

Bjarte Skredderhaug

A Threat to Operational Security

Analysing Information Disclosure to Third Parties
in Apps Used by Soldiers Aged 19-22

Master's thesis in Information Security, Experience-based

Supervisor: Maria Bartnes

Co-supervisor: Vivi Ringnes Berrefjord

December 2023

Bjarte Skredderhaug

A Threat to Operational Security

Analysing Information Disclosure to Third Parties in
Apps Used by Soldiers Aged 19-22

Master's thesis in Information Security, Experience-based
Supervisor: Maria Bartnes
Co-supervisor: Vivi Ringnes Berrefjord
December 2023

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Dept. of Information Security and Communication Technology



Norwegian University of
Science and Technology

Abstract

The extensive collection of user data through common mobile apps is a growing security risk for Norwegian soldiers. The data is sold at a market for profit, making it accessible to adversaries and introducing threats beyond ethical, privacy, legal, and regulatory concerns. Furthermore, the ecosystem of user data is entangled and understanding who collects what and with whom it is shared can be difficult. Several studies have previously analysed apps to determine what data is collected and shared. However, the damage potential of the data is yet to be explored for the Norwegian Armed Forces.

This thesis examines four apps used by Norwegian soldiers between the ages of 19 and 22, namely Eurosport, Strava, Tinder and ASKfm. It evaluates the quantity and sensitivity of the data collected and shared before applying this to a threat analysis. Furthermore, it explores which countermeasures can be taken to minimise the negative potential of the data sales. This thesis adopts a three-part approach: a content analysis of the apps' privacy policies, a static analysis of the application packages, and a dynamic analysis intercepting and decrypting the network traffic from the apps during use.

The analysis discovered that the four apps shared sensitive data such as name, email address, country, user interactions, location, logs, gender, age, language, and network information with third-party domains. The threat analysis further revealed that several military assets are vulnerable to threats such as extortion, localisation, and disruption due to this data sharing, potentially compromising the security of the subset. In summary, this thesis exposes the digital footprint of a Norwegian soldier in the age group 19-22 years and the harmful potential of the data shared. It emphasises the importance of minimising these risks by guiding users, improving routines, and advocating stronger legislation, making it crucial for policymakers, military officials, and anyone concerned with data privacy and security.

Sammendrag

Den omfattende innsamlingen av brukerdata gjennom vanlige apper utgjør en økende sikkerhetsrisiko for norske soldater. Dataene selges i et marked for profitt, noe som gir potensielle fiender tilgang til sensitiv informasjon. Dette skaper trusler utover etiske, personvern-, juridiske og regulatoriske bekymringer. Økosystemet for brukerdata er komplisert, og det er vanskelig å finne ut hvem som samler hvilke data, og hvem det deles med. Flere studier har tidligere analysert apper i forsøk på å avdekke hvilke data som samles inn og deles. Imidlertid er dataens skadepotensial ennå ikke utforsket for Forsvaret.

Denne oppgaven undersøker fire apper som brukes av soldater i alderen 19 til 22 år, nemlig Eurosport, Strava, Tinder og ASKfm. Den evaluerer mengden og sensitiviteten til data som samles inn og deles, og anvender dette i en trusselanalyse. Videre utforsker den ulike tiltak for å minimere skadepotensialet til dataen. Oppgaven benytter seg av en tredelt tilnærming: først en tematisk analyse av appenes personvernserklæringer, deretter en statistisk analyse av applikasjonspakken, og til slutt en dynamisk analyse som avlytter og dekrypterer nettverkstrafikken fra appene.

Analysen avdekket at de fire appene delte navn, e-postadresse, land, brukerinteraksjoner, lokasjonsdata, logger, kjønn, alder, språk og nettverksinformasjon med tredjepartsdomener. Trusselvurderingen konkluderte videre med at flere militære ressurser er sårbare for trusler som utpressing, lokalisering og forstyrrelse som følge av denne datadelingen. Disse truslene kan påvirke sikkerheten til utvalget i denne oppgaven. Oppsummert avdekker denne oppgaven det digitale fotavtrykket til en norsk soldat i aldersgruppen 19-22 år, samt dataens skadepotensial. Den legger vekt på viktigheten av å redusere disse risikoene gjennom å instruere brukere, forbedre rutiner og oppfordre til bedre lovgivning, noe som er avgjørende for beslutningstakere, militære ledere og alle som er opptatt av personvern og datasikkerhet.

Preface

This thesis, completed as part of the Experience-based master's program in Information Security at the Norwegian University of Science and Technology (NTNU), represents a significant milestone in my academic journey.

I was drawn to this thesis's subject by my role as an employee in the Norwegian Armed Forces and my personal experiences as a user of some of these apps. Motivated to broaden my technical horizons, I explored unfamiliar applications and technologies, acknowledging that overcoming challenges is essential to growth.

I want to thank my dedicated supervisors, Maria Bartnes and Vivi Ringnes Berrefjord, for their excellent support and sparring throughout this project and for prioritising me during busy times. Both of you were deliberately chosen for your unique insight into this field of study. I also wish to acknowledge Tor Erling Bjørstad at Mnemonic for his valuable advice on data collection.

Lastly, I am grateful to my family and significant other for their unwavering support and adaptability. Their encouragement allowed me to immerse myself fully in the writing process during this period. Thank you for being there every step of the way.

Contents

Abstract	i
Sammendrag	iii
Preface	v
Contents	vii
Figures	ix
Tables	xi
Code Listings	xiii
Acronyms	xv
Glossary	xvii
1 Introduction	1
1.1 Background and Motivation	1
1.2 Research Objectives	2
1.3 Scope and Limitations	2
1.4 Contributions	3
1.5 Thesis Structure	3
2 Background	5
2.1 The Ecosystem of User Data	5
2.1.1 The Product	5
2.1.2 The Market	6
2.1.3 Key Legal Pillars	8
2.2 The Threat to Military Operations	10
2.3 Data Collection	11
2.4 The Lay of the Land: Existing Measures to Limit Data Collections on Norwegian Soldiers	15
3 Method	17
3.1 Choice of Methods	17
3.2 Selection of Apps	20
3.3 Content Analysis of the Apps' Privacy Policies	21
3.4 Static Analysis	22
3.5 Dynamic Analysis	23
3.6 The Methodological Limitations	29
4 Results	31
4.1 Content Analysis	31
4.2 Static Analysis	34

4.3	Dynamic Analysis	40
5	Analysis and Discussion	55
5.1	Findings from the Content Analysis	55
5.2	Findings from the Static Analysis	55
5.3	Findings from the Dynamic Analysis	56
5.4	RQ1: How Data Sharing Affects Security	58
5.5	RQ2: Potential Countermeasures	61
5.6	Limitations	64
6	Conclusions and Future Work	67
6.1	Suggestion for Further Research	67
	Bibliography	69
A	Permission from NATO StratCOM COE	77
B	Master Agreement	81
C	Test Setup	87
D	Analysis Results	91

Figures

2.1	Data flow of user data	6
2.2	Sequence diagram showing traffic between Grindr, MoPub and third parties. Adapted from [42]	13
2.3	Traffic dump from Grindr app [42]	14
3.1	Research methodology roadmap	18
3.2	One soldier's attributes	23
3.3	Test setup [10]	25
3.4	Google Pixel 2 [62]	26
4.1	Comparison of embedded trackers and permissions	39
4.2	Starting the Eurosport app	41
4.3	Certificate un-pinning of Strava APK using APK-MITM	44
4.4	Starting the Strava app	44
4.5	Starting the Tinder app	47
4.6	Starting the ASKfm app	49
4.7	Overview of transmissions origin	52
4.8	Third parties within the first hour	52
4.9	Apps third parties	53
5.1	Attributes observed during the experiment	59
5.2	ASKfm interrupting the collection from Strava	65
C.1	Images uploaded to the Google account and Tinder	90
C.2	Screenshots	90
D.1	SANKEY diagram presenting the origin of the third parties	102

Tables

2.1	NATO’s comparison of white and the black markets for data [19] . . .	8
2.2	The OSI model [40]	12
3.1	The apps chosen for this experiment	21
3.2	The data categories studied in the content analysis	22
3.3	The fictitious soldier’s attributes (Image was AI-generated)	24
4.1	Results from content analysis of the apps’ privacy policies	32
4.2	List of partners from ASKfm’s privacy policy [66]	34
4.3	Comparison of data explicitly mentioned in the privacy policy to be collected by the apps	34
4.4	Results from static analysis with <i>Exodus</i>	35
5.1	Risk matrix	62
C.1	Apps tested for selection	89
D.1	Objects of the traffic intercepted from Eurosport	93
D.2	Objects of the traffic intercepted from Strava	98
D.3	Objects of the traffic intercepted from Tinder	100
D.4	Objects of the traffic intercepted from ASKfm	101

Code Listings

3.1	Installing MobSF	22
3.2	Commands run in SDK environment	27
4.1	User-Agent header field	40
4.2	Traffic observed from Eurosport to <code>liftoff.io</code>	41
4.3	Traffic observed from Eurosport to <code>careers.bupa.com.au</code>	42
4.4	Traffic observed from Eurosport to <code>arkoselabs.com</code>	43
4.5	Traffic observed from Strava to <code>branch.io</code>	45
4.6	Traffic observed from Strava to <code>appsflyer.com</code>	46
4.7	Traffic observed from Strava to <code>iterable.com</code>	47
4.8	Traffic observed from Tinder to <code>bugsnag.com</code>	48
4.9	Traffic observed from ASKfm to <code>pollfish.com</code>	50
4.10	Traffic observed from ASKfm to <code>mopub.com</code>	51

Acronyms

- AI** Artificial intelligence. 60, 90
- API** Application Programming Interface. 12, 40, 43, 64, 65
- APK** Android Application Package. 12, 14, 18, 21, 22, 34, 38, 39, 40, 43, 56, 59
- DIE** Data Interception Environment. 12, 14, 25, 26, 27
- GDPR** General Data Protection Regulation. 8, 9, 31, 63
- GEOINT** Geospatial Intelligence. 10
- GPS** Global Positioning System. 11, 12, 14, 65
- HTTP** Hypertext Transfer Protocol. 12, 13, 40, 51, 53, 57, 64, 91
- HUMINT** Human Intelligence. 10
- MITM** Man in the Middle. 15, 25
- MobSF** Mobile Security Framework. 14, 22, 38, 39, 56
- NATO** North Atlantic Treaty Organization. 1, 7, 58, 68
- NCC** The Norwegian Consumer Council. 8, 12
- OPSEC** Operational Security. 16, 56, 61
- PSYOP** Psychological Operations. 1
- RQ** Research Question. 2, 3, 17, 19, 61
- SDK** Software Development Kit. 12, 14, 21, 22, 27, 55, 56
- SIGINT** Signals Intelligence. 10
- SOCMINT** Social Media Intelligence. 10
- TECHINT** Technical Intelligence. 10
- TLS** Transport Layer Security. 12, 40

Glossary

Asset	Refers to any valuable resource, information, or component that an organisation seeks to protect.
Risk	The product of the severity and the probability of an incident. Often a result of the interaction between threats and vulnerabilities.
Security	Measures, protocols, and practices implemented to safeguard assets, systems, or individuals from various forms of harm. This also encompasses information security, ensuring the confidentiality, integrity, and availability of information [1].
Threat	Any source of danger or harm that has the capacity and/or intention to compromise security. Threats can manifest in diverse forms, ranging from external actors with malicious intent to natural disasters or technological failures.
Vulnerability	Refers to a weakness or flaw in a system, process, or entity that can be exploited.

Chapter 1

Introduction

1.1 Background and Motivation

In 2020, NRK bought information on Norwegians' movements from a British firm. The data originated from apps installed by the users themselves. NRK paid only 35.000kr, demonstrating how cheaply and quickly the data available could map the lives of unknowns. Among those tracked were several groups requiring special protection and military personnel at Rena camp. The data was so specific that journalists could surveil areas belonging to the special forces and one individual at the Intelligence Service stations in Northern Norway [2].

Today, smart devices leave digital footprints everywhere. As of 2020, 100% of Norwegians between the ages of 16 and 24 have access to a smartphone [3]. Their tracks are collected and repackaged for sale at a global market, including data on which advertisements they prefer, their hobbies, or something as simple as their age. Over time, information that may seem innocent can reveal patterns and coherence. From this, analysis can identify individuals from anonymised data or disclose information about preferences and personality.

The ecosystem of user data is entangled, and smartphone users are exposed to a complex threat landscape. Finding out who collects what and where information is shared is difficult. Despite initiatives such as the EU's GDPR, users often accept ambiguous terms for quicker access to services, possibly due to the diffuse consequences on the individual level [4]. User data, therefore, is abundant and readily accessible through data brokers.

In a military context, a soldier's information can be aggregated to unveil troop activity, with location data enabling precise targeting, planning manoeuvres, and intelligence gathering. Data leakage could further expose operations, removing the element of surprise, endangering personnel and their families performing counterintelligence, social engineering, and Psychological Operations (PSYOP) manipulating, demoralising, or intimidating military personnel to influence their actions and opinions [5–7]. According to an experiment conducted by North Atlantic Treaty Organization (NATO) StratCOM COE in 2019, an adversary can gather enough personal data on a soldier to influence their behaviour using targeted

messages [8]. This emphasises the need to restrict information flow vulnerable to exploitation, benefiting both soldier privacy and Armed Forces security. As one of the most digitalised countries in the world, the Norwegian Armed Forces' guidelines are inadequate, often relying on recommendations from local experts, resulting in deviations between branches [9].

Continuation of Existing Work

Parts of the introduction, background, theory, and method chapter are adapted from the author's report in IMT4205 Research Project Planning, the precursor to the master's thesis course [10]. The overall topic is, among others, inspired by a suggestion for further research by Hannah Ersdal and Sølvi Svendby Skjærstad in their master's thesis [4].

1.2 Research Objectives

This thesis will examine the quantity and sensitivity of data collected by a limited number of apps. This insight will inform a threat analysis, identifying assets, vulnerabilities, and threats. The analysis hopes to tell what countermeasures can be taken to increase resilience and reduce the threat to the Norwegian Armed Forces. These objectives are synthesised into two Research Questions (RQs):

- RQ1:** How do the selected apps' data sharing with third parties affect the security of Norwegian soldiers aged 19-22?
- RQ2:** What countermeasures can the military implement to address potential threats arising from apps' data sharing with third parties?

1.3 Scope and Limitations

This thesis is limited to data shared with third-party domains from four apps, namely Eurosport, Strava, Tinder, and ASKfm. The scope is limited to finding the threat this data poses to Norwegian soldiers between 19 and 22 years in a generic military context. The actual risk may vary, as this thesis has limited knowledge of systems within the Norwegian Armed Forces. All social and personal consequences caused by the information disclosure will not be studied. Neither will ethical and legal implications.

The experiment is based on data generated by a fictitious soldier's activity on a chosen set of apps due to practical and ethical reasons explained in Section 3.5. Consequently, generalising the findings across the entire population is difficult. Furthermore, this thesis will focus only on data shared directly from the apps to third parties. However, data can also be shared across third parties or via the app developer without being detected in this experiment.

This thesis intends to examine apps used by this subset, not to find apps that share the most sensitive data. This thesis is limited to the Android operating system,

using a Google Pixel mobile phone. Variations may apply to other mobile operating systems such as iOS. Furthermore, not all shared user data will be discovered or fully interpreted.

1.4 Contributions

This master thesis examines a handful of apps to decide which information is collected and to whom it is shared. Next, the data is grouped, interpreted, and analysed to create a holistic profile of the total footprint of an average Norwegian soldier in the age group 19-22 years.

This profile helps to determine the severity of the data and gives a broader understanding of the problem at hand in a Norwegian context. It contributes insight into a specific population segment and a particular country's unique circumstances. As a small but exceptionally resourceful country with a highly digitalised population, Norway provides an interesting case in digital risk assessment.

The work results in a recommendation of measures to be taken by the Norwegian Armed Forces to prevent data collection and sharing with third parties.

1.5 Thesis Structure

The following chapters are included in this thesis:

Chapter 2 - Background presents relevant theory and previous work in this field. Furthermore, it creates a theoretical foundation upon which the analysis can be built. The goal is to understand the ecosystem of user data and grasp the nature of technology in military operations.

Chapter 3 - Method explains the methodology for selecting apps, the content analysis, the static analysis, and the dynamic analysis: the prerequisites, tools, and techniques. Furthermore, the scientific methodology applied to sort, analyse, and interpret the data will be presented. Lastly, the chapter expands on the methodology's limitations and the data's trustworthiness.

Chapter 4 - Results presents the experiment's results and attempts to objectively visualise and explain the data.

Chapter 5 - Analysis and Discussion uses aggregated data to gain a perspective on the findings and their relevance in a military context while answering the two RQs.

Chapter 6 - Conclusions and Future Work summarises and restates the thesis. Key arguments are presented, with a perspective of the findings and suggestions for future studies.

Chapter 2

Background

This chapter will outline the contextual factors of the research by providing a solid presentation of the key elements framing the ecosystem of the user data market. Firstly, it will present the product and the market in which it is commercialised. This is necessary as the research addresses a rather novel economic exchange popularised by Shoshana Zuboff's term Surveillance Capitalism [11], meaning where the monetisation of the user experience becomes the driving force of the digital domain.

Then, this chapter will outline key legal principles operationalised through GDPR before researching the threat and potential of user data through Christl Wolfie's two reports, 'Corporate Surveillance in Everyday Life' [12] and 'Networks of Control' [13]. Lastly, relevant work examining data collection will be presented before reviewing recommendations and guidelines. Together, this will form a context of existing work and its significance in a military context.

Much of the related work and relevant background material were identified in the project planning course preceding the master thesis [10]. These are included below, as well as a few papers found or published after the project.

2.1 The Ecosystem of User Data

This section aims to describe the ecosystem of user data by exploring three key dimensions: the product, the market, and the existing legislation.

2.1.1 The Product

'If you are not paying for it, you're not the customer; you're the product being sold.' (Lewis, [14]).

In this thesis, user data refers to personal information shared willingly or unwillingly through apps and services. Depending on the collection method, data could be grouped into knowingly shared, observed and analysed data. The term includes metadata whose properties are relevant but primarily concerns personal

information collected from primary, secondary, or tertiary sources [15]. From the time the user installs an application to the time their data has been sold is illustrated in Figure 2.1.

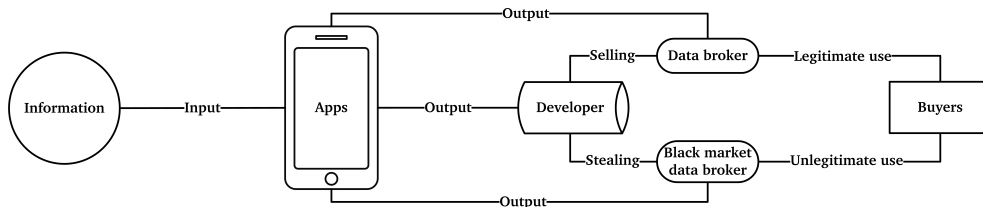


Figure 2.1: Data flow of user data

Surveillance capitalism is a market-driven process in which the commodity sold is users' data. Providers of complimentary online services, such as search engines and social media platforms, conduct extensive internet surveillance to acquire and generate personal data. This process involves the analysis of user online behaviours, encompassing but not limited to location- and demographic data, behaviour patterns, preferences, transactions, communications, and social media information. The resulting information serves commercial objectives, and individuals often remain unaware of the full scope of this surveillance [11].

An article published by Fleming *et al.* in the *Journal of Cybersecurity* discusses the potential costs, benefits and security risks posed to users sharing their personal information to commercial services. The article concludes that there is little correlation between users' privacy concerns and users' behaviour, referred to as the privacy paradox [16]. Often, users disclose information to services without considering the potential consequences of this data being sold [4]. Moreover, there is a pervasive lack of awareness that data stored in cookies on a trusted website can be sold to different actors years later.

The popular video-sharing app TikTok is a pertinent example, using an algorithm to recommend content for users. This algorithm allegedly bases its data on user interactions, including likes, scrolling patterns, and search history, to provide personalised content. This data is stored and possibly sold to primary- and third parties.

2.1.2 The Market

Dawson narrate that in the early days of the Internet, advertising paved the way to support platforms' ability to be 'free'. Customers then gave up certain data in exchange for access. In turn, these companies used the data for better target advertising to potential buyers [6].

Over time, companies realised the value of the data they possessed, leading them to explore opportunities for capitalisation. This trend saw pioneers such as Google and Facebook initiating efforts to monetise user data, a practice that has become increasingly common across the digital landscape. Today, many apps

and services share user data as part of their service provision or for optimisation through analytical insight [17].

User data has especially great value for the advertising industry, providing targeted advertisements that enable them to adapt to the market and achieve a more effective allocation of resources. Although the ethical implications of collecting and utilising user data are debatable, their use cases are usually legal. Furthermore, entertainment apps such as TikTok strive to provide content targeted to each user based on their preferences and interests. Credit card companies, mobile network firms and other services also use user data to establish a profile on each user to better understand their behaviour.

In addition to marketing and legitimate targeted communication, data brokers are known to sell information directly to criminals and rough actors [5]. The non-legitimate actors range from single criminals to nation-states. Their purpose may include intelligence operations, influence operations, targeted manipulation for financial or policy gain, and surveillance. An emerging market is a government entity using commercially available information to inform its decision-making. In 2017, the US Secret Service, the US federal police agency, reportedly paid \$2 million for access to the service Locate X, which aims to deliver location data collected from countless apps. Such information would typically require a court order, but outdated legislation allowed the purchase and use of data from such data brokers threatening human rights [18].

Today, the suppliers in the data market can be divided into first-party and third-party data brokers. Facebook and Google are examples of first-party data brokers, providing a service to users and then selling their data for profit. Third-party data brokers are not directly linked to the user and base their entire operation on buying or stealing data from first parties and then selling it to the highest bidder. The North Atlantic Treaty Organization (NATO) report 'Data Brokers and Security' provide a comprehensive framework for presenting the market of data brokers in a military context. As shown in Table 2.1 NATO, there are both legitimate and illegitimate third-party brokers.

Experian, CoreLogic, Epsilon, Acxiom, and LiveRamp are the most known data brokers. Experian for financial monitoring and is typically used for running background checks. CoreLogic for consumer information such as credit scores. Epsilon for behavioural data used for marketing. Acxiom for comprehensive data used for data portrait analysis. Lastly, LiveRamp specialises in offering a platform for data connectivity, a marketplace to buy and sell data.

However, in cases where journalists have tried to buy data about themselves, the data has often turned out to be highly inaccurate [20]. Thus, there is no guarantee for the quality of the data, making data reliability and validity challenging to verify for potential buyers.

Table 2.1: NATO's comparison of white and the black markets for data [19]

Characteristic	White market	Black market
Specialization	In most cases, traded data is used to understand and subsequently influence a target group.	As-a-service models unique to each customer; customised malware. Uses are likely to be illegal.
Market size by revenue	Approximately \$200 billion	Approximately \$1.5 trillion
Reliability	Both markets engage in questionable practices concerning legality and service. Sellers' reputations are vital for the business.	
Government regulation and impact	Government in constant confrontation with brokers, arguing that they operate on the verge of law.	Government in an arms race with cybercriminals, unable to curb the continual rapid expansion of the black market.
Sophistication	Technological innovation is a key factor in market development, the goal of which is to collect and effectively analyse as much information as possible.	Continuous market-driven improvements in security, anonymity, and more sophisticated ways to please customers.
Competition	Highly competitive market dominated by established strong players.	Easy to start a business in the ever-expanding black market; success depends on reputation and skills.

2.1.3 Key Legal Pillars

Many aspects could be included when addressing the legal framework for data brokers. This thesis is limited to two aspects. The first, General Data Protection Regulation (GDPR), is a comprehensive legal framework guiding the collection of personal data in the EU and Norway. The second aspect, the EOS committee's 2022 remarks regarding security services' opportunity to buy commercial intelligence, is highly relevant for this thesis as it addresses the critical question of militarising user data for national security purposes.

EU saw the challenges of enhancing data privacy in 2012 and spent four years developing what is known as GDPR, a European framework for data protection. In short, it aims to give the consumer ownership over their data, including the right to know what information is collected, access it, and correct or delete incorrect or unnecessary data. As mentioned by the NCC, when they read app terms for 32 hours straight, no one can expect the average user to read long and complex privacy policies for each app installed on their phone [21]. The legal concept of consent under GDPR requires that the users receive clear and easily understandable information about what they are consenting to.

GDPR also incorporates the aspect of data accountability, enforcing specific

rules on how organisations can collect, process, and store data. This also includes requirements for information security and reporting of privacy breaches. GDPR covers all personal data, meaning any information that can be used to identify an individual. This includes name, financial data, IP-, postal-, email address, health information, and more. This regulation was implemented in Norway on 20 July 2018, incorporating national rules with Norwegian adjustments in the ‘Lov om behandling av personopplysninger’ [22].

A legal framework is just as effective as its implementation. In this case, it is the company’s responsibility to comply with the legislation. Large tech companies have a mixed track record, and jurisprudence is still immature. However, some exciting processes are underway. As of 14 August 2023, Meta Platforms - which, among others, owns Facebook, WhatsApp, and Instagram were issued a daily fine of one million Norwegian crowns (NOK) by the Norwegian Data Protection Authority for misuse of user data. They had until 4 August 2023 to comply with the regulations and prove that they no longer use behavioural advertisement by targeting based on users’ location [23]. Meta now offers a paid subscription to stop receiving personalised ads and consequently stop the collection of data previously used to target ads [24]. The fulfilment of this pledged anonymity is yet to be seen.

