

MICHIGAN STATE
UNIVERSITY

Beta Presentation

Security Analytics Suite: Dataset Merger Tool The Capstone Experience

Team Avata

Jonny Dowdall

Paige Henderson

Matt Scheffler

Aasir Walajahi

Zac Wellmer

Department of Computer Science and Engineering

Michigan State University

Fall 2016



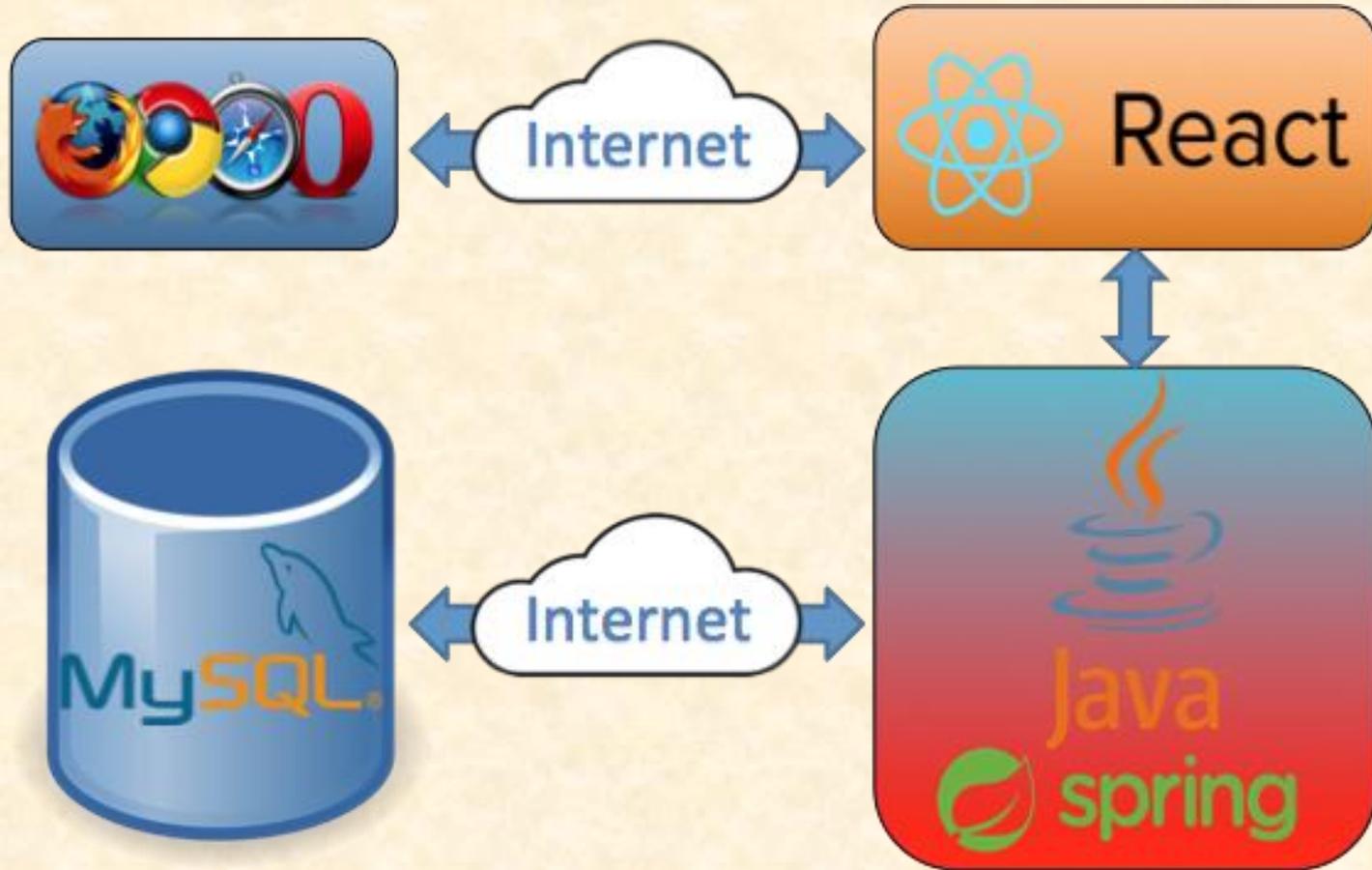
*From Students...
...to Professionals*

Project Overview

- Automatically identify and merge duplicate records within and across datasets.
- Allow the user to approve or disapprove the merging of similar records.
- Present the results as a report containing tables and charts.



