TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

Julius Lwanga 195665IVGM

ASSESSING THE PERCEPTION OF THE ESTONIAN COVID-19 CONTACT TRACING APP

An investigation into the adoption and impact (successes and failures of HOIA)

Master's thesis

Supervisor: Vishwajeet Pattanaik
M.Tech

Tallinn 2020

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Julius Lwanga 195665IVGM

Eesti COVID-19 kontaktjälitusrakenduse tajumise hindamine

Vastuvõtu ja mõju uurimine (Hoja edusammud ja läbikukkumised)

Juhendaja: Vishwajeet Pattanaik
M.Tech

Tallinn 2020

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Julius Lwanga

10.05.2021

14,275 words

Abstract

The COVID-19 pandemic has brought about huge societal and economic consequences across the globe. The continuous spread of the virus in different waves around the globe, forced governments and public health institutions to explore approaches on how to stop or slow the virus. One of these approaches is the use of new digital innovations known as Contact Tracing Applications (CTAs), which focus on the use of smartphones to detect residents' exposures to the virus as this can help slow the spread of COVID-19. One of those CTAs is Estonia's HOIA app which is based on a decentralised architecture and uses Apple/Google Exposure Notification APIs and a Bluetooth-based approach which provides for privacy and security. Despite the guarantee for privacy from the stakeholders of HOIA, the app's uptake has seen mixed results. This paper set out to assess the perception of HOIA using thematic analysis, textual/sentiment analysis, and survey analysis to show whether HOIA has been a success.

Keywords

Estonia, Contact Tracing Apps (CTAs), HOIA app, COVID-19, Perception, Success, Bluetooth-Based Approach, Decentralised Architecture, Privacy.

Annotatsioon

[Thesis title in Estonian]

[Tekst]

Lõputöö on kirjutatud [mis keeles] keeles ning sisaldab teksti [lehekülgede arv] leheküljel, [peatükkide arv] peatükki, [jooniste arv] joonist, [tabelite arv] tabelit.

List of abbreviations and terms

API	Application Programming Interface
BLE	Bluetooth Low Energy
COVID-19	Coronavirus disease
CTA	Contact Tracing App
DP-3T	Decentralised Privacy-Preserving Proximity Tracing
GPS	Global Positioning System
ID	Identification
MIT	Massachusetts Institute of Technology
NLP	Natural Language Processing
PACT	Private Automated Contact Tracing
PACT II	Privacy-Sensitive Protocols And Mechanisms for Mobile Contact Tracing
PII	Personally Identifiable Information
QTA	Qualitative Textual Analysis
RQ	Research Question
TCN	Temporary Contact Number
TEHIK	Welfare Information Systems Centre
TempID	Temporary ID
UUID	Universal Unique Identifier
Wi-Fi	Wireless Fidelity

Table of contents

List of figures

List of tables

1 Introduction

The COVID-19 pandemic has brought about huge societal and economic consequences across the globe. The continuous spread of the virus in different waves around the globe, forced governments and public health institutions to explore approaches on how to stop or slow the virus. One of these approaches is the use of new digital innovations known as Contact Tracing Applications CTAs, which focus on the use of smartphones to detect residents' exposures to the virus as this can help slow the spread of COVID-19 (Ferretti et al., 2020). CTAs automate the process of gathering data of close contacts of people who might have encountered individuals infected with the virus (Raman et al., 2021). The automation process is done to assess the risk of one getting the virus based on context, proximity, and duration of contact and this is done through Bluetooth or Global Positioning System (GPS) technologies (*ibid*).

The more traditional manual contact tracing approach initially adopted required many volunteers, consumed a lot of time, and struggled to identify all contacts before they caused further transmission. With the use of proximity sensors in smartphones, the digital approach of using CTAs offers a more efficient and responsive solution to slowing the spread of the virus (Rekanar et al., 2021). According to Hinch et al. (2020), who conducted a study in the United Kingdom (UK), the effectiveness of CTAs depends on the take-up of the apps across the population using it. The study recommends that the epidemic could be suppressed by 60% of the population using a contact tracing app (*ibid*). In Estonia, there are approximately 800,000 users of smartphones from a population of 1.325 million people (O'Dea, 2019), the UK study suggests that an app user baser of 608,000 (76.6% smartphone users) would suppress the spread of the virus in Estonia. As of 9th May 2021, the Estonian contact tracing app, HOIA has been downloaded 275,614 times with the 6,739 people having used it to mark themselves as infected with the virus (Terviseamet, 2021).

Like Estonia's HOIA, various CTAs have been released by countries around the world to help combat the spread of COVID-19 (Ahmed et al., 2020). Some of them have been designed with a centralized architecture, where all activities “are centred around a trusted server that supports storing encrypted information, analyses the risk of contacts and notifies the identified contacts” (*ibid*). Here, anyone who has access to the trusted server, can possibly know the identity of the app’s users. India’s Aarogya Setu, Australia’s COVIDSafe and Singapore’s TraceTogether are examples of this type of centralised architecture (Raman et al., 2021). On the other hand, some of the apps use a decentralized architecture which minimizes the information exchange and anonymizes user identification with the server (Bay et al., 2020). Switzerland’s SwissCOVID, Germany’s Corona-Warn-App and the Estonia’s HOIA are built based on this kind of architecture. There is also a hybrid model where only the personal identification and data of COVID-19 positive users or volunteers are saved on the servers (Raman et al., 2021). Apart from the technologies used, the adoption rates, and potential impact of the apps have been extremely varied across countries. In addition, the different architecture types through which the CTAs have been developed provide differing levels of security depending on the regulations of the countries they are rolled out in. Increased privacy has been deemed a fair trade-off for a decrease in the potential effectiveness of the app (Seto et al., 2021). They further cite the inability to download apps as a potential issue, because of the considerable number of people who do not own smartphones; this prevents people from benefitting from the Google and Apple Exposure Notification APIs that work on phones only released in the past five years, thereby excluding a significant part of society (*ibid*).

On the other hand, research on the adoption of technology similar to mobile phone CTAs has shown that users’ concern about data security and privacy can reduce user acceptance (Li et al., 2016). This has led to the design of CTAs especially in the Western world with a strong focus on guaranteeing privacy (Seto et al., 2021). Despite the guarantee of privacy and the potential to make the make process of contact tracing easier, according to Haggag et al. (2021) most people around the world are worried about using or even downloading CTAs. In countries where the use of CTAs is voluntary, few people have downloaded and used the apps because of privacy concerns (Muscato, 2021). Whilst there has been research on the challenges and successes of CTAs in other countries for example Germany, Switzerland, and Austria (Zimmerman et al., 2021), Australia (Garrett et al., 2021), India, Japan, Israel, UAE, France, Spain, Morocco, Brazil, Columbia, Canada, Qatar, Vietnam, Singapore (Raman et al., 2021), the United Kingdom (Williams et al., 2020), South Korea and the United States (Kim & Kwan, 2021). However, not many studies have been conducted with focus on the Estonian COVID-19 contact tracing app HOIA, at least to specifically understand its adoption rate and impact on the Estonian society.

1.1 Estonia's HOIA app

HOIA is a free app available for Android 6+ and iOS 13.5+ mobile phones that helps limit the spread of the COVID-19 virus, and it is the official COVID-19 contact tracing app for Estonia. It is owned by the Estonian Health Board and has been created in collaboration between the Estonian Health Board, the Ministry of Social Affairs, the Health and Welfare Information Systems' Centre and a voluntary consortium of Estonian companies (Health and Welfare Information Systems Centre, 2020a). The Estonian consortium composed of the following companies: Fujustu Estonia, Mobi Lab, Cybernetica, Guardtime, ASA Quality Services, Mooncascade, Heisi IT, FOB Solutions, Icefire, Velvet and Iglu (*ibid*).

The purpose of the app is to determine whether one has been in contact with someone who marked themselves as sick and notify them about the possible exposure with information on what to do next (Health and Welfare Information Systems Centre, 2020a). The app is used on a voluntary basis (*ibid*). The app is based on an anonymous contact diary that logs the various encounters via BLE (Martin et al., 2020). The Estonian Contract Tracing app, HOIA is based on a decentralised system architecture which uses the Apple/Google Exposure Notification APIs *(ibid*) and the DP-3T systems which provide security and privacy protections by "ensuring data minimization, preventing abuse of data, preventing tracking of user and the system dismantles itself at the end of the pandemic" (Troncoso, 2020). The app has been developed according to privacy preserving principles and the notifications sent through the app do not disclose the identity of the infected person. The source code of the app can be found on GitLab (Health and Welfare Information Systems Centre, 2020b).

When it comes to installation, the HOIA app requires no registration or personal information to install and use the app (Martin et al., 2020). During the app installation, a random Universal Unique Identifier (UUID) is generated by the app and then the app updates this Identification (ID) frequently (*ibid*). When two users are physically close, their smartphones send their current UUID to each other via Bluetooth Low Energy (BLE) (*ibid*). In regard to how the HOIA app functions, the app assesses the risk of the encounter based on its duration and the distance between the two smartphones. Phones onto which the app has been downloaded register Bluetooth signals from other nearby phones (Health and Welfare Information Systems Centre, 2020a). When the signal is "sufficiently close and long enough, an anonymous code referring to a close contact is stored in the phone" (*ibid*). When a person marks themselves as infected in the HOIA app, the anonymous codes on their phone is uploaded to a central server from where all users of the app can download them (*ibid*). When the phone of another user of the app compares the anonymous codes on the server and matches it with the anonymous coded previously stored on their phone, then the user is considered to be a close contact and notified on how to proceed (*ibid*). Information that would lead to indirect identification of the infected person is not revealed in any way to the other app users (*ibid*). When it comes to data processing, the details regarding the privacy policy of the app can be found in the document at Github (Health and Welfare Information Systems Centre, 2020c). The app including server is operated in the state Cloud Server in Estonia, managed by the Estonian Health and Welfare Information Systems Centre (TEHIK) and the servers are located in Estonia (*ibid)*.

1.2 Motivation for the research

The research motivation and idea of this paper comes from the author's interest in the success and failure of technological tools designed to tackle challenges that the society is facing today. In the age of the COVID-19 pandemic that has brought tremendous societal and economic challenges, it is a good thing to have new innovations that provide solutions to ease the crisis. CTAs were deemed to transform the process of tracing contacts who had the virus and complement the testing and isolation solutions to combat the spread of the Corona virus. However, whether CTAs have significantly changed the dynamics in the way the pandemic has been dealt with, requires further investigation. Hence, research is needed to study CTAs, their perception, and how they affect the success in slowing or stopping the spread of COVID-19.

In Estonia, a technology savvy nation, the launch of the contact tracing app which has a strong focus on privacy, should have helped significantly in easing the contact tracing process to help combat the spread of the virus and to complement the traditional manual methods. However, the app has brought about mixed results. This paper sets out to do a case study on the HOIA app through a Qualitative and Quantitative methods to assess the perception of the app in Estonia.

To better understand this, and to assess the perception of HOIA in Estonia, this paper sets out to conduct a case study analysis on the HOIA app through a Qualitative and Quantitative Research methods. The paper makes use of Thematic Analysis of open-ended interviews with the stakeholders to assess the key aspects that were taken into consideration when developing the app. It uses a thorough literature review to compare the architecture of HOIA, its features and user acceptance with other CTAs deployed around the world. Furthermore, the paper uses Qualitative Textual Analysis (QTA) of user reviews from Google Play Store and articles from news media to assess the perception of the public and media regarding the HOIA app. This is complimented finally with a Quantitative Analysis of residents' survey about the perception of the HOIA app. The results are then triangulated, and a discussion is provided to show whether HOIA has been a success by comparing its adoption and assessing its impact on society. Lessons are then drawn for both policy makers and the consortium of companies that developed the app. In the next section, a background of CTAs is provided.

2 Background

According to Harri (2020) "contact tracing is a time-proven tool used to limit the spread of infectious diseases that dates back to the 16th century". Contact tracing traditionally has been through interviewing infected individuals manually by the health authorities (*ibid*). The interviews are carried out with the aim of collecting information about the contacts the infected individuals have had (Ahmed et al., 2020). The information is then used by health officials to compute the risk score for each of the contacts based on the context, duration and proximity (*ibid*). In Estonia, the Estonian Health Board receives notifications on individuals who test positive for COVID-19 and officials from the board then call the individuals to inform them about their exposure to the virus, provide them with advice on what to do next, and to obtain information from them about their previous contacts and movements (*ibid*). This is done to find out any close contacts they may have had, so the board can notify them about the risk of infection that they pose to others.

However, it is challenging for people to accurately recall each person that they may have met in the last fourteen days and the infected individual may also have infected many other people for whom it would be difficult for them to identify (for example, unknown contacts on public transport, in grocery stores, and in public spaces). Furthermore, the interviews require a large number of health officials or volunteers trained in the process of manual contact tracing to effectively be able to trace all the people the infected individuals came in contact with. This makes it cumbersome and costly to run an efficient manual tracing process especially in situations where the virus spreads at fast rate. With this background, researchers have been focusing on "technological solutions to automate the contact tracing process with the aim of quickly and reliably identifying contacts that might be at significant infection risk" (Ahmed et al., 2020).

The widespread use of smartphones and their ability to keep track of their location for example through GPS and Wireless Fidelity (Wi-Fi) along with their in-built Bluetooth interface that allows communication and proximity detection with nearby smartphones, makes them ideal devices for automated and reliable contact tracing (*ibid*). This has resulted into the development and deployment of many smartphone CTAs for COVID-19 (Vaudenay, 2020). Many countries have adopted CTAs to stop or slow the spread of

COVID-19 and help with the process of contact tracing (Kostka & Habich-Sobiegalla, 2020). However, there has been a difference in preferences amongst the countries that have deployed CTAs mainly in terms of app design, data privacy and storage, involvement of private companies and research institutes in the development, and speed of adoption (*ibid*). For instance, most European countries opted for higher privacy-preserving CTAs relying on Bluetooth technology, like Austria, France and Germany (Toussaert, 2021). Other countries chose more centralized approaches, such as China, South Korea, Singapore, and Israel (Kostka & Habich-Sobiegalla, 2020). In the US, individual states adopted CTAs which predominantly relied on Bluetooth technology (*ibid*).

CTAs are designed to automate the process of gathering data of contacts of people with infected individuals, to identify the risks based on the context, proximity and duration of contact. Most CTAs are based on a Bluetooth-based approach or GPS-based approach to estimate exposure to the virus (Raman et al., 2021). The GPS-based approach allows for estimation of exposure related to surface transmission of disease and can help notify users if they were in a location shortly after a person infected with COVID-19, when the chance for exposure to the virus through commonly touched surfaces is high (Raman et al., 2021). It also enables users to import historical data which can help alert users to potential exposures from their location history (*ibid*). Finally, the GPS-based approach provides redacted anonymized GPS data to help public health officials follow the spread of disease within a community, and since it's able to record the user's location with small amounts of data, it is more scalable and easier to implement in regions with high data costs (*ibid*). GPS is, however, not suitable in built up places as the accuracy is low in built-up spaces (*ibid*).

Bluetooth available in smartphones, supports proximity estimation. The Bluetooth-based approach uses signal strength, which is reduced by walls and other barriers, to estimate the distance between users (Raman et al., 2021). In some places, such as a large, multi-floor building, this estimate more accurately reflects the chance of exposure to disease than a GPS-based approach (*ibid*). It also uses time-range-dependent, randomly generated numbers as IDs to ideally achieve relative anonymity and there is no potential to collect historical data from before the user downloaded the app (*ibid*). Furthermore, the Bluetooth-based approach requires the use of a compatible app by other users to record possible exposures and if an app is not widely adopted, the potential utility is limited (*ibid*). However, the distance estimation may vary based on surrounding objects and the orientation of the phone thus leading to false positive and negative alerts (*ibid*).

The introduction of COVID-19 CTAs has led to a debate about their architecture, data management, efficacy, privacy, and security (Vaudenay, 2020; Criddle & Kellion, 2020). Most of the apps claim to be privacy-preserving, that is to say, they do not reveal any Personally Identifiable Information (PII), identity, or location information of the contacts without explicit user permission. According to Redmiles, privacy concerns associated with contact tracing apps are one of the factors that influence their adoption (Redmiles, 2020). Privacy advocates for COVID-19 CTAs are also concerned with the extent to which the apps can be re-purposed to track their users, and how the data collected may be used when the current pandemic ends (Ahmed et al., 2020). Existing surveys on COVID-19 CTAs (Li & Gou, 2020; Reichert et al., 2021) are either not comprehensive or focus only on user privacy (Tang, 2020). Ahmed et al., go beyond the vulnerabilities and privacy issues in design of the COVID-19 CTAs, study their implementation and usability issues, and find that these factors too have an impact on the adoption of the CTAs (Ahmed et al., 2020).

CTAs according to Ahmed et al. (2020) are classified into three main architectures: centralised, decentralised, or hybrid; and each of these classifications has an implication on the data security, privacy, and data management of the app. The core functionalities of a centralised system are performed by a central server processing user data, which is managed by a health authority and can (subject to permissions) notify an infected user's contacts of exposure (Horvath et al., 2020). A decentralised system, on the other hand, has most of its core functionalities performed by users' devices including exposure notifications (*ibid*). The privacy implications of these two systems have often been discussed as a trade-off with other attributes (Cioroianu & Dal, 2020). The next section delves into the different architectures of CTAs.