Also, ChatGPT was banned in Italy in March 2023, and the Italian Data Protection Authority accused it of unlawfully collecting users’ data. They were given 20 days to sort the issue and successfully comply with EU’s GDPR, this shows that GDPR offers tools for the regulators to initiate change at the companies [25]. In August 2022, Google was fined A\$60 million in Australia for misleading Android users by ignoring device settings regarding collecting and storing location data [26]. Although there are equivalents, no nationwide GDPR law exists in countries outside of the EU. This effectively allows companies such as Facebook to treat personal data for some user groups as they please, affecting Norwegian soldiers abroad in personal and professional affairs.

The previous subsection, ‘The Market’, showed how the security apparatus might want to use the vast commercially available information to inform their work. It is therefore interesting to note that a report published by ‘Stortingets kontrollutvalg for etterretnings-, overvåkings- og sikkerhetstjeneste’ (EOS Committee) raises the question of whether buying metadata in large chunks violates the Intelligence Service Act [27] about intelligence gathering on citizens. The EOS Committee is a perpetual organisation appointed by the Norwegian Parliament. Its rationale is to control Norwegian agencies conducting intelligence, surveillance, and security services, guarding against the infringement of rights, use of excessive force, safeguarding the public interest and ensuring lawfulness.

The Norwegian Intelligence Service believes the method used for acquiring the data is decisive in whether they follow the law. The EOS Committee’s central question was, however, whether they were in possession of personal data intervening in an individual’s privacy. Precedent from The European Court of Human Rights (ECHR) shows that methods with such intervention require a foundation in law [28].

Open-source intelligence usually falls under ‘common liberty to act’ [29, pt. 8.5.4]. However, considering the amount of data from multiple sources, there is a risk of de-anonymising the user. Even though the user has willingly shared the information, the EOS Committee stated that buying chunks of data must be considered an intervention in an individual’s privacy. Therefore, the Norwegian Intelligence Service was advised to review its method and legislation [28].

2.2 The Threat to Military Operations

The threat and potential of user data are further explored in the report ‘Networks of Control’. In this report, Spiekermann and Christl studies the potential of analysing personal data to predict or even de-anonymise and re-identify individuals from anonymised data sets. These data sets could be anything from visited websites and browser fingerprints to keystroke and mouse dynamics [13].

Fundamental personal attributes such as name, gender, birth date, postal address, and ZIP code remain significant. Combining just two or three of these attributes can distinctly identify individuals with relatively high confidence [12]. Recent events demonstrate the re-identification of individuals using Netflix data, biometric data, or even just four apps installed on a user’s smartphone by cross-referencing with publicly available databases [13].

According to an article published in the *International Journal of Intelligence and CounterIntelligence* written by Caton, the traditional approach to gathering Technical Intelligence (TECHINT) involved combining Human Intelligence (HUMINT), Signals Intelligence (SIGINT), and Geospatial Intelligence (GEOINT) to fully interpret the adversaries’ technological capabilities. This manual approach proved to be insufficient, creating misleading or false intelligence. However, TECHINT has since evolved into an effective source of intelligence that can fill gaps left by the other types of intelligence [30].

In today’s smartphones and social media world, a new intelligence field called Social Media Intelligence (SOCMINT) has emerged. This field has proven cost-effective and relevant in gathering information from social media platforms. In 2023, a declassified report from the Office of the Director of National Intelligence outlines how the intelligence community may benefit from commercially available information and its subsequent commercially sourced intelligence [31, 32].

While the majority of collected user data is typically utilised for marketing purposes [33], military personnel, specifically those aged 19-22 in the Norwegian Armed Forces, face not only exposure to targeted marketing campaigns but also the potential for more immediate and critical consequences within a military context. In response to these complex challenges, Twetman and Bergmanis-Korats devised a risk taxonomy to facilitate a more structured and efficient understanding and management of the diverse risks. This taxonomy describes the resources, personnel, equipment, information, facilities, and activity and their vulnerability and threat. Below is a summary of their findings [5].

Personnel are a critical resource in any organisation but are naturally vulnerable to manipulation using personal information and preferences. Location data can present threats such as blackmail, blacklisting, doxing, and identity theft. User data collected from apps can also facilitate financial gain through sensitive information, such as identity theft, extortion, phishing, and credit card fraud. Financial consequences, however, fall outside the immediate threats this thesis focuses on. Twetman and Bergmanis-Korats suggests that foreign states purchase user data rather than gather their own intelligence, as this can be both more efficient and cost-saving [5, 10].

Furthermore, military equipment is becoming increasingly technological and brings with it new vulnerabilities. For example, the device ID can be correlated with personal information and geolocation to map usage patterns, communication infrastructure, and the equipment itself. Such information can further be used for intelligence on capacity and capability, manipulation, and sabotage [5, 10].

Lastly, military activities may be exposed through geolocation data and personal information collected from social media. Exploitation introduces the danger of disruption and interruption [5]. For example, GPS was first used in the Gulf War during Operation Desert Storm. It is now known as ‘the first space war’ in Kuwait and Iraq in 1991. This technology provided a revolutionary tactical advantage, providing a tremendous situational understanding and enabling precise manoeuvres on the battlefield [34]. Today, this data is particularly sensitive, as seen in Ukraine GPS data leaked to the enemy is acted on immediately, providing accurate and real-time targeting [35].

2.3 Data Collection

Brandtzaeg *et al.* studied in the article ‘Losing Control to Data-Hungry Apps: A Mixed-Methods Approach to Mobile App Privacy’ 21 popular social networking-, dating-, and fitness apps available on Android in Norway. They conducted a representative survey on Norwegian users’ privacy concerns. Their study showed that more than half of the participants had avoided downloading apps due to reluctance to share personal information. Furthermore, their analysis concluded that 19 of the 21 apps tested shared personal information to primary- and third-party domains, most of which are in the US. Their recommendation was for developers to include visualisations for the user to increase the transparency of personal data flows from mobile apps [36].

In 2019, Grundy *et al.* investigated the sharing of user data from top-rated mobile apps related to medicine for clinicians and consumers. Using traffic, content, and network analysis, they found that 79% of the sampled apps shared user data. The findings, however, are limited to medical apps, and the relevance to this thesis is the methodological framework, where a traffic analysis of the apps and, subsequently, a content and network analysis characterised the third parties [37].

Khatoon and Corcoran’s report on ‘Privacy concerns on Android devices’ demonstrates that Android offers Linux Kernel security on their operating system. Each

application must ask for permission through their Application Programming Interface (API) to access resources such as camera, microphone, Global Positioning System (GPS), network information, Bluetooth, user data and more. This is coded in the Android Application Package (APK) and can be modified using the Software Development Kit (SDK). According to the report, users can not know the consequences of permissions given under apps' installation, as most people have little understanding of privacy [38].

In December 2018, Privacy International, a London-based charity working for the global right to privacy, published a report showing how Android apps share data with Facebook. What was especially interesting was that data was shared, even though the users did not have Facebook. Almost 61% of all apps tested sent data to Facebook once opened, regardless of whether the Facebook app was installed on the phone. Furthermore, the data was detailed and sometimes sensitive. The test's setup was later published to help others repeat the experiment or continue the work. It was named Data Interception Environment (DIE). The setup consists of a virtualised computer based on the Kali Linux operating system, which captures and analyses the packets transmitted from the phone to the internet [39].

The OSI model in Table 2.2 conceptualises how systems are interconnected. The reference model is divided into seven different abstraction layers [40]. Hypertext Transfer Protocol (HTTP) is an application-level protocol for encoding and transporting information between clients and servers. Transport Layer Security (TLS) are often adopted to secure communication and encrypt transmissions between web applications and servers. A custom root certificate must be installed on the mobile device to bypass the TLS encryption. Furthermore, the traffic is sent through a HTTP proxy where packets can be intercepted in a readable format. Proxying refers to acting on behalf of someone else, in this case, routing traffic to an intermediary acting as the recipient.

Table 2.2: The OSI model [40]

Application layer
Presentation layer
Session layer
Transport layer
Network layer
Data link layer
Physical layer

Based on the same setup The Norwegian Consumer Council (NCC) in collaboration with Mnemonic, published in 2020 a report named 'Out of Control'. It examines how consumers are exploited by the AdTech industry, resulting in a record 65 million NOK fine for the dating app Grindr for illegally sharing personal information [33, 41]. The report was supplemented with a technical report published by Mnemonic, explaining how they collected and analysed app traffic data.

Mnemonic writes that the traffic data came in large volumes. As it was difficult to analyse and present understandably, they limited themselves to an ‘app-centric perspective’, studying to which parties app X transmit data. Mnemonic described the traffic as utilising different types of interaction and data elements. The interactions they observed were API integration, complex flows, or as part of resource fetch and tracking pixels. API calls usually send HTTP GET or POST to a third-party domain. In complex flows, the app typically calls a sequence of third-party endpoints, resource fetch, where the app requests a resource and sends data along with it, or data transmitted with a tracking pixel request. As for the data elements, Mnemonic gave a quick list of typical data elements collected, including advertising ID, IMEI, IP address, MAC address, GPS location, device information, device configuration, WiFi networks, app name, account information and lastly, user data [42].

One interesting phenomenon observed by Mnemonic is the use of intermediaries. Mnemonic’s concern was that MoPub, one of Grindr’s most important advertising partners mentioned in the privacy policy, acted as an intermediary, enabling sharing from Grindr to multiple third parties. The initial response presented in Figure 2.3 was followed by a response from MoPub with a JSON object that Grindr ran. The content specified where the app should send its subsequent request, in this case, to AppNexus. AppNexus is another advertisement provider that is not mentioned in Grindr’s privacy policy. This phenomenon can be critical in this thesis, as the official advertising suppliers can be planned for when stated in the privacy policy. On the other hand, unknown third parties are more challenging to consider. The timeline in Figure 2.2 is an adaptation of a sketch from Mnemonic.

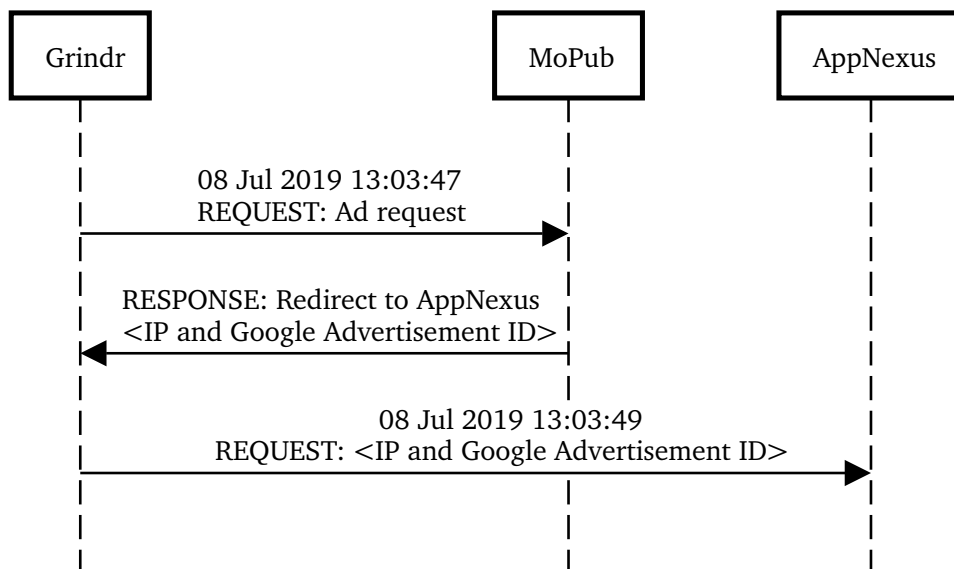


Figure 2.2: Sequence diagram showing traffic between Grindr, MoPub and third parties. Adapted from [42]

Figure 2.3 is adapted from Mnemonic's report 'Out of Control' and demonstrates a request from the Grindr app to MoPub, taken from the traffic dump. This is the opening message and contains several user data, for example, accurate GPS position, Android advertisement ID, app name and the user's sex and age. This is the data Privacy Internationals' DIE will seek to capture during the experiment [10].

```

POST /m/ad HTTP/1.1
accept-language: en-us
user-agent: Mozilla/5.0 (Linux; Android 8.1.0; Nexus 5X
Build/OPM7.181205.001; wv) AppleWebKit/537.36 (KHTML, like Gecko)
Version/4.0 Chrome/75.0.3770.101 Mobile Safari/537.36
Content-Type: application/json; charset=UTF-8
Host: ads.mopub.com
Connection: close
Accept-Encoding: gzip, deflate
Content-Length: 600

{
  "ll": "REDACTED, REDACTED",
  "vv": "0",
  "nv": "5.4.1",
  "dn": "LGE, Nexus 5X, bullhead",
  "sc": "2.625",
  "consented_vendor_list_version": "0",
  "current_consent_status": "explicit_yes",
  "consented_privacy_policy_version": "0",
  "id": "agltb3B1YilpbmNyDQsSBFNpdGUYtebmEgw",
  "udid": "ifa:52d0d5c2-e923-4b1b-bd67-d3b225795edb",
  "bundle": "com.grindrapp.android",
  "llsdk": "1",
  "gdpr_applies": "1",
  "lla": "51",
  "mr": "1",
  "llf": "138693",
  "h": "1920",
  "force_gdpr_applies": "0",
  "dnt": "0",
  "android_perms_ext_storage": "0",
  "o": "p",
  "q": "app_version:5.12.1",
  "ct": "2",
  "av": "5.12.1",
  "v": "6",
  "w": "1080",
  "z": "+0200",
  "user_data_q": "m_gender:m,m_age:34"
}

```

Figure 2.3: Traffic dump from Grindr app [42]

Together with Privacy Internationals' DIE, the tools *Exodus* [43] and Mobile Security Framework (MobSF) [44] can be used for static analysis of the apps' Android Application Packages (APKs). Each of these is well documented in Github. *Exodus* can be used to initially sort out the apps to investigate by looking at the required permissions to run and the number of third-party Software Development Kits (SDKs) embedded in the APKs, also known as trackers. MobSF can be used as a supplement to Privacy Internationals' DIE by analysing the application code. APK-MITM, a function integrated into DIE, can bypass countermeasures from the

developers, such as certificate pinning or anti-reversing techniques.

2.4 The Lay of the Land: Existing Measures to Limit Data Collections on Norwegian Soldiers

Previous research has investigated the sharing of user data from apps to third parties and how effective recommendations and guidelines are to govern users and their actions. However, most related work has studied apps from the perspective of privacy concerns, primarily with children, medicine, dating, or health apps in mind. It lacks the situational context of crisis and war. Furthermore, it does not study the age group 19-22, which in Norway includes the majority of infantry soldiers, the backbone of the Norwegian Armed Forces. Regardless, valuable insight and experience can be found.

For example, a study conducted by Shibchurn and Yan in 2014 attempted, through a survey with over 300 participants, to investigate how a financial reward affects users' willingness to share personal information. They concluded that users are more willing to share information if rewarded, but that it depends on the type of information and that false information is shared more frequently [45]. Research by Xiong *et al.* further studied the willingness to share personal information if certain privacy protection systems were in place. Their findings were that 62% of the participants decided to share information based on the protection mechanisms, 26% shared because it would be helpful for the service, and 22% with a lack of concern for their privacy. The majority who did not want to disclose personal information meant the information was too sensitive to share 37%, distrusted the protection mechanism 33%, and the last 30% worried about the risk of leakage through breaches or hack [46].

A parallel can be drawn to soldiers' need to participate on social platforms. Experts advise against the use of multiple services during military activity. However, it is still commonplace to locate the enemy during exercises through apps such as Snapchat and Tinder [47]. Some soldiers are not willing to exclude themselves from social platforms and choose to take risks with such services.

Van Kleek *et al.* researched using the same Man in the Middle (MITM) framework how revealing apps' collection of user data affected users' willingness to use their services and disclose personal information. Their study showed that users made different decisions with their updated permission interface and felt more confident about their choices [48]. This shows that a study on what data is collected from apps used by Norwegian soldiers may provide some changes in behaviour if presented correctly.

Except for the US Special Operations Command, the lack of guidance within the military to protect users' data from collection is also problematised by Dawson in 'Microtargeting as Information Warfare' [6]. She concludes that US soldiers have no existing policies or directives on removing their data from common databases, avoiding insecure email services or warning about the risk of installing Facebook

on personal devices. ‘There is no way for any individual to tackle the surveillance economy’ and collective efforts are necessary [6].

On that accord, Twetman and Bergmanis-Korats suggest five essential learning points for NATO: Firstly, recognise that consciousness is necessary but insufficient. Secondly, view data as critical infrastructure. Thirdly, control the data through training and standard operating procedures. Fourthly, use red teams to test practices and systems to identify and understand risk. And finally, harness the potential of the data in the battlefield [5]. These recommendations are based on a market analysis of raw data from information exchanges in Latvia. The recommendations aimed to ensure mission resolution for the allies and accommodate the risk with commercially available data [10].

The Norwegian Armed Forces has a security authority called the ‘Norwegian Defence Security Department’. In 2020, they published a policy on the use of social media when in service [49]. This policy, however, is ten pages long. While the security authority is responsible for the development, leaders in each department are responsible for ensuring that all employees are familiar with it. The Armed Forces recently stated a restriction on the app TikTok [50]. Until 1 December 2023, this recommendation only applies to work phones, and the effectiveness can be discussed.

The Armed Forces’ guidelines are primarily based on recommendations from local security advisers, who, during the first months of conscription, hold a 30-minute brief on the use of social media for the soldiers. A brief that typically scratches the surface of how problematic the use of certain apps would be in an actual war, without any follow-up [9]. Generally, this works for a short period before soldiers fall back into old habits. This practice also creates deviations across departments, depending on the awareness of the department’s managers and the organisation’s security culture.

This thesis aims to reduce the gap between research focused on analysing traffic data and research on data’s potential in a military context. Specifically, it is aimed at the Norwegian Armed Forces and concentrates on identifying the most severe threats to their soldiers. Hopefully, this thesis will aid in developing sound guidelines for young Norwegian soldiers and ensure better OPSEC in the future.

Chapter 3

Method

This chapter will address the methodology employed for this thesis, how data was collected, filtered, and analysed, and why these methods were appropriate and can produce relevant data for answering the RQs. Please note that this chapter is partially a continuation of the research project planning report written prior to this thesis [10].

3.1 Choice of Methods

Research Type

This thesis adopted an inductive approach to research, where theories and conclusions are derived from observations. This is opposed to a deductive approach, where theory guides data collection. Furthermore, this study was descriptive. It focused on the current state of the user data ecosystem to understand how the Norwegian Armed Forces can protect itself from leaking sensitive user data.

The experiment can further be classified as an observational study, applying a mixed-method research design. It was an observational study because the data was collected through monitoring and measuring actual behaviour rather than collecting data through methods such as a questionnaire or interview [51]. It used a 2-phase mixed-method design called explanatory sequential design, first performing a quantitative collection and analysis, then following up with a qualitative collection and analysis.

The first part of the experiment was quantitative as it collected measurable, quantifiable data in the form of traffic volume from an Android phone. To analyse the data and determine its significance in a military context some qualitative methods were necessary. It was, therefore, appropriate to apply a mixed-method research design. Together, these methods helped answer the two RQs.

Research Strategy

Maintaining objectivity and not seeking confirmation that apps illegally share user data was imperative. This thesis addressed a topic not previously explored in any subject of the master's degree program at NTNU, necessitating a thorough literature search. Furthermore, sound research methodology dictates thorough documentation of tools and techniques used in the experiment. Lastly, a content analysis of the intricate privacy policies of the apps was conducted to identify the data shared with third parties and their recipients. This was essential to understand the foundation on which users can make an informed decision, and how it compares to the observed network traffic.

Experiments as a method can be expensive and time-consuming and present ethical as well as practical challenges. In this thesis, however, there were enough advantages to outweigh these inconveniences. Incorporating a realistic experiment with controlled observation of data traffic improved the understanding of how data was collected. One holds greater control over the variables, minimises sources of error and increases the validity of the results compared to acquiring second-hand information. All covariates cannot be controlled, so this experiment had to consider some variation.

Roadmap

Figure 3.1 shows the stages of the methodology applied in this thesis. First, a content analysis was conducted on the apps' privacy policies. Then, a static analysis of the APKs. Lastly, the apps were analysed dynamically in a controlled environment. The apps were examined sequentially to create the most similar testing environment possible. Due to all the dynamic covariates in the setup causing differences in each collection, data was stored for later filtration and analysis.

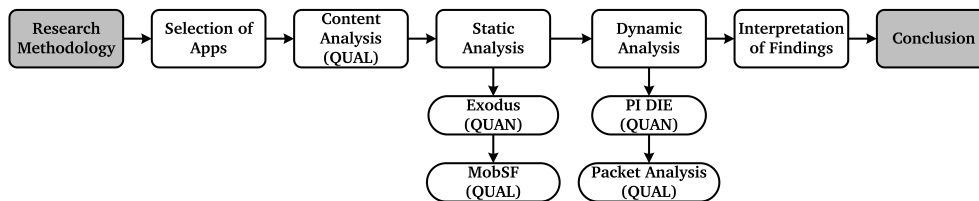


Figure 3.1: Research methodology roadmap

Sampling Strategy

As of 2022, the Norwegian Army had 8,463 employees, where 3,417 were military, 453 civilian and 4,593 conscripts [52]. Consequently, most Norwegian infantry forces are youth who serve for the first time, usually within the first three years of completing upper secondary school. Therefore, The majority will be aged 19 to 22, the subset chosen for this thesis. This group was also the largest consumer of

smartphones and apps in Norwegian society [3], and with their short tenure, has the most civilian mindset.

The selection of which apps to analyse was important, as a non-representative selection of apps could give an incorrect or skewed picture of reality. Findings would then not be representative. However, a probability sampling would have required some trustworthy data on the app usage of the chosen subset, which was unavailable at the time of writing. The chosen sampling method, therefore, became purposive sampling, a non-probability sampling technique that allowed for picking the apps with the highest probability of leaking data. One cannot guarantee that the apps selected represented the entire group of Norwegian soldiers aged 19-22. Hence, the findings may not be suitable for broad generalisations. However, the apps are known to be popular among this group, and any findings will demonstrate what is achievable and may, therefore, contribute to analysis and recommendation.

Theoretical Framework

In preparation for this thesis, a comprehensive review of existing literature was conducted, which included documentation, articles, and legal documents such as privacy policies.

The focus was on peer-reviewed English-language materials, complemented by select Norwegian articles, to establish a solid theoretical foundation and formulate relevant RQs. Much of the literature was from the industry itself, which has been particularly useful from the technical and military points of view. However, as such literature typically is written by employees from within the sector, special attention was paid to maintaining objectivity. It was essential to be aware of potential confirmation bias, particularly in cases where the literature described the marketing industry as immoral and redundant.

As a part of the literature review, the following journals were browsed for relevant articles: *Intelligence and National Security* [53], *International Journal of Intelligence and CounterIntelligence* [54], *Journal of Global Security Studies* [55], and *Journal of Cybersecurity* [56]. The main search database for this thesis was Bibsys's Oria. This search database includes search results from known databases such as IEEE Xplore, Science Direct, ACM, Gartner, and Springer Link, together with NTNU's library. Furthermore, digital search tools such as Google.com and Google Scholar supplemented Oria. A continuous evaluation of sources was necessary to ensure their trustworthiness. ChatGPT [57] has been used for inspiration, summarising literature, and rephrasing own text for better clarity.

Boolean operators were used with the following search strings: apps, user data, sharing, privacy, third parties and smartphone. Combining these strings with AND and OR operators yielded more precise search results, allowing for better sorting. Results were further filtered on peer-reviewed literature from the last five years to exclude outdated material due to the fast-paced development of the digital domain. Additionally, it was essential to be conscious that some literature may be confidential due to security.

3.2 Selection of Apps

The result of the experiment must be reproducible using the same procedure, which allows others to control and verify the findings. Producing data with reliability requires a standardised procedure. This thesis, therefore, strove to carry out a systematic selection of applications and analyse each application as equally as possible. The procedure also affects the validity of the findings. This experiment was therefore conducted as demonstrated in Figure 3.1. Furthermore, this thesis tried to use dual-tool verification whenever possible to ensure the validity of the different tools. Sparring with an employee at Mnemonic before conducting the experiment helped ensure that this thesis observed the most relevant data traffic.

The choice of apps would greatly influence the experiment. Finding credible sources for app demographics and statistics on the ages 19-22 in Norway quickly proved difficult. Google Play store provided some statistics without the ability to sort on population. Second-hand information was widespread but had little or no foundation in market analysis or statistics from Google. The selection, therefore, was based on the apps prone to disclosing information sensitive in a military context. Behavioural patterns and personal preferences were not as relevant as location data. Furthermore, the apps needed to be used by this thesis's subset.

The first step was to select potential apps to be examined which, preferably, stored information such as gender, religion, preferences, and health information, as this data poses the greatest threat in the event of a leak. The selection was further filtered using the Exodus Privacy tool [43] v1.28. This web-based service automatically unpacks application packages to provide an overview of embedded trackers and permissions required from the app. This analysis indicated which apps would be interesting to study further. Potential apps were then sorted into six main categories:





- | | |
|--|--|
| <p>1. Entertainment
Games, Photo, Video, and Music.</p> | <p>4. Social Networking
Communication, Dating,
Media-sharing, and Forums.</p> |
| <p>2. Lifestyle
Health, Training, Shopping,
Travelling, and Navigation.</p> | <p>5. News and Information
Magazines, Weather, and Sports.</p> |
| <p>3. Productivity
PDF-viewer, Notetaking,
and Education.</p> | <p>6. Finance and Utility
Banking, Transportation,
and Tools.</p> |

Although most apps collect data, some categories above were more prone to sharing information with third parties. As seen in the initial scan (Appendix C Table C.1), some categories had distinctively more trackers and permission requirements embedded than others. The last ten versions of the apps were considered. The app with the most embedded trackers and permissions was installed on the phone.