System Architecture



Run Page

The screenshot shows a web browser window with the title "Dataset Merger" and the URL "avataai.com/dataset-merger". The page features the Avata Intelligence logo at the top left. Below the logo, there are three main sections for configuration:

- Select Dates:** Includes "Start Date" (01/01/2000) and "End Date" (11/13/2016) input fields.
- Select Times:** Includes "Start Time" (12:00 AM) and "End Time" (11:00 PM) input fields.
- Select Data Sources:** Includes "Restaurants" (checked) and "Incidents" (unchecked) options.

A teal "Run Merge" button is located at the bottom left of the form area.



Run Page

The screenshot shows a web browser window titled "Dataset Merger" with the URL "avataai.com/dataset-merger". The page features the Avata Intelligence logo at the top. Below the logo, there are three main sections for configuration:

- Select Dates:** Includes "Start Date" (01/01/2000) and "End Date" (11/13/2016) input fields.
- Select Times:** Includes "Start Time" (12:00 AM) and "End Time" (11:00 PM) input fields.
- Select Data Sources:** Includes "Restaurants" (checked) and "Incidents" (unchecked) checkboxes.

At the bottom left, there is a teal "Merging" button and a loading spinner icon.



Merged Records Page

The screenshot shows a web browser window with the URL `avataai.com/merged-records`. The page features the Avata Intelligence logo and three navigation buttons: "Merged Records", "Similar Records", and "Analysis". The "Merged Records" button is active. Below the navigation, there are three distinct sections, each containing a table of merged records and a "Split Records" button.

Section 1: Bel-Air Hotel

Record ID	Restaurant_name	Address	Phone_number	Category	Source	City
5	Bel-Air Hotel	701 Stone Canyon Rd.	310-472-1211	Californian	zagats	Bel Air
366	Hotel Bel-Air	701 Stone Canyon Rd.	310/472-1211	Californian	fodors	Bel Air

Section 2: Citrus

Record ID	Restaurant_name	Address	Phone_number	Category	Source	City
21	Citrus	6703 Melrose Ave.	213-857-0034	Californian	zagats	Los Angeles
349	Citrus	6703 Melrose Ave.	213/857-0034	Californian	fodors	Los Angeles

Section 3: Granita

Record ID	Restaurant_name	Address	Phone_number	Category	Source	City
28	Granita	23725 W. Malibu Rd.	310-456-0488	Californian	zagats	Malibu
362	Granita	23725 W. Malibu Rd.	310/456-0488	Californian	fodors	Malibu



Similar Records Page

The screenshot shows a web browser window titled "Dataset Merger" with the URL "avataai.com/similar-records". The page features the Avata Intelligence logo and three navigation buttons: "Merged Records", "Similar Records", and "Analysis". The "Similar Records" button is active. Below the navigation, there are three tables, each representing a different restaurant. Each table has columns for Record ID, Restaurant_name, Address, Phone_number, Category, Source, and City. Below each table are two buttons: "Do Not Merge" and "Merge Records".