2.1 System Architecture for Contact Tracing Applications

According to Ahmed et al. (2020), “the type of architecture adopted for the data collection aspects of tracing apps has been a matter of much discussion due to both security and privacy concerns”. There are three distinct system architectures commonly used or proposed for developing COVID-19 CTAs, the centralised, decentralised, and the hybrid that combines features from both the centralised and decentralised architectures (*ibid*). The architectures are categorised based on the functionality and level of privacy preservation at the central server. In the centralised architecture, the server manages the security keys, generation of anonymous Identifications (IDs), contact risk analysis, and notification processes (*ibid*). All these roles are transferred to the devices in the decentralised architecture while the server acts simply as a bulletin board (*ibid*). The hybrid architecture tries to balance the load on the server and improve privacy preservation by splitting functionalities between the end-user device and the server (*ibid*).

2.1.1 Centralised Architecture

This is based on the Bluetrace protocol according to Bay et al., (2020). The initial requirement for the app is that a user has to pre-register with the central server (*ibid*). The server generates a privacy-preserving Temporary ID (TempID) for each device, and this TempID is then encrypted with a secret key (known only to the central server authority) and sent to the device (Ahmed et al., 2020). Devices exchange these TempIDs (in Bluetooth encounter messages) when they come in close contact with each other (*ibid*). Once a user tests positive, they can volunteer to upload all their stored encounter messages to the central server which then maps the TempIDs in these messages to individuals to identify at-risk contacts (*ibid*).

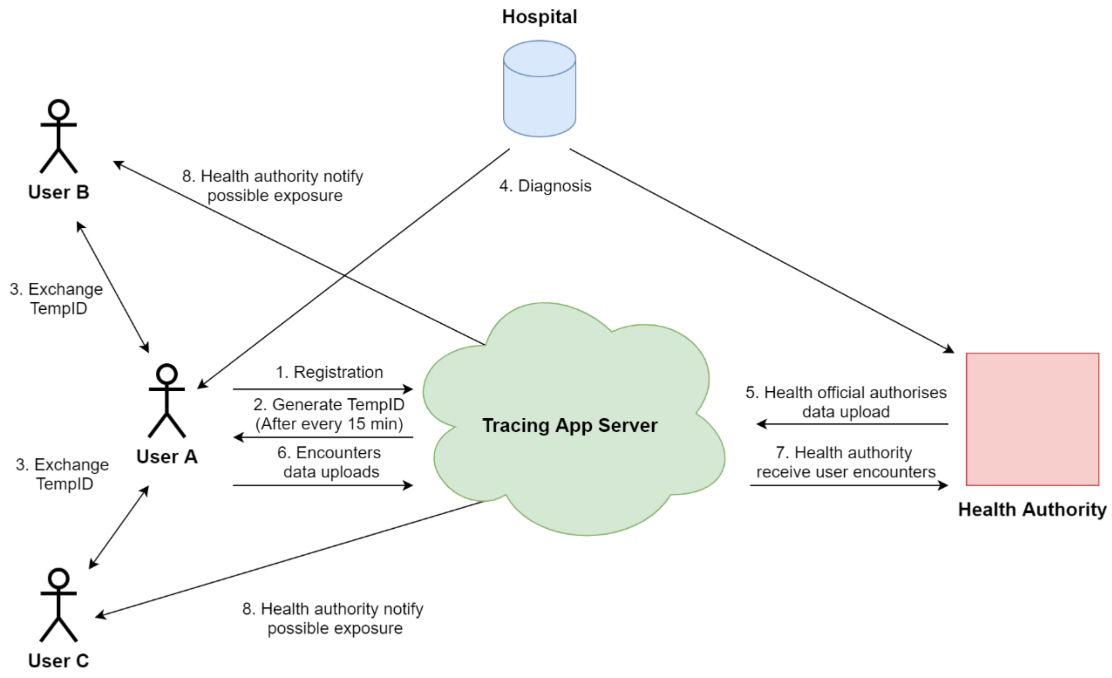


Figure 1. Centralised Architecture of CTAs (Ahmed et al., 2020)

In the centralised architecture (see *Figure 1*), the central server plays a key role in performing core functionalities such as storing encrypted PII, generating anonymous TempIDs, risk analysis, and notifications for close contacts (*ibid*). This accumulation of responsibilities has raised privacy concerns (Sowmiya et al., 2021). Additionally, the server is assumed trusted and some countries for example Australia have introduced strict privacy protection regulations to safeguard the use and life cycle of the collected data by the CTA (Greenleaf & Kemp, 2020).

2.1.2 Decentralised Architecture

The decentralised architecture (see *Figure 2*) in contrast to the centralised architecture makes a case for moving core functionalities to the user devices, leaving the server with minimal involvement in the contact tracing process (Ahmed et al., 2020). The idea is to enhance user privacy by generating anonymous identifiers at the user devices (keeping real user identities secret from the other users as well as the server) and processing the exposure notifications on individual devices instead of the centralised server (*ibid*).

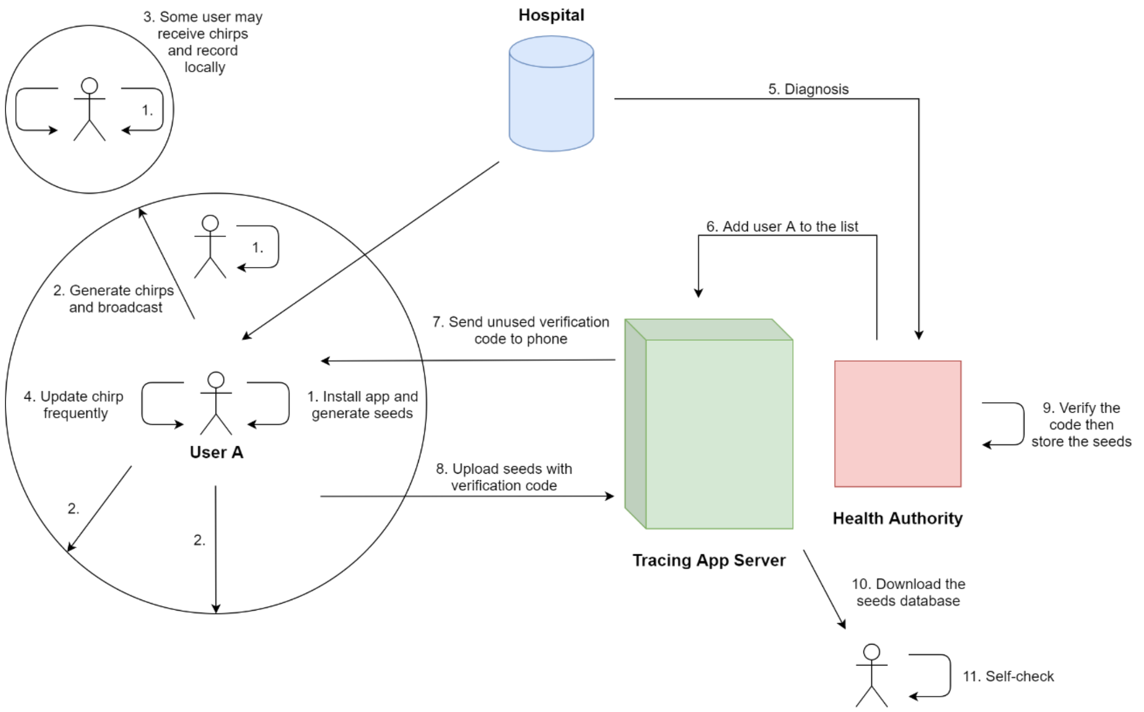


Figure 2. Decentralised Architecture of CTAs (Ahmed et al., 2020).

2.1.3 Hybrid architecture

The hybrid architecture (see *Figure 3*) is a mixture of the centralised and decentralised architectures. In the centralised architecture, the server performs all the complex tasks for example TempID calculations, encryption, decryption, risk analysis, and notifications of alerts for the at-risk contacts (*ibid*). All these functionalities are delegated to devices in the decentralised architecture, "keeping the server only as a bulletin board for lookup purposes" (*ibid*). The hybrid architecture proposes that these functionalities are shared between the server and the devices. In particular, the TempID generation and management remain decentralised, that is to say, they are handled by devices to ensure privacy and anonymisation, whilst the risk analysis and notifications should be the responsibility of the centralised server (*ibid*).

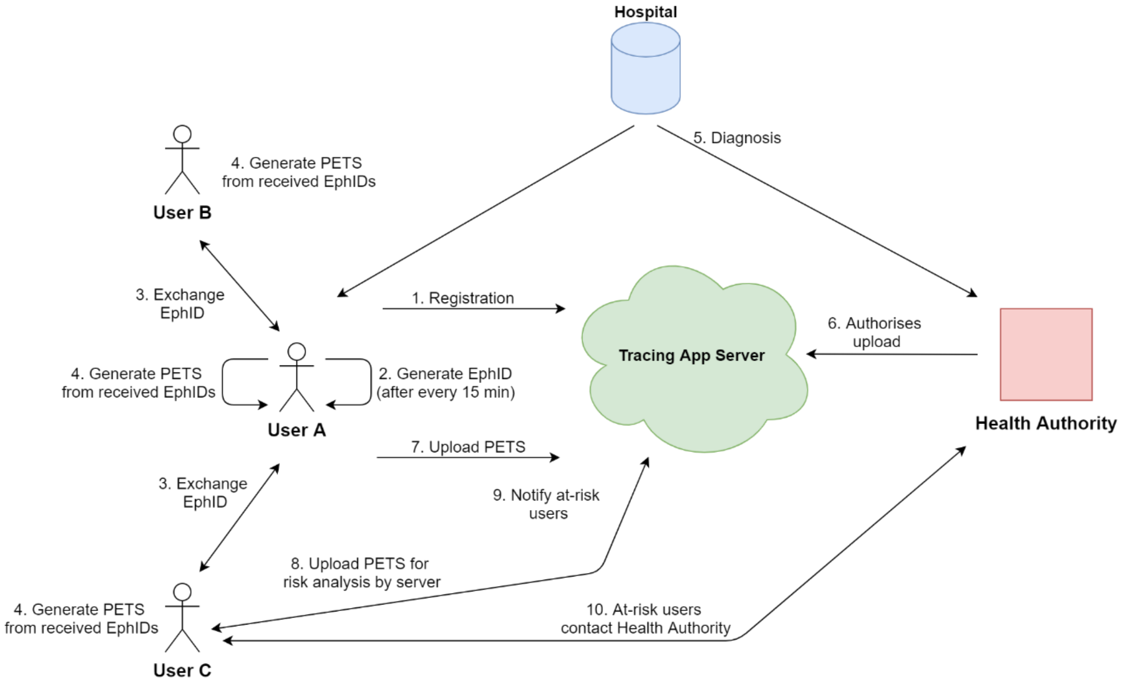


Figure 3. Hybrid Architecture of CTAs (Ahmed et al., 2020).

In summary, the advantage of using an architecture that pushes the risk analysis and notification process to the centralised server (centralised and hybrid architectures) is that health officials can decide the rate of notifications depending on the pandemic circumstances, for example, the availability of test kits (*ibid*). On the other hand, decentralised and hybrid architectures aim to keep the user identities secret from the central server hence, a server security breach in this architecture would result in lower information leakage, hence it being preferred for applications that priorities privacy. The Estonian contact tracing app in particular, uses the decentralised architecture because privacy was of high concern to the stake holders when developing and creating the app.

2.1.4 Classification of CTA Protocols based on Architecture

2.1.4.1 Centralised architecture

When it comes to the centralised architecture, CTAs have been developed based on Bluetrace protocol (Bay et al., 2020), The ROBERT protocol (Castelluccia et al., 2020) and those that use a combination of Bluetooth and GPS. The Singaporean Tracetogether app (Leith & Farrell, 2020) and the Australian CovidSafe app (Greenleaf & Kemp, 2020) are based on the Bluetrace protocol. The French StopCovid app is based on the ROBERT protocol and the Indian Aarogya Setu uses both Bluetooth and GPS.

2.1.4.2 Decentralised architecture

With the decentralised architecture, Apple and Google collaborated to support privacy-preserving contact tracing protocol by developing an exposure notification system and providing OS-level support (Leith & Farell, 2020). The Private Automated Contact Tracing (PACT) is another example of a protocol design based on the decentralised architecture and was developed through a research collaboration led by MIT (Chakraborty et al., 2020). Another protocol sharing the same name, "PACT II (Privacy-Sensitive Protocols and Mechanisms for Mobile Contact Tracing), developed by a team from the University of Washington" (Chan et al., 2020). The other protocols that use the decentralised architecture are Decentralised Privacy-Preserving Proximity Tracing (DP-3T) (Troncoso, 2020), Temporary Contact Number (TCN) (Coalition, 2020) developed by researchers from Stanford University and Waterloo University, Hamagen (Sun et al., 2021) developed by Israel's Ministry of Health, COVID Safe Paths (Ocheja et al., 2020) and Pronto-C2 (Avitabile et al., 2020) designed by researchers from the University of Salermo. The SwissCovid – DP 3T app is based on Apple/Google Exposure Notification Application Programming Interfaces (APIS), and the open-source app Corona-Warn-App from Germany is also based on the Apple/Google Exposure Notification APIs.

2.1.4.3 Hybrid Architecture

Hybrid protocols combine features of both centralised and decentralised architectures and here we have the DESIRE, Contra Corona and EpiOne (Ahmed et al., 2020). A summary of the different types contact tracing apps and protocols, is illustrated in *Figure 4*.

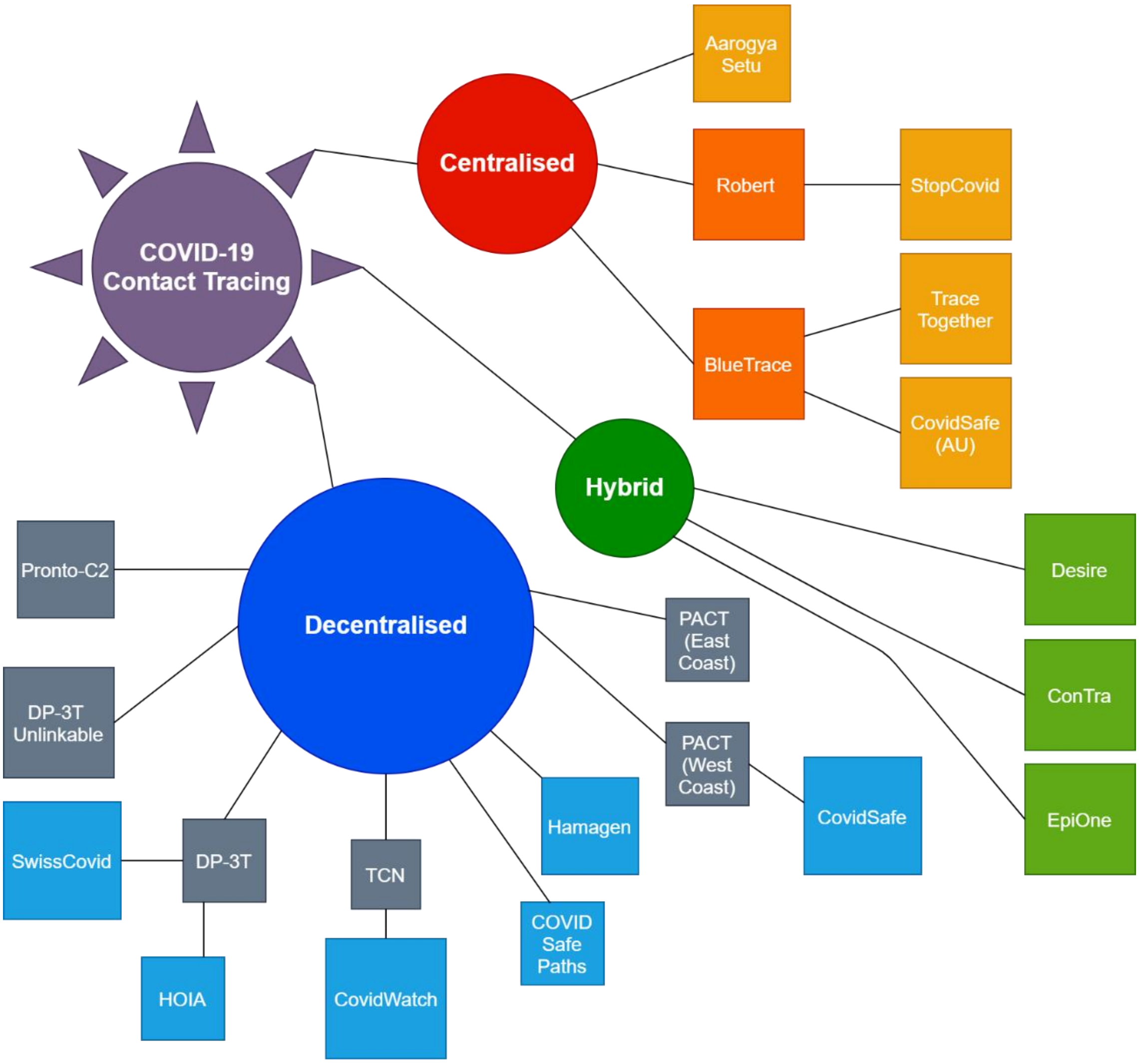


Figure 4. Summary of CTAs and protocols (Ahmed et al., 2020).

