MICHIGAN STATE UNIVERSITY

Alpha Presentation
Open Source Intel

The Capstone Experience

Team GM

Ben Buscarino
Will Crecelius
Taylor Zachar
Qiming Ren
Igli Ndoj

Department of Computer Science and Engineering
Michigan State University

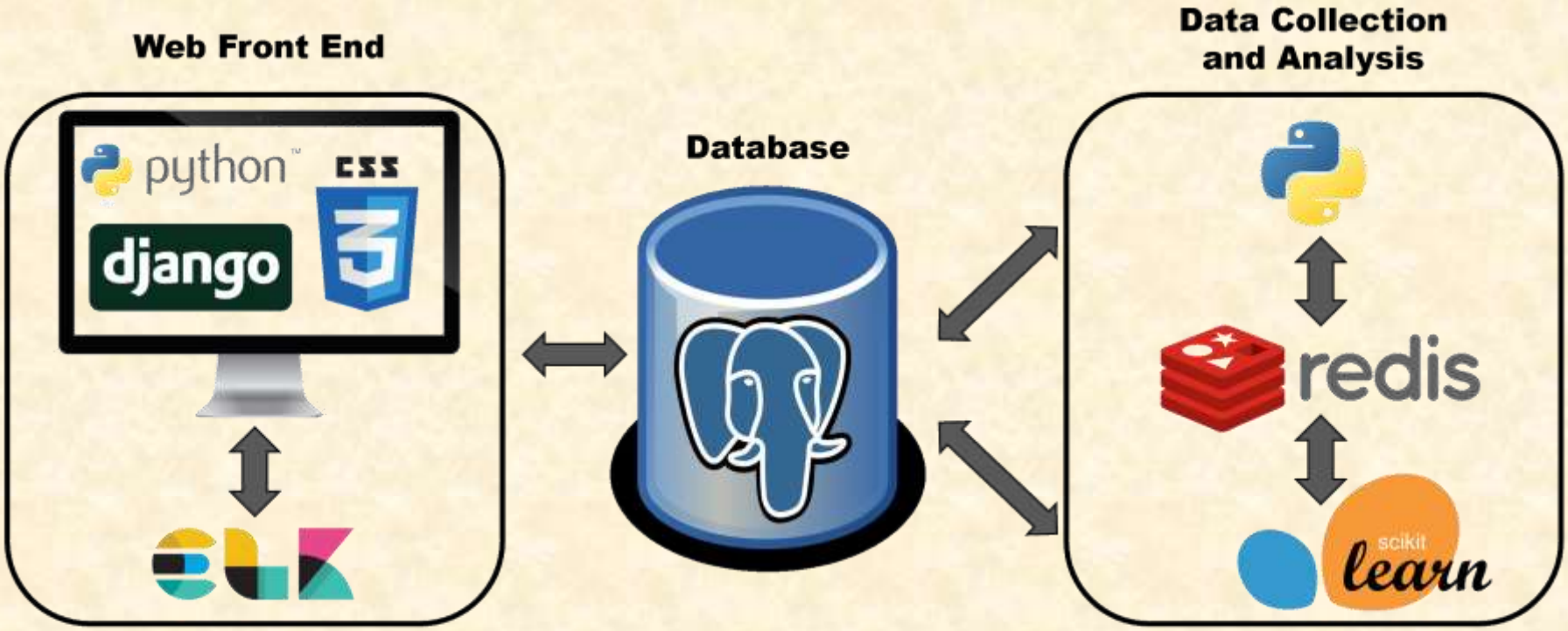
Spring 2020

From Students…
…to Professionals

Project Overview

- Scan public repositories such as GitHub, PasteBin, BitBucket for leaked GM intellectual property.
- Assign a confidence rating to a leak with machine learning.
- Display the leaks and a URL to the leak in a web application frontend.