The entertainment apps studied were Scrabble, Wordle, Candy Crush Saga, Eurosport, TV2 Play, YouTube, NRK TV, Viaplay, Netflix, Discovery, Amazon Prime, TikTok, and Spotify. The lifestyle category included apps such as Sleep Cycle, Headspace, Strava, Finn.no, Temu, Coop and Tise. In productivity, Adobe Acrobat, Goodnotes, Notion, Duolingo, Adblock Plus, Grammarly, Teams and Trello. The social networking group included Facebook, Messenger, WhatsApp, and Signal communication apps. The dating apps were Tinder, Happn, and Bumble. Media sharing and forums: Snapchat, Instagram, VSCO, ASKfm, BeReal, Pinterest, Reddit, and Discord. Norwegian news and weather apps such as VG, DB, NRK, TV2, Yr.no, VG Pent.no, and Storm were studied in the news category. Not surprisingly, commercial newspapers such as VG and Dagbladet were most extensive with many third-party SDKs and permissions embedded in the APK, as these are for-profit entities. Lastly, finance and utility included the apps DNB, Sparebank 1, Vipps, Uber, Ruter, Flashlight, and Calculator.

The next step was to map out in which categories users provide the most personal information. This led to excluding utilities, news, and productivity and moving forward with entertainment, games, lifestyle, and social networking. Strava quickly came to attention partially due to its media coverage in 2018 when ministers', soldiers', and key personnel's movements were exposed. Scanning for embedded trackers using *Exodus*, the following four apps presented in Table 3.1 were selected for further study.

Table 3.1: The apps chosen for this experiment

App	Name	Developer	Category	Downloads
	Eurosport	Eurosport	Entertainment	10,000,000+
	Strava	Strava Inc.	Lifestyle	50,000,000+
	Tinder	Tinder	Social Networking	100,000,000+
	ASKfm	Ask.fm	Social Networking	50,000,000+

3.3 Content Analysis of the Apps' Privacy Policies

The analysis of privacy policies was conducted to employ triangulation, which is finding separate data sources to validate and identify correlations with the results from the experiment. Furthermore, such an analysis would help identify recurring themes and patterns within policies to gain insights into the data-sharing practices of the selected apps. The following steps were undertaken to systematically study and extract relevant information from the apps' privacy policies concerning data collection and to whom it was shared.

This analysis was guided by two primary criteria: (1) explicit data collection descriptions and (2) recipients of shared information. The privacy policies of these

selected apps were retrieved directly from the respective app stores or official app developer websites. The newest versions of privacy policies were obtained for a representative analysis. The data stated to be shared was categorised into basic account information, technical information, profile, activity, use information, and other sources of information to identify patterns and variations among the selected apps. Lastly, the results from the content analysis were compared to the actual data flows to see any divergence between the terms consented to by the user and traffic observed during testing. This method provided a comprehensive understanding of data-sharing practices outlined in privacy policies, laying a foundation for subsequent study phases.

Table 3.2: The data categories studied in the content analysis

Category	Content
Basic account information	Name, gender, addresses, etc.
Technical information	IP address, device ID, login, etc.
Profile, activity, and use information	Location, time zone, clicks, etc.
Other sources of information	Surveys, market studies, etc.
Who the information is shared with	Third parties, law enforcement, etc.

3.4 Static Analysis

Static analysis was performed using *Exodus Privacy* tool and *MobSF*. *Exodus* scanned the APK for embedded third-party SDKs and required permissions, primarily for classifying the apps during selection but also for indicating what traffic to look for in the dynamic analysis. This tool is a web application and does not require any download or preparation.

MobSF offers more advanced capabilities for scanning and analysing the APKs. This program has a framework for binary-, security-, malware-, and reconnaissance analysis. It enables the examiner to dive deep into the foundations of the app. This program also offers capabilities for dynamic analysis and penetration testing; the latter were omitted in this thesis due to their lack of relevance as this thesis set out to describe apps' behaviour rather than test their security. The *MobSF* environment needed to be virtualised on the OS level in containers. The easiest way was to install Docker on the client and run the commands shown in Code listing 3.1.

```
docker pull opensecurity/mobile-security-framework-mobsf:latest
docker run -it --rm -p 8000:8000 opensecurity/mobile-security-framework-mobsf:
latest
```

Code listing 3.1: Installing *MobSF*

MobSF was then accessible via web-browser on this URL: <http://127.0.0.1:8000>. The apps in their respective versions were downloaded from <https://www.apkmirror.com>.

3.5 Dynamic Analysis

A dynamic analysis was performed to observe the apps' behaviour objectively and contribute to the triangulation.

The Profile

The experiment was based on one fictitious soldier. As demonstrated in Figure 3.2, one soldier has various attributes. These attributes tell something about the user's location, features, or actions. Listing all information related to an individual would not be convenient or feasible. Going clockwise on the illustration, this experiment focused on the attributes' basic info, gender, technical details, language, location, time zone, interactions, camera, gallery, contacts, payment info, logs, and health.

As shown in Table 3.3, relevant information was made up to see later which data was collected and observed in network packets.



Figure 3.2: One soldier's attributes

Traffic Capturing

This thesis considered four approaches. Firstly, a conceptual solution founded upon the data stated to be shared in the app's privacy policy, followed by the necessary presumptions regarding what data could be shared with third parties. Secondly, requesting own data collected by apps. Thirdly, create a fake user and generate data. Or lastly, purchase data sets from a data broker. The latter would quickly have become disproportionately expensive and, in this context, unethical as the data is likely to be genuine and contain sensitive information about undisclosed third parties. Using one's own data would also present the same privacy concerns. Therefore, this thesis's approach was to create a fake profile and extract data using a technical solution. At the advice of Dr Tor E. Bjørstad, the choice fell on Privacy International's technical framework.

Table 3.3: The fictitious soldier's attributes (Image was AI-generated)

General Information

Name	Axel Norland
Age	22
Gender	Male
Location	Oslo (59.9133301,10.7389701)
Nationality	Norwegian
Phone number	+4790909090

Employment Status

Occupation	Soldier
Employer	The Norwegian Armed Forces
Job Description	Infantry troops
Income Level	Lower-middle class
Education Level	Upper secondary school

Marital Status

Marital Status	Single
Number of Children	None

Personal Information

Date of Birth	01.07.2001
Height	180cm
Weight	80kg
Hair Colour	Blonde
Eye Colour	Blue
Skin Colour	White

Preferences

Sexual Preferences	Straight
Music Preferences	Rock
Food Preferences	Nut allergy
Hobbies	Motorsport
Political Affiliation	Labour Party
Religion	Christian

Online Profile

Email Address	axelnorland@gmail.com
Social Media Usernames	anorland
Password	Isolable321
IP Address	100.64.32.61
Time zone	CEST
System	Google Pixel 2 (Android)

Analysing traffic data from the apps required a setup. This thesis used the Privacy Internationals' Data Interception Environment [58] as a starting point, which captures traffic going from the app on the mobile to the central server using a method called Man in the Middle (MITM). In practice, one pretends to be the device's counterpart during communication and reads the data traffic between the devices. The setup is shown in Figure 3.3 and consisted of internet access, a computer running the virtual environment and a mobile, in this case, a virtualized Android phone [10].

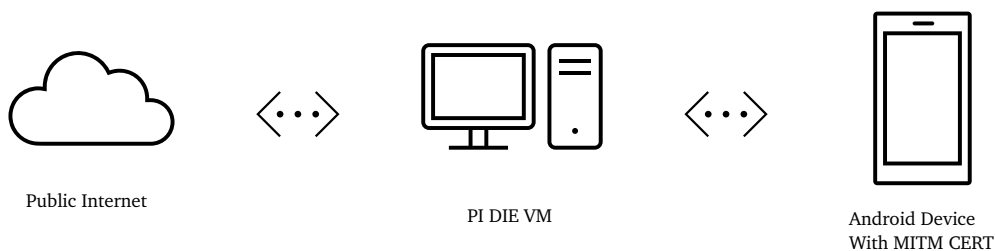


Figure 3.3: Test setup [10]

This study's independent variable was the extent of data sharing between selected apps and third parties. The dependent variable was the security of Norwegian soldiers aged 19-22. The study hypothesised that variations in the independent variable (data sharing) would influence the dependent variable (security). By monitoring and measuring the amount of traffic sent, one could analyse whether there were any connections between these variables.

The greatest challenge was apps that took active steps to prevent analysis. This can be done by certificate pinning, where the other party's certificate or public key is known and expected. It would, therefore, be difficult to pretend to be the other party. Another method is jailbreak detection, a technique where the app examines the environment in which it was run. Typically, this is applied as an anti-forensic measure when the developer wants to prevent the application from running in a debugging environment. In cases of certificate pinning, Mnemonic used the software Frida [59], but this required considerable effort [10].

The test was limited to Google's Android operating system. Primarily because it is the most used mobile operating system in the world [60], but also because Google is one of the most significant players in the advertising industry [61]. Furthermore, Android has a more open system architecture, which enables analysis. However, this does not mean that the findings were necessarily limited to the applications on Android [10].

The primary benefit of this setup was that it did not produce false positives. This was because all traffic captured was legitimate and can be traced back to its origin. The data was easily accessible and could be stored in different formats. Moreover, it can all be virtualised and requires no unique hardware. The significant drawback was that some traffic was likely not generated or captured. The setup was, therefore, still vulnerable to false negatives. Furthermore, there was a limited

set of test devices and physical locations where testing was performed. Additionally, some security protocols denied insight, and some apps could have complicated analysis by implementing hardening techniques.

DIE was run from VirtualBox along with the Google Pixel phone shown in Figures 3.4a and 3.4b, which was emulated using a software called Genymotion. Both these programs were run on MacOS using a MacBook Pro 2020. The choice fell on this phone and software because it was not resource-demanding and could be performed using most personal computers. Additionally, the Google Pixel is a simple phone, with Android architecture allowing the user more rights than typical smartphones.

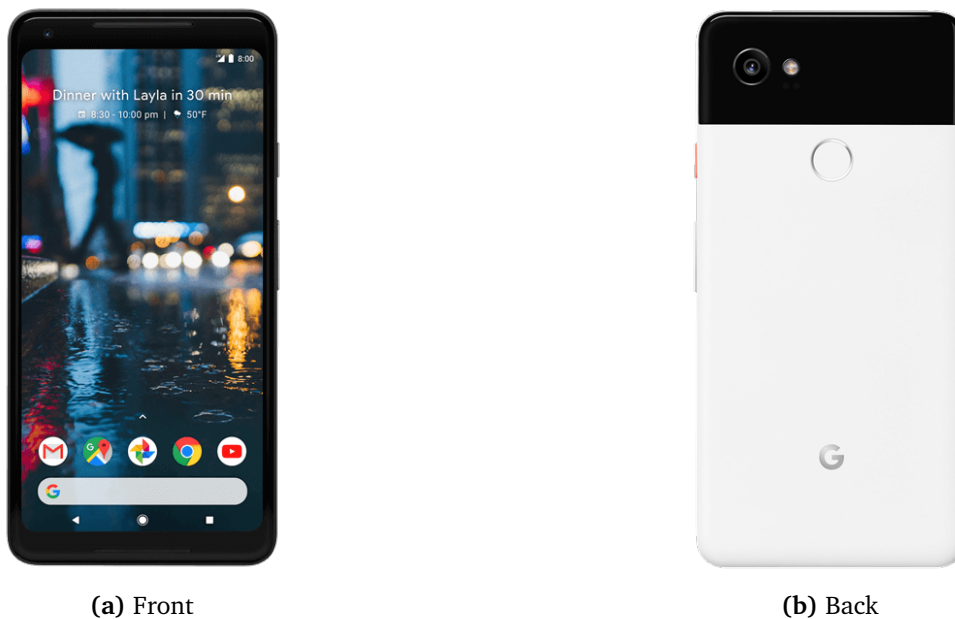


Figure 3.4: Google Pixel 2 [62]

The setup process was relatively simple. To achieve the same test environment as in this thesis, follow the steps below.

Setup of DIE:

1. Downloaded all .ova files from GitHub <https://github.com/privacyint/appdata-environment-desktop/releases>
2. Downloaded VirtualBox <https://www.virtualbox.org/wiki/Downloads>
3. Downloaded archiving software to unzip the files
4. Imported the .ova files into VirtualBox
 - a. Assigned desired resources to the computer
 - b. Changed network adapter from NAT to the internal network where traffic should be captured

Setup of Android Phone:

1. Downloaded Genymotion <https://www.genymotion.com/download/>
2. Generated a virtual Google Pixel 2
 - a. Assigned desired resources, in this thesis, the phone was reduced to two processors and 2048MB memory for better performance and less lag
 - b. Android version 12.0.0
 - c. Image version 3.0.2
3. Downloaded Open G Apps, equal to Google Play Store
4. Downloaded Chrome browser and the apps to be tested
5. Turned off system and apps auto-update function
6. Shut down the phone and changed network adaptor from NAT to internal
7. Connected to WiFi and verified the assigned IP address
8. GPS location was spoofed to be in the centre of Oslo

The Testing

1. On Android: Took a snapshot to get a known state for restoring after each testing
2. On DIE: Started and opened MITMProxy on localhost:8081
3. On Android: Opened browser and accessed mitm.it
 - a. Verified the interception and logging via MITMProxy
 - b. Downloaded the certificate for Android devices
4. On local machine: Downloaded the Android SDK Platform Tools and ran the commands in Code listing 3.2 to setup the certificate.

```

cd /Desktop/platform-tools
./adb devices
./adb shell
cd /sdcard/Download

In DIE:
sudo openssl x509 -inform PEM -subject_hash_old -in /root/.mitmproxy/mitmproxy-ca-
cert.pem | head -1

su
mount -o remount,rw /
cp /sdcard/Download/mitmproxy-ca-cert.pem /system/etc/security/cacerts/c8750f0d.0
chmod 644 /system/etc/security/cacerts/c8750f0d.0
chown root:root /system/etc/security/cacerts/c8750f0d.0
mount -o remount,ro /
reboot

```

Code listing 3.2: Commands run in SDK environment

Traffic was captured from the first start of the application plus one hour, including user registration and testing every app feature. The capture was stopped as the app was closed. However, data were produced continuously, so the collection

could not be done over several days. The amount of data would be too great. The aim was, therefore, to capture opening requests from the app to third parties and then over normal usage of the app for up to an hour. All the testing was done on the same day to ensure as equal a test environment as possible, including using the same manually installed certificate on the Android phone.

The procedure thus captured a snapshot of the phone, tested traffic interception, started a new capture, opened the application, registered using a Google account, provided the necessary information, tested every feature of the app, and let it run for one hour.

Traffic Filtering

Mnemonic captured 88 155 data transmissions in total across 216 different third-party domains. This shows a proper filtering mechanism was needed for the experiment to be completed within a reasonable time. Their advice to future researchers was to store log data in a readily accessible format. Filtering could include specific protocols, the app name, bundle ID, or advertising ID.

The two possible approaches were searching for information linking a packet to information disclosures or identifying and excluding irrelevant traffic established from a standby pattern. Either way, some software with filtering and searching capabilities must be utilised.

After the packets were captured and stored in the custom .mitm file format, attempts were made to export the raw data capture to .HAR or .CSV for further processing. However, the software did not provide this functionality and would have to be coded manually. The solution became to host a new MITMWeb session locally and import the stored captures. This made reading, filtering, and sorting the captured network traffic possible. The file could then be manually exported into JSON format for further processing.

The traffic captures were isolated entirely between the apps, which made it possible to avoid cross-contamination. Furthermore, the software allowed searching using regular expressions. This made it easy to find certain transmissions between known parties.

Traffic Analysis

The final step was to carry out an analysis of the data that had been collected. This analysis used quantitative and qualitative methods to determine the quantity and sensitivity of data collected. Firstly, a quantitative comparison of the amount of data shared with third parties was conducted. The amount of data shared did not necessarily connect with the number of transmissions to the various third parties. Therefore, the volume of transmissions to third parties was not examined. This method produced a data visualisation that could help identify patterns or trends, thus facilitating further interpretation of the findings. Secondly, qualitative in-depth analysis of the content of the traffic data determined which data was sent

to third parties. Applying mixed methods enabled a perspective on the topic from multiple angles, thereby enhancing the analysis's validity and quality.

The quantitative analysis laid the foundation for the qualitative analysis by organising the preliminary data. This sequence also enabled visualisation of the communication patterns, which increased the readers' understanding of user data flow within the ecosystem. This analysis used descriptive statistics to present the data clearly, as the goal was not to make inferences about larger populations such as inferential statistics.

As part of the qualitative analysis, one goal was to create a holistic profile of a Norwegian soldier based on the data obtained in the experiment. Data included activity time, location, personal information, communication, actions, and other relevant data that provided insight into the soldier's behaviour, patterns, and preferences. The qualitative approach helped uncover nuances, complexities and contextual factors that cannot necessarily be captured by purely quantitative data collection.

3.6 The Methodological Limitations

The chosen sampling method was prone to selection bias, potentially affecting the selection of apps. This could be personal biases overestimating the importance and use of certain apps or excluding some user groups. Consequently, this method could have influenced the representativeness of the findings.

The behaviour generated on the apps likely did not reflect the use of an authentic user within the subset. However, the behaviour was similar across all apps, which allows for accurate comparison. The Google Pixel 2 has an ARM x86 processor, further limiting the selection of apps considerably.

The manual analysis performed in this thesis was time-consuming. Studying the interactions between the applications and third parties was overwhelming. Thus, vital information may have been overlooked.

Some traffic was likely not generated or captured. Therefore, the setup remained vulnerable to false negatives. Furthermore, there was a limited set of test devices and physical locations where testing was performed. Some security protocols denied insight, and some apps could have implemented hardening techniques, making it challenging to analyse.

The methods used in this thesis relied on collecting technical data, which required functional hardware and software and an analyst with sufficient technical skills. Before this thesis, a risk assessment in the form of a risk matrix was made to identify the different risks for this thesis. A plan could be made to counter these risks and their influence on the result by pointing out the weak points of accomplishing this experiment.

The greatest challenge was traffic data, which was too difficult to decode and understand. Not comprehending what was shared through obfuscation or concealment indicates how actors operate in the grey zone.

Chapter 4

Results

Four apps were examined using the three methods: content-, static- and dynamic analysis. Firstly, the result from the content analysis is presented. Then, the static and dynamic analysis results with support from tables, diagrams, and screen captures. The analysis, interpretation, and evaluation of the result will be spared for Chapter 5's analysis and discussion.

4.1 Content Analysis

This section describes the results of the qualitative content analysis. GDPR requires the developers to create a privacy policy along with their apps, informing users on what and how they collect information [22, ch. III]. Companies must describe how they store and share information [22, ch. IV]. Table 4.1 compares what data is stated to be collected from the chosen apps and to whom it is shared. The app's entangled privacy policies are synthesised and attempted to be organised comprehensively. Only explicitly mentioned information is included.

The privacy policies studied are effective from the dates in the second row. Consequently, other researchers can analyse the same documents by following the citation and viewing the published history from the developer's websites. However, all companies reserve the right to change their privacy policies at any time. The analysis was carried out ultimo September 2023.

Table 4.1: Results from content analysis of the apps' privacy policies

Eurosport 2018-09-20 [63]	Strava 2023-06-30 [64]	Tinder 2023-02-24 [65]	ASKfm 2023-03-23 [66]
Which information is collected			
Basic account information			
<ul style="list-style-type: none"> • Name • Gender • Date of birth • Email address • Favourite sport, teams, and athletes • Marketing options • Language • Billing address • Payment info 	<ul style="list-style-type: none"> • Name • Gender • Date of birth • Email address • Username • Weight • Payment info 	<ul style="list-style-type: none"> • Gender • Date of birth • Phone number • Email address • Sexual orientation • Bio • Interests • Pictures/videos • Financial details 	<ul style="list-style-type: none"> • Name • Gender • Birth year • Phone number • Email address • Username • Password • Language • Postal address • Payment info
Technical information			
<ul style="list-style-type: none"> • IP address • Device identifier • Login information • Browser type and version • Browser plugins • OS and platform • Page response times • Download errors 	<ul style="list-style-type: none"> • Browser • Computer • Mobile device • Network info • Cookies • Analytic info • Internet service provider (ISP) • Referring/exit pages • Date and time • Number of clicks 	<ul style="list-style-type: none"> • IP address • Device ID • Device type • Apps settings and characteristics • App crashes • Advertisement ID • Identifiers associated with cookies • Web beacons, pixels, SDKs 	<ul style="list-style-type: none"> • IP address • Device identifiers • Browser type • Operating system • Platform type • Domain names • Error logs • Cookies • Referring/exit pages • Landing pages
Profile, activity, and use information			
<ul style="list-style-type: none"> • Location • Time zone • Websites visited • Products viewed and searched for • Length of visits to certain pages • Page interactions • Methods used to browse away • Video consumption and playback • Sign-in and out • Playback error • Browsing info 	<ul style="list-style-type: none"> • Geo-location • Date and time • Speed • Pace • Perceived exertion • Health info: <ul style="list-style-type: none"> ◦ Heart rate ◦ Power ◦ Cadence ◦ Weight • Content shared by the user • Contacts info 	<ul style="list-style-type: none"> • Precise geo-location • Photo verification data • Login, features used, actions, info presented, referring webpages address and ads • Interactions with other members • Info about other people • Customer care communication 	<ul style="list-style-type: none"> • Date and time • Number of clicks • Pages viewed • Time spent on pages • Info you provide, post, or allow us to access • Data collected on the user's friends: <ul style="list-style-type: none"> ◦ Contacts ◦ Last Name ◦ Email ◦ Location
Other sources of information			
<ul style="list-style-type: none"> • Surveys or competitions • Public sources • Third parties 	<ul style="list-style-type: none"> • Connected devices and apps • Other users • Third parties 	<ul style="list-style-type: none"> • Surveys, market studies, promotions, and events • Other members • Third parties 	
Who the information is shared with			
<ul style="list-style-type: none"> • Compliance with laws and legal proceedings • Third-party service providers • Group companies • Merger or acquisition • Social media platforms 	<ul style="list-style-type: none"> • Other users • Public information • Law enforcement • Public or governmental agencies • Private litigants • Service providers • Targeted advertising • Third parties 	<ul style="list-style-type: none"> • Other members • With law enforcement • To enforce legal rights • Service providers and partners • For corporate transactions • Affiliates • Sharing functionality 	<ul style="list-style-type: none"> • Other users • Public information • Law enforcement • Service providers • Affiliates • Third parties

Eurosport states that it collects and shares anonymised information but promises that no individual can be re-identified from information collected from their sites. Furthermore, they state they do not sell personal information to third parties. However, they also say that their website will contain links to and from partner networks, and they have no control over how these partners will use collected personal information.

Tinder collects and shares user data. However, they disclose that they do not sell users' personal data or use it for targeted advertising or profiling. Still, they mention in their privacy policy that they do not respond to the 'Do Not Track' option available for users. They justify this with the fact that not all browsers yet support this feature.

Strava from their privacy policy, Strava did reserve the right to use, sell, license, and share anonymised and aggregated user data. Strava is somewhat ambiguous in its privacy policy. Except for law enforcement and government agencies, they do not state actors with whom they share information. However, they say, 'We do not sell your personal information for monetary value' [64]. Aggregated data includes equipment, usage, demographics, routes, performance, and challenge participation and completion.

ASKfm declares that all anonymised information about users is not subject to their privacy policy and can be disclosed freely. However, they also say they do not sell or share personal information with third parties. Still, if users log in via third-party social networks, information will be collected from these services.

'We may obtain additional information about you from affiliates and may combine that information with information which we collect from or about you and information derived from any other product or service we or our affiliates provide.' [66]








































Similar to Tinder, ASKfm also ignores browser-initiated 'Do Not Track' signals, with the same arguments as justification. ASKfm is, however, the only app that states which information is shared with the different third parties, as shown in Table 4.2. Adsquare, Circulate, and Pangle are all AdTech companies. CleverDATA and Zeotap, on the other hand, focus on user analytics.

Table 4.2: List of partners from ASKfm’s privacy policy [66]

Company	User group	Data shared
Adsquare	Global	Age, gender, interests, and device IDs/cookies
Circulate	USA, CAN, FRA, and UK	Email (encrypted), and device IDs/cookies
CleverDATA	Global	Age, gender, interests, and device IDs/cookies
Zeotap	Global	Device IDs, age, gender, email (encrypted)
Pangle	Global	Age, gender, interests, and device IDs/cookies

Table 4.3 further illustrates and compares which data is stated to be collected from the different apps in the privacy policies. As shown, most attributes focused on in Chapter 3 are indeed collected when registering for the apps.

Table 4.3: Comparison of data explicitly mentioned in the privacy policy to be collected by the apps

App	Basic info	Gender	Tech details	Language	Location	Time zone	Interactions	Camera	Gallery	Contacts	Payment info	Logs	Health
													
													
													
													

4.2 Static Analysis

Running static analysis with *Exodus* provided information on embedded trackers and permissions in the APKs. Table 4.4 shows the result from the first app in Table 4.4a to the last in Table 4.4d. The left column lists the embedded trackers, and the right column lists the permissions required by the app.