The Estonian contact tracing app, HOIA is based on the decentralised architecture and uses Apple/Google Exposure Notification APIs and the Decentralised Privacy-Preserving Proximity Tracing (DP-3T) and a Bluetooth-based approach. In comparison with other apps that use the centralised or hybrid architecture, the HOIA app provides for privacy and data security. The next section discusses the state-of-the-art in regards to CTAs.

3 State-of-the-art

Several COVID19 CTAs have been developed and released with the goal of tracking and reducing the spread of COVID-19 to complement manual tracing efforts (Kaptchuk et al., 2020). CTAs have begun to gain attention from academic researchers. The potential use of contact tracing in the control of COVID-19 has been well established by studies (Ferretti et al., 2020; Hellewell et al., 2020). Initial successful rollouts in Singapore and South Korea in 2020, encouraged more to countries to introduce CTAs in order to complement their contact tracing processes (Whitelaw et al., 2020). Many of the CTAs vary widely in terms of transparency, data destruction (for privacy), and the technology employed (Raman et al., 2021). In examination of 50 CTAs available in Google Play Store, representing a number of government-issued COVID-19 tracing apps that are from both developing countries and developed countries, Sharma & Bashir (2020) found that most common functionalities of the apps are: "live maps and updates of confirmed cases; real-time location-based alerts; systems for monitoring and controlling home isolation and quarantine, direct reporting to government, and self-reporting of symptoms; and education about COVID-19".

Given their importance in preventing and stopping the spread of COVID-19, this section discusses the state-of-the-art in the existing literature. Most of the available research has focused on features, content, and technical characteristics of the CTAs (Weiß, et al., 2021). Some studies have analysed users' behavioural intentions to use these types of applications and adopted an extended unified theory of acceptance and use of technology (UTAUT2) (Ezzaouia & Bulchand-Gidumal, 2021). According to Kaiser et al. (2021), digital solutions have proven to be able to strengthen our healthcare system in emergency situations. However, they also mention that the use of such solutions for gaining control over the spread of COVID-19 has showed that "there are risks may come with the utilization of digital solutions" (*ibid*). These risks can sometimes pose a challenge for the uptake and use of the CTAs, for example in countries where the of CTAs is voluntary, the apps have struggled with low uptake and privacy concerns (Raskar et al., 2020).

A few examples include the COVIDSafe app by Australian government, which faced resistance from the public in general due to concerns about privacy and was not entirely successful in providing contact tracing due to lack of widespread adoption of the application (Thomas et al., 2020). India's Aarogya Setu app, which is the most downloaded contact tracing application, effectively collects personal information and constantly adds to location information in real time for its users (Dhar, 2020). On the other hand, the Corona-Warn-App used in Germany posed no major privacy concerns as it was developed with potential privacy infringements in mind. Data is depersonalized and stored locally to avoid any concerns in this regard (Kozyreva, 2021).

While many countries such as India have made contact-tracing apps practically compulsory for activities such as air travel (Dhar, 2020), other countries such as the UK have refrained from employing similar procedures due to public concerns over privacy (Pagliari, 2020; Rekanar, 2021; Thomas et al., 2020). There are clearly diverse approaches taken by governments during this pandemic when it comes to mobile contact tracing. However, in general, CTAs have been employed as a novel tool for curtailing the spread of this disease. The concerns raised by many with regards to privacy and data security in relation to these applications are valid and need to be discussed in detail (Raman et al., 2021).

Privacy-by-design principles such as data minimization and purpose limitation have guided the development of many CTAs (Ahmed et al., 2020). However, while privacy-preserving design may contribute to an app's public acceptance (Trang et al., 2020), it impedes evaluation of the usage of the app (Munzert et al., 2021). Mere download statistics by the organisations running the CTAs are silent about actual usage and user profiles (Reelfs et al., 2020) and while surveys can be used to measure usage patterns (Altman et al., 2020) reliance on self-reports of socially sensitive behaviour can generate reporting biases (Munzert & Selb, 2020).

Seto et al. (2021) mention that "while the level of privacy required for a COVID-19 CTA will depend on factors including whether it is voluntary, the underlying technology, and degree of data centralization, translation of those important safeguards into a user's perception of privacy will occur within the context of the norms and values of their country". Furthermore, they add that residents prefer a balanced (human plus digital) approach to contract tracing and that privacy concerns were not as influential on the choice of the digital app" (*ibid*). According to their study, they find that "when a trusted public health provider is involved in the development and deployment of the tracing app it can bring about the cooperation of the public necessary for its successful use in reducing the spread of infection” (*ibid*).

4 Research Questions

Given the above background and motivation, the purpose of this paper is to provide answers to the research questions and sub questions by conducting a case study of the HOIA app. This combines both qualitative and quantitative methods to answer the research questions. A thematic analysis of interviews with stakeholders of the HOIA app, a textual/sentiment analysis of the user reviews on Google Play Store and media articles on the HOIA app, and finally a survey analysis from the Estonian residents about the HOIA app are used to answer the research questions and the sub questions below:

RQ.1 What were the key aspects taken into consideration when developing the Estonian COVID-19 CTA, HOIA?

a. Why the specific design of the app?
b. Why focus hugely on privacy as an issue?

Here the key aspects taken into the consideration when developing the Estonian COVID-19 CTA will be explored. The questions will be answered using a thematic analysis of qualitative interviews with the stakeholders of the HOIA app.

RQ.2 How is the Estonian COVID-19 CTA, HOIA different from other apps?

a. What were the novel features of the HOIA app (if any)?

Here, the analysis will focus on identifying of how different the Estonian HOIA app is different from other CTAs used elsewhere in the world. What makes it unique, and what features make it special in the process of contact tracing. To answer these questions, a thematic analysis of interviews with the stakeholders of the app is used together with literature review of the state-of-the-art.

RQ.3 How does the HOIA app compare in regard to user acceptance (when compared to other apps)?

a. What is the stakeholders’ perception regarding the app? And why?
b. What is the residents' perception regarding the app? And why?

Here, the focus of analysis will be on the user acceptance of the Estonian COVID-19 CTA. How many people have downloaded and used the app in the process of contact tracing? The literature review is used to make the comparison to user acceptance. The stakeholder's and residents' perceptions will be explored in regard to HOIA app using a thematic analysis, textual/sentiment analysis and a survey analysis of the residents' perception of HOIA.

RQ.4 Has HOIA been a success?

a. What lessons could be learned from it?

Here, the analysis is done by triangulation and summarization of the results from the literature review in the state-of-the-art, the thematic analysis, the textual/sentiment analysis and survey analysis.

5 Research Methodology

This paper uses a case study of the Estonian CTA, HOIA to assess the perception of the app to show whether it has been successful and what lessons can be drawn from it. A case study is defined as "an empirical method aimed at investigating contemporary phenomena in their context" (Runeson et al., 2012). According to Yin, a case study is appropriate when "a 'how' or 'why' question is being asked about a contemporary set of events, over which the investigator has little or no control" (Yin, 2018). Furthermore, a case study may often be used to help "contribute to our knowledge of individual, group, organizational, social, political, and related phenomena" (*ibid*). Since the research questions are attempting to understand a new and contemporary process of using CTAs in tackling the COVID-19 pandemic which deals with organizational, political and social phenomena, a case study is used to answer the research questions.

In order to do this, first a literature review on the current state of CTAs was conducted to show how the use of CTAs to stop or slow the spread of COVID-19 came about. The literature review also explores the main approaches on which the many of the CTAs are based, Bluetooth-based approach and the GPS-based approach. It further discusses architecture of the CTAs and their impact on issues like privacy, data security and management. This provides a background on why the designers of the HOIA app chose to design it the way they did, and why they focused on the key aspects that they did. The discussion on the different architectures deployed for the various CTAs also shows how HOIA is different from some of the other CTAs rolled out elsewhere in the world.

Second, the paper uses qualitative and quantitative methods to analyse the data on the HOIA app collected through interviews, user reviews on Google Play Store and media articles about the perception of the HOIA app, and surveys about the app. For the qualitative part, the paper uses a thematic analysis to analyse the interviews with the stakeholders in order to assess key aspects that were taken into consideration when developing the app, and uses a semantic analysis of user reviews from google store and articles in the media to assess the perception of the public regarding the HOIA app. A qualitative case study "is an approach to research that facilitates exploration of a phenomenon within its context using a variety of data sources" (Baxter & Jack, 2008).

This allows for the issue that is being explored to be viewed through a variety of lenses which "allows for multiple facets of the phenomenon to be revealed and understood" (*ibid*).

For the quantitative part, a residents' survey on the perception of the HOIA app is done to find out what the Estonian residents' perception is about the app. The results from the three methods are the included in the discussion part to answer and the research questions to show whether HOIA has been a success. The lessons are then drawn for both policy makers and the consortium of companies that developed the app.

5.1 Data Collection

When conducting a case study, one way to improve the validity of the study is to utilize multiple sources of evidence (Runeson, 2012; Yin, 2018). The multiple sources of evidence allow the researcher to triangulate the data which results into stronger conclusions from the case study research. The purpose of this study is to assess the phenomena of the Estonian CTA, and since this is a new tool being rolled out, knowledge about the CTAs is evolving rapidly. To ensure that the information derived from this research accurately reflects the truth about the HOIA app, different research methods are used to study it in this paper. Triangulation helps to increase the validity, reliability and legitimation of the research findings (Moon, 2019). This paper uses different sources of collecting data: a literature review, interviews with experts, user reviews from Google Play Store and articles about the perception of HOIA in the media, and finally a survey about the perception of HOIA app from the public.

First, two interviews are conducted with experts. For the public sector, interviews are conducted from Estonian Health Board that owns the HOIA app. From the private sectors' side, the interview was conducted from an official from Iglu, one of the companies that provided the technical lead and design of the HOIA app. The interview questions explored the aspects of what was taken into consideration when designing the app, how the stakeholders think the public currently perceive the HOIA App, and whether they think the app has been a success. The interviews are then transcribed using YouTube, and the transcript inserted into a word document for analysis.

Second, user reviews about the HOIA app on Google Play Store are obtained manually and transferred to a word document. The reviews were then sorted manually to remove texts responses from the Welfare information Systems' Centre (TEHIK) which were responses to the, and not the perception of the public. The text in Estonian were then translated to English using a machine translation engine, *Neurotõlge* developed by the NLP lab at the University of Tartu (Fišel, 2020). The one comment in the Russian language was translated using Google Translate into English. The fully translated text was then put into a notepad document and run through a code designed using RStudio to obtain knowledge graphs on the text.

A search of the media articles was done on the with a focus on the perception of the HOIA app. This included opinion posts from experts, regarding their perception of the app. Articles that were giving information on how to download the app, and articles marketing the app from the stakeholders were excluded in this search. In the end, four articles were chosen to be included in the study and the links to the articles can be found in Appendix 6. The articles were then translated from Estonian to English using the machine translation engine, *Neurotõlge* developed by the NLP lab at the University of Tartu (Fišel, 2020). The fully translated text of the articles was then put into a notepad document and run through code on RStudio to obtain a textual/sentiment analysis on the text. The code snippet of the same can be found in Appendix 5.

Finally, a survey was sent out to the public to assess the perception of the residents about the HOIA app. A survey form was designed using Google Forms. The survey contained both open and close ended questions regarding the HOIA app. It was sent out through social media platforms (Facebook, Instagram, and emails). The survey was explicitly designed for residents of Estonia, and it was distributed in groups or to people who were residents of Estonia because the HOIA app can only be used by Estonian residents in Estonia at the moment. The collected data was then analysed to get a richer view of the perception and user acceptance from the public's side about the HOIA app.

5.2 Textual/Sentiment Analysis

Sentiment about a product or a service offered is valuable in this era more than ever for organisations (Mukherjee, 2021). Knowing whether their established customer base and potential customers are showing a positive or negative sentiment toward their product or service can be game-changing for the organisations (*ibid*).

5.2.1 Background

Textual/Sentiment analysis which is the "extraction of people's opinions and sentiments embedded in user-generated content" (Li et al., 2019); can be used to gain insights into the perception, reasoning and decision behind the people forming the opinions. According to Li et al, sentiment analysis of social media text intends to extract the sentiment information embedded in messages posted on social media websites (Li et al., 2019). Sentiment analysis can be classified into five classes: Machine Learning, Lexicon-Based, Hybrid, Graph-Based and Others (Giachanou & Crestani, 2016; Silva et al., 2016).

Machine-learning approaches are most used in textual QTA and besides traditional machine learning methods, ensemble classifiers are adopted to obtain more precise results when it comes to doing QTA for social media content (Li et al., 2019). However, the limitations of Machine learning methods, however, have limitations which are shown in two aspects (*ibid*). The first is that the performance of machine learning methods depends on the number of annotated samples and the second is that they are domain dependent (*ibid*). Hence, Lexicon-based methods are utilized because they do not require annotated data (*ibid*). They are however context independent and depend on static lists of words (*ibid*). Hybrid approaches compensate for the shortcomings of machine learning approaches and lexicon-based approaches (*ibid*). Furthermore, graph-based methods are proposed to utilize the social graph and its attributes, and they do not need large amounts of annotated data (*ibid*). However, "they are domain dependent because the sentiment lexicons and the connection graphs are domain specific" (*ibid*).

The user review page of an app is an important source of information about how people perceive it. For user reviews in this paper, data was collected primarily from Google Play Store rather than Apple App Store due to the former being a much larger repository of user downloads and usage statistics for the HOIA app. In this paper, a QTA R-code is used to analyse the text data from the user reviews on Google Play Store about the HOIA app and data from the media articles about the HOIA app. The code is then used to analyse the sentiment of the annotated data and visualized using knowledge graphs. The knowledge graphs are then used in the discussion part to assess whether the HOIA app has been a success.

5.2.2 Approach

In the paper, first, the sentiments of the input textual data are analysed. Second, key words that summarize the text are identified and third, an analysis of the keywords using knowledge graphs using different clustering methodologies. The input textual data is collected from the user reviews of the HOIA app on Google Play Store and the media articles on the HOIA app.

5.2.2.1 Input Data

The dataset folder contains two different segments of data. The first segment analysed was the user reviews of the HOIA app on Google Play Store and the second media articles on the HOIA app. For the media articles on HOIA, a search was conducted for articles about the HOIA app in the main media channels in Estonia, Delfi, and Postimees. Articles that were about informing people about the HOIA app were excluded. The study only focused on articles that were about the opinions of experts and journalists regarding the uptake and use of the HOIA app. The data segments were sorted manually and turned into a .txt document which was run into RStudio.

5.2.2.2 Basic Sentiment Analysis

To perform basic sentiment analysis, first the text is split into individual word referred to as tokens. The tokens which exist in the Bing and NRC lexicons are checked and the number of times the tokens occur in the text is counted. The identified tokens are then categorized as Positive/Negative (Bing) and Fear, Anger, Trust, Sadness, Disgust, Joy, Surprise, Anticipation (NRC). *Figure 5* and *Figure 11* show the words counted from Bing and NRC Lexicon in the input text for the textual data from the User reviews on the

Google Play Store and Media Articles, respectively. *Figure 6* and *Figure 12* show the Bing Lexicons in text contributing to different sentiments categorized as positive and negative. *Figure 7* and *Figure 13* show the NRC Lexicons in text contributing to different sentiments categorized under Fear, Anger, Trust, Sadness, Disgust, Joy, Surprise, Anticipation for the user reviews on the Google Play Store and Media Articles, respectively.

5.2.2.3 Keyword Extraction

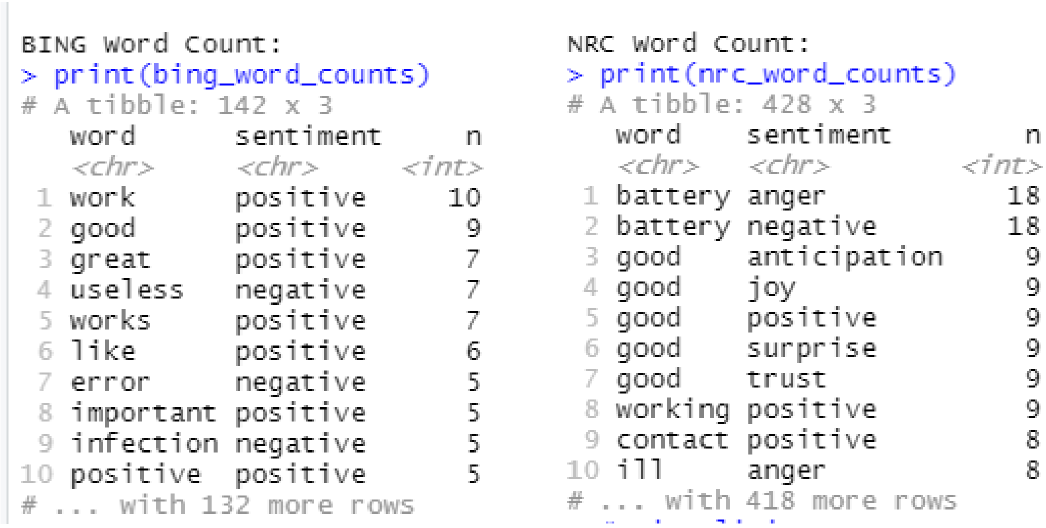
For this step we use two different approaches TextRank and KCore Retention. But before this step, the text is pre-processed using three steps:

- Tokenization: where the data is split the text into words, symbols, phrases or other expressive elements called tokens.
- PoS Tagging and Selection: where the tokens are marked with PoS (part of speech tags), based on both its definition as well as the context, that is to say the relationship with related and adjacent words in a sentence or phrase.
- Stopwords: Removal that filters out most common words.

