Creation of an Online Privacy Risk Quantification Tool

Autumn Redpath
u2208993

BSc Cyber Security Undergraduate Thesis

University of Warwick
2025

ABSTRACT

Online privacy risk quantification has been studied in some depth, but existing algorithms lack useable implementations. This project creates a novel algorithm with a corresponding tool, which is verified to be accessible and easy-to-use. This is especially important as one group that is most vulnerable to fraud is those who struggle with technology.

To the best of the author's knowledge, this research is the first to produce the necessary data and tooling to use the algorithm it establishes. Future research should focus on expanding and refining these datasets, using human-based research and real-life data.

PREFACE

This project was supervised by Sandy Taramonli and it sits within the Cyber Security Book of Knowledge (CyBOK) Knowledge Areas of *Risk Management & Governance*[1] and *Privacy & Online Rights*[2].

Some icons, available from `icons.getbootstrap.com`, are copyright © The Bootstrap Authors and licenced under the MIT licence.[3]

Some sections are noted "Beyond Scope", where the content technically goes beyond the initial scope of the project, but is deemed relevant enough to be included.

Thanks

Due to a wide array of personal and external circumstances, I was not able to complete this project to the extent that I set out to. I wish to extend thanks to Sandy Taramonli, WMG's mitigating circumstances panel, and my personal tutors for their support and understanding.

Thanks also to Olga Angelopoulou for initial guidance.

Notes on Typography

Coloured text is used to distinguish types of links; the colours are as follows.

- In-document links, such as footnote markers and cross-references.
- External links, primarily URLs.

[1] Pete Burnap, "Risk Management & Governance", in *The Cyber Security Body of Knowledge v1.1.0, 2021*, Version 1.1.1 (University of Bristol, 2021).

[2] Carmela Troncoso, "Privacy & Online Rights", in *The Cyber Security Body of Knowledge v1.1.0, 2021*, Version 1.0.2 (University of Bristol, 2021).

[3] See `github.com/twbs/icons/blob/main/LICENSE`.

The fonts used within this document are the following.

Main Text: Ubuntu[4]

Monospaced: JetBrains Mono[5]

Maths: Noto Sans Math[6]

Titles: Lato[7]

Finally, the following boxes are used to designate certain content.

> *A block quotation.*

Example 0.0.1

An example of some kind.

Listing 0.1 filename

```
Select code content from file ⟨filename⟩.
```

[4] Canonical Ltd., *Ubuntu* (version 0.83) [TrueType Font], design.ubuntu.com/font.

[5] JetBrains, *JetBrains Mono* (version 2.304) [TrueType Font], www.jetbrains.com/lp/mono.

[6] The Noto Project Authors, *Noto Sans Math* (version 3.0) [TrueType Font], github.com/notofonts/math.

[7] Łukasz Dziedzic, *Lato* (version 2.015) [TrueType Font], www.latofonts.com.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

API endpoint:

a specific URL that refers to an API resource.

Index set:

a set of items that each correspond to an item within a indexed family or indexed set, commonly $\mathbb{Z}^+$.

Indexed family:

a collection of items that can be indexed by some index set.

Indexed set:

a set of items that can be indexed by some index set.

Path:

a string referring to a file's location, locally.

Query paramater:

key-value pairs added on the end of a URL to (commonly) filter items.

ACRONYMS

Application Programming Interface (API):
an interface between two computer programs.

Entity Relationship Diagram (ERD):
aka., “Crow’s-Foot notation”, this is used to display relational database architecture.

Extensible Markup Language (XML):
a commonly-used text format for structured data.

HyperText Markup Language (HTML):
a text-based filetype used to represent webpages.

JavaScript Object Notation (JSON):
a commonly-used text format formed of lists and key-value sets.

Personally Identifiable Information (PII):
information that is related to an identified or identifiable individual (see Article 4(1) of the GDPR).

Uniform Resource Locator (URL):
the address of a specific resource; commonly formed from a protocol, hostname, path, and optional query paramaters.

User Experience (UX):
the overall experience of using a system.

User Interface (UI):
the interface that the user sees and interacts with.

CHAPTER

ONE

INTRODUCTION

> all men suppose that they themselves are altogether without fault or that their errors are few and mild and at great intervals[1]
>
> *The Diagnosis and Cure of the Soul's Passions*
> Galen, c. 165-175 CE

A key way to prevent digital fraud, and other cyber-dependent and cyber-enabled crimes, is by decreasing one's digital footprint.[2] That being said, however, only 69.63 % of people are confident in knowing "how to manage who has access to [their] personal data online".[3] Especially with 93.36 % of people using the internet,[4] understanding this digital footprint – and the risk it carries – is incredibly important.

Giving people a simple yet robust way to quantify this risk is a key step toward allowing them to understand their footprint and, therefore, become safer online. This is also reflected in the "target audience" of this tool – being those who are less "tech-savvy" and those more vulnerable to fraud.

In this Chapter, Section 1.1 explores the data behind the target audience of this tool and Section 1.2 describes the effects of poor online privacy. Finally, Section 1.3 outlines the objectives of this project and Section 1.4 addresses ethical concerns.

[1] Galen, "The Diagnosis and Cure of the Soul's Passions", in *Galen on the Passions and Errors of the Soul*, trans. Paul W. Harkins, with an introduction by Walther Riese (Ohio State University Press, 1963), 27–69, p. 29

[2] Action Fraud, "Protect yourself from fraud and cyber crime", www.actionfraud.police.uk/individual-protection, accessed 13 Nov. 2024; Gov.uk, "Stay safe on social media", stopthinkfraud.campaign.gov.uk/protect-yourself-from-fraud/protecting-against-online-fraud/stay-safe-on-social-media, accessed 13 Nov. 2024.

[3] Ofcom, *Adults' Media Literacy Core Survey 2024* [SPSS Data Tables] (20 Jan. 2025), IN3B.

[4] Ibid., S1.

1.1 Target Audience

As mentioned above, the target audience of this tools is those who are less tech-literate and more vulnerable to cybercrime.

The Crime Survey for England and Wales – as well as information from the National Fraud Intelligence Bureau – provides percentage breakdowns of fraud victims' demographics.[5] This data shows that, on average, women are more likely to be victims of fraud than men,[6] as well as those from mixed ethic groups, those who are separated or divorced, those who are unemployed or economically inactive, and disabled people.

Looking to the *Adults' Media Literacy Core Survey 2024*, one can see that these demographics generally have a lesser understanding of online privacy. For one, women – and especially older women – responded as being less confident in "knowing how to manage who has access to [their] personal data online".[7] Responding to the same question, those who describe their ethic background as non-white are generally less confident, as well as those struggling financially. An overview of these data points is given in Table 1.1.

1.2 Effects of Poor Online Privacy

For the context of this section, "poor online privacy" is taken to mean actions or a mindset that lead to the exposure of sensitive information online. This could be, for example, someone who thinks that they "have nothing to hide" and, therefore, discloses all of their personal information on their social media profile.

As highlighted by Get Safe Online,[8] fraudsters and identity thieves find personal information "invaluable". Additionally, so-called "oversharing" of some aspects of a person's identity can lead to threats, cyberstalking, harassment, and other harmful experiences/offences.[9]

[5] ONS Centre for Crime and Justice, *Appendix tables - Nature of fraud and computer misuse, year ending March 2024* [Spreadsheet], data from the Crime Survey for England and Wales and the National Fraud Intelligence Bureau (Office for National Statistics, 6 Nov. 2024), Table 7.

[6] Note: respondents are classified as only men or women, with apparently no option for non-binary identities.

[7] Ofcom, *Adults' Media Literacy Core Survey 2024*, IN3B.

[8] Get Safe Online, "Oversharing", `www.getsafeonline.org/personal/articles/oversharing`, accessed 26 Jan. 2025.

[9] Get Safe Online, "Online Gender-Based Violence", `www.getsafeonline.org/personal/articles/online-gender-based-violence`, accessed 26 Jan. 2025; Get Safe Online, "Online Abuse", `www.getsafeonline.org/personal/articles/online-abuse`, accessed 26 Jan. 2025.

Table 1.1: An overview of whether select demographics said they were confident in "knowing how to manage who has access to [their] personal data online", compared against the complement[†] of that demographic.[‡]

Demographic	**% Confident**	**% of Complement**	*Δ*
Women	71.01	78.40	7.39
65+ y.o. women	51.49	61.22	9.73
Struggling financially	63.57	78.05	14.48
Not in work/study	60.82	77.29	16.47
Ethnicity: Chinese	57.14	72.51	15.36
Ethnicity: Arab	53.85	72.69	18.84

[†] "Complement", here, means the inverse of the demographic in question – i.e., those who are not a part of the demographic.
[‡] Ofcom, *Adults' Media Literacy Core Survey 2024* [SPSS Data Tables] (20 Jan. 2025), IN3B.

Between June 2023-2024, fraud offences increased by 10 %, and computer misuse, which includes social media hacking, by 70 %.[10] Over the same time period, the Home Office recorded 126 146 instances of online harassment and stalking.[11]

Notwithstanding the emotional costs of experiencing fraud and related offences, it has been found that the cost of fraud against the individual was approximately £8.3 billion in 2023.[12]

1.3 Objectives

This project's goals are as follows, with the ultimate goal being an answer to the research question:

> *To what extent can a tool be created to allow any person to quantify their online privacy risk?*

[10] ONS Centre for Crime and Justice, *Crime in England and Wales, Appendix tables - year ending June 2024* [Data Tables] (Office for National Statistics, 24 Oct. 2024), Table C3.
[11] Ibid., Table C5.
[12] Tim Robinson et al., *Annual Fraud Indicator 2023* (Crowe, 2023), p. 5.

1. To create a single mathematical algorithm to quantify online privacy risk, inspired by existing literature and based on item sensitivity.
2. To create a web-based tool to automate the application of the algorithm on user-provided input.
3. To extend the web-based tool to enable automated input of user data.
4. To test the efficacy of the algorithm and web-based tool.

1.4 Ethical Concerns

Confirmation of waiver of ethical approval can be found in Appendix A.

An initial proposal of this project would have included integration with "web-scraping" tools, to actively search for information that the user had made public. However, due to legal,[13] ethical, and workload concerns, this was dropped.

The project detailed in this report integrates several measures to ensure against ethical breaches.

Firstly, no user data is collected. The application of the algorithm (see Chapter 5) collects no user data and operates entirely client-side. In order to avoid side-channel inference, all data is transferred to the user in one go – not just what they need based on their inputs. Even if such data was collected, it would not constitute personal data as it is not linked to an "identified or identifiable individual."[14] Additionally, while the data could be true to a person, it could also be someone "playing" with the tool. Nevertheless, nothing is collected.

A second concern is that of malicious use. Serious concerns were raised with the earlier proposal, as it could be used as a "doxxing" tool. This is no longer a concern as the "web-scraping" part was dropped. Still, a malicious actor could use the tool to simplify a process of figuring out which target(s) are most vulnerable. This is not a concern either, though, as many other quantification algorithms exist and *OPRiQ* does not offer any novel value to malicious actors.

[13] See Section 3.1 for discussion of the legal landscape.

[14] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC [2016] OJ L119 (GDPR), art. 4(1).

CHAPTER

TWO

METHODOLOGY

2.1 Online Privacy Risk Quantification Algorithm (Objective 1)

The algorithm will be mathematical in nature: as operations performed on sets of values. These operations, brought together, will then produce a value between 0 and 1, with 0 meaning “no risk” and 1 being “maximal risk”. Referring to the latter half of this objective, the reasoning behind decisions made in the algorithm will be based on existing literature, methods, and data. The narrative in Chapter 4 explains this in detail.

There is no identified, viable, alternate approach.

2.2 Web-Based Implementation (Objectives 2-3)

The web-based tool (implementation) will then utilise the Django framework,[1] due to its ease-of-use, database-model features, and extensibility. The frontend will necessarily be written with HTML, CSS, and JavaScript; the Bootstrap framework[2] allows easy customisation of this.

There are many alternate approaches to backend development – each with its own positives and negatives. Overall, the Python language will be used as it is the one with which the author has the most experience. In terms of Python-based frameworks,

[1] Django Software Foundation et al., *Django* (version 5.2) [high-level Python web framework] (2 Apr. 2025), `www.djangoproject.com`.

[2] Bootstrap Team et al., *Bootstrap* (version 5.3.3) [frontend toolkit] (19 May 2025), `getbootstrap.com`.

the author has the most experience with Django; Django also provides very helpful database models[3] and is simple to use/extend.

2.3 Testing (Objective 4)

Testing will be addressed throughout each chapter.

The algorithm can be tested through, first, logical tests outlined in the relevant chapter, and then comparison against other algorithms.

The implementation can be tested via automated accessibility test tools and manual testing.

[3] Django Software Foundation et al., “Django Documentation: Models and databases”, version 5.2, docs.djangoproject.com/en/5.2/topics/db.

CHAPTER

THREE

PRIVACY-FOCUSED LITERATURE

Each section title has respective keywords and key-phrases listed below.

It is also worth noting that the following comes from a primarily Western perspective, and does not reflect the views of the world as a whole. This is especially true in recognising the author's education, and Academic Imperialism as a whole. This topic is not covered here, but readers are directed to the likes of "In What Ways We Depend: Academic Dependency Theory and the Development of East Asian Sociology"[1], "Academic Dependency"[2], and "Academic Dependency and the Global Division of Labour in the Social Sciences"[3].

3.1 The Legal Landscape

Digital Privacy Law | Right to Privacy | GDPR

Before discussing literature in depth, it is important to recognise the legal landscape that applies to online privacy risk as a whole.

[1] Chengpang Lee and Ying Chen, "In What Ways We Depend: Academic Dependency Theory and the Development of East Asian Sociology", *Journal of Historical Sociology* (2022), 1–13. doi: 10.1002/johs.12358.

[2] Fernanda Beigel, "Academic Dependency", *Alternautas*, 2/1 (July 2015), 60–2. doi: 10.31273/alternautas.v2i1.1005.

[3] Syed Farid Alatas, "Academic Dependency and the Global Division of Labour in the Social Sciences", *Current Sociology*, 51/6 (2003), 599–613. doi: 10.1177/00113921030516003.

3.1.1 United Nations

From the perspective of the United Nations, the findings in “Guidelines for the Regulation of Computerized Personal Data Files: final report”[4] – also citing *The right to privacy: conference outcomes*[5] – establish a number of principles in the processing of people’s data. These principles are as follows.[6]

1. lawfulness and fairness,
2. accuracy,
3. purpose-specification,
4. interested-person access,
5. non-discrimination, and
7. security.

The principles are based on the *Universal Declaration of Human Rights*[7] and specifically Article 12:

> *No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.*

3.1.2 Europe

Within the European Union, the primary law relating to online privacy is the *General Data Protection Regulation*,[8] which sets out the rights of persons with regard to data protection (data processing principles). They are the following.

(a) lawfulness, fairness and transparency,

(b) purpose limitation,

(c) data minimisation,

[4] Louis Joinet, “Guidelines for the Regulation of Computerized Personal Data Files: final report”, E/CN.4/Sub.2/1988/22 (July 1988), adopted by Guidelines for the Regulation of Computerized Personnel Data Files (adopted 14 Oct. 1990) A/RES/45/95.

[5] *The right to privacy: conference outcomes* (Stockholm: International Commission of Jurists, 1967).

[6] Joinet, “Guidelines for the Regulation of Computerized Personal Data Files: final report”, Annex I.

[7] Universal Declaration of Human Rights (adopted 10 Dec. 1948) A/RES/3/217 A (UDHR).

[8] GDPR.

(d) accuracy,

(e) storage limitation,

(f) integrity and confidentiality, and

(g) accountability.

This regulation is founded on the *Convention for the Protection of Human Rights and Fundamental Freedoms*,[9] specifically Article 8 ("right to respect for private and family life") and associated case-law. The European Court of Human Rights has long held that digital information falls under the broad scope of Article 8, along with the autonomy and self-determination of that such data.[10] Article 8 is also read to protect information in the public domain.[11]

3.2 What is Privacy?

Defining privacy | The philosophy of privacy | Views on privacy

Moore[12] brings together a wide range of views on privacy into one definition; they write that the right to privacy is "a right to control access to and uses of—places, bodies, and personal information". For brevity, the views that Moore discusses are not repeated here. In the same vein, Bartoletti[13] considers privacy "a great collective good", yet also one that is "culturally constructed" and "[depends] on the time and place". Solove[14] and Negley[15] also write in depth on this topic, though Negley touches more on the philosophy of privacy.

One can also look to statute for a definition of privacy, as discussed above in Section 3.1. Again, this is not repeated in the name of brevity.[16]

[9] Convention for the Protection of Human Rights and Fundamental Freedoms (opened for signature 4 Nov. 1950, entered into force 3 Sept. 1953) 213 UNTS 221 (ECHR), though it should be noted that this convention applies more broadly than the GDPR, as it applies to the Council of Europe as a whole.

[10] See *L B v Hungary* (App no 36345/16) ECHR 2023, [103].

[11] *Satakunnan Markkinapörssi Oy And Satamedia Oy v Finland* (App no 931/13) ECHR 2017, [134].

[12] Adam D. Moore, "Defining Privacy", *Journal of Social Philosophy*, 39/3 (2008), 411–28 at p. 421.

[13] Ivana Bartoletti, *An Artificial Revolution: On Power, Politics and AI* (The Indigo Press, 2020).

[14] Daniel J. Solove, *Understanding Privacy*, Legal Studies Research Paper No. 420, Public Law Research Paper No. 420 (The George Washington University Law School, May 2008).

[15] Glenn Negley, "Philosophical Views on the Value of Privacy", *Law and Contemporary Problems*, 31 (1966), 319.

[16] Cf. Adrienn Lukács, *What is Privacy?: The History and Definition of Privacy* (University of Szeged), section 3; Janice Richardson, *Law and the Philosophy of Privacy* (27 Aug. 2015), chapter 2.

3.3 Quantifying Risk

Quantifying Risk | Calculating Risk

The following is separated by sector, as it is important to recognise the context in which the algorithms apply.

3.3.1 Business Risk

From a capital perspective, Sahinoglu[17] quantifies "final risk" as the product of vulnerability, threat, lack-of-countermeasures, and criticality. Further multiplying this by the capital cost gives the excepted cost of loss. This multiplicative effect is common in both qualitative and quantitative risk assessment methods.[18]

Meyer[19] brings together many project management risk metrics, mainly in terms of risk to timeframe, cost, and performance. They summarise the different measures into the following groups.

- Heuristic-based: based on expert opinion and/or pre-determined guidelines.
- Expected value: multiplying probability and cost.
- Probability distribution: fitting results to pre-defined distributions.
- Mathematical modelling.
- Interdependency: based on logical and contained dependencies.
- Empirical methods: based on learning from historical data.

3.3.2 Investment, Banking, and Trading

Perhaps one of the most well-known measures of risk is one's credit score: a measure of how trusted a person can be with money lent to them.[20] The FICO® score, currently at version 10, is the industry standard algorithm but, unfortunately for this report, it is almost entirely secret.[21]

[17] M. Sahinoglu, "Security meter: a practical decision-tree model to quantify risk", *IEEE Security & Privacy*, 3/3 (2005), 18–24. doi: 10.1109/MSP.2005.81.

[18] See, for example, Risk & Resilience Team, *Risk & Issues Register Template & Guidance* (University of Warwick, [13 June 2023]).

[19] Werner G. Meyer, "Quantifying risk: measuring the invisible", in *PMI® Global Congress 2015-EMEA*, Project Management Institute (London, England, 2015).

[20] Rosie Hamilton, "What is a good credit score?", Money Saving Expert (27 Nov. 2023), www.moneysavingexpert.com/credit-cards/what-is-a-good-credit-score.

[21] Salome Tabagari, "Credit scoring by logistic regression", Master's Thesis (University of Tartu, 2015).

3.4 Quantifying Online Privacy Risk

Quantifying Privacy Risk | Quantifying Online Privacy

Table 3.1 summarises certain characteristics of existing online privacy risk quantification algorithms; the following then provides a synopsis of each study.

Table 3.1: A comparison of existing online privacy risk quantification algorithms.

