Bachelor thesis
Computing Science

Radboud University

Investigating Usability Problems in Email Encryption Tools Based on IBE

Author:
Quoc An Ha
s4347420

First supervisor/assessor:
Dr. Hanna Schraffenberger
h.schraffenberger@cs.ru.nl

Second assessor:
Prof. Dr. Bart Jacobs
b.jacobs@cs.ru.nl

August 24, 2022

Abstract

E-mail end-to-end encryption, while readily available, nonetheless experiences a low adoption rate. This poses a real problem concerning the privacy of people using e-mail without end-to-end encryption, leaving researchers wondering why adoption rate is so low. Various studies indicated that this is because of usability problems in the encryption tools. More specifically, research has shown that one of the biggest problems pertains to users having to manage the public keys of contacts in those tools. With the advent of Identity-Based Encryption (IBE), which eliminates users having to manage public keys, we revisit the original notion of poor usability being a reason for the low adoption rate, and see if poor usability is also the case with IBE based tools. To this end, we performed usability evaluations in the forms of an heuristic evaluation and a usability test of two tools based on IBE: PostGuard and Voltage SecureMail. While we did find a variety of application specific usability problems, we did not uncover any overlapping usability problems specific to IBE.

Contents

Acknowledgements

First and foremost, I am incredibly grateful to my supervisor, Dr. Hanna Schraffenberger, for all her help and advice with this bachelor thesis, through which I have learned a great deal. I would also like to thank Merel Brandon, for helping out with the user test pilot, Daniel Ostkamp, for his technical support regarding getting PostGuard to work on my laptop, and Leon Botros, for assisting me in understanding PostGuard's inner workings. Then, I would like to thank Prof. Dr. Bart Jacobs, for his willingness to be my second assessor on such short notice and for his feedback on my drafts. Furthermore, I would like to thank Zoubeir Bellari, Selim Berntsen, Hang Le Ha, Ludwig Arntz, and Bram van Luyken for supporting me throughout this entire process. Lastly, my thanks goes out to all the participants who donated their time for the user study.

Chapter 1

Introduction

With over 3.9 billion users worldwide and over 300 billion emails sent yearly [1], email is one of the most used communication tools of the present day. However, curiously enough, the adoption of end-to-end encryption practices for the tool remain astonishingly low [2], even though the technology has been available to the public since 1991[1] with the release of Pretty Good Privacy (PGP).

Whitten and Tygar [3] postulated in their seminal paper "Why Johnny Can't Encrypt" that the low uptake was due to usability problems in the software, showing in their study that only 4 out of 12 participants were able to successfully encrypt an email using PGP. Since then, various studies have been conducted further investigating the usability of email encryption tools [4]–[8], revealing that public key management of contacts (see section 2.1.3.1) is one of the largest obstacles for the usability of end-to-end encryption. Public key management entails the end-user being required to maintain a collection of public keys for their contacts, which have to be obtained in a secure manner as well. Furthermore, if a contact changes their public key, the end-user needs to manually update the public key in their collection as well.

This problem of having to manage public keys was already recognized by Shamir [9] in 1984, who proposed the concept of "Identity-Based Encryption" (IBE), a type of public key cryptography where the public key consists of an arbitrary (identity) string instead of a large number. The advantage of using a string as a public key, is that it removes the necessity of acquiring any recipients' public key (i.e. their random number) before encrypting a message, since the sender can simply construct the recipients' identity string (e.g. "foo@bar.com"), thereby dissolving the need of managing and

[1] `https://en.wikipedia.org/wiki/Pretty_Good_Privacy#History`

distributing public keys at all.

Since then, various email encryption tools have adopted IBE (such as PostGuard[2], Voltage SecureMail[3], and FortiMail[4]), but very little research has been conducted on their usability. Therefore, since IBE theoretically solves the problem regarding public key distribution, and with research indicating that usability is a main factor limiting widespread adoption of email encryption, we aim to investigate what usability problems exist with end-to-end encryption tools based on IBE.

Additionally, we are also interested in how we can improve a free tool based on IBE called PostGuard, developed by a team from the Radboud University's interdisciplinary research hub on digitalization and society. While we will also evaluate another product named Voltage SecureMail (see section 2.5), our interest in PostGuard specifically comes from having the oppertunity to provide constructive feedback directly to PostGuard's developers through our affiliation with the university, thereby directly contributing to a free and secure Internet for all.

The research question we will be examining is: **"What, if any, are the usability problems regarding email encryption tools using IBE?"**. In order to help us answer our main research question, we formulated four subquestions:

1. What are the usability problems in PostGuard?
2. What are the usability problems in Voltage SecureMail?
3. Are there any overlapping usability problems when evaluating both PostGuard and Voltage SecureMail, that point towards the existence of general usability issues with IBE-based applications?
4. How can PostGuard's usability be improved?

The remainder of this thesis will be structured as follows: Chapter 2 will detail the preliminary knowledge required to understand the problem area and this research. Chapter 3 contains related work to this research. Chapter 4 will present one of our two evaluation methods, the heuristic evaluation, where we will describe the approach we took, present the results, and provide a subdiscussion of those results. Then, chapter 5 presents the other evaluation method, our user study. In this chapter we will also describe our approach, present the results, and additionally provide a subdiscussion of those results. Lastly, chapter 6 presents our discussion and conclusions, as well as detailing the limitations and providing suggestions for future work.

[2] `https://postguard.eu/`

[3] `https://www.microfocus.com/en-us/cyberres/data-privacy-protection/secure-mail`

[4] `https://www.fortinet.com/products/email-security`

Chapter 2

Preliminaries

This chapter will present the preliminary knowledge required in order to understand the research. We will first present the technical background, followed by introductions to usability and IRMA, before describing the two products we will be evaluating, PostGuard and Voltage.

2.1 Technical background

This section presents the related technical knowledge required in order to understand why email encryption is a problem. To fully understand why email security is a problem, it pays to understand how email came into existence. As such, we will present the technical knowledge in chronological order, as much as possible.

2.1.1 The internet

In order to understand why the current email protocol, SMTP, is insecure, we first have to detail how information is transmitted over the Internet.

Information on electrical networks is represented in bits, for example "1001", representing the number 9 in binary. Circuit-based networks open up a connection in the form of an electrical circuit between two or more nodes and simply "turn on" the electricity to transmit 1 and "turn off" to transmit 0 until it has processed the message. The modern Internet, however, is based on packet-switching networks, which envelops messages with additional meta information describing what to do with the message. This abstraction is important since it allowed for various abstractions of information transfer. A packet is simply a piece of information like our previous example "1001", but prepended with some other information called the header, e.g. "011001". Conventions in the protocols describe how the packets are organized (see

figure 2.1), such as the lengths of the various header fields and organization of the message, called the payload of the packet.

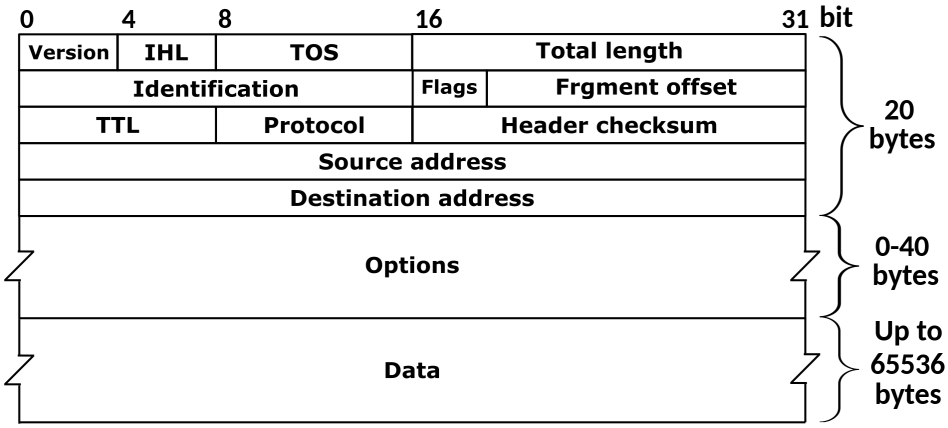

Figure 2.1: Layout of an IP packet [1]

The modern Internet is built upon this abstraction of packets, specified by the Internet protocol suite[2], also known as TCP/IP. The protocols are categorized into one of four layers, each layer implemented as a packet encapsulating the other layers above it (see figure 2.2).

The lowest layer, the one closest to the physical connection, is called the "Link Layer" and envelops local area networks, connections between two hosts on a local network (e.g. Ethernet and Wi-Fi). This layer functions similar to a connection in a circuit-switched network, differing only in that the message is sent in packets between two adjacent nodes[3].

The second layer is called the "Internet Layer", implemented as its own packet contained in the payload of the link layer packet (e.g. an ethernet packet) and hosts protocols such as the Internet Protocol (IP) and the Internet Control Message Protocol (ICMP) used for "pinging" other hosts. This layer provides the functionality required for information to pass between networks, such as from a desktop to an Internet Service Provider (ISP) to a cellular-connected phone.

The next layer is the "Transport Layer", again implemented as a packet inside an internet layer packet resulting in 3 nested packets so far, responsible

[1] By Michel Bakni - Postel, J. (Septemper 1981) RFC 791, Internet Protocol, DARPA Internet Program Protocol Specification, The Internet Society, p. 11 DOI: 10.17487/RFC0791., CC BY-SA 4.0, `https://commons.wikimedia.org/w/index.php?curid=79949694`

[2] `https://en.wikipedia.org/wiki/Internet_protocol_suite`

[3] At least, this is the case if you ignore that the message itself is nested with more packets

for the transport of information between applications on the hosts, encompassing protocols such as the Transmission Control Protocol (TCP) and the User Datagram Protocol (UDP). This abstraction allows for the streamlining of the actual data passed in the next and final layer, since a message at that level could be split up into multiple packets. In the case of TCP for example, it would ensure that the data packets arrive in order and that any lost or discarded packets are resent.

Lastly the final layer is named the "Application Layer" and is the one most relevant to our thesis since this is where email protocol lives. Additionally, other common protocols are the File Transfer Protocol (FTP), the Hypertext Transfer Protocol (HTTP), and the Domain Name System (DNS). At this stage, we have 4 nested packages, which concretely means that our original example of "1001" would be prepended with 4 headers.

All traffic sent over the Internet utilise all four layers.

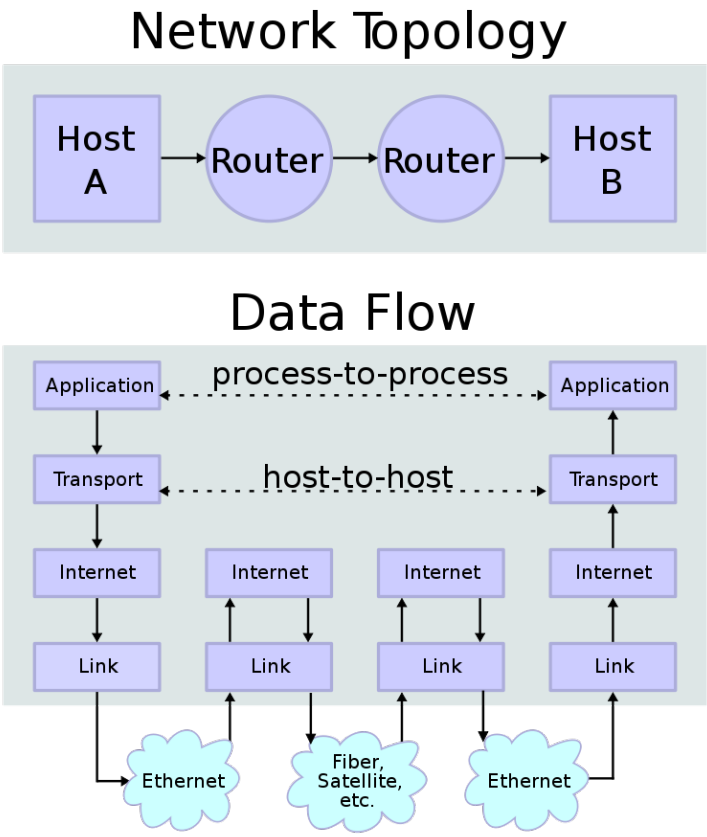

Figure 2.2: Diagram of the IP protocol stack. [4]

2.1.2 Simple mail transfer protocol

The premise of this thesis is that the current email protocol is insecure. In order to fully understand why, one first needs to understand how it works, which we will explain in this section. First we will elaborate on the actual specification of the protocol, whereafter we will summarize the security risks.

2.1.2.1 Specification

The specification of the email protocol is found in RFC 5321[5] and named the Simple Mail Transfer Protocol (SMTP). It is implemented by applications who need to send and receive emails, who are commonly called Mail Transfer Agents (MTA). MTAs receive the mail to be sent from so called Mail User Agents (MUA), which are distinct applications (e.g. Microsoft Outlook, webmail, and Mozilla Thunderbird) who compose the email according to RFC 5322. The MTA then figures out to which other MTA it has to send the mail by resolving the host names of the recipient email addresses. Noteworthy is that the domain name (the part after the "@" symbol in an email address) can point to the IP address of an inbound SMTP **relay** server, rather than the final delivery system. In that case, the relay SMTP server receives the email through the following steps as usual, before establishing a new SMTP connection to the next SMTP server to forward the email.

The sending MTA (henceforth called the smtp-client) then initiates an SMTP session over TCP with the receiving MTA (henceforth called the smtp-server) after which it is able to send SMTP specific commands. The smtp-client can then start to send mail transactions, which consist of three steps:

1. The smtp-client first sends the MAIL command that gives the sender identification.
2. Next, one or more RCPT commands are issued by the smtp-client, containing the recipient information.
3. Finally, a DATA command is initiated to issue the transfer of the mail body, ending the data with a line containing only a period, that is, the character sequence "<CRLF>.<CRLF>" (<CRLF> represents a line break and stands for "Carriage Return, Line Feed"), which also confirms the transaction.

[4]By en:User:Kbrose - Prior Wikipedia artwork by en:User:Cburnett, CC BY-SA 3.0, `https://commons.wikimedia.org/w/index.php?curid=1831900`

[5]Internet protocols are defined in "Request For Comments" documents, each identified by a number.

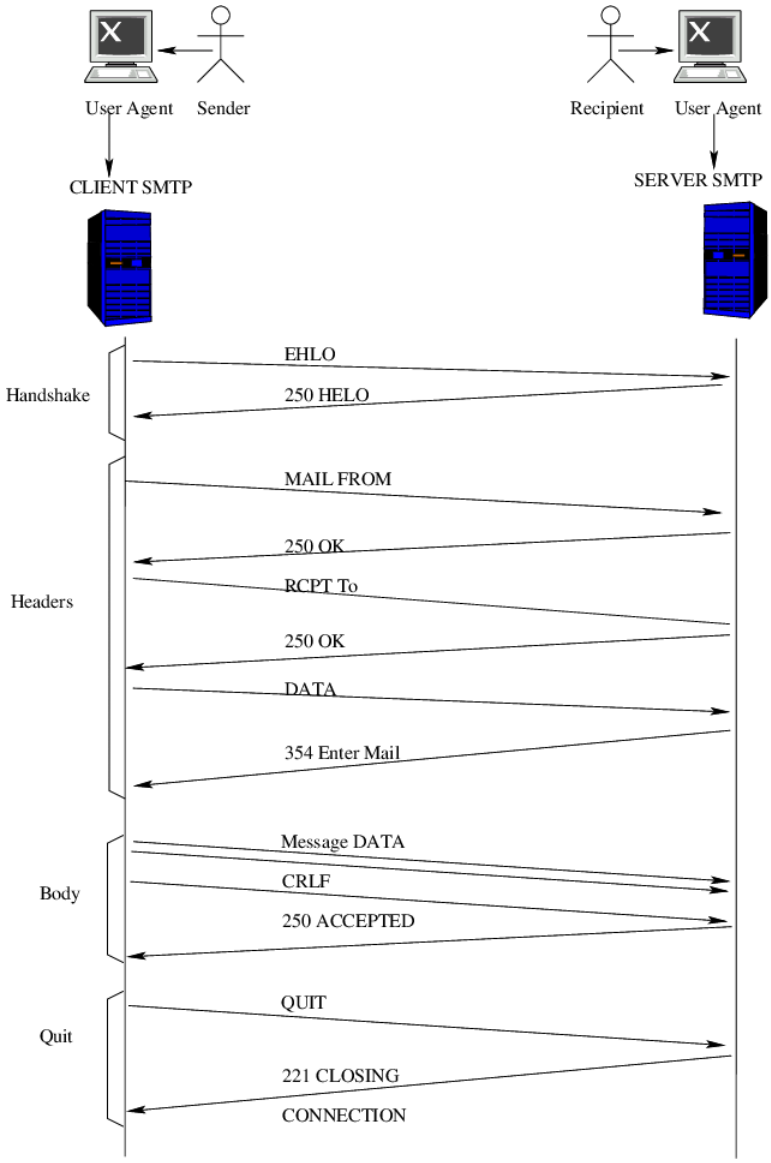

Figure 2.3: Flow of the SMTP protocol [10].

Upon reaching the final SMTP server, which acts as the first MTA for the recipient, the email is stored until polled by the MUA of the recipient. The polling is also defined by various protocols, the two most common ones being the Internet Message Access Protocol (IMAP) and the Post Office Protocol (POP).

2.1.2.2 Security risks

As mentioned before, the email protocol comes with an abundance of security flaws. It is noteworthy that the specification itself does not include any security measurements, which characterizes it as an inherently insecure protocol (in fact, section 7.1 of the specification explicitly mentions this). Concretely, the protocol is vulnerable to a concept called email spoofing (faking), which is the act of supplying a fake sender address and is relatively easy to do even for casual users. Additionally, the SMTP protocol sends its commands, which include the email data, in plain-text over TCP (also insecure) to the receiving SMTP server, allowing everyone observing the connection and specifically the transport layer to read the data. Lastly, there is no way of confirming the integrity of an email, which can happen through a malicious SMTP relay server or a Man-in-the-Middle attack. It follows that solutions are needed either internal or external to SMTP in order to securely send emails.

2.1.3 Email security

In this section we will describe various email security mechanisms currently existing, which are relevant since this thesis evaluates products whose goal is to secure emails.

We will introduce the following concepts in order:

- Public Key Encryption
- TLS Protocol
- End-to-End Encryption
- Email Authentication
- Identity-Based Encryption

2.1.3.1 Public-key cryptography

Public-key cryptography[6], also known as public key encryption, and asymmetric cryptography, describes encryption that uses a pair of two distinct keys of which either is the only feasible way to decrypt messages encrypted by the other. One is described as the public key, which describes the notion of the owner distributing that particular key while the other is described as the private key. The public and private key are theoretically cryptographically the same, but practically refer to two different types of files: the private key encompasses additional values used by the algorithm and additionally includes values to derive the public key.

[6] `https://en.wikipedia.org/wiki/Public-key_cryptography`

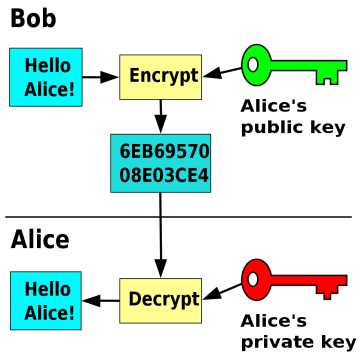

Figure 2.4: Encrypt/ Decrypt Diagram [7]

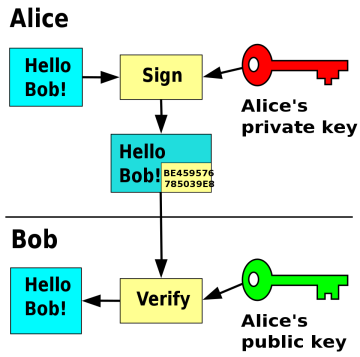

Figure 2.5: Digital Signature Diagram [8]

In the context of email encryption, public key cryptography can be used to provide authenticity, confidentiality, integrity, and non-repudiation. With authenticity meaning, an email was truly sent by the owner of the sender's email address. Confidentiality meaning, no other entity is able to read the contents of the email, other than the person whom it was intended for. Integrity meaning, the email was not tampered with; the received email is exactly the sent email. And lastly, with non-repudiation meaning, the sender not being able to deny they sent the message.

Someone attempting to send an encrypted email (see figure 2.4) requires the public key of the recipient in order to encrypt the message. The message can then be decrypted by the private key, which is only in possession of the recipients, thereby providing confidentiality. Furthermore, the sender can use a one-way hash to create a checksum of the message and encrypt it using their own private key to be send along with the encrypted message, called a digital signature (see figure 2.5). The recipient of the email can then use the sender's public key to decrypt the attached signature, revealing the hash of the original message and compare it to their own hash computation of the decrypted message, ensuring the authenticity of the sender and integrity of the message, additional to providing non-repudiation.

