

MICHIGAN STATE

UNIVERSITY

Alpha Presentation

Security Analytics Suite: Dataset Merger Tool

The Capstone Experience

Team Avata

Jonny Dowdall

Paige Henderson

Matt Scheffler

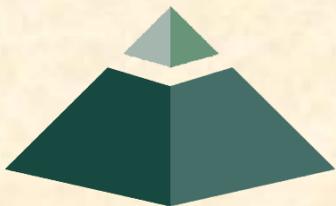
Aasir Walajahi

Zac Wellmer

Department of Computer Science and Engineering

Michigan State University

Fall 2016



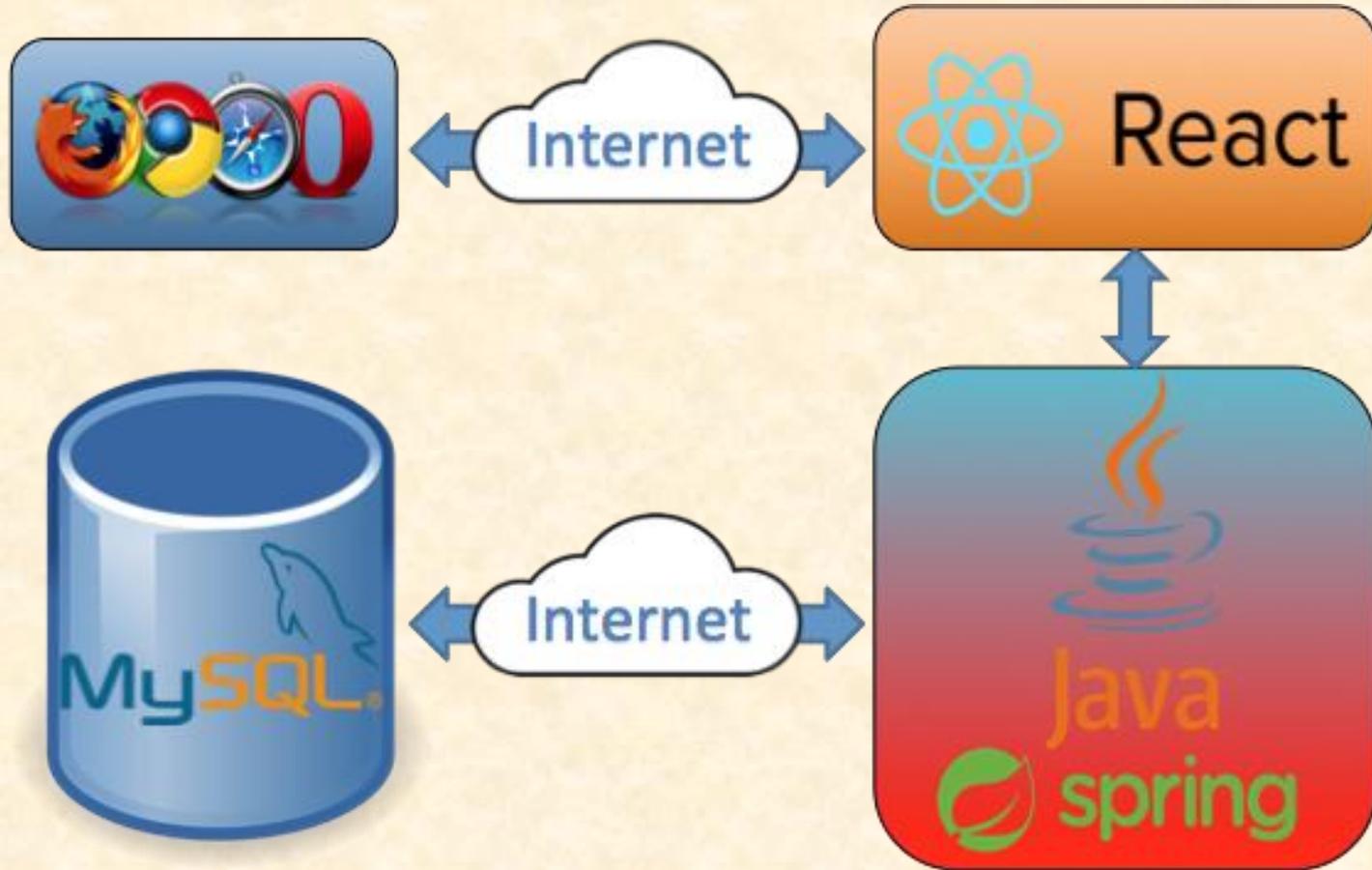
*From Students...
...to Professionals*

Project Overview

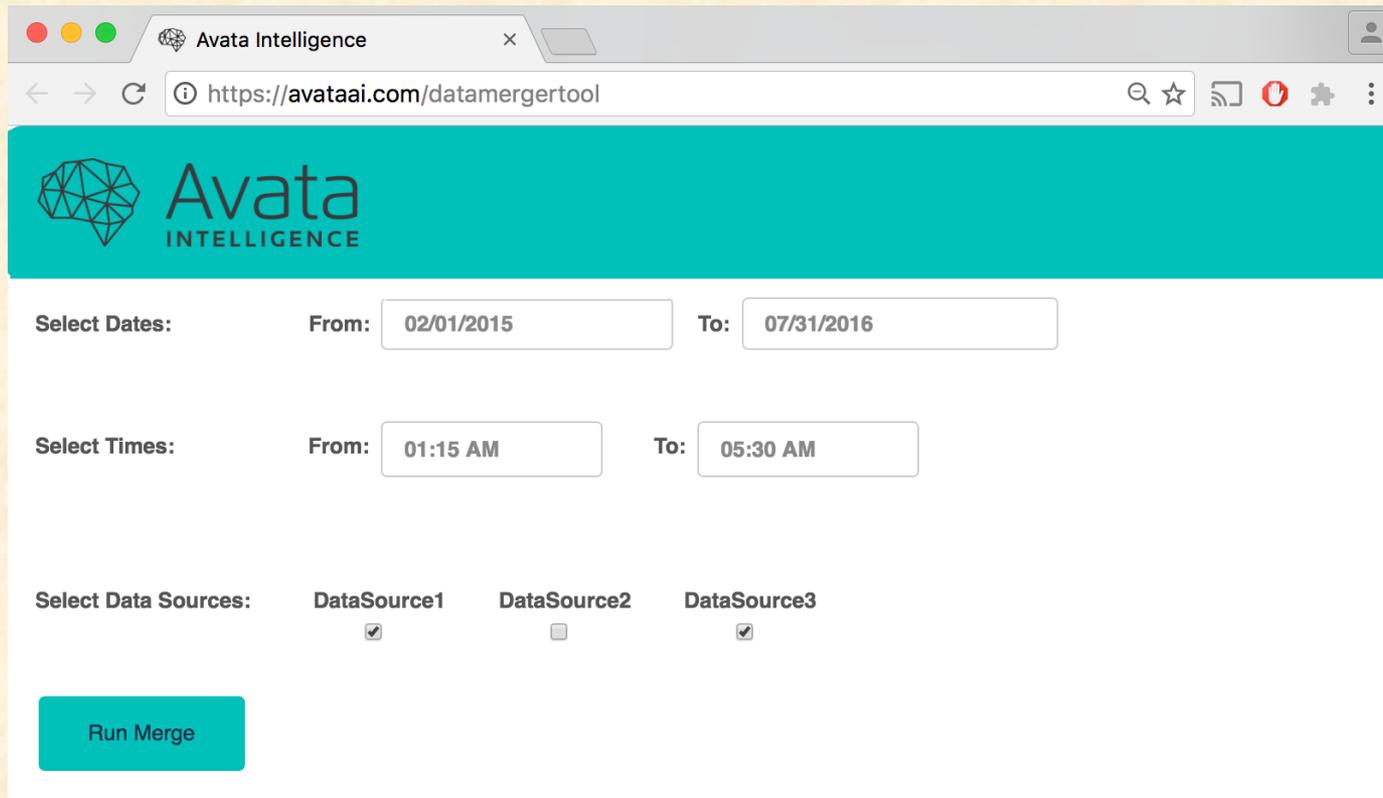
- Automatically identify and merge duplicate records within and across datasets.
- Allow the user to to approve or disapprove the merging of similar records.
- Present the results as a report containing tables and charts.



System Architecture



Landing Page



The screenshot shows a web browser window with the following elements:

- Browser Tab:** Avata Intelligence
- Address Bar:** <https://avataai.com/datamergertool>
- Header:** Avata INTELLIGENCE logo on a teal background.
- Select Dates:** From: To:
- Select Times:** From: To:
- Select Data Sources:**
 - DataSource1
 - DataSource2
 - DataSource3
- Action:**

Merge Progress

The screenshot shows a web browser window with the URL `https://avataai.com/datamergertool`. The page header features the Avata Intelligence logo. The main interface includes several configuration sections:

- Select Dates:** From: 02/01/2015, To: 07/31/2016