Title	***Re***	***Au***	***Po***	***Se***
Metrics for Privacy Assessment When Sharing Information in Online Social Networks	●	●	?⃝	?⃝
Network-aware privacy risk estimation in online social networks	●	●	●	●
A Framework for Computing the Privacy Scores of Users in Online Social Networks	●	○	◐	◐
Privacy Scoring of Social Network Users as a Service	○	●	●	○
Privometer: Privacy protection in social networks	◐	●	○	●
Social network privacy measurement and simulation	●	●	●	●
SONET: A SOcial NETwork Model for Privacy Monitoring and Ranking	◐	●	●	●
Privacy Scoring of Social Network User Profiles Through Risk Analysis	●	○	◐	◐
Towards quantifying and defining privacy metrics for online users	●	○	○	●

○ No ◐ Partially ?⃝ Implied that the user will ● Yes

Re Whether the algorithm considers an item's "reachability" or visibility.

Au Whether the algorithm considers the user's "audience" or the people who can view the user's item(s).

Po Whether the algorithm considers the user's own "policy" or tolerance.

IS Whether the algorithm considers the item's "sensitivity".

3.4.1 Metrics for Privacy Assessment When Sharing Information in Online Social Networks

Alemany et al.[22] propose two metrics to be used in tandem: Reachability and Audience, as defined below.

Reachability is defined as the probability of a user's message being seen by some proportion of the users a certain "distance" from them. This "distance" is based on the user's network – being how many "steps" or "jumps" another user is away from them. For example, the users that a person follows are considered to be a distance of 0 away.

Their audience metric uses the same ideas as above, and is defined as the proportion of users who are likely to see a user's message.

Alemany et al. argue that these metrics allow users to gauge the risk of posting a message, by comparing the values against one's internal preferences. It is also for the user to determine the requisite values, mainly being based on tractability and personal risk tolerance.

3.4.2 Network-aware privacy risk estimation in online social networks

Similar to the well-known "Pagerank" algorithm,[23] Pensa, Di Blasi, and Bioglio[24] consider the privacy scores of neighbours in computing one's own score.

The "intrinsic" privacy risk of some user is defined based on the sensitivity and visibility of information that they have shared. This score also considers the "distance" (see above) that a user is willing to share any one piece of information. Their "network-aware" privacy score is then defined as a distribution that satisfies their conditions, with a damping factor taken into account.

As with the previous study, this score carries issues of tractability – a common prob-

[22] Jose Alemany et al., "Metrics for Privacy Assessment When Sharing Information in Online Social Networks", *IEEE Access*, 7 (2019), 143631–45. doi: 10.1109/ACCESS.2019.2944723.

[23] Sergey Brin and Lawrence Page, "The anatomy of a large-scale hypertextual Web search engine", *Computer Networks and ISDN Systems*, 30/1 (1998), Proceedings of the Seventh International World Wide Web Conference, 107–17. doi: 10.1016/S0169-7552(98)00110-X.

[24] Ruggero G. Pensa, Gianpiero Di Blasi, and Livio Bioglio, "Network-aware privacy risk estimation in online social networks", *Social Network Analysis and Mining*, 9/1 (Apr. 2019). doi: 10.1007/s13278-019-0558-x.

lem in network-aware algorithms.[25] Calculating one's "true" risk score would require also calculating those of the near 8.2 billion people on earth,[26] and then fitting these to the distribution above.

3.4.3 A Framework for Computing the Privacy Scores of Users in Online Social Networks

Liu and Terzi[27] establish a privacy score based on the sensitivity and visibility of items within a network.

Each user in the network is considered to have a "profile" of some number of items, each of which have a "privacy level." This privacy level, similarly to above, represents the number of "steps" away from the user that they are willing to share the item.

Liu and Terzi provide two ways for computing sensitivities – a naïve method and an Item Response Theory (IRT) method. They then ran surveys to generate data to test these algorithms – comparing users' scores and attitudes across different regions.

3.4.4 Privometer: Privacy protection in social networks

Talukder et al.[28] develop a tool for Facebook that they call "Privometer", utilising item inference and the privacy scores of a user's friends.

Their model considers a malicious user who knows the attributes of each friend of a person. The model then employs inference functions to "guess" the attributes of the user. Privometer then suggests "self-sanitisation" actions – recommendations as to attributes the user's friends should hide.

Talukder et al. claim that Privometer is implemented in Facebook, via the *Facebook PHP Client API.*

[25] See, i.e., discussion in Anne Bouillard, "Trade-off between accuracy and tractability of Network Calculus in FIFO networks", *Performance Evaluation*, 153 (2022), 102250. doi: 10.1016/j.peva.2021.102250; S. Samaranayake, S. Blandin, and A. Bayen, "A tractable class of algorithms for reliable routing in stochastic networks", *Transportation Research Part C: Emerging Technologies*, 20/1 (2012), 199–217. doi: 10.1016/j.trc.2011.05.009.

[26] United Nations, *World Population Prospects 2024: Summary of Results* (UN DESA/POP/2024/TR/NO. 9, New York: Department of Economic and Social Affairs, Population Division, 2024), p. VII.

[27] Kun Liu and Evimaria Terzi, "A Framework for Computing the Privacy Scores of Users in Online Social Networks", *ACM Trans. Knowl. Discov. Data*, 5/1 (Dec. 2010). doi: 10.1145/1870096.1870102.

[28] Nilothpal Talukder et al., "Privometer: Privacy protection in social networks", in *2010 IEEE 26th International Conference on Data Engineering Workshops (ICDEW 2010)* (2010), 266–9. doi: 10.1109/ICDEW.2010.5452715.

3.4.5 SONET: A SOcial NETwork Model for Privacy Monitoring and Ranking

Nepali and Wang,[29] in their "SONET" model, construct a graph of connected users who all have their own attributes.

Their "privacy index" $PIDX$ is a percentage value based on the "privacy impact factors" of a user's known attributes. They then define privacy invasion as when $PIDX \geq T$ for some user-defined threshold T; otherwise, privacy is preserved.

Extending this $PIDX$ to consider visibility is done by adding the visibility measurement set. They do not define a concrete algorithm for computing this, but they do mention the use of "depth-first or breath-first traversals".

Nepali and Wang go on to explain how their model could be created from search engine data, and data from social networking sites directly. They also discuss the possibility of utilising data in deep-web sites. The author finds it pertinent to mention that utilising data from these sources may be in contravention of the law and ethical standards.

3.4.6 Social network privacy measurement and simulation

Wang, Nepali, and Nikolai[30] extend SONET with their "composite PIDX" function $c-PIDX$. This addition considers separation between users, in a similar way to other literature discussed.

3.4.7 Privacy Scoring of Social Network User Profiles Through Risk Analysis

Taking a Privacy Risk Analysis (PRA) approach, De and Imine[31] construct harm trees based on item inference and visibility.

[29] Raj Kumar Nepali and Yong Wang, "SONET: A SOcial NETwork Model for Privacy Monitoring and Ranking", in *2013 IEEE 33rd International Conference on Distributed Computing Systems Workshops* (2013), 162–6. doi: 10.1109/ICDCSW.2013.49.

[30] Yong Wang, Raj Kumar Nepali, and Jason Nikolai, "Social network privacy measurement and simulation", in *2014 International Conference on Computing, Networking and Communications (ICNC)* (2014), 802–6. doi: 10.1109/ICCNC.2014.6785440.

[31] Sourya Joyee De and Abdessamad Imine, "Privacy Scoring of Social Network User Profiles Through Risk Analysis", in *Risks and Security of Internet and Systems*, ed. Nora Cuppens et al. (Springer International Publishing, 2018), 227–43.

In their system, the root node of a tree represents a specific privacy harm, leaves are personal data attributes (name, gender, etc.), and connecting nodes are specific threats (such as unintended inference). These nodes are then connected by "AND", "OR", or "*k*-out-of-*n*" decisions. For example, in their Figure 4, De and Imine state that a user's birth date and year infer their date of birth. In a less explicit example, a user's date of birth, home address, and phone number all being revealed can lead to identity theft/fraud, they claim.

To decrease the size of these trees, they provide "pruning" methods based on item visibility and accuracy.

De and Imine claim that this mechanism could then be used as the base for a system to inform users of their privacy scores.

3.4.8 Towards quantifying and defining privacy metrics for online users

Also looking at visibility and sensitivity, similar to Nepali and Wang's *PIDX*, Blauw and Solms[32] calculate scores based on multiplying the these elements as well as the quantity. For sensitivity s, quantity q, and visibility v, the score would be sqv.

While the calculation is relatively simplistic, their view on visibility does seem unique. They extend the layers of "WWW"–"deep web"–"dark web" to their visibility metric, considering the amount of authentication required to access.

3.5 Gaps in Literature

To finish this chapter, it is important to identify the gaps in existing literature; that is what this section does.

As shown, there is a breath of research in the broad area of online privacy risk quantification. This research includes methods such as the visibility and sensitivity of attributes, the "audience" one reveals these attributes to, and a person's own comfort level. Some literature looks at the user in isolation, whereas other authors take into account the user's social network.

There are two main gaps in these studies, which appear to stem from the same cause.

[32] Frans F Blauw and Sebastiaan von Solms, "Towards quantifying and defining privacy metrics for online users", in *2017 IST-Africa Week Conference (IST-Africa)* (2017), 1–9. doi: 10.23919/ISTAFRICA.2017.8102366.

The first is that, with the exception of Talukder et al.,[33] no algorithms are actually implemented. Furthermore, Talukder et al. do *claim* that their algorithm is implemented on Facebook, but this implementation is not provided.

Secondly, the algorithms are quite theoretical and take an amount of mathematical knowledge to understand. For example, some algorithms above require knowledge of matrices, some network-based techniques, and others probability and associated theory. While this may not hinder implementation, it does mean that the average user may not understand why their score is “high” or “low” and how to change that.

These both likely stem from the fact that these algorithms originate from academic sources. Targetted at academics, these algorithms do not focus on being understandable by a lay-person or having an associated implementation. This is the gap that this project seeks to fill.

[33] Talukder et al., “Privometer: Privacy protection in social networks”.

CHAPTER

FOUR

CREATING THE ALGORITHM

This chapter establishes a novel online privacy risk quantification algorithm *OPRiQ*.

4.1 Algorithm Architecture

OPRiQ has two distinct parts:

1. Item sensitivity, and
2. Item inference.

Aside: Visibility

While other algorithms consider the "visibility" or "reachability" of an item, *OPRiQ* does not. The reason behind this decision is similar to that of the "motivated intruder" test of UK data protection law.

First mentioned in the case of *R (on the application of Department of Health) v Information Commissioner*,[1] then expanded on in *Information Commissioner v Magherafelt District Council*,[2] the motivated intruder test considers a person with no prior knowledge who will take all reasonable steps to attempt to identify an individual.[3] In the context of data protection, anonymised information is still considered "personal information" if a motivated intruder could identify, or de-anonymise, the individual.[4]

[1] *R (on the application of Department of Health) v Information Commissioner* [2011] EWHC 1430 (Admin).

[2] *Information Commissioner v Magherafelt District Council* [2012] UKUT 263 (AAC).

[3] See also *Information Commissioner v Miller* [2018] UKUT 229 (AAC).

[4] GDPR, Article 4(1).

Table 4.1: Notation used within *OPRiQ*.

Notation	**Explanation**
i	An item of information, such as a name.
s_i	The base "sensitivity" of an item i.
j_i	The user's value for item i.
$m_{i=j}$	The modifier to be applied to s_i if i is some value j.
$s'_{i=j}$	The modified sensitivity of i, s.t., $s'_{i=j} = s_i \cdot \left(1 + m_{i=j}\right)$.
r_i	Whether i is truly revealed – aka., "visible" or "public".
$\{i_0, i_1, \dots, i_n\} \rightarrow i$	The item i can be inferred from the items $\{i_0, i_1, \dots, i_n\}$.
$\{i_0, i_1, \dots, i_n\} \xrightarrow{p} i$	The item i can be inferred, with likelihood p, from the items $\{i_0, i_1, \dots, i_n\}$.
r'_i	The probability that item i can be considered revealed.
I	The set of all items $i \in I$.
J	The indexed family of the user's values of each corresponding item in I, s.t. $\forall i \in I,\ \exists j_i \in J$.
S	The indexed family of sensitivities, indexed by I, s.t. $\forall i,\ \exists s_i \in S$.
M	The indexed family of modifiers, indexed by I, J, s.t. $\forall i, j,\ \exists m_{i=j} \in M$.
R	The indexed family representing whether the user has truly revealed an item, indexed by I, s.t. $\forall i,\ \exists r_i \in R$.
IR	The set of inference tuples (A, i, p), where $A \xrightarrow{p} i$.

This is not to say there is no place for visibility in privacy risk assessments, but *OPRiQ* considers the worst-case scenario. This reflects the target use of the algorithm – being one to educate people and, especially, those who may be targets of fraud. Though fraudsters do not have unlimited finances, they are not low on finances either (see Section 1.2), and they are indeed motivated.

4.1.1 Sensitivity

"Item sensitivity" is used to refer to the value of a piece of information to an arbitrary attacker.

Example 4.1.1

A person's name is more sensitive than their age, but less sensitive than their home address.

This relatively simple model, as adopted in previous literature, may be easy to implement but does not consider how the value of an item can be societally reliant.

Example 4.1.2

A person's sexuality may be a sensitive piece of information itself, but a person being gay is, arguably, more sensitive than being straight.

Example 4.1.3

The sensitivity of one's political views would rely on the society in which they live – agreeing with democracy may be a normal view today, but a highly sensitive one under a dictatorship.

Within *OPRiQ*, base item sensitivity is expressed as a number $s_i \in [0, 1]$, for some item i. Where the item has no value to a malicious actor, $s_i = 0$; if the item is maximally valuable, $s_i = 1$. Then, a multiplier $m_{i=j} \in [-1, 1]$ can be applied for when the item i is the value j. The resultant sensitivity is expressed as $s'_{i=j}$, equal to $s_i \cdot (1 + m_{i=j})$.

Example 4.1.4

Suppose item a has sensitivity $s_a = .5$ and multipliers exist for values b, c such that $m_{a=b} = -.5$ and $m_{a=c} = .5$. The resultant scores would be $s'_{a=b} = .25$ and $s'_{a=c} = .75$.

Note that it is possible for some $s'_{i=j}$ to be greater than 1. This is acceptable, however, as, in such case, the sensitivity must be great and in need of special recognition.

4.1.2 Inference

As somewhat shown in De and Imine[5]'s harm trees, item inference refers to how certain information can be inferred from other items. This inference can be either total or probabilistic.

Example 4.1.5

Someone's birthday and age would reveal their date of birth.

Example 4.1.6

A person's name can imply their gender – though not certainly.

In *OPRiQ*, each item i is considered "truly revealed" if and only if r_i; items where $\overline{r_i}$. may be revealed through inference, however. Assume that some item i can be inferred from some set of items $\{i_0, i_1, \dots, i_n\}$, it is then true that $\{i_0, i_1, \dots, i_n\} \rightarrow i$. If the inference is probabilistic, for some probability $p \in (0, 1)$, then $\{i_0, i_1, \dots, i_n\} \xrightarrow{p} i$. The overall probability p of an item being considered revealed is notated as r'_i.

For where multiple sets $\Gamma_0, \Gamma_1, \dots, \Gamma_n$ all satisfy $\Gamma \xrightarrow{p} i$ for the same i, r'_i takes the maximal value of p; this can be expressed as below.

$$r'_i = \max\left(p_0 : \Gamma_0 \xrightarrow{p_0} i, p_1 : \Gamma_1 \xrightarrow{p_1} i, \dots, p_n : \Gamma_n \xrightarrow{p_n} i\right) \tag{4.1}$$

Finally, to clarify, $r_i \rightarrow r'_i$ but $r'_i \nrightarrow r_i$.

4.2 Sensitivity Database

Any implementation of *OPRiQ* necessarily requires a "sensitivity database" to provide the values for indexed families S and M.

[5] De and Imine, "Privacy Scoring of Social Network User Profiles Through Risk Analysis".

Example 4.2.1

If a user can select "name" as revealed, there *must* be a corresponding entry s_{name}.

Example 4.2.2

If a user can select a categorical item – such as age – then there *should* be entries in M for each category; for example $m_{\text{age=65}}$.

The key question, though, is how to determine these values. The following sets out various sources for, and justifications of, the values decided on.

Wording in the following

For most decisions, items (or sets thereof) were given a "base value" which could then be added to or subtracted from to arrive at the actual value of s_i. This "base value" is not to be confused with s_i, or any other values within *OPRiQ*, and is used solely for this establishment of sensitivities.

Example 4.2.3

Assume some set of items is given the base score .5, but some research suggests that an item in this set should be considered more sensitive. This item's score could then be the base score plus some modifier – i.e., .5 + .1 = .6.

4.2.1 Legal Perspective

For one, the *General Data Protection Regulation*[6] enumerates the following as "special categories" of Personally Identifiable Information (PII).

- Data revealing:
    - racial or ethnic origin,
    - political opinions,
    - religious or philosophical beliefs, or
    - trade union membership.

[6] GDPR, art. 9(1).

- Genetic data or biometric data for the purpose of identification.
- Data concerning a person's:
  - health,
  - sex life, or
  - sexual orientation.

As they are quite broad, but still recognised as sensitive, each of these sets are given a base value of .6.

PII "relating to criminal convictions and offences or related security measures"[7] is given even stricter protection. Therefore, these sets are given the base value .7.

Secondly, the *EA 2010*[8] establishes nine "protected characteristics"; these are below. Similarly to above, these are given a base of .6, with room for modification later.

- age,
- disability,
- gender reassignment,
- marriage and civil partnership,
- pregnancy and maternity,
- race,
- religion or belief,
- sex, and
- sexual orientation.

[7] GDPR, art. 10.

[8] Equality Act 2010, § 4.

Table 4.2: Dark web prices of certain information, collected from various sources.

Item	Cost
Credit Card with CVV	5 $ to 110 $
Cryptocurrency Account	20 $ to 2,650 $
Bank Account	200 $ to 1,000 $
Social Account Details	1 $ to 300 $
Name, Address, Email	5 $ to 15 $
Full Identity Profile	20 $ to 100 $
Medical Records	500 $
Passport Scan	5 $ to 125 $

4.2.2 Criminal Perspective

One can also look to the actual cost of PII to criminals. Some research has been done on this topic – collating dark-web prices of information;[9] the relevant data is given in Table 4.2.

Additionally, one could consider what information a malevolent actor needs in order to scam a person. As put by Andy,[10] the key information for scammers is as follows.

- Full name,
- Social security number,
- Date and place of birth,

[9] Miklos Zoltan, "Dark Web Price Index 2023", Privacy Affairs (23 Apr. 2023), www.privacyaffairs.com/dark-web-price-index-2023; Barry Elad, "Dark Web Statistics 2024", Enterprise Apps Today (19 Feb. 2024), www.enterpriseappstoday.com/stats/dark-web-statistics.html; Trustwave, "How Prices are Set on the Dark Web: Exploring the Economics of Cybercrime" (25 Nov. 2024), www.trustwave.com/en-us/resources/blogs/trustwave-blog/how-prices-are-set-on-the-dark-web-exploring-the-economics-of-cybercrime; Paul Bischoff, "Passports on the dark web: how much is yours worth?", Comparitech (19 Oct. 2018), www.comparitech.com/blog/vpn-privacy/passports-on-the-dark-web-how-much-is-yours-worth.

[10] Andy, "What Information Does A Scammer Need?", We Get Scammed For You (12 Dec. 2023), wegetscammedforyou.com/what-information-does-a-scammer-need.

- Bank account details,
- Bank account PINs,
- Card details (expiration dates and CVVs),
- Addresses,
- Phone numbers,
- ID numbers (i.e., passport numbers), and
- Affiliations, memberships, and job information.

Each item given by, or within those given by, Andy[11] is given a base score of .75 – recognising the immense value to criminals. The more costly items in Table 4.2 are then given a modification of +.1 to +.2, in recognition of their additional value. Items which form only part of one given above – such as one's given name being part of their full name – are modified by –.2.

4.2.3 Using Previous Literature

Finally, one can look to other academic and professional literature.

Guide to Protecting the Confidentiality of Personally Identifiable Information

McCallister, Grance, and Scarfone's *Guide to Protecting the Confidentiality of Personally Identifiable Information*[12] enumerates certain types of PII. Those not mentioned above are given below.

- IP and MAC addresses,
- Personal photos,
- Voice data,
- Vehicle registrations,

[11] Andy, *What Information Does A Scammer Need?*

[12] Erika McCallister, Tim Grance, and Karen Scarfone, *Guide to Protecting the Confidentiality of Personally Identifiable Information*, Special Publication 800-122 (National Institute of Standards and Technology, Apr. 2010).