Asymmetric cryptography is opposed to symmetric cryptography, in which decryption and encryption both use the same key. We will elaborate on assymetric crypthography in the following section.

[7] By Davidgothberg - Own work, Public Domain, `https://commons.wikimedia.org/w/index.php?curid=1028460`

[8] By FlippyFlink - Combined changed the image `https://en.wikipedia.org/wiki/File:Public_key_encryption.svg` from encryption to signing., CC BY-SA 4.0, `https://commons.wikimedia.org/w/index.php?curid=78867393`

Public key infrastructure
Asymmetric cryptography is also called public key cryptography since it relies on the trust of the public keys actually belonging to the corresponding people. However, how can one trust that a public key actually belongs to a certain person or entity? Currently, there are two main models through which this trust is established:

1. The centralized Certificate Authority (CA) model, and
2. the decentralized Web of Trust (WoT) model.

Both models use the concept of a "certificate", which is an electronic document that includes information about the public key, information about the identity of its owner, a digital signature of the entity that has verified the certificate's content, and possibly other information depending on the implementation choices of the certificate issuer[7].

Since the WoT model is only used in OpenPGP (an open standard for encryption), we will postpone its explanation until section 2.1.3.3, where we describe how OpenPGP works. The CA model, however, is used in both TLS and S/MIME and thus makes sense to explain now before either one is introduced.

The CA model solves the trust problem by introducing third parties which are globally assumed to be trustable, called **Root** Certificate Authorities. These Root CAs all have corresponding public keys, which are included within operating systems and browsers, and can be used to verify messages signed by their private key. These Root CAs use their private key to sign public keys of regular CAs (after manual verification), thereby allowing anybody to verify the public key of regular CAs by means of the Root CAs' public key (see figure 2.6). Then, these regular CAs provide services signing public keys and creating certificates for the public. These certificates then contain the public key of the "public" user, a signature from the regular CA, and the certificate of that CA itself. If anybody then wants to verify an arbitrary certificate issued by one of the regular CA's, they simply verify the certificate of the regular CA using the public key of the root CA. If valid, that means they can trust that regular CA, whereafter they can use the public key of the regular CA to verify the signature on the "public" users' certificate. Again, if valid, this means they can trust that that public key belongs to the credentials given in the certificate, since they can trust the regular CA.

A disadvantage of using certificates is that, an encrypting entity is limited to the identity specified with the certificate they have of the receiving entity. This is in contrast to IBE (introduced in section 2.1.3.5), where an

[7] `https://en.wikipedia.org/wiki/Public_key_certificate`

encrypting entity can specify any identity for the recipient as they see fit, since with (attribute-based techniques in) IBE the identity essentially *is* the public key (and where the private key can be acquired after the public key has been used).

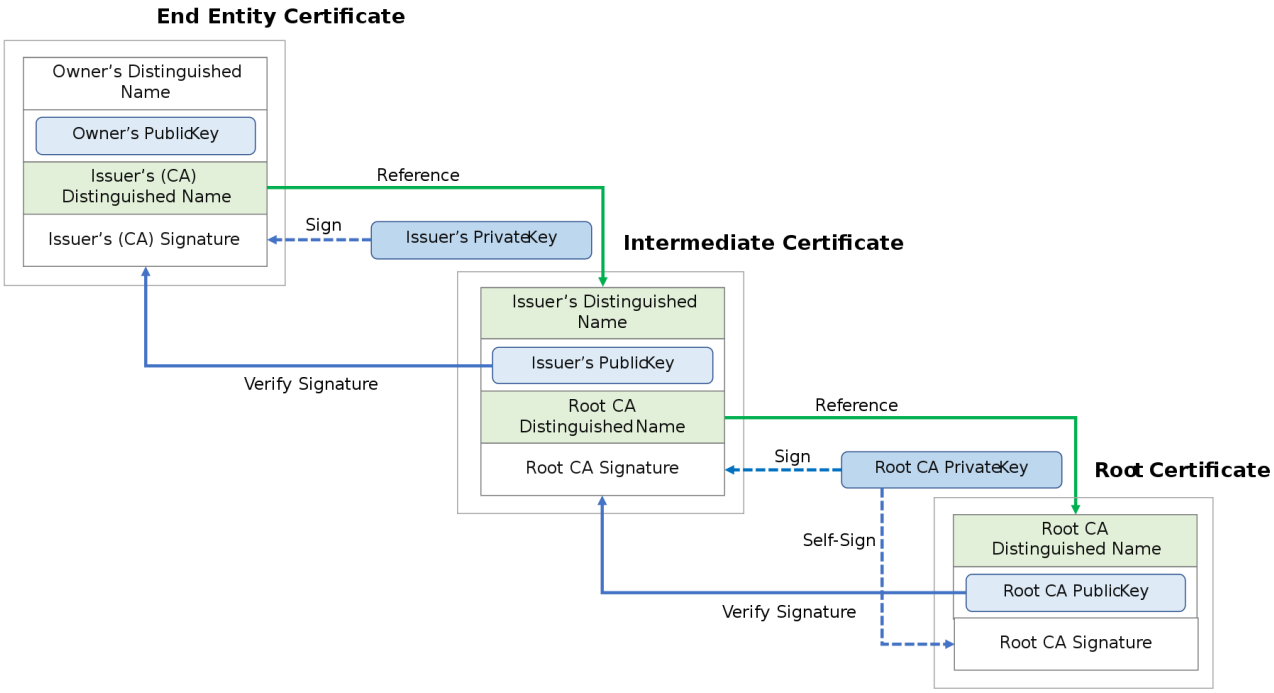

Figure 2.6: Certificate Authority Model Diagram[8]

So instead of explicitly trusting individual public keys, the trust originates from a set of so called **Root** Certificate Authorities, which is then further distributed towards lower certificate authorities. Noteworthy is that when using this system, trust is placed in the certificate authorities to check and validate that public keys in the certificates also truly belong to the the credentials in the certificates.

Disadvantages troubling any PKE system is the loss of private keys, rendering the encrypted data inaccessible.

[8]By Yuhkih - Own work, CC BY-SA 4.0, `https://commons.wikimedia.org/w/index.php?curid=94177019`

2.1.3.2 TLS protocol

The Transport Layer Security (TLS), formerly known as the Secure Sockets Layer (SSL) protocol, is a cryptographic protocol designed to provide communications security over a computer network[9]. While the name implies that it works on the transport layer, as defined by the Internet protocol suite, it actually runs on top of another transport protocol such as TCP or UDP, classifying it as an application layer protocol. However, applications usually wrap all of their actual communication inside TLS packets, making it function similarly to a protocol on the transport layer.

The protocol is designed to provide confidentiality, integrity and authenticity using asymmetric cryptography. It facilitates this by means of a handshake[10] between a client and a server over an unsecure connection, during which a symmetric encryption method is established for the actual encryption of the communication using the servers' (PKI) certificate to ensure the confidentiality of the established symmetric key and authenticity of the server. Afterwards, all communication is secured by the chosen symmetric encryption method.

STARTTLS[11], also known as "opportunistic TLS", is the term used to denote TLS being used to secure an otherwise unsecure protocol sending its data in plain-text, such as SMTP. It is derived from several protocols that use "STARTTLS" as command names for this purpose, including IMAP, POP3, and SMTP.

When used in conjunction with SMTP, TLS provides "hop-by-hop" encryption due to the SMTP being defined between hosts providing for the possibility of multiple SMTP relay servers, which means the TLS connection is terminated at each relay server.

STARTTLS is now recommended to always be used for POP, IMAP, SMTP Submission, and all other protocols used between an Mail User Agent and Mail Service Provider[12], resulting in widespread adoption. However, we want to emphasize that while there is widespread adoption for STARTTLS, intermediate points (relay SMTP servers, see section 2.1.2) can still read the contents of the email passing through them.

2.1.3.3 End-to-end encryption

In end-to-end encryption (E2EE), the data is encrypted and decrypted at the source and the destination of the information transfer using asymmetric

[9] `https://en.wikipedia.org/wiki/Transport_Layer_Security`

[10] A negotiation through the exchange of information establishing the parameters of a communication link; a mini-protocol.

[11] `https://en.wikipedia.org/wiki/Opportunistic_TLS`

[12] RFC 8314: `https://datatracker.ietf.org/doc/html/rfc8314#section-3`

cryptography. This means that the message is sent ciphered, which can then be sent over plain-text channels without reducing confidentiality and integrity, and additionally also granting authenticity and non-repudation if combined with a digital signature. Again, as explained in section 2.1.3.1, the main concern is how to establish trust in the distributed public keys, which now have to be handled by the end-users themselves.

There are two commonly used protocols for E2EE email encryption:

S/MIME
Secure/Multipurpose Internet Mail Extensions (S/MIME) is an email encryption protocol based on asymmetric cryptography with its trust model based on CAs. Being based on certificate authorities means that users can simply mail each other their digital certificates, which can then be verified using the root certificates packed with the clients browser or operating system. It additionally is well integrated with popular MUAs, such as Microsoft Outlook, and Gmail/Google Workspace, which often provide additional integrated access to Lightweight Directory Access Protocols (LDAP), which provide a service akin to a digital phone book, storing usernames, passwords, emails and additional information including the possibility of certificates.

OpenPGP
Another email E2EE protocol is the OpenPGP standard, based on the proprietary Pretty Good Privacy (PGP) encryption program, not to be confused with the GNU Privacy Guard (GPG) which is an open source implementation of OpenPGP. It is, like S/MIME, based on asymmetric cryptography and also has widespread adoption in MUAs.

OpenPGP leverages the use of the Web of Trust (WoT), which is a trust model relying on the small worlds effect on any given network, where most nodes can be reached from every other node by a small number of hops and steps (see figure 2.7).

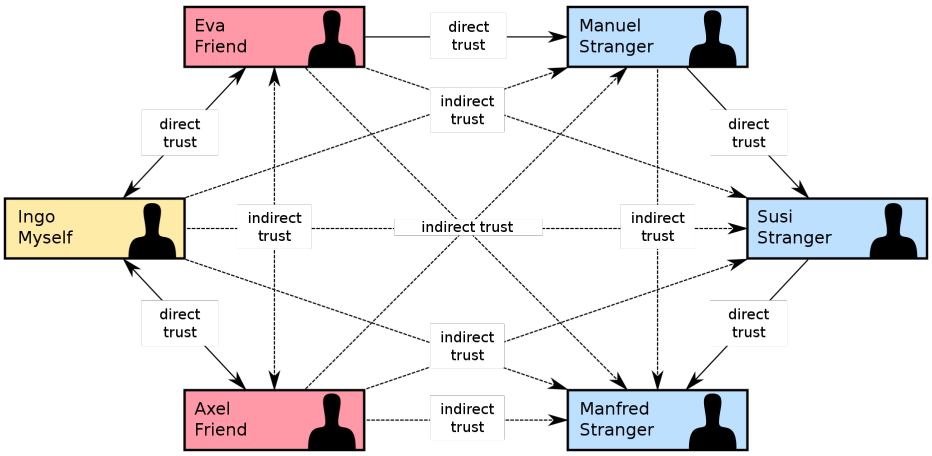

Figure 2.7: Diagram of the Web of Trust [13]

It makes use of OpenPGP certificates which encapsulate identifying information along with one or more public keys, which can be signed by other users, effectively endorsing that the public key actually belongs to the described person.

Whenever someone then attempts to validate a public key, the OpenPGP software will run an algorithm checking whether: they have either signed it personally, it has been signed by someone they fully trust (present in their current WoT) or it has been signed by three marginally trusted keys in their WoT.

This process requires people to actively go out and increase their WoT, which is typically done through key signing parties where people bring their public key on a piece of paper and hand it to other people for verification in combination with some form of identification.

Furthermore, PGP has been shown to be too complex in usage for the common user to be gain widespread adoption [3], [7].

2.1.3.4 Email authentication

In order to combat malicious modifications of the email header fields, used in email spoofing and email bounce attacks, various mechanisms have been developed that authenticate emails used in the sender header fields. They all leverage DNS records, configurable by the administrative owners of domains, to include authentication instructions queryable by receiving MTAs in order to accept or block the incoming mail.

[13] By Kku - Own work, CC BY-SA 4.0, `https://commons.wikimedia.org/w/index.php?curid=80652637`

Sender policy framework
The Sender Policy Framework (SPF) utilizes TXT records to identify allowed hosts to send emails from that domain. It is a simple mechanism where the record simply contains the domain name, or ip addresses from which the sender address can be sent.

DomainKeys identified mail
The DomainKeys Identified Mail (DKIM) leverages public key encryption in order to authenticate mails. It does this by attaching a digital signature verifying the header fields on outgoing mails while simultaneously publishing the corresponding public key in a DNS record paired with the senders domain name. A receiving MTA can then verify the received email by performing a DNS lookup on the host of the given header field in order to obtain the public key, which can then be used to decrypt the signature and verify the included hashes.

DMARC
The Domain-based Message Authentication, Reporting and Conformance (DMARC) protocol is an extension of both SPF and DKIM. It works similarly by way of setting a DNS record indicating the configured parameters, which include the activation of SPF and/or DKIM on that domain. Among other features, DMARC most notably allows for reports to be sent to the domain owner for the tracking of malicious actors.

2.1.3.5 Identity-based encryption

Our research focuses on the usability problems in email encryption tools based on Identity-Based Encryption (IBE), and as such, it is vital that the reader understands this concept in order to appreciate the purpose of this research.

IBE's main advantage in our context is that it eliminates the need for a public key distribution infrastructure. Concretely, IBE allows for anyone to **generate** any public key for any identity-string using the Master Public Key of the Private Key Generator (PKG). Through this mechanism, anyone can encrypt a message intended for any arbitrary identity-strings, even without the potential recipients having obtained the corresponding private key beforehand. If a recipient wants to decrypt a message, but cannot since they do not have the private key, they can simply query the PKG for the private key (which the PKG can also generate using its Master Private Key).

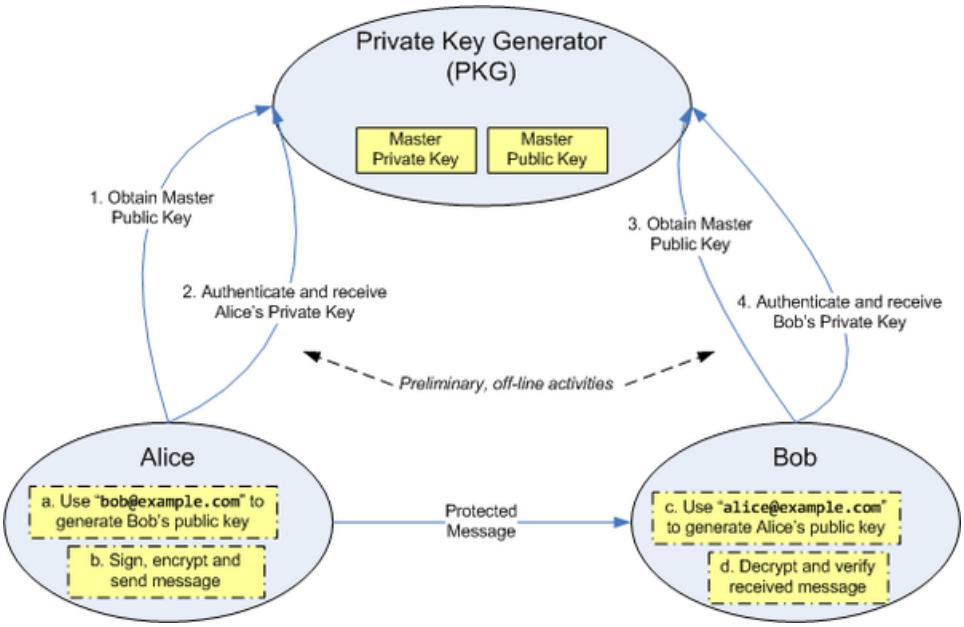

Figure 2.8: IBE steps diagram[14]

Obtaining the private key, however, should involve the PKG verifying that the identity-string actually belongs to the recipient, or in other words authenticating the recipient. However, IBE does not try to deal with the prior mentioned authentication problem for obtaining a private key, but leaves it open for implementations to choose their own authentication scheme. PostGuard, for example, uses IRMA to provide authentication (more on IRMA follows in section 2.3), while Voltage (more on Voltage in section 2.5) uses an email/password authentication mechanism.

The disadvantage of IBE is that the trust is moved entirely to the PKG, since it can generate all public and private keys for all identity-strings on command.

2.2 Usability and user experience

Poor usability has been postulated to be the main reason behind the lack of adoption of E2EE for email [3], but what is usability exactly?

Usability has been defined in ISO 9241-11 as: "The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". Any system should logically be designed to be effective and efficient, but noteworthy

[14]Source: `https://en.wikipedia.org/wiki/Identity-based_encryption`

here is the use of the word "satisfaction", which is inherently personal to each individual, making it a difficult criteria to measure. On this matter, several quantitative techniques have been developed in order to measure usability, one of the most popular being the System Usability Scale (SUS). We will first describe the System Usability Scale, followed by the heuristic evaluation method.

2.2.1 System usability scale

This research will make use of the System Usability Scale (SUS), proposed by John Brooke [11] in 1996. As the name implies, the scale measures the usability of a given system and aims to do this in a short but effective manner. It contains only 10 Likert-type scale questions, each phrased alternating positive and negative in order to stimulate careful deliberation of every question, avoiding response biases. These questions, predetermined by Brooke, are general enough to apply to any application, however, studies have shown that minor variations do not change the reliability of the scale [12]–[14].

The scale yields a single number, representing a composite measure of the usability of the system studied. Each question contributes a score in the range of 0 to 4, calculated differently for positive and negative questions. Positive questions' contributions are derived by subtracting the answer's scale position by 1, while negative questions' contributions are obtained by 5 minus the answer's scale position. This results in a score ranging from 0 to 40, which is then multiplied by 2.5 in order to get a score ranging from 0 to 100. The final score calculation formula is then as follows:

$$f(x) = \begin{cases} 1 & \text{iff answer to question } x \text{ has scale position 1} \\ 2 & \text{iff answer to question } x \text{ has scale position 2} \\ 3 & \text{iff answer to question } x \text{ has scale position 3} \\ 4 & \text{iff answer to question } x \text{ has scale position 4} \\ 5 & \text{iff answer to question } x \text{ has scale position 5} \end{cases} \tag{2.1}$$

$$\text{SUS score} = 2.5 * \sum_{n=1}^{10} \begin{cases} 5 - f(n) & \text{if } n \equiv 0 \pmod 2 \\ f(n) - 1 & \text{if } n \equiv 1 \pmod 2 \end{cases} \tag{2.2}$$

where $f(x)$ is a function returning the scale position of an answer to a question.

Our motivation for using SUS is due to its proven status; reported to be an industry standard in addition to being cited in over 1200 publications [12],

[13]. Furthermore, it has been shown that it is able to get reliable results with a small sample size of 8 to 12 participants [15], ideal for the small scale of a bachelor thesis like this one. The small amount of questions required also streamlines the experiment for each participant, since participants might already be frustrated after using an unknown system and would not want to fill in a lengthy questionnaire.

Any interpretation of the final score should be cautious not to compare it to percentages due to it being fitted within the range of 0 to 100. In order to facilitate for a better interpretation of the final score, we will map the SUS score to an Adjective Rating Scale proposed by Bangor et al. [16], which maps score ranges to the adjectives: Worst Imaginable, Awful, Poor, Ok, Good, Excellent and Best Imaginable.

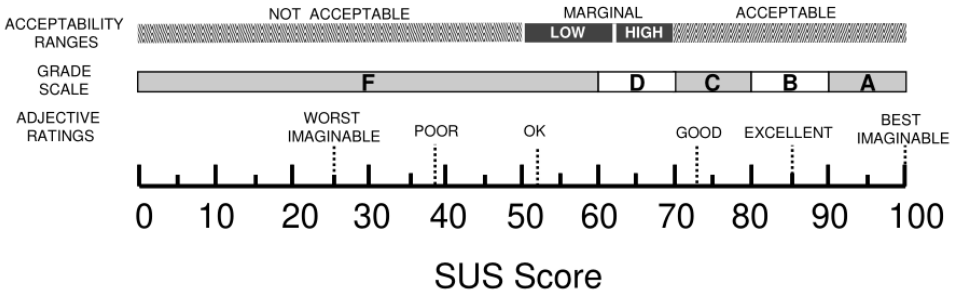

Figure 2.9: A comparison of the adjective ratings, acceptability scores, and school grading scales, in relation to the average SUS score[16]

.

Lastly, we note that SUS does not produce any diagnostic results, meaning that the questionnaire does not reveal what specifically is wrong with the system.

2.2.2 Heuristic evaluation

In this section we will provide an introduction to the heuristic evaluation method, since we will apply this technique in our research to analyze the usability of email encryption tools based on IBE.

