****

**School of Computing**

**Final Year {Engineering / Research / Study} Project**

**Project Initiation Document**

**{your name}**

**Maximising privacy, anonymity and reduction of digital footprint over the internet**

# Basic details

|  |  |
| --- | --- |
| Student name: |  |
| Draft project title: |  |
| Course and year: |  |
| Project supervisor: |  |
| Client organisation: |  |
| Client contact name: |  |

# 

# Degree suitability

The proposal "**Maximising privacy, anonymity and reduction of digital footprint over the internet**" meet the criteria for my current course for a Software Engineering degree. The idea tackles a critical issue in today's over internet: the need for improved digital security against rising data breaches.

# Outline of the project environment and problem to be solved

*Explain the context of the project - a bit like setting the scene:*

|  |  |  |
| --- | --- | --- |
| *For engineering* ***projects without a client****:* | *For projects* ***with a client****:* | *For* ***research*** *or* ***study*** *projects:* |
|  |  | *Since internet became such a widespread means of communication, digital privacy has been a source of concern. A digital footprint is a personal information that is available online. While many people have a digital footprint, only some, if any, understand what it entails or how to control it. Technology and personal information are becoming increasingly inextricably linked as technology gets more incorporated into daily activities. Personal information refers to facts about a person, such as race or buying habits. Many organisations regard shopping practices as personal information & spend money tracking or predicting individual purchases. However, more conventional kinds of personal information include facts such as gender, DOB, and residence town. With so much personal information available, not all of it is equally useful or uniquely personal.* *This research aims to enlighten and apprise clients about the breadth and depth of their digital footprint and educate how they maximise privacy and anonymity and reduce their digital footprint over the internet.* |

# Project aim and objectives

This research aims to enlighten and apprise clients about the breadth and depth of their digital footprint and educate how they maximise privacy and anonymity and reduce their digital footprint over the internet.

Here are the main objectives of our research.

* Identifying the amount of sensitive data that is exposed nowadays and then evaluating the effectiveness of that exposure.
* Proposed maximize privacy and anonymity and reduce measures for mitigating digital footprint over the internet and improve overall business functions.
* Sensitivity ranking for different areas of personal information to examine the effect of sensitive data.

# Project constraints

The main restriction of the concept is the need for an appropriate implementation for the research demands, such as setup the lab.

# Facilities and resources

The main requirement for the successful completion of this research project are mentioned below:

Environment: Windows OS 64 Bit

* License: Openly licensed data and instruments
* Hardware Equipment: The directed computer resources, including high-speed internet with 8GB RAM.
* University Resources, scholarly materials, and IEEE are examined.

# Log of risks

|  |  |  |
| --- | --- | --- |
| **Description** | **Impact** | **Mitigation/Avoidance** |
| Technical risk | Isolation failure, Resource exhaustion, Intercepting data in transit, Insecure of Ineffective deletion of data | Network monitoring to spot early signs of failure. |
| Legal risks | Issues due to jurisdiction, Data protection risks | It would help if you guaranteed that the only persons who have access to this data are those who might and should. Row-level security is required to guarantee that you can give quick access to the information for analysis while also controlling what users can view. |
| Unauthorized access | Steal files, data, and other informations | Instructing clients to include pictures, numbers, and a combination of things makes passwords more difficult to crack. |

# Project deliverables

# Chapter 1 Introduction:

Humans are social creatures. To feel a part of and operate well in society, we form relationships, and revealing information about ourselves is a natural part of this communication process. In the physical world, we guard our privacy while exchanging information with someone or society or engaging in unique or communal activities. We are more accustomed to regulating what we hide or disclose in the real world, but protecting privacy in the digital context, where we are still working to integrate, is significantly more difficult (Moore, 2012). Because businesses are participating in the internet world, they are responsible for most internet platforms and other relevant online services.

Netizens leave a massive digital footprint, whether deliberately or unwittingly, by leaving individually identifiable data from their daily activities of producing, sharing, & accessing material via online platforms (Arakerimath, 2015).

## 1.1 Brief History of Web Tracking & Targeted Advertising using digital footprinting

History's netizens are bombarded with a slew of online activity monitors. In truth, the history of data security and personal privacy is as ancient as the commercial internet itself. Since commercialized broadband internet in the mid-nineties, data privacy issues have been on an ever-increasing rise. As digital technology gets more complicated, this tendency continues (Yu, 2022).

The creation of two apps, Gopher2 or the World Wide Web (WWW), in the early 1990s (Peters, 2022) cleared the path for everyone to use the internet. Soon after, the internet evolved into "a global system integrating thousands of computers and putting them at the heart of a new transmission channel."

## 1.2 What is a digital footprint, so how does it affect your internet privacy?

Whenever you go online, you create a digital footprint. These digital footprints have advantages and disadvantages. You may discover a website you visited several days ago using preserved browser history. Cookies on websites enhance your online activities to be more personalised & accessible. However, with each digital footprint, you lose some level of privacy. We cover how your online privacy is tied to your digital footprints in the article and how to keep your personal files secure, even while utilising vulnerable public wifi (Ye, 2022).

