



CyberHuskies

PII Safeguard - A PII Discovery/Redaction Browser Extension Tool
iSchool Capstone 2024

Introduction



Tom Jubin
PM, Cybersecurity



Carlos Cura
Data Science, M.L.



Sabrina Jahed
UX/UI Design



Noah Logman
Full-Stack Development



Blake Sketchley
Front-End Development



PROBLEM STATEMENT AND CONTEXT

What is PII?

Personally identifiable information (PII) is defined as any information connected to a specific individual that can be used to uncover that individual's identity, such as their social security number, full name, email address or phone number.

Why is it important?

legal compliance, cyber-hygiene, personal privacy, sensitive data security, limited damage from a data breach

How might users and organizations have the ability to discover, identify, and redact Personally Identifiable Information (PII) from documents at scale so that they can safeguard sensitive data to protect privacy and remain compliant?

KEY RESEARCH





Key Research

FERPA - Family Educational Rights and Privacy Act, governs access to educational records by public entities such as employers, publicly funded educational institutions, and foreign governments - common violations include failure to secure information, and leakage without consent, whether intentionally or not. HIPAA similar but for medical records

Our biggest concern was legal compliance, as privacy violations are one of the biggest legal headaches for companies and people in general. There were 4500 known breaches between 2005 and 2015 - this is old data, and if this is the amount of known breaches, there are definitely many more that went unreported. Providing a way for the average person to take data security into their own hands is crucial to reducing the damage caused by data breaches.

USER PERSONAS



ALEX PATTERSON, 32

- Cybersecurity Professional
- wants to prevent damage of future breaches
 - wants to integrate with existing workflows



JESSICA HELEN, 28

- HR Specialist
- company has changing regulations
 - wants to protect employee data




USER PERSONAS

User Needs:

- Changing regulations means we need a new software tool
- A precise PII detecting program to enhance the security of sensitive employee data since current systems lacked necessary precision for foolproof protection
- A program that seamlessly integrates into HR workflows to enhance data security and ensure compliance with privacy regulations

ETHICAL CONSIDERATIONS

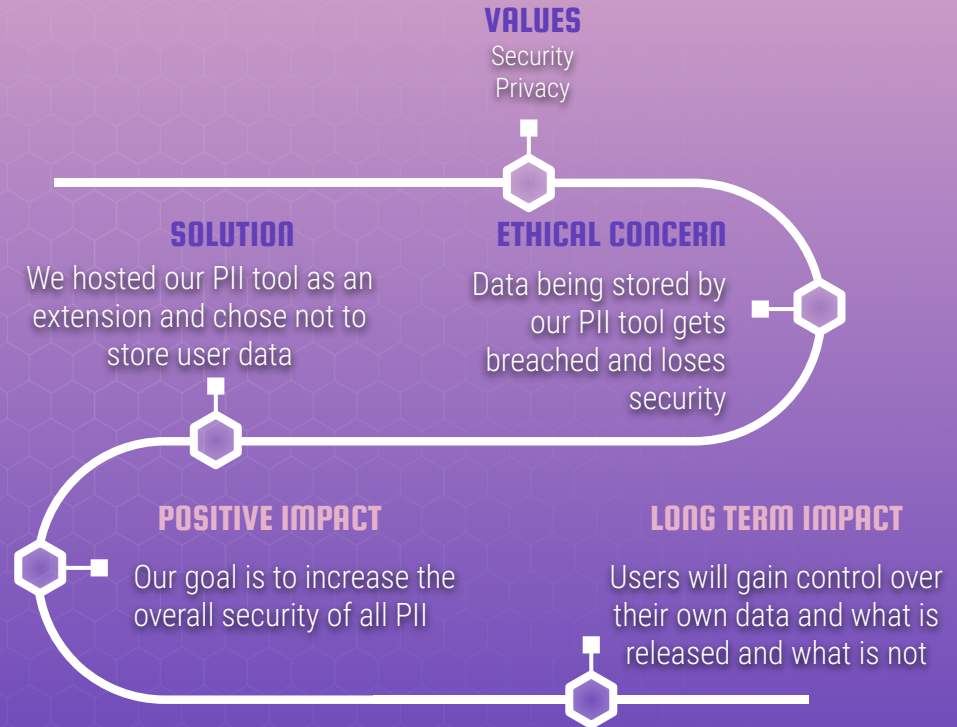


The main values in designing our project are privacy and security, this translates to a prominent ethical concern of data being stored by our PII tool potentially getting breached and personal data getting stolen.