- Place of birth,
- Weight,
- Education data, and
- Financial information.

These categories are quite broad and varied; for example, financial information could lead to all kinds of fraud, while disclosure of one's weight carries very little risk (though, for some, it may be a very personal disclosure). Taking this into account, the majority of these are given a base value of .7 with modifications of −.2 to +.2, with the exception of weight which is given the value .1.

Handbook for Safeguarding Sensitive PII

Similarly, U.S. DHL's *Handbook for Safeguarding Sensitive PII*[13] lists examples of PII. Again, those not mentioned yet are given below.

- Immigration/citizenship status,
- Passwords, and
- Security question information.

These are all given a sensitivity of .9, recognising their immense value.

What is personal data?

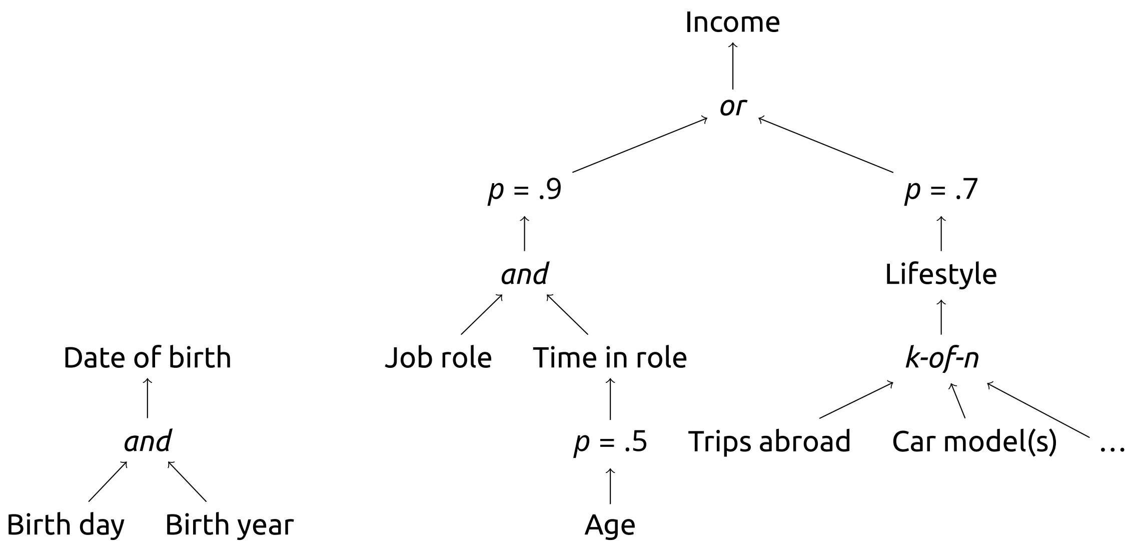
The ICO[14] guide "*What is personal data?*"[15] also gives examples of PII; as above, new items are given below.

- Location data,
- Usernames, and
- Pseudonyms.

[13] U.S. Department of Homeland Security, *Handbook for Safeguarding Sensitive PII*, Revision 3, Privacy Policy Directive 047-01-007 (4 Dec. 2017).

[14] Note that the ICO's function is to, among other things, interpret the Data Protection Act 2018 and GDPR (as retained in UK law by virtue of European Union (Withdrawal) Act 2018, § 3); therefore, much of their guidance echoes work already cited.

[15] Information Commissioner's Office, "What is personal data?", `ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/personal-information-what-is-it/what-is-personal-data/what-is-personal-data`, accessed 11 May 2025.



(a) A simple absolute inference.

(b) A more complex, and probabilistic, inference; probabilities are approximate.

Figure 4.1: Example inference trees for a person's attributes.

4.3 Inference Trees

As mentioned prior, certain items can be "inferred" from sets of others; this inference can be absolute or probabilistic and can be modelled through so-called "inference trees". Take, for example, the trees in Fig. 4.1 which show how certain information leads to the inference of other information. While these trees could be connected and displayed as a network of directional graphs, this is not done here for simplicity.

In reality, and as explained in Section 4.1.2, these trees are expressed as sets $\{i_0, i_1, \dots, i_n\} \xrightarrow{p} i$, where $i_0, i_1, \dots, i_n, i \in I$ and $0 \leq p \leq 1$ is the probability of inference. The purpose of displaying them as trees here is to aid in visualisation; this is in line with the project's goals.

As with Section 4.2, the bulk of this section is on determining these inference tuples.

4.3.1 Absolute Inferences

Some inferences are absolute – notated prior as $\{i_0, i_1, \dots, i_n\} \rightarrow i$. An example of this is how someone's day and year of birth, together, reveal their date of birth. These absolute inferences, based on the items identified in Section 4.2, are displayed as trees in Fig. 4.2.

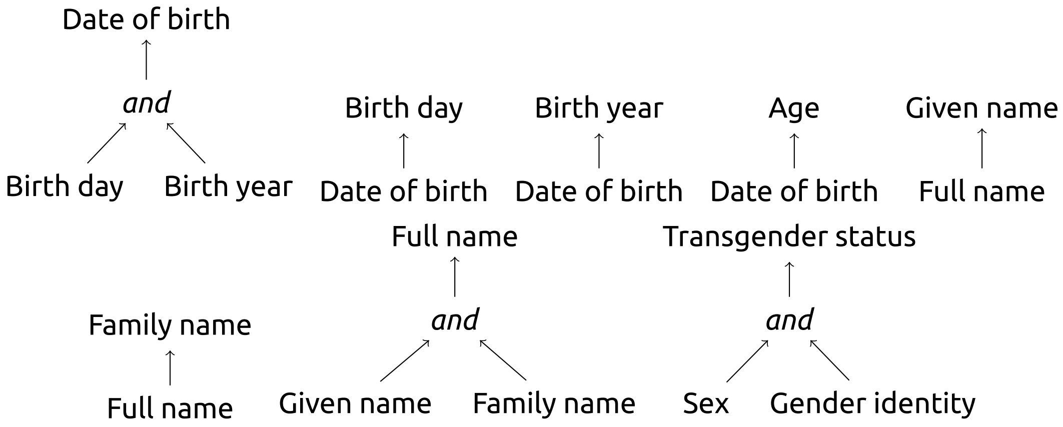


Figure 4.2: Absolute inference trees for a person's attributes.

4.3.2 Probabilistic Inferences

The remainder of inferences are probabilistic, and are given in Fig. 4.3. Some of these are not worth discussion, as they are relatively simplistic ($\{\text{age}\} \overset{.9}{\to}$ birth year, for example). Others, however, do deserve justification.

Firstly, one's IP address can, in some cases, reveal their address or at least their approximate location.[16] Secondly, gender identity and sex both infer each other with a certainty of 93.5 %, following the 2021 census data.[17]

4.4 Modifiers

The final set of values to determine is M – the set of modifiers; in determining these, each item in Section 4.2 is considered.

Names

Full name, given name, family name

There are no specific name values that would justify a modification of sensitivity.

[16] See generally Ovidiu Dan, Vaibhav Parikh, and Brian D. Davison, "IP Geolocation through Reverse DNS", *ACM Trans. Internet Technol.* 22/1 (Oct. 2021). doi: 10.1145/3457611; Christopher Luna, "How accurate is IP geolocation?", MaxMind (1 July 2021), blog.maxmind.com/2021/07/how-accurate-is-ip-geolocation.

[17] Office for National Statistics, "Gender identity, England and Wales: Census 2021", statistical bulletin (6 Jan. 2023), www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/genderidentity/bulletins/genderidentityenglandandwales/census2021.

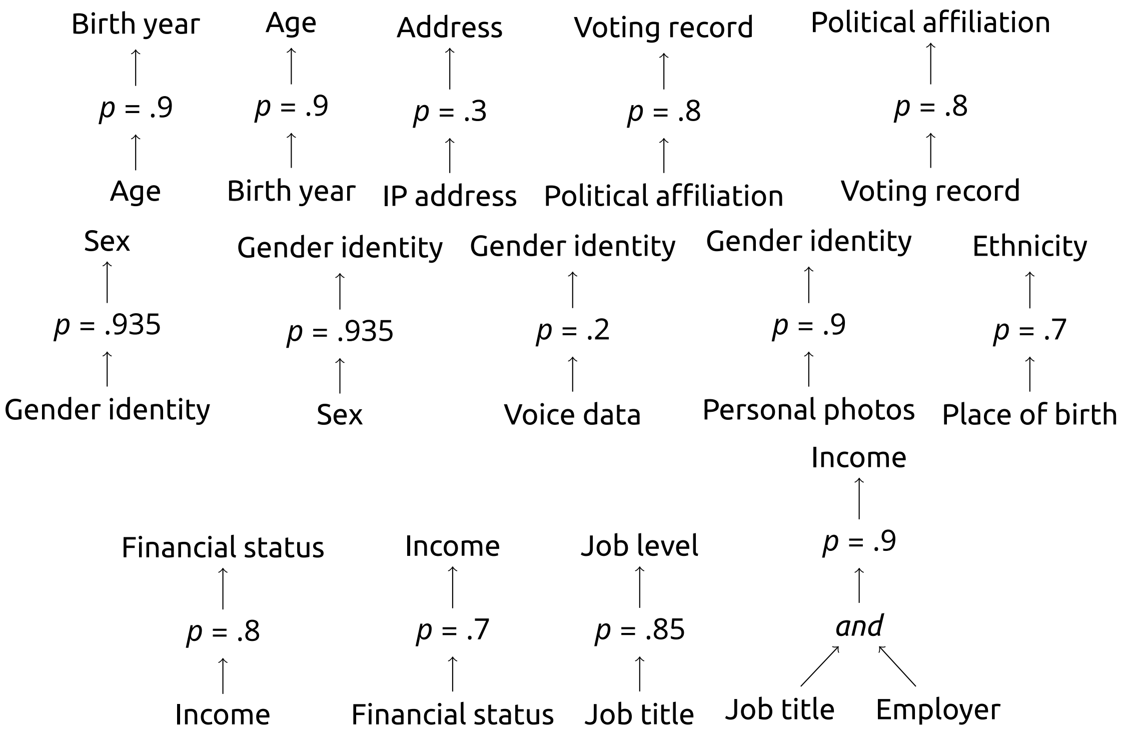


Figure 4.3: Probabilistic inference trees for a person's attributes.

Dates

Date of birth, birth day, year of birth

As above, there are no birthday-related values that would justify a modification of sensitivity.

Protected Characteristics &c.

Age, disability, transgender status, marriage/civil partnership, race, ethnicity, religion/faith, sex, gender identity, sexuality

ONS data referred to prior[18] gives percentage breakdowns of fraud victims by personal characteristics. To convert these percentages to modifications, one can follow Eq. (4.2) – taking the specific percentage as a multiple of the baseline, then modifying it to fit approximately $-1 \leq m \leq 1$. For these, let $P(i = j)$ be the proportion of victims of fraud with characteristic $i = j$ (i.e., gender = male); $P(*)$ refers to the proportion of all people.

[18] ONS Centre for Crime and Justice, *Appendix tables - Nature of fraud and computer misuse, year ending March 2024*.

$$m_{i=j} = \frac{P(i=j)}{P(*)} - 1 \tag{4.2}$$

Thankfully, the proportions $\frac{P(i=j)}{P(*)}$, for most i, j in the dataset, fit approximately $[0, 2]$. This means that the only adjustment is subtracting one, and no multiplication is needed.

These final m values for the characteristics in the ONS data are given in Table C.1.

Addresses

Address, email address, IP address, MAC address, place of birth, location data

If an IP address is registered to a VPN supplier,[19] it loses most of its value to a threat actor. Therefore, an appropriate value for $m_{\text{IP=VPN}}$ would be $-.9$, as the data still has some – albeit now minimal – value.

Financial Data

Bank details, bank card details, bank card expiration date, bank card cvv, financial status

Financial status on the extremes can certainly draw attention of scammers. For example, a person in need of money may be more vulnerable to monetary scams; similarly, someone with a large disposable income may be seen as a high-value target. Therefore, both conditions are given a modifier of .5.

Other than financial status, there are no specific values for these items that would justify a modification of sensitivity.

Work Data

Job title, employer, job level, income, education

Job title and education are covered by the ONS data discussed previously; one's job level is also somewhat covered by virtue of the former. One's income can carry some value in the eyes of threat actors – in deciding how to target someone, for example – but this does not provide for discernable modifications.

Unlike these, a person's employer certainly carries value. It is difficult to produce a list of all employers who could be seen as valuable to a malicious actor, though, so *OPRiQ* leaves this to the user with $m_{\text{employer=high-value target}} = .5$.

[19] See, i.e., Mathew Heard et al., *VPN & datacenter IPs* (14 May 2025), github.com/X4BNet/lists_vpn.

Societal

Club membership, voting record, political affiliation, immigration/citizenship status, sex life, criminal record, national/government ID number, vehicle registration number, health data

Somewhat similarly to “employer,” above, modifiers for these items rely on broad classifications that are left to the user to apply.

Firstly, the following are given modifiers of “unpopular” (.2), “deeply unpopular” (.35) and “persecuted” (.5).

- Club membership,
- Voting record, and
- Political affiliation.

“Criminal record” is given modifiers of “minimal” (–.5), “moderate” (.1), and “major” (.5). Finally, “health data” is given modifiers of “somewhat revealing” (.4) and “majorly revealing” (.6).

The remainder do not have any specific values that would justify a modifier.

Other

Password, security question information, username/pseudonym, phone number, personal photo, voice data/recordings, weight

There are no specific values for these items that would justify a modification of sensitivity.

4.5 Bringing It All Together

OPRiQ, taking into account the above, can be summarised as the function in Eq. (4.3).

$$\mathit{OPRiQ} : (I, J, S, M, R, IR) \mapsto [0, 1] \tag{4.3}$$

This can then be expanded into that in Eq. (4.4), using the parts discussed above and having calculated r_i' for each $i \in I$ prior. Eq. (4.4) also requires some combinatory functions α and β.

$$OPRiQ = \alpha\left[\beta\left(s'_{i=j_i}, r'_i\right) \mid i \in \mathcal{I}\right] \tag{4.4}$$

Now the question comes of defining α and β.

α: the Outer Combination Function

The definition of α depends on the desired "shape" of the online privacy risk score graph. For a proper definition of α, let Λ be the indexed family of values passed to the function such that $\forall\lambda \in \Lambda$, $\lambda \in [0, 1]$. $\overline{\Lambda}$ is therefore the mean value of Λ s.t. $\overline{\Lambda} = \frac{\sum\Lambda}{|\Lambda|}$.

Using $\overline{\Lambda}$, the graphs in Fig. 4.4 show different "shapes" that the function α could take, with the caveat that the equations in the legend are simplified. Simply adding the individual scores (the line $\alpha(\Lambda) = \overline{\Lambda}$ in Fig. 4.4) is common in literature,[20]. The use of entropy is also common, to represent the uncertainty present in privacy metrics.[21] As *OPRiQ* is supposed to be understandable to a lay-person, it does not use entropy functions.

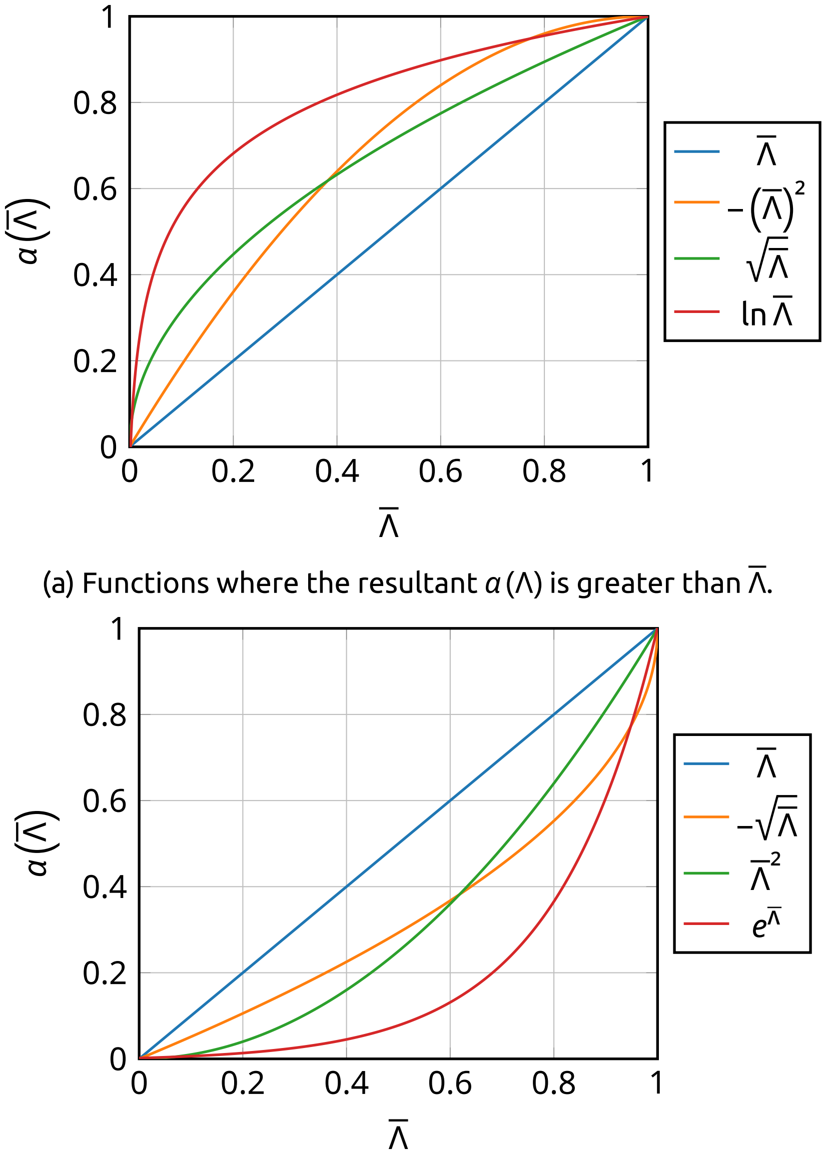
This all being said, *OPRiQ* takes a logarithmic approach ($\ln\overline{\Lambda}$ in Fig. 4.4a). The primary reason for this is that privacy risk intuitively raises quickly, then slows, with respect to the amount of information known. In the eyes of an attacker, the increase in value from no information to a small amount of information is great. The respective increase from a large amount of information to even more is not as great.

Example 4.5.1

In executing a phishing attack, an attacker could target a person's place of work or club that they attend, among other things. Knowing both a person's place of work and club does not lend much utility to this attack, as opposed to just knowing one.

[20] See Mishtu Banerjee et al., "Quantifying Privacy Violations", in *Secure Data Management*, ed. Willem Jonker and Milan Petković (Berlin, Heidelberg: Springer Berlin Heidelberg, 2011), 1–17. doi: 10.1007/978-3-642-23556-6_1; Manar Alohaly and Hassan Takabi, "Better Privacy Indicators: A New Approach to Quantification of Privacy Policies", in *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)* (Denver, CO: USENIX Association, June 2016).

[21] See Isabel Wagner and David Eckhoff, "Technical Privacy Metrics: A Systematic Survey", *ACM Comput. Surv.* 51/3 (June 2018). doi: 10.1145/3168389; Razi Arshad and Muhammad Rizwan Asghar, "Characterisation and Quantification of User Privacy: Key Challenges, Regulations, and Future Directions", to be published in *IEEE Communications Surveys & Tutorials* (2024). doi: 10.1109/COMST.2024.3519861.



(a) Functions where the resultant $\alpha(\Lambda)$ is greater than $\overline{\Lambda}$.

(b) The respective inverse functions of each in Fig. 4.4a.

Figure 4.4: Example functions for $\alpha(\Lambda)$, based on $\overline{\Lambda}$; legend equations simplified.