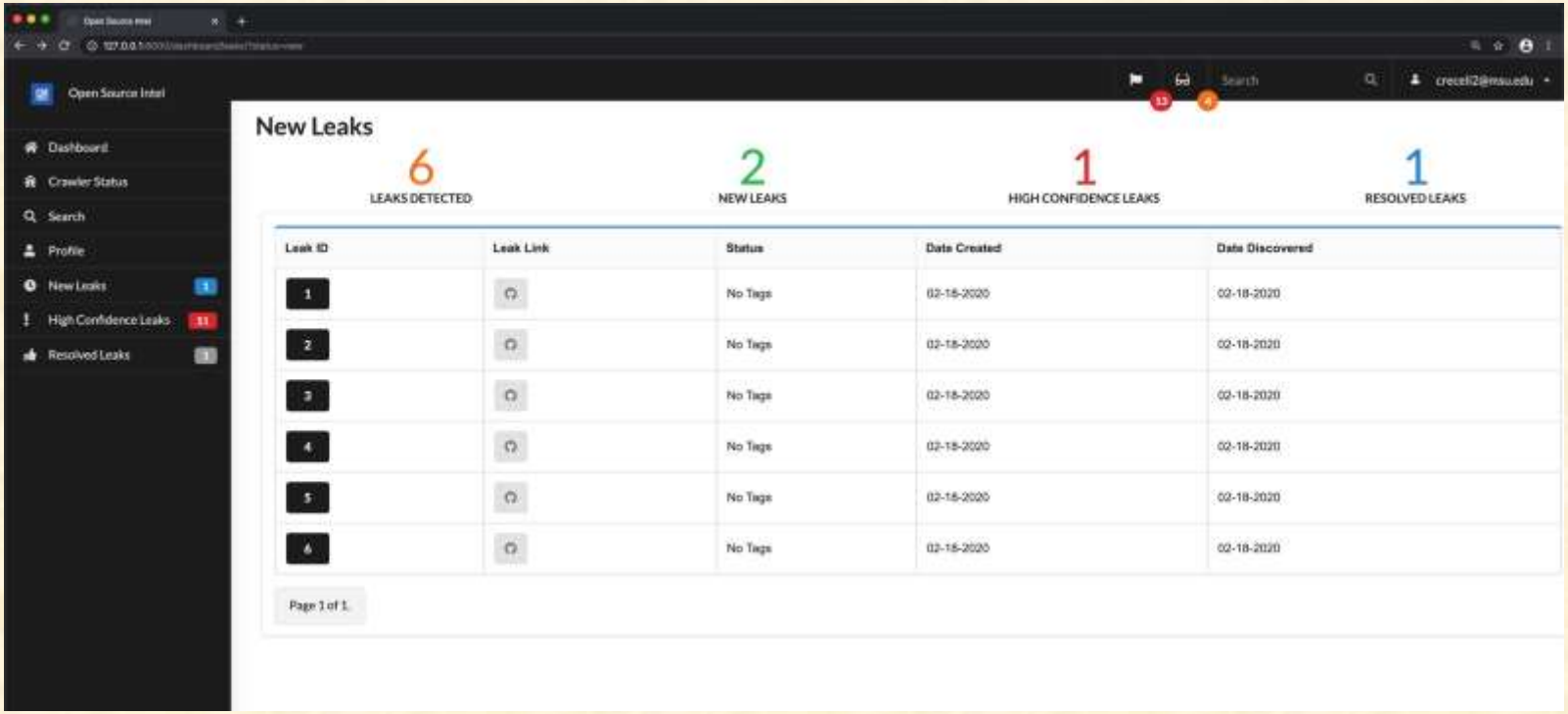
System Architecture



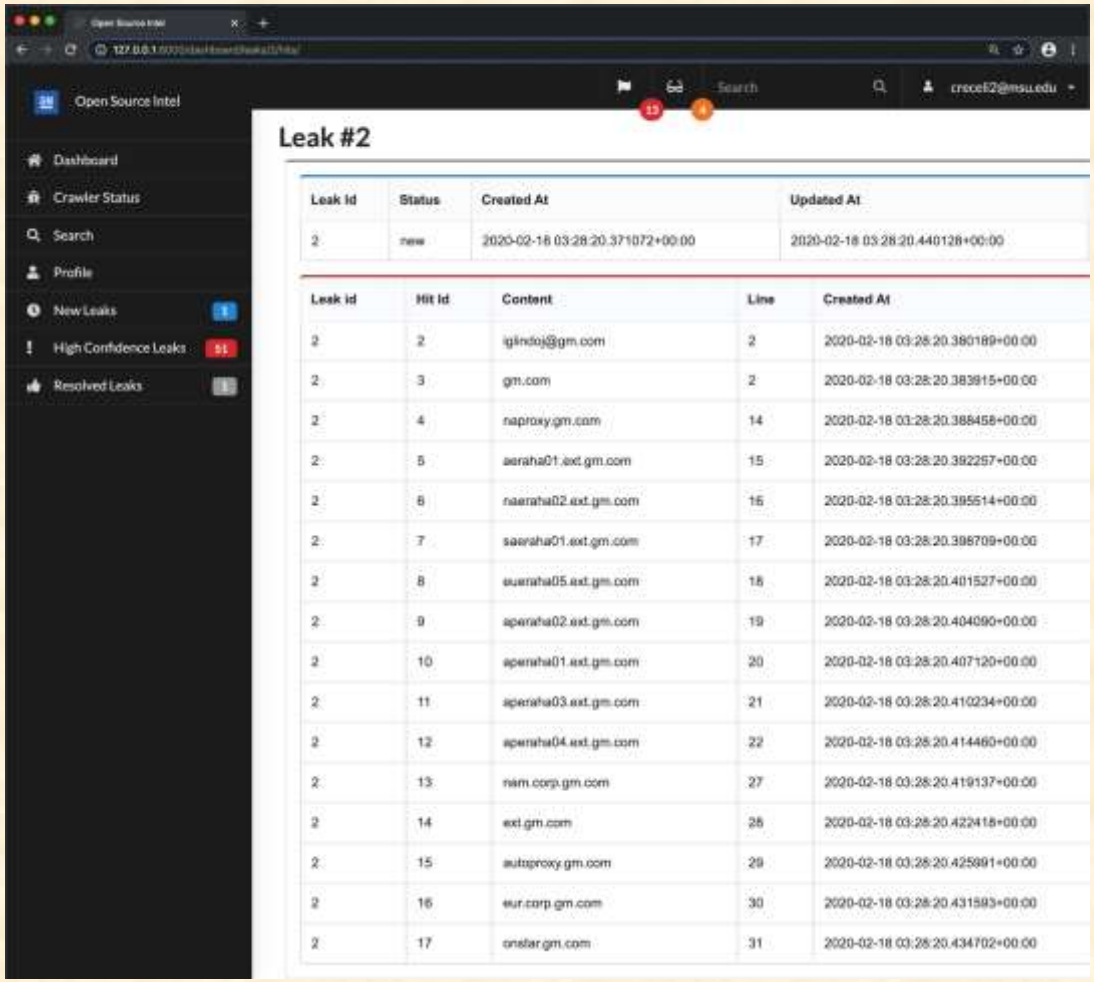
Dashboard


Dashboard
22
2.3 MILLION
1231
123
Leak Breakdown by Type
Leaks Discovered over Time