The heuristic evaluation method is popularized by Nielsen and Molich [17] in 1990. It is an inexpensive and quick method of finding usability problems by following heuristics (or guidelines) while looking at an interface. In their study, Nielsen and Molich [17] showed that one evaluator finds between 20 and 51% of all usability problems, indicating that just one evaluator is not sufficient. However, by aggregating the results of three to five evaluators, the method appeared to find about two thirds of all usability problems, which is good for such an inexpensive method.

While their original paper [18] included only 9 heuristics, we use the widely used updated version of 2005 containing 10 heuristics, widely referred to as "Nielsens' 10 Heuristics" [19].

Nielsen' 10 heuristics [19]
(The following enumeration is sourced entirely from [19].)

1. **Visibility of system status**
   The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

2. **Match between system and the real world**
   The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.

3. **User control and freedom**
   Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.

4. **Consistency and standards**
   Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

5. **Error prevention**
   Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

6. **Recognition rather than recall**
   Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

7. **Flexibility and efficiency of use**
   Accelerators – unseen by the novice user – may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

8. **Aesthetic and minimalist design**
   Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with

the relevant units of information and diminishes their relative visibility.

9. **Help users recognize, diagnose, and recover from errors**
   Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

10. **Help and documentation**
    Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

2.2.3 User experience

While our study focuses mainly on usability, user experience is significantly related enough to usability that evaluating it in some aspects can be worthwhile. To this end, we will use Product Reaction Cards in our usability test, which we will introduce after first describing user experience in the following paragraph.

User Experience (UX) is a term coined by Donald Norman while working at Apple Computer Inc [20]. The need for this term arose due to his notion that human interface and usability were too narrow [21]. He meant for the term to cover all aspects of the person's experience with a system including industrial design graphics, the interface, the physical interaction and the manual [21]. However, in practice, a definite definition of the term was found difficult to establish [22], although one definition (ISO 9241-210:2010) was found to be a promising start [23]: "A person's perceptions and responses resulting from the use and/or anticipated use of a product, system or service [23, p. 727] [24]". User experience is not about how good or fancy an product looks, but rather transcends the product, it is truly about the experience [25].

2.2.3.1 Product reaction cards

Since we will be using product reaction cards in our usability test, we will introduce this concept in the following paragraphs.

Product reaction cards were developed by a team at Microsoft in order to measure the user experience, or specifically the"desirability" of a product [26]. It consists of a set of 118 cards, comprising 40% negative and 60% positive words. The procedure on using them is as follows. A user, after having evaluated a system, is asked to pick the words (spread out on a surface beforehand) that best describe the product, or how using the product made them feel. Once the user picked all words they found relevant, they are subsequently asked to narrow their selection down to the five most important ones. Then, the researcher or moderator of the evaluation, inquires about

the reasoning behind their five cards.

The value of this method is not necessarily the words that are picked, but rather the ensuing discussion between the researcher and the user, about the user's reasons for the words. In other words this process aims to provide qualitative insights into the product, by creating an environment for the user to talk about their experience. Although, quantitative data can be derived from this technique, for example by counting the positive and negative cards.

2.3 IRMA

We will now describe the IRMA identity platform, used by PostGuard as an authentication mechanism for proving possession of the email address attribute.

The name IRMA is an acronym for *I Reveal My Attributes*, and is under active development by the *Privacy by Design Foundation*[15] and the SIDN[16] domain name registry. The platform consists of a set of open source software projects, which implement the IBM Idemix attribute-based credential scheme [27], [28]. Central in IRMA is their mobile application (also called IRMA), which holds the attributes of a user, and acts as a passport used for authentication. Attributes can be anything, given that there is an entity that verifies and issues those attributes. An example of an attribute could be that someone lives in Amsterdam, which would need to be verified and issued by the Amsterdam municipality. Another example, more relevant to this research, is the attribute of ownership of a specific email address, which is verified and issued by SIDN.

This attribute-based credential scheme allows for anonymous authentication and signing based on an arbitrary string. The anonymity feature comes from the ability of a user disclosing nothing more than possession of the attribute (which is signed by the attribute-issuer), facilitated by zero-knowledge proofs as detailed in the Idemix specification. This notion of attribute-based credentials enhances the privacy of users by only disclosing information that is relevant, such as a predicate on age being older than 18 instead of their actual age.

Furthermore, IRMA stores its data (the signed attributes) decentralized, namely on your phone which is queried on authentication, leaving the user in control of their data.

To use IRMA[17] as an authentication method, the user installs the IRMA application on their mobile phone, which acts as their passport containing

[15] `https://privacybydesign.foundation/`
[16] `https://www.sidn.nl/`
[17] `https://irma.app/docs/what-is-irma/`

their attributes. The user can then interact with attribute-issuers, who upon verification of possession of the attribute, provide a digital signature of the attribute for the user to be stored in their passport. Then (not necessarily in this order), the user can interact with an application that needs the verification of an attribute, which it can obtain by requesting a session with the IRMA server. In short, the IRMA server responds with a session token and additional parameters, which is representable as a QR-Code. The application then displays this QR-Code to the user, who scans it using their passport (the IRMA mobile application). The mobile phone, who now has the session id, connects to the session created on the IRMA server, initiating the authentication process and prompts the user for permission to share the attributes. When both succeed, the IRMA server sends only the necessary attributes to the requesting application, concluding the authentication process. A diagram of this process can be found at Figure 2.3.

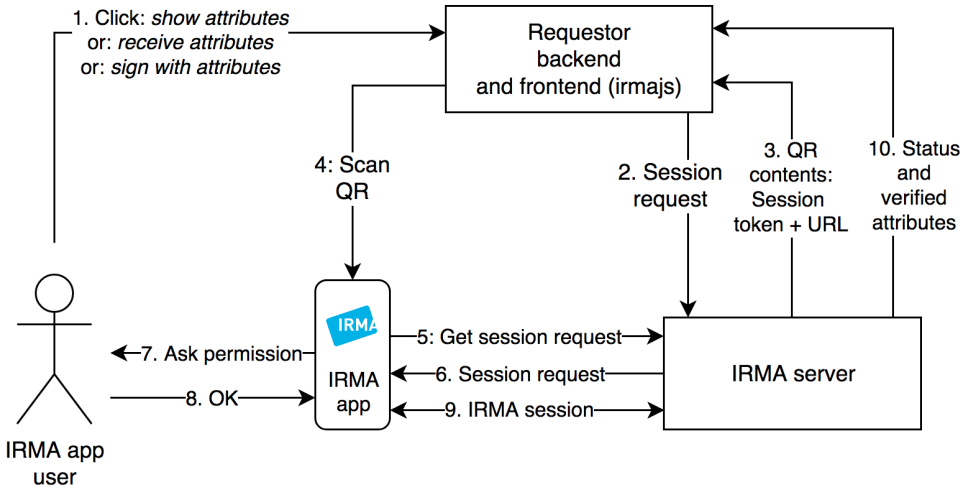

Figure 2.10: IRMA session diagram [18]

2.4 PostGuard

The first product based on IBE (see section 2.1.3.5) that we will be evaluating is PostGuard. PostGuard[19], formerly known as IRMASeal, is an email encryption tool developed by a team at iHub[20] at the Radboud University. The tool focuses on the usability, accessibility, and user experience of free and open source encryption technology.

[18] Source: `https://irma.app/docs/what-is-irma/`
[19] `https://postguard.eu/`
[20] `https://ihub.ru.nl/`

By leveraging IBE, users no longer require managing public keys themselves. Instead, they obtain the keys automated and on-demand through the email client software and the PKG. The only additional action required by the user (excluding the setup of the tool), is to authenticate themselves by scanning a QR-Code when decrypting an email. This QR-Code is a consequence of the team's choice of authentication: PostGuard uses the IRMA identity-platform (see section 2.3), although other options are possible within the specifications of IBE.

The tool concretely offers the free service of a PKG, which provides the Master Public Key[21] to the public, and retrieval of the private keys. Additionally, PostGuard provides several decrypting methods, currently consisting of a webpage[22] and various email add-ins/plugins (Thunderbird, Outlook, and Outlook Web). Encryption is provided through the add-ins only.

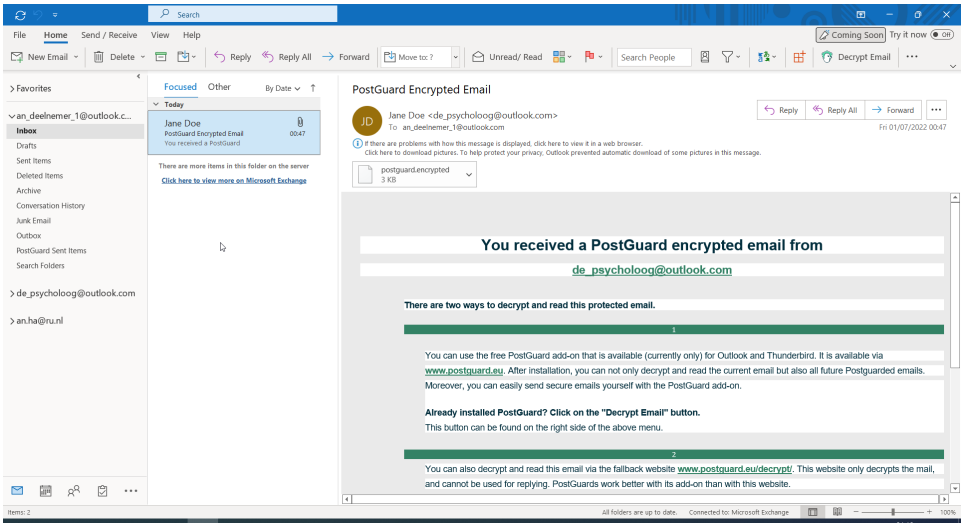

Figure 2.11: PostGuard integrated within the Outlook application (the button "Decrypt Email" in the top-right corner).

2.5 Voltage SecureMail

The other product based on IBE that we will be evaluating is Voltage SecureMail. Voltage is a Software-as-a-Service based encryption solution rolling its own IBE solution called "Micro Focus Voltage Identity-Based Encryption".

[21] `https://main.irmaseal-pkg.ihub.ru.nl/v2/parameters`
[22] `https://postguard.eu/decrypt/`

It uses[23] the AES block cipher in CBC mode for message encryption[24], IBE for key wrapping and public key exchange, a standard elliptic curve-based algorithm[25] for public key operations, and an S/MIME-based message structure.

Voltage concretely provides a software bundle for Windows, including their "Voltage Encryption Manager", which manages the identities and public keys, as well as presumably handling the cryptography. This bundle also installs an Outlook plugin that encrypts and decrypts emails using IBE and their cryptographic implementations. Additionally, they also provide a way to send and receive encrypted email without any downloads required, aptly named, the "Zero Download Messenger" (ZDM).

Furthermore, the authentication mechanism Voltage uses for their IBE solution is through an email/password scheme over SSL.

Lastly, we would have liked to report on who runs the PKG of Voltage's IBE solution, as well as the up-time statistics of the service, and confidentiality guarantees, however, unfortunately, we were not able to find this information on their website or in their documentations.

[23] `https://www.microfocus.com/media/white-paper/rethink_email_encryption_eight_best_practices_for_success_wp.pdf`

[24] Explaining block ciphers is outside the scope of this thesis, refer to `https://en.wikipedia.org/wiki/Block_cipher#Rijndael_/_AES`, and `https://en.wikipedia.org/wiki/Block_cipher_mode_of_operation#Cipher_block_chaining_(CBC)` for more information.

[25] Explaining elliptic-curve cryptography is outside the scope of this thesis, refer to `https://en.wikipedia.org/wiki/Elliptic-curve_cryptography` for more information.

Chapter 3

Related work

In this chapter we will discuss the existing research related to this research. We will first look at research studying the usability of email encryption tools, whereafter we will other explanations than usability for the low adoption of email encryption. Lastly we will end with a section discussing existing work specifically related to PostGuard.

3.1 Usability of email encryption tools

The first usability study focusing on security was conducted by Whitten and Tygar [3] in 1999, a time in which the importance of (internet) security began to gain traction. They argued that most security failures were caused by configuration errors rather than flaws in the underlying technology, specifying the reason for the failures as a user interface design problem. In order to learn more about security user interfaces, they searched for existing software representing the then-best user interface design for security, resulting in their choice to evaluate of PGP 5.0. PGP was interesting not only because of it's well designed interface by general consumer software standards, but also because it leveraged the popular concept of public-key management. They first acknowledged that a general definition of usability for security was required in order to evaluate any software on it:

Definition 3.1.1. Security software is usable if the people who are expected to use it [3]:

- Are reliably made aware of the security tasks they need to perform
- Are able to figure out how to successfully perform those tasks
- Don't make dangerous errors
- Are sufficiently comfortable with the interface to continue using it

Whitten and Tygar [3] subsequently applied that definition to the context of PGP 5.0 in order to formulate specific criteria for PGP to be usable. With the evaluation criteria in hand, they opted for two evaluation methods: an informal cognitive walkthrough in combination with evaluation heuristics (based on the aforementioned usability definition), and a user test performed in a laboratory. The cognitive walkthrough, a simulation of system-use by the evaluators as if they were novice users, identified various interface design flaws, such as metaphors concealing the concept of private and public keys and encryption algorithms, components hidden behind menu's (keyserver), and irreversible user actions without adequate warnings. Their user test culminated with only a third of their 12 participants being able to successfully sign and encrypt an email, with a quarter of them accidentally fatally exposing their private key in the process. It became apparent that participants did not understand the public key model well enough in order to develop intuitive and correct usage of the program. The combined assessments from the two methods drove Whitten and Tygar to deem PGP 5.0 unusable for people who are not already knowledgeable in the area of security.

Whitten and Tygar's [3] paper eventually proved to be seminal and encouraged other researchers to do much work in the field of usability and security.

A related follow-up study was conducted by Garfinkel and Miller [8] in 2005 titled "Johnny 2: A User Test Of Key Continuity Management with S/MIME and Outlook Express". The authors observed that S/MIME along with its widespread integration in email clients eliminated many of the usability failings found by Whitten and Tygar. Signing and encrypting emails in these programs had been reduced to one single click, displaying a warning when no corresponding public key was present. Additionally, S/MIME included a rudimentary version of key distribution: digital signatures included the sender's public key, verifiable by the recipient's chain of trust, followed by automatic import to the recipient's address book. Garfinkel and Miller, however, noted that S/MIME still required users to obtain their public key somehow, and argued that many users were still unwilling to go through that vital process. Corroborating this notion, they based their "Johnny 2" paper on a reinterpretation of Whitten and Tygar's paper: namely that the usability problems were caused by the public key model and not due to interface of PGP 5.0. Following that argument, they opted to research a method that automated various actions related to key management, aptly named Key Continuity Management (KCM) software. KCM radically simplified the security model by trusting each individual public key, similar to SSH[1], absolving the need for trust chains. Email clients implementing KCM would

[1] A SSH server generates a host key fingerprint upon setup, which is presented to each connecting client. Each client then explicitly confirms this fingerprint the first time it encounters the fingerprint. Then, the client is notified every time the host key fingerprint changes.

automatically create a self-signed key pair for a newly used sender address, which was then attached to all outgoing emails from that respective address. Additionally, these keys would also be used for signing the outgoing emails from that respective email address. A receiving client then automatically imported any attached key to be used for verifying inbound email from that address, and encrypting outgoing mails to that address. Once established, receiving and sending secure mail would be fully automatic, only requiring action when a known public key has changed. Namely, when a known public key changed, the user was notified and prompted to accept or decline the new key, which they would have to evaluate themselves. Normally, as with SSH, a change in the public key would only happen if the sender used a different machine, or in the worst-case scenario by an attack of some kind, and was left up to the user to make that evaluation.

Their study design opted to resemble Whitten and Tygar's [3] experiment as much as possible, but due to KCM requiring no actions to encrypt or decrypt content, they changed the scenario (which originally required a participant to generate a personal keypair, obtain public keys of contacts, and use those encrypt and send them an email) to include an attack instead. This required the users to successfully detect that attack and to properly withhold the sensitive information.

Garfinkel and Miller [8] found in their results that while KCM significantly increases security, "it is not the panacea to the mail security problem". KCM still required users to independently trust new identities, since it provided no additional utilities for this.

Seven years after Whitten and Tygar's seminal paper [3], Sheng et al. [7] followed up with a similar study in 2006, but instead using PGP 9 and Outlook Express 6.0 in order to find out whether the new version had solved the initial problems. They had their participants perform the following tasks: create a key pair, get public keys, verify public keys, encrypt an email, sign an email, decrypt an email, verify a digital signature, and save a backup of public and private keys. While they found improvements regarding the creation and distribution of public keys, none of the participants were able to encrypt any outgoing emails, in addition to no one being able to digitally sign a message. Furthermore, no user was able to verify their keys, noting confusion about the reasoning to do so.

Sheng's findings indicated that public key management still remained too complex for casual users, corroborating and validating Whitten's findings still seven years later.

In 2017, Bai et al. [29] researched the opinion of average users about the tradeoffs between centralized and decentralized key-distribution models. Different from other studies focussing on the design of usable encryption

interfaces, they instead examined users' opinions of the encryption models' underlying properties. They found that users perceived the use of a centralized key-directory service to be considerably more convenient than manually exchanging and managing public keys themselves.

In 2018, Ruoti et al. [30] conducted the first A/B evaluation study comparing shared secrets (passwords), public key directories (PKD), and IBE. The authors opted for a within-subjects study, using a paired participant methodology where two familiar people were tasked with sending an encrypted email to the other. In order to limit confounding factors between the evaluation of each key management system, they created MessageGuard: a Gmail browser-addon providing end-to-end encryption for email, which featured an architecture allowing for easy replacement of the key management system used for the encryption. Using this system, they then implemented IBE, password (shared secret), and PKD versions to be used in the study. Using SUS as their quantitative measurement, they found PKDs and IBEs to be more usable than shared secrets. Additionally, they compared their systems against publicly available systems, finding their own system ultimately superior in all cases. Most notably for us is that their IBE variant trumped Voltage Mail with 14.64 points on the SUS scale. Lastly, while stressing that this result was not statistically significant, IBE did appear to be slightly ahead when the participants were asked their favorite system.

3.2 Other reasons for low adoption

As mentioned earlier, usability was shown to be one main of the reasons for the low adoption rate of end-to-end email encryption [3]–[8]. While usability is the main focus of this research, it is interesting to examine other reasons for the low adoption rate, since ultimately the goal of any tool is to actually be used.

In 2014, Renaud et al. [31] looked for other reasons than poor usability to explain the low adoption rate of end-to-end encryption. To that end, they conducted a qualitative study examining the end-user's mental model of emails and email security. They postulated 7 explanations, displayed in table 3.1, for why people do not use end-to-end encryption, deduced from a natural progression ranging from awareness, to understanding, to acting. They then conducted a user study and a literature review, attempting to find support for those explanations. Their results confirmed 6 out of 7 explanations, as seen in table 3.1.

Proposed Explanation	Literature	Participant statements
1. No Awareness		(✓)
2. No Concern	✓	✓
3. Misconceptions of How to Protect	✓	✓
4. No Perceived Need to Take Action	✓	✓
5. Needs to Take Action But Does Not Know How to Act	✓	✓
6. Inability to use E2E Encryption	✓	
7. Becoming Side-Tracked	✓	

Table 3.1: Findings of support through literature or interviews [31]

In 2017, Abu-Salma et al. [32] noticed an uprising in popular communication tools embracing end-to-end encryption, such as Whatsapp, iMessage, Signal, and Telegram. With so many people having encountered secure communication methods, they wondered how much users understood the protection these tools offered. To that end, they conducted a study interviewing 60 participants about their experience and perceptions with the security features of the tools. They found, amongst other results, that the vast majority did not understand vital aspects of end-to-end encryption, which limited their motivation to adopt and use them.

In 2019, Dechand et al.[33] also investigated the perception of end-to-end encryption following the mass adoption by popular messaging tools. Their key findings, amongst others, were that users did not trust encryption, lack awareness and did not feel targeted. Additionally, they also found that users' tend to underestimate the power of cryptography, holding the notion that anything can be decrypted by a sufficiently equipped hacker.

In 2020, Reuter et al. [34] investigated two questions: 1) why are users hesitating to use email end-to-end encryption and 2) which usability issues exist that hinder users from securing their daily email communication using end-to-end encryption. To this end, they conducted an online survey in addition to a user test. They found that, while users were generally aware of the importance of email encryption, over 60% reported to be unaware of the existence of encryption technologies. Furthermore, they found that users were overwhelmed with the management of public keys.