As an aside: this behaviour can be expressed as in Eqs. (4.5) and (4.6).

$$\left.\frac{\mathrm{d}\alpha}{\mathrm{d}\overline{\Lambda}}\right|_{\overline{\Lambda}=i} > \left.\frac{\mathrm{d}\alpha}{\mathrm{d}\overline{\Lambda}}\right|_{\overline{\Lambda}=j} .\ \forall i < j \tag{4.5}$$

$$\frac{\mathrm{d}^2\alpha}{\mathrm{d}\overline{\Lambda}^2} < 0.\ \forall\overline{\Lambda} \tag{4.6}$$

As previously mentioned in relation to Fig. 4.4, the actual equations are more complex than those in the legend. In this case, $\alpha = \ln\overline{\Lambda}$ requires four separate constants; α can then be expressed as follows.

$$\alpha = a \cdot \ln\left(b \cdot \overline{\Lambda} + c\right) + d \tag{4.7}$$

In order to decide on these constants' values, some preliminary conditions can be recalled and set. Firstly, *OPRiQ* – and therefore α – must evaluate to a number between 0 and 1; this can be expressed as follows.

$$\alpha\left(\overline{\Lambda}\right) = \begin{cases} 0 : & \overline{\Lambda} = 0 \\ n : & \overline{\Lambda} = 1 \end{cases} \tag{4.8}$$

Eq. (4.8) can be alternately expressed as the following.

$$0 = a \cdot \ln(b \cdot 0 + c) + d \tag{4.9}$$
$$0 = a \cdot \ln(c) + d \tag{4.10}$$
$$-d = a \cdot \ln(c) \tag{4.11}$$

$$1 = a \cdot \ln(b \cdot 1 + c) + d \tag{4.12}$$
$$1 = a \cdot \ln(b + c) + d \tag{4.13}$$
$$1 - d = a \cdot \ln(b + c) \tag{4.14}$$

$$1 + a \cdot \ln(c) = a \cdot \ln(b + c) \tag{4.15}$$
$$\frac{1}{a} + \ln(c) = \ln(b + c) \tag{4.16}$$

As $\ln(n)$ is only defined for $n > 0$, $c > 0$ and $b + c > 0$. In reality, b is not necessary as its modifications are handled by a, c, and d. Representing $\frac{1}{a}$ as n leaves the following.

$$n + \ln(c) = \ln(1 + c) \tag{4.17}$$
$$e^{n+\ln(c)} = 1 + c \tag{4.18}$$
$$e^n \cdot e^{\ln(c)} = 1 + c \tag{4.19}$$
$$e^n \cdot c = 1 + c \tag{4.20}$$
$$e^n \cdot c - c = 1 \tag{4.21}$$
$$c(e^n - 1) = 1 \tag{4.22}$$
$$c = \frac{1}{e^n - 1} \tag{4.23}$$

Then, recalling Eq. (4.11), one can use this c to evaluate d.

$$-d = n \cdot \ln(c) \tag{4.24}$$
$$d = -\frac{1}{n} \cdot \ln\left(\frac{1}{e^n - 1}\right) \tag{4.25}$$
$$d = -\frac{1}{n} \cdot -\ln(e^n - 1) \tag{4.26}$$
$$d = \frac{\ln(e^n - 1)}{n} \tag{4.27}$$

This leaves a as the following.

$$a = \frac{1}{n} \cdot \left[\ln\left(\overline{\Lambda} + \frac{1}{e^n - 1}\right) + \ln(e^n - 1)\right] \tag{4.28}$$
$$a = \frac{1}{n} \cdot \ln\left[(e^n - 1) \cdot \left(\overline{\Lambda} + \frac{1}{e^n - 1}\right)\right] \tag{4.29}$$
$$a = \frac{1}{n} \cdot \ln\left[(e^n - 1) \cdot \overline{\Lambda} + 1\right] \tag{4.30}$$

Finally, n must be determined. Figure 4.5 shows how a would look with various values for n; as shown, n does not change the intercepts (0, 0) or (1, 1), but solely the “shape” of the graph.

$n = 3$ was chosen for *OPRiQ*, to maintain the idea of risk rising quickly, then slowing, but not to a degree where most users have very high scores. This final a is displayed in Fig. 4.6.

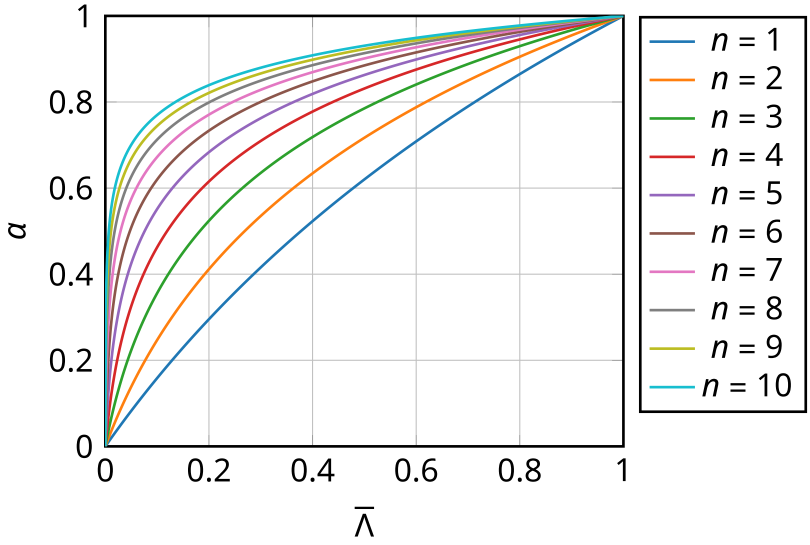


Figure 4.5: Graphs for α, as n varies.

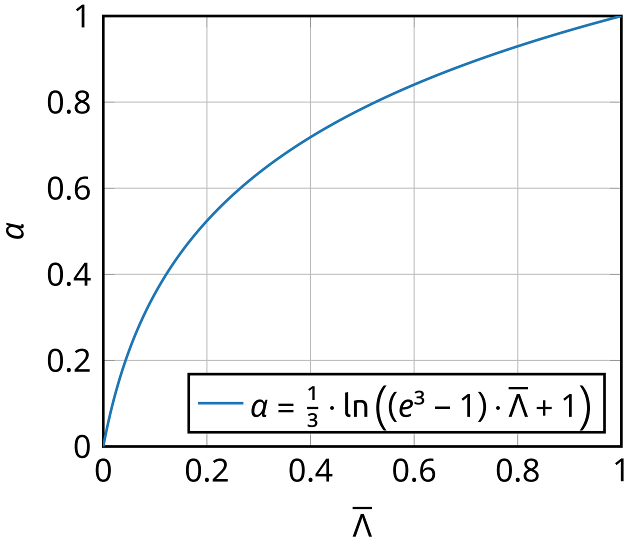


Figure 4.6: The final graph for α and, therefore, *OPRiQ*.

Derivative Proof

Beyond Scope

As an aside, it can be proven that this α satisfies Eqs. (4.5) and (4.6) mentioned earlier. This proof is given in Appendix H.

β: the Inner Combination Function

As mentioned above, $\beta(s', r') \in [0, 1]$ for all s', r'. Further, and as with α, deciding the "shape" of β is key to defining β itself; the key difference from α is that β takes two distinct arguments.

Some preliminary conditions of β are those in Eq. (4.31); the real question, though, is that of $0 < r' < 1$.

$$\beta(s', r') = \begin{cases} 0 : & r' = 0 \\ s' : & r' = 1 \end{cases} \tag{4.31}$$

One approach would be to multiply both arguments, such that $\beta(s', r') = s' \cdot r'$. While this may appear simplistic, it may be the most appropriate.

As repeated prior, one aim of *OPRiQ* is to be understandable by a lay-person; keeping its respective parts as simplistic as possible is key to this aim. Therefore, taking into account the relative complexity of α above, having $\beta(s', r') = s' \cdot r'$ seems appropriate. Further, this multiplicative effect is common in existing practice.[22]

To conclude on β:

$$\beta(s', r') = s' \cdot r' \tag{4.32}$$

Using α and β

Therefore, taking into account the above and noting that it is possible for some $s'_{i=j}$ to be greater than 1, *OPRiQ* can be expressed as follows.

$$OPRiQ = \left[\frac{1}{3} \cdot \ln\left(\left(e^3 - 1\right) \cdot \frac{\sum_{i \in I} \left(s'_{i=j_i} \cdot r'_i\right)}{|I|} + 1\right)\right]_0^1 \tag{4.33}$$

[22] See Fahad Usmani, "What is Expected Monetary Value in Project Management?", PM Study Circle (27 Sept. 2024), pmstudycircle.com/expected-monetary-value-emv.

This may appear confusing to some, though, so it can be simplified to that in Eq. (4.34), with the caveat that it is a simplified version.

$$OPRiQ \sim \ln\left[\frac{\sum_{i\in I}(s'_{i=j_i}\cdot r'_i)}{|I|}\right] \tag{4.34}$$

4.6 Testing

To test the algorithm, its output can be compared against other similar calculations.

4.6.1 Test Data

To aid comparison, several test datasets of R were created; they are as follows.

- An "empty" set, where $\overline{r_i}.\ \forall i \in I$ – referred to as $D_{\varnothing}$.
- A "full" set, where $r_i.\ \forall i \in I$ – referred to as D_*.
- Five "random" sets, generated by the script in Listing G.1 – referred to as D_n for $n \in \{1, 2, 3, 4, 5\}$.

As no prior research developed datasets similar to S, M, or IR, the values created prior are used for all comparisons.

4.6.2 Algorithm Comparisons

For the purpose of this section, algorithms are "compatible" with *OPRiQ* if their output can be compared in some way.

Those Incompatible

Firstly, Alemany et al.[23] produce both reachability and audience metrics – not a risk metric compatible with *OPRiQ*. Secondly, Pensa, Di Blasi, and Bioglio[24]'s scores re-

[23] Alemany et al., "Metrics for Privacy Assessment When Sharing Information in Online Social Networks".

[24] Pensa, Di Blasi, and Bioglio, "Network-aware privacy risk estimation in online social networks".

quire a populated user graph to calculate and are, therefore, incompatible. Further, Talukder et al.[25] provide no implementation for their “Privometer Reading”.

Those Compatible

Liu and Terzi[26]’s dichotomous *PR* can be compared with *OPRiQ*. The polytomous case, and *Pr_IRT*, are not compatible with *OPRiQ*, as they require further data.

For reference, *PR* is defined as in Eq. (4.35), for some user j; for comparison, $V(i, j)$ – the visibility function – is taken to return 1 if the item is “revealed” and 0 otherwise. This is inline with an example given by the authors, and is comparable *OPRiQ*’s *R*. Making this substitution and converting sensitivity to the same notation as in *OPRiQ* gives the equation in Eq. (4.36).

$$PR = \sum_{i=1}^{n} \beta_i \times V(i, j) \tag{4.35}$$

$$PR' = \sum_{i \in I} s_i \times r_i \tag{4.36}$$

The calculation of Nepali and Wang[27]’s privacy index *PIDX* is identical to *PR* above – the sum of revealed items’ sensitivities.

While Wang, Nepali, and Nikolai[28] do extend *SONET*’s *PIDX*, the resultant calculation does not change in a way that would make comparison any different.

Similarly to the above, Blauw and Solms[29]’s score multiplies sensitivity, quantity, and visibility. In this context, this becomes a sum of all sensitivities of revealed items.

Overall, these four compatible algorithms all – in essence – result in the sum of sensitivities.

4.6.3 Results

Figures 4.7a and 4.7b show the differences between *OPRiQ* and other scores discussed above. While the differences in Fig. 4.7a may appear drastic, this is by design.

As discussed in the design, *OPRiQ* captures the fact that risk increases quickly, then

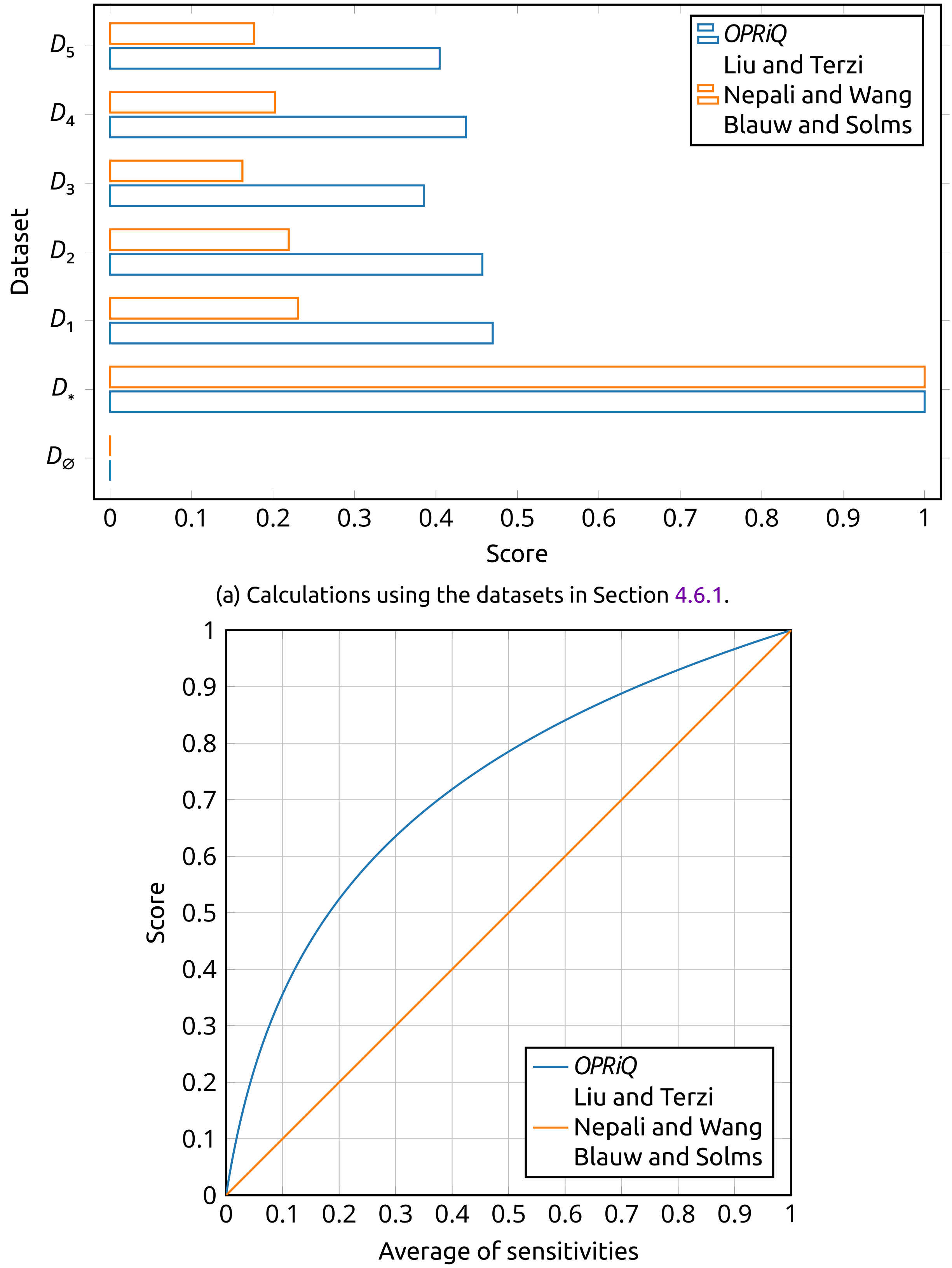
[25] Talukder et al., “Privometer: Privacy protection in social networks”.
[26] Liu and Terzi, “A Framework for Computing the Privacy Scores of Users in Online Social Networks”.
[27] Nepali and Wang, “SONET: A SOcial NETwork Model for Privacy Monitoring and Ranking”.
[28] Wang, Nepali, and Nikolai, “Social network privacy measurement and simulation”.
[29] Blauw and Solms, “Towards quantifying and defining privacy metrics for online users”.

slows. Other research has incorporated logarithmic functions,[30] but not in a way that is compatible with *OPRiQ*. As further discussed above, the algorithms that can be compared with *OPRiQ* – being the majority of previous research – all take a linear approach. Again, this is where *OPRiQ* diverges from the majority.

Overall, there is not much critique to draw from comparison, as most other literature does not apply logarithmic techniques.

The practical implications of these results are that *OPRiQ* scores, in the vast majority of cases, will be greater than any other algorithm. However, it is unlikely that users will be comparing scores in the first place, as *OPRiQ* appears to be the first algorithm to have an actual implementation. Furthermore, a logarithmic perspective is more realistic and intuitive than linear, as discussed prior.

[30] See, i.e., Liu and Terzi, "A Framework for Computing the Privacy Scores of Users in Online Social Networks", Section 5.2.



(a) Calculations using the datasets in Section 4.6.1.

(b) Calculations in relation to the average of revealed items' sensitivities.

Figure 4.7: Other algorithms compared with *OPRiQ*, referred to by author and scaled to [0, 1].

CHAPTER

FIVE

WEB-BASED IMPLEMENTATION

This chapter creates a web-based implementation of *OPRiQ*.

5.1 Design

The implementation can be easily split into the "backend" and "frontend".

5.1.1 Backend

The "backend" of a website receives requests and returns responses; this can be split into API endpoints, the database, and routing.

API Endpoints

Beyond Scope

The use of API endpoints here is to allow for future extensibility of the website. These API endpoints (as set out in Appendix D) will facilitate access to the database, for – among other things – the sets of values I, S, M, and IR. In these, path-based arguments are used to select individual items (including whole sets), and query paramaters are used to filter such sets.

Database

The database, in this case, refers to the relational database that stores the sets of values I, S, M, and IR. An ERD for the intended design of this database is given in Fig. 5.1.

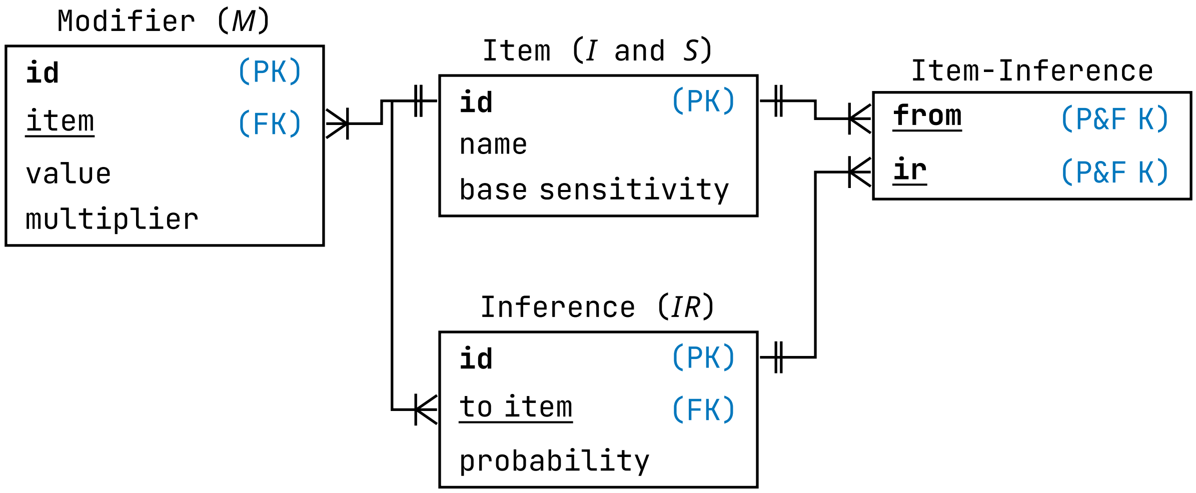


Figure 5.1: An ERD describing the planned database table layout.

The tables “`Item`” and “`Modifier`” are quite self-explanatory – mapping directly to the sets S (and I) and M, respectively. The table “`Item-Inference`” exists to resolve the many-to-many relationship between I and IR, with the values of IR then being split between “`Item-Inference`” and “`Inference`”. This many-to-many resolution then brings the database design in line with Codd’s first “normal form”.[1] The intended design is also compliant with the second “normal form”, but not the third.[2] This is due to the use of “IDs” within all tables, bar “`Item-Inference`”, for convenience, following industry-norms.[3]

Routing

As the site will be relatively simplistic, the key three paths are as follows.

/ The home page, where the user can calculate their score.

`/about` A page describing how the algorithm works.

`/api/` API endpoints referred to prior.

[1] E. F. Codd, “A Relational Model of Data for Large Shared Data Banks”, *Communications of the ACM* 13/6 (June 1970). doi: `10.1145/362384.362685`, section 1.4.

[2] As given by E. F. Codd, “Further Normalization of the Data Base Relational Model”, *Communications of the ACM* RJ909 (Aug. 1971).