The solution we decided on was to not store any user data, and instead, have them host it on their device as an extension. For example, like grammarly.

The positive impact our product would have is increasing the general security of PII data by essentially giving the power of detection back to its users

This translates to a longer term impact which is that users will have more control over what data is released and what is not



CONCEPT VALIDATION



FEASIBILITY

Browser extension
discovery tool instead
of website/app



USER VALIDATION

Provide many use
cases to cater to the
needs of dealing with
PII



LESSONS LEARNED

Simplicity is key for
easier access and
being intuitive

KEY TAKEAWAYS



SOLUTION

Chrome extension for PII tracking and redacting



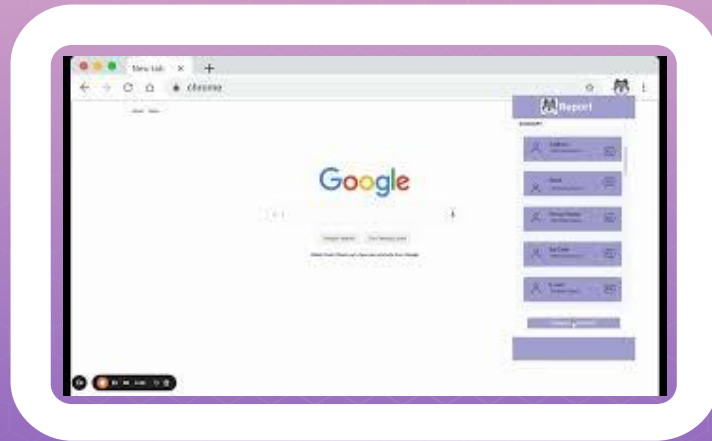
MEASURES

PII Identification
Data policy compliance

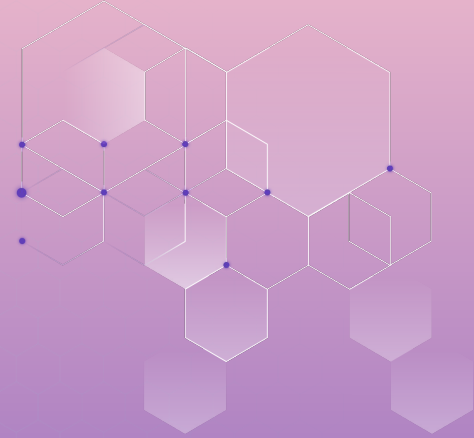
ETHICS

Privacy Protection
Compliance with
Regulations
Support Systems

DEMO



CyberHuskies PII SafeGuard Chrome Browser
Extension Software Tool Demo Video
https://youtu.be/WU_IeHY_c8s?si=RcZd_lfpZ7xKnrdR





CyberHusky

UPLOAD



Drag & drop files or [Browse](#)

Supported formats: CSV, JSON, PDF, XLS

UPLOAD FILES



CyberHusky

UPLOAD



Drag & drop files or [Browse](#)

Supported formats: CSV, JSON, PDF, XLS

UPLOAD FILES

Uploading - 1/1 files

Patients.csv

SELECT FILES

Main page 3



CyberHusky

UPLOAD



UPLOADING...