3.3 Previous work relating to PostGuard

In 2022, Starren et al. [35] conducted a bachelor thesis research on a prior version of PostGuard (then called IRMASeal), comparing its usability di-

rectly to PGP. To this end, they performed a cognitive walkthrough using the tools and conducted a between-subjects experiment evaluating users using both tools. Their selected email client, contrasting to this research, was Thunderbird instead of Outlook. Furthermore, since IRMASeal was not usable at the moment, they performed the evaluations using a clickable mockup created by the developers instead of the actual product. They were able to gather 9 participants for their user study, and asked them to use IRMASEAL to encrypt and decrypt emails, and to use PGP to encrypt emails (since PGP automatically decrypted emails, requiring no user intervention). They found that, while every participant was able to successfully encrypt and decrypt emails using IRMASeal, only five out of nine participants could decrypt using PGP. Furthermore, they used the SUS usability measurement to gather quantitative data on the usability, resulting in a score of 85.3 for IRMASeal, and 46.1 for PGP.

3.4 Take-aways

We will now shortly describe what we gathered from the previously discussed related work.

We found that usability tests are the most used evaluation methods for usability [3], [7], [8], [29], [30], providing both quantitative and qualitative insights into problems with the software. Furthermore, we observe a trend in the studies focussing more and more on the usability of the public key management systems [29], [30], as postulated by Garfinkel and Miller [8], which strengthens our motivation to research the usability of IBE. Lastly, we also learned that the user's perception plays an important role in the adoption of encryption tools [31]–[33].

Chapter 4

Heuristic evaluation

In this chapter, we will conduct a heuristic evaluation of both PostGuard and Voltage SecureMail in order to identify any possible usability problems (see section 2.2.2). We will adhere to the widely used 10 heuristics presented by Nielsen in 2005 [19] (see section 2.2.2).

We first started by identifying the tasks involved in using the products and formulated four tasks:

1. Decryption without installation of any software.
2. Installation of the Outlook add-in.
3. Sending an encrypted email.
4. Receiving and decrypting an encrypted email.

We then stepped through the products, performing these tasks, while keeping the 10 heuristics in mind, attempting to identify all usability problems.

We want to mention that, while a heuristic evaluation should preferably be conducted by 3 to 5 evaluators [17], this research only had access to one evaluator (namely, the author of this paper).

Additionally, we assume during the evaluation for PostGuard that the user (whom we are pretending to be) has never heard of both PostGuard and IRMA, thereby evaluating how a completely new user would progress through the software. We assume the same for Voltage; we imagine ourselves as a user completely new to the product.

The remainder of this chapter will be structured as follows: We will first detail the evaluation for PostGuard, subdivided by each task. For each task, we will first describe what actions are needed to complete the task, whereafter we will describe the problems found while executing those tasks.

Afterwards, we will present Voltage in the same manner. Lastly, we will present a preliminary subdiscussion of the results found, to be further used in our discussion and conclusions in chapter 6.

4.1 PostGuard

4.1.1 Decryption without installation

Upon receiving their first encrypted email without any prior knowledge, we assume that most users would prefer to not make any permanent changes to their system in order to read the encrypted message. As such, we will take a close look at any usability problems accomplishing this. We will first detail the involved general steps before listing any usability problems encountered.

Decrypting a PostGuard encrypted message starts with the receipt of the default encrypted message email (see figure 4.1). The email always has the subject "PostGuard Encrypted Email", comes with a file attachment named "postguard.encrypted", and details in its message body the instructions for decrypting the original message. The email mentions the sender's email address, and provides instructions on two ways for decrypting (through the fallback website, and by using the add-in). It furthermore contains two paragraphs informing recipients about what PostGuard is, what IRMA is, and how they need to use IRMA to prove that they are the intended recipient. Additionally, the email provides two links to the App store (iOS) and Google Play (Android).

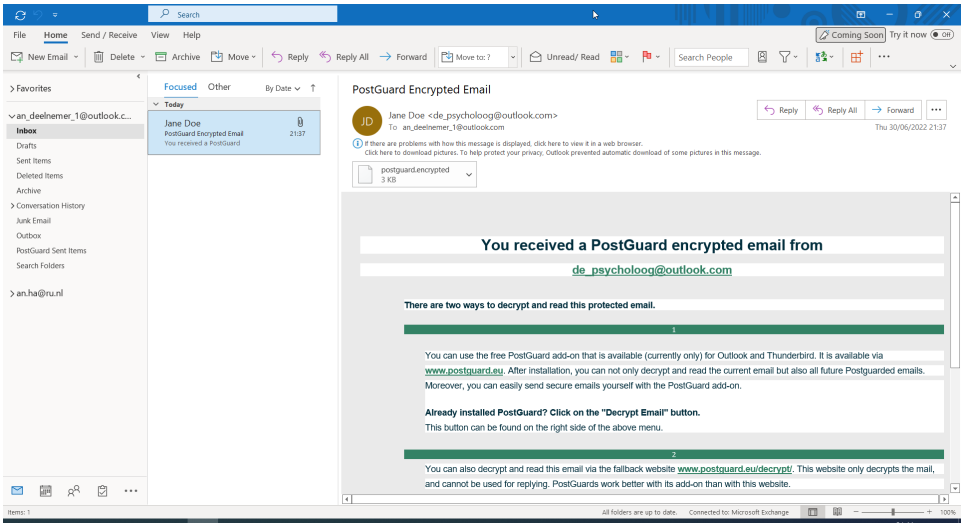

Figure 4.1: PostGuard Email

In order to decrypt, the user has to either install the add-in, or use the

fallback website at `postguard.eu/decrypt` (see figure 4.2). Currently installing the add-in is not available to the public: the product is still in alpha testing and requires an email request to obtain access. Consequently, the fallback website is the only publicly available decryption method.

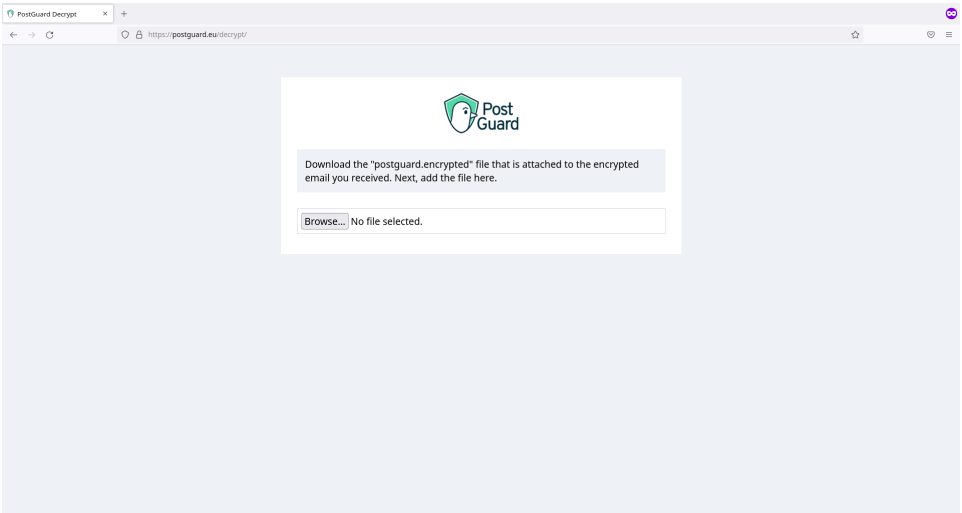

Figure 4.2: PostGuard Fallback Website

The fallback page contains one paragraph of instructions with a file picker button below it. Selecting a "postguard.encrypted" file through the picker presents a modal window with a QR-Code and a reference to IRMA (see figure 4.3). The user is then required to use the IRMA app, loaded with the respective email address attribute, to scan the QR-Code which then presents the original email message upon success (see figure 4.4).

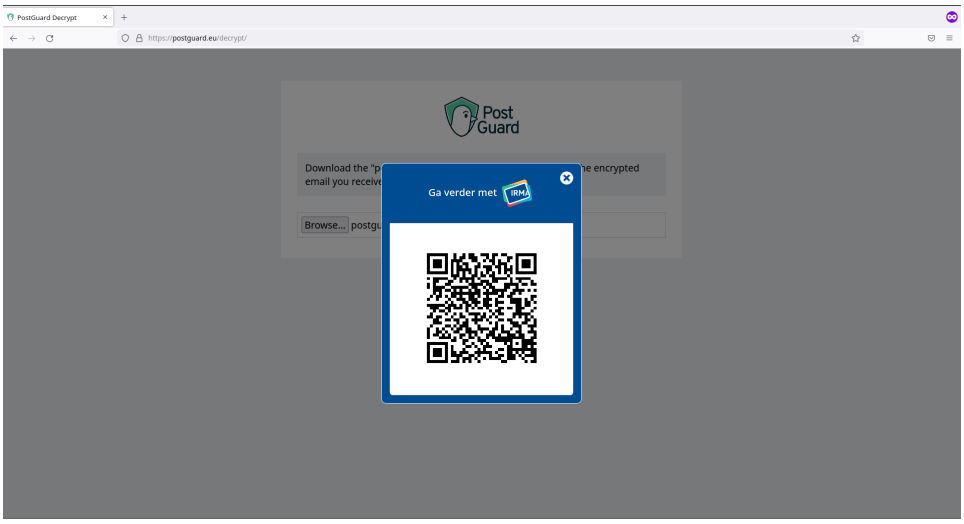

Figure 4.3: PostGuard QR-Code modal window.

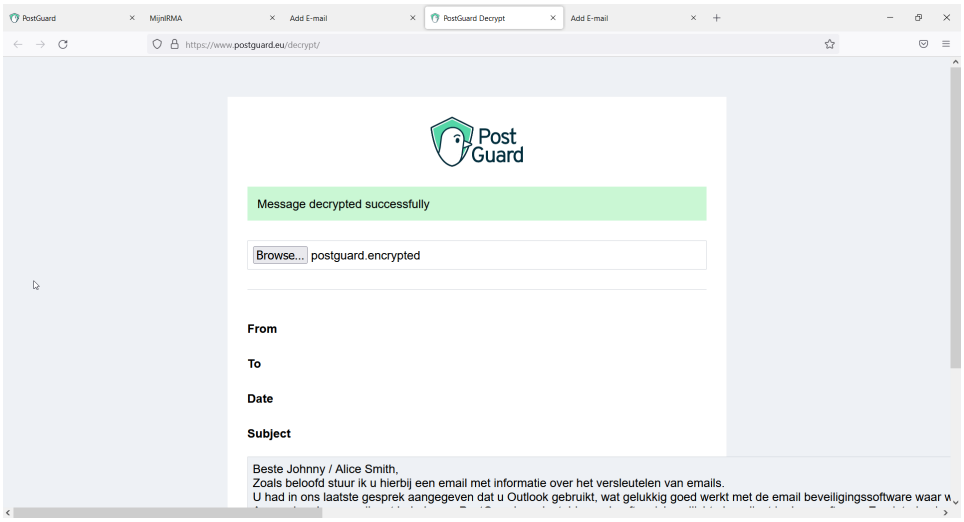

Figure 4.4: The fallback website displaying an email after successfully authenticating.

4.1.1.1 Results for PostGuard's subtask 1

Problem 1	
Context	Email content of the default PostGuard information email.
Description	While the email does mention all components required in order to decrypt the email (a link to `https://www.postguard.eu/` and `https://irma.app/`), it is not focused on the user's goal (namely decrypting). Navigation to each of the websites requires exploration on the user's part in order to find the specific information required for decrypting.
Related Heuristics	• Help and documentation
Suggestions	Consider providing a complete overview of all required steps for new users in one place.

Problem 2	
Context	Fallback page at `postguard.eu/decrypt/`
Description	There is no information on the page about the purpose of the page, instead the user is required to either infer it from the instructions, or to recall it from the email or `postguard.eu`.
Related Heuristics	• Recognition rather than recall • Help and documentation
Suggestions	Consider providing information about PostGuard and IRMA on this page for new users. Since the fallback website might be the first encounter with PostGuard for impatient users, consider providing an overview of all required steps including setup and usage of IRMA.

Problem 3	
Context	Picking a file at `postguard.eu/decrypt`
Description	Picking a file subsequently shows a modal containing a QR-Code. This shows an assumption of the system that: 1) the user possesses the IRMA app and 2) has loaded the email address pertaining to the encrypted message. If one of those two conditions are not fulfilled, the action cannot be completed.
Related Heuristics	• Error Prevention • Help and documentation
Suggestions	Consider a confirmation screen inquiring about those conditions before displaying the QR-Code, or displaying these conditions complementary to the QR-Code.

Problem 4	
Context	QR-Code pop-up at `postguard.eu/decrypt/`
Description	The pop-up containing the QR-Code does not provide concrete instructions on how to proceed. This is problematic for new users, in that they have to find out themselves where to look for information, or recall that information from the email (if they actually read it). Furthermore, it depends on recurring users memorizing the actions required to proceed (although we do acknowledge that this is actually quite recognizable with the QR-Code alone).
Related Heuristics	• Recognition rather than recall • Help and documentation
Suggestions	Consider providing them a way to obtain the knowledge required in order to decrypt the message.

Problem 5	
Context	Fallback website at `postguard.eu/decrypt`.
Description	The initial instructions on the website are in English, but the information (including the time-out message, see problem 6) given in the QR-Code modal window are in Dutch.
Related Heuristics	• Consistency and standards
Suggestions	Provide information in a consistent language, optionally also provide a choice for multiple languages.

Problem 6	
Context	QR-Code modal window at `postguard.eu/decrypt`.
Description	When the QR-Code has not been scanned within a certain amount of time, the QR-Code disappears with a non-descriptive message: "Sorry! We did not hear anything from you for too long."(translated from Dutch), with a link underneath labeled "Try again". This message does not explain why an action is required within a specific timeframe, and does not mention to have the IRMA app ready before trying again.
Related Heuristics	• Help users recognize, diagnose, and recover from errors
Suggestions	Consider providing a concrete explanation as to why the timeout occurred and inform the user to have the IRMA app ready with their email attribute loaded. Furthermore, consider a more concrete label for the "Try again" link (e.g. "Generate a new QR-Code").

Problem 7	
Context	Fallback page at `postguard.eu/decrypt/`
Description	The file picker button labelled "Browse" does not disappear when displaying the decrypted email content, but stays visible with the name of the selected file next to it. It is not clear what happens when the button is clicked and a new file is chosen without trying. Upon attempting to select another file, it becomes evident that the user is able to decrypt another message through the filepicker. However, this functionality is also provided through the button on the bottom labelled "Decrypt another email". While this functionality might provision advanced users in decrypting multiple messages quickly, it might also add additional cognitive load to new users. The heuristic for advanced users would be "Flexibility and efficiency of use" with the countering heuristic for novice users being "Aesthetic and minimal design".
Related Heuristics	• Visibility of system status • Flexibility and efficiency of use • Aesthetic and minimalist design
Suggestions	Consider replacing this input group with a clear message about which file has been decrypted without a button.

4.1.2 Installation of the Outlook add-in

Installation of the Outlook add-in provides an integrated method of decrypting PostGuard encrypted messages. Additionally, the add-in is the only method of sending an encrypted message through the service. The instructions are currently available at `postguard.eu/install_instructions` as a hidden link (no mention of the link on any other page).

Installation occurs through the add-in manager provided by Outlook. Currently, users have to add a custom add-in by means of a URL, instead of retrieving it from the official Outlook add-in repository. We acknowledge that this is because of the alpha stage the product is in, and note that it will likely change in the future.

4.1.2.1 Results for PostGuard's subtask 2

Problem 8	
Context	Instruction page at `postguard.eu/install_instructions.html`
Description	The instructions page for the add-on does not provide an overview of all the compatible email clients. Users have to scroll through the entire page in order to find out which clients are compatible.
Related Heuristics	• Help and documentation
Suggestions	Consider adding a table of contents at the top of the page or creating dedicated pages for each client.

Problem 9	
Context	Add-in instruction page at `postguard.eu/install_instructions.html`
Description	The instructions on the page do not provide any screenshots of the add-in within the respective email clients. Thereby requiring the user to remember the instructions, instead of having a visual cue.
Related Heuristics	• Recognition rather than recall • Help and documentation
Suggestions	Consider adding images of the email client's interface in order to assist the user in the installation process.

4.1.3 Sending an encrypted email.

The Outlook add-in provides an additional button in the ribbon of the email compose window labeled "Send encrypted Email" (see figure 4.1). Clicking on this button sends the current composing email securely.

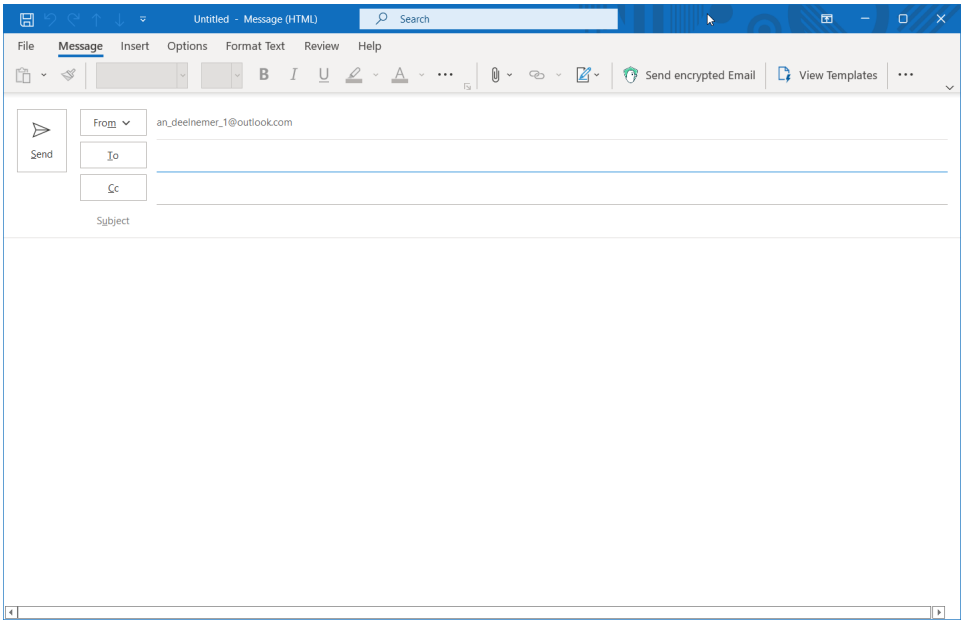

Figure 4.5: Email compose window with PostGuard's button in the ribbon

4.1.3.1 Results for PostGuard's subtask 3

Problem 10	
Context	Email compose window in Outlook
Description	Users unaware of how the add-in works might click on the default "Send" button instead of the "Send encrypted Email" button.
Related Heuristics	• Error Prevention
Suggestions	Consider adding a warning the first time an unencrypted email is sent, warning the user that it will not be secure while providing a way to hide the warning in the future.

4.1.4 Receiving and decrypting an encrypted email.

Decrypting a PostGuard encrypted email is done through a button labeled "Decrypt Email" on Outlook's ribbon. When clicking, the add-in prompts for a QR-Code to be scanned with the IRMA app. Once the identity has been verified, it is subsequently cached until 4 AM, after which the user has to re-identify themself.

4.1.4.1 Results for PostGuard's subtask 4

We found no usability problems regarding this task.

4.2 Voltage

4.2.1 Decryption without installation

The default email that comes with a Voltage SecureMail encrypted message provides two options for decryption on a desktop.

The first method is to download and open the attachment called "message_zdm.html". This static (and offline) HTML page contains a button labeled "Click to Read Message" which takes you to the (online) Voltage Zero Download Messenger (ZDM) which is their online tool for reading and sending encrypted email. The ZDM prompts for the creation of an account if none has been registered for the receiving email address. If none existed and the user created one, a verification email is sent containing a verification link that doubles as a link to view the original message. If one exists, then the user is prompted to enter their password, followed by the display of the original message.

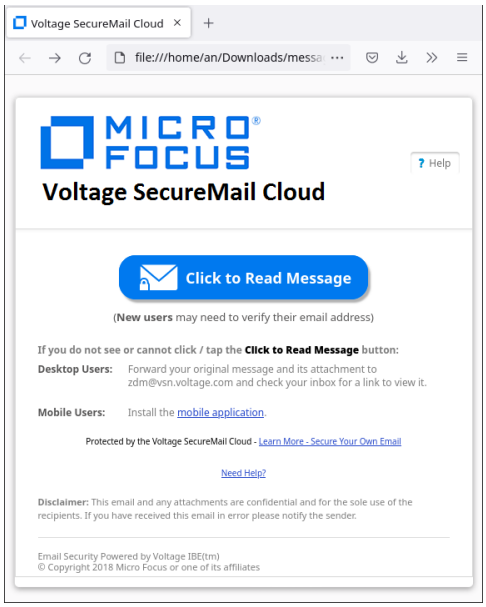

Figure 4.6: Contents of the file "message_zdm.html".