Table 4.4: Results from static analysis with *Exodus*

(a)

Eurosport Version 5.41.0 [67]

Trackers	Permissions
ABTasty	ACCESS_NETWORK_STATE
AdColony	ACCESS_WIFI_STATE
Adincube	INTERNET
AerServ	WAKE_LOCK
AppLovin (MAX and SparkLabs)	SET_ALARM
Appnext	C2D_MESSAGE
AppsFlyer	RECEIVE
Chartbeat	BIND_GET_INSTALL_REFERRER_SERVICE
ChartBoost	
ComScore	
Conviva	
Demdex	
Facebook Ads, -Analytics, -Login	
Facebook Places, -Share	
Flurry	
FreeWheel	
Google AdMob	
Google Analytics	
Google CrashLytics	
Google Firebase Analytics	
Google Tag Manager	
IAB Open Measurement	
Inmobi	
Integral Ad Science	
JW Player	
Millennial Media	
Moat	
myTarget	
New Relic	
Nielsen	
Ogury Presage	
OpenTelemetry (OpenCensus, OpenTracing)	
Teads	
Twitter MoPub	
Unity3d Ads	
Vungle	
39 trackers	8 permissions

(b)

Strava Version 154.9 [68]

Trackers	Permissions
Adjust	ACCESS_FINE_LOCATION
Branch	ACCESS_MEDIA_LOCATION
Bugsnap	ACCESS_NETWORK_STATE
Facebook Analytics	ACCESS_WIFI_STATE
Facebook Login	ACTIVITY_RECOGNITION
Facebook Places	BLUETOOTH
Facebook Share	BLUETOOTH_ADMIN
Google AdMob	BODY_SENSORS
Google CrashLytics	FOREGROUND_SERVICE
Mapbox	GET_ACCOUNTS
	GET_TASKS
	INTERNET
	READ_CONTACTS
	READ_EXTERNAL_STORAGE
	RECEIVE_BOOT_COMPLETED
	REORDER_TASKS
	USE_CREDENTIALS
	VIBRATE
	WAKE_LOCK
	WRITE_EXTERNAL_STORAGE
	BILLING
	RECEIVE
	BIND_GET_INSTALL_REFERRER_SERVICE
	READ_GSERVICES
	C2D_MESSAGE
10 trackers	25 permissions

(c)
Tinder Version 14.17.0 [69]

Trackers	Permissions
Amplitude	ACCESS_COARSE_LOCATION
AppsFlyer	ACCESS_FINE_LOCATION
Branch	ACCESS_NETWORK_STATE
Bugsnag	ACCESS_WIFI_STATE
Facebook Analytics	BLUETOOTH
Facebook Login	CAMERA
Facebook Share	CHANGE_WIFI_STATE
Google AdMob	FOREGROUND_SERVICE
Google CrashLytics	INTERNET
Google Firebase Analytics	MODIFY_AUDIO_SETTINGS
IAB Open Measurement	POST_NOTIFICATIONS
Tinder Analytics	READ_CONTACTS
Unity3d Ads	READ_EXTERNAL_STORAGE
	READ_MEDIA_IMAGES
	READ_MEDIA_VIDEO
	READ_PHONE_STATE
	RECEIVE_BOOT_COMPLETED
	RECORD_AUDIO
	STORAGE
	USE_BIOMETRIC
	USE_FINGERPRINT
	VIBRATE
	WAKE_LOCK
	WRITE_EXTERNAL_STORAGE
	BILLING
	RECEIVE
	BIND_GET_INSTALL_REFERRER_SERVICE
	AD_ID
	READ_GSERVICES
	DYN_RECEIVER_NOT_EXP_PERMISSION
	MAPS_RECEIVE
13 trackers	31 permissions

(d)
ASKfm Version 4.75 [70]

Trackers	Permissions
AppMetrica	ACCESS_NETWORK_STATE
AppMonet	ACCESS_WIFI_STATE
Appnext	CAMERA
AppsFlyer	FOREGROUND_SERVICE
Criteo	INTERNET
Facebook Ads	READ_APP_BADGE
Facebook Analytics	RECEIVE_BOOT_COMPLETED
Facebook Login	VIBRATE
Facebook Share	WAKE_LOCK
Google AdMob	WRITE_EXTERNAL_STORAGE
Google CrashLytics	UPDATE_COUNT
Google Firebase Analytics	BILLING
Huawei Mobile Services (HMS) Core	BROADCAST
IAB Open Measurement	ACCESS
myTarget	RECEIVE
Pangle	BIND_GET_INSTALL_REFERRER_SERVICE
PubNative	READ_SETTINGS
Smaato	UPDATE_SHORTCUT
Smart	CHANGE_BADGE
Twitter MoPub	READ_SETTINGS
Unity3d Ads	WRITE_SETTINGS
Verizon Ads	UPDATE_BADGE
Vkontakte SDK	READ_SETTINGS
Vungle	WRITE_SETTINGS
Yandex Ad	READ
	WRITE
	BROADCAST_BADGE
	PROVIDER_INSERT_BADGE
	msa
	BADGE_COUNT_READ
	BADGE_COUNT_WRITE
25 trackers	31 permissions

After analysing all the APKs, the apps are compared in Figure 4.1. Eurosport exhibits the highest number of integrated trackers, nearly four times more than Strava. However, Eurosport also requires fewer permissions from the device than Tinder and ASKfm. A closer examination of Table 4.4a through Table 4.4d reveals that several of these permissions are unambiguous and necessary for the apps to function as intended. Nevertheless, allowing an app access comes with risks. The dynamic analysis demonstrates that data can be transmitted to numerous third parties.

Static analysis was also conducted using the software MobSF. However, this analysis did not result in any new significant findings. The results correlated well with *Exodus* and worked well as a dual-tool verification. Since the result

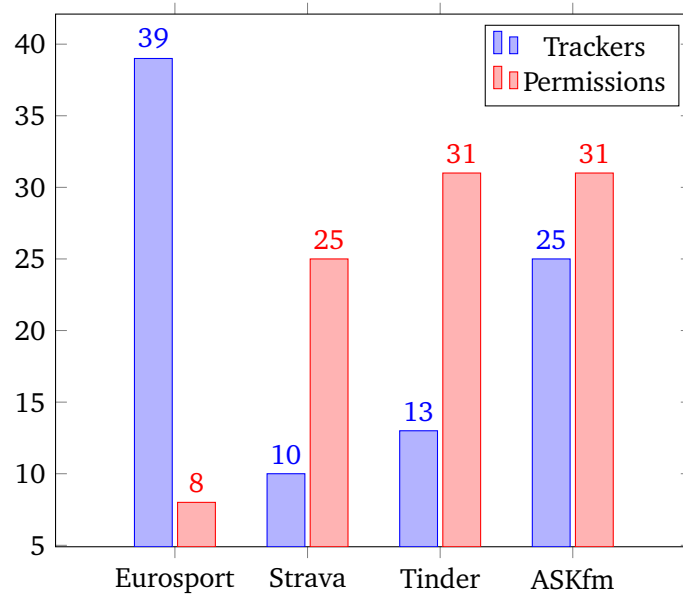


Figure 4.1: Comparison of embedded trackers and permissions

corresponds with *Exodus*, it is not presented twice to avoid redundancy. However, a short note of the results is included below.

For **Eurosport**, the same permissions were found using MobSF. The APKID analysis found some indications of anti-VM code and obfuscating code in `classes.dex`, `classes2.dex`, and `classes3.dex`. `classes4.dex` also contained the anti-debug code `Debug.isDebuggerConnected()`. The network security of the app was overall adequate but had some poor configuration to internal domains, and the third-party domains `itaipu.rugbyrama.fr`, `sdk.adincube.com`, `*.scorecardresearch.com` allowing clear text traffic to flow between them. The MobSF analysis found 30 trackers embedded in the APK, compared to *Exodus*'s 39. The trackers missing were AdColony, Appnext, ChartBoost, Flurry, Inmobi, Millennial Media, myTarget, Unity3d Ads, and Vungle.

Strava showed all the same embedded permissions as with *Exodus*. Strava also had some indications of anti-VM code in `classes.dex`, `classes4.dex`, and in `classes5.dex` but did not have any code for obfuscation or anti-debugging. Network security allowed clear text traffic between the app and `connect.garmin.com`. MobSF found nine trackers, missing Google CrashLytics compared with the *Exodus* scan.

MobSF was unable to analyse **Tinder**'s APK, even after trying other architectures and versions of the app, using both the online version and locally hosted version of MobSF.

Running **ASKfm**'s APK through MobSF provided the same embedded permissions as *Exodus*. This APK had in `assets/audience_network.dex` the anti-debug code `Debug.isDebuggerConnected()`. In `classes.dex` anti-VM code `Build.SERIAL`,

-.FINGERPRINT, -.MODEL, -.MANUFACTURER, -.PRODUCT, -.HARDWARE, SIM operator, network operator, ro.harware, ro.kernel.quemu, in classes2.dex Build.TAGS, and in classes3.dex Build.BOARD, subscriber ID. ASKfm allows for some clear text traffic between internal domains. Lastly, it found 24 trackers embedded in the APK, compared to 25 with *Exodus*, leaving out Huawei Mobile Services (HMS) Core.

4.3 Dynamic Analysis

This section presents the results from the dynamic part of the experiment. Running the apps on a phone while intercepting the network traffic from the app to its recipients. This quantitative method observes and compares the number of third parties for each app. However, as addressed in Chapter 3, the transmission volume to each third-party domain was not examined. Additionally, this experiment was intended to include a qualitative analysis of each traffic packet. The intention was to use a proxy and a rooted Android device installed with a custom certificate to bypass TLS and capture traffic in a readable format. This was mostly successful, but some packets remained unreadable due to the app's security measures.

Please note that all communications to the app domain are excluded from the dynamic analysis and do not appear among the results, as this thesis only examines third-party interactions. Traffic to third parties was monitored for one hour. The analysis found that sensitive data was shared from all apps, but primarily to the developer's own API. Furthermore, only the body of the HTTP packets are studied, and information transmitted in other ways was not captured. Packing information in the HTTP header is one way, as this contains information including user-agent, language, fetch-site, and connection. All graphs are drawn using the tools Excel and RawGraphs in combination with the Latex package TikZ. The 'User-Agent' header field appears below in Code listing 4.1 as this field is similar in all transmissions.

```
User-Agent: Mozilla/5.0 (Linux; Android 12; Pixel 2 Build/SQ1D.220205.004; wv)
AppleWebKit/537.36 (KHTML, like Gecko) Version/4.0 Chrome/91.0.4472.114 Mobile
Safari/537.36
```

Code listing 4.1: User-Agent header field

Eurosport was the first app to be tested. Figure 4.2 shows the initialisation screen of the app. The first time this app was started, they alerted us of an updated privacy policy.

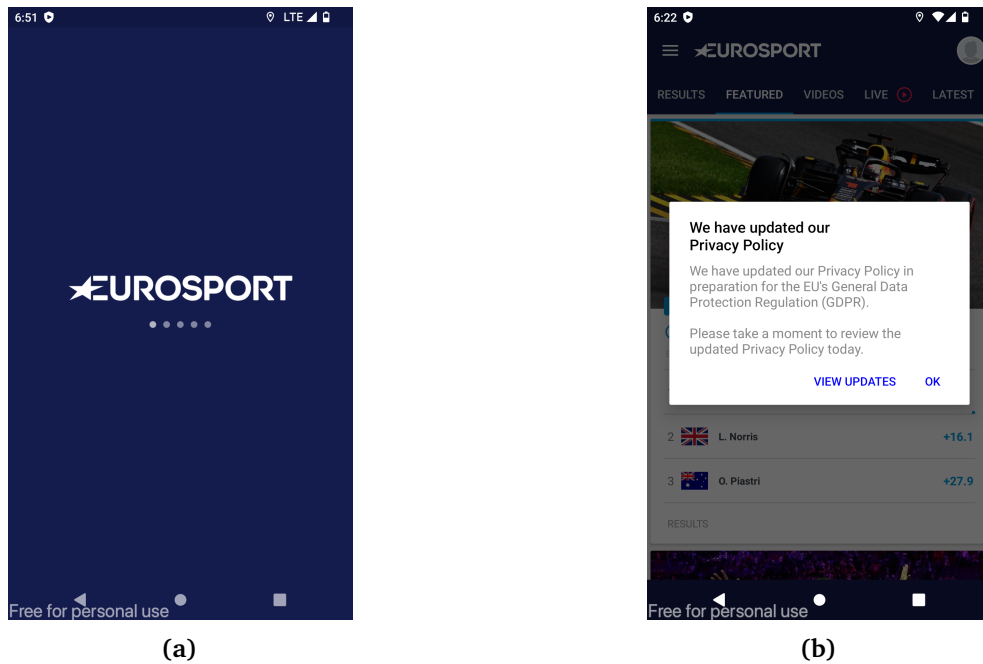


Figure 4.2: Starting the Eurosport app

Code listing 4.2 presents one of the GET requests sent to `liftoff.io`, referred to by `adx.g.doubleclick.net`. The packet contained parameters including advertisement ID, channel ID, auction ID, and origin.

```
GET https://impression-asia.liftoff.io/doubleclick/beam?ad_group_id=170916&
channel_id=16&creative_id=134457&auction_id=13bbb16cb8e9afb57d6709fb6da4be7a&
loid=DjAKivGwY1J2e4KGtzo0&origin=haggler-doubleclick15099 HTTP/1.1
Host: impression-asia.liftoff.io
Connection: keep-alive
Accept: image/avif,image/webp,image/apng,image/svg+xml,image/*,*/*;q=0.8
X-Requested-With: com.eurosport
Sec-Fetch-Site: cross-site
Sec-Fetch-Mode: no-cors
Sec-Fetch-Dest: image
Referer: https://adx.g.doubleclick.net/
Accept-Encoding: identity
Accept-Language: no-NO,no;q=0.9,nb-NO;q=0.8,nb;q=0.7,en-US;q=0.6,en;q=0.5
{
  ad_group_id: 170916
  channel_id: 16
  creative_id: 134457
  auction_id: 13bbb16cb8e9afb57d6709fb6da4be7a
  loid: DjAKivGwY1J2e4KGtzo0
  origin: haggler-doubleclick15099
}
```

Code listing 4.2: Traffic observed from Eurosport to `liftoff.io`

Code listing 4.3 presents one of the packets sent to `careers.bupa.com.au` which contained information about the user's location.

```

POST https://careers.bupa.com.au/search-jobs/SetSearchRequestGeoLocation?lat=null&
lon=null&IsUsingGeolocation=true&hasHtml5GeoError=true&geoType=ip2ifnohtml5
HTTP/1.1
Host: careers.bupa.com.au
Connection: keep-alive
Content-Length: 0
Accept: */*
X-Requested-With: XMLHttpRequest
Origin: https://careers.bupa.com.au
Sec-Fetch-Site: same-origin
Sec-Fetch-Mode: cors
Sec-Fetch-Dest: empty
Referer: https://careers.bupa.com.au/heartofhealthcare?s_cid
=[]:[30416047]:[9037693]:[374701674]:[197191631]&dclid=CMDV48aAz4EDFX-
qZgIdfpIFiw
Accept-Encoding: identity
Accept-Language: no-NO,no;q=0.9,nb-NO;q=0.8,nb;q=0.7,en-US;q=0.6,en;q=0.5
Cookie: SearchVisitorId=e3aa3ba1-0f5d-d422-0d5a-deba66293a05; _ga=GA1
.3.671438780.1695962502; _gid=GA1.3.1738438802.1695962502; SearchSessionId={%22
SearchSessionId%22:%2252735084-ece3-15c7-2222-ab2ee989455d%22%2C%22
ImpressionParentId%22:%22%22%2C%22ViewParentId%22:%22%22%2C%22
GoogleSearchRequestId%22:%22%22%2C%22GoogleJobId%22:%22%22%2C%22Created
%22:%221695962502699%22}; _gat=1
{
  lat:          null
  lon:          null
  IsUsingGeolocation: true
  hasHtml5GeoError: true
  geoType:     ip2ifnohtml5
}

```

Code listing 4.3: Traffic observed from Eurosport to careers.bupa.com.au

Code listing 4.4 presents one of the POST requests sent to arkoselabs.com. The packet contained parameters including category and action, which potentially describe activity on the user device.

```

POST https://client-api.arkoselabs.com/fc/a/ HTTP/2.0
content-length: 196
accept: */*
cache-control: no-cache
x-newrelic-timestamp: 169596100825289
x-requested-with: XMLHttpRequest
x-requested-id: {"ct":"n/gWa3GRe6RxCtCcs8gAEQ==","iv":"5
    d36dee101ea8ec0908bafb9afd2b6df","s":"bfd61486ac7f1a3c"}
content-type: application/x-www-form-urlencoded; charset=UTF-8
origin: https://client-api.arkoselabs.com
sec-fetch-site: same-origin
sec-fetch-mode: cors
sec-fetch-dest: empty
referer: https://client-api.arkoselabs.com/fc/gc/?token=2551789444646e186
    .3866181403&r=ap-southeast-2&meta=3&metabgclr=%23ffffff&metaiconclr=%23757575&
    guitextcolor=%23000000&pk=FE296399-FDEA-2EA2-8CD5-50F6E3157ECA&at=40&rid=68&ag
    =101&cdn_url=https%3A%2F%2Fclient-api.arkoselabs.com%2Fcdn%2Ffc&lurl=https%3A%2
    F%2Faudio-ap-southeast-2.arkoselabs.com&surl=https%3A%2F%2Fclient-api.
    arkoselabs.com&smurl=https%3A%2F%2Fclient-api.arkoselabs.com%2Fcdn%2Ffc%2
    Fassets%2Fstyle-manager
accept-encoding: identity
accept-language: no-NO;q=0.9,nb-NO;q=0.8,nb;q=0.7,en-US;q=0.6,en;q=0.5
{
  sid:                ap-southeast-2
  game_token:         405651652ce914724.2893340903
  session_token:      2551789444646e186.3866181403
  game_type:          3
  render_type:        canvas
  category:           begin app
  action:             user clicked verify
  analytics_tier:     40
}

```

Code listing 4.4: Traffic observed from Eurosport to arkoselabs.com

Strava was the most challenging app to intercept. Initially, no traffic was readable, and the app did not work as intended. This was due to the developer's implementation of certificate pinning. Removing this feature from the APK before installing the app was necessary to intercept the traffic. There are several ways to do this. In this experiment, a tool called APK-MITM was used. This is demonstrated in Figure 4.3. After the removal of the certificate pinning, the app functioned as intended. However, the base map was missing, as shown in Figure 4.4. Research showed that the complication was caused by Google Maps, which requires an API key restricted to Strava's certificate. A new custom API key would solve this issue, but this was not pertinent for the experiment as traffic was intercepted successfully. However, this app only communicated with a fraction of Eurosport's number of third parties Eurosport communicated with.

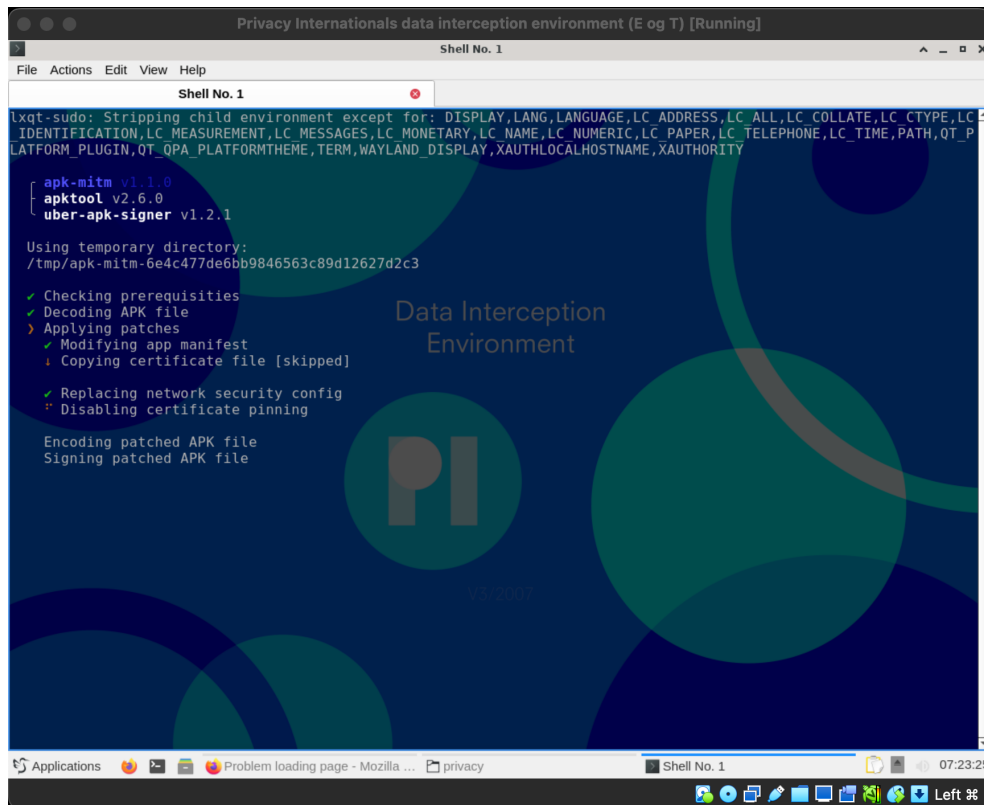
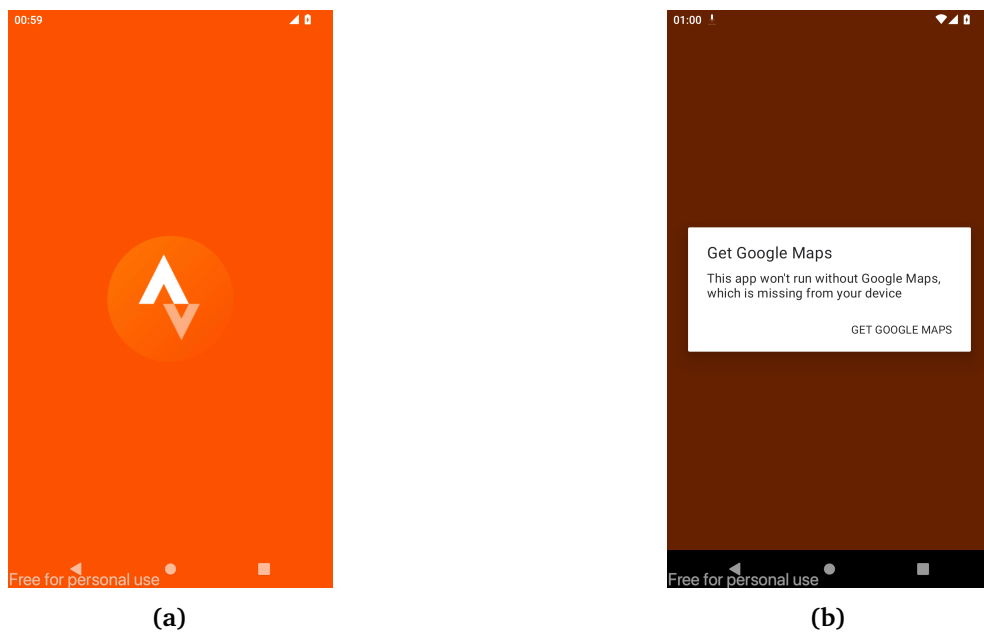


Figure 4.3: Certificate un-pinning of Strava APK using APK-MITM



(a)

(b)

Figure 4.4: Starting the Strava app

Code listing 4.5 presents one of the packets sent from Strava to `branch.io`. The packet contained device information, language, country, advertisement ID, and network information.

```
POST https://api.branch.io/v1/open HTTP/1.1
Content-Type: application/json
Accept: application/json
Host: api.branch.io
Connection: Keep-Alive
Accept-Encoding: identity
Content-Length: 942
{
  "app_version": "154.9",
  "branch_key": "key_live_lmpPsfj2DP8CfII4rmzfiemerte7sgwm",
  "brand": "Genymobile",
  "cd": {
    "mv": "-1",
    "pn": "com.strava"
  },
  "country": "NO",
  "debug": false,
  "device_fingerprint_id": "1234391181132195067",
  "environment": "FULL_APP",
  "external_intent_uri": "strava://dashboard",
  "facebook_app_link_checked": false,
  "first_install_time": 1695969018572,
  "google_advertising_id": "9057c569-6524-48a8-b9de-ba540e0716e2",
  "hardware_id": "33839551d0119e58",
  "identity_id": "1236190999731158629",
  "instrumentation": {
    "v1/open-qwt": "0"
  },
  "is_hardware_id_real": true,
  "is_referrable": 1,
  "language": "nb",
  "lat_val": 0,
  "latest_install_time": 1695969018572,
  "latest_update_time": 1695969018572,
  "local_ip": "100.64.32.53",
  "metadata": {},
  "model": "Pixel 2",
  "os": "Android",
  "os_version": 31,
  "previous_update_time": 1695969018572,
  "retryNumber": 0,
  "screen_dpi": 420,
  "screen_height": 1794,
  "screen_width": 1080,
  "sdk": "android2.19.3",
  "ui_mode": "UI_MODE_TYPE_NORMAL",
  "update": 1,
  "wifi": true
}
```

Code listing 4.5: Traffic observed from Strava to `branch.io`

Code listing 4.6 presents one of the packets sent from Strava to `appsflyer.com`. The packet contained several objects of information about the user and the device.

```

POST https://events.appsflyer.com/api/v4/androidevent?buildnumber=4.8.11&app_id=com
.askfm HTTP/1.1
Content-Length: 1342
Content-Type: application/json
Host: events.appsflyer.com
Connection: Keep-Alive
Accept-Encoding: identity
{
  "advertiserId": "9057c569-6524-48a8-b9de-ba540e0716e2",
  "advertiserIdEnabled": "true",
  "af_events_api": "1",
  "af_preinstalled": "false",
  "af_timestamp": "1695971137434",
  "af_v": "40b022c25443fdaefe74beed671d0ed72ae2a624",
  "af_v2": "9fd88cf95dcb527d6b1d7c62fd9c581498304552",
  "app_version_code": "1253",
  "app_version_name": "4.25",
  "appsflyerKey": "BJET4dcg6k72DBvdE5XYLP",
  "brand": "Google",
  "carrier": "Android",
  "cksm_v1": "549c1793ce1443b99f8fab7399616fa180",
  "counter": "27",
  "country": "",
  "date1": "2023-09-24_123138+0000",
  "date2": "2023-09-24_123138+0000",
  "device": "vbox86p",
  "deviceData": {
    "arch": "",
    "build_display_id": "vbox86p-userdebug 12 SQ1D.220205.004 119 test-keys",
    "cpu_abi": "x86_64",
    "cpu_abi2": "",
    "dim": {"d_dpi": "420","size": "2","x_px": "1080","xdp": "420.0","y_px":
      "1794","ydp": "420.0"}
  },
  "deviceType": "userdebug",
  "eventName": "App Opened",
  "eventValue": "{}",
  "firstLaunchDate": "2023-09-24_123159+0000",
  "iaecounter": "16",
  "installDate": "2023-09-24_123138+0000",
  "isFirstCall": "true",
  "isGaidWithGps": "true",
  "lang": "norsk nynorsk",
  "lang_code": "nn",
  "model": "Pixel 2",
  "network": "MOBILE",
  "operator": "T-Mobile",
  "platformextension": "android_native",
  "prev_event": "{\"prev_event_timestamp\":\"1695960650676\",\"prev_event_value
    \":\"{}\", \"prev_event_name\": \"App Opened\"}",
  "product": "vbox86p",
  "registeredUninstall": false,
  "sdk": "31",
  "uid": "1695558719403-1431824616624252343"
}

```

Code listing 4.6: Traffic observed from Strava to appsflyer.com

Code listing 4.7 presents one of the packets sent from Strava to iterable.com.