PageRank function is then used to identify the keywords from the list of pre-processed tokens. The top one-third list of keywords based on the token's occurrence count is printed as the keywords identified using TextRank. The subgraph with the finally selected TextRank is represented in *Figure 9* and *Figure 15* for the User reviews on the Google Play Store and Media Articles, respectively.

The coreness of the nodes in the token's graph is then calculated. The nodes with the maximum coreness values are presented as the selected keywords as shown in *Figure 8* and *Figure 14* for the User reviews on the Google Play Store and Media Articles, respectively. The induced subgraph with the selected keywords is presented in *Figure 10* and *Figure 16* for the User reviews on the Google Play Store and Media Articles, respectively as a knowledge graph of the KCore keywords.

5.2.2.4 Sentiment Analysis (User reviews from Google Play Store)



```
BING Word Count:
> print(bing_word_counts)
# A tibble: 142 x 3
   word      sentiment     n
   <chr>     <chr>     <int>
 1 work      positive     10
 2 good      positive      9
 3 great     positive      7
 4 useless   negative      7
 5 works     positive      7
 6 like      positive      6
 7 error     negative      5
 8 important positive      5
 9 infection negative      5
10 positive  positive      5
# ... with 132 more rows
```

```
NRC Word Count:
> print(nrc_word_counts)
# A tibble: 428 x 3
   word    sentiment        n
   <chr>   <chr>        <int>
 1 battery anger           18
 2 battery negative        18
 3 good    anticipation     9
 4 good    joy              9
 5 good    positive         9
 6 good    surprise         9
 7 good    trust            9
 8 working positive         9
 9 contact positive         8
10 ill     anger            8
# ... with 418 more rows
```

Figure 5. Words from Bing and NRC Lexicon in the input text

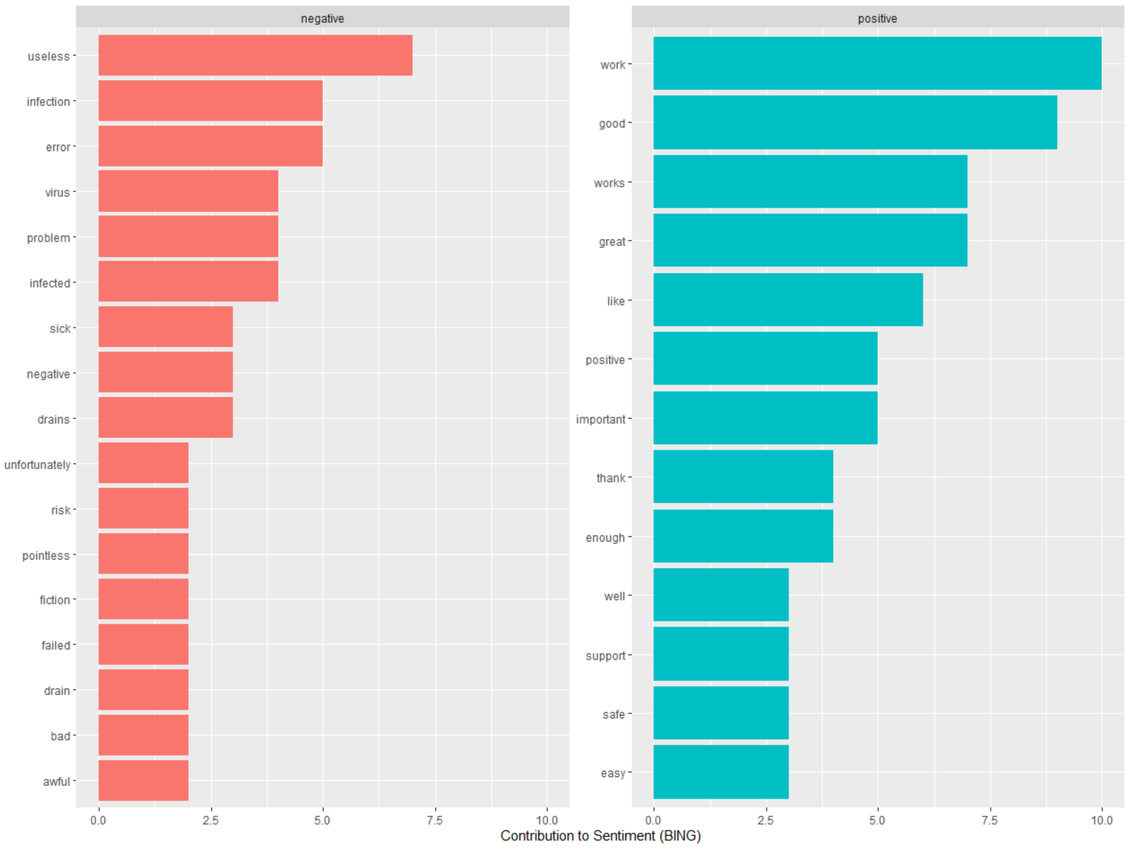


Figure 6. Bing lexicons in text contributing to different sentiments

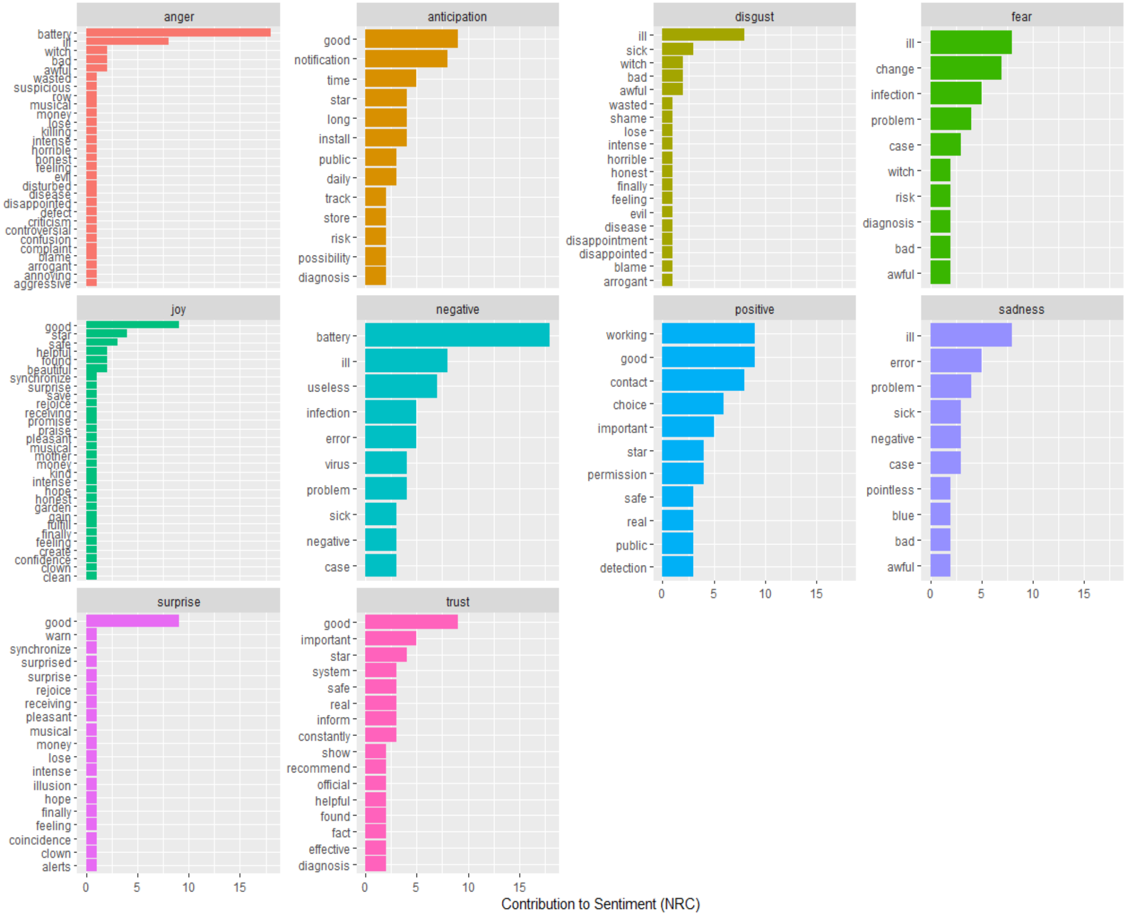


Figure 7. NRC Lexicons in text contributing to different sentiments

```
TextRank Keywords:
> print(names(final_keywords))
  [1] "i"             "app"          "language"      "bluetooth"     "people"        "phone"         "battery"
  [8] "use"           "location"     "only"          "dont"          "estonian"      "many"          "work"
 [15] "other"         "english"      "more"          "im"            "users"         "great"         "notification"
 [22] "good"          "etc"          "useless"       "day"           "need"          "much"          "huawei"
 [29] "important"     "idea"         "contact"       "change"        "cant"          "choice"        "services"
 [36] "works"         "version"      "data"          "help"          "google"        "application"   "edit"
 [43] "ill"           "os"           "infection"     "time"          "settings"      "positive"      "daily"
 [50] "covid"         "close"        "android"       "switch"        "same"          "technical"     "little"
 [57] "function"      "play"         "selection"     "switching"     "error"         "cannot"        "system"
 [64] "simple"        "icon"         "easy"          "hoia"          "devices"       "long"          "drains"
 [71] "home"          "addition"     "full"          "able"          "number"        "estonia"       "real"
 [78] "negative"      "fine"         "public"        "tracking"      "enough"        "least"         "everything"
 [85] "example"       "open"         "support"       "uninstall"     "problem"       "video"         "virus"
 [92] "infected"      "lot"          "advertising"   "id"            "star"          "smart"         "sick"
 [99] "developer"     "something"    "first"         "danger"        "information"   "fiction"       "permission"
[106] "new"           "someone"      "corona"        "thing"         "screen"        "fast"          "days"
[113] "design"        "possible"     "test"          "potential"     "beautiful"     "gps"           "questionnaire"
[120] "detection"     "next"         "pointless"     "doesnt"        "point"         "several"       "hours"
[127] "everyday"      "witch"        "second"        "update"        "old"           "blue"          "call"
[134] "own"           "access"       "privacy"       "stars"         "effective"     "moment"        "way"
[141] "youtube"       "iphone"       "button"        "notifications" "activated"     "fact"          "media"
[148] "okay"          "person"       "official"      "contacts"      "move"          "half"          "message"
[155] "program"       "country"      "user"          "newer"         "untested"      "optimization"  "owner"
[162] "controversial" "innovative"   "oneplus"       "schedule"      "station"       "suspicious"    "necessary"
[169] "pity"          "pit"          "everyone"      "matter"        "apps"          "device"        "minutes"
[176] "order"         "usual"        "awful"         "3x"            "notice"
>

KCore Keywords:
> print(names(verticesHavingMaxCoreness))
 [1] "i"            "dont"         "app"          "need"         "bluetooth"    "location"     "first"
 [8] "drains"       "battery"      "second"       "google"       "phone"        "use"          "hoia"
[15] "notification" "day"          "enough"       "much"         "more"         "people"       "positive"
[22] "users"        "only"         "good"         "covid"        "tracking"     "long"         "useless"
[29] "great"        "idea"         "able"         "contact"      "android"      "many"         "im"
[36] "time"         "work"         "ill"          "permission"   "application"  "etc"          "works"
[43] "other"        "human"        "close"        "lot"          "cant"         "switch"       "language"
[50] "os"           "settings"     "english"      "change"       "design"       "someone"      "little"
[57] "home"         "huawei"       "easy"         "error"        "data"         "doesnt"       "estonian"
[64] "choice"       "access"       "beautiful"
```

Figure 8. Keywords identified from input text (using TextRank and KCore algorithm)

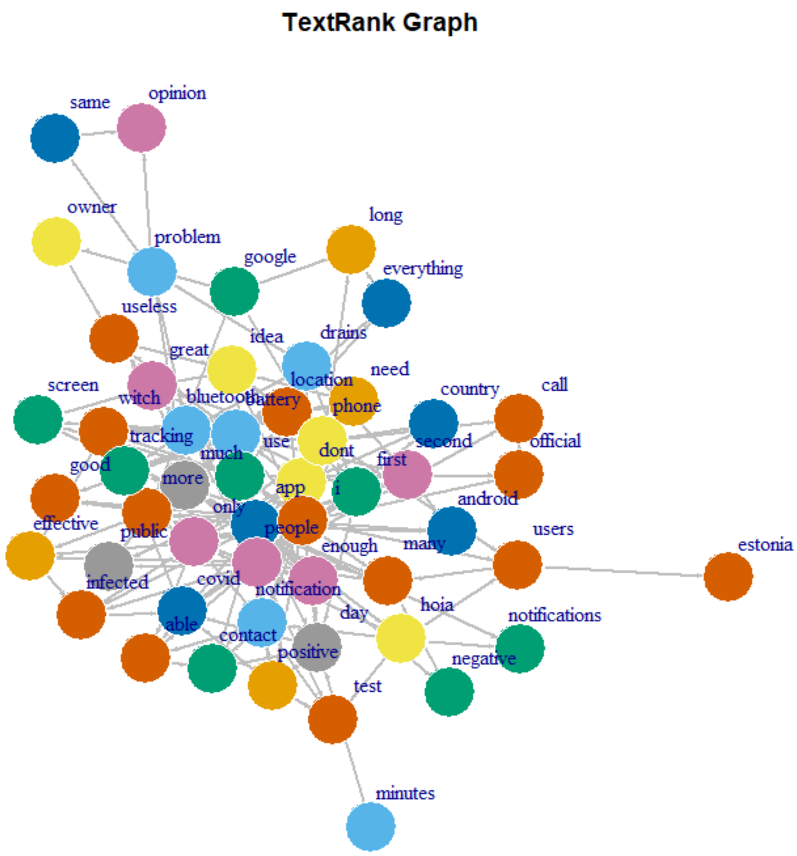


Figure 9. Knowledge graph of TextRank keywords

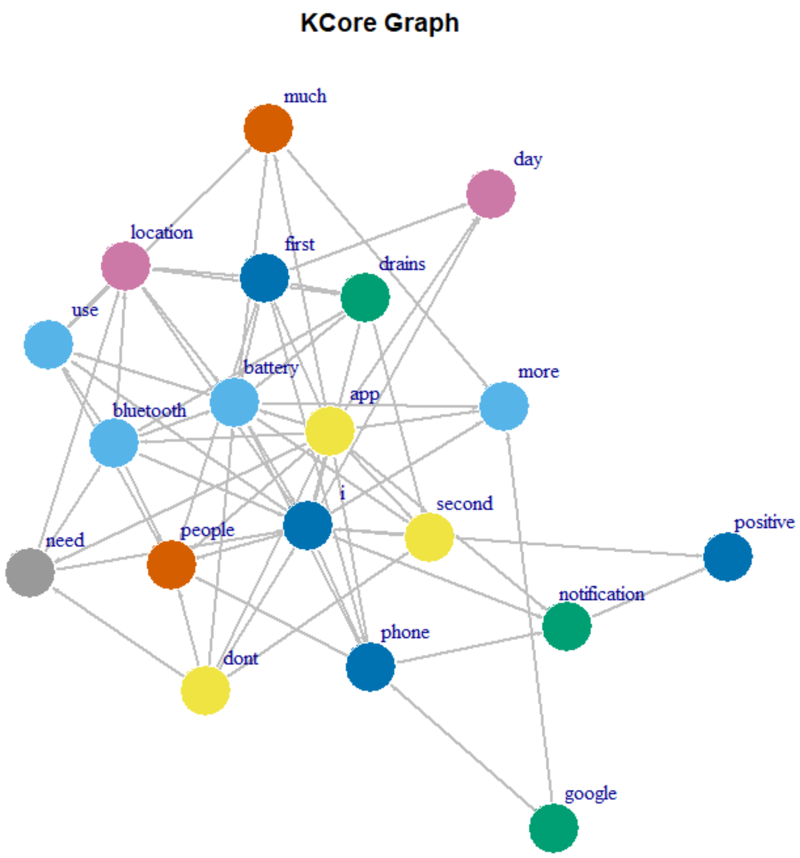


Figure 10. Knowledge graph of KCore keywords

5.2.2.5 Sentiment Analysis (Media articles)

```
BING Word Count:
> print(bing_word_counts)
# A tibble: 107 x 3
   word     sentiment     n
   <chr>    <chr>     <int>
 1 positive positive     10
 2 smart    positive      7
 3 virus    negative      7
 4 clear    positive      6
 5 failed   negative      6
 6 good     positive      6
 7 support  positive      6
 8 crisis   negative      4
 9 free     positive      4
10 infected negative      4
# ... with 97 more rows
```

```
NRC Word Count:
> print(nrc_word_counts)
# A tibble: 356 x 3
   word          sentiment        n
   <chr>         <chr>        <int>
 1 board         anticipation    14
 2 mobile        anticipation    11
 3 contact       positive         8
 4 public        anticipation     8
 5 public        positive         8
 6 solution      positive         7
 7 virus         negative         7
 8 communication trust            6
 9 good          anticipation     6
10 good          joy              6
# ... with 346 more rows
```