The other method is to forward the default email to "zdm@vsn.voltage.com" which in turn sends another email with a link to their ZDM. Similarly to the first method, the user is prompted through the link for an account creation or their password, displaying the original message upon completion.

4.2.1.1 Results for Voltage's subtask 1

Problem 1	
Context	The ZDM account creation page.
Description	The page does not provide any information as to what the purpose of the web page is, but simply instructs "Create a password to continue".
Related Heuristics	• Visibility of system status
Suggestions	Consider adding information about that they are creating an account, as well as information as to why an account is needed.

4.2.2 Installation of the Outlook add-in

The Voltage add-in for Outlook is a COM/VSTO type add-in, as opposed to the more novel Web add-in type that PostGuard utilizes. The installation

of COM/VSTO add-ins happen through 'regular' Windows setup files and allows the add-in to run code locally as opposed to within a Web environment.

The Voltage add-in for Outlook comes with the "Voltage SecureMail and SecureFile" software setup. This setup installs the "Voltage Encryption Manager" which handles the identities for the SecureMail add-in and SecureFile software. Upon installation, an extra button is added in the email composing window labeled "Send Secure" which in turn, upon clicking sends the current composing email securely.

4.2.2.1 Results for Voltage's subtask 2

Problem 2	
Context	The setup program installing the Voltage software.
Description	The setup software never mentions what features it actually installs. Furthermore, the website does not mention this information anywhere as well.
Related Heuristics	• Visibility of system status
Suggestions	Consider adding this information clearly in one of the setup dialogues.

4.2.3 Sending an encrypted email

Sending an encrypted email is facilitated through the added button "Send Secure" in the email composing window (see figure 4.8). Clicking this button sends the current composing email securely.

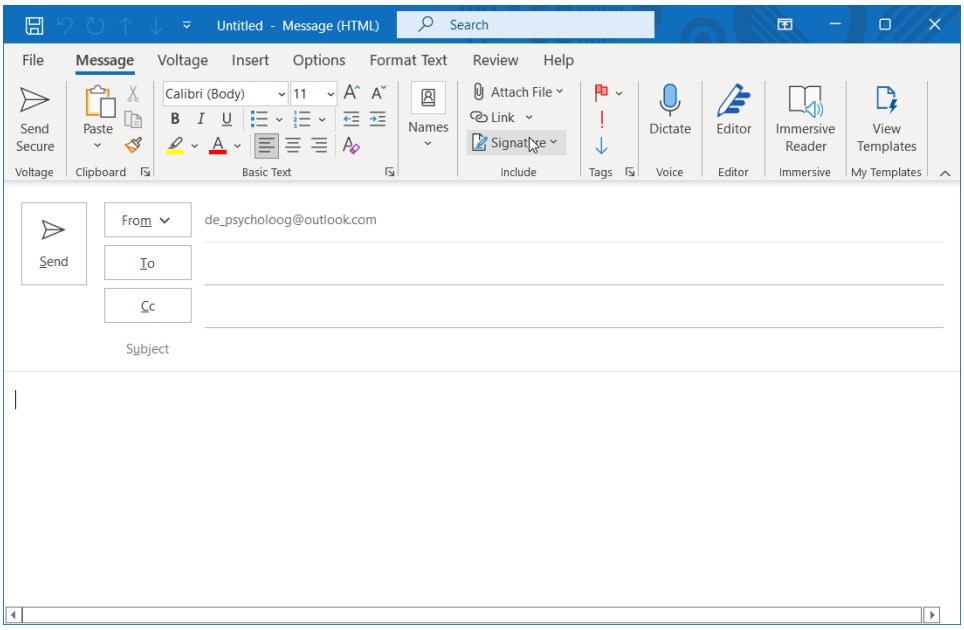

Figure 4.7: Email compose window with Voltage's button in the ribbon.

4.2.3.1 Results for Voltage's subtask 3

Problem 3	
Context	Email compose window in Outlook.
Description	Users unaware of how the add-in works might click on the default "Send" button instead of the "Send Secure" button whilst assuming that their message is sent secure.
Related Heuristics	• Error prevention.
Suggestions	Consider adding a warning the first time an unencrypted email is sent, warning the user that it will not be secure while providing a way to hide the warning in the future.

4.2.4 Receiving and decrypting an encrypted email

Emails encrypted using Voltage are automatically decrypted upon opening the email. The add-in automatically checks for the correct identity within the external program "Voltage Encryption Manager", prompting authentication if not found.

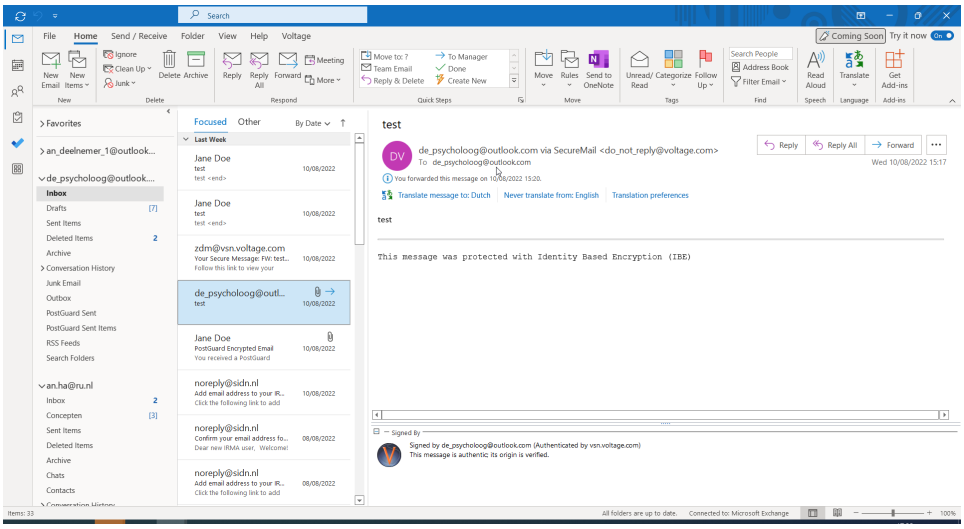

Figure 4.8: An automatically decrypted email by Voltage.

4.2.4.1 Results for Voltage's subtask 4

Problem 4	
Context	Clicking on a Voltage encrypted email without having the correct identity.
Description	When the identity has not been found, Voltage prompts for authentication through several dialogues. However, these dialogues do not mention the reason for which the actions are required.
Related Heuristics	• Help and documentation
Suggestions	Consider adding the purpose of the actions in the dialogues.

4.3 Subdiscussion

We will now present a preliminary (sub)discussion of the results of these heuristic evaluations for each product. The intention of this section is to digest these results and thereby partially answer our first three research subquestions in the context of the heuristic evaluation method. Additionally, the findings presented here will serve as keypoints to be quoted and used as arguments in the overarching discussion in chapter 6.

We found a total of 14 usability problems through the heuristic evaluation method, with 10 belonging to PostGuard and 4 to Voltage SecureMail. We will first discuss the findings of PostGuard, followed by Voltage, before ending with any overlapping findings.

4.3.1 PostGuard

When we look at the results of the heuristic evaluation of PostGuard, we can see that most of the problems (9 out of 10) contribute (albeit negatively) to the overall experience of a new user getting acquainted with the system.

We found that the very first email a user receives from PostGuard already is inadequate by providing insufficient information, a pattern also prominently seen on the fallback website and it's QR-Code pop-up. Furthermore we also found some minor inconveniences on the web page detailing the add-in installation. It is crucial that these problems are solved, since a first impression is important and can deter an otherwise happy user in adopting the tool for frequent use.

Inversely, we conclude that, once familiar with the system, users will encounter few usability frustrations with the system. Supporting this notion is the observation that 6 out of 10 problems are related to Nielsens' "Help and documentation" heuristic, since familiarity with the system dissolves the need for instructions. Moreover, once a user has installed the accompanying add-in, as we assume most frequent users of PostGuard will do, then 8 of the 10 usability problems will cease to be relevant since they pertain specifically to the fallback website or the add-in installation instructions. Specifically, only one problem was found specific to the core features of the Outlook add-in, which is an ambiguity in functionality of the button to send an encrypted email. This ambiguity might lead to a user inadvertently sending an unencrypted email meant to be secure, which is the polar opposite of what the software is trying to do. However, since this button along with the scanning of the QR-Code are the only two actions required for frequent use, we expect the user to quickly memorize and internalize this functionality. In other words, we dare say that this usability problem is insignificant for advanced users.

As such, we finally conclude from this heuristic evaluation that PostGuard has various usability problems for new users, possibly hurting adoption in the future if not addressed, but is expected to perform well for users familiar with the software.

4.3.2 Voltage

Of the 4 problems we identified with the heuristic evaluation method applied on Voltage, we found only 1 to be of major importance, while considering the other three to be minor inconveniences.

The only major problem is regarding the possibility of a user sending an unencrypted email that was intended to be encrypted, which defeats the purpose of the tool.

Then, we found three minor problems, which we regard minor because they do not actually inhibit the usage of the tool: the problems are all insufficient in informing the users about why an action is required, but nonetheless do effectively guide the user in what actions they have to perform. The first of these is that there is no explanation as to why a user should create an account and what it is used for. The second instance is whilst importing a new identity into the Voltage Encryption software. The dialogue is simple enough to be error-prone, but there is again no explanation as to what an identity is and why it is important. The last problem is concerning the software installation of the tool. Voltage does not provide any information as to what is being installed and what it is used for. While the user can function effectively without the missing knowledge addressed above, we argue that it is important to educate users about how security tools work in order to be motivated to actually use them.

With one usability problem and only three minor problems, we conclude from this heuristic evaluation that Voltage should perform well for both new users and advanced users.

4.3.3 Overlapping problems

We will now discuss any overlapping issues found between the two products.

Analyzing the overlapping problems of both products, we can see first and foremost, that they both lack in providing information to the end-user. We most clearly saw this regarding information on usage instructions for both products, both in concrete instructions and justification for their actions (Why does a user have to create an account for Voltage?).

We extend this problem to that both applications do not provide any justification as to why their products are useful and needed. Both products rely on the semantics of the word “security” in order to sway potential users.

Finally, both tools do not inform the user when an email has been sent, or will be sent, insecurely. Consequently, both tools do not allow for error recovery if an email that was supposed to be encrypted was sent unencrypted.

Chapter 5

Usability Test

This chapter will detail the usability test conducted in order to find usability issues relating to IBE-based email encryption software. Contrasting to the heuristic evaluation, this method involves actual participants, thereby providing us with genuine insights from non-expert perspectives. Furthermore, while we initially considered a within-subjects approach, time constraints required us to opt for a between-subjects evaluation instead. Concretely, this meant that each participant only evaluated one tool, instead of both tools.

5.1 Test setup

The study was designed to simulate a real-world scenario requiring the use of secure messaging through email (we will describe the scenario and test design in section 5.3). The tests were conducted between the time period of July 12th 2022 to August 7th 2022. We managed to gather 12 willing participants, of which 2 instances failed to complete the test (due to a software bug and insufficient English), resulting in a dataset consisting of 10 participants. Each session took about 90 to 120 minutes, with about 60 to 90 minutes being assigned to the test itself and the remaining time being used for questionnaires and interviews.

The tests were conducted at the residences of the participants. Upon arrival and initial greeting, we set up the materials (described in section 5.4) after which we started with the pre-test phase of the study. The pre-test phase encompassed the presentation of information about the structure of the study, the signing of a consent form by the participant, a demographics questionnaire, a primer on the thinking-aloud method (literally describing their thought process out loud) which we asked them to apply during the test, and the sending of a test email using Outlook to the researchers' email

account. The test itself was then conducted by the participant with the researcher observing and taking notes, additionally the researcher provided help after a waiting a reasonable amount of time in case the participant got stuck. During the test, the researcher would halt the participant at certain milestones in order to ask them a few questions about the test so far. After completion of the test, the participants were asked to fill in a post-test questionnaire and a system usability scale questionnaire (see section 2.2.1). Furthermore, a post-test interview was conducted, after which the experiment concluded with an evaluation of the participants' impression of the software utilizing product reaction cards (see section 2.2.3.1).

Additionally, a screen and audio recording was taken from the moment the test started to the end of the product reaction cards, which allowed for additional review of each session afterwards.

5.2 Demographics

We recruited 12 voluntary participants for our study from our private circle, however, one participant appeared to have insufficient English comprehension in order to follow the software instructions (which were only available in English) and one evaluation session was halted due to a bug in the PostGuard software, resulting in a data set of 10 participants.

Participants were one-third female: male (7; 70%), female (3; 30%). The participants skewed towards the ages between 26 to 39 years old: 18 to 25 years old (3; 30%), 26 to 39 years old (7;70%). The participants were also diverse in educational background: high school (3; 30%), MBO (2; 20%), HBO (2: 20%), WO Bachelor (2: 20%), WO Master (1; 10%). Only one participant had a background in an information technology related field (1; 10%). Contrasting the low percentage of IT-related backgrounds, the majority of participants provided a satisfactory explanation of the meaning of the term "encryption": (6; 60%). A satisfactory explanation of the term "encryption" was evaluated loosely: any mention of the term "versleuteling" (the Dutch translation for encryption) or "beveiliging" (the Dutch translation of "security") regarding information was regarded satisfactory, since these terms indicate that they at least know what the purpose and context of the term is. The participants furthermore mostly used email every day: a few times per day (6: 60%), a few times per week (3; 30%), a few times per month (1; 10%). The overwhelming majority had used Outlook as their email client (participants could enter multiple familiar clients): Outlook (9; 90%), Gmail (4; 40%), iOS Mail (2; 20%). Lastly, none of the participants reported to be familiar with either software products tested (PostGuard and Voltage).

5.3 Scenario and test design

The participants were asked to play along in a scenario where they emulated a person who had just been to a session with their psychologist. In that session, their psychologist proposed to set up encrypted emails facilitating any acute need of communication, for example in the case of a mental breakdown.

The participants were assigned alternating to use either PostGuard or Voltage SecureMail in their email exchange with their psychologist. In both cases, the psychologist (played by the researcher) would send an encrypted email using either product, resulting in the participant receiving the default information email sent out by the products detailing how to decrypt the accompanying encrypted message. PostGuard would replace the original subject of the email with "PostGuard Encrypted Email", while Voltage kept the original subject of the email. Both products however, replaced the original email content with their decryption instructions. After the participants successfully decrypted the email without installing any software, they were asked by the psychologist in the email to also install the respective Outlook plugin/add-in and to send back an encrypted email containing the date of their next appointment. After sending the date in an encrypted email, the psychologist would reply with a confirmation that the message had been received, concluding the usability test.

During the test, the researcher would interrupt the participant at various milestones in order to ask them questions about the process so far. These milestones were designed such that the interruptions would impact their natural working flow as little as possible. These milestones, named after the actions performed by the participants, are:

1. Decryption using the browser.
   Interruption would occur whenever the initial email was decrypted.
2. Installation of the plugin/add-in and sending an encrypted email.
   Interruption after successfully sending the encrypted email.
3. Receiving of an encrypted email.
   Interruption after having read the last received email.

The two initial milestones concerning the initial decryption and the installation of the plugin/add-in allowed for a usability evaluation of users completely unfamiliar with the products, providing insights into any issues regarding adoption of the products. The latter two actions, sending and receiving encrypted mail, further tested the usability of regular usage.

5.4 Materials

In this section, we will describe the materials that were used in the usability tests. We will first give a broad overview of the materials, followed by a more detailed description of the forms used. All forms can be found in the appendix.

The materials that were used in this study consisted of two laptops, an Android mobile phone, a notepad to take notes on, and the study forms containing the script, questionnaires, interview questions, and product reaction cards. Additionally, we created an email address for each participant to be used in the scenario. One of the laptops was prepared with Outlook and configured with the participant's email, to be used by the participant during the test. The other laptop was used by the researcher in order to send the emails as the psychologist. The forms consisted of a bundle intended for the participant containing an information letter detailing the test, a consent form, a task description, a demographic questionnaire, test-specific questionnaires, the SUS questionnaire, and a datasheet containing their "character's" information. The datasheet concretely contained the name of their psychologist, their character's first and last name, their character's email address and password, and the time and date of their next appointment with the psychologist. The other forms were for the researcher and contained the script to be followed as well as the interview questions and the product reaction cards form. The mobile phone was used for users assigned to the PostGuard group, which was to be used with IRMA. The phone did not come pre-installed with IRMA and thus required the participant to install it themselves.

5.4.1 Questionnaires, interviews, and product reaction cards

We will now describe the various forms used in the tests, ordered chronologically with respect to the tests.

5.4.1.1 Demographic questionnaire

Before starting the task, the participants were asked to fill in a demographics questionnaire in order to map their background. In this questionnaire, we focused on their prior experience with the tools they would have to use: Windows, Android, Outlook, PostGuard, and Voltage. This data is useful when interpreting the results, and possibly identify causes for issues that might pop up.

5.4.1.2 Task interviews

During the task, we interrupted the participant at predetermined milestones (designed to disturb the natural flow as little as possible), to ask them questions about the process so far, in order to obtain qualitative data about the products. With these questions, we attempted to identify any frustration points experienced during the task, which translate to usability issues. We also posed questions about their understanding of the steps they had taken in order to obtain an image of their mental model of the product.

5.4.1.3 Post task questionnaire

After completion of the task, we asked the participants to fill in a short questionnaire consisting of two questions. These questions inquired about the ease-of-use of the product and conversely the amount of frustration they had while performing the task. This provided us with some quantitative data for comparison between the products.

5.4.1.4 System usability scale

After the post task questionnaire, we asked the participants to fill in the System Usability Scale (SUS) questionnaire pertaining to statements about the software accompanied by Likert-type scale responses for the participants to indicate their agreement.

We did not use the original English SUS statements presented in [11], but used a Dutch translation of the statements, since we also conducted the tests in Dutch. We obtained the translation from [36] and did not change words ourselves.

5.4.1.5 Post task interview

After the SUS questions, we asked the participants questions about the product in general. More specifically, we asked them what they liked or disliked about the product and asked them to elaborate. We also inquired about the trust that they would place in the product, again illuminating their perception of the product. Furthermore, we questioned them whether they would use the product in their daily life, to see what value they place in securing their emails and thus the value of the product. Lastly, we explicitly asked them whether there was something they would change in the product, as a catch-all question to highlight any missed issues.

5.4.1.6 Product reaction cards

Lastly, the participants were asked to pick any number of cards from the 118 product reaction cards (see section 2.2.3.1) spread out on a suitable surface

(the table or floor were used). After an initial selection, they were then asked to filter the selection down to exactly five cards if the number of cards exceeded five, and to pick additional cards to bring the total up to five if the number of cards was less than five. We then noted down their choices and asked for them to elaborate on the picks, providing us both quantitative and qualitative data.

We did not use the original English words, since the tests were conducted in Dutch. Instead, we used a translated version, with the cards presenting both the English word alongside the Dutch translation. This translated version was provided to us by our supervisor, who had the version available from a previous usability study. Additionally, a few translations were adapted by the author of this paper. These cards can be found in appendix F.

5.5 Study development

We initially designed the study to be a within-subjects type study, however, after conducting a pilot with our first supervisor and one of PostGuard's designers (Merel Brandon), we decided to limit each participant to one product and to turn it into a between-subjects study. This decision was made due to the amount of time it took to complete the test for one product: with the thinking-aloud process and the intermediate interviews taking up a lot of time alongside the test, it took in total about 60 minutes just to evaluate one product. Thus, since we have limited resources and it was deemed too much to ask 3 hours of each participant, we opted for a between-subjects approach instead.

5.6 Results

This section will describe the results we found for each product and relate to our first two research subquestions: 1) what are the problems with PostGuard? And 2) what are the problems with Voltage SecureMail? We intentionally omit discussion about any overlapping qualities, reserving that topic for the subdiscussion answering subquestion 3 in section 5.7 and more broadly in chapter 6.

We will present the results for each product by result-type: beginning with the timings, followed by the SUS scores, after that the observed usability problems and software issues, then the interview results, finally ending with the Product Reaction Cards.

5.6.1 Timings

We extracted the timings of various subtasks from the tests using the screen and audio recordings made during the sessions. These subtasks are based on the milestones described in section 5.3, however, since we were also interested in how long the installation process itself took, we split the "milestone" of "Installation of the plugin/add-in and sending an encrypted email" into two distinct parts. We extracted the timestamps for the beginning and end of each subtask, for each participant, providing us with the duration of the subtasks. The subtasks we have identified are:

1. Decryption using the browser.
2. Installation of the plugin/add-in.
3. Sending of an encrypted email.
4. Decrypting an encrypted email within Outlook.

Unfortunately, an error occurred in the setup: for some participants we forgot to hook up a mouse, resulting in the user navigating using the touchpad which might impact the speed at which people read and click.

We will first present the timing results for PostGuard, followed by the results for Voltage.