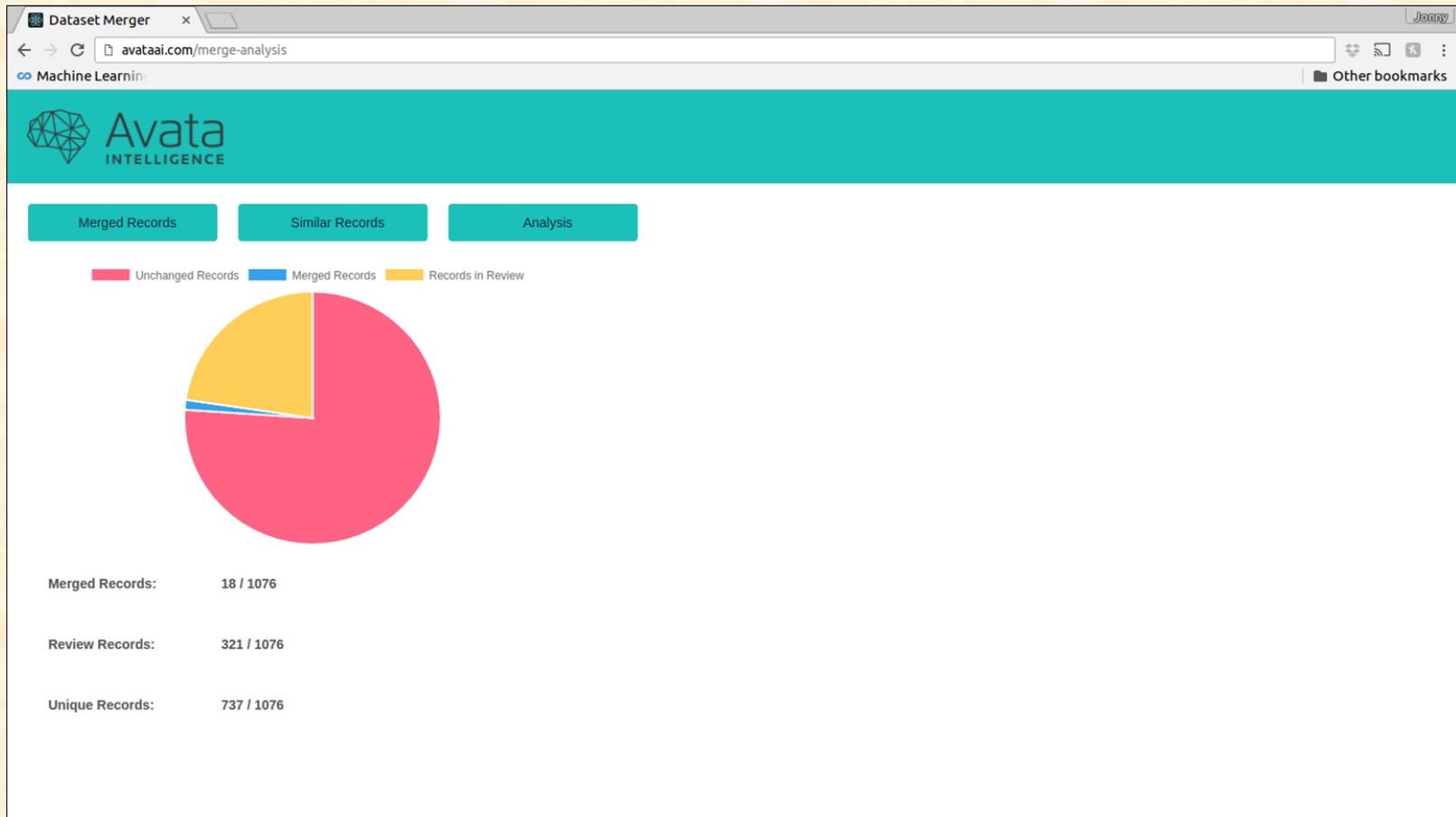
Record ID	Restaurant_name	Address	Phone_number	Category	Source	City
1	Arnie Morton's of Chicago	435 S. La Cienega Blvd.	310-246-1501	Steakhouses	zagats	Los Angeles
331	Arnie Morton's of Chicago	435 S. La Cienega Blvd.	310/246-1501	American	fodors	Los Angeles

Record ID	Restaurant_name	Address	Phone_number	Category	Source	City
2	Art's Deli	12224 Ventura Blvd.	818-762-1221	Delis	zagats	Studio City
332	Art's Delicatessen	12224 Ventura Blvd.	818/762-1221	American	fodors	Studio City

