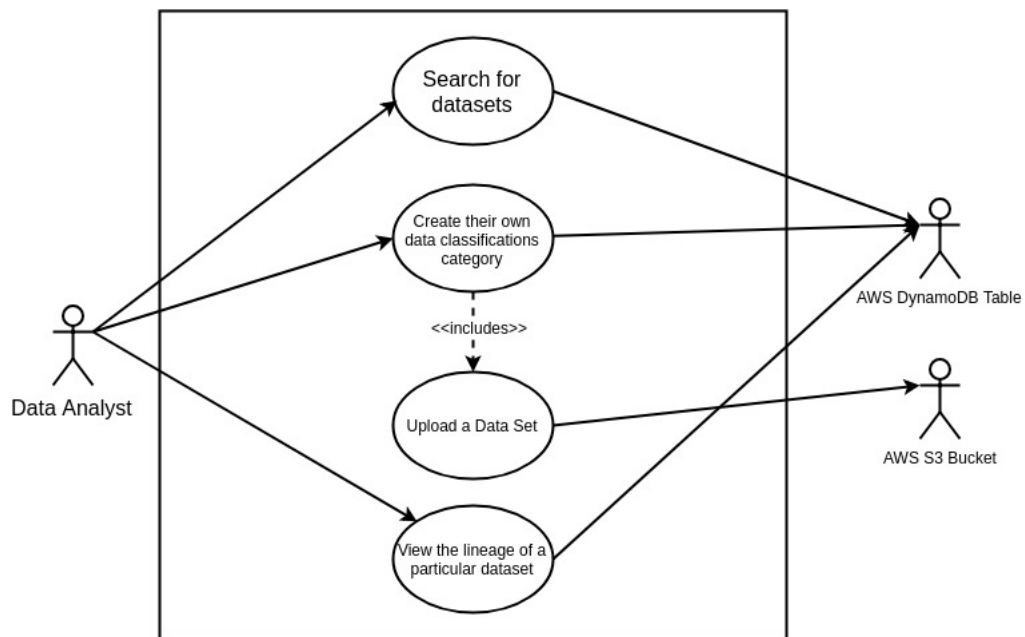**Kent** Use Case Diagram

Kent Architecture Diagram

Tanner Activity Diagram

Tanner Sequence Diagram

Victoria Activity Diagram

Victoria Use Case Diagram

4   Test

TBD

5   Issues

TBD

A   Glossary

TBD

B   Issues List

TBD