UPLOADING FILES

Uploading - 1/1 files

Patients.csv

Uploaded

EmployeeInfo.xls

NewOffers.PDF

SELECT FILES

Selection



CyberHusky

Select PII to Identify



Personal

Personal Name

Medical

Person Role

Financial

Email Address

Security

Home Address

Zip Code

SCAN FILES

Selection Medical



CyberHusky

Select PII to Identify



Personal

Body Structure

Medical

Dosage

Financial

Diagnosis

Security

Family Relation

Gene/Protein

Medication

Symptom

Treatment

Prescription

SCAN FILES

Selection Financial



CyberHusky

Select PII to Identify



Personal

Account Number

Medical

Credit Card

Financial

Debit Card

Security

IBAN

Routing Number

SCAN FILES

Selection Security



CyberHusky

Select PII to Identify



Personal

IP Address

Medical

Financial

Security

SCAN FILES



Report

SUMMARY



Social Security (SSN)

4 Entities Found
Revealing 1 entity

Credit Card

4 Entities Found



Marital Status

4 Entities Found



Person's Name

4 Entities Found



Zip Code

4 Entities Found



E-mail

25 Entities Found



DOWNLOAD REPORT

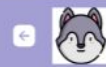


CyberHusky

SUMMARY - ZIP-CODE VIEW

 Zip-Code 98112 46182 72913 09251 15478

SAVE CHANGES

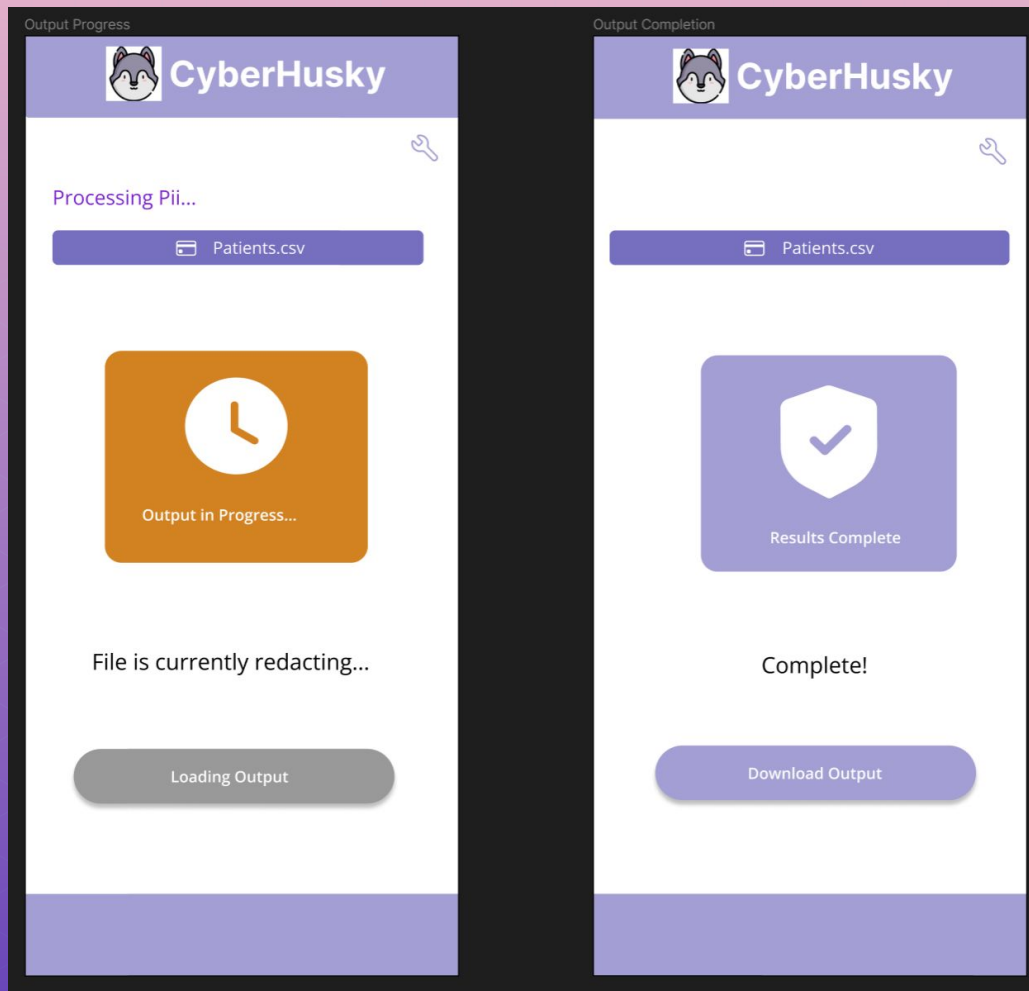
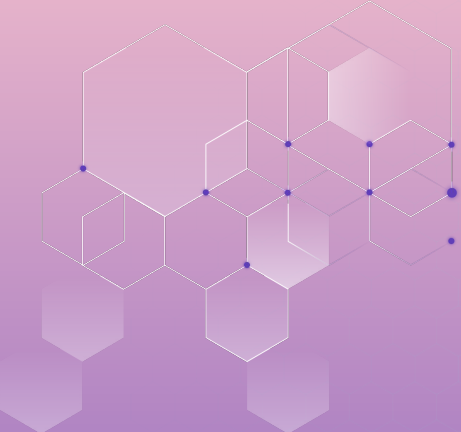