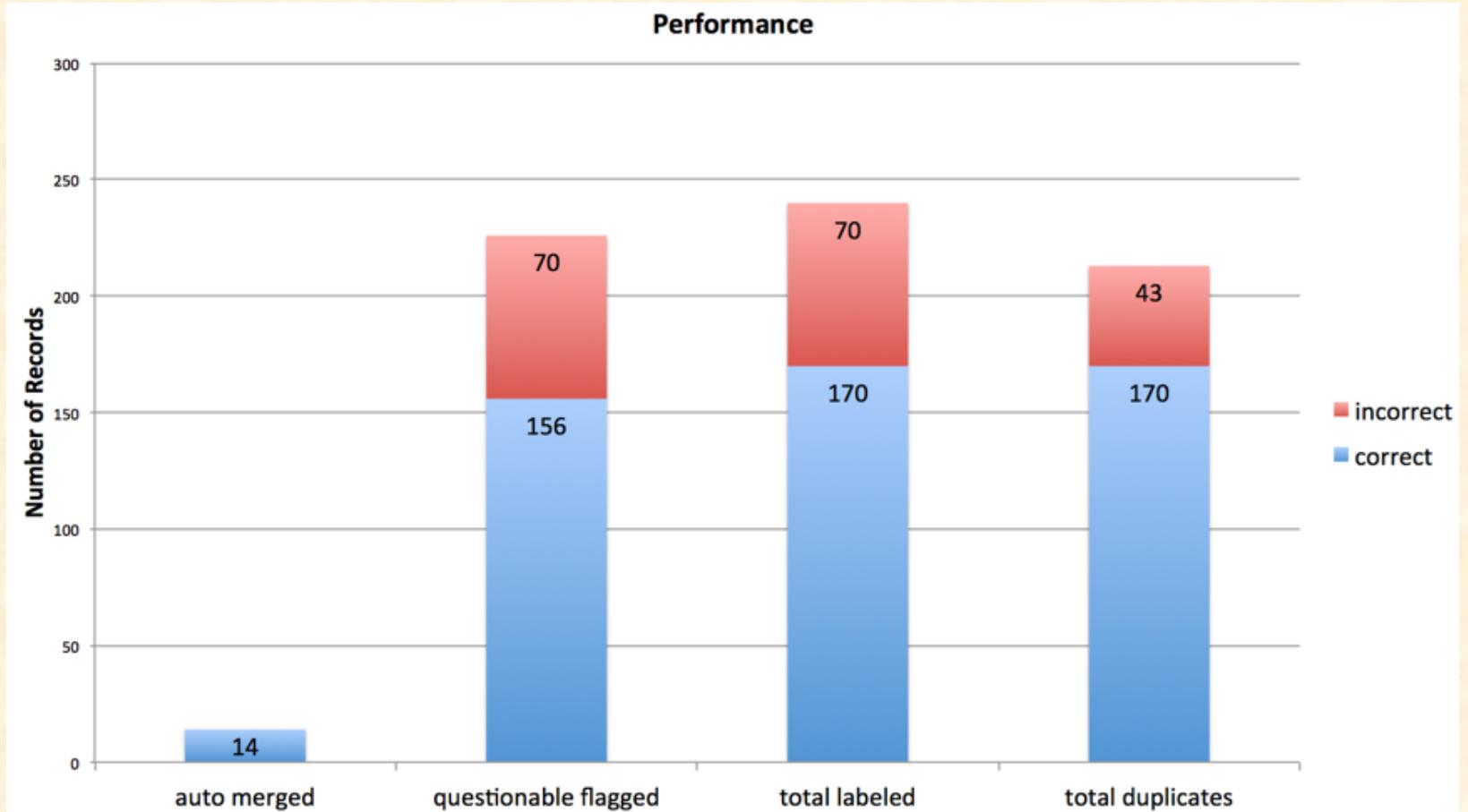
Record ID	Restaurant_name	Address	Phone_number	Category	Source	City
6	Belvedere The	9882 Little Santa Monica Blvd.	310-788-2306	Pacific New Wave	zagats	Beverly Hills
16	Cafe Blanc	9777 Little Santa Monica Blvd.	310-888-0108	Pacific New Wave	zagats	Beverly Hills



Analysis Page



Results



What's left to do?

- Improve SQL performance and security
- Improve parameterized query UX
- Styling fixes
- Add more statistics for analysis
- Finish testing
- Minor bug fixes



Questions?

?

?

?

?

?

?

?

?

?