## 1.3 The rises of third parties cookies

Cookies, particularly 3rdparties  cookies, are essential to the contemporary Internet but pose a risk to privacy protection (Janné, 2022).Third-party cookies are stored on a user's mobile or desktop computer. Brands and advertisers employ them to monitor your internet activity to provide users with specific and relevant adverts.Google's move to exclude third-party cookies has flashed heated controversy in the internet advertising business. Some have hailed the action, while others have condemned it (Murthy, 2022).

## 1.4 Motivation of this research

The fast growth in data breaches & misuse, user privacy & the implementation of data privacy and security legislation in many countries are strong indicators of the rising relevance of data privacy concerns, and digital advertising practices are at the centre of this complicated subject. Data privacy concerns how data is obtained and maintained, what it is used for, and whether it is given to a third party (Abrantes, 2021). These data are remnants of internet users' digital actions in most circumstances. Personal privacy in the technological sense is critical even though privacy is still a human right guaranteed by law, whether online or offline. Furthermore, the right to privacy is not null and void (Levy, 2021).

## 1.5 Problems in current solution:

Technology has advanced to become increasingly personal and "intelligent" as consumers have accepted it over time. People have grown to trust technology as a result of this personalisation (Orlova, 2021). They are progressively revealing more of their data without recognising that they are leaving a deeper & larger digital footprint. People still want to utilise technology, although they wouldn't want their data revealed. Privacy security and anonymity are no more only problems for offenders or the wealthy and famous; they touch everyone. There is now a disconnect between consumers' faith in intrusive technology and their comprehension of the quantity of private data those technologies publicly divulge. Although the awareness level of privacy has been increasing in recent years, little is known about how to monitor, regulate, and declassify material that is currently available online (Loney-Howes, 2021).

## 1.6 Research Questions:

RQ1: How much information can be gathered about a person with few details?

RQ2: How can the size & influence of a digital footprint be quantified using multiple performance measures?

RQ3: How 3rd parties cookies work in collecting data for digital footprint?

RQ4: Why does your digital footprint important?

## 1.7 Aim and objectives

This research aims to enlighten and apprise clients about the breadth and depth of their digital footprint and educate how they maximize privacy and anonymity and reduce their digital footprint over the internet.

Here are the main objectives of our research.

* Identifying the amount of sensitive data that is exposed nowadays and then evaluating the effectiveness of that exposure.
* Proposed maximize privacy and anonymity and reduce measures for mitigating digital footprint over the internet and improve overall business functions.
* Sensitivity ranking for different areas of personal information to examine the effect of sensitive data.

## 1.8 Project approach

In this research, respond to the RQ, we must devise a strategy (or select a method) for developing the study design. The research technique specifies gathering enough data to answer the research questions. There are three types of research methods to select from:

(1) qualitative

(2) quantitative

(3) mixed methods.

This project investigate the digital footprints over the internet of a representative sample of the population who use online services and then analyze that data for themes and developments in how layers link to form a digital footprint. Naturally, some people find it simpler to discover information if they have public accounts on Facebook, Instagram and Twitter. Other users will have more difficulty discovering the same information because of the private profile. As a result, unique profiles of the same digital footprint may emerge for different groups, such as millennial, youth, and the middle-aged.

The quantitative research approach has been chosen for this study.

Explain How a digital footprint is used to track you online

Give steps to reduce your digital footprint over the internet

# Chapter 2 Literature Reviews

## Deep analysis of Literature reviews

This chapter starts with a description of the fundamental principles needed to comprehend digital footprints and their origins. To begin, the many sorts of digital footprints are addressed. Following that, several famous digital footprint research are discussed.

, Claims that the privacy of digital footprint material, which has significant commercial value, should be recognized as a semi-fundamental human right. Furthermore, when to use a DNT opt-out or default method is determined by the type of private information involved. The practical consequences indicate a technological balance between digital footprint privacy protection and commercial marketing uses. Authors are one of the earliest to do so since this significant issue is relatively new in application domains, and no fundamental academic study has developed a thorough theoretical legal foundation. Aside from its uniqueness, this research contributes to the field by offering a theoretically feasible technique for both digital footprint privacy protection & marketing gains. research aims to motivate current digital disparities frameworks to incorporate this new dimension.

The literature examination demonstrates how various groups may benefit or suffer due to their digital footprints. A specific focus should be placed on people on the periphery, such as users from low socioeconomic backgrounds. We further analysis on following categories.

## 2.1 Online Personal Information Breach

Information leaked might or might not be harmful, but a vicious attacker may exploit that data over several sites and link it to a single user. Based on research on connecting pseudonyms for social media sites, "in the best-case scenario, an attacker will locate over 60 per cent of the overall user's Facebook account.

## 2.2 Digital footprints using online behavior

Most studies on Internet users do not consider liking, favouriting, following, or commenting on social networking sites to be online content production, although they add to digital footprints. Browsing histories, web searches, buying habits, and geographical information are sensitive data that contribute to digital footprints even when concealed from users. Platform algorithms play a crucial role in the generation of such data (Valanarasu, 2021). They not only push users to disclose personal information by filling out profiles and forms but also create digital footprints from every online action. Notably, due to data mining efforts, such data offer value for social media firms.