5.6.1.1 PostGuard

Participant	**Subtask 1**	**Subtask 2**	**Subtask 3**	**Subtask 4**
29311	21:38	07:06	12:26	00:17
65570	22:39	04:28	07:45	02:58
30850	19:26	05:00	10:07	01:01
73720	06:03	03:33	02:52	01:39
27601	15:36	06:03	03:53	00:53
Mean	17:04	05:14	07:25	01:22
Std.Dev	06:44	01:23	04:03	01:01

Table 5.1: Timings for the PostGuard group

First, we re-iterate that the timing results include the overhead introduced by the thinking-aloud process. As such, these results are not to be interpreted as realistic timings of how long each task takes. We will now present the results for each subtask in order.

Looking at table 5.1 we can see that subtask 1, involving the decryption

of an email using the browser, took on average 17 minutes and 4 seconds. While the tasks themselves are not comparable, its sheer duration invites a deeper inspection since such a lengthy time-requirement for the fulfillment of a single goal might in itself be a usability problem. Upon closer examination, it appears clearly that the extra time is taken up by the process of setting up the IRMA application and the subsequent loading of the respective email address attribute. Every participant evaluating this product attempted to decrypt the message without having the IRMA application ready, which made their action ineffective and required them to execute it again. Additionally, upon encountering the QR-Code in the previously mentioned (fruitless) attempt to decrypt, every participant was also confused as to what their next step should have been, adding to the time of the task. This result indicates that there might be a lot of ground to gain here timewise by guiding the users to be sure to have IRMA before initiating the decryption process.

Subtask 2 involved the installation of the Outlook add-in and took an average of 5 minutes and 14 seconds. We interpret this result to not be extraordinary considering the thinking-aloud process paired with having to read and follow instructions.

Subtask 3 involved sending an encrypted email using PostGuard and took an average of 7 minutes and 25 seconds. Noteworthy is the standard deviation of 4 minutes and 3 seconds, indicating that some users took much longer than the quickest user (by definition of standard deviation). Upon closer inspection of all cases, we discover that in the two cases taking the longest the extra time is explained due to the participant erroneously sending the email unencrypted, as the result of wrongly interpreting the usage of the "Send encrypted Email" button. In the third-longest case (07:45), the participant appeared to be very meticulous in following and reconfirming the correctness of each step they were taking.

Subtask 4 involved decrypting an encrypted email, concretely only requiring clicking the "Decrypt Email" button and scanning the QR-Code. The average time this task took was 1 minute and 22 seconds, which was within our expectations due to the little amount of steps involved.

5.6.1.2 Voltage

Participant	Subtask 1	Subtask 2	Subtask 3	Subtask 4
88303	05:08	05:23	08:10	00:20
14982	06:05	06:35	07:41	01:00
11144	07:17	03:00	06:31	01:40
99754	08:44	03:03	08:38	00:44
94595	09:56	03:12	05:21	01:01
Mean	07:26	04:15	07:16	00:57
Std.Dev.	01:56	01:39	01:20	00:29

Table 5.2: Timings for the Voltage Group

We mention again for clarity that these timings are not to be interpreted as realistic timings for these tasks due to the overhead introduced by the thinking-aloud process. Additionally, subtask 2 included the download of a 25mb setup file, which added ∼10 seconds for some participants, while the internet connection was fast enough for others, to be instant.

Subtask 1 involved reading the encrypted email using Voltage's Zero Download Messenger (ZDM), requiring the creation of an account before being allowed to read the email. This task took a mean time of 7 minutes and 26 seconds.

Subtask 2 encompassed downloading and executing the setup file installing the Voltage encryption software and took a mean time of 4 minutes and 15 seconds.

Subtask 3 asked the user to send an encrypted email to the psychologist using the Outlook integration and took a mean time of 7 minutes and 16 seconds.

Subtask 4 required the user to receive an encrypted email and to read it. It took a mean time of 57 seconds for participants to complete this. Since task 3 imported the user's identity into the Voltage Encryption Manager, decryption would happen automatically given the message is not marked as spam. Variability in the participants is explained (in addition to reading-speed) through whether they marked the Voltage email address (do_not_reply@voltage.com) as non-spam. If not, they had to move it to the inbox whereas the others could automatically read the message.

5.6.2 SUS scores

Following the task itself, the participants were asked to fill in the System Usability Questions. We will first present the SUS score for PostGuard, additionally mapping them to the adjective rating scale [16].

5.6.2.1 PostGuard

Participant	**SUS Score**
29311	35.0
65570	90.0
30850	52.5
73720	67.5
27601	42.5
Mean	57.5
Std. Dev.	21.86

Table 5.3: SUS Scores for the PostGuard group

In table 5.3 we can see that PostGuard received a mean score of 57.5, which maps to the adjective "Ok", the grade letter "F" and a low marginal acceptability score [13].

5.6.2.2 Voltage

Participant	**SUS Score**
88303	77.5
14982	80.0
11144	90.0
99754	47.5
94595	97.5
Mean	78.5
Std.Dev.	19.08

Table 5.4: SUS Scores for the Voltage group

Table 5.4 presents the mean of Voltage's SUS score being 78.5, which maps to the adjective "Good", the grade letter C, and an acceptable acceptability score [13].

5.6.3 Usability Problems and Software Issues

We will now describe any usability problems and software issues identified through our observations and notes taken during the tests. We will begin with PostGuard and follow with Voltage.

5.6.3.1 PostGuard

PostGuard, still being in its alpha stage, was expected to have a software issues and usability problems, which we did encounter during the test. We will detail them in the following paragraphs, beginning with the usability problems followed by the software issues.

The first usability problem most participants encountered was upon the display of the QR-Code on the fallback page (`postguard.eu/decrypt/`) after selecting the encrypted attachment file (see figure 4.2). Participants typically arrived at the decryption page as being the first link they followed from the email content, without having read the information about IRMA. This, paired with that no information on IRMA is given in the modal window (see figure 4.3), resulted in confusion as what to do next.

The second observed usability problem is more related to IRMA than PostGuard, but is still worth mentioning nonetheless. The problem occurs after entering the email address to be loaded as an attribute in the IRMA app and consequently clicking the link in the email that is received. The link is labeled "Load the email address attribute" and upon clicking and loading the page in the browser, displays a QR-code to be scanned with the app. The issue is that the user's phone is still open with the last instructions required in order to finish loading the attribute, with the last step (no. 4) prompting, "Go through the steps that are offered" (see figure 5.1). However, no steps are actually offered in either the email body or the on the QR-code popup modal. The user has to deduct themselves that they have to return to the home screen of the IRMA app and notice the "Scan QR" button on the bottom right.

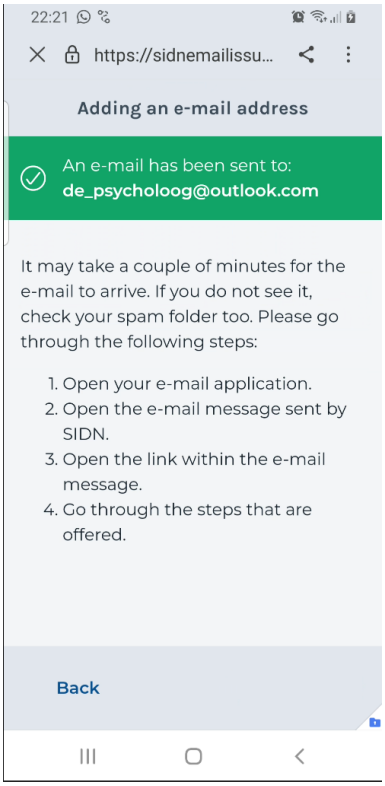

Figure 5.1: The IRMA app instructions after entering an email address to be loaded as an attribute.

The third usability problem is that the fallback website does not break long sentences to the next line within the email body display box (see figure 4.4). Instead, it widens the page, requiring the user to scroll horizontally to read the sentence.

The fourth usability problem encountered is regarding the installation instructions of the add-in at `postguard.eu/install_instructions.html`. Almost every participant (4; 80%) started out by following the instructions for Thunderbird, downloading the .xpi file and subsequently finding out that they cannot follow the instruction.

The fifth usability problem entails the discovery of the "Get Add-ins" button in Outlook used to manage add-ins. Two participants had trouble finding the button in the midst of all other buttons on the ribbon.

The sixth and last usability problem we observed pertains to the interpreted function of the "Send encrypted Email" button when composing a new email. A few participants (3; 60%) voiced that they were unsure whether the button would open a new email composing window, or whether it would send the

current composing email. Additionally, 2 participants (40%) erroneously sent the email unencrypted.

Then, we found two software issues while running the task. The first software issue is set up when choosing a file at the fallback page at `postguard.eu/decrypt/`, viewing the QR-code and then dismissing the modal by clicking on the X icon. The bug then occurs when choosing yet another file, without refreshing the page: upon selection of a file the second time, nothing happens.

The second software issue is a more intricate one, which occurred during the first participant evaluating PostGuard, resulting in the halting of the test due to the researcher's inability to fix it on the spot. Upon clicking the "Send Encrypted Email" button, a message is displayed saying "PostGuard outlook add-in is working on your Send encrypted Email request." and stays there indefinitely. We postulated that it was due to the caching of logged-in Microsoft accounts by Windows itself and attempted to circumvent this mechanism by creating a separate user account for each subsequent participant testing PostGuard. This seemed to properly bypass the mechanism, until at one instance, during the setup of the separate account, the researcher forgot to uncheck the option "Use this account for all Microsoft applications" while activating Outlook. Subsequent creation of another user account and ensuring the checkbox was turned off resulted in the add-in working properly again.

5.6.3.2 Voltage SecureMail

The participants also encountered a few usability problems with Voltage and additionally found one software issue. We will detail them in that order respectively below.

The first problem encountered was that participants entered the password of their Outlook email address on the account creation page instead of generating a new one. This indicates that they were not aware of the purpose of the page and could be dangerous depending on how Voltage SecureMail stores passwords. We do acknowledge however, that this problem could be due to the participants partaking in a test and not being an actual real life scenario.

The second usability problem happens after installation of the Outlook plugin. Part of the installation process involves closing the Outlook application and all browser applications. This resulted in the participants losing track of what they were doing or forgetting what they were supposed to do. We acknowledge here that this problem might not be relevant in a real scenario, since the closing is only a problem because the test depends on the instructions in the decrypted email.

Additionally, participants attempted to view a demo video presented by Voltage in their Quick Start Guide, however, the page hosting the video did not show a video.

Finally every email originating from Voltage domains ended up in the spam/junk folder. Not all participants moved the email to the inbox before reading, but attempted to digest the email without any formatting while still being marked as spam. Eventually, the participants attempted to click on the links and received a warning to first move the message to the inbox, which they all did.

Following the usability problems we found one software issue: after installation of the plugin and upon restart of Outlook, Outlook would sometimes disable the freshly installed plugin, citing the reason being that it was too slow. The user has to manually re-activate the plugin again within the Outlook settings.

5.6.4 Interview results

The interview allowed the participants to elaborate on any problems they had encountered and provided us with qualitative insights. We will present those insights now, beginning with PostGuard and followed by Voltage.

5.6.4.1 PostGuard

The most noteworthy assessment through the interviews was that nearly all participants reported that the instructions were unclear (4; 80%). It even prompted one participant to say, "I'm under the impression that I accomplished the task through sheer luck.", another participant mentioned "It [the instructions] could have been clearer, for instance by means of a concrete step-by-step plan, instead of just 'Download the app'". Those participants subsequently all mentioned to improve on the instructions by giving a clear-cut step-by-step plan. One also mentioned to use screenshots so that users can look for visual cues as where to click.

More specifically, the participants mentioned to have the most difficulties with IRMA: "The beginning went smoothly, but it became unclear when I had to use the IRMA app.", "The IRMA app can be made more simple, there is a lot going on the screen", and "The IRMA app did not automatically go back to the main screen so that I could open the QR-scanner.".

To the question of whether they would use PostGuard themselves, one answered, "No, because I don't send sensitive information", implying they would if they did handle sensitive information. Another participant responded with "Yes 100%, since I work in healthcare and handle a lot of personal data.". The remaining three participants reported that they would

not use it personally, but would use it in a business setting.

When asked about whether they could explain how PostGuard works, most participants outlined the structure of the actions they took, but none of them mentioned the term end-to-end encryption or IBE, which are (in our opinion) the most vital aspects of PostGuard.

Lastly, when inquired about whether they trusted PostGuard or not, all participants responded positively. Three participants stated their reasoning being due to the aim of the product: surely the product most be safe since it involves security. A quote from one of the participants is: "Yes, seeing how it involves encryption, I assume that it has to be safe.".

5.6.4.2 Voltage

Participants mentioned that the overall design did not look very trustworthy: "The interface seemed a bit sketchy, a bit outdated, which made me automatically suspect that it was not trustworthy.". Two participants (40%) made this remark explicitly.

Additionally, one participant recommended to display instructions right after installing the plugin, since the setup closes the Outlook plugin and the browser. It would have helped guide them towards use of the plug-in without having to navigate to the documentation page again.

Furthermore, two participants recommended making the button more clear in usage. As one participant answered to the question what they would change: "Make the button more clear". Related to this issue another participant mentioned that they were unsure whether the email was sent secure successfully and would have liked a confirmation of any kind.

When asked whether they would use Voltage in their daily lives, four out of five participants reported to be willing provided they had a reason. This is in line with the remaining participant answering no: "No, not necessary since I don't send any sensitive information". The results to this question could thus be interpreted as that all participants would be willing to use it, if they saw a legitimate reason to do so.

On the question whether they could explain how Voltage works, no participant mentioned anything about the most crucial security benefit, end-to-end encryption, neither did they mention anything about IBE.

Lastly, it seemed that the participants all trusted the product, but all of them provided indirect reasons for doing so. Four out of five participants reported to trust the product since they trust their psychologist, and the other participant based their trust due to the software's aim being safety "Yes, because the software stands for safety."

5.6.5 Product reaction cards

At the very end of the test, we spread out the 118 product reaction cards (with Dutch translations, see section 5.4.1.6) and asked the participants to first pick as many cards they thought were relevant to the product they were evaluating. Upon completion, we then asked them to filter down (or pick more) cards to come to a count of exactly five cards. We then noted which cards these were and asked them to elaborate on their choice. See section 2.2.3.1 for more information.

We will first present a quantitative analysis of these results, whereafter we will discuss any remarkable reasons the participants gave as elaboration to the words.

5.6.5.1 Quantitative results

We first present a simple word count paired with the nature of the word (either positive or negative), for each product. We will use the polarity in order to calculate the positive/negative ratio of words picked, providing an indication of usability satisfaction. The Dutch translation is included within parenthesis behind each (english) word.

PostGuard		
Word	Count	Polarity
Trustworthy (Betrouwbaar)	2	+
Confusing (Verwarrend)	1	-
Useful (Nuttig)	2	+
Too technical (Te technisch)	1	-
Time-consuming (Tijdrovend)	1	-
Confident (Zelfverzekered)	1	+
High quality (Hoge kwaliteit)	1	+
Secure (Veilig)	3	+
Valuable (Waardevol)	2	+
Efficient (Efficient)	2	+
Difficult (Moeilijk)	1	-
Relevant (Relevant)	1	+
Integrated (Geintegreerd)	1	+
Empowering (Mogelijkheden gevend)	1	+
Essential (Essentieel)	1	+
Fast (Snel)	1	+
Effective (Effectief)	1	+
Hard to use (Moeilijk te gebruiken)	1	-
Businesslike (Zakelijk)	1	+

Table 5.5: Product Reaction Cards picked for PostGuard.

Voltage		
Word	Count	Polarity
Usable (Betrouwbaar)	1	+
Easy to use (Makkelijk te gebruiken)	3	+
Reliable (Degelijk)	1	+
Simplistic (Simplistisch)	1	+
Useful (Nuttig)	3	+
Relevant (Relevant)	1	+
Fast (Snel)	1	+
Clear (Duidelijk)	2	+
Professional (Professioneel)	1	+
Straightforward (Evident)	1	+
Secure (Veilig)	3	+
Familiar (Vertrouwd)	1	+
Frustrating (Frustrerend)	1	-
Trustworthy (Betrouwbaar)	2	+
Valuable (Waardevol)	1	+
Businesslike (Zakelijk)	1	+
Accessible (Toegankelijk)	1	+

Table 5.6: Product Reaction Cards picked for Voltage SecureMail.

Ratio of positive and negative cards		
Product	Positive words	Negative words
PostGuard	20 (80%)	5 (20%)
Voltage	24 (96%)	1 (4%)

Table 5.7: Ratio of positive and negative cards for PostGuard and Voltage

5.6.6 Qualitative results

We asked each participant to elaborate on their choices, eliciting comments that help gain a deeper understanding of how they feel about the product. We will first present the reasoning behind the negative cards, followed by a selection of positive cards which we think are remarkable.

Word	Elaboration
PostGuard	
Confusing (Verwarrend)	"Because you constantly have to go from one app to the other."
Too technical (Te technisch)	"You have to do a lot, especially elderly people will not understand this"
Time-consuming (Tijdrovend)	"The whole process takes too long, you want to be able to read your email quickly."
Difficult (Moeilijk)	"The installation process was clumsy. It feels unfinished and is too difficult for the average person."
Hard to use (Moeilijk te gebruiken)	"The installation is difficult to use, but the usage [of the add-in] itself was good."
Voltage	
Frustrating (Frustrerend)	"Because I did not expect certain things, like the email ending up in the junk folder."

Table 5.8: Elaborations to the negative cards picked.

Word	Elaboration
PostGuard	
Secure (Veilig)	"It facilitates sending data securely." "Self-explanatory since it involves secure emails" "This is what it [PostGuard] does, it enables sending data securely. I can't say anything about the software itself."
Trustworthy (Betrouwbaar)	"It looked trustworthy, and the software is about security." "It looks safe with the app. The IRMA mechanism works dependable [betrouwbaar]. Reminds me of banks."
Voltage	
Secure (Veilig)	"Self-explanatory, that's what it is all about." "The verification email feels safe." "Because it's about safety."
Trustworthy (Betrouwbaar)	"Because you can trust that your email is sent safely." "Because it is dependable [betrouwbaar] in usage."

Table 5.9: Elaborations to the positive cards picked.

5.7 Subdiscussion

This section presents a preliminary (sub)discussion of the results of each product, linking together the findings found across the various evaluation methods (timings, SUS, observations, interviews, etc.) in preparation for the overarching discussion in chapter 6. Furthermore, this section will also directly provide answers to our first three subquestions in the context of the user study.

We will first discuss the results of PostGuard, followed by those from Voltage before finally ending with any overlapping problems.

5.7.1 PostGuard

PostGuard's mean SUS score of 57.5 makes it a candidate for increased scrutiny and continued improvement, and should be judged to be marginal at best [13], and thus a closer inspection is warranted as to why this is the case, for which we will use the results from our other metrics. Upon observing the other metrics, it becomes apparant that it is highly probable that the bad impressions come from the complex setup phase entailing the setup of IRMA and scanning of the QR-Codes. This is reflected in the timing scores for subtask 1, which consistently took up more than 40% of the total time required for the tasks for each participant. Additionally, this is is corroborated by the interview results where the participants mentioned a lot of friction with the instructions (regarding PostGuard and IRMA) and the usage of IRMA itself. Furthermore, we can also see this again with the observed usability problems described in section 5.6.3, where 5 out of the 6 problems are related to the fallback website and add-in installation instructions. Then lastly, this is also reflected by the results of the Product Reaction Cards. The ratio of positive to negative words consists of 80:20, of which the elaborations depict a clear picture that the negative words were chosen because of the setup process: "The installation is difficult to use, but the usage [of the add-in] itself was good.". We therefore hypothesize that this low SUS score is a consequence of the difficulties encountered while using PostGuard for the first time in conjunction with IRMA, but in order to determine this accurately, an additional study would be required that isolates the setup phase from the recurring phase. If our hypothesis is true, then it would be clear that this problem is only relevant for novice users unfamiliar with the system and would not pose a problem for advanced users. However, in order to become an advanced user, users would first have to be willing to go (and be capable of going) through the setup, so it is vital this problem is addressed.

Another problem worth mentioning is the confusion about how the "Send encrypted Email" button works. Two participants erroneously sent an email,

intended to be encrypted, unencrypted.

Seeing how PostGuard's biggest issue is within the setup phase, and thus irrelevant for "normal" use, leaving only some ambiguity over the functionality of the "Send Encrypted Email" button as its only concrete problem, we conclude that PostGuard has a substantial amount of usability problems for novice users, but does not have any major issues for advanced users.

