TOPIC: *Where's My Email? Mapping Your Online Accounts*

Authors:

Abid Banthanal (banthanalabid17@gmail.com)

Rajashekar (marajashekar24@gmail.com)

Darshan N (darshan220044@gmail.com)

VijethM (vijeth.m.2004@gmail.com)

Abstract:

In today's digital landscape, individuals create online accounts across numerous platforms, resulting in an extensive and often unmanaged digital footprint. Over time, many users lose track of where their email addresses have been used, increasing the risk of data breaches, identity theft, and privacy violations. This project, ***"Where's My Email? Mapping Your Online Accounts,"*** aims to develop a system that enables users to identify and monitor online accounts associated with their email addresses.

By leveraging data breach repositories, open-source intelligence (OSINT) techniques, and email-based lookup tools, the system offers detailed insights into linked accounts, potential vulnerabilities, and exposed credentials. The project seeks to enhance cybersecurity awareness by empowering users with greater visibility over their digital presence, ultimately supporting informed decision-making in account management and proactive risk mitigation.

Introduction:

In the modern digital era, individuals interact with countless online platforms, services, and applications, often using a single email address as the central identifier. Whether signing up for social media accounts, newsletters, e-commerce platforms, or cloud services, the email address becomes the primary gateway to one's digital identity. As the number of registered accounts grows over time, users tend to forget where their credentials have been used, leaving behind a trail of unused or forgotten accounts.

This phenomenon contributes to a growing issue: the **unmanaged digital footprint**. Many of these forgotten accounts may still contain personal data, remain vulnerable to cyberattacks, or be involved in data breaches without the user's awareness. In addition, users may unknowingly continue using compromised email addresses linked to sensitive services, further escalating security and privacy risks.

Several large-scale data breaches have made headlines in recent years, exposing millions of email addresses and passwords. Platforms like *HaveIBeenPwned,* and *Firefox Monitor* have emerged as breach notification tools, allowing users to check if their email addresses have been compromised in known breaches. However,

these tools are limited in scope—they do not offer full visibility into where an email has been used, nor do they provide a structured mapping of all associated online accounts.

In this context, the need for a system that maps user email addresses to their corresponding online accounts becomes evident. Such a system could play a critical role in enhancing cybersecurity hygiene by helping users identify forgotten accounts, recognize exposed credentials, and act before potential damage occurs.

The project titled **"Where's My Email? Mapping Your Online Accounts"** is designed to address this gap by leveraging publicly available data, breach repositories, and open-source intelligence (OSINT) techniques to provide a detailed view of a user's email-based account activity. The goal is not only to improve user awareness but also to empower them with actionable insights to take control of their digital presence.

1.1 Problem Statement & Objectives:

The modern internet user relies heavily on email addresses to create, access, and manage a wide variety of online services. From social media platforms and e-commerce websites to productivity tools and cloud-based storage, an email address acts as a unique identifier—serving not only for authentication but also for communication and account recovery. Over time, users may sign up for dozens or even hundreds of websites using a single email, many of which are later forgotten or abandoned.

This gradual accumulation of accounts contributes to what is known as a **digital footprint**—the trace a user leaves behind through their online activities. While some of these traces are intentional and active, others remain dormant, leaving personal data exposed or at risk. The unmanaged nature of these forgotten accounts increases the potential for data breaches, phishing attacks, identity theft, and privacy violations. In particular, when users are unaware of where their email addresses have been used, they are less likely to take preventive actions such as deleting inactive accounts, updating passwords, or enabling security features like two-factor authentication.

While tools like *HaveIBeenPwned* and *Firefox Monitor* offer breach notifications for known data leaks, they do not provide a comprehensive view of where a particular email address has been used online. Similarly, OSINT tools such as *theHarvester*, *Maigret*, and *Recon-ng* have shown promise in email and username tracing, but these tools are often built for cybersecurity professionals rather than general users, making them complex and inaccessible for the average individual.

This project, titled **"Where's My Email? Mapping Your Online Accounts,"** seeks to address these limitations by offering a system that actively traces, compiles, and presents the online services associated with a user's email address. By integrating data breach databases, public information scraping, and open-

source intelligence techniques, the system empowers users to visualize their digital presence, identify risks, and manage their online identities more securely.