The packet contained device information, email address and user ID.

```
POST https://api.iterable.com/api/users/registerDeviceToken... HTTP/2.0
time-offset-seconds: 7200
ot-tracer-traceid: d1f6740b854ffe59
ot-tracer-spanid: 79d91644933bd170
ot-tracer-sampled: true
content-type: application/json; charset=UTF-8
content-length: 300
accept-encoding: identity
{
  "device": {
    "applicationName": "STRAVA_ANDROID",
    "platform": "GCM",
    "token": "f7kLDfpzeDc:APA91bHmcoU5rCfZpiYtmbED5W-Z-3260
      TcvoKuiLDWJBZ8TWRpF2XptwCxS4oibnk08KEuuzlpSz7_MZTwK74F-uBMpsm0Y-
      MQQEVyfkqQRh0DqDs-6cmCwhI4Gq1JTdabDWzPL59dx"
  },
  "email": "axelnorland@gmail.com",
  "preferUserId": true,
  "userId": "125252019"
}
```

Code listing 4.7: Traffic observed from Strava to iterable.com

The collection on Tinder, as shown in Figure 4.5, required some parameters for a complete registration. As listed below, these could mostly be taken out of the table created in the methodology.

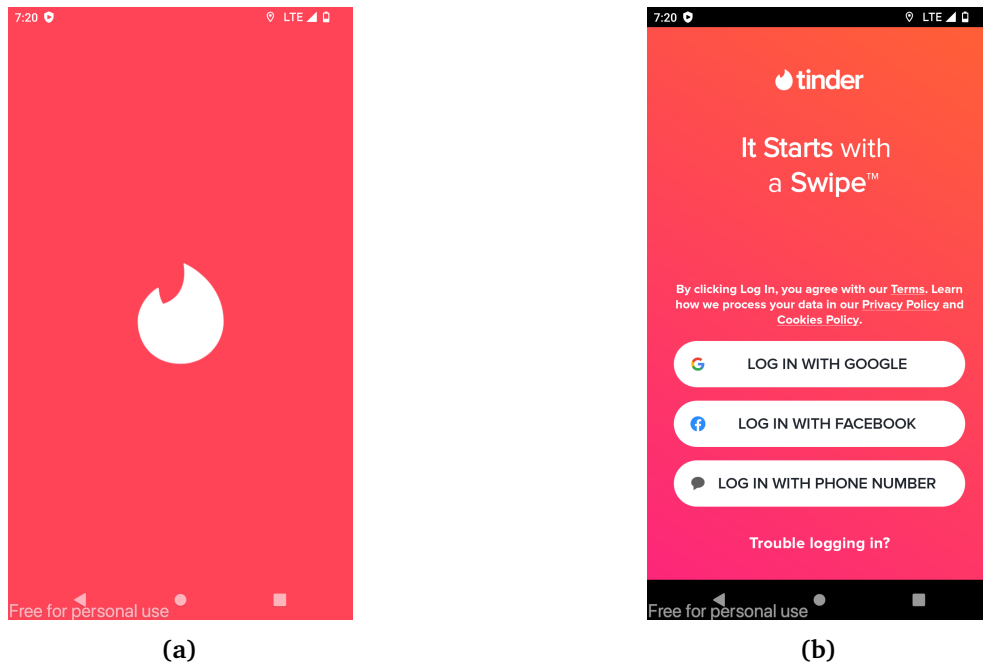


Figure 4.5: Starting the Tinder app

- Maximum distance 80km
- Age range 18-33
- Searching for women
- Casual relationship
- Elvebakken upper secondary school
- Gym, Rock, Active Lifestyle, MC Racing, Motorsport
- Added the two pictures shown in Figure C.1

Code listing 4.8 presents one of the packets sent from Tinder to bugsnag.com. The packet contained device-and user information.

```

POST https://sessions.bugsnag.com/ HTTP/1.1
Bugsnag-Integrity: sha1 74b593aa664846ea3932401e944a472429de24e8
Bugsnag-Payload-Version: 1.0
Bugsnag-API-Key: 745b354d173a082bd8bbf7a1501df2e6
Content-Type: application/json
Bugsnag-Sent-At: 2023-09-29T03:05:08.250Z
Content-Length: 762
Host: sessions.bugsnag.com
Connection: Keep-Alive
Accept-Encoding: identity
{
  "app": {
    "binaryArch": "x86_64",
    "buildUUID": "c752ea05-3599-4aeb-8fd8-47f1e4a09eaa",
    "id": "com.tinder",
    "releaseStage": "production",
    "type": "android",
    "version": "14.17.0",
    "versionCode": 14170092
  },
  "device": {
    "cpuAbi": ["x86_64", "x86"],
    "id": "2364fbd8-6994-4480-ad55-d3be360ec5b1",
    "jailbroken": true,
    "locale": "nb_NO",
    "manufacturer": "Genymobile",
    "model": "Pixel 2",
    "osName": "android",
    "osVersion": "12",
    "runtimeVersions": {
      "androidApiLevel": "31",
      "osBuild": "vbox86p-userdebug 12 SQ1D.220205.004 119 test-keys"
    },
    "totalMemory": 2069975040
  },
  "notifier": {
    "name": "Android Bugsnag Notifier",
    "url": "https://bugsnap.com",
    "version": "5.31.1"
  },
  "sessions": {
    "id": "3652c8e8-f8f7-496a-ba16-2500353615aa",
    "startedAt": "2023-09-29T03:05:08.234Z",
    "user": {
      "id": "2364fbd8-6994-4480-ad55-d3be360ec5b1"
    }
  }
}

```

Code listing 4.8: Traffic observed from Tinder to bugsnag.com

The data collection was successful with fully readable traffic packets, including sensitive details such as sexual preferences, pictures, location and more. Again, the information was primarily sent to internal domains. Similar to Strava, Tinder conveyed the same information to `branch.io`. In addition, there were several requests to AppsFlyer, which remained unreadable after decryption.

ASKfm initialised in Figure 4.6, proved in the static analysis to be one of the apps that share the most data with third parties. This was confirmed in the dynamic analysis, where the app communicated with several advertising and analytic companies.

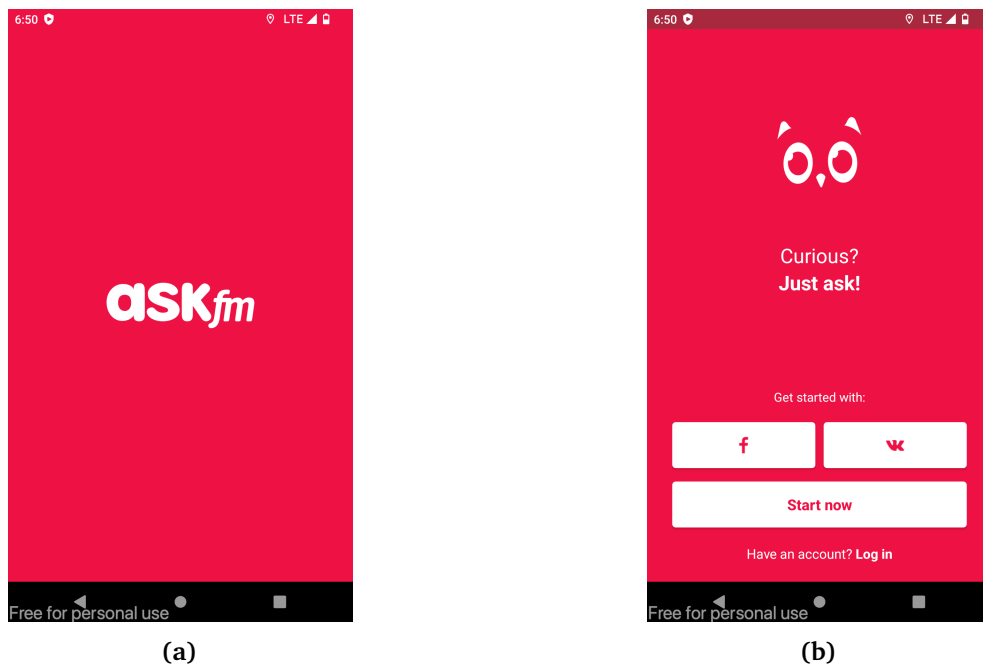


Figure 4.6: Starting the ASKfm app

Code listing 4.9 presents one of the packets sent from ASKfm to `pollfish.com`. The packet contained a parameter telling it not to encrypt the body, device information, whether the user is roaming, gender, year of birth, and a timestamp.

```

POST https://wss.pollfish.com/v2/device/register HTTP/1.1
Connection: Close
Content-Type: application/x-www-form-urlencoded
Host: wss.pollfish.com
Accept-Encoding: identity
Content-Length: 1555
{
  dontencrypt: true
  "device_descr": "Pixel 2 (vbox86p)",
  "provider": "Android",
  "provider_mcc": "310",
  "provider_mnc": "260",
  "nfc_enabled": "false",
  "position": "1",
  "usr_agent": "\Dalvik\\2.1.0 (Linux; U; Android 12; Pixel 2 Build\\SQ1D
    .220205.004)\",
  "custom_init": "true",
  "nfc_exists": "false",
  "app_id": "com.askfm",
  "board": "\unknown\"",
  "brand": "\Google\"",
  "target": "26",
  "os": "0",
  "os_ver": "31",
  "scr_h": "1794",
  "scr_w": "1080",
  "manufacturer": "Genymobile",
  "app_version": "4.251253",
  "con_type": "WIFI",
  "video": "true",
  "locale": "nn",
  "scr_size": "4.985714285714286",
  "is_roaming": "false",
  "accessibility_enabled": "false",
  "developer_enabled": "true",
  "install_non_market_apps": "true",
  "hardware_accelerated": "false",
  "api_key": "6eef8804-b822-4ec4-b5d6-63168f1be98d",
  "device_id": "9057c569-6524-48a8-b9de-ba540e0716e2",
  "opt_out": "false",
  "survey_format": "0",
  "version": "18",
  "debug": "false",
  "google_play": "true",
  "gender": "2",
  "year_of_birth": "2001",
  "timestamp": "1695959215126",
  "encryption": "A422267BB67AD2F6920EFA64A1573EF820871F40D9B32B7C170FF63BE0332DC9"
}

```

Code listing 4.9: Traffic observed from ASKfm to pollfish.com

Code listing 4.10 shows one of the transmissions between ASKfm and MoPub. This packet shared information on the user, including host device, consent and 'Do Not Track' status, age, gender, and advertising ID.

```

POST https://ads.mopub.com/m/ad HTTP/1.1
accept-language: nb-no
Content-Type: application/json; charset=UTF-8
Host: ads.mopub.com
Connection: Keep-Alive
Accept-Encoding: identity
Content-Length: 430
{
  "android_perms_ext_storage": "0",
  "av": "4.25",
  "bundle": "com.askfm",
  "cn": "Android",
  "ct": "2",
  "current_consent_status": "unknown",
  "dn": "Genymobile,Pixel 2,vbox86p",
  "dnt": "0",
  "force_gdpr_applies": "0",
  "h": "1920",
  "id": "197a68ccda884626aed4afd9f62573cd",
  "mcc": "310", "mnc": "260", "mr": "1",
  "nv": "5.3.0",
  "o": "p",
  "q": "m_age:22,m_gender:m",
  "sc": "2.625",
  "udid": "mopub:6f59c9cc-98b1-41bd-b9db-ca2bf76a4269",
  "v": "6", "vv": "3",
  "w": "1080",
  "z": "+0200"
}

```

Code listing 4.10: Traffic observed from ASKfm to mopub.com

Figure 4.7 illustrates the locations of the third parties communicated with. Not all third parties showed their location on the HTTP header; hence, this illustration is not exhaustive. Figure D.1 shows which apps communicated where and can be found in Appendix D.

In summary, the apps communicated with the number of third parties shown in Figure 4.8. Eurosport, which communicated with 46 third-party domains in only one hour, stands out. The other three were quite similar, although they had different advertising partners.

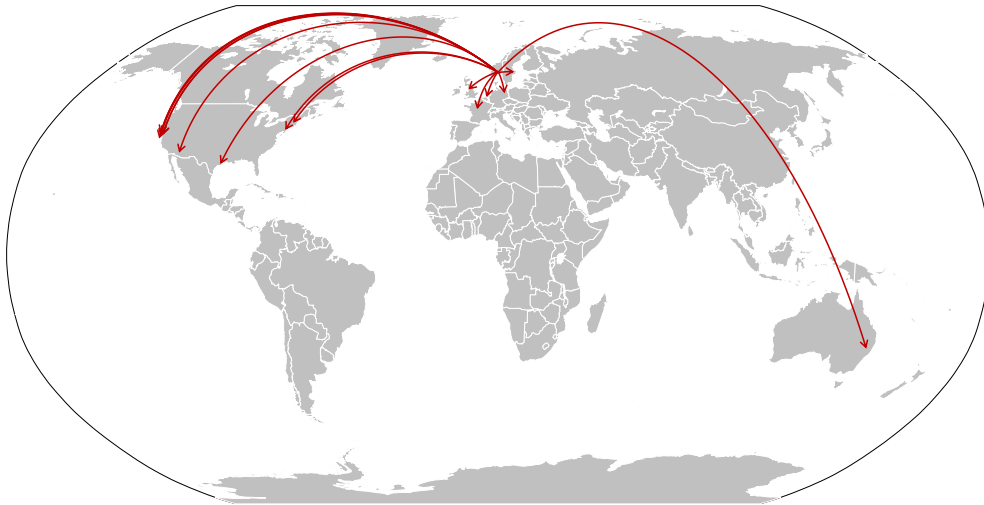


Figure 4.7: Overview of transmissions origin

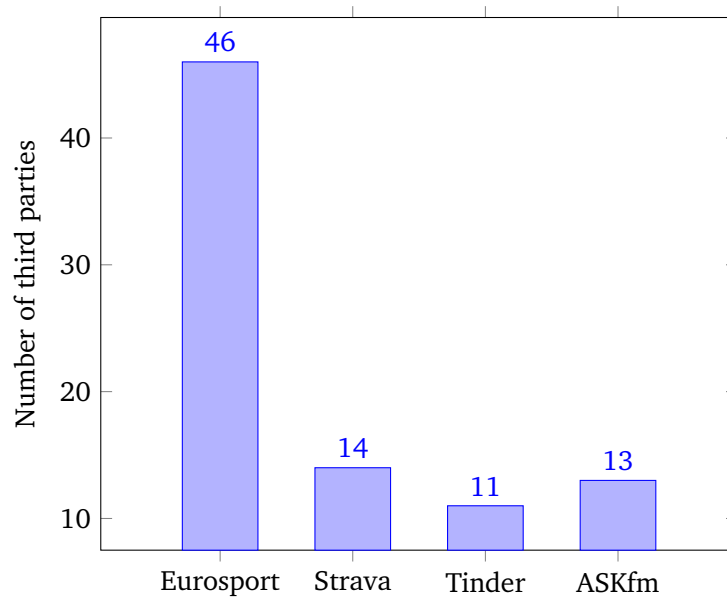


Figure 4.8: Third parties within the first hour

Figure 4.9 shows all the applications and their third parties combined in one Sankey diagram to illustrate which third parties are involved in the communication between the various apps. Facebook received data from all apps, and so did Google, when including Google Analytics.

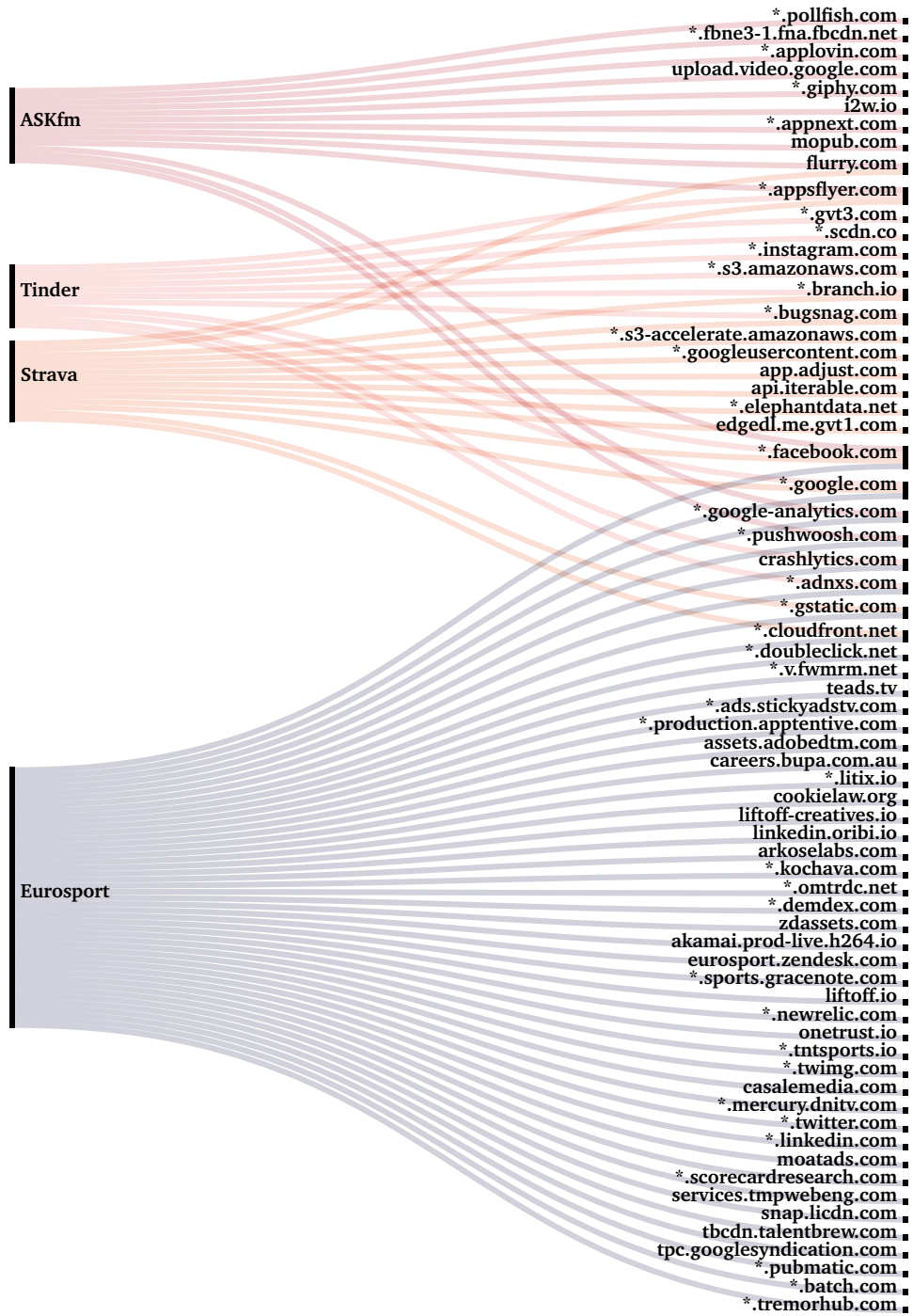


Figure 4.9: Apps third parties

Due to inconvenience, the JSON objects of the HTTP body are presented in Appendix D. Synthesising the data would corrupt the original message, and the original form is too lengthy to include.

Chapter 5

Analysis and Discussion

This chapter analyses and interprets the results presented in Chapter 4. Then, it evaluates the findings to answer the research questions. Lastly, it concludes by elaborating and evaluating the results' limitations, possible errors, and shortcomings, providing an understanding of the research outcomes.

5.1 Findings from the Content Analysis

The apps' privacy policies were studied to evaluate what the producers themselves disclose as being collected and shared. The results revealed a consistent trend among all apps. They were found to collect and share anonymised and aggregated information, including basic account information, technical information, and profile, activity and use information. The information collected is shared with the public, other app users, law enforcement, service providers, or partners. The findings of this thesis are consistent with the study conducted by Brandtzaeg *et al.* [36]. However, while none of the 21 apps examined in Brandtzaeg *et al.* disclosed their third parties, one of the apps examined in this thesis, ASKfm, did mention the third parties with whom they share data.

Importantly, all privacy policies said they do not sell personal data to third parties. However, most relinquished the responsibility for data collected by partners and how this data is used, stored, and shared. Tinder and ASKfm disclosed that they ignore 'Do Not Track' signals, a function integrated to reduce the collection of personal data. Unfortunately, there is no universal legislation making this illegal. Additionally, all apps reserve the right to change their privacy policy at any time.

5.2 Findings from the Static Analysis

The static analysis revealed the embedded third-party SDKs, permissions, and suspicious code in the apps. The apps' number of trackers ranged from 10 to 39, in ascending order of Strava (10), Tinder (13), ASKfm (25), and Eurosport (39). Eurosport required the least permissions from the device, only 8 of which

most accessed network configuration. Strava required 25 permissions to run, while Tinder and ASKfm required 31. Additionally, all apps had trackers from Facebook, Google, and advertising and analytics companies. Most permissions could be rationalised as providing necessary information to the app's services, such as `ACCESS_FINE_LOCATION` in Strava and `CAMERA` in Tinder.

Anti-VM code was found in Eurosport, Strava, and ASKfm. This code checks if the app runs inside a virtual machine and alters its behaviour accordingly. Eurosport and ASKfm also contained `Debug.isDebuggerConnected()`, which can be interpreted as anti-debug code. Additionally, Eurosport had obfuscation code. These findings are suspicious, as the developer wants to hinder the analysis of the applications. However, these observations do not prove that the apps collect or share more user data than initially assumed.

MobSF was unable to analyse Tinder's APK. Initially, the hypothesis was that the problem stemmed from conflicting naming conventions, as the software failed to recognise the file as an APK. However, after verifying its content and trying other versions of the app that are compatible with x64 architecture, it is reasonable to assume Tinder has implemented some functionality to avoid being disassembled and analysed.

While other studies, such as 'Out of Control', stated to use tools for static analysis, their results were not explicitly mentioned in the report [17]. It is, therefore, difficult to contextualise the findings of this thesis. Perhaps this thesis can serve as a reference for future research when studying application packages.

5.3 Findings from the Dynamic Analysis

There is an apparent coherence between the number of trackers in the APKs, and the observed number of third parties interacted with in the dynamic analysis. Thus, a more significant result could have been derived using a static analysis tool such as *Exodus* to find apps with the most trackers across all categories and user demographics. However, this thesis did not intend to find the app that shares the most sensitive data. Instead, the analysis seeks to examine apps used by the subset of the population and determine their impact on OPSEC. Not all parties with embedded SDKs received data. Either because the testing did not provoke the sharing, or some could be receiving data mediated via other integrated SDKs.

Looking exclusively at third-party communications, Eurosport sent data to 46 unique domains, Strava to 14, Tinder to 11, and ASKfm to 13. Most of the third parties were located in the US. However, five are in Europe, and one is in Australia. Well-known companies such as Facebook and Google are the most frequently observed. Lesser-known companies such as Adnxs, AppsFlyer, Arkose Labs, Branch, Bugsnag, Flurry, Iterable, Liftoff, MoPub, Pollfish, and Pushwoosh are also present. Interestingly, these lesser-known companies are all in the ad tech industry, except Arkose Labs, which specialises in cyber defence, and Bugsnag, which works in error monitoring and reporting.