Figure 11. Words from Bing and NRC Lexicon in the input text

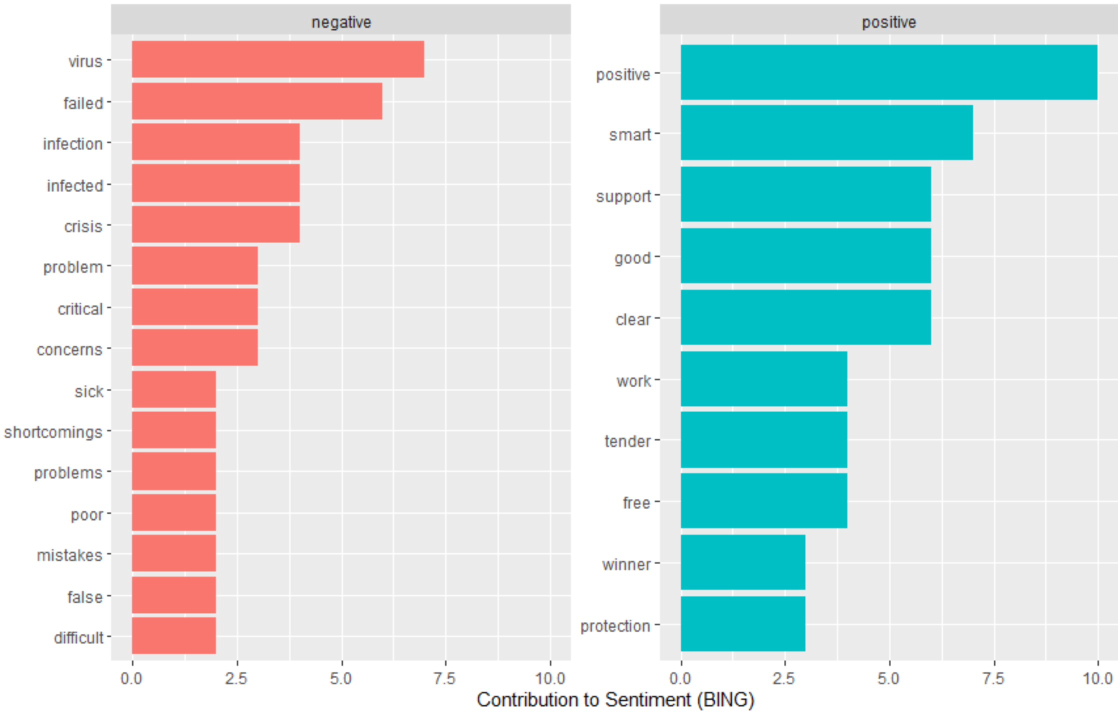


Figure 12. Bing Lexicons in text contributing to different sentiments

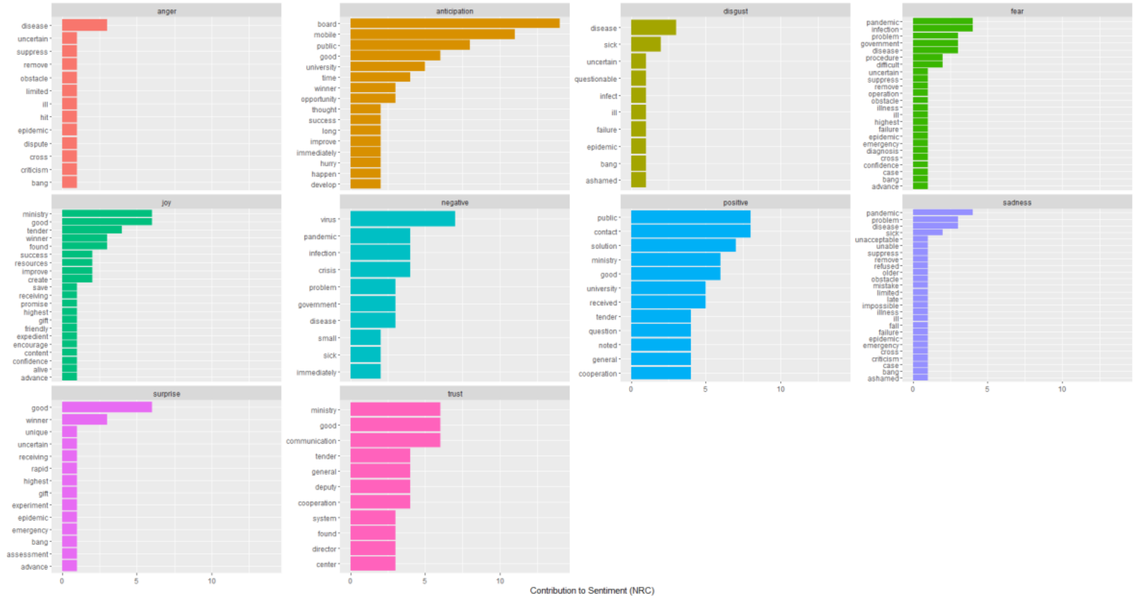


Figure 13. NRC Lexicons in text contributing to different sentiments

```
TextRank Keywords:
> print(names(final_Keywords))
  [1] "hoia"             "application"       "app"                  "health"
  [5] "procurement"      "people"            "tehik"                "development"
  [9] "killar"           "board"             "state"                "estonia"
 [13] "positive"         "close"             "more"                 "public"
 [17] "register"         "charge"            "contact"              "kã¤ã¤ramees"
 [21] "companies"        "department"        "users"                "mobile"
 [25] "total"            "communication"     "science"              "very"
 [29] "security"         "potential"         "solution"             "virus"
 [33] "university"       "corona"            "clear"                "cooperation"
 [37] "vilo"             "astok"             "mobileid"             "technical"
 [41] "support"          "such"              "systems"              "person"
 [45] "situation"        "owner"             "marko"                "vaccination"
 [49] "solutions"        "smartid"           "other"                "ministry"
 [53] "coronavirus"      "good"              "example"              "times"
 [57] "further"          "cases"             "current"              "software"
 [61] "possible"         "welfare"           "order"                "percent"
 [65] "system"           "today"             "lanno"                "tracking"
 [69] "work"             "merits"            "process"              "october"
 [73] "last"             "information"       "addition"             "contacts"
 [77] "social"           "pandemic"          "deputy"               "developments"
 [81] "general"          "secretary"         "identification"       "technology"
 [85] "additional"       "resources"         "help"                 "priisalu"
 [89] "viik"             "estonian"          "data"                 "free"
 [93] "first"            "downloaded"        "several"              "professor"
 [97] "applications"     "privacy"           "beginning"            "family"
[101] "coronary"         "infection"         "taltech"              "number"
[105] "sample"           "problem"           "specific"             "tender"
[109] "reasons"          "whole"             "andres"               "various"
[113] "minister"         "time"              "concerns"             "codes"
[117] "center"           "affairs"           "carrier"              "director"
[121] "months"           "technological"     "capability"           "critical"
[125] "smartphone"       "voluntary"         "end"                  "crisis"
[129] "tools"            "vaccine"           "less"                 "similar"
[133] "doctor"           "entire"            "jaak"                 "tenders"
[137] "decision"         "announcements"     "most"                 "experts"
[141] "thousands"        "academy"           "hannes"               "institute"
[145] "infected"         "responsible"       "sector"               "applicationâ€™s"
[149] "private"          "thing"             "month"                "question"
[153] "vicinity"         "advance"           "someone"              "disputes"
[157] "reaction"         "mechanism"         "situations"           "better"
[161] "need"             "aware"             "terms"                "testing"
[165] "small"            "procedure"         "maintenance"          "manager"
[169] "fraction"         "patients"          "answer"               "field"
[173] "disease"          "protection"        "sotsiaalministeeriumi" "kersti"
[177] "fact"             "repair"            "urmas"                "part"
[181] "viilma"           "same"              "population"           "governmentâ€™s"
[185] "winner"           "isamaa"            "mps"                  "page"

KCore Keywords:
> print(names(verticesHavingMaxCoreness))
 [1] "hoia"          "app"           "development"   "communication" "department"    "application"   "charge"
 [8] "very"          "health"        "tehik"         "procurement"   "state"         "public"        "companies"
[15] "owner"         "further"
```

Figure 14. Keywords identified from input text (using TextRank and KCore algorithm)

Figure 15. Knowledge graph of TextRank keywords

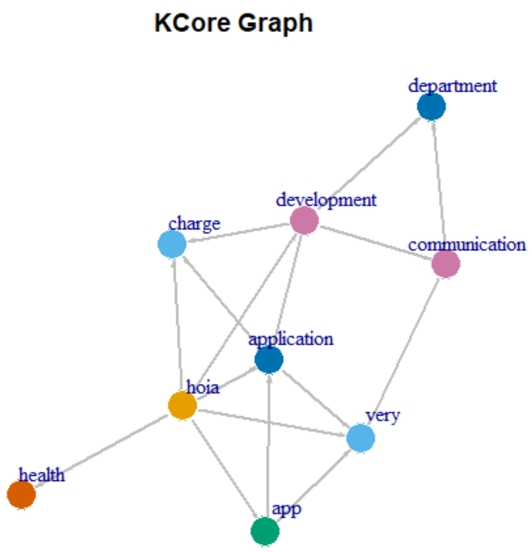


Figure 16. Knowledge graph of KCore keywords

5.2.3 Results

5.2.3.1 User reviews from Google Play Store

The analysis of the from the BING word count for user reviews from Google Play Store shows that the sentiment of the HOIA app is positive with words like "work, good, works, great, like, positive and important" scoring 5 out of 10 in the contribution to the positive sentiment as shown in *Figure* 5. On the other hand, the words "useless, error and infection score highly" in the negative sentiment towards HOIA shown in *Figure 5*. The analysis from the NRC word count shows "battery" as the word associated with negative and angry sentiments as shown in the NRC Lexicons in text contributing to different sentiments in *Figure 7*. From the knowledge graph of the KCore words (*Figure10*), it shows that "the app's use of Bluetooth location drains the battery" and this can be associated to the negative and angry sentiments as show in *(ibid).*

5.2.3.2 Media articles

The analysis of the BING word count for from the media articles shows that the sentiment of the HOIA app positive with words like "positive, smart, support, good, clear" scoring 5 out of 10 in the contribution to the positive sentiment as shown in *Figure 12*. On the other hand, the words "virus, and failed" also score above 5 out of 10 in their contribution towards the negative sentiment as shown in *(ibid)*. The analysis from the NRC word count shows anticipation, trust and positive and negative sentiments towards the HOIA app (*Figure 13*). to the words "board, mobile, public and good" as shown in *Figure 13.* The negative sentiment is mainly associated with the virus, the positive sentiment with "public, and contact" *(ibid)*. Finally, the trust sentiment is mainly associated with ministry, good and communication. From the knowledge graph of the KCore words (*Figure 16*), it shows that the sentiment in the media is positive and anticipatory towards the development and communication of the HOIA app.

In summary, from the sentiment analysis, this paper shows that sentiment towards the HOIA app is mainly positive from the analysis of the user reviews on Google Play Store and the media articles about HOIA. From the media articles, the sentiment is positive and anticipatory with trust in regard to the development and communication of the app. However, from the user reviews on Google Play Store, show that "the app's use of Bluetooth location drains the battery" and this can be associated to the negative and angry sentiments.

5.3 Thematic Analysis

5.3.1 Background

A thematic analysis is a method for analysing qualitative data that entails searching across a data set to identify, analyse, and report repeated patterns (Braun and Clarke, 2021). A thematic analysis is used in this paper because it strives to identify patterns of themes in the interview data collected from the stakeholders about the HOIA app. The method is used for describing data and it also involves interpretation in the processes of selecting codes and constructing themes to make meaning out of the data (Kirger & Vapio, 2020). According to Gavin (2008), "it is a flexible method which you can use both for explorative studies, where you don't have a clear idea of what patterns you are searching for, as well as for more deductive studies, where you know exactly what you are interested in". Thematic analysis describes an iterative process as to how to go from messy data to a map of the most important themes in the data (Kiger &Varpio, 2020). The process contains six steps: "familiarize yourself with your data, assign preliminary codes to your data in order to describe the content, search for patterns or themes in your codes across the different interviews, review themes, define and name themes and finally produce a report "(Gavin, 2008).

5.3.2 Approach

In this paper, a thematic analysis was done on the data from the interviews that were conducted with the stakeholders of the HOIA app. A list of interviewees may be found in Appendix 4 and the interview questions in Appendix 5. The transcripts of the interviews were obtained from YouTube and inserted into a word document. The data was then sorted out manually to remove errors in spelling and typos, and a process of familiarization with the data was done. Preliminary codes (given below) were then assigned to the data to describe its content. A search for themes and patterns was carried out to see what the main theme of the interviews was. This then led to a review of themes searched which were then defined and named and inserted into a report. The report is then used in the discussion part to assess whether the HOIA app has been a success.

<table>
<tr><th>Themes</th><th>Codes</th></tr>
<tr><td rowspan="2">Privacy: Communication/Messaging about privacy of the app</td><td>Privacy/Marketing the app/Media negative about the app</td></tr>
<tr><td>Marketing and communication/Privacy/Fear of abuse of data</td></tr>
<tr><td>Technical issues: Challenges with use of the app</td><td>Codes Technical issues with the app/Usability/ Regrets about app features (language)</td></tr>
</table>

	Language issues/Technical Issues with the app/Improvements of the app
The app is great and helpful	Positive about app/App helpful
	Ease of use of the app
Research done before the app was released	Research done on app before launch
	Research done on the app
Challenges making HOIA/Ownership of the app	App made voluntarily/Changing circumstances/Own app
	Circumstances changing/Challenges making app
Ease of use of the app	Design
	Codes
HOIA vs other apps	Codes HOIA vs Other apps/Architecture

Table 1. Table showing themes and codes for thematic analysis

5.3.3 Interview Themes

The analysis produced seven themes.

5.3.3.1 Privacy: Communication/messaging about privacy of the app

Both participants in the interview reported that privacy was of utmost importance when it came to designing the app. This translated into the communication about the app from Stakeholders focusing on privacy, as they attempted to convince the public that their data was safe when they used the app.

Laura: “What I know was done before the huge part was about its safety”

Kerstin-Gertrud: “The main message was you know it's safe it doesn't leak, we don't know who you are the government doesn't know, we don't track you, we're not here to take notes into our secret vaults of something you know so that was the meaning.”

Kerstin-Gertrud: "When I came in the sort of privacy was a really big thing and I thought it that really positive and I think that my privacy got some more uses because when something is sort of labelled under the government's name no one really wants to sort of put the personal data in there so you know making it as private as possible and sort of making it clear that it's hidden under eyes and no one will know that it has been you"

Kerstin-Gertrud: "It comes from the government you know it comes up people are more sensitive towards topics that have to do with the government and there are also different rules what we can collect, and we can't collect so it's really limited so we can't take your data just because you know we want to that's not correct."

5.3.3.2 Technical issues: Challenges with use of the app

The participants in the interviews also reported that the app had some technical issues which affected the perception of the people when it came to downloading and using it. There was also a challenge with the language of the app which affected the app users.
Kerstin-Gertrud: "It may not work of course there is always you know technical difficulties. We have boosted the app's sensitivity so now it's able to track more those contacts."

Laura: "there are three things that I really regret first is the language selection that's this in July that was something that it came out from user testing but in the team there was a discussion that a correct way to do a language selection is it's based on the language of the phone yeah and that the understandable solution the information there was that everybody can change their language so it sounded like a good idea but we realized in the first day that there are some cases where a person wants to have a phone in English but that app still in Estonian. This was one big mistake."

Laura: "The other thing is we have developed it very like securely that means that font puts conversion in a testing phase like ill so we can't have like a live testing so somebody is like ill and we can check it in the complete flow in a live environment so for my surprise because we are using like Apple and iOS technical solution we had some this page that's really famous for newspapers that failed so badly it's the one that the application isn't working the functionality from iOS and android isn't working but you have marked your selfie this is one I’m really embarrassed and it's based on that I didn’t know that the information will happen on the same time. I understood that if you have done marked yourself feel then in 24 hours you will get that that application stops working not that same time so that was my like this is the most embarrassing thing I have done in years.”

Laura: “One more was from google and yes they did a change that we had a really big drop box and we didn’t understand what it was and if we got in the lives then they added a screen this kind of screen this right one hand, very nice they used wording like share and we are talking about really anonymous information that the share related like share to friends on Facebook, so people were not really afraid if they saw this so to find this found this then we had we added like an intro page before that to explain what it is and the worst case is that most of the phones will have this pop-up still in English if even if your main language is Estonian and Russian so this is like these were the like the scary moments of the project.”

5.3.3.3 The app is great and helpful

The participants in the interviews also noted that the app was great and helpful in helping with the process of Contact Tracing in Estonia.

Kerstin-Gertrud: “In January we reached a huge number of positive cover positives in Estonia and that's the huge number of close contacts and HOIA actually was really helpful and sort of finding out those contacts.”

Kerstin-Gertrud: “we were at some point in time in a situation when HOIA was there first, and we came in second. So, people were saying, oh yeah I already got notified about the exposure from HOIA.”

Laura: “If you saw my first research, there are people who do not have smart phones. Now we can add that google ended the relationship with Huawei there are those people who cannot use it so we have a big group of people who cannot actually use the application, so I think based on that the download numbers are good.”

Laura: “350,000 to 400,000 people, so even based on that we are not thinking about 1.2 million of every Estonian because we have like age limits, phone limits, then the number of downloads is very good because the number of people who have decided to like participate is good based on that. I feel like we have done something that has it had like an effect, that's something we have talked about with different like team members.”