1.2 Problem Statement

As the internet continues to integrate deeper into daily life, individuals are creating online accounts across a wide range of services. From social networking and cloud storage to e-commerce and educational platforms, these registrations typically require just one element of identity: an email address. Over time, this leads to an extensive yet often unmonitored digital footprint, with users losing track of the various platforms where their email has been registered.

This lack of visibility exposes users to serious privacy and security risks. Forgotten or unused accounts may still contain sensitive personal information such as passwords, addresses, contact details, and even financial data. If any of these services experience a data breach, the user may not even be aware that their credentials have been exposed. In worst-case scenarios, these credentials may be sold on the dark web or used in credential stuffing attacks across other platforms.

While tools like *HaveIBeenPwned* provide valuable breach alerts, they are reactive in nature and limited in scope. Users still do not receive a comprehensive overview of where their email has been used, how many accounts exist under it, or which of these accounts are still active, inactive, or potentially vulnerable.

There is, therefore, a strong need for a centralized, user-friendly system that helps users trace and manage their digital presence by mapping their online accounts linked to their email addresses. A solution that not only highlights breach exposures but also improves visibility, encourages cleanup of unused accounts, and promotes proactive cybersecurity behavior is the need of the hour.

1.3 Objectives of the Study

The primary objectives of this literature survey and subsequent project are as follows:

1. **To study and evaluate** the existing tools, techniques, and research methodologies related to email-based account discovery and digital footprint analysis.
2. **To explore the potential of OSINT (Open-Source Intelligence)** in identifying online accounts and services associated with a user's email.
3. **To analyze existing breach alert services** such as *HaveIBeenPwned*, *DeHashed*, and *Firefox Monitor* in terms of their capabilities and limitations.
4. **To identify research gaps** in current solutions that fail to offer a complete map of a user's online presence.

5. **To justify the need for a system** that consolidates information about account exposure, breach incidents, and unused account risks.
6. **To lay the foundation** for designing a privacy-aware, user-friendly solution that empowers users to take control of their digital identity.

1.4 Methodology

The methodology adopted for this literature survey is structured to identify, evaluate, and compare existing research papers, tools, and techniques relevant to email-based account discovery, digital footprint tracking, and data breach awareness. The goal is to analyze the current state of knowledge in the field and identify gaps that justify the need for the proposed system.

1.4.1 Research Approach

A qualitative research approach was followed, focusing on the critical analysis of peer-reviewed articles, cybersecurity case studies, whitepapers, and OSINT documentation. The selection was guided by relevance to the core themes of this project, including:

- Email-based identity mapping
- Open-source intelligence (OSINT) tools and frameworks
- Data breaches and credential exposure
- Online account security and user privacy
- User behavior in digital identity management

1.4.2 Source of Data

The following sources were used to gather literature:

- **Academic Databases:** IEEE Xplore, SpringerLink, Elsevier (ScienceDirect), ACM Digital Library
- **Security Forums & Whitepapers:** OWASP, Kaspersky Labs, SANS Institute, etc.
- **Open-source Tools Documentation:** Official GitHub repositories and usage guides of tools such as *theHarvester*, *Maigret*, *Recon-ng*, and *Sherlock*

- **Breach Lookup Platforms:** *HaveIBeenPwned*, *DeHashed*, and *Firefox Monitor*
- **Google Scholar and Semantic Scholar:** For discovering recent and cited research articles

1.4.3 Keyword Strategy

To ensure comprehensive coverage, the following keywords and phrases were used while searching for relevant studies:

- "Email-based account discovery"
- "OSINT tools for cybersecurity"
- "Digital footprint mapping"
- "Data breach analysis"
- "Cyber hygiene and privacy"
- "Online account exposure"
- "Forgotten accounts and privacy risk"

1.4.4 Selection Criteria

The selection of papers and sources was based on:

- **Relevance:** Direct connection to email discovery, OSINT, and digital footprint
- **Recency:** Preference to works published after 2015 for relevance to modern threats
- **Credibility:** Peer-reviewed and cited publications, official documentation, or reputable cybersecurity blogs

1.5 Literature Survey

Understanding where an individual's email has been used online and which platforms it is linked to has become increasingly important in the field of cybersecurity. A significant amount of research has been conducted around email-based breaches, digital footprint analysis, and OSINT methodologies for data tracing. This section presents a thematic literature review of existing tools, platforms, and research studies that contribute to this domain.