In addition to the usability problems, we also observed some interesting aspects towards the participants' stances regarding PostGuard and security in general. First of all, the participants voiced to only see need for encryption if they were actually handling sensitive data, which most of them claimed not to do so currently. Secondly, none of the participants mentioned either end-to-end encryption or IBE when asked how the product works, while these are key components of PostGuard. Lastly, participants seemed to source their trust for the product not through direct product qualities, but rather through indirect relations, such as that it was recommended by their psychologist or because the product operates in the field of security.

5.7.2 Voltage

When examining the results for Voltage, we see few major problems, which corroborates with the SUS score of 78.5 and a 96:4 ratio of positive versus negative words with the Product Reaction Cards. We will first describe the two major problems we found, before addressing the minor problems. Finally, we will end with a few related aspects regarding the users' stance towards the product.

The first major problem observed is that every email sent through Voltage ended up in the spam folder unless the address had explicitly been trusted. Secondly, while all participants noticed the button "Send secure" and used it correctly, two participants expressed that they would wanted to button to be more clear.

The other, more minor, problems entail informing the user about the mechanisms involved with email encryption. One of those situations occurs upon the prompt to create an account, which comes out of the blue without any briefing beforehand or during the process. Participants, unaware that they were creating a new account, initiated to enter the password of their Outlook email, which essentially provided Voltage their credentials. Another situation involves the setup of the Voltage Encryption software, where after installation no instructions are provided while Outlook itself is closed simultaneously with all browsers.

Lastly we were able to gain some insights on the participants mental model towards the product from the interview results. First of all, all participants

reported that they appreciated the security Voltage provides and would use it in their daily lives, but only if they would have a reason to do so (i.e. handle sensitive information). Second, they seemed to place trust in the product through recommendation or assumptions instead of facts about the product, such as that the product is inherently a security tool and thus should be trustable. Finally, none of the participants mentioned either end-to-end encryption or identity-based encryption when asked how the product worked.

5.7.3 Overlapping problems

When examining both products for similar problems, we concretely identify one problem regarding the user interface design: both products inadequately inform the user over the usage of the button used to send the encrypted mail, consequently, both products also do not prevent the user from accidentally sending an email insecurely.

Additional to concrete usability problems, we identified that participants from both groups sourced their trust from indirect arguments, such as through the recommendation of a trusted person or because the product deals in security. Furthermore, we also saw that all participants did appreciate the protection of their privacy and thus placed value in taking up usage of email security tools. However, even though they trusted the developers and the software, none of the participants prominently mentioned the essential factors of the products: end-to-end encryption or identity-based encryption.

Chapter 6

Discussion and conclusions

In this chapter we will discuss the results and present our conclusions, but first, we will reiterate our research questions for clarity.

In the introduction of this thesis we presented our main research question: **"What, if any, are the usability problems regarding email encryption tools using IBE?"**. In order to help answer this question, we formulated four subquestions:

1. What are the usability problems in PostGuard?
2. What are the usability problems in Voltage SecureMail?
3. Are there any overlapping usability problems when evaluating both PostGuard and Voltage SecureMail, that point towards the existence of general usability issues with IBE-based applications?
4. How can PostGuard's usability be improved?

We will first present our findings and conclusions, followed by our recommendations for PostGuard, thereby answering our research (sub)question(s). Afterwards, we will discuss the limitations pertaining to this study, before ending with our suggestions for future work.

6.1 Findings of our methods

With respect to the research question and subquestions, we applied two methods: 1) the heuristic evaluation, the application of heuristics on a product through an expert's perspective, and 2) a user study, an evaluation of users using the product.

We have discussed both methods individually in sections 4.3 and 5.7, and

will now first cross-examine them, before presenting our conclusions in conjunction with related work.

6.1.1 Cross-examining the findings of the two methods

A heuristic evaluation is in many ways very similar to a user study, for an expert evaluator in essence performs the same exploratory actions as a regular participant, except with the evaluator, through their experience, being considerably more vigilant than a participant. This can be seen in our results for the two methods, as they intersect in the major key findings.

The heuristic evaluation showed us that PostGuard contained various problems regarding providing proper and clear instructions involving the setup of the tool, but once set up correctly, would pose little problems for frequent use. This problem was prominently observed through the user study as well, as nearly all participants voiced frustrations about the instructions for setting up IRMA and scanning the QR-Codes. However, supporting the conclusion of the heuristic evaluation, not a single participant would change anything about the mechanics of the tool, noting only that they would want better instructions.

Furthermore, the heuristic evaluation of PostGuard revealed one major user interface problem related to the usage of the "Send encrypted Email" button, risking the user erroneously sending an unencrypted email. Again, this was also observed in the user study, as all participants had some degree of trouble with understanding the button on first use.

A similar observation can be made from the results of the two methods for Voltage. The heuristic evaluation revealed only one usability problem, about Voltage not preventing the user from unintentionally sending an email unencrypted. While the participants of the user study did not make this mistake explicitly, they did express uncertainty about how the usage of the button. Both results indicate that the software warrants clearer instructions regarding the use of the "Send secure" button.

The two methods revealing similar results for both products is promising, providing confidence in our findings. With the user study confirming the heuristic evaluation, the results of the evaluation can now easily be converted into an issue list for the developers of the tools, backed by the results of the user study.

6.1.2 Conclusions

We will now present our conclusions as to what usability problems were found regarding PostGuard and Voltage, after which we will give our thoughts on the usability problems regarding IBE in general. Then, we will compare

our SUS scores to the scores found in related work pertaining to non-IBE based tools. Lastly, we will discuss our findings in relation to reasons for the low adoption rate of e-mail encryption.

6.1.2.1 Usability of PostGuard

To answer our first subquestion, we will briefly describe our findings regarding usability problems in PostGuard.

PostGuard displayed a fair amount of problems, however, most of them entailed the instructions and information provided to the user rather than actual user interface problems. Specifically, the fallback website does not provide sufficiently detailed instructions, a problem also present in the email sent with an encrypted email. Furthermore, we only found one actual user interface problem, which entails the "Send Encrypted Email" button, which is easily overlooked and ambiguous in how it works.

6.1.2.2 Usability of Voltage

To answer our second subquestion, we will briefly describe our findings regarding usability problems in Voltage.

There are a few usability problems present regarding informing the users about why they were required to perform certain actions, however, the actions themselves were clear and intuitive and thus these problems can be seen as minor. Additionally, all emails coming from the Voltage domain were marked as spam, which is an inconvenience until the user marks the domain as trusted. Furthermore, Voltage also contained one actual user interface problem, also entailing their "Send Secure" button, which is easily overlooked, as well as ambiguous in how it works.

6.1.2.3 Usability of identity-based encryption

Looking at the results of the heuristic evaluations and the user study, we see that while they do contain some problems, all of them are application specific and cannot be accredited to the underlying IBE mechanisms.

While PostGuard's problems are related to the setup and usage of IRMA, it mainly involves providing better instructions rather than poor usability of the required actions themselves. Furthermore, in support of PostGuard's issues not being generalizable to IBE are the results showing that Voltage does not have these issues.

The only concrete usability problem both products have, entail the usage of the button to send an email encrypted, but this problem could very possible also present itself in encryption solutions not employing IBE. This answers our third research question.

We therefore conclude that there appears to be no generalizable usability problems regarding email encryption tools employing IBE, which answers our main research question.

6.1.2.4 Comparison of SUS scores to non-IBE based tools

We will now mention a few words comparing the SUS scores of PostGuard and Voltage with SUS scores from related work. To reiterate, in our study PostGuard received a mean SUS score of 57.5 and Voltage received a score of 78.5. Contrastingly, in the usability study comparing IRMASeal and PGP by Starren et al. [35], PGP received a score of 46.1, which indicates that IBE performs better than PGP. Additionally, IRMASeal received a mean SUS score of 85.3, contrasting our result of 57.5 for PostGuard. We find this to be within expectations, seeing how the IRMA app was preloaded on their researcher's phone, preloaded with the email attributes. Furthermore Ruoti et al. [5] analyzed the usability of a tool based on PGP, which received a SUS score of 34.5, again supporting the idea that IBE is more usable than PGP.

6.1.2.5 Reasons for the low adoption rate of email encryption

Whitten and Tygar's seminal paper [3] first hypothesized in 1999 that the low adoption rate of email encryption was due to poor usability. This notion was embraced, and several follow-up studies were conducted evaluating this [5], [7]. However, our study has found little usability problems for email encryption tools utilizing IBE, yet these tools do not seem to have gained widespread adoption, contradicting their hypothesis and at least indicates that the solution to the problem, in addition to usability, should also be sought elsewhere. This is supported by the study of Abu-Salma et al. [32] titled *Obstacles to the Adoption of Secure Communication Tools*, where one of their key qualitative insights is: "Usability is not the primary obstacle to adoption [32, p. 137]".

Furthermore, the post-test interviews for each product revealed that all users would use end-to-end encryption if they saw the need for it, but would not necessarily adopt it for their personal use. Concretely, many participants stated that they would use it in a business context as opposed to a private context, implying that business information is more sensitive than their private information. This indicates that the participants are at least aware of the risk of their privacy being invaded, but do not take action. Existing work by Renaud et al. [31] supports this finding. They explicitly looked for reasons for non-adoption other than usability and were able to confirm six additional reasons ranging from a natural progression beginning with awareness, to understanding to acting (refer to section 3.2 for more details). Their second explanation, stating "They are aware of the possibility of pri-

vacy violation of their emails, but do not take any action for a variety of different reasons, perhaps because it does not *concern* them [31, p. 246]" in particular, corroborates with our observation.

6.2 PostGuard recommendations

We will now present concrete recommendations for improvement for PostGuard, thereby answering our fourth research (sub)question. These recommendations are inherently tied to the results of the heuristic evaluation and the user study; they have been extracted from them, thereby also inherently leaving out minor details. Consequently, since we also have brevity in mind, we will present our major recommendations here, but kindly refer the reader to the results of each method in chapter 4 and section 5.6 for more details.

First of all, we recommend adding integrated instructions within Outlook upon installation of the add-in, for example by highlighting the two buttons added, thereby confirming the user has actually noticed them. In the same train of thought, a warning message[1] might be appropriate the first time a user sends an unencrypted email after installing the add-in.

Additionally, we recommend providing a complete step-by-step plan in the default decryption email (possibly by linking to a tutorial page) detailing all required actions, including the setup of IRMA. The aim is to provide the user a single source of truth for them to refer back to in case they get stuck, without having to navigate through various resources.

Then, we also recommend providing more information on the fallback website and in the QR-Code modal window, facilitating impatient users who clicked as soon as they saw the decrypt link.

The next recommendation is a minor one: add a table of contents to the webpage detailing the installation instructions for the various email clients available. Additionally, we recommend adding screenshots here of the buttons' they need to click. Visual cues are easier to follow then written instructions for the average user.

Lastly, we recommend PostGuard to explicitly educate curious users as to **why** email encryption is needed and **how** PostGuard works (in conjunction with IRMA) since as of now, there are no reasons given for a user already using email encryption to use PostGuard instead.

Then, while IRMA is not the focus of this study, since a lot of participants had trouble with IRMA we nonetheless want to express a recommendation. We recommend designing the mobile application so that the QR-Code

[1] The on-send feature for Outlook add-ins might be useful at `https://docs.microsoft.com/en-us/office/dev/add-ins/outlook/outlook-on-send-addins`.

scanner opens up on the home screen. Currently, the home screen centers an image, the text "Your data securely on your mobile", and a button labeled "Add more data", with the currently verified attributes displayed in a stack on the bottom. The only concrete actionable elements are the button, and the display of the possessed cards. We argue, however, that the most frequent use-case for the application is to verify possession of presented attributes, or more concretely, to scan a QR-Code. Thus, automatically opening up the camera first of all primes new users for this action, and furthermore enables advanced users to identify themselves more quickly. Subsequently, we reason that the actions of adding an attribute, or more general managing the attributes, is a secondary activity to be performed only once in a while. Moreover, we observed that if a user scans a QR-Code asking for an attribute not in their possession (concretely this was an email address), the app automatically prompts whether the user wants to add it, opening up the same user interface for when a user manually adds an attribute. As such, we argue that optimizing this "scanner-first" user flow would linearize the user flow, simplifying first-time use while also speeding up frequent use.

6.3 Limitations

In the following section, we will detail the various limitations regarding this research.

The first limitation is that while our research questions aims to find general usability problems regarding IBE, we only were able to evaluate two products. An analysis of a wider range of products could have possibly yielded better conclusive findings.

The second limitation is the amount of expert evaluators that conducted the heuristic evaluation. Research has shown that one individual evaluator only finds between 20 and 51% of all usability problems, however, aggregating the results of 3 to 5 evaluators provides an accurate image [17]. Unfortunately, with our research we only had access to 1 evaluator: the author of this paper. Thus, it would have been more optimal if we did have access to more evaluators.

Furthermore, in addition to the commonly present limitation of a small sample size also present in this research, the demographics of our user study consisted entirely of people younger than 39 years old. One could argue that email encryption is important for all people regardless of age, therefore making our sample population not entirely representative.

Then, another limitation is that the PostGuard application is not ready for the public yet, while Voltage is, which is reflected back in the results through the amount of problems associated with either application. Concretely, this

prevented users from acquiring the PostGuard add-in installation instructions through normal means (i.e. finding instructions on the website), and required them to specifically follow instructions through the email of the psychologist, which is not necessarily the normal flow a user goes through. Evidently, it would be best to evaluate the product from an end-user's perspective with the product actually finished. Nonetheless, conducting the research anyway has provided, in our opinion, a good amount of feedback for PostGuard to improve upon, thereby answering our fourth research subquestion.

6.3.1 Mistakes

This study has been conducted as a bachelor thesis, which aims are for the student to learn how to conduct proper research. As such, it is not surprising that we actually made a few mistakes, which we will detail now.

The biggest mistake was made while constructing the questions for the interviews with the user study. The questions were not constructed enough with the potential answers in mind, resulting in various questions being answered with the same information. For example, the first question after decrypting the first email from the psychologist was "How did this process go and why?", and prompted the participant to detail all problems they encountered. This made the more specific follow-up questions quite redundant, since the participant already addressed those issues in the first question. This mistake could have been prevented if we had first fully conducted the heuristic evaluation, before designing the user test, which would have given us better background for constructing the questions. In this instance, however, we began with the biggest obstacle first, the user test, before simultaneously conducting the heurstic evaluation with the actual user tests with the participants. Additionally, because the heuristic evaluation was conducted during the time we took the tests, the heuristic evaluations (or rather the evaluator) are bound to have been influenced by the impressions of the first few participants, affecting the honesty of the results.

Furthermore, there was a feeling on the part of the researcher to rush the tests due to guilt regarding asking so much free time from the participants, which was inflated by the clearly redundant questions as described in the paragraph above. While this mainly comes from the unavoidable nature of recruiting participants for a bachelor thesis, it still affects the results and deserves to be mentioned here.

Then, we noticed another thing about the interviews which we would handle differently in the future, which is regarding the forms used in the interviews. To be short, we simply allocated too few lines for the answers to be written on, resulting in the researcher instinctively rephrasing the answers to be as

concise as possible. It would have been better for the results if less rephrasing was done. Additionally, in hindsight, using pen and paper was not the most effective method time-wise, taking the answers by typing on the computer would have saved a lot of time due to writing speed, in addition to keeping rephrasing to a minimum.

6.4 Future work

We will now give our thoughts for future work, complementing of course the limitations and mistakes presented above, of which it is evident that future work should try to avoid.

This study specifically analyzed the usability of email encryption tools using IBE within an actual email client (Outlook), however, we found that there were few user interface situations present involving the tools for frequent use. Concretely, both Voltage and PostGuard only required the user to click a different button to send the email, and similarly for decryption. As such, we recommend that any future usability study of email encryption using IBE, should rather focus on the usability of the authentication mechanisms used (PostGuard: IRMA, Voltage: email address/password), than the integration of the tools within email clients itself. It would be interesting to see future work mapping out all the various authentication mechanisms used for a wider amount of tools and specifically evaluate those.

Furthermore, our conclusions indicated that usability is not the main problem for adoption of email encryption, but instead showed that participants simply saw no reason to use it. Specifically, the participants claimed that they, at that moment, did not handle data sensitive enough to legitimize the use of encryption; a claim worth investigating by itself. As such, we recommend that future work should also investigate users' motivations regarding encryption rather than focusing on the usability of the products alone.

Then, since our findings indicated that the main problems with PostGuard were with the setup concerning IRMA, we recommend future work to investigate other authentication methods for PostGuard as well.

Lastly, as already mentioned in the limitations, we only evaluated 2 products using IBE and as such would recommend future work to analyze bigger amount of products in order to find better generalizable conclusions.

6.5 Final words

With our research, we attempted to identify any usability problems related specifically to IBE. To that matter, we deployed two evaluation methods, a heuristic evaluation, and a user study of two products (PostGuard and

Voltage). We identified individual problems of each tool, after which we evaluated the overlapping problems in order to prune application specific problems. We found that IBE involves very few security steps (for PostGuard this is scanning a QR-Code, and for Voltage simply logging in), which implicitly has a positive impact on usability in general. We conclude that IBE has no concrete usability problems pertaining to the protocol itself, thereby making it a strong alternative to other, more conventional, email encryption methods.

References

[1] S. Radicati, "Email statistics report, 2019–2023," *The Radicati Group, Inc.*, 2019. [Online]. Available: `www.radicati.com/wp/wp-content/uploads/2018/12/Email-Statistics-Report-2019-2023-Executive-Summary.pdf`.

[2] C. Stransky, O. Wiese, V. Roth, Y. Acar, and S. Fahl, "27 years and 81 million opportunities later: Investigating the use of email encryption for an entire university," in *43rd IEEE Symp. on Security and Privacy*, May 2022. [Online]. Available: `https://publications.cispa.saarland/3601/`.

[3] A. Whitten and J. D. Tygar, "Why johnny can't encrypt: A usability evaluation of pgp 5.0.," in *USENIX Security Symp.*, vol. 348, 1999, pp. 169–184.

[4] S. Ruoti, N. Kim, B. Burgon, T. Van Der Horst, and K. Seamons, "Confused johnny: When automatic encryption leads to confusion and mistakes," in *Proc. of the 9th Symp. on Usable Privacy and Security*, 2013, pp. 1–12.

[5] S. Ruoti, J. Andersen, D. Zappala, and K. Seamons, "Why johnny still, still can't encrypt: Evaluating the usability of a modern pgp client," *arXiv preprint arXiv:1510.08555*, 2015.

[6] S. Ruoti *et al.*, ""we're on the same page": A usability study of secure email using pairs of novice users," in *Proc. of the 2016 CHI Conf. on Human Factors in Computing Systems*, 2016, pp. 4298–4308.

[7] S. Sheng, L. Broderick, C. A. Koranda, and J. J. Hyland, "Why johnny still can't encrypt: Evaluating the usability of email encryption software," in *Symp. On Usable Privacy and Security*, ACM, 2006, pp. 3–4.

[8] S. L. Garfinkel and R. C. Miller, "Johnny 2: A user test of key continuity management with s/mime and outlook express," in *Proc. of the Symp. on Usable privacy and security*, 2005, pp. 13–24.

[9] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Workshop on the theory and application of cryptographic techniques*, Springer, 1984, pp. 47–53.

[10] S. Kaushik, P. Ammann, D. Wijesekera, W. Winsborough, and R. Ritchey, "A policy driven approach to email services," in *Proc. of the 5th IEEE Int. Workshop on Policies for Distributed Systems and Networks*, 2004, pp. 169–178. DOI: 10.1109/POLICY.2004.1309163.

[11] J. Brooke, "Sus: A "quick and dirty" usability scale," in *Usability evaluation in industry*, Taylor and Francis, 1996, pp. 189–194.

[12] J. Brooke, "Sus: A retrospective," *Journal of usability studies*, vol. 8, no. 2, pp. 29–40, 2013.

[13] A. Bangor, P. T. Kortum, and J. T. Miller, "An empirical evaluation of the system usability scale," *Intl. Journal of Human–Computer Interaction*, vol. 24, no. 6, pp. 574–594, 2008.

[14] J. Sauro, *A practical guide to the system usability scale: Background, benchmarks & best practices*. Measuring Usability LLC, 2011.

[15] T. S. Tullis and J. N. Stetson, "A comparison of questionnaires for assessing website usability," in *Usability professional association conference*, Minneapolis, USA, vol. 1, 2004, pp. 1–12.

[16] A. Bangor, P. Kortum, and J. Miller, "Determining what individual sus scores mean: Adding an adjective rating scale," *Journal of usability studies*, vol. 4, no. 3, pp. 114–123, 2009.

[17] J. Nielsen and R. Molich, "Heuristic evaluation of user interfaces," in *Proc. of the SIGCHI Conf. on Human factors in computing systems*, 1990, pp. 249–256.

[18] R. Molich and J. Nielsen, "Improving a human-computer dialogue," *Communications of the ACM*, vol. 33, no. 3, pp. 338–348, 1990.