[3] See, i.e., Django Software Foundation et al., “Django Documentation: Models”, version 5.2, `docs.djangoproject.com/en/5.2/topics/db/models`; for points on this debate, see Stack Overflow users, “How do you like your primary keys?” (19 Jan. 2017), `stackoverflow.com/questions/404040/how-do-you-like-your-primary-keys`, accessed 9 May 2025.

5.1.2 Frontend

The “frontend” refers to what the user receives and sees; this is split into User Interface (UI) and User Experience (UX), application of the algorithm (calculation), and accessibility.

UI & UX

UI and UX refer generally to how the user interacts with a system. In line with the project’s goals, this tool should prioritise ease-of-use above all. Following BS EN ISO 9241-11:2018[4], which defines usability, and its constituent parts, as follows, the design should focus on being effective, efficient, and satisfactory.

> Usability focuses on the effectiveness, efficiency and satisfaction of the user’s interaction with the object of interest. […]
>
> Effectiveness is the accuracy and completeness with which users achieve specified goals. […]
>
> Efficiency is the resources used in relation to the results achieved. […]
>
> Satisfaction is the extent to which the user’s physical, cognitive and emotional responses that result from use of a system, product or service meet user’s needs and expectations. […]

To ensure that the tool is effective, the algorithm’s implementation should be proper and use the correct values. These values should reflect those in the database – being updated where necessary.

For efficiency, the user should be able to calculate their risk with the fewest “clicks”, redirections, and other interactions possible.[5] To achieve this, the following should be used.

- The score should update automatically, when the user updates their selection.
- The “homepage” should host the calculation, such that the user need not navigate through the site to perform the calculation.
- All items should be displayed as buttons, so the user only has to click on them to select them.

[4] BS EN ISO 9241-11:2018, *Ergonomics of human-system interaction: Part 11 – Usability: Definitions and concepts* (British Standards Institute) (henceforth BS EN ISO 9241-11:2018).

[5] Noting BS EN ISO 9241-11:2018, 6.3.2–6.3.3.

- Only the appropriate modifiers should be shown, as to not confuse the user.

Satisfaction is generally covered by the above; an “about” page should also be provided to give the user more information about the algorithm.

Calculation

To ensure trust and privacy, the calculation must take place client-side and transfer no data to the server. Therefore, the values of all sets must be transferred to the client, and *OPRiQ* be implemented in JavaScript.

Accessibility

The accessibility of any implementation can be assessed against *WCAG 2.2*;[6] this is implementation-specific, so isn’t discussed further here.

5.2 Implementation

Now that the website has been designed, it can be implemented. Django[7] was used for this implementation, due to its ease of use and the author’s familiarity. The remainder of this section covers the separate parts of the implementation, split into backend and frontend.

5.2.1 Backend

API Endpoints

Beyond Scope

While API endpoints are not explicitly within the scope of this project, they do allow for future development and extensibility.

API endpoints are simple to implement with Django’s model framework (more on this below). These API endpoints retrieve the relevant data from the database and return it in the JavaScript Object Notation (JSON) formats given in Appendix D. While this

[6] *Web Content Accessibility Guidelines*, ed. Alastair Campbell et al., version 2.2 (W3C Recommendation; World Wide Web Consortium, 5 Oct. 2023), `www.w3.org/TR/WCAG22` (henceforth WCAG 2.2).

[7] Django Software Foundation et al., *Django*.

does require a database transaction for each request, it allows for the most up-to-date information to be returned every time. This could be optimised through a form of “hook” that updates a cached value in line with the database, but this is considered far out of scope.

The MathML[8] endpoint, being an outlier, simply returns the Extensible Markup Language (XML) file shown in Listing F.1.

Database

Django’s model framework[9] allows for an abstracted view of database tables. For example, take the definitions in Listing 5.2 below, where the columns are defined as fields within the class. These models then provide abstracted access to the underlying data, as shown in Listing 5.1. It is worth noting that the table “`Item-Inference`” is not explicitly defined, as the many-to-many relationship is handled by Django.

Listing 5.1 `opriq/views.py`

```python
def index(request):
    return render(request, "home.html.django", context={
        "items": Item.objects.all(),
        "mods": Modifier.objects.all(),
        "inferences": Inference.objects.all()
    })
```

Listing 5.2 `opriq/models.py`

```python
class Item(models.Model):
    id = models.AutoField(primary_key=True)
    name = models.CharField(max_length=128, unique=True, null=False,
     ↪  blank=False)
    base_sensitivity = models.FloatField()

    [...]


class Inference(models.Model):
    id = models.AutoField(primary_key=True)
    from_items = models.ManyToManyField(Item,
     ↪  related_name='inferences_coming_from')
```

[8] See BS ISO/IEC 40314:2016, *Mathematical Markup Language (MathML) Version 3.0 2nd Edition* (British Standards Institute) (henceforth BS ISO/IEC 40314:2016).

[9] Django Software Foundation et al., *Django Documentation*.

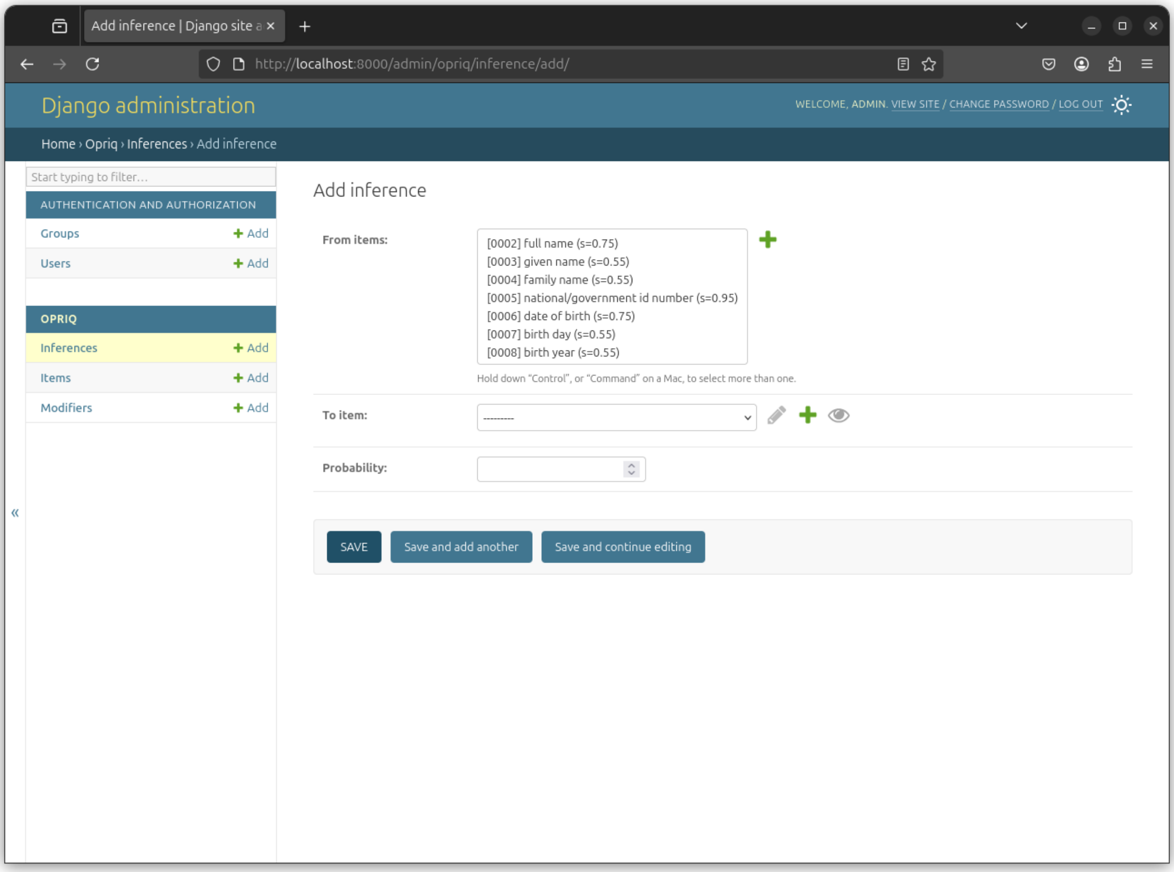
```
    to_item = models.ForeignKey(Item, on_delete=models.CASCADE,
     ↪  related_name='inferences_pointing_to')
    probability = models.FloatField()

    [...]


class Modifier(models.Model):
    id = models.AutoField(primary_key=True)
    item = models.ForeignKey(Item, on_delete=models.CASCADE)
    value = models.CharField(max_length=128, null=False, blank=False)
    multiplier = models.FloatField()

    [...]
```

Registering these models with the Django admin system[10] allows for simple inputting of values, as shown in Fig. 5.2.


Add inference | Django site a ×
+
Django administration
WELCOME, ADMIN. VIEW SITE / CHANGE PASSWORD / LOG OUT
Home › Opriq › Inferences › Add inference
Start typing to filter...
AUTHENTICATION AND AUTHORIZATION
Groups + Add
Users + Add
OPRIQ
Inferences + Add
Items + Add
Modifiers + Add
«
SAVE
Save and add another
Save and continue editing

http://localhost:8000/admin/opriq/inference/add/

Add inference

From items:
[0002] full name (s=0.75)
[0003] given name (s=0.55)
[0004] family name (s=0.55)
[0005] national/government id number (s=0.95)
[0006] date of birth (s=0.75)
[0007] birth day (s=0.55)
[0008] birth year (s=0.55)

Hold down "Control", or "Command" on a Mac, to select more than one.

To item: ---------

Probability:

Figure 5.2: Example usage of the admin interface to add an inference entry.

[10] Django Software Foundation et al., "Django Documentation: The Django admin site", version 5.2, docs.djangoproject.com/en/5.2/ref/contrib/admin.

Routing

Routing is simplistic with Django; as shown in Listing 5.3, a route is formed from a path string and view function. An example of a view function is given in Listing 5.1, which returns a rendered HyperText Markup Language (HTML) file for the main page.

Listing 5.3 opriq/urls.py

```
urlpatterns = [
    path('admin/', admin.site.urls),
    path('', index, name='index'),
    path('about', about, name='about'),
    [...]
]
```

5.2.2 Frontend

UI & UX

Use of the Bootstrap framework[11] simplifies UI design. As noted in the design section, the main page is that which hosts the calculation.

To allow for further development, the values are loaded into JavaScript objects; this could be extended in the future to use the API endpoints. These are then used, with HTML templates,[12] to build the UI. An example template is given in Listing 5.4.

Listing 5.4 templates/home.html.django

```
<div class="container-fluid" id="items"></div>
<template id="item-template">
    <div class="form-check-inline py-2">
        <input class="btn-check" type="checkbox">
        <label class="btn btn-outline-primary"></label>
    </div>
</template>
```

A simple header and footer, titles, and a “theme switcher”[13] are included to finalise

[11] Bootstrap Team et al., *Bootstrap*.

[12] Mozilla et al., “<template>: The Content Template element”, developer.mozilla.org/en-US/docs/Web/HTML/Reference/Elements/template, accessed 23 May 2025.

[13] Bootstrap Team, *JavaScript theme switcher* [Color mode toggler for Bootstrap’s docs], getbootstrap.com/docs/5.3/customize/color-modes.

the UI. With all of this, the main website page is shown in Fig. 5.3a. Using the same base, an "about" page is also available, as shown in Fig. 5.3b.

Calculation

Calculation is handled by two JavaScript functions: "`handle_item_click(id)`" and "`calculate()`".

`handle_item_click` is run each time that an item is clicked. Primarily, it toggles the corresponding entry in "`item_checked`". It then updates the modifiers and inferences to show only those that are relevant. Finally, it runs "`calculate()`".

`calculate` updates the user's score by running calculating *OPRiQ* based on their inputs. It first establishes a running sum of revealed items, along with any selected modifications, and then adds the inferred items. The final score is then calculated using this running sum, and displayed to the user as a percentage, rounded to two decimal places. For clarity, this final stage is shown in Listing 5.5.

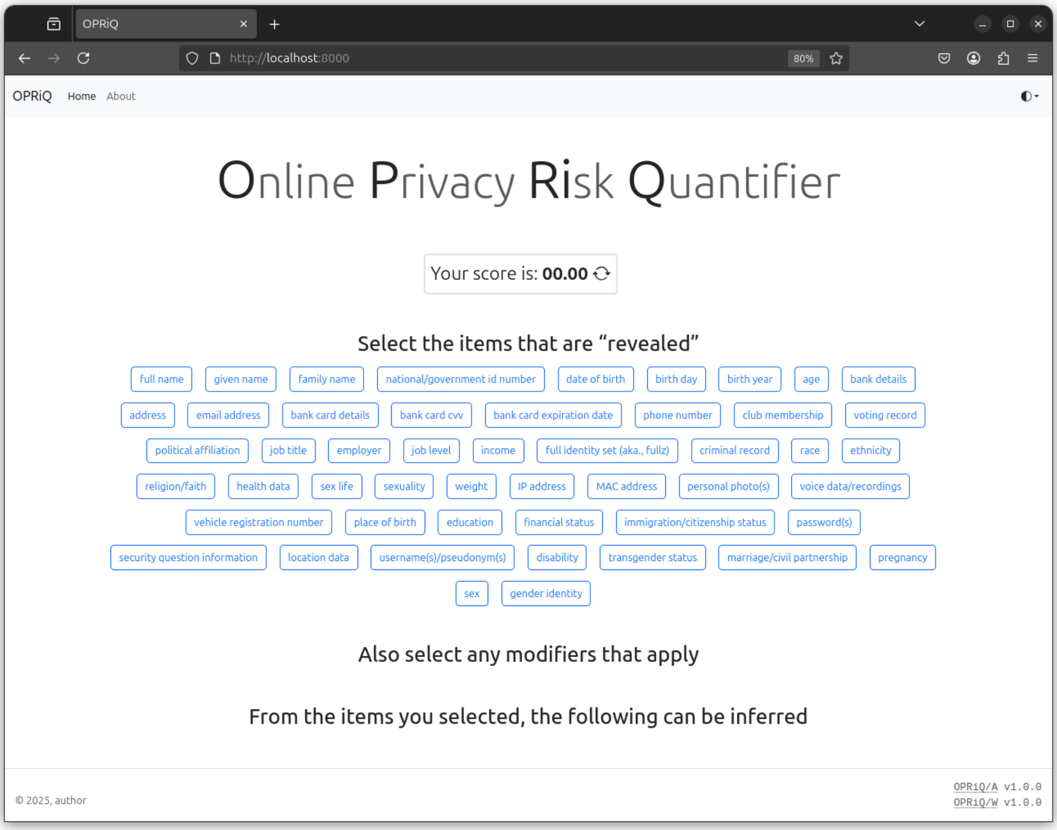
Listing 5.5 templates/home.html.django

```
let final_score = (
    Math.log((Math.exp(3) - 1) * (running_sum / num_items) + 1)
) / 3;
document.getElementById("score").innerText = (final_score *
 ↪  100).toFixed(2);
```

Accessibility

Use of *AInspector*,[14] shown in Fig. 5.4, shows that the site only has one violation of WCAG 2.2. This violation is that the "`h1`" (heading) text does not contain part of the "`title`" element's content. However, this is due to the site's title being the acronym "OPRiQ", which is expanded in the heading. Therefore, the site is in compliance with WCAG 2.2.

[14] Jon Gunderson, *AInspector* (version 3.1) [firefox extension for analysing accessibility compliance] (23 May 2025), `ainspector.disability.illinois.edu`.



Online Privacy Risk Quantifier

Your score is: **00.00**

Select the items that are “revealed”

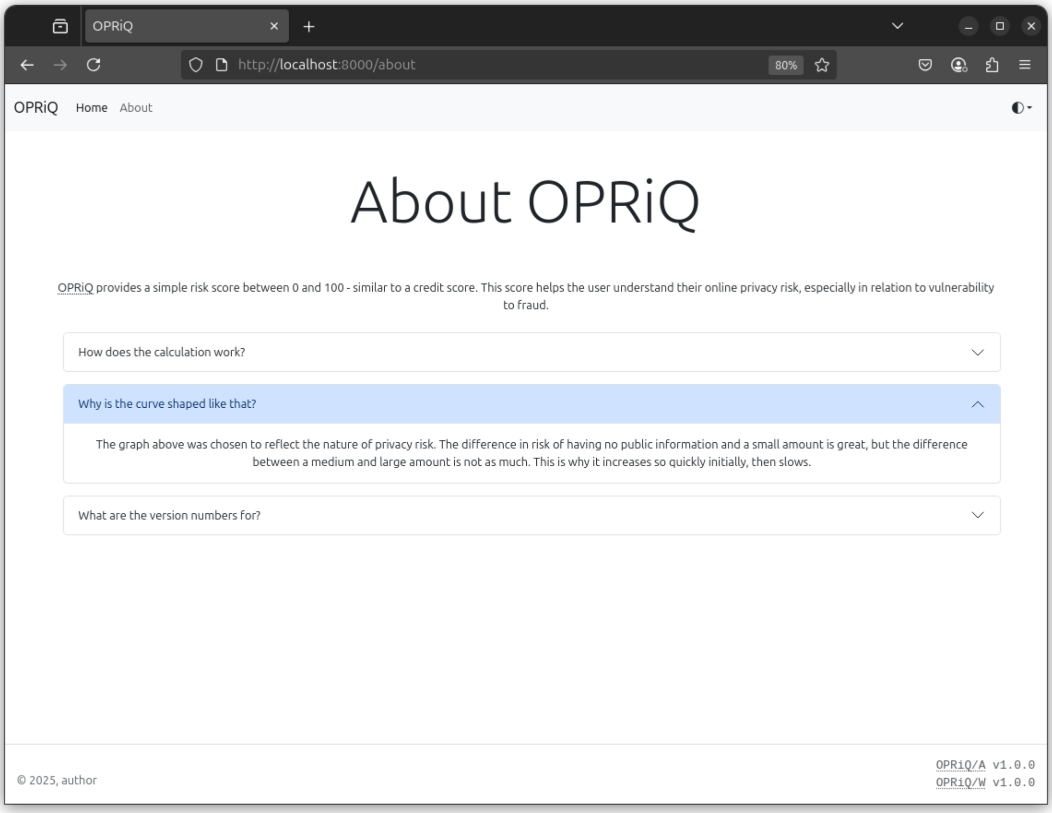
Also select any modifiers that apply

From the items you selected, the following can be inferred

© 2025, author

OPRiQ/A v1.0.0
OPRiQ/W v1.0.0

(a) The primary page.



About OPRiQ

OPRiQ provides a simple risk score between 0 and 100 - similar to a credit score. This score helps the user understand their online privacy risk, especially in relation to vulnerability to fraud.

How does the calculation work?

Why is the curve shaped like that?

The graph above was chosen to reflect the nature of privacy risk. The difference in risk of having no public information and a small amount is great, but the difference between a medium and large amount is not as much. This is why it increases so quickly initially, then slows.

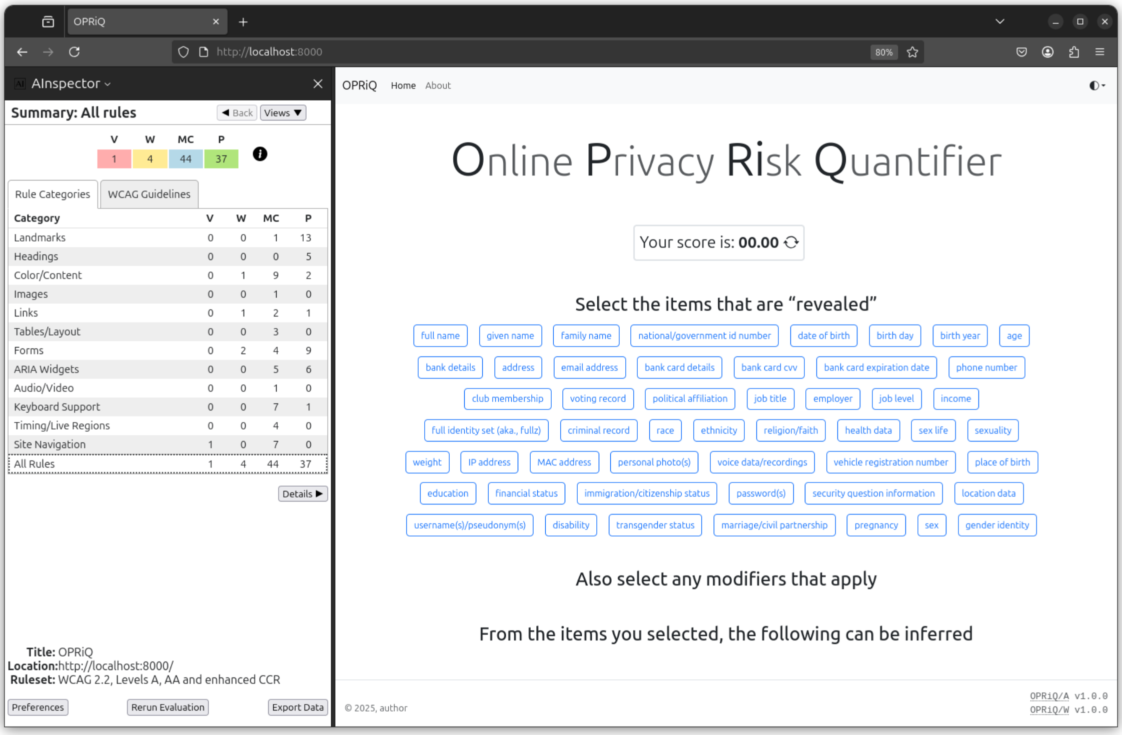
What are the version numbers for?