CyberHusky

SUMMARY - HOME ADDRESS VIEW

 Address 8371 MLK Jr Way East 3900 East Side Holly Jr 4167 Rainier Beach Ave 8167 Rainier Vista Park 4317 South Tacoma Drive Place

SAVE CHANGES



fname	lname	email	address	city	state	country	postcode	age	last_visit	next_visit	height	weight	heart_rate	diagnosis	prescription
Liam	XXXX	XXXXXX@XX.XXX	XXXX XXXX XXX,	XXXXX	XXX XXXX	XXXXXXXX XXXXXX	XXXXX	23	10/15/2022	11/19/2022	73	180	84	Anxiety	Zoloft
Noah	XXXX	XXXXXX@XXX.XXX	XXX XXX XX,	XXXXXXXX	XXXXXXXXXXXXXXXX	XXXXXXXX XXXXXX	XXXXX	49	10/03/2022	11/07/2022	69	160	78	High Blood Pressure	Lisinopril
William	XXXXXX	XXXXXXXX@XXX.XXX	XXX XXXXXXXXXXX XXX,	XXX XXXXX	XXXXXX	XXXXXXXX XXXXXX	XXXXX	21	10/10/2022	11/14/2022	74	170	97	Migraines	Imitrex
James	XXXXXXXX	XXXXXXXXXX@XXXXX.XXX	XXXX XXX XX,	XXX XXXX	XXX XXXX	XXXXXXXX XXXXXX	XXXXX	28	10/22/2022	11/26/2022	72	175	88	Diabetes	Metformin
Robert	XXXXXXXX	XXXXXXXXXX@XXXXXXXX.XXX	XXX X XXXX XX,	XXX XXXX	XXX XXXX	XXXXXXXX XXXXXX	XXXXX	71	10/08/2022	11/12/2022	68	185	67	Arthritis	Celebrex
Michael	XXXXX	XXXXXX@XXXX.XX	XXXX XXXXX XX X,	XXXXXXXX	XXXXXX	XXXXXX	XXX XXX	36	10/16/2022	11/20/2022	70	178	80	Depression	Prozac
David	XXXXXX	XXXXXXXX@XXX.XXX	XXXX XXXXXXXXXXX XXX XX,	XXXXXXXXXX	X.X.	XXXXXX XXXXXX	XXXXX	75	10/06/2022	11/10/2022	75	190	74	Back Pain	Ibuprofen
Daniel	XXXXXX	XXXXXXXX@XXXXXXXX.XXX	XXX XXXX XX,	XXXXXXXXXXXXXX	XXXXXXXXXXXXXX	XXXXXX XXXXXX	XXXXX	32	10/12/2022	11/16/2022	71	170	76	Allergies	Zyrtec
Alexander	XXXXXX	XXXXXX@XXXXXXXX.XXX	XXX XXXX XXX,	XXXXX	XXXXXX	XXXXXX XXXXXX	XXXXX	65	10/04/2022	11/08/2022	67	150	83	Cancer	Chemotherapy
Matthew	XXXXXXXX	XXXXXXXXXX@XXXXXXXXXX.XXX	XXX XXXX XX,	XXX XXXX	XXXXXXXXXXXX	XXXXXX XXXXXX	XXXXX	26	10/14/2022	11/18/2022	72	180	95	Asthma	Inhaler
Joseph	XXXXX	XXXXXXXX@XXX.XXX	XXX X XXXXX XX,	XXXXXX	XXXXXX	XXXXXX XXXXXX	XXXXX	40	10/18/2022	11/22/2022	73	175	90	Eczema	Topical Steroid Cream