1.5.1 Data Breach Monitoring Tools

Several online platforms have emerged to help users check if their email addresses have been compromised in known data breaches. These tools work by collecting publicly leaked databases from various breach incidents and allowing users to query their email addresses to check for exposure.

HaveIBeenPwned (Troy Hunt, 2013–Present)

HaveIBeenPwned (HIBP) is one of the most widely used data breach checking platforms. It aggregates publicly disclosed breaches and allows users to search their email addresses or domains. The tool not only lists breached websites but also discloses the type of data exposed—such as passwords, phone numbers, or location data.

Limitations:

- Only includes publicly known breaches
- Offers limited insights into active accounts, focusing instead on historical exposure

DE Hashed

DE Hashed offers a more advanced (and partially paid) service that scans through large dumps of leaked data, allowing searches by username, email, phone number, or even IP address. It is primarily used by cybersecurity professionals and law enforcement.

Limitations:

- Not user-friendly for general users
- Access to deeper information often requires a subscription

Firefox Monitor

Developed by Mozilla, Firefox Monitor integrates HIBP data and offers breach alerts directly to Firefox users. It is simple and easy to use, making it accessible for the average internet user.

Limitations:

- Limited to breach notifications
- Does not map or trace the email across all account types

These tools are vital in raising user awareness about breaches. However, they are **reactive** in nature. They inform users **after** the breach has occurred and do not provide a complete overview of where a user's email has been registered.

1.5.2 OSINT Tools for Email and Account Discovery

Open-Source Intelligence (OSINT) tools are widely used in cybersecurity investigations to gather publicly available data from the internet. When applied to email tracking, these tools can help identify usernames,

platforms, domains, and even social media accounts associated with a given email address. They form the technical backbone for any system attempting to map an individual's online presence.

a) theHarvester

Overview:

theHarvester is a powerful reconnaissance tool used to extract emails, subdomains, hostnames, and IPs from public sources like search engines, PGP key servers, and DNS records. Although it is not designed exclusively for email mapping, it provides a good starting point for identifying where an email might be referenced.

Use Case in Context:

By collecting all mentions of a particular email address on public-facing platforms, *theHarvester* can indicate potential account usage or exposure.

Limitations:

- Requires technical knowledge to use
- Data might include false positives or outdated results
- Not specifically tailored to account mapping

b) Maigret

Overview:

Maigret is a specialized OSINT tool that searches for accounts associated with a given username or email across over 500 websites, including social media platforms, coding forums, and blogging sites.

Use Case in Context:

This tool aligns very closely with the goal of this project. It automates the process of checking for registered accounts across a wide variety of sites and is ideal for identifying where an email or alias might be active.

Limitations:

- Cannot confirm account activity (only presence)
- May fail on sites with anti-bot or login protections

c) Sherlock

Overview:

Sherlock is another OSINT tool similar to Maigret, focused mainly on social media usernames. Although it

does not support email queries directly, it's often used to complement email-based scans by correlating usernames.

Use Case in Context:

Useful when an email address reveals a username (via breach data), which can then be used in Sherlock to expand the identity profile.

Limitations:

- Does not support email input
- Primarily used for usernames, not email tracking

d) Recon-ng

Overview:

Recon-ng is a full-featured web reconnaissance framework. It offers a modular interface where you can load different modules (like Metasploit) to gather information based on email addresses, domains, or IPs.

Use Case in Context:

When customized with the right modules, Recon-ng can extract information from public data sources, social media, and even breached credential dumps.