© 2025, author

OPRiQ/A v1.0.0
OPRiQ/W v1.0.0

(b) The about page, with one FAQ expanded.

Figure 5.3: Pages of the website.



Summary: All rules

V	W	MC	P
1	4	44	37

Category	V	W	MC	P
Landmarks	0	0	1	13
Headings	0	0	0	5
Color/Content	0	1	9	2
Images	0	0	1	0
Links	0	1	2	1
Tables/Layout	0	0	3	0
Forms	0	2	4	9
ARIA Widgets	0	0	5	6
Audio/Video	0	0	1	0
Keyboard Support	0	0	7	1
Timing/Live Regions	0	0	4	0
Site Navigation	1	0	7	0
All Rules	1	4	44	37

Title: OPRiQ
Location:http://localhost:8000/
Ruleset: WCAG 2.2, Levels A, AA and enhanced CCR

Online Privacy Risk Quantifier

Your score is: **00.00**

Select the items that are “revealed”

Also select any modifiers that apply

From the items you selected, the following can be inferred

© 2025, author

OPRiQ/A v1.0.0
OPRiQ/W v1.0.0

Figure 5.4: Use of AInspector on the homepage of the website.

5.3 Automated Input of Items

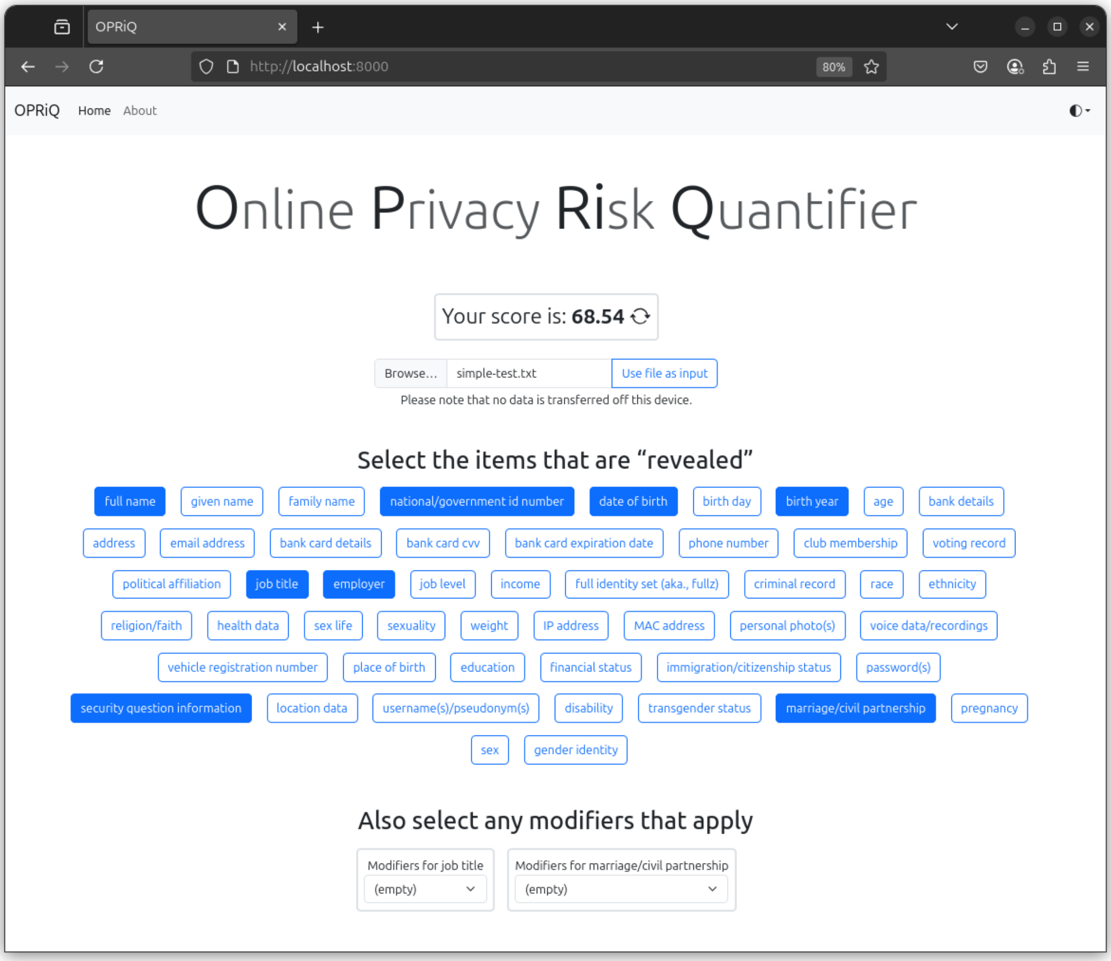
“Automated input of user data” refers to allowing users to process a structured file as inputs to the tool. In line with the ethical and legal concerns noted prior (see Section 1.4), and the conditions of the waiver of ethical approval, this data cannot be transferred outside of the user’s system. Therefore, all processing must take place in the user’s browser.

5.3.1 Prototype

An initial prototype can be created to interpret a set input structure. This example considers a file that is simply a comma-separated list of item IDs. For simplicity, modifiers are not included, but the user can still utilise the UI to input thee.

When provided as input, the corresponding button-clicks can be simulated and the remainder left to the existing system. This would also allow the user to see how the file was interpreted, and easily modify the input.

This prototype, when provided with the file content in Listing 5.6, causes the site to display Fig. 5.5.



Online Privacy Risk Quantifier

Your score is: **68.54**

Browse… simple-test.txt Use file as input

Please note that no data is transferred off this device.

Select the items that are “revealed”

Selected: full name, national/government id number, date of birth, birth year, job title, employer, security question information, marriage/civil partnership

Also select any modifiers that apply

Modifiers for job title: (empty)

Modifiers for marriage/civil partnership: (empty)

Figure 5.5: The result of using the file in Listing 5.6 as input; remainder of page not shown due to size.

Listing 5.6 simple-test.txt

```
2,5,6,8,20,21,43,48
```

5.3.2 ⟨Character⟩-Separated Values

To extend to allow for many formats, regular expressions can be used to match to a processing function. To illustrate: a regular expression for the above example could be `/(\d+,)+\d+/`.

Extending the comma-separated value example above to use other separator characters is relatively simple. Listing 5.7 shows a preliminary list of regular expression patterns, with two functions to process the corresponding content.

Listing 5.7 templates/home.html.django

```
let format_map = [
    [/(\d+,)+\d+/gs, csv_ids],
    [/(\d+\|)+\d+/gs, (i) => { return csv_ids(i, "|") }],
    [/(\d+:)+\d+/gs, (i) => { return csv_ids(i, ":") }],
    [/(\d+\t)+\d+/gs, (i) => { return csv_ids(i, "\t") }],
    [/(\d+\n)+\d+/gs, (i) => { return csv_ids(i, "\n") }],

    [/([\w ]+,)+[\w ]+/gs, csv_names],
    [/([\w ]+\|)+[\w ]+/gs, (i) => { return csv_names(i, "|") }],
    [/([\w ]+:)+[\w ]+/gs, (i) => { return csv_names(i, ":") }],
    [/([\w ]+\t)+[\w ]+/gs, (i) => { return csv_names(i, "\t") }],
    [/([\w ]+\n)+[\w ]+/gs, (i) => { return csv_names(i, "\n") }],
];
```

5.3.3 JSON-Based Content

Another common format to consider is JSON. Instead of detecting this via regular expressions, one can simply attempt to parse it. If this parsing succeeds, it must necessarily be valid JSON; otherwise, it is not and further detection can commence. Thankfully, JavaScript offers a built-in interface for JSON objects.

Listing 5.8 shows how JavaScript can be used to attempt to parse the file content as JSON, handling the error if not possible.

Listing 5.8 templates/home.html.django

```
let obj;
try {
    obj = JSON.parse(content);
} catch (e) {
    console.log(e);
    return;
}
```

As most JSON implementations support only arrays and dictionaries at the topmost level,[15] one only needs to check these types.

[15] Tim Bray, *The JavaScript Object Notation (JSON) Data Interchange Format*, RFC 8259, Dec. 2017. doi: 10.17487/RFC8259 at 2; see, i.e., Mozilla et al., "JSON", developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Global_Objects/JSON, accessed 26 May 2025.

Arrays

Using the “`obj`” above, where it is an array, each item can be handled depending on its type. For simplicity, it is assumed that numbers are IDs, strings are names, and objects are handled separately. Where the type does not fit into one of these categories, an error is returned.

While numbers and strings are relatively simple to parse, objects present further complexity. For these, likely property keys are searched for and, where one is found, its value is parsed as above. These likely keys are shown in Listing 5.9, where searching is case insensitive.

Listing 5.9 `/templates/home.html.django`

```
let id_keys = ["id", "item id", "item_id", "item-id"];
let name_keys = ["name", "item name", "item_name", "item-name"];
```

Dictionaries

Where the topmost item is not an array, and therefore a dictionary, one can search for common keys – similar to above in Listing 5.9. The corresponding values can then be interpreted as arrays, with the same logic as above.

CHAPTER

SIX

CONCLUSION

This report set out to create a novel algorithm to quantify online privacy risk (Objective 1) and provide a usable implementation (Objectives 2 and 3). These objectives have, in the author's opinion, been fulfilled.

For one, an algorithm *OPRiQ* has been created, along with the necessary data to support its implementation. Then, the web-based implementation has also been created, with the ability to automatically input data. Each of these has been tested, finalising the achievement of all objectives.

6.1 Results & Evaluation

While there are a wide range of existing online privacy risk quantification algorithms, the calculation established prior – *OPRiQ* – takes a somewhat different approach. Instead of presenting risk as linear, *OPRiQ* takes a logarithmic approach, where risk increases quickly then slows. Logarithmic techniques are not uncommon in risk measurement, but seem uncommon in online privacy risk quantification.

OPRiQ appears to be the first of its kind to provide a usable implementation. This is a considerable step forwards in the area of privacy risk quantification, as it allows any person to easily generate a score that represents their vulnerability.

As this implementation appears to be the first, it cannot be compared to previous work. However, it does accurately implement the algorithm, passes accessibility testing, allows for automated input, and handles data securely.

6.2 Discussion & Future Work

While this is a step forward in online privacy risk quantification, there is still work to be done.

For one, more research is needed to explore people's perception of online privacy risk. While a logarithmic approach appears logical, this may not align with people's actual feelings.

Secondly, and in the same vein, research is needed to extend and refine the data used in *OPRiQ*'s implementation. For reference and clarity, these are the item's sensitivities, modifiers, and inferences. These were established for *OPRiQ* through the analysis of legal resources, research, and other data; however, this could be greatly improved through human-based research.

REFERENCES

Action Fraud, "Protect yourself from fraud and cyber crime", www.actionfraud.police.uk/individual-protection, accessed 13 Nov. 2024.

Alatas, Syed Farid, "Academic Dependency and the Global Division of Labour in the Social Sciences", *Current Sociology*, 51/6 (2003), 599–613. doi: 10.1177/00113921030516003.

Alemany, Jose, et al., "Metrics for Privacy Assessment When Sharing Information in Online Social Networks", *IEEE Access*, 7 (2019), 143631–45. doi: 10.1109/ACCESS.2019.2944723.

Alohaly, Manar, and Takabi, Hassan, "Better Privacy Indicators: A New Approach to Quantification of Privacy Policies", in *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)* (Denver, CO: USENIX Association, June 2016).

Andy, "What Information Does A Scammer Need?", We Get Scammed For You (12 Dec. 2023), wegetscammedforyou.com/what-information-does-a-scammer-need.

Arshad, Razi, and Asghar, Muhammad Rizwan, "Characterisation and Quantification of User Privacy: Key Challenges, Regulations, and Future Directions", to be published in *IEEE Communications Surveys & Tutorials* (2024). doi: 10.1109/COMST.2024.3519861.

Banerjee, Mishtu, et al., "Quantifying Privacy Violations", in *Secure Data Management*, ed. Willem Jonker and Milan Petković (Berlin, Heidelberg: Springer Berlin Heidelberg, 2011), 1–17. doi: 10.1007/978-3-642-23556-6_1.

Bartoletti, Ivana, *An Artificial Revolution: On Power, Politics and AI* (The Indigo Press, 2020).

Beigel, Fernanda, "Academic Dependency", *Alternautas*, 2/1 (July 2015), 60–2. doi: 10.31273/alternautas.v2i1.1005.

Bischoff, Paul, "Passports on the dark web: how much is yours worth?", Comparitech (19 Oct. 2018), www.comparitech.com/blog/vpn-privacy/passports-on-the-dark-web-how-much-is-yours-worth.

Blauw, Frans F, and Solms, Sebastiaan von, "Towards quantifying and defining privacy metrics for online users", in *2017 IST-Africa Week Conference (IST-Africa)* (2017), 1–9. doi: 10.23919/ISTAFRICA.2017.8102366.

Bootstrap Team et al., *Bootstrap* (version 5.3.3) [frontend toolkit] (19 May 2025), getbootstrap.com.

Bootstrap Team, *JavaScript theme switcher* [Color mode toggler for Bootstrap's docs], getbootstrap.com/docs/5.3/customize/color-modes.

Bouillard, Anne, "Trade-off between accuracy and tractability of Network Calculus in FIFO networks", *Performance Evaluation*, 153 (2022), 102250. doi: 10.1016/j.peva.2021.102250.

Bray, Tim, *The JavaScript Object Notation (JSON) Data Interchange Format*, RFC 8259, Dec. 2017. doi: 10.17487/RFC8259.

Brin, Sergey, and Page, Lawrence, "The anatomy of a large-scale hypertextual Web search engine", *Computer Networks and ISDN Systems*, 30/1 (1998), Proceedings of the Seventh International World Wide Web Conference, 107–17. doi: 10.1016/S0169-7552(98)00110-X.

BS EN ISO 9241-11:2018, *Ergonomics of human-system interaction: Part 11 – Usability: Definitions and concepts* (British Standards Institute).

BS ISO/IEC 40314:2016, *Mathematical Markup Language (MathML) Version 3.0 2nd Edition* (British Standards Institute).

Burnap, Pete, "Risk Management & Governance", in *The Cyber Security Body of Knowledge v1.1.0, 2021*, Version 1.1.1 (University of Bristol, 2021).

Canonical Ltd., *Ubuntu* (version 0.83) [TrueType Font], design.ubuntu.com/font.

Codd, E. F., "A Relational Model of Data for Large Shared Data Banks", *Communications of the ACM* 13/6 (June 1970). doi: 10.1145/362384.362685.

—— "Further Normalization of the Data Base Relational Model", *Communications of the ACM* RJ909 (Aug. 1971).

Convention for the Protection of Human Rights and Fundamental Freedoms (opened for signature 4 Nov. 1950, entered into force 3 Sept. 1953) 213 UNTS 221.

Dan, Ovidiu, Parikh, Vaibhav, and Davison, Brian D., "IP Geolocation through Reverse DNS", *ACM Trans. Internet Technol.* 22/1 (Oct. 2021). doi: 10.1145/3457611.

Data Protection Act 2018.

De, Sourya Joyee, and Imine, Abdessamad, "Privacy Scoring of Social Network User Profiles Through Risk Analysis", in *Risks and Security of Internet and Systems*, ed. Nora Cuppens et al. (Springer International Publishing, 2018), 227–43.

Django Software Foundation et al., *Django* (version 5.2) [high-level Python web framework] (2 Apr. 2025), www.djangoproject.com.

—— "Django Documentation: Models and databases", version 5.2, docs.djangoproject.com/en/5.2/topics/db.

Django Software Foundation et al., “Django Documentation: Models”, version 5.2, docs.djangoproject.com/en/5.2/topics/db/models.

—— “Django Documentation: The Django admin site”, version 5.2, docs.djangoproject.com/en/5.2/ref/contrib/admin.

Dziedzic, Łukasz, *Lato* (version 2.015) [TrueType Font], www.latofonts.com.

Elad, Barry, “Dark Web Statistics 2024”, Enterprise Apps Today (19 Feb. 2024), www.enterpriseappstoday.com/stats/dark-web-statistics.html.

Equality Act 2010.

European Union (Withdrawal) Act 2018.

Galen, “The Diagnosis and Cure of the Soul’s Passions”, in *Galen on the Passions and Errors of the Soul*, trans. Paul W. Harkins, with an introduction by Walther Riese (Ohio State University Press, 1963), 27–69.

Get Safe Online, “Online Abuse”, www.getsafeonline.org/personal/articles/online-abuse, accessed 26 Jan. 2025.

—— “Online Gender-Based Violence”, www.getsafeonline.org/personal/articles/online-gender-based-violence, accessed 26 Jan. 2025.

—— “Oversharing”, www.getsafeonline.org/personal/articles/oversharing, accessed 26 Jan. 2025.

Gov.uk, “Stay safe on social media”, stopthinkfraud.campaign.gov.uk/protect-yourself-from-fraud/protecting-against-online-fraud/stay-safe-on-social-media, accessed 13 Nov. 2024.

Guidelines for the Regulation of Computerized Personnel Data Files (adopted 14 Oct. 1990) A/RES/45/95.

Gunderson, Jon, *AInspector* (version 3.1) [firefox extension for analysing accessibility compliance] (23 May 2025), ainspector.disability.illinois.edu.

Hamilton, Rosie, “What is a good credit score?”, Money Saving Expert (27 Nov. 2023), www.moneysavingexpert.com/credit-cards/what-is-a-good-credit-score.

Heard, Mathew, et al., *VPN & datacenter IPs* (14 May 2025), github.com/X4BNet/lists_vpn.

Information Commissioner v Magherafelt District Council [2012] UKUT 263 (AAC).

Information Commissioner v Miller [2018] UKUT 229 (AAC).

Information Commissioner’s Office, “What is personal data?”, ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/personal-information-what-is-it/what-is-personal-data/what-is-personal-data, accessed 11 May 2025.

JetBrains, *JetBrains Mono* (version 2.304) [TrueType Font], www.jetbrains.com/lp/mono.

Joinet, Louis, "Guidelines for the Regulation of Computerized Personal Data Files: final report", E/CN.4/Sub.2/1988/22 (July 1988), adopted by Guidelines for the Regulation of Computerized Personnel Data Files (adopted 14 Oct. 1990) A/RES/45/95.

L B v Hungary (App no 36345/16) ECHR 2023.

Lee, Chengpang, and Chen, Ying, "In What Ways We Depend: Academic Dependency Theory and the Development of East Asian Sociology", *Journal of Historical Sociology* (2022), 1–13. doi: 10.1002/johs.12358.

Liu, Kun, and Terzi, Evimaria, "A Framework for Computing the Privacy Scores of Users in Online Social Networks", *ACM Trans. Knowl. Discov. Data*, 5/1 (Dec. 2010). doi: 10.1145/1870096.1870102.

Lukács, Adrienn, *What is Privacy?: The History and Definition of Privacy* (University of Szeged).

Luna, Christopher, "How accurate is IP geolocation?", MaxMind (1 July 2021), blog.maxmind.com/2021/07/how-accurate-is-ip-geolocation.