Some of the transmissions to the domain `teads.tv` included an ‘Action’ object comprising parameters such as `adAvailable-success`, `adReached`, `adReceived`, `adSlotVisible`, `click`, `complete`, and `impression`. This data suggests that `teads.tv` is involved with delivering ads to Eurosport and delivering potential user interactions. One of the advertisements presented in the app came from the domain `careers.bupa.com.au`. Among the information shared to this domain was a packet containing information about the user’s location, as shown in Code listing 4.3. Code listing 4.2 shows one of the requests sent to Liffort from Eurosport. This packet contained `adx.g.doubleclick` as the referer in the HTTP header. The body included elements such as `ad_group_ID` and `auction_ID`, which supports the theory that this is a request for an advertisement from the app. Mnemonic also associated Liffort with participating in MoPub’s mediation technique, but since the app did not communicate with MoPub during the testing, this is ruled out. In this case, `doubleclick.net` is more likely to be the mediator, as this also had transmissions including `mediation_fill_status`. Lastly, requests sent to ArkoseLabs contained user activities such as `app status` and `user click` as shown in Code listing 4.4.

Strava shared information with several third parties, including AppsFlyer, Iterable, and Branch. AppsFlyer received information about the user and the device as shown in Code listing 4.6. Iterable email address and user ID. Branch, advertisement ID, country, language, and IP address. Strava had implemented certificate pinning and initially denied the connection to `mitmproxy` as it was expecting a different certificate from the host. However, this was successfully removed, and the result suggests that the certificate pinning did not affect the data collection. The developer most likely implemented this as a security measure to enhance the trust and integrity of the app and server communication.

Tinder shared most of the same information as Strava. In addition, it communicated in large volumes with AppsFlyer, which Mnemonic proved received date of birth, gender advertising ID, location, target gender, and age. These packets remained unreadable after decryption in this experiment but had the same characteristics as those captured by Mnemonic. Therefore, these packets could likely have contained the same information.

ASKfm communicated with numerous third parties and shared the same information with several. The companies receiving the most data were MoPub, Pollfish, and AppsFlyer. MoPub received device information, age, gender, and advertising ID. Pollfish received encryption, roaming, device ID, gender, birth year, and time. The request to AppsFlyer contained the same layout as the request sent by Strava.

Please note that all transmissions to internal domains were excluded from the analysis. So was the volume of transmissions to each third party. This is because most recurring packets were either re-transmissions, standby traffic, or a standard form transmitted back and forth. Neither of which provided any information relevant to this thesis.

In summary, the findings from the dynamic analysis suggest that little has changed since Privacy International’s study in 2018 [39]. However, this thesis establishes that this observation extends to apps typically used by young Norwegian

soldiers today. Thus, it bridges the gap between previous research and the demand for determining the severity of the data within the context of the Norwegian military.

Addressing this thesis's research questions involves examining the broader impact of the experiment's results. This includes assessing security implications for young soldiers and, by extension, the Norwegian military, leading to specific recommendations for building awareness and resistance.

5.4 RQ1: How Data Sharing Affects Security

How do the selected apps' data sharing with third parties affect the security of Norwegian soldiers aged 19-22?

This section explores the security implications for the subset arising from the selected apps' data sharing with third parties. This is done by a threat analysis, identifying relevant assets, their vulnerabilities and threats. Please note that the ethical and privacy implications of third-party tracking do not fall under the scope of this thesis. Neither is the apps' compliance with legal and regulatory frameworks.

Assets

The risk taxonomy provided by NATO StratCOM [5], presented on Page 10, identifies a broad range of assets. These assets encompass personnel, equipment, information, facilities, and activities, all of which will affect the security of Norwegian soldiers aged 19-22.

Vulnerabilities

Each asset has specific vulnerabilities, either inherent or due to its environment, which exposes them to exploitation. Personnel, including the subset directly, are vulnerable to the exposure of personal details, preferences or otherwise leverageable information. In addition, personnel are vulnerable to the exposure of location data, financial details, and proprietary information. Indirectly, the subset is vulnerable to equipment malfunction in either software or hardware and disclosure of sensitive details. Information affecting the security of soldiers aged 19-22 is vulnerable to loss of integrity and confidentiality. Facilities in which the subset works, lives, or visits are vulnerable to localisation and unauthorised access. Lastly, activities in which the group participates are vulnerable to being located and disrupted.

Threats Caused by the Selected Apps' Data Sharing

Twetman and Bergmanis-Korats' report and its corresponding risk taxonomy highlights numerous threat to military personnel. Referring to Table 2.1 and its delineation between white- and black markets for data, the vulnerability of sensitive

information reaching the public becomes evident, particularly as major information exchanges have experienced hacking incidents in recent years, leading to the leakage of sensitive user data. The nature and severity of the data's threats can be viewed from various perspectives. This thesis focuses on an operational military context.

Publicly available data can threaten military operational and organisational integrity. User data can, for example, be used together with location data to identify key personnel who can become victims of targeted attacks. A simple mailing list for military personnel can be acquired in the legitimate data market. In contrast, the black market typically offers data adapted for illegal purposes such as spear phishing attacks. Data aggregation can also reveal sensitive information about facilities such as NRK, revealing a man working at one of the Norwegian intelligence service's stations in Northern Norway [2]. And the Strava heat map back in early 2018, when military bases worldwide were mapped out using fitness tracking data [71].

The holistic profile of the fictitious soldier in Figure 5.1 demonstrates the footprint of an average Norwegian soldier in the age group 19-22 years from these four apps alone, excluding data sent to internal domains. Furthermore, this is solely based on the observed traffic from the first hour of the app running and does not include the data stated in the privacy policies nor the data hardcoded in the apps APKs.

The four apps were observed sharing name, email address, country, gender, age, network information, language, location data, time zone, user interactions, and logs. Arguably, more attributes may have been observed if the apps were tested over a greater period. This data is, according to Spiekermann and Christl, sufficient for re-identifying individuals [13].



Figure 5.1: Attributes observed during the experiment

Additionally, a realistic user has more than four apps installed on their phone. However, sensitive information shared by merely one app is sufficient to pose an equivalent threat, resulting in increased knowledge for the adversary. This underscores the risks associated with the overall tracking volume and the potentially significant impact of individual applications.

The results from the different analytical methods correlated well while also providing supplementary perspectives and depth. The methodological strategy effectively served the purpose of triangulation, with all three data sets closely aligning. Nevertheless, not all third parties identified in the static analysis were contacted during the dynamic analysis, and not all the information specified in the privacy policies was observed in the traffic analysis.

This thesis's results correlate with earlier research. Studies going 2-5 years back find a stronger data flow to third parties. This was anticipated as increased scrutiny and economic fines will likely affect companies' operations.

Some apps claim only to sell anonymised and aggregated information without the ability to identify the user. However, as the examples in Chapter 2 showed, the anonymisation promised by the apps is not always adequate, as some information may be deanonymized by correlating multiple data sources and applying sufficient efforts. The anonymised data still introduces some risk, and some of the information shared is not anonymised at all.

Therefore, privacy policies claiming anonymity are no guarantee, particularly in the face of significant advances in Artificial intelligence (AI), which excel correlating data sets. This threat increases with an individual's public exposure, rendering the next generation of higher-ranking military officers prone to extortion, intelligence gathering, and counterintelligence by more sophisticated adversaries.

The apps were observed sharing personal information, which often, through multi-source analysis, exposes personnel to threats such as extortion, manipulation, impersonation, and intelligence gathering. The age group studied in this thesis is the next generation of military leaders, who will serve as key personnel within their respective branches over time. Device ID and network information were also observed, exposing equipment to intelligence on capabilities and mapping communication patterns that can be used for sabotage.

The sensitive information shared from the apps is at risk of data theft and exposure, compromising the confidentiality and integrity of military data. Facilities could face threats related to the localisation of sensitive or secret sites, the identification of personnel working in specific facilities, and unauthorised access through impersonation. Lastly, activities could be compromised by tracking personnel movement, pinpointing operations, and the potential disruption of activities.

The data can also have operational and tactical implications. In the context of intelligence, this comprehensive profile of each soldier could be collected, analysed, and presented to inform decisions on the battlefield. Potentially locating enemy positions on the battlefield through geo-tracking or other digital markers enables precise surveillance from afar. Equally, coordinates collected through apps may disclose troop movements or targets for a weapon system. Therefore, the main concern is that lack of data control can become an accelerator [15].

The attempt by journalists to purchase personal data, as discussed in Chapter 2, frequently resulted in highly inaccurate information. Buyers face challenges in verifying the reliability and validity of such data, posing a risk of acquiring bad intelligence. This is worth noting, as it highlights the potential for adversaries

to access inaccurate data, even when obtained through a data broker. Therefore, information shared by the apps does not necessarily end up with the enemy.

The experiment demonstrated that the most sensitive data was shared with internal domains. However, this also holds its inherent risks. Its internal nature may rationalise more comprehensive collection and profiling. Secondly, in times of crisis, governments may instruct companies to share user information. This sets the stage for a new dimension of supply chain vulnerabilities, giving associations to the recent debates regarding government officials such as the Minister of Justice's use of TikTok. Data sharing is defined by a multifaceted threat understanding of where the destination and circumstances impact the potential risks and consequences.

In summary, data sharing from apps to third parties affects the security of Norwegian soldiers aged 19-22. The impact extends beyond direct threats, such as identification and extortion or manipulation, originating from sharing sensitive information. It also manifests indirectly through the threats this data poses to their equipment, information, facilities, and activities. This matters to the Norwegian Armed Forces and should be addressed to preserve OPSEC, and consequently, the security of soldiers aged 19-22. See Table 5.1 for a more extensive presentation of the identified threats.

5.5 RQ2: Potential Countermeasures

What countermeasures can the military implement to address potential threats arising from apps' data sharing with third parties?

To answer what countermeasures the military can implement, it is necessary to study the findings of RQ1. The study of the four apps revealed a noteworthy sharing of user data to third parties. The extent and sensitivity of the data raise profound concerns about its implications in a military context. It also proves the necessity of exploring countermeasures to safeguard the digital autonomy and, consequently, the soldiers of the Norwegian Armed Forces.

When discussing countermeasures, it is essential to have an understanding of risk. The identified threats are unlikely to be entirely eradicated. Instead, countermeasures aim to reduce the risk associated with an event to an acceptable level. This level could be defined as risk tolerance, which ultimately must be established by the organisation and its leaders. Especially in war, some significant risks may need to be embraced to attain desired outcomes. However, during peacetime, this tolerance should be lower. Nevertheless, one should always be aware of the risks and minimize them to the greatest extent possible.

The different threats can be sorted and arranged as shown in Table 5.1. Each event is sorted on its severity and probability, ranging from one to three, which multiplied produces a risk from one to nine. A countermeasure is then applied, leaving only a residual risk for each event. In presenting the risk matrix, it is essential to note that the assessments, including the numerical values, identified

threats, and corresponding mitigations, are subjective. Despite being grounded in available data and contextual considerations, the estimates are not definite.

Table 5.1: Risk matrix

Threat	S	P	R	Countermeasure	Residual Risk
Personnel					
Extortion	3	2	6	Routine for reporting and handling extortion attempts	4
Manipulation	3	1	3	Training and awareness briefings	2
Impersonation	1	3	3	Implement multi-factor authentication	1
Equipment					
Revealing capabilities	3	2	6	Restrict access to device info	3
Revealing communication patterns	2	3	6	Implement secure communication protocols	3
Information					
Data theft	3	2	6	Encrypt sensitive data, implement data access controls	3
Data leaks	2	3	6	'Need to know', data management	3
Facilities					
Localisation	2	3	6	Not allowed to bring devices to facilities	4
Identification of personnel	3	3	9	Data minimisation, use of VPN	4
Unauthorised access	3	3	9	Implement physical security measures, enhance access controls	4
Activities					
Tracking of personnel movement	2	3	6	Not carrying personal devices	5
Pinpointing operations	2	2	4	Implement geofencing or similar mitigation techniques	4
Disruption	3	2	6	Develop contingency plans	4

The countermeasures in Table 5.1 offer suggestions for mitigating the specific threats outlined in Section 5.4. However, some broader countermeasures must be implemented to address the vulnerabilities produced by the overall data sharing with third parties. One may categorise the myriad of effective countermeasures into organisational, user-centred, and technical measures. The implementation of a user policy can exemplify an organisational measure. A user-centred measure can be to limit the amount of sensitive data disclosed to the apps. A technical measure can be to enforce data minimisation, only allowing apps to access data necessary for its function.

In this case, data classification and governance are in place when discussing organisational measures. An elemental policy for the use of social media also exists, even though this is ten pages long. One must be able to assume that the

Armed Forces routines include proper vendor assessment, ensuring third-party vendors' data handling complies with internal policies and laws such as GDPR before entering agreements. Policy enforcement should also be in place, building consistent data and security policies across the organisation. The main concern is that most of these measures apply to work phones used in service. Reducing data shared from one device is insufficient when a personal device could share the same information without being encompassed by the organisation's policy.

Among the technical measures, ensuring adequate encryption, access controls, network segmentation, and data loss prevention should be expected in the security apparatus. However, implementation and maintenance can be complicated in large organisations with diverse populations collaborating across physical and technical architecture.

User-centred countermeasures are crucial. Often, countermeasures focus on the technology. However, the users are often the weakest link in the security chain. Tools such as user training, awareness, privacy settings, two-factor authentication, and reporting data leaks can be worthwhile to avoid costly data leaks.

Users are forced to play by the developer's rules, as all apps reserve the right to deny access to their service if users refrain from sharing information. Consequently, these common recommendations for protecting one's privacy should be considered. Firstly, avoid downloading unnecessary apps. Secondly, read the data security section of each app downloaded. Thirdly, avoid providing personal information. Lastly, ask services to delete user data when finished using it. For individuals with more flexibility, this approach is sensible. However, for employees with rigid job requirements, it can be more challenging.

Even though all the correct mechanisms are in place, individuals may still struggle to reduce the amount of data shared to a satisfactory level. For example, there are options to restrict data collection. From the users' perspective, this is a great and working feature, but as shown in the content analysis, both Tinder and ASKfm ignore and do not respond to this request. Stronger legislation is, therefore, necessary to protect user data. There is no point in having these features when companies lawfully ignore user prompts. From the examples of Meta in Norway, ChatGPT in Italy, and Google in Australia presented in Chapter 2, one can see that the legislation effectively helps regulate user data collection or at least hinders the companies from benefiting from the data collected.

Aligning these measures with the findings from the experiment shows that not all countermeasures would entirely mitigate the threat. Therefore, the most effective measure would be to refrain from carrying personal devices during service. However, this is a somewhat drastic and unrealistic solution for the broad strokes of the military. Even so, there will be a residual risk of collecting personal data outside of work hours. A compromise for protecting data pertinent to the military essentially comes down to routines established and enforced by the organisation that prioritises data protection.

As discussed, sensitive information shared by merely one app is sufficient to pose an equivalent threat, resulting in increased knowledge for the adversary. This

nuanced understanding of risk will help inform the development of resilience strategies. Countering broad risks requires comprehensive strategies such as policy implementation and awareness-building. While for specific app risks, imposing restrictions and targeted measures could be sufficient. The chosen approach depends on the perceived nature of the threat, influencing effective risk mitigation.

The implications of this thesis and the recommended course of action for the Norwegian Armed Forces can be grouped into organisational, user-centred, and technical measures. Firstly, to prevent data collection and sharing with third parties, the Armed Forces should continue their preventive measures of security clearances both for employees and partners. However, their social media policy must be revised, shortened, and clarified to ensure users' understanding and compliance.

Secondly, as a user-centred recommendation, the Norwegian Armed Forces should standardise their education program for young soldiers. With the findings of this thesis in mind, this training should contain what effective technical measures can be applied to the device and what measures are considered ineffective. The training should focus on data minimisation, only allowing apps access to what is necessary. Furthermore, a guide on how to remove their data from common databases and how to evaluate which apps are safe to use. Lastly, the training should allow soldiers to recognise malicious exploitation attempts, such as phishing attacks. The goal would be to empower soldiers with the skills necessary to make informed decisions about application usage for a sustainable effect.

Thirdly, the Armed Forces should maintain a technical infrastructure for sufficient security, using up-to-date encryption, secure communication protocols, and access control mechanisms. Furthermore, they should continue the implementation of multi-factor authentication and consider approaches such as geofencing around sensitive activities. This needs to be in place for there to be any point in instructing individuals on how to reduce the sharing of their data.

5.6 Limitations

This thesis examined four apps over one hour of use. The limited number of apps studied allows for significant variation in the quantity and sensitivity of data shared with third parties. Hence, these results cannot be generalised to all apps, a broader population than the one examined, or exclusively stated applicable to the subset of this thesis.

As discussed, the selection of apps ideally should have been based on objective criteria such as popularity. However, for the reasons presented, this was not feasible. The findings still indicate data collection and sharing within the target groups' popular apps.

The analysis of the traffic packets includes only the HTTP body. Some information could also be retrieved via the HTTP header, such as connection, user-agent, and requested language. Most traffic captured during the dynamic analysis went to internal domains and Application Programming Interfaces (APIs). There is a possibility that data could be shared via these domains, making the transmissions

seem legitimate from the perspective in this thesis but resulting in data being shared, nonetheless. These communications were not studied. Furthermore, the JSON objects containing user data were often called misleading names. This made it challenging to detect where sensitive information was shared.

Capturing data from location-dependent apps such as Strava and Tinder was more challenging due to the virtualised test environment. During the interception of Strava, some workouts were simulated using GPS spoofing. This seemingly worked fine for imitating running sessions but did not include pairing the app with accessories such as a heart rate monitor, which could have created more data to be shared. Moreover, a custom API key was not generated, and Google Maps did not present the base map correctly during testing, which unlikely affects the result. Tinder was only tested on one location, mentioned in Table 3.3.

This thesis also had some technical weaknesses. Firstly, the phone had a 32-bit processor, reducing the number of compatible apps. Secondly, during the dynamic testing, all apps were installed simultaneously and not uninstalled after testing. This led to some background-process interruptions as shown in Figure 5.2.

The screenshot displays a network traffic analysis tool interface. On the left, a table lists network requests with columns for Path, Method, Status, Size, and Time. The right pane shows the details of a selected response, including headers and a JSON body. The JSON body contains tracking-related information such as 'advertiser_id', 'application_tracking_enabled', and 'application_package_name'.

Path	Method	Status	Size	Time
https://c.strava.com/com.snowplowanalytics....	POST	200	2.0kb	1s
https://c.strava.com/com.snowplowanalytics....	POST	200	6.2kb	282ms
https://api.branch.io/v1/close	POST	200	603b	212ms
https://c.strava.com/com.snowplowanalytics....	POST	200	6.9kb	311ms
https://tracing-collector.strava.com/api/v2/rep...	POST	200	922b	296ms
https://graph.facebook.com/v3.1/129215213...	GET	200	638b	331ms
https://graph.facebook.com/v3.1/129215213...	POST	200	2.4kb	811ms
https://events.appsflyer.com/api/v4/androide...	POST	400	1.3kb	352ms
https://t.appsflyer.com/api/v4/androidevent?b...	POST	400	1.2kb	1s
https://data.flurry.com/aap.do	POST	200	2.0kb	336ms
https://api.ask.fm/config?rt=1&ts=1695971137	GET	200	39.5kb	666ms
https://data.flurry.com/aap.do	POST	200	409b	274ms
https://data.flurry.com/aap.do	POST	200	1.0kb	311ms
https://blackhole.pushwoosh.com/67735/regi...	POST	200	593b	445ms
https://blackhole.pushwoosh.com/67735/getl...	POST	200	209b	427ms
https://graph.facebook.com/v3.1/129215213...	POST	200	1.6kb	412ms

```

Request  Response  Details
Content-Type: content/unknown
[{"_eventName": "fb_sdk_initialize", "_eventName_md5": "d470d22f23
--312ndDfv2rTH1S1sAbouNdArYfORhtTPEefj3q2f
Content-Disposition: form-data; name="event"
CUSTOM_APP_EVENTS
--312ndDfv2rTH1S1sAbouNdArYfORhtTPEefj3q2f
Content-Disposition: form-data; name="advertiser_id"
9057c569-6524-48a8-b9de-ba540e0716e2
--312ndDfv2rTH1S1sAbouNdArYfORhtTPEefj3q2f
Content-Disposition: form-data; name="advertiser_tracking_enabled"
true
--312ndDfv2rTH1S1sAbouNdArYfORhtTPEefj3q2f
Content-Disposition: form-data; name="anon_id"
XZ819bd975-bbc4-479f-af91-e1bb6650cc29
--312ndDfv2rTH1S1sAbouNdArYfORhtTPEefj3q2f
Content-Disposition: form-data; name="application_tracking_enabled"
true
--312ndDfv2rTH1S1sAbouNdArYfORhtTPEefj3q2f
Content-Disposition: form-data; name="extinfo"
["a2", "com.askfm", 1253, "4.25", "12", "Pixel 2", "nn", "CEST", "Android"
--312ndDfv2rTH1S1sAbouNdArYfORhtTPEefj3q2f
Content-Disposition: form-data; name="application_package_name"
com.askfm
--312ndDfv2rTH1S1sAbouNdArYfORhtTPEefj3q2f
View: auto  Couldn't parse: falling back to Raw
transparent mode  showhost  raw-top  anticache  anticomp  0.0.0.0:8080  v7.4
Applications  mitmproxy - Mozilla Firefox  Progress  08:0

```

Figure 5.2: ASKfm interrupting the collection from Strava

Chapter 6

Conclusions and Future Work

Android apps' data collection and sharing with third parties concerns the Norwegian Armed Forces. A handful of apps have been examined to decide which information is collected and to whom it is shared. Both the quantity and sensitivity of data formed a threat analysis, expanding on the potential of the data in a military context. Furthermore, some countermeasures have been discussed to increase resilience and minimise the threat.

Throughout the experiment, user data was observed being transmitted to numerous third parties. Some represent well-known advertising companies. Others are giant social media platforms or analytic companies. The uncertainty regarding which third parties might share sensitive information with adversaries amid a conflict poses a significant risk. The user data included sensitive details about the users' attributes such as age and location, actions such as clicks and page views, and preferences in language and ads. Consequently, this data poses potential threats to Norwegian Armed Forces soldiers, including risks of extortion, localisation, and disruption.

In conclusion, the most efficient measure is to avoid carrying personal devices during service due to limited control over information collection by specific apps. However, considering practical constraints such as budget and organisation, this may not be realistic. A compromise may, therefore, be to establish solid routines, guide and warn users in managing their data, and encourage stronger legislation. Any first step must include a realistic understanding of organisational and individual assets and vulnerabilities. Given the complex and evolving threat landscape shaped by surveillance capitalism, today's vulnerabilities may differ from tomorrow's.

6.1 Suggestion for Further Research

Hopefully, this thesis has helped inform a more comprehensive understanding of assets and vulnerabilities. It has identified data sharing and why this is a military risk. However, to build sound countermeasures, a fuller understanding of what our

adversaries seek to accomplish and how that can go about it is necessary. Therefore, work must be done regarding threat-understanding.

This study focuses on a specific subset, namely soldiers aged 19-22. Future work should consider a broader scope by encompassing all personnel within the Norwegian Armed Forces to validate the findings across the entire population.

Furthermore, it would be interesting to see whether the challenges identified in this study extend to military personnel of other NATO member nations. Or if unique factors contribute to making other nation's military less prone to information disclosure.

Lastly, it would also be interesting to see the amount of data shared by a realistic number of apps on a device. This would, combined with testing the apps over a more extended period, offer a valuable dimension to the research.

Bibliography

- [1] 'Information security, cybersecurity and privacy protection,' International Organization for Standardization, Geneva, CH, Standard, Oct. 2022, ISO/IEC 27001:2022.
- [2] M. Gundersen, Ø. B. Skille, H. Lied, M. Grafsrønningen, and H. K. Jansson, 'Norske offiserer og soldater avslørt av mobilen,' 2020, Accessed 19. April 2023. [Online]. Available: <https://www.nrk.no/norge/xl/norske-offiserer-og-soldater-avslort-av-mobilen-1.14890424>.
- [3] Y. Walther-Zhang, 12344: *Bruk av smarttelefon og sikkerhet knyttet til denne (prosent), etter alder, statistikkvariabel, år og kjønn*, 2020. [Online]. Available: <https://www.ssb.no/statbank/table/12344/tableViewLayout1/>.
- [4] H. Ersdal and S. S. Skjærstad, 'Privacy and social media: Do users really care?' Department of Telematics, NTNU – Norwegian University of Science and Technology, master's thesis, May 2016.
- [5] H. Twetman and G. Bergmanis-Korats, 'Data brokers and security,' NATO Strategic Communications Centre of Excellence, 2021, ISBN: 978-9934-564-31-4. [Online]. Available: https://stratcomcoe.org/pdfjs/?file=/publications/download/data_brokers_and_security_20-01-2020.pdf?zoom=page-fit.
- [6] J. Dawson, 'Microtargeting as information warfare,' *The Cyber Defense Review*, vol. 6, no. 1, pp. 63–80, 2021.
- [7] J. Scott, 'Metadata: The most potent weapon in this cyberwar the new cyber-kinetic-meta war support icit,' Jul. 2017. DOI: 10.13140/RG.2.2.26334.95040.
- [8] S. Bay and N. Biteniece, 'The current digital arena and its risks to serving military personnel,' NATO Strategic Communications Centre of Excellence, 2019, ISBN: 978-9934-564-39-0. [Online]. Available: <https://stratcomcoe.org/publications/the-current-digital-arena-and-its-risks-to-serving-military-personnel/102>.
- [9] S. Kampesæter, *Rekrutter vil ha tydeligere føringer for sosiale medier*, Accessed 10. May 2023, 2020. [Online]. Available: <https://forsvaretsforum.no/cyber-sosiale-medier-teknologi/rekrutter-vil-ha-tydeligere-foringer-for-sosiale-medier/158307>.