Limitations:

- High learning curve
- Primarily used by penetration testers and researchers

1.5.3 Academic Studies on Digital Footprint & Privacy

While tools and frameworks provide practical approaches, several academic researchers have contributed to understanding how digital footprints, especially through email usage, affect privacy and security. These studies emphasize the long-term implications of uncontrolled data exposure and highlight the need for better account management practices.

a) "The Web Never Forgets: Persistent Tracking Mechanisms in the Wild" (Acar et al., 2014)

Summary:

This paper highlights how online trackers and third-party services can continuously monitor a user's activity using various identifiers, including email addresses. While the focus is on cookies and fingerprinting, it indirectly shows how accounts linked to emails can create long-term digital trails.

Relevance to Project:

Demonstrates the lasting nature of digital presence, making forgotten email-linked accounts a long-term privacy concern.

b) "An Empirical Study of Deleted Usernames and Account Reuse on the Web" (Thomas et al., 2016)

Summary:

This study analyzes how abandoned or deleted usernames (and emails) can be reclaimed or reused on various platforms, leading to impersonation, spoofing, and identity theft.

Relevance to Project:

Highlights the importance of identifying old or inactive email-based accounts, which are often overlooked by users.

c) "Understanding the Privacy Practices of Email Service Providers" (Ruoti et al., 2020)

Summary:

The research explores how email service providers handle privacy, metadata, and security. It reveals the inconsistency in privacy settings across services and emphasizes how users unknowingly expose more than they realize.

Relevance to Project:

Reinforces the idea that email is not just a login method — it can also leak metadata, identity, and associated accounts.

d) "The Forgotten: Accounts that Never Log Out" (User Behavior Study by Cybersecurity Ventures, 2021)

Summary:

This industry report found that over 60% of users have more than 20 unused or forgotten online accounts, many of which are still linked to their primary email. These dormant accounts are rarely monitored or secured.

Relevance to Project:

Directly supports the project's objective — there is a clear gap in tools that help users monitor where their email has been used over time.

Key Takeaway from Literature

While tools can surface technical evidence of email usage, academic studies validate the psychological and behavioral patterns of users — they forget, abandon, or ignore many accounts. Combined, both aspects show the necessity of building a tool that:

- Maps email-based account history
- Notifies users of dormant accounts
- Helps prevent risks tied to forgotten digital traces

1.5.4 Comparative Analysis of Existing Tools and Studies

To better understand the scope and limitations of existing solutions related to email-based account discovery, a comparative analysis was performed. The table below compares popular tools, frameworks, and studies based on their capabilities, ease of use, focus area, and relevance to the project's goals.

📊 Comparative Table:

Tool/Study	Email Mapping	Breach Alerts	OSINT-based	User Friendly	Free Access
HaveIBeenPwned	❌	☑	❌	☑	☑
DeHashed	☑	☑	☑	❌	❌ (limited)
Firefox Monitor	❌	☑	❌	☑	☑
theHarvester	☑	❌	☑	❌	☑
Maigret	☑	❌	☑	⚠ Moderate	☑
Sherlock	❌ (Username only)	❌	☑	⚠ Moderate	☑
Recon-ng	☑	❌	☑	❌ (advanced use)	☑
Academic Studies	☑ (theoretical)	☑ (risk focus)	⚠ Indirect	☑ (conceptual)	☑

Insights from the Comparative Study:

- **No single tool** offers full email-based account mapping with both breach alert integration and OSINT intelligence in one place.
- **Most OSINT tools** are technical and command-line based, limiting accessibility for non-tech-savvy users.
- **Breach tools are reactive**, not proactive. They notify after leaks occur but don't prevent or identify at-risk, dormant accounts.

- **Academic research** supports the behavioral findings — people forget their digital traces, which remain online as vulnerabilities.

These observations strongly support the need for a system that:

- Maps **all online accounts** linked to an email
- Flags potentially risky, unused, or old accounts
- Combines breach data + OSINT + a clean user interface
- Helps users take **proactive steps** in managing their digital identity

1.6 Identified Research Gaps

Despite the growing number of cybersecurity tools and studies aimed at data breach awareness and identity protection, significant gaps remain in the current ecosystem when it comes to comprehensive email-based account mapping. This section outlines key limitations observed in existing work and tools.