Platforms and 3rdparty service providers analyse, identify, and find the meaning of such data for behavioural predictions, monitoring, and advertising. If available and examined with sufficient technologies, digital footprints provide a precise picture of an Internet user. For example, using Facebook Likes, a readily known digital trace, to predict personal variables such as behavioural, ethnic, sexual preference, religious and political beliefs, personality traits, or drug users with high accuracy.

## 2.3 Personal Information types and classes

"Appropriately categorising information is the basis for establishing formation control plan and sharing information," according to research at Beijing University (Shi, 2007). This type of personal information is according to its importance and sensitivity (Norberg, 2007). Value is established by factors such as reliability.

## 2.4 Personal Information Cost

According to a new study, internet privacy practices are not contradictory, but rather the result of various privacy beliefs (Gencoglu, 2015).

# Chapter 3 Methodology

3.1 Research Methodology followed (Qualitative/Quantitative)

3.2 Project Design

3.3 Data collection

# Chapter 4 Data collection and analysis for Digital footprint

4.1 Difference between active and passive footprints

4.2 using 3rd party’s data collection and analysis

4.3 using social media data collection and analysis

4.5 Reduce your Digital footprint

4.6 Anonymity footprints and their challenges

# Chapter 5 Conclusion and future work

5.1 Conclusions

5.2 Future Works

# Project approach

* This project investigate the digital footprints over the internet of a representative sample of the population who use online services and then analyze that data for themes and developments in how layers link to form a digital footprint. Naturally, some people find it simpler to discover information if they have public accounts on Facebook, Instagram and Twitter. Other users will have more difficulty discovering the same information because of the private profile. As a result, unique profiles of the same digital footprint may emerge for different groups, such as millennial, youth, and the middle-aged.
* The quantitative research approach has been chosen for this study.
* Explain How a digital footprint is used to track you online
* Give steps to reduce your digital footprint over the internet

# Project plan

The mainphases of the proposal are:

* Problems identification in the current system
* Deep analysis on literature reviews
* Research on 3rd parties cookies
* research on social media footprint
* Potential impact when digital footprint is used
* write a report with references

# Supervisor Meetings

The following strategies will be used as the primary process for meeting or interacting with the supervisor:

* In-person meetings
* Video conferencing via Zoom
* Email communication while the supervisor is away for a lengthy period.

# Legal, ethical, professional, social issues (mandatory)

**1. Legal and ethical issues**

GDPR

The Computer Misuse Act (CMA) 1990

DPA

Investigatory Powers Act and RIPA-2016

Privacy

**2. Professional and social issues**

* Reputation damage

**Appendix A: Ethics certificate**

**Appendix B: Gantt chart**

The following is the project timetable, milestone, and Gantt chart:



References:

# Bibliography

# Bibliography

Abrantes, B. F. (2021). Digital footprints wrangling–Are analytics used for better or worse?: Evidence of Danish data governance practices on tech-based services.

Arakerimath, A. a. (2015). Digital footprint: Pros, cons, and future. *International Journal of Latest Technology in Engineering 4, no. 10* .

Cheng, F.-C. a. (2018). The do not track mechanism for digital footprint privacy protection in marketing applications. *Journal of Business Economics and Management 19, no. 2*.

Gencoglu, O. H. (2015). Collecting a citizen's digital footprint for health data mining. *In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 7626-7629*.

Janné, M. a. (2022). Effects of employing third-party logistics arrangements in construction projects. *Production planning & control 33, no. 1*.

Levy, Y. a. (2021). Introducing the concept of cybersecurity footprint. *Information & Computer Security*.

Loney-Howes, R. K. (2021). "Digital footprints of# MeToo. *Feminist Media Studies* .

Micheli, M. C. (2018). Digital footprints: an emerging dimension of digital inequality. *Journal of Information, Communication and Ethics in Society*.

Moore, S. C. (2012). Digital Footprints on the Internet. *International Journal of Childbirth Education*.

Murthy, N. a. (2022). Does openness increase vulnerability to digital frauds? Observing social media digital footprints to analyse risk and legal factors for banks. *international Journal of Law and Management ahead-of-print*.

Norberg, P. A. (2007). The Privacy Paradox: Personal Information . *Journal of consumer affairs 41, no. 1*.

Orlova, E. V. (2021). Methodology and models for individuals’ creditworthiness management using digital footprint data and machine learning methods. *Mathematics 9, no. 15*.

Peters, U. (2022). Reclaiming control: Extended mindreading and the tracking of digital footprints. *Social Epistemology 36, no. 3*.

Valanarasu, M. R. (2021). Comparative analysis for personality prediction by digital footprints in social media. *Journal of Information Technology 3, no. 02*.

Ye, Y. (2022). Analysis on the Control of Data Collection, Privacy, and Usage. *In 2021 International Conference on Social Development and Media Communication (SDMC 2021), pp. 997-1000. Atlantis Press*.

Yu, C. T. (2022). Analyzing the structure of tourism destination network based on digital footprints: taking Guilin, China as a case. *Data Technologies and Applications ahead-of-print* .