- [10] B. Skredderhaug, 'Deling av brukerdata fra android-apper til tredjeparter,' Department of Information Security, Communication Technology, NTNU – Norwegian University of Science, and Technology, Project report in IMT4205, May 2023.
- [11] S. Zuboff, 'Surveillance capitalism and the challenge of collective action,' *New Labor Forum*, vol. 28, no. 1, pp. 10–29, 2019. DOI: 10.1177/1095796018819461. eprint: <https://doi.org/10.1177/1095796018819461>. [Online]. Available: <https://doi.org/10.1177/1095796018819461>.
- [12] W. Christl, K. Kopp, and P. U. Riechert, 'Corporate surveillance in everyday life,' *Cracked Labs*, vol. 6, 2017.
- [13] S. Spiekermann and W. Christl, *Networks of Control – A Report on Corporate Surveillance, Digital Tracking*. Jan. 2016, ISBN: 978-3-7089-1473-2.
- [14] A. Lewis, *User-driven discontent*, Accessed 10. May 2023, 2010. [Online]. Available: <http://www.metafilter.com/95152/Userdriven-discontent#3256046>.
- [15] V. R. Wilhelmsen, 'Kommersiell sporing – nasjonal risiko,' *Internasjonal Politikk*, vol. 80, no. 1, pp. 53–77, 2022. DOI: 10.23865/intpol.v80.3096. [Online]. Available: <https://tidsskriftet-ip.no/index.php/intpol/article/view/3096>.
- [16] P. Fleming, S. G. Edwards, A. P. Bayliss, and C. R. Seger, 'Tell me more, tell me more: repeated personal data requests increase disclosure,' *Journal of Cybersecurity*, vol. 9, no. 1, tyad005, Mar. 2023, ISSN: 2057-2085. DOI: 10.1093/cybsec/tyad005. eprint: <https://academic.oup.com/cybersecurity/article-pdf/9/1/tyad005/50476426/tyad005.pdf>. [Online]. Available: <https://doi.org/10.1093/cybsec/tyad005>.
- [17] Forbrukerrådet, 'Out of control how consumers are exploited by the online advertising industry,' 2020, Accessed 10. May 2023. [Online]. Available: <https://storage02.forbrukerradet.no/media/2020/01/2020-01-14-out-of-control-final-version.pdf>.
- [18] K. Cox, 'Secret service buys location data that would otherwise need a warrant,' 2020, Accessed 24. April 2023. [Online]. Available: <https://arstechnica.com/tech-policy/2020/08/secret-service-other-agencies-buy-access-to-mobile-phone-location-data/>.
- [19] H. Twetman and G. Bergmanis-Korats, 'Data brokers and security,' 2021, Table 2, Comparing the white and the black markets for data; p. 17. [Online]. Available: https://stratcomcoe.org/pdfjs/?file=/publications/download/data_brokers_and_security_20-01-2020.pdf?zoom=page-fit.
- [20] C. R. Miller, 'I bought a report on everything that's known about me online,' 2017, Accessed 25. April 2023. [Online]. Available: <https://www.theatlantic.com/technology/archive/2017/06/online-data-brokers/529281/>.

- [21] F. Myrstad, *The consumer council and friends read app terms for 32 hours*, Accessed 26. September 2023. [Online]. Available: <https://www.forbrukerradet.no/appfail-en>.
- [22] Personopplysningsloven, *Lov om behandling av personopplysninger*. Lovdata, 2018, LOV-2018-06-15-38. [Online]. Available: <https://lovdata.no/dokument/NL/lov/2018-06-15-38>.
- [23] G. Fouche, 'Norway fines facebook owner meta over privacy breaches,' 2023, Accessed 04. September 2023. [Online]. Available: <https://www.reuters.com/technology/norway-data-regulator-fine-meta-over-privacy-breaches-2023-08-07/>.
- [24] Meta, 'Facebook and instagram to offer subscription for no ads in europe,' 2023, Accessed 20. November 2023. [Online]. Available: <https://about.fb.com/news/2023/10/facebook-and-instagram-to-offer-subscription-for-no-ads-in-europe/>.
- [25] S. Mukherjee, E. Pollina, and R. More, 'Italy's chatgpt ban attracts eu privacy regulators,' 2023, Accessed 04. September 2023. [Online]. Available: <https://www.reuters.com/technology/germany-principle-could-block-chat-gpt-if-needed-data-protection-chief-2023-04-03/>.
- [26] N. Mittal, 'Australian court orders google to pay \$43 million for misleading users,' 2022, Accessed 04. September 2023. [Online]. Available: <https://www.reuters.com/technology/google-pay-427-mln-penalties-misleading-users-australian-watchdog-2022-08-12/>.
- [27] Etterretningstjenesteloven, *Lov om Etterretningstjenesten*. Lovdata, 2020, LOV-2020-06-19-77. [Online]. Available: <https://lovdata.no/dokument/LTI/lov/2020-06-19-77>.
- [28] A. Aas-Hansen, K. K. Devold, M. M. Kleppa, E. J. Husabø, C. B. Øvald, J. A. Ellingsen, O. Lysne, and H. Magnusson, 'Eos-utvalgets årsmelding for 2022,' 2023. [Online]. Available: <https://eos-utvalget.no/wp-content/uploads/2023/06/EOS-utvalgets-arsmelding-for-2022.pdf>.
- [29] Regjeringen, *Proposisjoner til Stortinget 80 L (2019–2020)*. 2020, LOV-2020-06-19-77. [Online]. Available: <https://www.regjeringen.no/no/dokumenter/prop.-80-l-20192020/id2698600/?ch=8>.
- [30] J. Caton, 'Dancing on the razor's edge: A foundational review of iot exploitation and defense through the lens of techint collection,' *International Journal of Intelligence and CounterIntelligence*, vol. 33, no. 3, pp. 540–555, 2020. DOI: 10.1080/08850607.2020.1720007. eprint: <https://doi.org/10.1080/08850607.2020.1720007>. [Online]. Available: <https://doi.org/10.1080/08850607.2020.1720007>.

- [31] '(u) report to the director of national intelligence,' *Office of the Director of National Intelligence Senior Advisory Group Panel on Commercially Available Information*, 2022, Accessed 18. September 2023. [Online]. Available: <https://www.dni.gov/files/ODNI/documents/assessments/ODNI-Declassified-Report-on-CAI-January2022.pdf>.
- [32] K. V. Rønn and S. O. Søre, 'Is social media intelligence private? privacy in public and the nature of social media intelligence,' *Intelligence and National Security*, vol. 34, no. 3, pp. 362–378, 2019. DOI: 10.1080/02684527.2019.1553701. eprint: <https://doi.org/10.1080/02684527.2019.1553701>. [Online]. Available: <https://doi.org/10.1080/02684527.2019.1553701>.
- [33] Ø. H. Kaldestad, 'Grindr dømt til å betale 65 millioner kroner etter klage fra forbrukerrådet,' 2021, Accessed 20. April 2023. [Online]. Available: <https://www.forbrukerradet.no/siste-nytt/grindr-domt-til-a-betale-65-millioner-kroner-etter-klage-fra-forbrukerradet/>.
- [34] W. J. Perry, 'Desert storm and deterrence,' *Foreign Aff.*, vol. 70, p. 66, 1990.
- [35] W. C. B. Michael N. Schmitt, 'Ukraine symposium – are civilians reporting with cell phones directly participating in hostilities?' *Articles of War*, Nov. 2022. [Online]. Available: <https://lieber.westpoint.edu/civilians-reporting-cell-phones-direct-participation-hostilities/>.
- [36] P. B. Brandtzaeg, A. Pultier, and G. M. Moen, 'Losing control to data-hungry apps: A mixed-methods approach to mobile app privacy,' *Social Science Computer Review*, vol. 37, no. 4, pp. 466–488, 2019. DOI: 10.1177/0894439318777706. eprint: <https://doi.org/10.1177/0894439318777706>. [Online]. Available: <https://doi.org/10.1177/0894439318777706>.
- [37] Q. Grundy, K. Chiu, F. Held, A. Continella, L. Bero, and R. Holz, 'Data sharing practices of medicines related apps and the mobile ecosystem: Traffic, content, and network analysis,' *bmj*, vol. 364, 2019. DOI: 10.1136/bmj.l920.
- [38] A. Khatoun and P. Corcoran, 'Privacy concerns on android devices,' in *2017 IEEE International Conference on Consumer Electronics (ICCE)*, 2017, pp. 149–152. DOI: 10.1109/ICCE.2017.7889265.
- [39] Privacy International, 'How apps on android share data with facebook,' 2018. [Online]. Available: <https://www.privacyinternational.org/sites/default/files/2018-12/How%20Apps%20on%20Android%20Share%20Data%20with%20Facebook%20-%20Privacy%20International%202018.pdf>.
- [40] 'Information technology, open systems interconnection,' International Organization for Standardization, Geneva, CH, Standard, Nov. 1994, ISO/IEC 7498-1:1994.
- [41] P. B. Brandtzæg, 'Datingapp deler sensitive helsedata,' Accessed 26. September 2023. [Online]. Available: <https://www.sintef.no/siste-nytt/2018/datingapp-deler-sensitive-helsedata/>.

- [42] A. Claesson and T. E. Bjørstad, 'Technical report "out of control" – a review of data sharing by popular mobile apps,' 2020, Accessed 10. May 2023. [Online]. Available: <https://storage02.forbrukerradet.no/media/2020/01/mnemonic-security-test-report-v1.0.pdf>.
- [43] Exodus, *The privacy audit platform for android applications*, 2023. [Online]. Available: <https://reports.exodus-privacy.eu.org/en/>.
- [44] M. D. i. Ajin Abraham India Magaofei china and V. N. france, *Mobile security framework (mobsf)*, 2022. [Online]. Available: <https://github.com/MobSF/Mobile-Security-Framework-MobSF>.
- [45] J. Shibchurn and X. B. Yan, 'Investigating effects of monetary reward on information disclosure by online social networks users,' in *2014 47th Hawaii International Conference on System Sciences*, 2014, pp. 1725–1734. DOI: 10.1109/HICSS.2014.220.
- [46] A. Xiong, T. Wang, N. Li, and S. Jha, 'Towards effective differential privacy communication for users' data sharing decision and comprehension,' in *2020 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2020, pp. 392–410. DOI: 10.1109/SP40000.2020.00088.
- [47] J. Karlsen, *Soldaters tinder-bruk på øvelse bekymrer*, Accessed 10. November 2023, 2019. [Online]. Available: <https://www.forsvaretsforum.no/soldaters-tinder-bruk-pa-ovelse-bekymrer/104826>.
- [48] M. Van Kleek, I. Liccardi, R. Binns, J. Zhao, D. J. Weitzner, and N. Shadbolt, 'Better the devil you know: Exposing the data sharing practices of smart-phone apps,' in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 5208–5220. DOI: 10.1145/3025453.3025556.
- [49] K. H. K. Herland, 'Policy for bruk av sosiale medier i forsvaret,' *Særskilt instruks for sjef Forsvarets sikkerhetsavdeling av 2020-03-01 pkt. 3 a.*, Ugradert.
- [50] S. B. Kaalaas, *Forsvaret forbyr tiktok på telefoner som brukes i tjeneste*, Accessed 10. May 2023, 2023. [Online]. Available: <https://forsvaretsforum.no/datasikkerhet-sosiale-medier/forsvaret-forbyr-tiktok-pa-telefoner-som-brukes-i-tjeneste/318780>.
- [51] P. Leedy and J. Ormrod, *Practical research: Planning and design, global edition*, 12th ed. London, England: Pearson Education, Apr. 2020.
- [52] Forsvaret, *Forsvarets årsrapport 2022*, Accessed 18. September 2023, 2022. [Online]. Available: [https://www.forsvaret.no/aktuelt-og-presse/publikasjoner/forsvarets-arsrapport/\(U\)_Forsvarets_Arsrapport_2022.pdf/_attachment/inline/4f35ffdf-160c-4b00-b9a7-bff5c04bee97:15a95807a777185c168808eb0190879f995702a9/\(U\)_Forsvarets_Arsrapport_2022.pdf](https://www.forsvaret.no/aktuelt-og-presse/publikasjoner/forsvarets-arsrapport/(U)_Forsvarets_Arsrapport_2022.pdf/_attachment/inline/4f35ffdf-160c-4b00-b9a7-bff5c04bee97:15a95807a777185c168808eb0190879f995702a9/(U)_Forsvarets_Arsrapport_2022.pdf).
- [53] M. Phythian, 'Intelligence and national security,' Online ISSN: 1743-9019. [Online]. Available: <https://www.tandfonline.com/action/journalInformation?journalCode=fint20>.

- [54] J. Goldman, 'International journal of intelligence and counterintelligence,' Online ISSN: 1521-0561. [Online]. Available: <https://www.tandfonline.com/action/journalInformation?journalCode=ujic20>.
- [55] J. G. Asaf Siniver, 'Journal of global security studies,' Online ISSN: 2057-3170. [Online]. Available: <https://academic.oup.com/jogss/pages/About>.
- [56] D. P. Tyler Moore, 'Journal of cybersecurity,' Online ISSN: 2057-2093. [Online]. Available: <https://academic.oup.com/cybersecurity/pages/About>.
- [57] OpenAI, *Chatgpt-3.5: Language models for text generation*, <https://chat.openai.com/>, 2023.
- [58] Privacy International, *Data interception environment*, 2021. [Online]. Available: <https://github.com/privacyint/appdata-environment-desktop/issues>.
- [59] O. A. V. Ravnås, *Dynamic instrumentation toolkit for developers, reverse-engineers, and security researchers*. 2023. [Online]. Available: <https://frida.re>.
- [60] R. Reith and N. Popal, *Smartphone market share*, Accessed 20. March 2023, 2023. [Online]. Available: <https://www.idc.com/promo/smartphone-market-share>.
- [61] C. Chin, *It's the perfect time to break up google's ad-tech monopoly*, Accessed 20. March 2023, 2023. [Online]. Available: <https://www.csis.org/analysis/its-perfect-time-break-googles-ad-tech-monopoly>.
- [62] Android, *Google pixel 2*, Accessed 10. September 2023. [Online]. Available: https://www.android.com/intl/en_au/phones/google-pixel-2/.
- [63] Eurosport, *Privacy policy*, Accessed 24. September 2023. [Online]. Available: https://www.eurosport.com/eurosport/legal/2020-2021/privacy-policy-eu-version-english-language_sto7660197/story.shtml.
- [64] Strava, *Privacy policy*, Accessed 28. September 2023. [Online]. Available: <https://www.strava.com/legal/privacy>.
- [65] Tinder, *Privacy policy*, Accessed 24. September 2023. [Online]. Available: <https://policies.tinder.com/privacy/intl/en/>.
- [66] ASKfm, *Privacy policy*, Accessed 28. September 2023. [Online]. Available: <https://about.ask.fm/legal/en/privacy.html>.
- [67] Exodus, *Eurosport*, 2023. [Online]. Available: <https://reports.exodus-privacy.eu.org/en/reports/101359/>.
- [68] Exodus, *Strava*, 2023. [Online]. Available: <https://reports.exodus-privacy.eu.org/en/reports/134749/>.
- [69] Exodus, *Tinder*, 2023. [Online]. Available: <https://reports.exodus-privacy.eu.org/en/reports/384714/>.

- [70] Exodus, *Askfm*, 2023. [Online]. Available: <https://reports.exodus-privacy.eu.org/en/reports/189998/>.
- [71] D. Ingram, 'Exclusive: Fitness app strava overhauls map that revealed military positions,' 2018, Accessed 10. October 2023. [Online]. Available: <https://www.reuters.com/article/us-strava-privacy-exclusive-idUSKCN1GP1WE>.

Appendix A

Permission from NATO StratCOM COE

Attached below is written permission from one of the authors of the article 'Data Brokers and Security' [5] to reuse Table 2.1 in this thesis.

Fra: Gundars Bergmanis-Korats Gundars.Bergmanis@stratcomcoe.org
Emne: Fw: Permission to reuse table from article
Dato: 3. oktober 2023 kl. 23:53
Til: bjartsk@stud.ntnu.no
Kopi: Info Info@stratcomcoe.org

Dear Bjarte Skredderhaug,

As an author, I hereby permit you to reuse the table. As you already indicated, please cite the source and authors of the research.

Kind regards,
Gundars

Gundars Bergmanis-Korats, Ph.D. | Principal Scientist
NATO Strategic Communications Centre of Excellence
11B Kalnciema Street, Riga, LV-1048, Latvia
Office: +371 6733-5498

From: Bjarte Skredderhaug <bjartsk@stud.ntnu.no>
Sent: Wednesday, September 27, 2023 5:02 AM
To: Info <Info@stratcomcoe.org>
Subject: Permission to reuse table from article

Permission to reuse a table published in the article «Data brokers and security» by NATO StratCom COE

My name is Bjarte Skredderhaug and I am currently writing my master thesis in Information Security on the topic of apps sharing user data to third parties, while I work in the Norwegian Armed Forces.

I found your article «Data brokers and security» relevant to my project and was wondering if I could include table 2: "NATO's comparison of white and the black markets for data» as a part of my chapter on related work?

The table will be made available for users to read and download for private use in accordance with copyright law.

I will refer correctly to you article and will be happy to insert any set phrase that you would require.

Thank you for your cooperation!

Bjarte Skredderhaug
+4798067751

Šis ir NATO StratCOM IC e-pasta sūtījums un paredzēts tajā norādītajam adresātam. Ja Jūs neesat šī sūtījuma adresāts vai persona, kas tiesīga šo sūtījumu saņemt, lūdzu, informējiet nosūtītāju un izdzēsiet šo e-pasta sūtījumu. Informējam, ka jebkāda šī sūtījuma satura izpaušana, kopēšana, izplatīšana vai darbība, pamatojoties uz tajā ietverto informāciju, ir aizliegta un var būt pretlikumīga.

This is the e-mail message of the NATO StratCOM COE and its contents are intended solely for the attention and use of the intended recipient. If you are not the intended recipient or the person eligible to receive this message, please contact the sender and delete this e-mail message from your system. You are hereby notified that any disclosure, copying, distribution or taking any action in reliance on the contents of this information is strictly prohibited and may be unlawful.

Appendix B

Master Agreement

The master agreement for this thesis is appended, providing the original definition of the thesis.

Master`s Agreement / Main Thesis Agreement

Faculty	Faculty of Information Technology and Electrical Engineering
Institute	Department of Information Security and Communication Technology
Programme Code	MISEB
Course Code	IMT4905

Personal Information	
Surname, First Name	Skredderhaug, Bjarte
Date of Birth	26.11.1998
Email	bjartsk@stud.ntnu.no

Supervision and Co-authors	
Supervisor	Maria Bartnes
Co-supervisors (if applicable)	, Vivi Ringnes Berrefjord
Co-authors (if applicable)	

The Master`s thesis	
Starting Date	14.08.2023
Submission Deadline	15.12.2023
Thesis Working Title	Deling av brukerdata fra Android-apper til tredjeparter
Problem Description	Cyberdomenet er i dag utsatt for et sammensatt og kompleks trusselbilde, hvor aktører fra enslige hackere til statssponsede grupper kan ha fiendtlige formål. Data om brukere selges på informasjonsbørser og informasjonen har aldri vært lettere tilgjengelig. I en militær kontekst kan deling av brukerdata utgjøre forskjellen mellom liv og død. Retningslinjene for mobilbruk i Forsvaret er i dag ikke tilfredsstillende (Kampeseter, 2020). Forskningsspørsmål: 1. Hvilke brukerdata deles med tredjeparter fra de utvalgte appene? 2. Hvilken trussel utgjør denne dataen for norske soldater, og hvilke grep kan Forsvaret ta?

Risk Assessment and Data Management	
Will you conduct a Risk Assessment?	Yes
If “Yes”, Is the Risk Assessment Conducted?	Yes
Will you Apply for Data Management? (REK*, NSD**)	No
Will You Write a Confidentiality Agreement?	No
If “Yes”, Is the Confidentiality Agreement Conducted?	No

* REK -- <https://rekportalen.no/>

** Norwegian Centre for Research Data (<https://nsd.no/nsd/english/index.html>)

Topics to be included in the Master`s Degree (if applicable)
IMT4110, IMT4130, IMT4116, IMT4203, IMT4114, IMT4113, IMT4205, ETTFO-STRA, IMT4905.

Guidelines – Rights and Obligations

Purpose

The Master's Agreement/ Main Thesis Agreement is an agreement between the student, supervisor, and department. The agreement regulates supervision conditions, scope, nature, and responsibilities concerning the thesis.

The study programme and the thesis are regulated by the Universities and University Colleges Act, NTNU's study regulations, and the current curriculum for the study programme.

Supervision

The student is responsible for

- Arranging the supervision within the framework provided by the agreement.
- Preparing a plan of progress in cooperation with the supervisor, including a supervision schedule.
- Keeping track of the counselling hours.
- Providing the supervisor with the necessary written material in a timely manner before the supervision.
- Keeping the institute and supervisor informed of any delays.
- Adding fellow student(s) to the agreement, if the thesis has more than one author.

The supervisor is responsible for

- Clarifying expectations and how the supervision should take place.
- Ensuring that any necessary approvals are acquired (REC, ethics, privacy).
- Advising on the demarcation of the topic and the thesis statement to ensure that the work is feasible within agreed upon time frame.
- Discussing and evaluating hypotheses and methods.
- Advising on literature, source material, data, documentation, and resource requirements.
- Discussing the layout of the thesis with the student (disposition, linguistic form, etcetera).
- Discussing the results and the interpretation of them.
- Staying informed about the work progress and assist the student if necessary.
- Together with the student, keeping track of supervision hours spent.

The institute is responsible for

- Ensuring that the agreement is entered into.
- Find and appoint supervisor(s).
- Enter into an agreement with another department / faculty / institution if there is an external co-supervisor.
- In cooperation with the supervisor, keep an overview of the student's progress, the number of supervision hours spent, and assist if the student is delayed by appointment.
- Appoint a new supervisor and arrange for a new agreement if:
 - The supervisor will be absent due to research term, illness, travel, etcetera.
 - The student or supervisor requests to terminate the agreement due to lack of adherence from either party.
 - Other circumstances where it is appropriate with a new supervisor.
- Notify the student when the agreement terminates.
- Inform supervisors about the responsibility for safeguarding ethical issues, privacy and guidance ethics
- Should the cooperation between student and supervisor become problematic, either party may apply to the department to be freed from the agreement. In such occurrence, the department must appoint a new supervisor

This Master`s agreement must be signed when the guidelines have been reviewed.

Signatures

Bjarte Skredderhaug
Student

14.08.2023
Digitally approved

Maria Bartnes
Supervisor

14.08.2023
Digitally approved

Hilde Bakke
Department

13.09.2023
Digitally approved

Appendix C

Test Setup

The apps were selected based on preliminary search results on embedded trackers and permissions using *Exodus* as shown Table C.1. Furthermore, the images used for the accounts are appended with screenshots during the test setup's configuration.

Table C.1: Apps tested for selection

Entertainment			Social networking		
App	Trackers	Permissions	App	Trackers	Permissions
Scrabble GO	36	17	Facebook	0	68
Wordle	15	9	Messenger	5	69
Candy Crush Saga	10	13	WhatsApp	1	74
Eurosport	39	8	Signal	0	73
Youtube	2	48	Tinder	13	31
NRK TV	6	16	Happn	7	23
Viaplay	6	13	Bumble	6	48
Netflix	2	20	Snapchat	2	64
Discovery+	8	15	Instagram	3	58
Amazon Prime	5	41	VSCO	9	22
TikTok	5	38	ASKfm	25	31
Spotify	6	43	BeReal	5	38
			Pinterest	5	28
			Reddit	5	47
			Discord	2	27

Lifestyle			News and information		
App	Trackers	Permissions	App	Trackers	Permissions
Sleep Cycle	3	6	VG	9	15
Headspace	3	18	DB	1	30
Strava	10	25	NRK	5	6
Finn	7	20	YR	5	14
Temu	5	14	VG Pent.no	4	5
Coop	5	8	Storm 24/7	0	4
Tise	8	21			

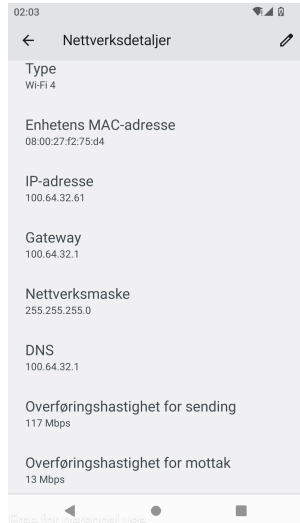
Productivity			Finance and utility		
App	Trackers	Permissions	App	Trackers	Permissions
Adobe Acrobat	8	25	DNB	2	14
Notion	2	28	Sparebank 1	3	5
Duolingo	10	35	Vipps	3	18
AdBlock Plus	0	4	Uber	4	34
Grammarly	3	13	Ruter	1	17
Teams	4	61	Flashlight	14	18
Trello	3	23	Calculator	1	5

The profile of the soldier is mentioned in the methodology chapter. Figure C.1 shows the images used for creating the necessary profiles. The profile images were created using AI and are not an actual individual; neither is it copyrighted.