1.6.1 Incomplete Coverage of Email Usage

Most current solutions, such as *HaveIBeenPwned* or *Firefox Monitor*, only inform users when their email appears in a **known breach database**. These platforms do not reveal a list of all websites or services where the email has been used. As a result, users are unaware of dormant, legacy, or obscure accounts that may still be active.

Lack of Unified Framework

Tools like *theHarvester*, *Maigret*, and *Recon-ng* offer fragmented features. One might scan social media accounts, while another might search for breached credentials. However, there is **no unified system** that brings together:

- OSINT scans
- Breach data
- Visual mapping
- User-friendly recommendations

Technical Barriers for Non-Experts

Many of the best tools in this domain are command-line based and designed for cybersecurity professionals. There is a **lack of accessible tools** that average internet users or students can use without prior training. This technical barrier prevents widespread adoption.

Reactive vs. Proactive Security

Most breach-related tools operate **after** an incident has occurred. There are **very few proactive systems** that encourage users to clean up or secure their online footprint **before** a breach happens. Users need solutions that help them:

- Identify forgotten accounts
- Delete unused registrations
- Secure active ones

No Visualization or Mapping Feature

While tools exist to search and list potential linked accounts, none of them provide a **graphical visualization or intuitive dashboard** showing where and how the email is being used. This makes understanding one's digital presence more difficult.

Relevance to the Proposed System

The gaps identified directly support the need for the proposed project:
"Where's My Email? Mapping Your Online Accounts" aims to be an all-in-one tool that bridges these shortcomings by:

- Merging breach data + OSINT in a single workflow
- Making it usable even for non-technical individuals
- Visualizing the digital footprint in a user-friendly way
- Helping users manage or clean up their accounts proactively

1.7 Link to the Proposed System

Building upon the reviewed literature, existing tools, and identified gaps, the proposed project **"Where's My Email? Mapping Your Online Accounts"** introduces a novel approach to digital footprint visibility. The system integrates capabilities from OSINT tools, breach databases, and user interface principles to offer a centralized, accessible platform for everyday users concerned about their online presence.

Bridging Gaps with Innovation

The survey clearly indicates that while there are multiple standalone tools for specific tasks — such as breach alerts or social media scanning — **no single tool combines everything** in a user-centric, proactive solution. This project is designed to bridge that gap by offering:

- **Email-based account discovery** using open-source intelligence
- **Breach status analysis** through integration with public databases (e.g., HIBP)
- **Cross-platform linkage** of accounts that share the same email or alias
- **Dashboard-style visualization** of active, inactive, and possibly forgotten accounts
- **Account risk classification** based on age, exposure, and password reuse signals

Proposed Architecture at a Glance

The system will follow a modular architecture:

1. **Input Layer:** User enters one or more email addresses
2. **Discovery Layer:** Queries APIs, OSINT tools (Maigret, Recon-ng modules), and breach sources
3. **Analysis Layer:** Processes matches, filters duplicates, classifies risk levels
4. **Visualization Layer:** Interactive UI dashboard showing linked services, breach history, and suggested actions
5. **Recommendation Engine (Optional):** Flags suspicious or old accounts for review or deletion

Why This System is Unique

- **Proactive, not reactive**: Warns users of potential risks before incidents occur
- **No advanced tech skills needed**: Designed for students, professionals, and everyday users
- **Visual overviews**: Helps users clearly understand and manage their digital footprint
- **Scalable and modular**: Future upgrades can include phone number, username, or IP mapping

1.8 Future Scope & Opportunities

The scope of "Where's My Email? Mapping Your Online Accounts" goes beyond its initial goal of identifying and visualizing online accounts linked to a user's email. As cybersecurity threats evolve and user

data continues to be spread across the internet, the system can be enhanced in various directions to become a complete **digital identity control centre**.

1.8.1 Multi-Identity Tracking

- **Usernames & Aliases:** Future iterations can allow users to input commonly used usernames, which can then be correlated with the email scan for a deeper footprint analysis.
- **Phone Number Linking:** Integration with services that track accounts via phone numbers can provide more complete results, especially for services that use OTP logins (e.g., WhatsApp, PayTM, etc.).