McCallister, Erika, Grance, Tim, and Scarfone, Karen, *Guide to Protecting the Confidentiality of Personally Identifiable Information*, Special Publication 800-122 (National Institute of Standards and Technology, Apr. 2010).

Meyer, Werner G., "Quantifying risk: measuring the invisible", in *PMI® Global Congress 2015-EMEA*, Project Management Institute (London, England, 2015).

Moore, Adam D., "Defining Privacy", *Journal of Social Philosophy*, 39/3 (2008), 411–28.

Mozilla et al., "<template>: The Content Template element", developer.mozilla.org/en-US/docs/Web/HTML/Reference/Elements/template, accessed 23 May 2025.

—— "JSON", developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Global_Objects/JSON, accessed 26 May 2025.

Negley, Glenn, "Philosophical Views on the Value of Privacy", *Law and Contemporary Problems*, 31 (1966), 319.

Nepali, Raj Kumar, and Wang, Yong, "SONET: A SOcial NETwork Model for Privacy Monitoring and Ranking", in *2013 IEEE 33rd International Conference on Distributed Computing Systems Workshops* (2013), 162–6. doi: 10.1109/ICDCSW.2013.49.

Ofcom, *Adults' Media Literacy Core Survey 2024* [SPSS Data Tables] (20 Jan. 2025).

Office for National Statistics, "Gender identity, England and Wales: Census 2021", statistical bulletin (6 Jan. 2023), www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/genderidentity/bulletins/genderidentityenglandandwales census2021.

ONS Centre for Crime and Justice, *Appendix tables - Nature of fraud and computer misuse, year ending March 2024* [Spreadsheet], data from the Crime Survey for England and Wales and the National Fraud Intelligence Bureau (Office for National Statistics, 6 Nov. 2024).

—— *Crime in England and Wales, Appendix tables - year ending June 2024* [Data Tables] (Office for National Statistics, 24 Oct. 2024).

Pensa, Ruggero G., Di Blasi, Gianpiero, and Bioglio, Livio, "Network-aware privacy risk estimation in online social networks", *Social Network Analysis and Mining*, 9/1 (Apr. 2019). doi: 10.1007/s13278-019-0558-x.

R (on the application of Department of Health) v Information Commissioner [2011] EWHC 1430 (Admin).

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC [2016] OJ L119.

Richardson, Janice, *Law and the Philosophy of Privacy* (27 Aug. 2015).

Robinson, Tim, et al., *Annual Fraud Indicator 2023* (Crowe, 2023).

Sahinoglu, M., "Security meter: a practical decision-tree model to quantify risk", *IEEE Security & Privacy*, 3/3 (2005), 18–24. doi: 10.1109/MSP.2005.81.

Samaranayake, S., Blandin, S., and Bayen, A., "A tractable class of algorithms for reliable routing in stochastic networks", *Transportation Research Part C: Emerging Technologies*, 20/1 (2012), 199–217. doi: 10.1016/j.trc.2011.05.009.

Satakunnan Markkinapörssi Oy And Satamedia Oy v Finland (App no 931/13) ECHR 2017.

Solove, Daniel J., *Understanding Privacy*, Legal Studies Research Paper No. 420, Public Law Research Paper No. 420 (The George Washington University Law School, May 2008).

Stack Overflow users, "How do you like your primary keys?" (19 Jan. 2017), stackoverflow.com/questions/404040/how-do-you-like-your-primary-keys, accessed 9 May 2025.

Tabagari, Salome, "Credit scoring by logistic regression", Master's Thesis (University of Tartu, 2015).

Talukder, Nilothpal, et al., "Privometer: Privacy protection in social networks", in *2010 IEEE 26th International Conference on Data Engineering Workshops (ICDEW 2010)* (2010), 266–9. doi: 10.1109/ICDEW.2010.5452715.

Team, Risk & Resilience, *Risk & Issues Register Template & Guidance* (University of Warwick, [13 June 2023]).

The Noto Project Authors, *Noto Sans Math* (version 3.0) [TrueType Font], github.com/notofonts/math.

The right to privacy: conference outcomes (Stockholm: International Commission of Jurists, 1967).

Troncoso, Carmela, “Privacy & Online Rights”, in *The Cyber Security Body of Knowledge v1.1.0, 2021*, Version 1.0.2 (University of Bristol, 2021).

Trustwave, “How Prices are Set on the Dark Web: Exploring the Economics of Cybercrime” (25 Nov. 2024), www.trustwave.com/en-us/resources/blogs/trustwave-blog/how-prices-are-set-on-the-dark-web-exploring-the-economics-of-cybercrime.

U.S. Department of Homeland Security, *Handbook for Safeguarding Sensitive PII*, Revision 3, Privacy Policy Directive 047-01-007 (4 Dec. 2017).

United Nations, *World Population Prospects 2024: Summary of Results* (UN DESA/POP/2024/TR/NO. New York: Department of Economic and Social Affairs, Population Division, 2024).

Universal Declaration of Human Rights (adopted 10 Dec. 1948) A/RES/3/217 A.

Usmani, Fahad, “What is Expected Monetary Value in Project Management?”, PM Study Circle (27 Sept. 2024), pmstudycircle.com/expected-monetary-value-emv.

Vidyalakshmi, B.S., Wong, Raymond K., and Chi, Chi-Hung, “Privacy Scoring of Social Network Users as a Service”, in *2015 IEEE International Conference on Services Computing* (2015), 218–25. doi: 10.1109/SCC.2015.38.

Wagner, Isabel, and Eckhoff, David, “Technical Privacy Metrics: A Systematic Survey”, *ACM Comput. Surv.* 51/3 (June 2018). doi: 10.1145/3168389.

Wang, Yong, Nepali, Raj Kumar, and Nikolai, Jason, “Social network privacy measurement and simulation”, in *2014 International Conference on Computing, Networking and Communications (ICNC)* (2014), 802–6. doi: 10.1109/ICCNC.2014.6785440.

Web Content Accessibility Guidelines, ed. Alastair Campbell et al., version 2.2 (W3C Recommendation; World Wide Web Consortium, 5 Oct. 2023), www.w3.org/TR/WCAG22.

Zoltan, Miklos, “Dark Web Price Index 2023”, Privacy Affairs (23 Apr. 2023), www.privacyaffairs.com/dark-web-price-index-2023.

APPENDIX

A

ETHICAL APPROVAL

A.1 Ethical Approval Waiver Email

Date of Approval: 18/02/2025

Dear Autumn

Warwick ID number: 2208993

Project: Creation of an Online Privacy Risk Quantification Tool

Your ethics application number is WMG-2025-FTBSc-R_22312BYGDrGb7tv

Your supervisor Sandy Taramonli has recommended to the Cyber Ethics Panel the following outcome: Ethical approval be waived for this research.

The Cyber Ethics Panel has confirmed this outcome.

You are reminded that you must now adhere to the answers and detail given in the completed ethics form. If anything changes in your research such that any of your answers change to the form for which you received an ethical waiver for, then you must contact your supervisor to check if you need to reapply for or update your ethical waiver before you proceed with data collection.

When you submit your dissertation, please write N/A against the ethical approval field on the cover page of the submission and include a copy of this email into the Appendices of your dissertation.

Kind regards,

WMG Projects Team

APPENDIX

B

TECHNICAL LITERATURE REVIEW

This chapter technical, in-depth, review of online privacy risk quantification literature. For more of an overview, see Chapter 3.

Metrics for Privacy Assessment When Sharing Information in Online Social Networks

Alemany et al.[1] propose two metrics to be used in tandem: Reachability ($Re(a_i, l, r)$) and Audience ($Au(a_i, l)$ and $Au_G(a_i, l)$), as defined below.

The reachability $Re(a_i, l, r)$ is defined as the probability of a_i's message being seen by the proportion r of users in the layer l. N is the network of users in the social network. $L_{a_i}(l)$ is the set of users in N who are l "steps" away from a_i, such that $L_{a_i}(1)$ are those directly connected to a_i. Finally, $\Gamma(a_i, a_j) = 1$ if a message from a_i is seen by a_j and 0 otherwise; using this Γ, Γ' and Γ'' are defined as in Eqs. (B.1) and (B.2).

$$\Gamma' = \left\{\gamma \in \Gamma \,\middle|\, \exists a_j : \gamma(a_i, a_j) = 1\right\} \tag{B.1}$$

$$\Gamma'' = \left\{\gamma \in \Gamma \,\middle|\, \frac{\sum\limits_{a_j \in L_{a_i}(l)} \gamma(a_i, a_j)}{\left|L_{a_i}(l)\right|} \geq r\right\} \tag{B.2}$$

The reachability metric is then defined as below.

[1] Alemany et al., "Metrics for Privacy Assessment When Sharing Information in Online Social Networks".

$$Re(a_i, l, r) = \frac{|\Gamma''|}{|\Gamma'|} \tag{B.3}$$

Their audience metric uses the same definitions as above, and is defined as in Eqs. (B.5) and (B.6). A is extracted for brevity, though Alemany et al. chose not to present it in this way.

$$A = \sum_{a_j \in L_{a_i}(l)} \gamma(a_i, a_j) \tag{B.4}$$

$$Au(a_i, l) = \frac{\sum_{\gamma \in \Gamma'} \left(\frac{A}{|L_{a_i}(l)|} \right)}{|\Gamma'|} \tag{B.5}$$

$$Au_G(a_i, l) = \frac{\sum_{\gamma \in \Gamma'} \left(\frac{A}{|N|} \right)}{|\Gamma'|} \tag{B.6}$$

Alemany et al. argue that these metrics allow users to gauge the risk of posting a message, by comparing the values $Re(a_i, l, r)$ and $Au(a_i, l)$ against one's internal preferences. It is also for the user to determine the values of l and r, with l mainly being based on tractability and r being based on their risk tolerance. In addition, they suggest Au_G as it gives the user a "more global picture of the risk" when comparing to Au.

Network-aware privacy risk estimation in online social networks

Similar to the well-known "Pagerank" algorithm,[2] Pensa, Di Blasi, and Bioglio[3] consider the privacy scores of neighbours in computing one's own score.

The "intrinsic" privacy risk $\rho_p(v_i)$ for some user v_i is defined by Eqs. (B.7) to (B.12). For these definitions, there are n users and m "topics" or items; v_i is then the ith user and t_j is the jth topic. Further, σ_j is the sensitivity of topic t_j and υ_{ij} is the visibility of t_j due to v_i. Finally, r_{ij} represents the number of "levels" that user v_i is willing to share topic t_j with; for example $r_{ij} = 0$ means v_i wants to keep t_j private, $r_{ij} = 1$ means it can be shared with direct friends, and so on.

[2] Brin and Page, "The anatomy of a large-scale hypertextual Web search engine".

[3] Pensa, Di Blasi, and Bioglio, "Network-aware privacy risk estimation in online social networks".

$$\rho_p(v_i) = \sum_{j=1}^{m} \rho'_p\left(v_i, t_j\right) \tag{B.7}$$

$$\rho'_p(v_i, t_j) = \frac{\rho_p\left(v_i, t_j\right)}{\max\limits_{v_k \in V} \rho_p\left(v_k, t_j\right)} \tag{B.8}$$

$$\rho_p(v_i, t_j) = \sum_{h=0}^{\ell} \sigma_{jh} \times \upsilon_{ijh} \tag{B.9}$$

$$\sigma_{jh} = \frac{1}{2}\left(\frac{n - \sum_{i=1}^{n} \mathbb{1}\left(r_{ij} \geq h\right)}{n} + \frac{n - \sum_{i=1}^{n} \mathbb{1}\left(r_{ij} \geq h + 1\right)}{n}\right) \tag{B.10}$$

$$\mathbb{1} : \begin{cases} \mathbb{1}(A) = 1 \iff A. \\ \mathbb{1}(A) = 0 \iff \overline{A}. \end{cases} \tag{B.11}$$

$$\upsilon_{ijh} = \frac{\sum_{i=1}^{n} \mathbb{1}\left(r_{ij} = h\right)}{n} \times \frac{\sum_{j=1}^{m} \mathbb{1}\left(r_{ij} = h\right)}{m} \times h \tag{B.12}$$

Their "network-aware" privacy score is then defined as a distribution satisfying the following, where $\boldsymbol{\rho} = \left[\rho_p(v_1), \dots, \rho_p(v_n)\right]^{\mathsf{T}}$ and d is a damping factor such that $0 < d < 1$.

$$\mathbf{P} = d\mathbf{A}^{\mathsf{T}}\mathbf{P} + \frac{1 - d}{\sum\limits_{k=1}^{n} \rho_p(v_k)}\boldsymbol{\rho} \tag{B.13}$$

As with the previous study,[4] this score carries issues of tractability – a common problem in network-aware algorithms.[5] Calculating one's "true" risk score would require also calculating those of the near 8.2 billion people on earth,[6] and then fitting these to the distribution above.

A Framework for Computing the Privacy Scores of Users in Online Social Networks

Liu and Terzi[7] establish a privacy score based on the sensitivity and visibility of items within a network.

[4] Alemany et al., "Metrics for Privacy Assessment When Sharing Information in Online Social Networks".

[5] See, i.e., discussion in Bouillard, "Trade-off between accuracy and tractability of Network Calculus in FIFO networks"; Samaranayake, Blandin, and Bayen, "A tractable class of algorithms for reliable routing in stochastic networks".

[6] United Nations, *World Population Prospects 2024*, p. VII.

[7] Liu and Terzi, "A Framework for Computing the Privacy Scores of Users in Online Social Networks".

Each user $j \in \{1, \ldots, N\}$ in the network G has a profile of n items. Each of these items is assigned a "privacy level" in either $\{0, 1\}$ or $\mathbb{N}$ depending on the system (dichotomous and polytomous, respectively). The value $\mathrm{R}(i, j)$ is the privacy level of item i for user j. For example, for a dichotomous system, $\mathrm{R}(i, j) = 0$ means the user does not want to share that item and $\mathrm{R}(i, j) = 1$ means they do. For the polytomous system, the value of $\mathrm{R}(i, j)$ is the number of "steps" the user is comfortable sharing that item with – similar to Alemany et al.'s $L_{a_i}(l)$, above.

Then, the sensitivity of any item i is defined as β_i and the visibility is $V(i, j) = Prob\,\{R(i, j) = 1\}$ for some user j. Using these, the privacy score for a user j by consequence of an item i is $PR(i, j) = \beta_i \otimes V(i, j)$, where $\otimes$ is an arbitrary combination function. This can be extended to a full user risk score quite simply, as in Eqs. (B.14) to (B.15).

$$
\begin{aligned}
PR(j) &= \sum_{i=1}^{n} PR(i, j) && \text{(B.14)} \\
&= \sum_{i=1}^{n} (\beta_i \otimes V(i, j)) && \text{(B.15)}
\end{aligned}
$$

As it stands, this is only applicable in the dichotomous case; to extend to the polytomous system, a parameter k is added, as well as further refinements. Visibility becomes $V(i, j, k) = Prob\,\{\mathrm{R}(i, j) = k\} \times f_j(k)$, where f_j is some function of k such that $f_j(0) = 0$. Sensitivity is also β_{ik}, for the same i and k.

Extending the dichotomous system, the polytomous user risk score is in Eq. (B.16).

$$
PR(j) = \sum_{i=1}^{n} \sum_{k=0}^{l} (\beta_{ik} \times V(i, j, k)) \qquad \text{(B.16)}
$$

Liu and Terzi provide two ways for computing β_{ik} – a naïve method and an Item Response Theory (IRT) method. They then ran surveys to generate data to test these algorithms – comparing users' scores and attitudes across different regions.

Privometer: Privacy protection in social networks

Talukder et al.[8] develop a tool for Facebook that they call "Privometer", utilising item inference and the privacy scores of a user's friends.

Their model considers a malicious user who knows the attributes of each friend of a

[8] Talukder et al., "Privometer: Privacy protection in social networks".

person. The model then employs inference functions to "guess" the attributes of the user. Privometer then suggests "self-sanitisation" actions – recommendations as to attributes the user's friends should hide.

Talukder et al. claim that Privometer is implemented in Facebook, via the *Facebook PHP Client API.*

SONET: A SOcial NETwork Model for Privacy Monitoring and Ranking

Nepali and Wang,[9] in their "SONET" model, construct a graph of connected users who all have their own attributes.

Their "privacy index", as defined in Eq. (B.19), is a percentage value based on the "privacy impact factors" of a user's known attributes. For the below, L_k is the subset of attributes that are known for the user k and S_k is the corresponding subset of impact factors.

$$W_{L_k} = \sum S_k \tag{B.17}$$

$$W = \sum S^{10} \tag{B.18}$$

$$PIDX = \frac{W_{L_k}}{W} \times 100 \tag{B.19}$$

They then define privacy invasion as when $PIDX \geq T$ for some user-defined threshold T; otherwise, privacy is preserved.

Extending this $PIDX$ to consider visibility is done by adding the visibility measurement set V_k, as shown in Eq. (B.20). They do not define a concrete algorithm for computing this, but they do mention the use of "depth-first or breath-first traversals".

$$W_{L_k} = \sum_{i=0}^{n} (S_{ki} V_{ki}) \tag{B.20}$$

Nepali and Wang go on to explain how their model could be created from search engine data, and data from social networking sites directly. They also discuss the

[9] Nepali and Wang, "SONET: A SOcial NETwork Model for Privacy Monitoring and Ranking".

[10] While this is not explicitly stated by Nepali and Wang, it is implied.

possibility of utilising data in deep-web sites. The author finds it pertinent to mention that utilising data from these sources may be in contravention of the law and ethical standards.

Social network privacy measurement and simulation

Wang, Nepali, and Nikolai[11] extend SONET[12] with their "composite PIDX" function $c - PIDX$, as defined in Eq. (B.23).

$$w - PIDX(i, j) = \frac{\sum_{t=1}^{n} s_t g_j(a_{jt}, d_{ij})}{\sum_{t=1}^{n} s_j} \times 100 \tag{B.21}$$

$$m - PIDX(i, j) = \max\left(s_1 g_j(a_{j1}, d_{ij}), s_2 g_j(a_{j2}, d_{ij}), \dots, s_n g_j(a_{jn}, d_{ij})\right) \times 100 \tag{B.22}$$

$$c - PIDX(i, j) = m - PIDX(i, j) + (100 - m - PIDX(i, j)) \times \frac{w - PIDX(i, j)}{100} \tag{B.23}$$

Where s_t, g_i, a_{jt}, d_{ij} are defined as follows.

s_t is the sensitivity of item $t \in [0, n]$ s.t. $s_t \in S$ – the set of sensitivities.

g_i is some user A_i's "attribute visibility function".

a_{jt} is attribute $t \in [0, n]$ of user A_j.

d_{ij} is the degrees of separation between user A_i and A_j.

Privacy Scoring of Social Network User Profiles Through Risk Analysis

Taking a Privacy Risk Analysis (PRA) approach, De and Imine[13] construct harm trees based on item inference and visibility.

In their system, the root node of a tree represents a specific privacy harm, leaves are personal data attributes (name, gender, etc.), and connecting nodes are specific threats (such as unintended inference). These nodes are then connected by "AND",

[11] Wang, Nepali, and Nikolai, "Social network privacy measurement and simulation".
[12] Nepali and Wang, "SONET: A SOcial NETwork Model for Privacy Monitoring and Ranking".
[13] De and Imine, "Privacy Scoring of Social Network User Profiles Through Risk Analysis".

"OR", or "*k*-out-of-*n*" decisions. For example, in their Figure 4, De and Imine state that a user's birth date and year infer their date of birth. In a less explicit example, a user's date of birth, home address, and phone number all being revealed can lead to identity theft/fraud, they claim.

To decrease the size of these trees, they provide "pruning" methods based on item visibility and accuracy.

De and Imine claim that this mechanism could then be used as the base for a system to inform users of their privacy scores.

Towards quantifying and defining privacy metrics for online users

Also looking at visibility and sensitivity, similar to Nepali and Wang[14]'s $PIDX$, Blauw and Solms[15] calculate scores based on multiplying the these elements as well as the quantity. For sensitivity s, quantity q, and visibility v, the score would be sqv.

While the calculation is relatively simplistic, their view on visibility does seem unique. They extend the layers of "WWW"–"deep web"–"dark web" to their visibility metric, considering the amount of authentication required to access.

[14] Nepali and Wang, "SONET: A SOcial NETwork Model for Privacy Monitoring and Ranking".

[15] Blauw and Solms, "Towards quantifying and defining privacy metrics for online users".