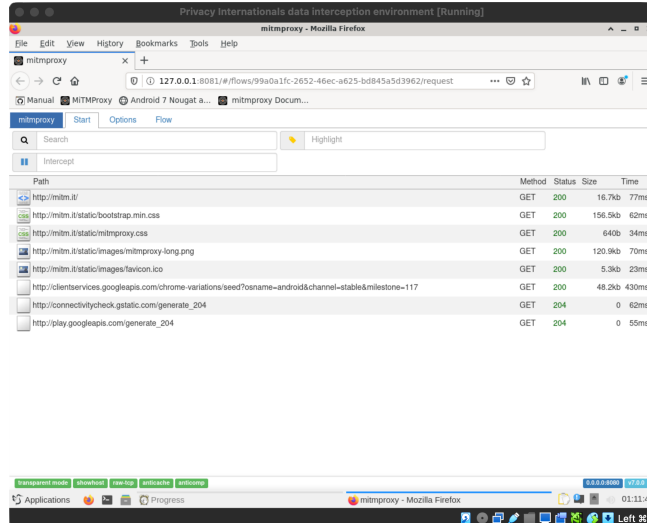


Figure C.1: Images uploaded to the Google account and Tinder

Figure C.2a presents the network settings active on the phone during the testing. Figure C.2b shows how mitmproxy, during setup, was able to intercept network traffic from the phone's web browser on mitm.it.



(a) Network settings on the Android device



(b) Configuring the test environment

Figure C.2: Screenshots

Appendix D

Analysis Results

The raw HTTP bodies were too lengthy to include in Chapter 4. Appended is, therefore, all JSON objects transmitted to all third-party domains during the experiment.

Figure D.1 illustrates together with Figure 4.7 the locations of the third parties communicated with and which apps communicated where. As mentioned, this is not an exhaustive representation as not all requests had a location in their HTTP header.

Table D.1: Objects of the traffic intercepted from Eurosport

***.doubleclick.net**
 _gid, _mv, _r, _u, _v, adurl, ai, aip, an, apm_app_id, apm_app_type, bisch, blev, blob, canm, cap, caps, carrier, cid, content_url, correlator, cust_params, description_url, dv, eid, env, fbs_aeid, fbs_aiid, format, gdfp_req, gjid, gl, gmp_app_id, gsb, guci, heap_free, heap_max, heap_total, hl, impl, is_lat, is_sidewinder, iu, jid, js, jsv, label, lite, lv, m_ast, mediation_fill_status, ms, msid, mv, n, net, ogsb, output, preqs, preqs_in_session, q, rdps, region, request_id, sai, seq_num, sig, sigh, smc_index, sst, submodel, support_transparent_background, sz, t, target_api, tid, time_in_session, u_audio, u_h, u_sd, u_so, u_tz, u_w, uach_m, unviewed_position_start, url, urlfix, v, vnm, wv_count, xai, z

***.v.fwmrm.net**
 _fw_did_google_advertising_id, _fw_dpr, _fw_gdpr, _fw_gdpr_consent, _fw_player_height, _fw_player_width, adid, arid, auth, cn, competition, et, event, f, init, iw, kv, limittracking, n, os, os_version, r, reid, s, slid, sport, t, tpos, uxct, uxnw, uxss, video_targeting

teads.tv
 action, apiFrameworks, app_bundle_id, app_name, appId, appVersion, auctid, auction_currency, auction_price, browser, carrier, checksum, cid, country, cph, crevenue, crevenue_advertising, crevenue_curr, crevenue_platform_fee, cs, cts, deviceFamily, deviceType, dr, env, f, fms, fv, gdprlab, gid, inte, locale, network, omidPn, os, osVersion, p, pageId, payload, pfid, pid, piv, pscid, psid, random, rcid, referer, revenue, revenue_advertising, revenue_curr, revenue_platform_fee, scid, screenHeight, screenWidth, sdkEngineVersion, sdkIntegrationType, sdkVersion, sid, slot, srevenue, srevenue_curr, srevenue_fp, studio_cid, sv, tag, ts, userId, ut, vid, windowReferrerUrl

***.ads.stickyadstv.com**
 _fw_content_category, _fw_content_genre, _fw_gdpr, _fw_gdpr_consent, appBundle, appName, appStoreUrl, deviceIfa, loc, playerSize, protocolVersion, reqType, zoneId

***.google.com**
 app, app_ver, auid, cert, delete, device, frm, gclid, gclsrc, gcm_ver, info, plat, sender, target_ver, tfd, tft, unreg_cause, url, X-app_ver, X-app_ver_name, X-appid, X-cliv, X-firebase-app-name-hash, X-gmp_app_id, X-gmsv, X-Goog-Firebase-Installations-Auth, X-osv, X-scope, X-subtype

***.production.apptentive.com**
 app_release, board, bootloader_version, brand, build_id, build_type, carrier, client_created_at, client_created_at_utc_offset, cpu, current_carrier, custom_data, debug, device, identifier, inheriting_styles, integration_config, label, locale_country_code, locale_language_code, locale_raw, manufacturer, min_sdk_version, model, network_type, nonce, os_api_level, os_build, os_name, os_version, overriding_styles, person, product, radio_version, sdk_distribution, sdk_distribution_version, sdk_nonce, sdk_platform, sdk_version, target_sdk_version, type, utc_offset, uuid, version_code, version_name

***.google-analytics.com**
 _gmsv, _s, _v, a, adid, aid, aiid, an, ate, av, cd, cd2, cd5, cid, ea, ec, el, ht, id, pv, qt, rv, sr, t, tid, uid, ul, v

assets.adobedtm.com
 Unintelligible

***.pushwoosh.com**

Continued on next page

Table D.1 – Continued from previous page

application, device_type, hwid, language, userId, v
careers.bupa.com.au
applyUrl, candidateCardPageId, categoryIds, dclid, geoType, hasHtml5GeoError, IsUsingGeolocation, jobFeedId, jobId, jobOrganizationId, lat, locationIds, lon, mobileApplyUrl, s_cid, searchAnalyticsCurrentJobId, url, urlReferrer
*.litix.io
ake, dcuva, e, fnm, mapve, mem, memve, mvrld, pht, piiti, pinid, pispa, pmxpinm, pmxpive, pnm, ppgti, pphti, psqno, pswnm, pswve, pve, pwd, qbyld, qcule, qhn, qlbbi, qmddu, qmdstti, qrpen, qrphe, qrpst, qty, qur, qviht, qviwd, rtt_ms sex, sid, sst, transmission_timestamp, ualnm, ualve, ucxty, udvnm, udvmo, uosar, uosfm, uosve, uti, uusid, vcetty, vdn, vdu, vid, visli, vsmtty, vsobi, vsofp, vsoht, vsour, vsowd, vtt, wty, xavrqth xavrqth, xctpbti, xid, xmaphps, xreco, xredu, xrqco, xsqno, xtcltpbti, xtlgd, xwati, ypyid
*.adnxs.com
aaid, an_audit, appid, bdifs, bdfref, bdtop, bh, bstk, bw, e, ft, id, jm, LimitAdTrackingEnabled, nmt, ph, pl, psa, pw, px, py, referrer, s, sh, sid, size, sv, sw, tag_id, tv, type, ua, vd, wh, ww, x
cookielaaw.org
Unintelligible
liftoff-creatives.io
ad_group_id, channel_id, creative_id, auction_id, loid, origin
linkedin.oribi.io
Unintelligible
arkoselabs.com
action, ag, analytics_tier, at, bda, bio, cache_key, category, cdn_url, challenge, data[status], game_token, game_type, gameToken, guess, guitextcolor, lang, lurl, meta, metabgclr, metaiconclr, onload, pk, public_key, r, render_type, rid, rnd, session_token, sessionToken, sid, site, smurl, style_theme, surl, token, userbrowser
*.facebook.com
Unintelligible
*.kochava.com
action, adid, app_limit_tracking, app_name, app_short_string, app_version, architecture, attempt_count, battery_level, battery_status, bms, consent, count, device, device_cores, device_limit_tracking, device_orientation, disp_h, disp_w, duration, experiencecloudid, huawei_referrer, identity_link, install_begin_time, install_referrer, installed_date, installer_package, instant_app, is_genuine, kochava_app_id, kochava_device_id, language, last_install, locale, manufacturer, marketingcloudvisitorid, metrics, min_api, network_conn_type, notifications_enabled, nt_id, os_version, package, platform, product_name, referrer, referrer_click_time, required, screen_brightness, screen_dpi, screen_inches, sdk_build_date, sdk_protocol, sdk_version, send_date, state, state_active, state_active_count, status, target_api, time, timezone, ui_mode, uptime, url, usertime, volume
*.gstatic.com
Unintelligible
*.cloudfront.net
Unintelligible
*.omtrdc.net

Continued on next page

Table D.1 – Continued from previous page

.a, .c, .cid, .DSID_20914, a., aamb, aamlh, accountStatus, adobeEcid, analytics.enableSSL, analytics.reportSuite, analytics.trackingServer, AppID, as, authState, brand, buildVersion, c., CarrierName, ce, channel, cid., competition, contentOwner, contentType, contentPosition, contentSiteSection, contentSubSection, contentSubSection2, contentSubSection3, cp, DailyEngUserEvent, DayOfWeek, DeviceName, discipline, DSID_20914, DSID_20914., embedded_status, environment, eventType, family, flag, format, gender, HourOfDay, id, InstallDate, InstallEvent, internalaction, language, Launches, LaunchEvent, locale, loginStatus, magazine, media.ad.podFriendlyName, media.ad.podIndex, media.ad.podSecond, media.channel, media.contentType, media.downloaded, media.id, media.length, media.libraryVersion, media.name, media.playerName, media.resume, media.streamType, mediaId, mid, MonthlyEngUserEvent, ndh, notificationStatus, offer_type, offerType, OSVersion, pageName, pageTitle, pageUniqueId, params, participants, pe, pev2, platform, playerTime, playhead, playlistId, playType, product, profileID, Resolution, round, RunMode, season, spoilerStatus, sponsoredFlag, sport, sportEvent, subscriptionUserID, t, TimeSinceLaunch, transmission_type, trigger, ts, videoformat, videoType, visitor.aamLocationHint, visitor.customerIDs, visitor.marketingCloudOrgId, visitor.marketingCloudUserId
*.demdex.com d_blob, d_cid_ic, d_mid, d_orgid, d_rtbd, d_ver, dcs_region, device_consent
zdassets.com Unintelligible
akamai.prod-live.h264.io Unintelligible
eurosport.zendesk.com Unintelligible
crashlytics.com 1120225978 11632926720 129516972, 2532122624, actionType, activityHistory, analyticsEvents, androidClientInfo, appBuild, appInfo, applicationBuild, appName, appToken, appVersion, architecture, build_version, buildId, bundleId, bytesReceived, bytesSent, carrier, category, cause, className, clientInfo, clientType, connectionType, contentType, country, crashed, dataToken, device, deviceInfo, deviceManufacturer, deviceModel, deviceName, deviceId, diskAvailable, display_version, eventTimeMs, eventType, eventUptimeMs, exception, fileName, fingerprint, guid, hardware, id, instance, lastInteraction, lineNumber, locale, logEvent, logSourceName, manufacturer, mccMnc, memoryUsage, memUsageMb, methodName, model, modelName, name, networkConnectionInfo, networkStatus, networkType, newRelicVersion, nr.responseBody, obfuscated, orientation, osBuild, osMajorVersion, osName, osVersion, platform, platformVersion, priority, processId, product, protocolVersion, qosTier, requestDomain, requestMethod, requestPath, requestTimeMs, requestUptimeMs, requestUrl, responseTime, runTime, screenResolution, sdkVersion, sessionAttributes, sessionDuration, sessionId, source, sourceExtension, sourceExtensionJsonProto3, stack, state, statusCode, threadId, threadNumber, threads, timeSinceLoad, timestamp, timezoneOffsetSeconds, trace.id, uuid
*.sports.gracernote.com Unintelligible
liftoff.io

Continued on next page

Table D.1 – Continued from previous page

ad_group_id, channel_id, creative_id, auction_id, loid, origin
*.newrelic.com actionType, bytesReceived, bytesSent, category, connectionType, eventType, guid, id, platform, platformVersion, requestDomain, requestMethod, requestPath, requestUrl, responseTime, size, statusCode, timeSinceLoad, timestamp, trace.id
onetrust.io Country, Id, identifier, InteractionType, isAnonymous, requestInformation, syncGroup, tcStringV2, test, TransactionType, UserAgent
*.tntsports.io Unintelligible
*.twimg.com Unintelligible
casalemedia.com app.bundle, app.content.contentrating, app.content.genre, app.content.title, app.storeurl, device.ifa, device.ip, device.ua, ext.prebid.storedrequest.id, id, imp.0.ext.prebid.storedrequest.id, imp.0.id, imp.0.video.h, imp.0.video.w, regs.ext.gdpr, requesttype, user.ext.consent
*.mercury.dnity.com Unintelligible
*.twitter.com Unintelligible
*.linkedin.com cookiesTest, fmt, liSync, pid, redirect, time, url, v
moatads.com ar, bd, bo, bq, cb, cm, cs, cu, d, de, dnt, e, em, en, f, fd, fs, gu, gw, hp, hq, hr, hs, ht, hu, i, id, ih, ii, it, iw, j, jk, jm, kq, ll, lm, ln, lv, m, na, pe, pxm, q, r, sgs, t, ti, vb, vz, wf, yl, ym, zGSRC, zl, zMoat_ad_entity_id, zMoat_connection, zMoat_connection_entity_id, zMoat_domain, zMoat_pid, zMoat_subdomain, zMoat_wid, zMoatAuctID, zMoatOrigSlicer1, zMoatOrigSlicer2, zp
*.scorecardresearch.com c1, c12, c2, name, ns_ak, ns_ap_ais, ns_ap_an, ns_ap_ar, ns_ap_as, ns_ap_bi, ns_ap_bt, ns_ap_bv, ns_ap_cfg, ns_ap_cs, ns_ap_csf, ns_ap_das, ns_ap_dbt, ns_ap_device, ns_ap_dft, ns_ap_dit, ns_ap_ec, ns_ap_env, ns_ap_ev, ns_ap_fg, ns_ap_ft, ns_ap_gs, ns_ap_i3, ns_ap_id, ns_ap_install, ns_ap_it, ns_ap_jb, ns_ap_lang, ns_ap_pfm, ns_ap_pfv, ns_ap_pn, ns_ap_po, ns_ap_pv, ns_ap_res, ns_ap_sd, ns_ap_smv, ns_ap_sv, ns_ap_usage, ns_ap_ut, ns_ap_ver, ns_radio, ns_ts, ns_type
services.tmpwebeng.com Unintelligible
snap.licdn.com Unintelligible
tbcn.talentbrew.com Unintelligible
tpc.googlesyndication.com adk, app, avms, bin, bs, cr, ec, ffslot, id, io2, isd, itpl, la, le5-xQ, lsd, mc, mcvt, met, mtop, mtos, n, nXlftA, p, pbe, r, reach, rpt, rs, rst, sft, sig, spb, st, sv, tid, tos, tv, v, vae, vs, vu, wmsd, xai
*.pubmatic.com

Continued on next page

Table D.1 – Continued from previous page

adId, adtype, bundle, kadpageurl, kadudid, kadudidhash, kadudidtype, placement, pubId, sec, siteId, storeurl, vadFmt, vcom, vfmt, vh, vmaxbtr, vmaxl, vminl, vplay, vpos, vskip, vskipdelay, vtype, vw
*.batch.com
Unintelligible
*.tremorhub.com
adCode, appBundle, appName, appStore, appStoreId, deviceDNT, deviceId, deviceUA, gdpr, gdpr_consent, pchain, playerHeight, playerWidth, schain

Table D.2: Objects of the traffic intercepted from Strava

*.gstatic.com
Unintelligible
*.google.com
_internal_experimental_sets, @os, @updater, accept_locale, acceptformat, app, app_ver, appid, arch, avx, cert, channel, cohort, cohortname, dedup, device, dg, download_time_ms, downloaded, downloader, enabled, event, eventresult, eventtype, fp, gcm_ver, hw, info, installdate, ismachine, k, lang, mav, milestone, mp, nacl_arch, nextfp, nextversion, os, osname, p, package, packages, phymemory, ping, ping_freshness, plat, platform, previousfp, previousversion, prodversion, protocol, rd, request, requestid, sender, sessionid, sse, sse2, sse3, sse41, sse42, ssse3, tag, target_ver, total, updatecheck, updaterversion, url, version, X-app_ver, X-app_ver_name, X-appid, X-cliv, X-gmp_app_id, X-gmsv, X-osv, X-scope, X-subtype
*.appsflyer.com
advertiserId, advertiserIdEnabled, af_events_api, af_preinstalled, af_timestamp, af_v, af_v2, app_version_code, app_version_name, appsflyerKey, arch, brand, btch, btl, build_display_id, carrier, cksm_v1, counter, country, cpu_abi, cpu_abi2, d_dpi, date1, date2, device, deviceData, deviceType, dim, eventName, eventValue, firstLaunchDate, iaecounter, installDate, isFirstCall, isGaidWithGps, lang, lang_code, model, network, operator, platformextension, prev_event, product, registeredUninstall, sdk, size, timepassedsinclastlaunch, uid, x_px, xdp, y_px, ydp
*.s3-accelerate.amazonaws.com
Format, jfif_version, jfif_density, jfif_unit, Size
*.bugsnag.com
"releaseStage", "type", "version", "versionCode", "device", "cpuAbi", "jailbroken", "manufacturer", "model", "osName", "osVersion", "androidApiLevel", "osBuild", "name", "url", "version", "id", "startedAt", "user", "id"
*.googleusercontent.com
Unintelligible
*.facebook.com
fields, format, sdk_version, sdk, platform
*.cloudfront.net
Unintelligible
flurry.com
Unintelligible
*.branch.io
app_version, branch_key, brand, cd, country, debug, device_fingerprint_id, environment, facebook_app_link_checked, first_install_time, google_advertising_id, hardware_id, identity, identity_id, install_begin_ts, instrumentation, is_hardware_id_real, is_referrable, language, lat_val, latest_install_time, latest_update_time, local_ip, metadata, model, mv, os, os_version, pn, previous_update_time, retryNumber, screen_dpi, screen_height, screen_width, sdk, session_id, ui_mode, update, v1/close-brtt, v1/close-qwt, v1/install-brtt, v1/install-qwt, v1/open-brtt, v1/open-qwt, v1/profile-qwt, wifi
app.adjust.com

Continued on next page

Table D.2 – Continued from previous page

android_uuid, api_level, app_token, app_version, callback_params, country, created_at, device_manufacturer, device_name, device_type, display_height, display_width, environment, event_count, event_token, gps_adid, language, last_interval, needs_response_details, os_name, os_version, package_name, screen_density, screen_format, screen_size, sent_at, session_count, session_length, subsession_count, time_spent, tracking_enabled
api.iterable.com
applicationName, email, platform, preferUserId, token, userId
*.elephantdata.net
data, hdle, sign, pkey,
edgedl.me.gvt1.com
Unintelligible

Table D.3: Objects of the traffic intercepted from Tinder

*.google.com
Unintelligible
*.branch.io
data, device_fingerprint_id, identity_id, link, session_id, uri_skip_list, version
*.gvt3.com
age, elapsed_time, method, phase, protocol, referrer, sampling_fraction, server_ip, status_code, type, type, url, user_agent, age, elapsed_time, method, phase, protocol, referrer, sampling_fraction, server_ip, status_code, type, type, url, user_agent
crashlytics.com
x-crashlytics-developer-token, x-crashlytics-device-model, x-crashlytics-installation-id, x-crashlytics-os-display-version, x-crashlytics-api-client-version, user-agent, x-crashlytics-api-client-type, x-crashlytics-google-app-id, x-crashlytics-os-build-version, instance, build_version, display_version, source
*.facebook.com
access_token, fields, format, sdk_version, sdk, platform
*.scdn.co
Unintelligible
*.appsflyer.com
Unintelligible
*.instagram.com
_nc_cat, ccb, _nc_sid, _nc_ohc, _nc_ht, edm, oh, oe
*.bugsnag.com
app, binaryArch, buildUUID, id, releaseStage, type, version, versionCode, device, cpuAbi, id, jailbroken, locale, manufacturer, model, osName, osVersion, androidApiLevel, osBuild, totalMemory, name, url, version, id, startedAt, user, id
*.adnxs.com
an_audit, bh, bw, d, d0, d100, d25, d50, d75, e, ft, ic, id, jm, nvt, pd, pl, px, py, referrer, s, sf, sid, sv, tag_id, tv, type, ua, vd, x
*.s3.amazonaws.com
Unintelligible

Table D.4: Objects of the traffic intercepted from ASKfm

<p>*.pollfish.com accessibility_enabled, api_key, app_id, app_version, board, brand, con_type, custom_init, debug, developer_enabled, device_descr, device_id, encryption, gender, google_play, hardware_accelerated, install_non_market_apps, is_roaming, locale, manufacturer, nfc_enabled, nfc_exists, opt_out, os, os_ver, position, provider, provider_mcc, provider_mnc, scr_h, scr_size, scr_w, survey_format, target, timestamp, usr_agent, version, video, year_of_birth</p>
<p>*.appsflyer.com advertiserId, advertiserIdEnabled, af_events_api, af_preinstalled, af_timestamp, af_v, af_v2, app_version_code, app_version_name, appsflyerKey, arch, brand, btch, btl, build_display_id, carrier, cksm_v1, counter, country, cpu_abi, cpu_abi2, d_dpi, date1, date2, device, deviceType, dim, firstLaunchDate, iaecounter, installDate, isFirstCall, isGaidWithGps, lang, lang_code, model, network, operator, platformextension, product, registeredUninstall, sdk, size, timepassedsincelastlaunch, uid, x_px, xdp, y_px, ydp</p>
<p>*.facebook.com _rdt, api_key, app_id, auth_type, cancel_url, client_id, default_audience, display, e2e, error_code, kid_directed_site, locale, next, pl_dbl, redirect_uri, refsrc, response_type, return_scopes, scope, sdk, signed_next, skip_api_login, state</p>
<p>*.fbne3-1.fna.fbcdn.net Unintelligible</p>
<p>*.google-analytics.com Unintelligible</p>
<p>*.applovin.com app_info, app_name, app_version, applovin_sdk_version, first_install, ic, installed_at, installer_name, package_name, device_info, adns, adnsd, adr, brand, brand_name, dnt, gy, hardware, idfa, locale, model, orientation_lock, os, revision, sdk_version, sim, type, tz_offset, volume, wvvc, sc, stats, TaskFetchBasicSettings_count, TaskFetchBasicSettings_time</p>
<p>upload.video.google.com Unintelligible</p>
<p>*.giphy.com Unintelligible</p>
<p>i2w.io Unintelligible</p>
<p>flurry.com Unintelligible</p>
<p>*.appnext.com Unintelligible</p>
<p>*.pushwoosh.com android_package, app_version, application, attributes, device_model, device_name, event, hwid, idfa, jailbroken, language, os_version, push_token, sounds, timestampCurrent, timestampUTC, timezone, userId, v</p>
<p>mopub.com assets, av, cn, ct, current_consent_status, force_gdpr_applies, h, MAGIC_NO, mcc, mnc, o, sc, st, uidid, w, z</p>

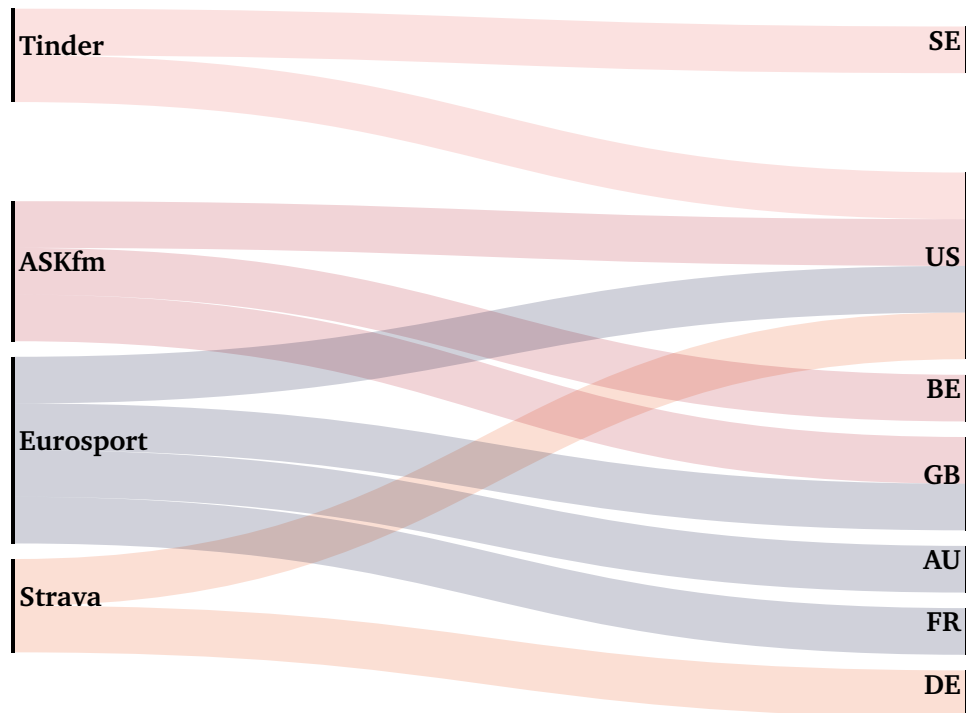


Figure D.1: SANKEY diagram presenting the origin of the third parties



 **NTNU**

Norwegian University of
Science and Technology