1.8.2 AI-Powered Risk Scoring

- **Behavioral Analysis:** Use machine learning to analyze which types of accounts are likely to be abandoned or risky.
- **Risk Scoring:** Assign risk levels to accounts based on time of last use, password reuse, breach history, and service reputation.

1.8.3 Cross-Platform Integrations

- **Browser Extensions:** Allow real-time email usage tracking as users sign up on new websites, helping them keep an updated map.
- **Cloud Backup Integration:** Sync data with secure cloud platforms to back up account information securely (optional).

1.8.4 Privacy-First Account Clean-Up

- **Automated Deactivation Requests:** Tools to guide users through deleting or deactivating unwanted accounts, while preserving privacy.
- **Anonymization Alerts:** Warn users if their credentials are indexed by search engines or public data leaks.

1.8.5 Enterprise/Academic Use Case

- **Training Module for Students:** A version of the tool can be adapted for educational purposes to teach students about digital hygiene and responsible online behavior.
- **Corporate Audits:** Organizations can offer employees a way to track exposure of official email addresses for internal cybersecurity audits.

1.9 Conclusion

In today's hyper-connected world, where email addresses serve as gateways to countless digital services, the importance of maintaining visibility and control over one's online presence cannot be overstated. The review of existing tools, platforms, and research clearly reveals that while various solutions exist for data breach notifications and OSINT-based scanning, there remains a **critical gap** in providing users with a **centralized, proactive, and user-friendly system** for mapping their email-linked accounts.

The tools examined — from *HaveIBeenPwned* and *DeHashed* to *Maigret* and *theHarvester* — serve specific purposes but fall short in delivering a holistic picture. Similarly, academic research highlights the psychological and technical challenges users face when managing digital footprints, particularly in the absence of tools tailored for this purpose. These gaps, combined with the growing threat of credential stuffing, identity theft, and dormant account exploitation, demand a new kind of solution.

The proposed system, **"Where's My Email? Mapping Your Online Accounts,"** directly addresses these issues. By leveraging OSINT techniques, breach databases, and intelligent classification mechanisms, it empowers users with actionable insights and control. Unlike traditional tools, it focuses not just on post-breach alerts, but on proactive identity management — helping users reduce digital clutter, minimize exposure, and strengthen their cybersecurity posture.

As digital identity becomes a core concern for individuals and organizations alike, projects like this stand at the intersection of usability, security, and awareness. This literature survey provides a solid foundation for further development, encouraging a smarter, safer, and more informed interaction with the digital world.

2.0 References

1. T. Hunt, Have I Been Pwned, [Online]. Available: https://haveibeenpwned.com

2. DeHashed, Search Engine for Leaked Data, [Online]. Available: https://www.dehashed.com

3. Mozilla, Firefox Monitor, [Online]. Available: https://monitor.firefox.co

4. theHarvester, GitHub Repository. [Online]. Available: https://github.com/laramies/theHarvester

5. Maigret, GitHub Repository. [Online]. Available: https://github.com/soxoj/maigret

6. Sherlock, GitHub Repository. [Online]. Available: https://github.com/sherlock-project/sherlock

7. Recon-ng, GitHub Repository. [Online]. Available: https://github.com/lanmaster53/recon-ng

8. G. Acar et al, "The Web Never Forgets: Persistent Tracking Mechanisms in the Wild," in Proceedings of the ACM SIGSAC 2014

9. D. Thomas et al, "Account Reuse and Impersonation on the Web," in USENIX Security Symposium, 2016.

10. S. Ruoti et al, "Understanding the Privacy Practices of Email Providers," IEEE Security & Privacy, vol. 18, no. 2, 2020

11. Cybersecurity Ventures, "The Forgotten Accounts Report," 2021.

12. OWASP, Open Source Intelligence (OSINT) Resources, [Online]. Available: https://owasp.org/www-community/OSINT_Tool

13. SANS Institute, "Digital Footprint and Privacy Risk Analysis," [Online Whitepaper], 2022

14. Kaspersky Labs, "Data Leaks and Email Exploits in 2020s," [Online]. Available: https://www.kaspersky.com/blog/data-leaks-analysis

15. SpringerLink, Digital Identity Research Collection Online]. Available: https://link.springer.com