APPENDIX

C

MODIFICATION VALUES BASED ON ONS DATA

Table C.1: Modification values $m_{i=j}$ for the characteristics given in *Appendix tables - Nature of fraud and computer misuse, year ending March 2024*.

Characteristic	Value	Mod (4 d.p.)
Age	16-24	–0.3652
Age	25-34	–0.1011
Age	35-44	0.0944
Age	45-54	0.2126
Age	55-64	0.2049
Age	65-74	0.0232
Age	75+	–0.2071
Men	Men	–0.1169
Men	Men 16-24	–0.4554
Men	Men 25-34	–0.2312
Men	Men 35-44	–0.0402
Men	Men 45-54	0.0423
Men	Men 55-64	0.0396
Men	Men 65-74	–0.0377
Men	Men 75+	–0.2094
Women	Women	0.1129
Women	Women 16-24	–0.2698
Women	Women 25-34	0.0325
Women	Women 35-44	0.2258
Women	Women 45-54	0.3784
Women	Women 55-64	0.3635
Women	Women 65-74	0.0796
Women	Women 75+	–0.2052
Ethnic group	White	0.0316

(continues)

Table C.1: (continued)

Characteristic	Value	Mod (4 d.p.)
Ethnic group	English/Welsh/Scottish/Northern Irish/British	0.0379
Ethnic group	Irish	0.0908
Ethnic group	Any other white background	–0.0645
Ethnic group	Mixed/Multiple	–0.0043
Ethnic group	White and Black Caribbean	–0.0616
Ethnic group	White and Black African	–0.3333
Ethnic group	White and Asian	0.3593
Ethnic group	Any other Mixed/Multiple ethnic background	–0.2260
Ethnic group	Asian/Asian British	–0.2629
Ethnic group	Indian	–0.1176
Ethnic group	Pakistani	–0.3234
Ethnic group	Bangladeshi	–0.6285
Ethnic group	Chinese	–0.5406
Ethnic group	Any other Asian background	–0.1475
Ethnic group	Black/African/Caribbean/Black British	–0.0704
Ethnic group	African	–0.1913
Ethnic group	Caribbean	–0.0625
Ethnic group	Any other Black/African/Caribbean background	1.2951
Ethnic group	Other ethnic group	–0.1725
Ethnic group	Arab	–0.2168
Ethnic group	Any other ethnic group	–0.1473
Country of birth	Born in the UK	0.0226
Country of birth	Not born in the UK	–0.1104
Marital status	Married/civil partnered	0.0366
Marital status	Cohabiting	0.0638
Marital status	Single	–0.1615
Marital status	Separated	0.5364
Marital status	Divorced/legally dissolved partnership	0.4720
Marital status	Widowed	–0.2255
Employment status	In employment	0.0694
Employment status	Unemployed	0.1181
Employment status	Economically inactive	–0.1367
Employment status	Economically inactive: Student	–0.2556
Employment status	Economically inactive: Looking after family/home	–0.3942
Employment status	Economically inactive: Long-term/temporarily sick/ill	0.3649
Employment status	Economically inactive: Retired	–0.1258

(continues)

Table C.1: (continued)

Characteristic	Value	Mod (4 d.p.)
Employment status	Economically inactive: Other inactive	–0.6703
Occupation	Managerial and professional occupations	0.1855
Occupation	Intermediate occupations	–0.0155
Occupation	Routine and manual occupations	–0.1330
Occupation	Never worked and long-term unemployed	–0.4485
Occupation	Full-time students	–0.2791
Occupation	Not classified	0.2807
Highest qualification	Degree or diploma	0.1668
Highest qualification	Apprenticeship or A/AS level	0.0074
Highest qualification	O level/GCSE	–0.0731
Highest qualification	Other	–0.2274
Highest qualification	None	–0.4302
Disability	Not disabled	–0.0780
Disability	Disabled	0.3172
Religion	No religion	–0.0175
Religion	Christian	0.0413
Religion	Buddhist	0.3751
Religion	Hindu	–0.0085
Religion	Jewish	0.0414
Religion	Muslim	–0.3604
Religion	Sikh	–0.5202
Religion	Other	1.0234
Sexual orientation	Heterosexual/straight	0.0237
Sexual orientation	Gay/Lesbian	–0.0106
Sexual orientation	Bisexual	0.0423
Sexual orientation	Other	0.2007
Gender identity	Gender identity the same as sex registered at birth	0.0561
Gender identity	Gender identity different from sex registered at birth	–0.6545

APPENDIX

D

INITIAL API ENDPOINTS

GET `/api/i` `json`

Returns the list of items in I, in the following JSON format.

```
[
    {
        "id": ⟨int⟩,
        "name": ⟨string⟩,
        "base sensitivity": ⟨int⟩
    }, ...
]
```

GET `/api/i/⟨id⟩` `json`

Returns the item in I, with the id ⟨id⟩, in the following JSON format.

```
{
    "name": ⟨string⟩,
    "base sensitivity": ⟨int⟩
}
```

GET `/api/m` `json`

Returns the list of modifications in *M*, in the following JSON format.

```
[
    {
        "item id": ⟨int⟩,
        "value": ⟨string⟩,
        "modifier": ⟨float⟩
    }, ...
]
```

GET `/api/m?item=⟨id⟩` `json`

Returns the list of modifications for the item with the id ⟨id⟩, in the following JSON format.

```
[
    {
        "value": ⟨string⟩,
        "modifier": ⟨float⟩
    }, ...
]
```

GET `/api/ir` `json`

Returns the list of inference tuples in *IR*, in the following JSON format.

```
[
    {
        "from items": [⟨int⟩, ...],
        "to item": ⟨int⟩,
        "probability": ⟨float⟩
    }, ...
]
```

GET `/api/ir?to=⟨id⟩` json

Returns the list of inference tuples that infer the item with the id ⟨id⟩, in the following JSON format.

```
[
    {
        "from items": [⟨int⟩, ...],
        "probability": ⟨float⟩
    }, ...
]
```

GET `/api/opriq` xml

Returns a *MathML* representation of *OPRiQ*, in line with BS ISO/IEC 40314:2016.

APPENDIX

E

DATABASE AND SET VALUES

The following tables are auto-generated from the Django database.

Table E.1: Values for table `Item`.

Id	Name (*i*)	Base Sensitivity (s_i)
02	full name	0.750
03	given name	0.550
04	family name	0.550
05	national/government id number	0.950
06	date of birth	0.750
07	birth day	0.550
08	birth year	0.550
09	age	0.600
10	bank details	0.950
11	address	0.850
12	email address	0.550
13	bank card details	0.950
14	bank card cvv	0.750
15	bank card expiration date	0.750
16	phone number	0.750
17	club membership	0.700
18	voting record	0.700
19	political affiliation	0.700
20	job title	0.750
21	employer	0.750
22	job level	0.750
23	income	0.750
24	full identity set (aka., fullz)	1.000
25	criminal record	0.700
	(continues)	

Table E.1: (continued)

Id	**Name** (*i*)	**Base Sensitivity** (s_i)
26	race	0.600
27	ethnicity	0.600
28	religion/faith	0.600
29	health data	0.850
30	sex life	0.700
31	sexuality	0.700
32	weight	0.100
33	IP address	0.850
34	MAC address	0.500
35	personal photo(s)	0.800
36	voice data/recordings	0.800
37	vehicle registration number	0.750
38	place of birth	0.800
39	education	0.500
40	financial status	0.850
41	immigration/citizenship status	0.900
42	password(s)	0.900
43	security question information	0.900
44	location data	0.800
45	username(s)/pseudonym(s)	0.400
46	disability	0.600
47	transgender status	0.600
48	marriage/civil partnership	0.600
49	pregnancy	0.600
50	sex	0.600
51	gender identity	0.600

Table E.2: Values for table `Inference` and `Inference-Item`.

Id	**From**	**To**	**Probability**
01	given name, family name	full name	1.000
02	birth year, birth day	date of birth	1.000
03	date of birth	birth day	1.000
04	date of birth	birth year	1.000
05	age	birth year	0.900
06	birth year	age	0.900
07	voting record	political affiliation	0.800
08	political affiliation	voting record	0.800

(continues)

Table E.2: (continued)

Id	From	To	Probability
09	IP address	address	0.300
10	date of birth	age	1.000
11	full name	given name	1.000
12	full name	family name	1.000
13	sex, gender identity	transgender status	1.000
14	gender identity	sex	0.935
15	gender identity	sex	0.935
16	voice data/recordings	gender identity	0.200
17	personal photo(s)	gender identity	0.900
18	place of birth	ethnicity	0.700
19	income	financial status	0.800
20	financial status	income	0.700
21	job title	job level	0.850
22	job title, employer	income	0.900

Table E.3: Values for table `Modifier`.

Id	Item	Multiplier	Value
1	9	−0.3652	16-24
2	9	−1.011	25-34
3	9	0.0944	35-44
4	9	0.2126	45-54
5	9	0.2049	55-64
6	9	0.0232	65-74
7	9	−0.2071	75+
8	51	−0.1169	man
9	51	0.1129	woman
10	27	0.0379	white british
11	27	0.0908	white irish
12	27	−0.0645	white - other
13	27	−0.0616	white and black caribbean
14	27	−0.3333	white and black african
15	27	0.3593	white and asian
16	27	−0.226	mixed/multiple - other
17	27	−0.1176	indian
18	27	−0.3234	pakistani
19	27	−0.6285	bangladeshi
20	27	−0.5406	chinese

(continues)

Table E.3: (continued)

Id	Item	Multiplier	Value
21	27	−0.1475	asian - other
22	27	−0.1913	african
23	27	−0.0625	caribbean
24	27	1.2951	black/african - other
25	27	−0.2168	arab
26	27	−0.1473	other
27	48	0.0366	married/civil partnered
28	48	0.0638	cohabiting
29	48	−0.1615	single
30	48	0.5364	separated
31	48	0.472	divorced/legally dissolved
32	48	−0.2255	widowed
33	20	0.1855	managerial
34	20	−0.133	routine/manual
35	20	−0.4485	unemployed
36	20	−0.2791	student
37	20	0.2807	other
38	39	0.1668	degree
39	39	0.0074	no more than a/as-level
40	39	−0.0731	no more than o-level/gcse
41	39	−0.2274	other
42	39	−0.4302	none
43	46	−0.078	not disabled
44	46	0.3172	disabled
45	28	−0.0175	none
46	28	0.0413	christian
47	28	0.3751	buddhist
48	28	−0.0085	hindu
49	28	0.0414	jewish
50	28	−0.5604	muslim
51	28	−0.5202	sikh
52	28	1.0234	other
53	31	0.0237	straight
54	31	−0.0106	gay/lesbian
55	31	0.0423	bisexual
56	31	0.2007	other
57	47	0.0561	cisgender
58	47	−0.6545	transgender
59	33	−0.9	VPN address

(continues)

Table E.3: (continued)

Id	Item	Multiplier	Value
60	21	0.5	high-value target
61	40	0.5	large disposable income
62	40	0.5	in poverty
63	17	0.2	unpopular
64	17	0.35	deeply unpopular
65	17	0.5	persecuted
66	18	0.2	unpopular
67	18	0.35	deeply unpopular
68	18	0.5	persecuted
69	19	0.2	unpopular
70	19	0.35	deeply unpopular
71	19	0.5	persecuted
72	25	−0.5	minimal
73	25	0.1	moderate
74	25	0.5	major
75	29	0.4	somewhat revealing
76	29	0.6	majorly revealing

APPENDIX

F

MATHML MARKUP OF *OPRiQ*

Listing F.1 static/opriq.mathml

```
<math mode="display" xmlns="http://www.w3.org/1998/Math/MathML">
<semantics>
    <!-- First child of <semantics> is displayed -->
    <mrow>
        <mi>OPRiQ</mi>
        <mo>=</mo>
        <mfrac>
            <mn>1</mn>
            <mn>3</mn>
        </mfrac>
        <mo>&#x22C5;</mo>  <!-- Center dot -->
        <mo fence="true">[</mo>
            <mo fence="true">(</mo>
                <msup>
                    <mi>e</mi>
                    <mn>3</mn>
                </msup>
                <mo>-</mo>
                <mn>1</mn>
            <mo fence="true">)</mo>
            <mo>&#x22C5;</mo>  <!-- Center dot -->
            <mfrac>
                <mrow>
                    <munder>
                        <mo largeop="true" movablelimits="true">∑</mo>
                        <mrow>
                            <mi>i</mi>
                            <mo>∈</mo>
                            <mi>I</mi>
                        </mrow>
                    </munder>
                    <mo fence="true">(</mo>
                        <msub>
                            <mrow>
                                <mi>s</mi>
                                <mo>′</mo>
                            </mrow>
                            <mrow>
                                <mi>i</mi>
                                <mo>=</mo>
                                <msub>
                                    <mi>j</mi>
                                    <mi>i</mi>
                                </msub>
                            </mrow>
                        </msub>
                        <mo>□</mo>
                        <msub>
                            <mrow>
                                <mi>r</mi>
                                <mo>′</mo>
                            </mrow>
                            <mi>i</mi>
                        </msub>
                    <mo fence="true">)</mo>
                </mrow>
                <mrow>
                    <mo fence="true">|</mo>
                    <mi>I</mi>
                    <mo fence="true">|</mo>
                </mrow>
            </mfrac>
            <mo>+</mo>
            <mn>1</mn>
        <mo fence="true">]</mo>
    </mrow>

    <!-- Remainder are not displayed to the user -->
    <annotation encoding="application/x-tex">
        OPRiQ = \frac{1}{3}\cdot
         ↪  \ln\left[\left(e^3-1\right)\cdot\frac{\displaystyle
         ↪  \sum_{i\in I} \left( s^\prime_{i=j_i} \cdot r^\prime_i
         ↪  \right)}{\left|I\right|} + 1\right]
    </annotation>
</semantics>
</math>
```

APPENDIX

G

TEST DATASET GENERATION

Listing G.1 mk_datasets.py

```python
from random import choices
from pathlib import Path

BASE_DIR = Path(__file__).resolve().parent

with open(BASE_DIR / "opriq_item.csv") as f:
    SENSITIVITIES = [
        i.split(",")[2] for i in f.read().split("\n")[1:-1]
    ]

with open(BASE_DIR / "empty.csv", "w") as f:
    f.write("0")

with open(BASE_DIR / "full.csv", "w") as f:
    f.write(",".join(SENSITIVITIES))

for i in range(1,6):
    with open(BASE_DIR / f"random-{i}.csv", "w") as f:
        f.write(",".join(set(choices(SENSITIVITIES, k=30))))
```

APPENDIX

H

DERIVATIVE PROOF

Definition 1. $\overline{\Lambda}$ is a number s.t. $0 \leq \overline{\Lambda} \leq 1$.

Definition 2. α is a function of $\overline{\Lambda}$ s.t. $\alpha = \frac{1}{3} \cdot \ln\left((e^3 - 1) \cdot \overline{\Lambda} + 1\right)$.

Proposition 1.

$$\frac{\mathrm{d}^2\alpha}{\mathrm{d}\overline{\Lambda}^2} < 0.\ \forall\overline{\Lambda}$$

Proof.

$$\begin{aligned}
\frac{\mathrm{d}^2\alpha}{\mathrm{d}\overline{\Lambda}^2} &= \frac{\mathrm{d}}{\mathrm{d}\overline{\Lambda}}\frac{\mathrm{d}\alpha}{\mathrm{d}\overline{\Lambda}} && \text{(H.1)}\\
&= \frac{\mathrm{d}}{\mathrm{d}\overline{\Lambda}}\frac{\mathrm{d}}{\mathrm{d}\overline{\Lambda}}\left[\frac{1}{3} \cdot \ln\left((e^3 - 1) \cdot \overline{\Lambda} + 1\right)\right] && \text{(H.2)}\\
&= \frac{1}{3} \cdot \frac{\mathrm{d}}{\mathrm{d}\overline{\Lambda}}\frac{\mathrm{d}}{\mathrm{d}\overline{\Lambda}}\left[\ln\left((e^3 - 1) \cdot \overline{\Lambda} + 1\right)\right] && \text{(H.3)}
\end{aligned}$$

To begin, the question of the first derivative.

Lemma 1 (The Chain Rule)**.** For some function $f(x) = g\,(h(x))$, where g and h are differentiable, it holds that $f'(x) = g'\,(h(x)) \cdot h'(x)$.

Lemma 2.

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln x = \frac{1}{x}$$

Therefore – via Lemmas 1 and 2 – it holds that, for some differentiable function $j(x)$:

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln\left(j(x)\right) = \frac{\frac{\mathrm{d}j}{\mathrm{d}x}}{j\,(x)}$$

Therefore:

$$\frac{\mathrm{d}}{\mathrm{d}\overline{\Lambda}}\left[\ln\left((e^3-1)\cdot\overline{\Lambda}+1\right)\right] = \frac{\frac{\mathrm{d}}{\mathrm{d}\overline{\Lambda}}\left[(e^3-1)\cdot\overline{\Lambda}+1\right]}{(e^3-1)\cdot\overline{\Lambda}+1} \tag{H.4}$$

$$= \frac{e^3-1}{(e^3-1)\cdot\overline{\Lambda}+1} \tag{H.5}$$

Now the question of the second derivative.

Lemma 3 (The Quotient Rule)**.** For some function $f(x) = \frac{g(x)}{h(x)}$, where g and h are differentiable, it holds that:

$$\frac{g'(x)\cdot h(x) - g(x)\cdot h'(x)}{h(x)^2}$$

Therefore, using this Lemma 3:

$$\frac{\mathrm{d}}{\mathrm{d}\overline{\Lambda}}\left[\frac{e^3-1}{(e^3-1)\cdot\overline{\Lambda}+1}\right] = \left(e^3-1\right)\cdot\frac{\mathrm{d}}{\mathrm{d}\overline{\Lambda}}\left[\frac{1}{(e^3-1)\cdot\overline{\Lambda}+1}\right] \tag{H.6}$$

$$= \left(e^3-1\right)\cdot\frac{\frac{\mathrm{d}}{\mathrm{d}\overline{\Lambda}}[1]\cdot\left[(e^3-1)\cdot\overline{\Lambda}+1\right] - 1\cdot\frac{\mathrm{d}}{\mathrm{d}\overline{\Lambda}}\left[(e^3-1)\cdot\overline{\Lambda}+1\right]}{\left[(e^3-1)\cdot\overline{\Lambda}+1\right]^2} \tag{H.7}$$

$$= \left(e^3-1\right)\cdot\frac{0-\frac{\mathrm{d}}{\mathrm{d}\overline{\Lambda}}\left[(e^3-1)\cdot\overline{\Lambda}+1\right]}{\left[(e^3-1)\cdot\overline{\Lambda}+1\right]^2} \tag{H.8}$$

$$= \left(e^3-1\right)\cdot\frac{-e^3+1}{\left[(e^3-1)\cdot\overline{\Lambda}\right]^2+2\cdot\left[(e^3-1)\cdot\overline{\Lambda}\right]+1} \tag{H.9}$$

Therefore, without further simplification:

$$\frac{\mathrm{d}^2\alpha}{\mathrm{d}\overline{\Lambda}^2} = \frac{1}{3}\cdot\left(e^3-1\right)\cdot\frac{-e^3+1}{\left[(e^3-1)\cdot\overline{\Lambda}\right]^2+2\cdot\left[(e^3-1)\cdot\overline{\Lambda}\right]+1} \tag{H.10}$$

As the proposition to be proven only concerns "sign" of the result – being whether it is positive, negative, or zero – one can consider these categories. For the following, (+) represents a positive value, (–) a negative value, (+0) is a value that is either positive or zero, and (–0) is the opposite.

$$\begin{aligned}
\frac{\mathrm{d}^2\alpha}{\mathrm{d}\overline{\Lambda}^2} &= (+)\cdot(+)\cdot\frac{(-)}{(+0)^2+(+)\cdot(+0)+(+)} && \text{(H.11)}\\
&= (+)\cdot\frac{(-)}{(+0)+(+0)+(+)} && \text{(H.12)}\\
&= (+)\cdot\frac{(-)}{(+)} && \text{(H.13)}\\
&= (-) && \text{(H.14)}
\end{aligned}$$

Therefore, it follows that $\frac{\mathrm{d}^2\alpha}{\mathrm{d}\overline{\Lambda}^2} < 0.\ \forall\overline{\Lambda}$. QED