- Select Times:** From: 01:15 AM, To: 05:30 AM
- Select Data Sources:** DataSource1 (checked), DataSource2 (unchecked), DataSource3 (checked)

A teal **Run Merge** button is positioned below the configuration options. The **Progress:** section shows a light blue progress bar that is approximately 5% full. Below the progress bar, the status is summarized as follows:

Category	Count	Action
Processed	200 / 5000	
Merged	10 / 5000	Review Merges
Review	20 / 5000	Review Records

Merged Records

The screenshot shows a web browser window with the URL <https://avataai.com/datamergertool>. The page features the Avata Intelligence logo and three main navigation buttons: "Merged Records", "Review Records", and "Analysis". Below these buttons, there are two tables of merged records, each with an "UnMerge Records" button.

Table 1:

Records	Name	Gpa	Address	Date
Record 1	John Doe	3.5	1234 Wilson Road	11-4-1990
Record 2	John Doe	NULL	1234 Wilson Rd.	11/4/1990
Record 3	J. Doe	3.5	NULL	NULL

Table 2:

Records	Name	Gpa	Address	Date
Record 4	Dave Smith	2.9	456 Farm Ln.	4-3-1995
Record 5	Dave S.	2.9	456 Farm Lane	4-3-1995



What's left to do?

- Create Review Records Page
 - Allow user to view confidence score
- Create Analysis Page
 - Visually represent data as graph or table
- Fine tune merge algorithm
- Parameterized queries from the front-end
- Dynamically update



Questions?

?

?

?

?

?

?

?

?

?