5.3.3.4 Research was done before HOIA was released

The participants in the interview also reported that research was done on the app before it was launched out to the Estonian public, and one of the things that came out like privacy, the need for the app in society are in the excerpts below:

Kerstin-Gertrud: “Research was done…. there were questionnaires, by the Social ministry people to see whether why people would download HOIA or why they had downloaded ‘CTAs’ everywhere.”

Kerstin-Gertrud: “Reasons why people would then download where at that time the main sort of reason for people was either they thought it was a risk in sense of you know its personal data and they didn't quite understand the sort of concept of it's really anonymized so people would think they were afraid, you know it's not safe for something or their data will be leaked, or they will be followed.”

Kerstin-Gertrud: “people's thought they don't need it because they said you know it was the research was done during the first drop, so they're locked down after period, so people were saying like we don't move around I don't need it or so they didn't quite understand that.”

Laura: “We started with the research. We did the interviews with people about what would they think HOIA would do, what, how, would it work, what are the problems to people, would people use it, what are they afraid of, and we had I think more than 12 interviews just to get that knowledge what people think before there was HOIA before there was a name.”

Laura: “We got the idea that the very important thing is the safety and anonymity, people don't want that their medical data will get somewhere.”

Laura: “We did a bigger questionnaire. We got the idea of we had like an open question for people how to answer what these problems would be and, what they are scared of, and we had the open questions analysed.”

Laura: “The next one is privacy and the ones that don't have any questions they are okay to use it. Next one is people who are worried about their personal data, some people think that this is not an application that we need…. then a lot of people had a question how it works, and something we are like working in not only on the beginning of the project but also in the future, so this is being also here people are worried about the state.”

Laura: “One answers in that group I would maybe bring out one that's an anxiety to get the notification there were like a group of people who know that they don't want to get the notification. Understanding that safety is very important like I showed here that people are afraid that their data will goes somewhere that they don't want, and they don't want to get that their name is associated with illness, and they lived in a small city, and everybody knew that they were the ones that had the sickness, so it was really important and in safety.”

Laura: “We did a lot of like usability testing on the application on the prototype.”
Laura: “I understood from the research that there are very big different groups that will never use it like people who don't think thatCOVID-19 exists. I got those people also in the interviews who told me that in a different view understand that they will never use it.”

5.3.3.5 Challenges making HOIA/ Ownership of the app

The participants in the interview reported that there were challenges in the process of making the HOIA app, and they also clarified on who owns the app, the Estonian Health Board.

Kerstin-Gertrud: “I’m an epidemiologist in the health board so I’m the product owner of the legal sense.”

Kerstin-Gertrud: “Again midway product owner, so there was like another product owner before me, so he had the vision I came in I even I didn't even meet him, so it was sort of left and we left somewhere so I picked it up okay.”

Kerstin-Gertrud: “Initially, no sort of contracts you know everybody was doing it voluntarily and the time when I got in there were really few people left who had time for it.”

Kerstin-Gertrud: “Yeah and of course, you know take that into consideration that, and you know the language options came out, and you know things might you know again in the beginning when they weren't really finding contracts and you know people were doing out of their free time eventually you know some things take time and they take longer time when they are done out of someone else's free time.”

Laura: “I was the one leading the app design part, so it was a teamwork I had a lot of people doing user testing.”

Laura: “I’m not even arguing that we had so little time like that was the first period of that we worked so that all of the weekend all of the night so and it would be in an ultrasound interview that we would have done.”

Laura: “I will talk from my point of view it's a project that my company doesn't get any money so my maybe time is more limited how much I can uh give to the project”

Laura: “I think always there is an option to do things better so if you have more maybe time to put on a project maybe there would be a better solution.”

Laura: “But in the first months nobody thought about the projects that gave them money everybody gave like more than 100 to do it to get it live, and it was based on people’s vacations free time weekends everything and so I would really like point out that it’s not only that if we do some kind of project we have like a work we have like ours we have like nine to five we are working.”

5.3.3.6 Ease of use of the app

Another theme was the ease of use of the HOIA app reported by one of the participants.

Kerstin-Gertrud: “It's really so easy, it's nothing complicated you don't have to go through excessive signing in process. It's there you don't even have to look at it every day let you know if you have to look at.”

5.3.3.7 HOIA in comparison to other CTAs

The final theme that was reported was a comparison of HOIA in regard to other CTAs in terms of architecture.

Laura: "We are similar or the same even because we are using the google and iOS solution. I think one good solution that we did was that we did not go with the centralized system that some of the countries went. So, we do not have its central database, we have like had a priority of security so high that the conversations and like a project first step was very like how you can hack it, how can you lose your data, so all of the details were described, and I think that's one of the things that we our security guys did a very good job that we decentralized we thought about every small thing."

Laura: "Some of the countries have gone in a way that they have added a lot of additional functionality to this application this is something we have talked about I don't know."

5.3.4 Results

The analysis of the interviews conducted with a representative from the Estonian Health Board and Iglu, one of the companies in the consortium that worked on developing the app produce seven themes. This section provides a report on the data collected from the themes in regard to the perception of the HOIA app.

First, research was carried out about the HOIA app before it was launched out to find out through interviews, questionnaires and surveys with people to understand whether they would use the app, download it, and usability testing on the prototype. Insights from this research guided the stakeholders in the process of designing the app, and when it came to rolling it out to the public. From the research carried out by the stakeholders, privacy and anonymity were of high priority. This was reflected in terms of the architecture the stakeholders chose to use for the HOIA App, the decentralised architecture and the use of Apple/Google Exposure Notification APIs. In regard to privacy, the focus of the stakeholders was to design an app that was secure and did not collect any personal identification data from the users of the app. This can be linked to the theme about research being done on the app before being released. This focus on privacy was also a key message in marketing the app.

Second, technical issues technical issues of the app, in regard to users was a main point of concern during the interviews with the stakeholders. There were issues about the language selection of the app. The app initially was designed to use the language setting of the phone, and users could not access it in Estonian or Russian if their phone setting was in English. This led to a frustration amongst users, and from the stakeholder's point of view, it affected the perception of the app.

Third, from the theme of challenges in making the app, there were some issues with it being designed voluntarily by a consortium of private companies together with the government. This was indicated in the responses from the interviews, as some of them had to work in their free time on the app as they were not being paid. This also limited the amount time which was put into project by the stakeholders as they had to balance commitments elsewhere with working on the app.

Furthermore, another theme that came out was the ease of use of the app, and the app being great and helpful from the perspective of the stakeholders. In their perspective, the app was easy to use, and it helped in the process of tracing close contacts during the pandemic. From the research carried out about the app, the stakeholders were targeting about 350,000 to 400,000 potential users of the app, and from the 275,614 downloads, as of 9th May 2021, the app has had an effect in society.

In summary, from the thematic analysis of the interviews with the stakeholders, perception of the HOIA app is that its great and helpful with some challenges. The main focus during the design process of the app, and after the launch of the app, has been on privacy. This is reflected in the decentralized system of architecture they chose for the app. However, technical challenges with the app in regard to the language choice in the app, and usability have affected the perception of the app in a negative way from the public's side. These technical issues could have been a result of the challenges faced by the stakeholders working on the app voluntarily without contracts.

5.4 Resident Survey

5.4.1 Background

Quantitative research methods are concerned with collecting and analysing data that is structured and can be represented numerically (Bloomfield & Fisher, 2019). One of the central goals is to build accurate and reliable measurements that allow for statistical analysis. Because quantitative research focuses on data that can be measured, it is very effective at answering the "what" or "how" of a given situation (*ibid*). Questions are direct, quantifiable, and often contain phrases such as what percentage? what proportion? to what extent? how many? how much? In this paper, the quantitative research seeks to demonstrate to what extent the HOIA app has been used and accessed and compliment the results from the qualitative research to show whether HOIA has been a success. This is because data does not provide evidence for why populations think, feel, or act in certain ways (Bamberger, 2000).

5.4.2 Approach

Since it is challenging to obtain data on the statistics about the use of HOIA because of its privacy by design default, this study carried out a survey of the Estonian residents to seek answers on whether they have heard, downloaded or used the HOIA App, and how important privacy is to them.

The data used for this research were collected via a google survey form. The form was distributed through social media channels (Facebook, Instagram, and WhatsApp) targeting residents of Estonia in different parts of the country. In addition, solicitations were also distributed via the authors' Facebook and Instagram accounts. In the end, 253 people participated in the online survey. The survey form and questions are found in Appendix 2.

5.4.2.1 Sample size determination

The sample size was determined by the accuracy required for the national estimates, resources and operational constraints. The following formula was used to calculate the sample size, n.

$$n_0 = \frac{z_\alpha^2 * p(1-p)}{\delta^2}$$

where,

z_α^2 =1.96, the abscissa of the normal distribution cure at $\alpha = 95\%$

P=0.5, the proportion in the sample that will yield the optimal sample size of app users

δ^2 - the margin of error to be tolerated in the sample, taken to be 6.16%

Using the above formula, the effective sample size for the survey with a Relative Margin of Error (RME) of 6.16% yielded a national sample of 267 respondents. The sample size was adjusted to the number of people who have the smartphones with the capability to use the Android 6+ and iOS 13.5+, which is 400,000 from the figures from the interviews from the research conducted by the stakeholders of the HOIA app.

The sample was adjusted using the following formula below:

$$n = \frac{n_0}{(1 + (n_0 - 1)/N)}$$

Some of the forms were sent back with no data in them, and the total number of participants in the survey was n = 253 which still gives a margin of error of 6.16%.

The next section provides a result of the survey form from the respondents.

5.4.3 Results

The survey was conducted with a view to acquire information to show whether there is association between people who heard about the HOIA app, and those that downloaded it, and how important privacy was to them. The purpose of the first question of importance in the survey was to obtain information about what percentage of people knew about the HOIA app. The results show that 94.9% of the people had heard about the app, with only 5.1% not having heard of it. This shows that most Estonian residents know about the HOIA app.

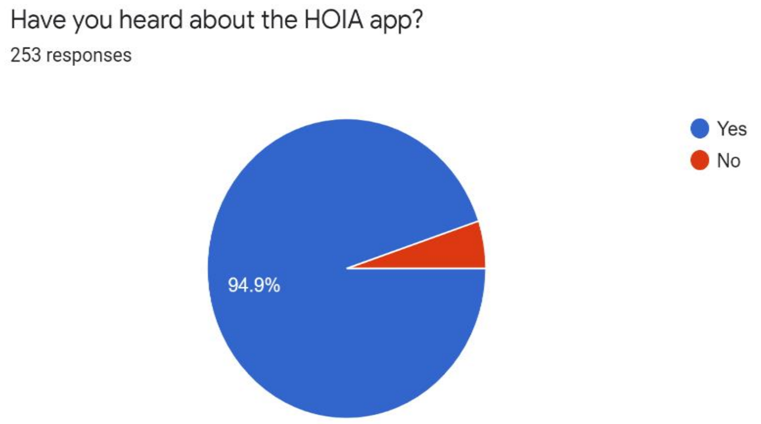


Figure 17. Have you heard about the HOIA app?

The purpose of the second question of importance from the survey was to find out how many people had downloaded the app to their phone to obtain information to carry out analysis to show if there is association between knowing about the app and downloading it. 56.9% of the respondents downloaded the app, and 43.1% did not.

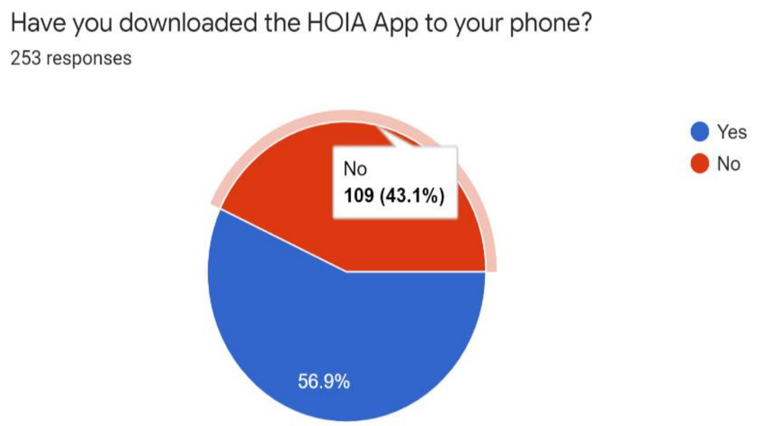


Figure 18. Have you downloaded the app to your phone?

Statistical analysis was done on the data using a statistical tool, STATA to show if there is significant association between those who heard about the HOIA app, and those who downloaded it using the Pearson's Chi-Square test. The results obtained are showed in *Figure 19* below:

Have you heard about the app?	Have you downloaded the app to your phone??					
	No		Yes		Total	
	n	%	n	%	n	%
No	12	92.31	1	7.69	14	100
Yes	97	40.59	142	59.41	239	100
Total	109	43.25	143	56.75	253	100
	Pearson chi2(1) = 13.4378 Pr = 0.000					

Figure 19. Pearson's Chi-Square test for association between knowledge of HOIA and downloading the app

From the results in *Figure 19*, Pearson's chi-squared value is 13.4378, and the probability is 0.00. This shows that there is a statistically significant association between those who heard about the HOIA app, and those who downloaded it. 59.41% of the respondents who heard about the app downloaded it, which shows the 6 in 10 people who heard about the HOIA app download it. This is important as it shows that the residents' respond to the communication about the HOIA app and do want to use it to help in the process of contact tracing.

Finally, the purpose of the final question from the survey was find out how many people thought the information they provided to HOIA was secure in terms of privacy. This was done to measure the perception of the Estonian residents on the HOIA app. 232 of the 253 responded to this question as shown in *Figure 20.* 54.7% think that the information they provide to the HOIA app is secure. 17.2% think that information is not secure, and 28% were unsure whether the information is secure or not. This shows that a significant part of the Estonian residents thinks that the information they fill in into the app in to help with the contact tracing process is secure, hence the HOIA app is perceived as being secure.

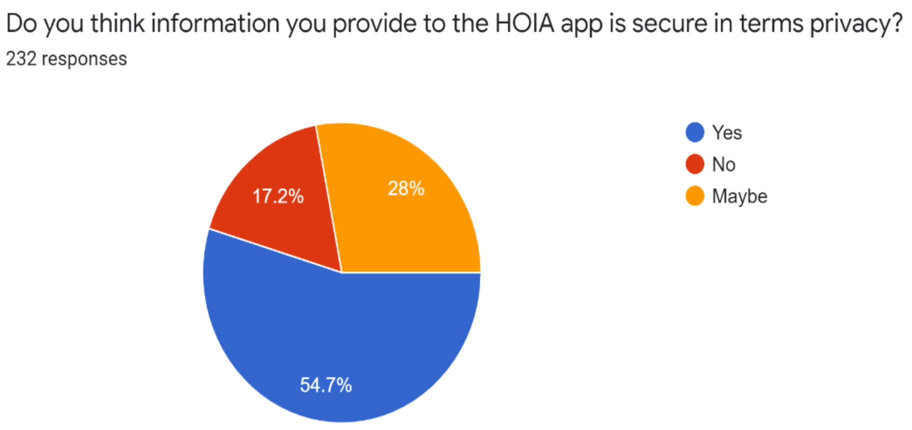


Figure 20. Do you think information you provide to the HOIA app is secure in terms of privacy?

In summary, the survey that was conducted shows that majority of the Estonian residents have heard about the HOIA app, and a significant percentage of those who have heard about it downloaded it on to their phones. The survey also shows that majority of the respondents think that the information they provide into the app is secure in terms of privacy. This shows that the perception of the HOIA app as a tool amongst the Estonian residents is positive. However, an explanation is needed as to why the perception is positive, and he download numbers were high, but the number of people actively using the app was low. This is done in the section below through a triangulation of data collected from the three research methods used in this paper.

6 Discussion

In this section a discussion of the results from the thematic analysis from interviews from the stakeholders, textual and sentiment analysis on user reviews from Google Play Store, and media articles about the HOIA app, and a survey analysis of the residents' perception of the HOIA app are discussed and answers to the research questions R1, R2, R3 and R4 are provided.

The first research question set in this paper, RQ.1 is "What were the key aspects taken into consideration when developing the Estonian COVID-19 contact tracing app, HOIA?" and the sub question "a. Why the specific the design of the app, and why focus hugely on privacy as an issue?"

RQ.1 is answered by the results from the thematic analysis of interviews from the stakeholders of the HOIA app. According to the responses from the interviews, the research was done on the app before it was launched to investigate whether the Estonian residents needed the app, what their concerns were in case a contact tracing app were to be rolled out, and what mattered the most to them. The results from the research according to the interviews indicated that privacy was a high priority. This was reflected in the specific design of the app to use a decentralised system of architecture, and Apple/Google Exposure Notification APIs as they were privacy preserving.

The second research question set in this paper, RQ.2 is "How is the Estonian COVID-19 tracing app, HOIA different from other apps?" and the sub question "a. What were the novel features of the HOIA app (if any)?"

RQ.2 is also answered by the results from the thematic analysis of interviews with the stakeholders of the app is used and the literature review in the state-of-the-art. From the literature review, Estonian Contract Tracing app, HOIA is based on a decentralised system architecture which uses the Apple/Google Exposure Notification APIs, and the DP-3T systems which provides security and privacy protections by "ensuring data minimization, preventing abuse of data, preventing tracking of user" (Troncoso, 2020). This separates it from CTAs that use the centralised and hybrid systems of architecture which do not focus on privacy. However, from the thematic analysis, the HOIA app is similar to other CTAs that use the decentralised system and are built on BLE and use Apple/Google Exposure Notification APIs.