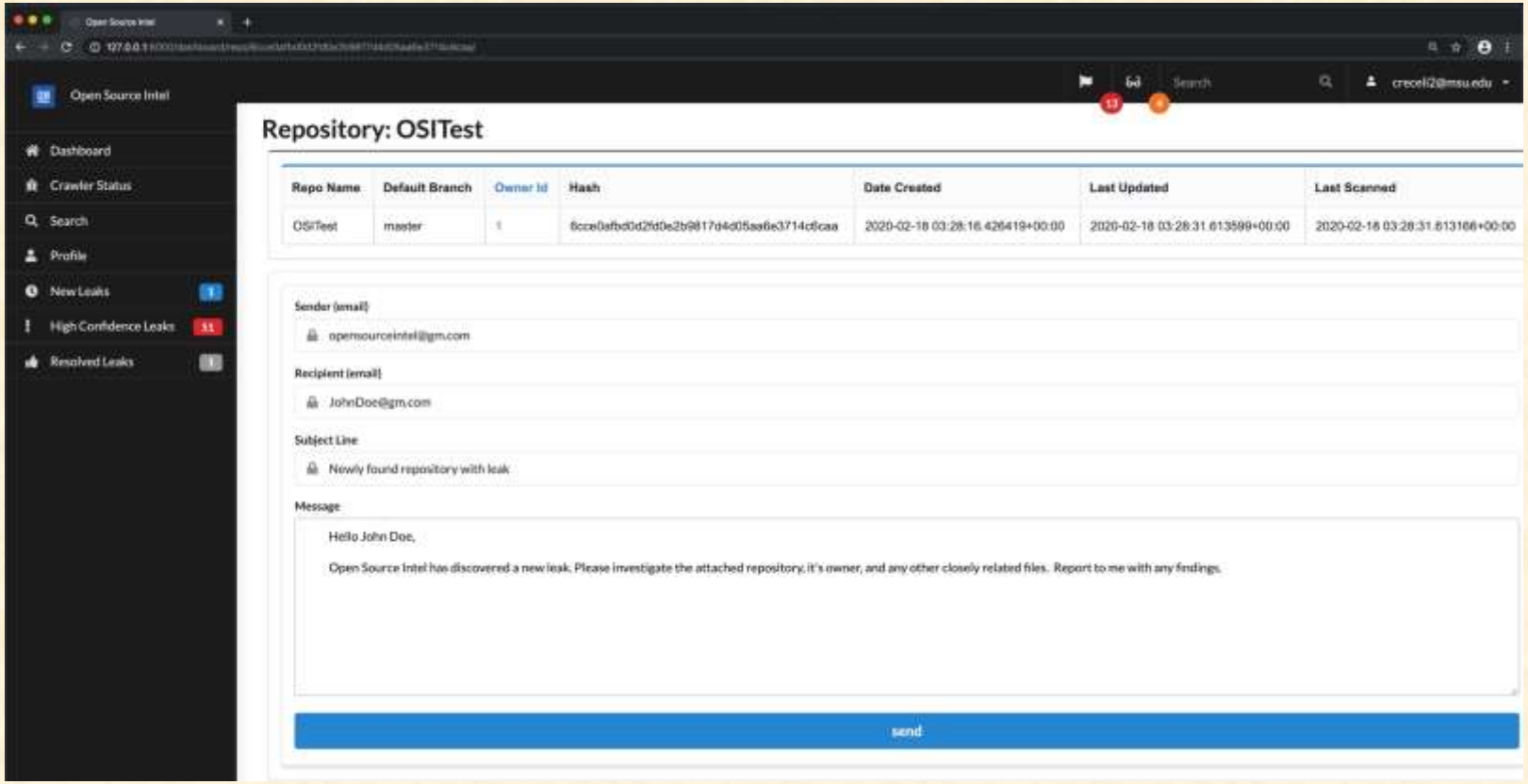
Leaks Page


Open Source Intel
Search
New Leaks
6
LEAKS DETECTED
2
NEW LEAKS
1
HIGH CONFIDENCE LEAKS
1
RESOLVED LEAKS
Dashboard
Crawler Status
Search
Profile
New Leaks
High Confidence Leaks
Resolved Leaks
Leak ID
Leak Link
Status
Date Created
Date Discovered
No Tags
02-18-2020
Page 1 of 1.

Single Leak Page



Repository Page



What's left to do?

- Scan and collect data from other public web open source hosts (BitBucket, PasteBin, etc.)
- Continue training and finetuning the machine learning module.
- Implement notification system and inter-user communication.

Questions?