Andrew	XXXXXX	XXXXXXXX@XXX.XXX	XXX XXXXXXXXXXX XX,	XXXXXXXX	XXXXXXXX	XXXXXX XXXXXX	XXXXX	48	10/07/2022	11/11/2022	74	165	101	Glaucoma	Timolol
Kevin	XXXXX	XXXXXXXX@XXXXXXXXXX.XXX	XXX XXXX XX,	XXXXXXXX	XXXXXX	XXXXXX XXXXXX	XXXXX	20	10/11/2022	11/15/2022	69	155	82	ADHD	Ritalin
Benjamin	XXXXXX	XXXXXXXXXX@XXXXXXXXXXXX.XXX	XXX XXXXXXX XX,	XXX XXXXXXX	XXXXXXXXXX	XXXXXX XXXXXX	XXXXX	51	10/09/2022	11/13/2022	68	180	70	Crohn's Disease	Humira
Ryan	XXXXXXXX	XXXXXXXXXXXX@XXXXXXXXXXXX.XXX	XXX XXXX XX,	XXXXXX	XXXXX	XXXXXX XXXXXX	XXXXX	25	10/17/2022	11/21/2022	75	175	87	Anxiety	Lexapro
Adam	XXXXXX	XXXXXXXX@XXXXXXXX.XXX	XXX XXXX XX,	XXXXXXXXXX	XXXX XXXXXXXXX	XXXXXX XXXXXX	XXXXX	69	10/05/2022	11/09/2022	71	195	65	Osteoarthritis	Glucosamine
Christopher	XXXXXX	XXXXXXXX@XXX.XXX	XXX XXXXX XX XX XXXX,	XXX XXXX	XXXXXX XXXX	XXXXXX XXXXXX	XXXXX	33	10/19/2022	11/23/2022	67	160	79	Sinusitis	Amoxicillin
Cameron	XXXX	XXXXXX@XXXX.XXX	XXX XXXX XXX,	XXX XXXX	XXX XXXX	XXXXXX XXXXXX	XXXXX	31	11/02/2022	12/06/2022	65	150	72	IBS	Bentyl
Sebastian	XXXXX	XXXXXX@XXXXX.XXX	XXX XXXX XX,	XXXX XXXXXXX	XXXXXXXXXXXX	XXXXXX XXXXXX	XXXXX	24	11/04/2022	12/08/2022	73	170	85	Acne	Retin-A
Gabriel	XXXX	XXXXXX@XXXXXXXX.XXX	XXX XXX XXX,	XXXXXX	XXXXXXXXXXXX	XXXXXX XXXXXX	XXXXX	43	11/06/2022	12/10/2022	76	180	90	GERD	Prilosec
Jack	XXXXXX	XXXXXXXXXX@XXX.XXX	XXX XXXX XX,	XXXXXX	XXXXX	XXXXXX XXXXXX	XXXXX	63	11/08/2022	12/12/2022	72	165	74	COPD	Inhaler
Liam	XXXXXXXX	XX@XXXXX.XXX	XXX XXXX XXX,	XXX XXXX	XXXXXXXXXXXX	XXXXXX XXXXXX	XXXXX	34	1/11/2023	2/15/2023	68	145	80	Depression	Zoloft
Noah	XXXXX	XX@XXXXXXXX.XXX	XXX X. XXXXX XX.	XXXXXXXX	XXXXXXXX	XXXXXX XXXXXX	XXXXX	48	1/13/2023	2/17/2023	70	170	75	Migraines	Imitrex
William	XXXXX	XX@XXX.XXX	XXX XXXXXXXXXXX XXX XX,	XXXXXXXXXX	X.X.	XXXXXX XXXXXX	XXXXX	22	1/15/2023	2/19/2023	75	190	82	Asthma	Albuterol