The third research question set in this paper, RQ.3 is “How does the HOIA app compare in regard to user acceptance (when compared to other apps)?” and the sub questions, “a. What is the stakeholders’ perception regarding the app? And why?” and “b. What is the residents' perception regarding the app? And why?”.

RQ.3 is answered using the literature review. The first sub question (a) is answered using the thematic analysis and the second sub question (b) is answered data collected from both the textual and sentiment analysis of the user reviews from Google Play Store and the media articles, and the survey analysis of the residents’ perception of the HOIA app. The literature review shows that in terms of user acceptance, CTAs have struggled around the world especially in countries where they are not mandatory to be used. This is because users are sceptical about privacy issues regarding the apps (Raskar et al., 2020). In regard to the Estonian contact tracing app, HOIA, data from the survey analysis shows that majority of the Estonian residents perceive the HOIA app in a positive way. This is indicated by the high number of downloads (275,614) as of 9th May 2021. However, the number of people who have marked themselves as sick with the virus (6,739) has been low (Terviseamet, 2021). This shows that the user acceptance of HOIA too has struggled.

When it comes to the perception of the stakeholders of the HOIA app, a thematic analysis of the interviews is used. The perception of HOIA app from the stakeholder’s point of view is that its great and helpful with some challenges. The main focus during the design process of the app, and after the launch of the app, was on privacy of the app. This was reflected in the decentralized system of architecture they chose for the app. However, the stakeholders also agree that technical challenges with the app regarding the language choice in the app, and usability have affected the perception of the app in a negative way from the residents’ side. The data from the thematic analysis also indicates that the technical issues could have been a result of the challenges faced by the stakeholders working on the app voluntarily without contracts and in their free time. This might have resulted in them missing out on the important technical aspects in regard to the choice of language in the app and usability issues which has had a negative impact on the user acceptance and perception of the app.

When it comes to the perception of the residents, data from the survey analysis is used. The data shows that the residents heard about the app, and this resulted into a significant number of people downloading the app. The data from the survey analysis also shows that the residents' perception of the app in regard to privacy is good as a majority of them think that the information they send to provide into the app is secure. However, from the data in the textual/sentiment analysis of the user reviews on Google Play Store, the residents' sentiments towards the app are negative and angry associated with technical issues like the "the app's use of Bluetooth location drains the battery". This shows why the app has been download, but few people actively use it.

Finally, fourth research question set in the paper, RQ.4 is "Has HOIA been a success?" and the sub question "a. What lessons could be learned from it?"

RQ.4 is answered done by a summary of the results from the literature review in the state-of-the-art, the thematic analysis, the textual/sentiment analysis and survey analysis. The HOIA app was designed to complement the process of manual contact tracing in Estonia. Data from the thematic analysis of interviews with the stakeholders shows that privacy was of high priority in the process of designing and rolling out the app, and therefore, the decentralised system of architecture was chosen because its privacy preserving. Despite some challenges during the designing process and the technical issues, the app has been great and helpful in the process of contact tracing from the perspective of the stakeholders. The textual/sentiment analysis for the user reviews on Google Play Store and the media articles about the HOIA app was mainly positive, as were the results from the survey analysis on the residents' perception of the HOIA app. The results from the survey analysis showed that a majority of Estonian residents had heard about the HOIA App, and there was significant association between having heard about the app and downloading it. This is indicated in the high number of people (275,614) who have downloaded the HOIA app. However, there was a negative and angry sentiments because of the "app's use of Bluetooth location which drains the battery" too from the textual/sentiment analysis of the user reviews on Google Play Store. Combining this with the technical and language choice issues with the app, it can be seen why the number of using the app to highlight themselves as sick has been low. In summary, despite the success in having the residents know about the app, there have been challenges with getting to people to use the app which has resulted in mixed results for the HOIA app.

The first lesson that could be learned is offering contracts to stakeholders when working on designing the project. This allows for those working on the project to fully invest themselves into dealing with all issues regarding all aspects of the design process, so that quick solutions are available to any challenges that arise. The other lesson that could be learned is investing in educating people how to use the app, instead of only asking them to download it.

7 Conclusion

In conclusion, this paper set out to assess the perception of the Estonian Contact Tracing App, HOIA through by using both quantitative and qualitative methods. The qualitative methods used a thematic analysis of interviews that were conducted with the stakeholders of the HOIA app a textual/sentiment analysis of the user reviews from Google Play Store and media articles was carried out to assess the perception of the HOIA app. The quantitative method focused on the use of a survey to obtain the residents' perception of the HOIA app. The results from the triangulation of these three methods show that perception of the HOIA app from the Estonian society is positive. This is shown by the number of people who have heard about the app and downloaded it. However, there were also negative perceptions of the app. These arise from the technical issues and language choices of the app which made its usability challenging. The app's use of Bluetooth location which drains the phone battery of a user also affected the apps usability in a negative way. These challenges show why the app has been downloaded by a large number of Estonian residents', but the number of people who actively use it for the contact tracing process has been low. In summary, the perception of the HOIA app has been mainly positive. However, usability challenges have hindered its uptake and use as tool to complement the contact tracing process in Estonia.

Lessons can be drawn from this for the stakeholders and policy makers. First, there should be contracts for projects that work on tools that are designed for public use. This makes it easier for those involved to be more work more efficiently on making the tools that meet the need of the residents. Second, more education on how to use the tools should be rolled out, and the message should go beyond asking people to download the tools. This way, people can easily know how things work, and how to have any usability issues fixed in a way that does not make them disregard using the tool again.

7.1 Limitations

The main challenge during the study was obtaining data on the HOIA app. Since the app is designed to be privacy preserving, data on the usage of the app is not collected. This made it difficult to choose methods on how to collect data and study the app. Another limitation was the finding scientific research specifically on the HOIA app, as it has not been studied extensively yet.

7.2 Future work

The purpose of this paper was to assess the perception of the Estonian COVID-19 Contact Tracing Applications from the perspective of the stakeholders who designed the app, and the residents who use the app. Further research could also be conducted on exploring how the voluntary collaboration between the public and private sectors affects the end process of the project the two entities work on.

References

Ahmed, N., Michelin, R. A., Xue, W., Ruj, S., Malaney, R., Kanhere, S. S., Seneviratne, A., Hu, W., Janicke, H., & Jha, S. K. (2020). A Survey of COVID-19 Contact Tracing Apps. *IEEE Access*, *8*, 134577–134601. https://doi.org/10.1109/access.2020.3010226.

Altmann, S., Milsom, L., Zillessen, H., Blasone, R., Gerdon, F., Bach, R., Kreuter, F., Nosenzo, D., Toussaert, S., & Abeler, J. (2020). Acceptability of App-Based Contact Tracing for COVID-19: Cross-Country Survey Evidence. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3590505

Avitabile, G., Botta, V., Iovino, V., & Visconti, I. (2020). Towards Defeating Mass Surveillance and SARS-CoV-2: The Pronto-C2 Fully Decentralized Automatic Contact Tracing System. *IACR Cryptol. ePrint Arch*, 493.

Bamberger, M. (Ed.). (2000). *Integrating quantitative and qualitative research in development projects: Lessons from the field*. The World Bank.

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The qualitative report*, *13*(4), 544-559.

Bay, J., Kek, J., Tan, A., & Hau, C.S. (2020). BlueTrace: A privacy-preserving protocol for community-driven contact tracing across borders.

Bloomfield, J., & Fisher, M. J. (2019). Quantitative research design. *Journal of the Australasian Rehabilitation Nurses Association*, *22*(2), 27-30.

Braun, V., & Clarke, V. (2021). Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research*, *21*(1), 37-47.

Castelluccia, C., Bielova, N., Boutet, A., Cunche, M., Lauradoux, C., Le Métayer, D., & Roca, V. (2020). ROBERT: ROBust and privacy-presERving proximity Tracing.

Chakraborty, Pranab, Subhamoy Maitra, Mridul Nandi, and Suprita Talnikar. (2020). "Decentralized Contact Tracing Protocols." In *Contact Tracing in Post-Covid World*, pp. 71-104. Springer, Singapore.

Chan, J., Foster, D., Gollakota, S., Horvitz, E., Jaeger, J., Kakade, S., Kohno, T., Langford, J., Larson, J., Singanamalla, S., Sunshine, J., & Tessaro, S. (2020). PACT: Privacy-Sensitive Protocols And Mechanisms for Mobile Contact Tracing. *ArXiv*, *abs/2004.03544*.

Cioroianu, I., & Dal, A. (2020). What is missing from the online debate around COVID-19 digital tools?

Criddle, C., & Kelion, L. (2020, May 7). Coronavirus contact-tracing: World split between two types of app. *BBC News*. https://www.bbc.com/news/technology-52355028.html

Dhar, T. (2020). Aarogya Setu-Carrying Your Privacy in Your Hands? *Available at SSRN 3614506*.

Ezzaouia, I., & Bulchand-Gidumal, J. (2021). A Model to Predict Users' Intentions to Adopt Contact-Tracing Apps for Prevention from COVID-19. In *Information and Communication Technologies in Tourism 2021* (pp. 543-548). Springer, Cham.

Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., Parker, M., Bonsall, D., & Fraser, C. (2020). Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. https://doi.org/10.1101/2020.03.08.20032946

Fišel, M. (2020). Neurotõlge: Adaptive, Multilingual and Reliable Machine Translation for Estonian (1.01.2018−31.12.2020). https://www.etis.ee/Portal/Projects/Display/b30268d1-d05c-4f19-a6c4-3b4b9fee5f1c?lang=ENG (accessed 07/05/2021).

Garrett, P. M., White, J. P., Lewandowsky, S., Kashima, Y., Perfors, A., Little, D. R., & Dennis, S. (2021). The acceptability and uptake of smartphone tracking for COVID-19 in Australia. *PloS one*, *16*(1), e0244827.

Gavin, H. (2008). Thematic analysis. *Understanding research methods and statistics in psychology*, 273-282.

Giachanou, A., & Crestani, F. (2016). Like It or Not. *ACM Computing Surveys*, *49*(2), 1–41. https://doi.org/10.1145/2938640

Goddard, C. (2011). *Semantic analysis: A practical introduction*. Oxford University Press.

Greenleaf, G., & Kemp, K. (2020). Australia's COVIDSafe experiment, Phase III: legislation for trust in contact tracing.

Haggag, O., Haggag, S., Grundy, J., & Abdelrazek, M. (2021). COVID-19 vs Social Media Apps: Does Privacy Really Matter? arXiv preprint arXiv:2103.01779.

Harri, K. (2020). Creating HOIA. https://lab.mobi/articles/HOIA-covid19-app. (accessed 14/03/2021).

Health and Welfare Information Systems Centre (TEHIK). (2020a). "How does HOIA function?" https://HOIA.me/en/ (accessed 18/05/2021).

Health and Welfare Information Systems Centre (TEHIK). (2020b). "Hoia source code," 2020, https://koodivaramu.eesti.ee/tehik/hoia. (accessed 07/05/2021).

Health and Welfare Information Systems Centre (TEHIK). (2020c). “HOIA phone application privacy policy,” https://hoia.me/privacy/. (accessed 07/05/2021).

Hellewell, J., Abbott, S., Gimma, A., Bosse, N. I., Jarvis, C. I., Russell, T. W., Munday, J. D., Kucharski, A. J., Edmunds, W. J., Funk, S., Eggo, R. M., Sun, F., Flasche, S., Quilty, B. J., Davies, N., Liu, Y., Clifford, S., Klepac, P., Jit, M., Van Zandvoort, K. (2020). Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health, 8*(4). https://doi.org/10.1016/s2214-109x(20)30074-7

Hinch, R., Probert, W. J., Nurtay, A., Kendall, M., Wymant, C., Hall, M., Lythgoe, K., Cruz, A. B., Zhao, L., Stewart, A., Ferretti, L., Montero, D., Warren, J., Mather, N., Abueg, M., Wu, N., Finkelstein, A., Bonsall, D. G., Abeler-Dörner, L., & Fraser, C. (2020). OpenABM-Covid19 - an agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing. https://doi.org/10.1101/2020.09.16.20195925

Horvath, L., Banducci, S., & James, O. (2020). Citizens’ Attitudes to Contact Tracing Apps. *Journal of Experimental Political Science*, 1-13.

Kaiser, F. K., Wiens, M., & Schultmann, F. (2021). Use of digital healthcare solutions for care delivery during a pandemic-chances and (cyber) risks referring to the example of the COVID-19 pandemic. *Health and Technology*, 1-13.

Kaptchuk, G., Goldstein, D.J., Hargittai, E., Hofman, J., & Redmiles, E.M. (2020). How good is good enough for COVID19 apps? The influence of benefits, accuracy, and privacy on willingness to adopt. *ArXiv, abs/2005.04343*.

Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical teacher*, *42*(8), 846-854.

Kim, J., & Kwan, M. P. (2021). An Examination of People’s Privacy Concerns, Perceptions of Social Benefits, and Acceptance of COVID-19 Mitigation Measures That Harness Location Information: A Comparative Study of the US and South Korea. *ISPRS International Journal of Geo-Information*, *10*(1), 25.

Kostka, G., & Habich-Sobiegalla, S. (2020). In Times of Crisis: Public Perceptions Towards COVID-19 Contact Tracing Apps in China, Germany and the US. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3693783

Kozyreva, A., Lorenz-Spreen, P., Lewandowsky, S., Garrett, P., Herzog, S., Pachur, T., & Hertwig, R. (2021). Public Perceptions of COVID-19 Digital Contact Tracing Technologies During the Pandemic in Germany.

Leith, D. J., & Farrell, S. (2020). Coronavirus Contact Tracing App Privacy: What Data Is Shared by the Singapore OpenTrace App? *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, 80–96. https://doi.org/10.1007/978-3-030-63086-7_6

Li, H., Wu, J., Gao, Y., & Shi, Y. (2016). Examining individuals' adoption of healthcare wearable devices: An empirical study from privacy calculus perspective. *International journal of medical informatics*, *88*, 8-17.

LI, J., & Guo, X. (2020). Global Deployment Mappings and Challenges of Contact-tracing Apps for COVID-19. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3609516

Li, Z., Fan, Y., Jiang, B., Lei, T., & Liu, W. (2019). A survey on QTA and opinion mining for social multimedia. *Multimedia Tools and Applications*, *78*(6), 6939-6967.

Martin, T., Karopoulos, G., Hernández-Ramos, J. L., Kambourakis, G., & Nai Fovino, I. (2020). Demystifying COVID-19 digital contact tracing: A survey on frameworks and mobile apps. *Wireless Communications and Mobile Computing*, *2020*.

Moon, M. D. (2019). Triangulation: A method to increase validity, reliability, and legitimation in clinical research. *Journal of Emergency Nursing*, *45*(1), 103-105.

Mukherjee, S. (2021). Sentiment analysis. In *ML. NET Revealed* (pp. 113-127). Apress, Berkeley, CA.

Munzert, S., Selb, P., Gohdes, A., Stoetzer, L. F., & Lowe, W. (2021). Tracking and promoting the usage of a COVID-19 contact tracing app. *Nature Human Behaviour*, *5*(2), 247-255.

Munzert, S., & Selb, P. (2020). Can we directly survey adherence to non-pharmaceutical interventions? In *Survey Research Methods 14(2)*, 205-209.

Muscato, L. (2021, February 11). *The UK's covid app made a serious difference during the winter surge*. MIT Technology Review. https://www.technologyreview.com/2021/02/11/1018010/uk-exposure-notification-contact-tracing-app-succes. (accessed 15/03/2021)

Ocheja, P., Cao, Y., Ding, S., & Yoshikawa, M. (2020). Quantifying the Privacy-Utility Trade-offs in COVID-19 Contact Tracing Apps. *arXiv preprint arXiv:2012.13061*.

O'Dea, S. (2019, November 22). *Estonia: number of smartphone users 2015-2022*. Statista. https://www.statista.com/statistics/566096. (accessed 09/05/2021)

Pagliari, C. (2020). The ethics and value of contact tracing apps: International insights and implications for Scotland's COVID-19 response. *Journal of Global Health*, *10*(2). https://doi.org/10.7189/jogh.10.020103

Raman, R., Achuthan, K., Vinuesa, R., & Nedungadi, P. (2021). COVIDTAS COVID-19 Tracing App Scale—An Evaluation Framework. *Sustainability*, *13*(5), 2912.

Raskar, R., Nadeau, G., Werner, J., Barbar, R., Mehra, A., Harp, G., & Louisy, K. (2020). COVID-19 Contact-Tracing Mobile Apps: Evaluation and Assessment for Decision Makers. *arXiv preprint arXiv:2006.05812*.

Redmiles, E. M. (2020). User concerns & tradeoffs in technology-facilitated contact tracing. *arXiv preprint arXiv:2004.13219*.

Reelfs, J. H., Hohlfeld, O., & Poese, I. (2020). Corona-Warn-App: Tracing the Start of the Official COVID-19 Exposure Notification App for Germany. *arXiv preprint arXiv:2008.07370*.

Reichert, L., Brack, S., & Scheuermann, B. (2021). A survey of automatic contact tracing approaches using Bluetooth Low Energy. *ACM Transactions on Computing for Healthcare*, *2*(2), 1-33.