James	XXXXX	XX@XXX.XXX	XXX X. XXXXX XX.	XXXXXXXX	XXXX XXXXXXXXX	XXXXXX XXXXXX	XXXXX	56	1/17/2023	2/21/2023	68	200	77	High Blood Pressure	Lisinopril
Alexander	XXXXXX	XX@XXXXXXXX.XXX	XXX XXXXXXX XXX,	XXXXXX	XXXXXXXXXXXXXX	XXXXXX XXXXXX	XXXXX	47	1/19/2023	2/23/2023	73	160	90	Eczema	Triamcinolone
Michael	XXXXXX	XX@XXXXXXXX.XXX	XXX XXXX XX.	XXXXX	XXXXXX	XXXXXX XXXXXX	XXXXX	36	1/21/2023	2/25/2023	67	180	73	Back Pain	Ibuprofen
Daniel	XXXXX	XX@XXX.XXX	XXX XXXXX XX.	XXXXXX	XXXXXXXXXX	XXXXXX XXXXXX	XXXXX	29	1/23/2023	2/27/2023	72	190	88	Diabetes	Metformin
Matthew	XXXXXXXX	XX@XXX.XXX	XXX XXXX XX.	XXXXXX	XXXXX	XXXXXX XXXXXX	XXXXX	65	1/25/2023	3/1/2023	76	200	70	Arthritis	Celebrex
Elijah	XXXXXX	XX@XXXXXXXXXX.XXX	XXX XXXXXXX XX.	XXXXX	XXXXXX	XXXXXX XXXXXX	XXXXX	21	1/27/2023	3/3/2023	73	155	85	Anxiety	Buspirone
Joseph	XXXXXXXX	XX@XXXXX.XXX	XXX XX XXXX XX.	XX XXXX	XXXXX	XXXXXX XXXXXX	XXXXX	38	1/29/2023	3/5/2023	72	185	71	GERD	Omeprazole
Andrew	XXXXXXXXXX	XX@XXXXX.XXX	XXX XXXXX XX.	XXX XXXXXXX	XXXXX	XXXXXX XXXXXX	XXXXX	40	2/1/2023	3/7/2023	67	140	90	Insomnia	Zolpidem
Benjamin	XXXXX	XX@XXXXXXXXXX.XXX	XXX XXXXXXX XX.	XXX XXXXXXX	XXXXXXXXXXXX	XXXXXX XXXXXX	XXXXX	67	2/3/2023	3/9/2023	69	175	74	Glaucoma	Travatan
Samuel	XXXXXXXX	XX@XXX.XXX	XXX XXXX XX.	XXXXXX	XXXXXXXXXX	XXXXXX XXXXXX	XXXXX	52	2/5/2023	3/11/2023	70	190	81	High Cholesterol	Lipitor
David	XXXXXX	XX@XXXXXXXX.XXX	XXX XXXXXXX XXX.	XXXXXXXXXXXX	XXX XXXXXXX	XXXXXX XXXXXX	XXXXX	36	2/7/2023	3/13/2023	72	150	88	IBS	Donnatal
Anthony	XXXXXX	XX@XXXXXXXX.XXX	XXX XXXX XX.	XXXXXX	XXXXX	XXXXXX XXXXXX	XXXXX	59	2/9/2023	3/15/2023	68	170	79	Prostate Cancer	Tamsulosin

NEXT STEPS



CLOSING PROJECT

We have decided to closeout our project



CURRENT AGENDA

With immediate post-graduation plans, we will put this project on pause



LOOKING FORWARD

As AI develops further, we plan to revisit our project in the next 5 years hopefully



THANK YOU!