Rekanar, K., O'Keeffe, I. R., Buckley, S., Abbas, M., Beecham, S., Chochlov, M., & Buckley, J. (2021). Sentiment analysis of user feedback on the HSE's Covid-19 contact tracing app. *Irish Journal of Medical Science (1971-)*, 1-10.

Runeson, P., Host, M., Rainer, A., & Regnell, B. (2012). Case study research in software engineering: Guidelines and examples. John Wiley & Sons.

Seto, E., Challa, P., & Ware, P. (2021). Adoption of COVID-19 Contact Tracing Apps: A Balance Between Privacy and Effectiveness. *Journal of Medical Internet Research*, *23*(3).

Sharma, T., & Bashir, M. (2020). Use of apps in the COVID-19 response and the loss of privacy protection. *Nature Medicine*, *26*(8), 1165-1167.

Silva, N. F. F. D., Coletta, L. F., & Hruschka, E. R. (2016). A survey and comparative study of tweet QTA via semi-supervised learning. *ACM Computing Surveys (CSUR)*, *49*(1), 1-26.

Sowmiya, B., Abhijith, V. S., Sudersan, S., Sundar, R. S. J., Thangavel, M., & Varalakshmi, P. (2021). A Survey on Security and Privacy Issues in Contact Tracing Application of Covid-19. *SN computer science*, *2*(3), 1-11.

Sun, R., Wang, W., Xue, M., Tyson, G., Camtepe, S., & Ranasinghe, D. C. (2021). An Empirical Assessment of Global COVID-19 Contact Tracing Applications. *2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE)*. https://doi.org/10.1109/icse43902.2021.00101

Tang, Q. (2020). Privacy-preserving contact tracing: current solutions and open questions. *arXiv preprint arXiv:2004.06818*.

Terviseamet. (2021). Coronavirus dataset. https://www.terviseamet.ee/en/coronavirus-dataset (Accessed 09/05/2021).

Thomas, R., Michaleff, Z., Greenwood, H., Abukmail, E., & Glasziou, P. (2020). More than privacy: Australians' concerns and misconceptions about the COVIDSafe App: a short report. *medRxiv*.

Trang, S., Trenz, M., Weiger, W. H., Tarafdar, M., & Cheung, C. M. K. (2020). One app to trace them all? Examining app specifications for mass acceptance of contact-tracing apps. *European Journal of Information Systems*, *29*(4), 415–428. https://doi.org/10.1080/0960085x.2020.1784046

Troncoso, C., Payer, M., Hubaux, J. P., Salathé, M., Larus, J. R., Lueks, W., & Pereira, J. (2020). Decentralized Privacy-Preserving Proximity Tracing. *IEEE Data Engineering Bulletin*, *43*(2), 36-66.

Vaudenay, S. (2020). *Centralized or decentralized? The contact tracing dilemma* (No. CONF).

Weiß, J. P., Esdar, M., & Hübner, U. (2021). Analyzing the Essential Attributes of Nationally Issued COVID-19 Contact Tracing Apps: Open-Source Intelligence Approach and Content Analysis. *JMIR mHealth and uHealth*, *9*(3), e27232.

Whitelaw, S., Mamas, M. A., Topol, E. & Topol, H. G. (2020). Applications of digital technology in COVID-19 pandemic planning and response. *Lancet Digit. Health* **2**, E435–E440

Williams, S. N., Armitage, C. J., Tampe, T., & Dienes, K. (2020). Public attitudes towards COVID-19 contact tracing apps: A UK-based focus group study. https://doi.org/10.1101/2020.05.14.20102269

Yin, R. K. (2018). Case Study Research Design and Methods (6th Ed.). Thousand Oaks, CA: Sage Publishing.

Zimmermann, B. M., Fiske, A., Prainsack, B., Hangel, N., McLennan, S., & Buyx, A. (2021). Early Perceptions of COVID-19 Contact Tracing Apps in German-Speaking Countries: Comparative Mixed Methods Study. *Journal of medical Internet research*, *23*(2), e25525.

Appendix 1 – Non-exclusive licence for reproduction and publication of a graduation thesis[1]

I Julius Lwanga

1. Grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis "[Thesis title]" , supervised by [Supervisor's name]
   1.1. to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;
   1.2. to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.
2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.
3. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

10.05.2021

1 The non-exclusive licence is not valid during the validity of access restriction indicated in the student's application for restriction on access to the graduation thesis that has been signed by the school's dean, except in case of the university's right to reproduce the thesis for preservation purposes only. If a graduation thesis is based on the joint creative activity of two or more persons and the co-author(s) has/have not granted, by the set deadline, the student defending his/her graduation thesis consent to reproduce and publish the graduation thesis in compliance with clauses 1.1 and 1.2 of the non-exclusive licence, the non-exclusive license shall not be valid for the period.

Appendix 2 – HOIA Survey Questions

HOIA APP Survey

Tere! I am conducting research to assess the perception and user acceptance of the HOIA App among Estonian residents. I would love to hear from you about the use of the app in tackling the COVID-19 pandemic in Estonia. It is an anonymous survey, and the information will be used for the purpose of writing a thesis. The survey should only take 3 minutes, and your responses are completely anonymous. You can only take the survey once. If you have any questions about the survey, please email me: julius.lwanga@hotmail.com. Your input is greatly appreciated!

* Kohustuslik

1. Gender *:
   - Female
   - Male
   - Prefer not to say
2. Living area
   - Harju County
   - Hiiu County
   - Ida-Viru County
   - Jõgeva County
   - Järva County
   - Lääne County
   - Lääne-Viru County

- Põlva County
- Pärnu County
- Rapla County
- Saare County
- Tartu County
- Valga County
- Viljandi County
- Võru County

3. Age
   - Below 13
   - 13-19
   - 20-30
   - 30-40
   - 40-50
   - 50-60
   - Above 60

4. Which type of smartphone do you use? *
   - Android
   - Apple
   - Don't own a phone

5. Have you heard about the HOIA app? *
   - Yes
   - No

6. Have you downloaded the HOIA App to your phone? *
   - Yes
   - No

7. If yes, how would you rate your experience with the HOIA app?

- Very dissatisfied
- Dissatisfied
- Neither agree nor disagree
- Satisfied
- Very Satisfied

8. Do you use the HOIA App to establish contact with infected individuals?
   - Yes
   - No
   - Sometimes
9. If yes, why? If no, why?

   …………………………………..
10. Has the HOIA app helped protect you against the virus or inform you about close contacts?
    - Yes
    - No
    - Maybe
11. Do you think information you provide to the HOIA app is secure in terms privacy?
    - Yes
    - No
    - Maybe
12. Do you read the privacy policy of the apps you use?
    - Yes
    - No
    - Maybe
13. Do you think the citizens should be worried about the data collected by the HOIA app and how it is used?
    - Yes
    - No

- Maybe

14. Who should be able to access the information gained from the HOIA app?
    - Individuals who respond to trace COVID-19 infection
    - Those who run COVID-19 infection tracing technology
    - Those who financially support COVID-19 Infection tracing Technology
    - Communities in which COVID-19 infection tracing technology is performed
    - Government/public sector
    - The private sector
    - COVID-19 tracing technology software companies
15. Do you think the HOIA App was a good way for the government to spend the taxpayer’s money?
    - Yes
    - No
    - Maybe
16. Do you know that the HOIA App was a collaboration between the private and public sectors?
    - Yes
    - No
    - Maybe
17. Do you think the HOIA app has helped in tackling the COVID-19 situation in Estonia?
    - Yes
    - No
    - Maybe
18. What impressed you about the HOIA App?
    - Design
    - Easy to use
    - Language of the app
    - Privacy

19. What is the most important thing for you when it comes to the HOIA App?

- Privacy of information
- Voluntary participation
- User interface
- Ease/difficulty of use
- Phone battery usage

20. What do you think could be added to the HOIA App to make it better?

……………………………………..

Appendix 3 – Interview Questions

1. What were the key aspects taken into consideration when developing COVID-19 Contact Tracing Apps? Were there any specific design choices you had to make while designing the app? If so, what? Was focus on privacy a key component of the app design process? If you had to redesign/redevelop the app, what things would you change?

2. In your opinion, how is the HOIA app different from other COVID-19 Contact Tracing apps from around the world? What makes HOIA unique? Is HOIA's technology different or better than others? If so, what are these features? Do you see it as a breakthrough technology in tackling COVID-19 in Estonia? Could you please elaborate? And why?

3. In your opinion, how would you rate the uptake in terms of usage of the HOIA App? Has your team studied the app's user acceptance when compared to others? What was your perception regarding the app 'before it was rolled out'/'after it was rolled out'? Do you think it has been a success in tackling the spread of COVID-19?

4. In your opinion, how do you think the residents perceive the HOIA App? Do you think it adds value to them in helping stop the spread of COVID-19? If so, why? Based on your experience, could you please describe the role of user's perception when designing the HOIA app? Were surveys about user acceptance carried out before the designing the HOIA APP or during the process of designing the app?

5. What key factors do you think, should be kept in mind when designing apps when typically designing health/medicine related app? Lastly, do you think HOIA has been a success? And, why so?

Appendix 4 – List of Interviewees

1. **Kerstin-Gertrud** – *Epidemiologist*, Estonian Health Board, Estonia.

2. **Laura Asu** – Team Lead and UX/UI analyst, Iglu Digital Agency, Estonia.

Appendix 5 – RStudio Code

Code Snippet 2

```
# tokenize
tokens <- tibble(text = fileText) %>% unnest_tokens(word, text)

# get BING sentiment from text:
bing_tokens <- tokens %>%
  inner_join(get_sentiments("bing")) %>% # pull out only sentiment words
  count(sentiment) %>% # count the # of positive & negative words
  spread(sentiment, n, fill = 0) %>% # made data wide rather than narrow
  mutate(sentiment = positive - negative) # # of positive words - # of negative words

# Visualizing
pBing <- plot_ly(
  x = c(names(bing_tokens)),
  y = c(as.numeric(bing_tokens)),
  type = "bar") %>%
  layout(
  title = "Bing Sentiment Value"
  )
pBing

# Word Count
bing_word_counts <- tokens %>%
  inner_join(get_sentiments("bing")) %>%
  count(word, sentiment, sort = TRUE) %>%
  ungroup()

cat("\nBING Word Count:\n")
print(bing_word_counts)

# Visualizing
bing_word_counts %>%
  group_by(sentiment) %>%
  top_n(10) %>%
  ungroup() %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(word, n, fill = sentiment)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~sentiment, scales = "free_y") +
  labs(y = "Contribution to Sentiment (BING)",
       x = NULL) +
  coord_flip()
```

Code Snippet 2

```r
# get NRC sentiment from text:
nrc_tokens <- tokens %>%
  inner_join(get_sentiments("nrc")) %>% # pull out only sentiment words
  count(sentiment) %>% # count the # of positive & negative words
  spread(sentiment, n, fill = 0) %>% # made data wide rather than narrow
  mutate(sentiment = positive - negative) # # of positive words - # of negative words

# Visualizing
pNrc <- plot_ly(
  x = c(names(nrc_tokens)),
  y = c(as.numeric(nrc_tokens)),
  type = "bar") %>%
  layout(
    title = "NRC Sentiment Value"
  )
pNrc

# Word Count
nrc_word_counts <- tokens %>%
  inner_join(get_sentiments("nrc")) %>%
  count(word, sentiment, sort = TRUE) %>%
  ungroup()

cat("\nNRC Word Count:\n")
print(nrc_word_counts)

# Visualizing
nrc_word_counts %>%
  group_by(sentiment) %>%
  top_n(10) %>%
  ungroup() %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(word, n, fill = sentiment)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~sentiment, scales = "free_y") +
  labs(y = "Contribution to Sentiment (NRC)",
       x = NULL) +
  coord_flip()
```

Code Snippet 3

```r
# Pre-Processing
corp <- tm_map(corp, removeWords, c("'s"))
corp <- tm_map(corp, stripWhitespace) # removing extra spaces (keeping only single space)
corp <- tm_map(corp, tolower) # transforming to lower case
corp <- tm_map(corp, removePunctuation)
words_with_punctuation <- SplitText(as.character(corp[[1]]))

# GRAPH CONSTRUCTION
words <- SplitText(as.character(corp[[1]])) # tokenization
tagged_text <- tagPOS(corp[[1]])
tagged_words <- SplitText(as.character(tagged_text))

# Keep only NN (Nouns) & JJ (Adjectives) tagged words

tagged_words <-c(SelectTaggedWords(tagged_words,"/NN"),SelectTaggedWords(tagged_words,"/JJ"))
#Remove tags

tagged_words <- RemoveTags(tagged_words)
selected_words <- unique(tagged_words)  # Keeping unique words.
```

Code Snippet 4

```
# Text Rank
keywords_list<- page.rank(textIgraph, algo="prpack", directed=FALSE)$vector # function from igraph
package package

# POST-PROCESSING
nodes_num <- length(nodes(text_graph))
keywords_num <- round(nodes_num/3)  # a third of the number of vertices in the graph.

ordered_keywords<- keywords_list[order(keywords_list,decreasing=TRUE)]
final_Keywords<- head(ordered_keywords,keywords_num)

# Finding TextRank of the text graph
textRank <- induced.subgraph(graph=textIgraph,vids=names(final_Keywords))

#Plotting TextRank Graph
visIgraph(textRank) %>%
  visNodes(size = 25, shape = "circle") %>%
  visIgraphLayout(layout = "layout_nicely") %>%
  visOptions(highlightNearest = TRUE,
             nodesIdSelection = TRUE) %>%
  visInteraction(keyboard = TRUE)

# Keyword List
cat("TextRank Keywords:\n")
print(names(final_Keywords))

trankEdgeList <- as.data.frame(get.edgelist(textRank, names=TRUE))

# Plotting TextRank Graph
pTrank <- simpleNetwork(trankEdgeList,
  Source = 1,                 # column number of source
  Target = 2,                 # column number of target
  charge = -100,                # numeric value indicating either the strength of the node repulsion
(negative value) or attraction (positive value)
  fontSize = 14,               # size of the node names
  fontFamily = "serif",       # font og node names
  linkColour = "#A9CCE3",        # colour of edges, MUST be a common colour for the whole graph
  nodeColour = "#1F618D",     # colour of nodes, MUST be a common colour for the whole graph
  opacity = 1.0,              # opacity of nodes. 0=transparent. 1=no transparency
  zoom = T                    # Can you zoom on the figure?
)
pTrank
```

Code Snippet 5

```r
# Coreness
coreness <- graph.coreness(textIgraph_kcore)
maxCoreness <- max(coreness)
verticesHavingMaxCoreness <- which(coreness == maxCoreness)

# Finding Kcore of the text graph
kcore <- induced.subgraph(graph=textIgraph_kcore,vids=verticesHavingMaxCoreness)

# Plotting Kcore Graph
visIgraph(kcore) %>%
  visNodes(size = 25, shape = "circle") %>%
  visIgraphLayout(layout = "layout_nicely") %>%
  visOptions(highlightNearest = TRUE,
             nodesIdSelection = TRUE) %>%
  visInteraction(keyboard = TRUE)

# Keyword List
cat("\nKCore Keywords:\n")
print(names(verticesHavingMaxCoreness))

kcoreEdgeList <- as.data.frame(get.edgelist(kcore, names=TRUE))

# Plotting Kcore Graph
pKcore <- simpleNetwork(kcoreEdgeList,
  Source = 1,                  # column number of source
  Target = 2,                  # column number of target
  charge = -100,                 # numeric value indicating either the strength of the node repulsion
(negative value) or attraction (positive value)
  fontSize = 14,                # size of the node names
  fontFamily = "serif",        # font og node names
  linkColour = "#A9CCE3",         # colour of edges, MUST be a common colour for the whole graph
  nodeColour = "#1F618D",      # colour of nodes, MUST be a common colour for the whole graph
  opacity = 1.0,               # opacity of nodes. 0=transparent. 1=no transparency
  zoom = T                     # Can you zoom on the figure?
)
pKcore
```

Appendix 6 – Links to Articles Chosen for QTA

- **Henry, L.A. (2021, February 22). Series of flops or how HOIA failed**. ***Postimees.*** https://news.postimees.ee/7185647/series-of-flops-or-how-hoia-failed. (accessed 15/03/2021)

- **Raimo, P. (2021, January 17).** ***HOIA äpp ei täida eesmärki. Parandused on vaidlustesse takerdunud***. ***Delfi Eesti Päevaleht***. https://epl.delfi.ee/artikkel/92234633/hoia-app-ei-taida-eesmarki-parandused-on-vaidlustesse-takerdunud. (accessed 15/03/2021)

- **Aivar, P. (2021, March 8).** ***Poolt ja vastu: kas HOIA äppi on sellisel kujul veel üldse vaja? Delfi Forte***. https://forte.delfi.ee/artikkel/92777189/poolt-ja-vastu-kas-hoia-appi-on-sellisel-kujul-veel-uldse-vaja. (accessed 15/03/2021)

- **Sven, R. (2021, March 16).** ***HOIA äpp on ebaõnnestunud, andmete käsitsi sisestamine lausa piinlik***. ***Postimees Tehnika.*** https://tehnika.postimees.ee/7203227/linnar-viik-hoia-app-on-ebaonnestunud-andmete-kasitsi-sisestamine-lausa-piinlik. (accessed 15/03/2021)