[19] J. Nielsen, *Ten usability heuristics*, 2005. [Online]. Available: http://designingwebinterfaces.com/6-tips-for-a-great-flex-ux-part-5.

[20] D. Norman, J. Miller, and A. Henderson, "What you see, some of what's in the future, and how we go about doing it: Hi at apple computer," in *Conference companion on Human factors in computing systems*, 1995, p. 155.

[21] P. Merholz, *Peter in conversation with don norman about ux & innovation*, Adaptive Path, 2007. [Online]. Available: https://web.archive.org/web/20190313233623/http://adaptivepath.org/ideas/e000862/.

[22] E. Law, V. Roto, A. P. Vermeeren, J. Kort, and M. Hassenzahl, "Towards a shared definition of user experience," in *CHI'08 extended abstracts on Human factors in computing systems*, 2008, pp. 2395–2398.

[23] E. L.-C. Law, V. Roto, M. Hassenzahl, A. P. Vermeeren, and J. Kort, "Understanding, scoping and defining user experience: A survey approach," in *Proc. of the SIGCHI conference on human factors in computing systems*, 2009, pp. 719–728.

[24] "Ergonomics of human-system interaction — part 210: Human-centred design for interactive systems," International Organization for Standardization, Geneva, CH, Standard, 2010.

[25] M. Hassenzahl, "User experience and experience design," in *The encyclopedia of human-computer interaction*, 2nd ed. The Interaction Design Foundation Aarhus, Denmark, 2013. [Online]. Available: `https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/user-experience-and-experience-design`.

[26] J. Benedek and T. Miner, "Measuring desirability: New methods for evaluating desirability in a usability lab setting," *Proceedings of Usability Professionals Association*, vol. 2003, no. 8-12, p. 57, 2002.

[27] J. Camenisch, "Specification of the identity mixer cryptographic library, version 2.3.4," IBM Research, Zurich, Tech. Rep., Feb. 2012. [Online]. Available: `https://dominoweb.draco.res.ibm.com/reports/rz3730_revised.pdf`.

[28] J. Camenisch and A. Lysyanskaya, "A signature scheme with efficient protocols," in *International Conference on Security in Communication Networks*, Springer, 2002, pp. 268–289.

[29] W. Bai, D. Kim, M. Namara, Y. Qian, P. G. Kelley, and M. L. Mazurek, "Balancing security and usability in encrypted email," *IEEE Internet Computing*, vol. 21, no. 3, pp. 30–38, 2017.

[30] S. Ruoti, J. Andersen, T. Monson, D. Zappala, and K. Seamons, "A comparative usability study of key management in secure email," in *14th Symp. on Usable Privacy and Security (SOUPS 2018)*, 2018, pp. 375–394.

[31] K. Renaud, M. Volkamer, and A. Renkema-Padmos, "Why doesn't jane protect her privacy?" In *International Symp. on Privacy Enhancing Technologies*, Springer, 2014, pp. 244–262.

[32] R. Abu-Salma, M. A. Sasse, J. Bonneau, A. Danilova, A. Naiakshina, and M. Smith, "Obstacles to the adoption of secure communication tools," in *2017 IEEE Symp. on Security and Privacy (SP)*, IEEE, 2017, pp. 137–153. DOI: `10.1109/SP.2017.65`.

[33] S. Dechand, A. Naiakshina, A. Danilova, and M. Smith, "In encryption we don't trust: The effect of end-to-end encryption to the masses on user perception," in *2019 IEEE European Symp. on Security and Privacy (EuroS&P)*, IEEE, 2019, pp. 401–415.

[34] A. Reuter, K. Boudaoud, M. Winckler, A. Abdelmaksoud, and W. Lemrazzeq, "Secure email-a usability study," in *Int. Conf. on Financial Cryptography and Data Security*, Springer, 2020, pp. 36–46.

[35] N. Starren, H. Schraffenberger, and B. Jacobs, "Johnny can encrypt? a usability study of irmaseal," 2022. [Online]. Available: `https://www.cs.ru.nl/bachelors-theses/2022/Niels_Starren___1020246___Johnny_can_Encrypt_-_A_Usability_Study_of_IRMAseal.pdf`.

[36] –, “System usability scale voor meten gebruiksvriendelijkheid,” usersense.nl. [Online]. Available: https://www.usersense.nl/usability-testing/system-usability-scale-sus.

Appendix A

Information letter

Informatiebrief

IBE USABILITY STUDY

Introductie

Wij vragen u om mee te doen aan een bachelor scriptie onderzoek. Meedoen is vrijwillig. Om mee te doen is uw schriftelijke toestemming nodig. Voordat u beslist of u wilt meedoen aan dit onderzoek, krijgt u uitleg over wat het onderzoek inhoudt. Lees deze informatie rustig door en vraag de onderzoeker uitleg als u vragen heeft.

Beschrijving en doel van het onderzoek

In dit onderzoek testen wij de gebruiksvriendelijkheid van een tool om beveiligde e-mails te versturen. Deze tool wordt momenteel ontwikkeld op de Radboud Universiteit. Naast de gebruiksvriendelijkheid van deze tool zelf, testen wij ook de gebruiksvriendelijkheid van een al bestaande oplossing voor beveiligde e-mails. De evaluaties van deze twee tools zullen dan met elkaar vergeleken worden om sterktes en zwaktes te identificeren. De resultaten van dit onderzoek zullen beschreven worden in de bachelor scriptie en zullen verder dienen als feedback bij de verdere ontwikkeling van de tool.

Wat wordt er van u verwacht?

In dit onderzoek gaat u in een sessie van ongeveer 60 minuten twee email tools gebruiken om een mail met gevoelige informatie veilig (versleuteld) te sturen. U wordt ook gevraagd om een versleutelde mail te ontsleutelen. Na het gebruik van beide tools wordt u verwacht enkele vragen te beantwoorden over uw ervaring met deze tools. U voert deze taken uit op een laptop van een onderzoeker. U gebruikt telkens een speciaal voor dit onderzoek ingerichte mailaccount, niet uw persoonlijke mailaccount. De gevoelige informatie wordt door de onderzoeker beschikbaar gesteld en is daadwerkelijk ook niet gevoelig.

Risico’s en ongemak

Wij verwachten geen risico’s of ongemakken. Het is goed mogelijk dat u niet in staat zult zijn om sommige taken binnen de tijdslimiet uit te voeren. Ook deze informatie is zeer waardevol voor ons onderzoek. Wij testen niet u, maar de gebruiksvriendelijkheid van de tools.

Welke gegevens worden er verzameld?

Wij vragen uw naam op het toestemmingsformulier, en tijdens het onderzoek geanonimiseerd uw leeftijdscategorie, nationaliteit, geslacht, educatieniveau, en ervaringen met gerelateerde concepten. Verder zullen wij tijdens het onderzoek van u vragen om hardop te denken. Hier zullen wij een audio opname van maken alsmede handgeschreven of getypte notities over uw handelingen en opmerkingen. Naast de audio opname zullen wij ook een beeldopname maken van de computer waar u de test op uit voert. Ook noteren wij of het gelukt was om de taken succesvol af te ronden binnen de tijdslimiet en zo ja, hoelang het duurde. Ten slotte wordt u gevraagd enkele vragen te beantwoorden over uw ervaringen met- en meningen over de tools.

Vrijwilligheid

U doet vrijwillig mee aan dit onderzoek. U kunt op elk moment tijdens het onderzoek uw deelname stopzetten en uw toestemming intrekken. U hoeft niet aan te geven waarom u stopt. Stoppen met de deelname heeft geen enkel gevolg voor u.

Wat gebeurt er met mijn gegevens?

De onderzoeksgegevens die we in dit onderzoek verzamelen, zullen door ons gebruikt worden voor de bachelor scriptie en voor de verdere ontwikkeling van de email tool. We publiceren de onderzoeksresultaten in de bachelor scriptie. Verder zullen we de gegevens en resultaten delen met de ontwikkelaars van de tool.

Persoonsgegevens die verzameld worden, blijven vertrouwelijk. Als we gegevens met andere onderzoekers delen, kunnen deze niet tot u herleid worden. Het door u ondertekende toestemmingsformulier zal gedurende 10 jaar na afronding van het onderzoek bewaard worden. Uw geanonimiseerde onderzoeksgegevens worden bewaard tot ten minste 10 jaar na het afronden van het onderzoek.

U kunt tot een maand na deelname ook uw onderzoeksgegevens en persoonsgegevens laten verwijderen. Dit kunt u doen door een mail te sturen naar REDACTED@REDACTED.REDACTED.

We bewaren alle onderzoeks- en persoonsgegevens op beveiligde wijze volgens de richtlijnen van de Radboud Universiteit.

Heeft u vragen over het onderzoek?

Als u vragen heeft of meer informatie over het onderzoek wilt hebben, kunt u contact opnemen via de contactgegevens onderaan deze brief. Ook kunt u de scriptiebegeleider (Hanna Schraffenberger, REDACTED@REDACTED.REDACTED) contacteren.

Toestemmingsverklaring

Als u aan dit onderzoek mee wilt doen, vragen we u de toestemmingsverklaring te ondertekenen. Door uw schriftelijke toestemming geeft u aan dat u de informatie heeft begrepen en instemt met deelname aan het onderzoek.

Met vriendelijke groet,
Quoc An Ha, REDACTED@REDACTED.REDACTED [REDACTED_PHONE_NUMBER]

Appendix B

Consent form

Toestemmingsverklaring

Voor deelname aan het wetenschappelijk onderzoek:
IBE USABILITY STUDY

Vink de juiste vakjes aan	***Ja***	***Nee***
Ik bevestig dat ik naar behoren ben ingelicht over het onderzoek, zowel schriftelijk als mondeling.	☐	☐
Ik bevestig dat ik de informatiebrief ibe-study_informatiebrief.pdf heb gelezen.	☐	☐
Ik bevestig dat ik de mogelijkheid heb gehad om vragen te stellen en mijn vragen naar behoren zijn beantwoord.	☐	☐
Ik bevestig dat ik ruimschoots de kans heb gekregen om goed na te denken over deelname aan deze studie.	☐	☐
Ik bevestig dat ik vrijwillig deelneem aan deze studie.	☐	☐
Ik begrijp dat ik het recht heb mijn toestemming tijdens het onderzoek in te trekken zonder opgave van redenen en zonder vrees voor nadelige gevolgen.	☐	☐
Ik begrijp dat ik het recht heb om mijn toestemming in te trekken en daarmee mijn persoonsgegevens en onderzoeksgegevens te laten wissen tot 1 maand na afloop van het onderzoek door contact op te nemen met Quoc An Ha op [REDACTED_PHONE_NUMBER] of REDACTED@REDACTED.REDACTED.	☐	☐
Ik begrijp dat mijn persoonsgegevens worden verwerkt in overeenstemming met de privacyverklaring van de Radboud Universiteit (https://www.ru.nl/english/vaste-onderdelen/privacy-statement-radboud-university/).	☐	☐
Ik ben akkoord dat mijn persoonlijke en/of onderzoeksgegevens in het kader van dit onderzoek zullen worden verkregen voor wetenschappelijke doeleinden en gedurende 10 jaar bewaard zullen blijven.	☐	☐
Ik ben akkoord dat de getekende toestemmingsverklaring voor 10 jaar bewaard zal worden.	☐	☐
Ik ben akkoord dat mijn persoonsgegevens, die voor administratieve redenen zijn gekoppeld aan mijn onderzoeksgegevens, maximaal tot 1 maand na het afronden van het onderzoek bewaard mogen blijven.	☐	☐
Ik ben akkoord dat toezichthoudende autoriteiten mijn persoons- en onderzoeksgegevens mogen inzien met het oog op controle van het onderzoek.	☐	☐
Ik geef toestemming voor het verwerken van de volgende (bijzondere) persoonsgegevens over mij: leeftijdscategorie, geslacht, nationaliteit, educatieniveau en ervaringen met gerelateerde concepten.	☐	☐
Ik stem in dat tijdens het experiment video-opnamen van het computerscherm gemaakt worden. Deze opnamen worden gebruikt voor de analyse van de gebruiksvriendelijkheid van de software in het onderzoek door Quoc An Ha.	☐	☐
Ik stem in dat tijdens het experiment audio-opnamen gemaakt worden. Deze opnamen worden gebruikt voor de analyse van het gedachteproces van de deelnemer tijdens het onderzoek door Quoc An Ha.	☐	☐
Ik stem in dat verkregen, mogelijk identificeerbare, video/audio opnames openbaar gemaakt worden in de context van mijn onderzoek en/of om de verbetering van veilige email software te bevorderen.	☐	☐
Ik ben minimaal 18 jaar oud.	☐	☐

Ik begrijp dat ik, om aan het onderzoek te mogen deelnemen, **"Ja"** moet beantwoorden op alle bovenstaande punten.

Ik ga akkoord met deelname aan deze studie.

Naam:

Datum:

Handtekening:

In te vullen door de onderzoeker:

Ik, de ondergetekende, verklaar dat de hierboven genoemde persoon geschreven en gesproken is geïnformeerd over het bovengenoemde onderzoek.
Naam:

Positie, onderzoeksinstelling:

Datum:

Handtekening:

Appendix C

Participant form

IBE USABILITY STUDY

Deelnemersformulier

Juli 2022
Onderzoeker: Quoc An Ha
Gerandomiseerd deelnemersnummer:

Introductie onderzoek

Hoe is dit onderzoek gestructureerd?

Allereerst zullen we controleren of u de bijbehorende informatiebrief en toestemmingsverklaring gelezen heeft. Om het onderzoek door te laten gaan, is het een voorwaarde dat u het eens bent met de toestemmingsverklaring en deze getekend heeft. **We willen benadrukken dat u vrij bent om deze niet te tekenen als u dat niet wilt, hiervoor hoeft u ook geen reden te geven.**

Vervolgens vangen we aan met het onderzoek. Deze bestaat uit een pre-taak vragenlijst, gevolgd door een introductie tot het concept van hardop nadenken en ook tot het emailprogramma Outlook. Daarna voert u de taak daadwerkelijk uit. We sluiten vervolgens af met een post-taak procedure. Deze omvat een enquête, een interview, en een evaluatiemethode genaamd “Product Reaction Cards”.

Het onderzoek duurt ongeveer een uur. U kunt stoppen wanneer u wilt zonder opgaaf van reden.

Informatiebrief

Als u de informatiebrief over dit onderzoek nog niet gelezen heeft, doet u dat dan nu. Vraag de onderzoeker naar deze brief als u deze niet heeft.

Toestemmingsverklaring

Heeft u de toestemmingsverklaring al gelezen en getekend? Zo niet en wilt u dat wel doen? Vraag de onderzoeker dan hiernaar.

Kennismaking hardop denken

Tijdens het onderzoek zullen we u vragen hardop te denken. Om dit te demonstreren zullen we u een instructievideo laten zien.

Kennismaking Outlook

Voor het onderzoek is het nodig dat u bekend bent met het Outlook emailprogramma. Stuurt u alstublieft een email naar REDACTED@REDACTED.REDACTED met daarin een begroeting, terwijl u het hardop denken concept toepast.

Introductie scenario

Het is een onstuimige periode wat ertoe heeft geleid dat u sinds kort bij een psycholoog, Jane Doe, loopt. Tijdens jullie laatste gesprek heeft mevrouw Doe benadrukt dat u haar altijd kunt emailen voor acute problemen. Om haar veilig vertrouwelijke informatie te kunnen emailen, moeten emails echter wel versleuteld zijn. Ze heeft beloofd om u een email te sturen vanuit haar emailadres "REDACTED@REDACTED.REDACTED", die uitlegt hoe dit werkt.

Taak instructie

Lees de email van de psycholoog en handel naar eigen inzicht. We vragen u om tijdens het onderzoek zo veel mogelijk hardop te denken, zodat we inzicht in uw gedachteproces krijgen. De onderzoeker zal u mogelijk op bepaalde punten onderbreken om wat vragen te stellen. U bent klaar zodra de psycholoog u een email heeft gestuurd met dat het proces goed opgezet is.

Gegevensoverzicht

Hieronder staat informatie die u mogelijk nodig heeft.

Datum naam	Waarde	Beschrijving
Naam psycholoog	Jane Doe	De naam van uw psycholoog
Emailadres		Uw gesimuleerde eigen emailadres tijdens deze test
Emailadres wachtwoord		Het wachtwoord van uw emailadres
Voornaam	Alex	Uw voornaam tijdens de test
Achternaam	Smith	Uw achternaam tijdens de test
Datum afspraak	16 Augustus 2022	De datum van uw volgende afspraak met uw psycholoog
Tijdstip afspraak	14:30	Het tijdstip van uw volgende afspraak met uw psycholoog

Appendix D

Participant questions forms

D.1 PostGuard

IBE USABILITY STUDY

Vragenformulier

Juli 2022
Onderzoeker: Quoc An Ha
Gerandomiseerd deelnemersnummer:
Groep A

PRE-TEST QUESTIONNAIRE

Welke van de volgende opties beschrijft uw leeftijd het beste?

☐ 18 tot 25 ☐ 26 tot 39 ☐ 40 tot 59 ☐ 60 tot 74 ☐ 75 en ouder

Met welk geslacht identificeert u zich?

☐ Vrouw ☐ Man ☐ Anders

☐ Wil ik niet zeggen

Wat is uw nationaliteit?

_____

☐ Wil ik niet zeggen

Wat is uw hoogst genoten (afgeronde) opleiding?

☐ Middelbare school ☐ MBO ☐ HBO

☐ WO Bachelor ☐ WO Master

☐ Anders, namelijk: _____

☐ Wil ik niet zeggen

Heeft u een achtergrond in de informatietechnologie of andere gerelateerde gebieden?

☐ Ja ☐ Nee

☐ Wil ik niet zeggen

Op een schaal van 1 tot 5, hoe geïnteresseerd bent u in privacy en veiligheid?

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

☐ Wil ik niet zeggen

Op een schaal van 1 tot 5, hoeveel ervaring heeft u met Windows?

1 = helemaal geen ervaring, 5 = heel veel ervaring

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

☐ Wil ik niet zeggen

Op een schaal van 1 tot 5, hoeveel ervaring heeft u met Android?

1 = helemaal geen ervaring, 5 = heel veel ervaring

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

☐ Wil ik niet zeggen

Hoe vaak heeft u gewerkt met de Outlook emailapplicatie?

1 = nooit, 2 = een aantal keer, 3 = jaarlijks, 4 = wekelijks, 5 = dagelijks

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

☐ Wil ik niet zeggen

Hoe vaak gebruikt u gemiddeld email?

☐ Een aantal keer per dag

☐ Een aantal keer per week

☐ Een aantal keer per maand

☐ Een aantal keer per jaar

☐ Nooit

Als u emails verstuurt, verstuurt u deze dan via de computer of mobiel? Meerdere vakjes mogelijk.

☐ Computer ☐ Mobiel

Welke email applicaties gebruikt u?

_____

☐ Wil ik niet zeggen

Heeft u ooit gebruik gemaakt van de IRMA app, gemaakt door de stichting Privacy by Design?

☐ Ja ☐ Nee

Heeft u ooit gebruik gemaakt van Voltage SecureMail?

☐ Ja ☐ Nee

Weet u wat de term "encryptie" inhoudt?

☐ Ja, namelijk:

_____

_____

☐ Nee

POST-POSTGUARD VRAGEN

QUESTIONNAIRE

Op een schaal van 1 tot 5, hoe soepel verliep deze taak voor u?

1 = helemaal niet soepel, 5 = heel erg soepel

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Op een schaal van 1 tot 5, hoeveel frustratie had u tijdens deze taak?

1 = helemaal geen frustratie, 2 = heel veel frustratie,

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

SYSTEM USABILITY SCALE

DIGITAL EQUIPMENT CORPORATION, 1986.

Kruist u alstublieft het vak aan wat het beste uw reactie op de uitdrukkingen beschrijft. Met “product” bedoelen wij de PostGuard software.

Uitdrukking	Helemaal mee oneens 1	2	3	4	Helemaal mee eens 5
1. Ik denk dat ik dit product frequent zou willen gebruiken.					
2. Ik vond het onnodig ingewikkeld.					
3. Ik vond het product makkelijk te gebruiken.					
4. Ik denk dat ik technische support nodig heb om het product te gebruiken.					
5. Ik vond de verschillende functies van het product goed met elkaar geïntegreerd.					
6. Ik vond dat er te veel tegenstrijdigheden in het product zaten.					
7. Ik kan me voorstellen dat de meeste mensen snel met het product overweg kunnen.					
8. Ik vond het product omslachtig in gebruik.					
9. Ik voelde me zelfverzekerd tijdens het gebruik van het product.					
10. Ik moest veel over het product leren voordat ik het goed kon gebruiken.					

D.2 Voltage

IBE USABILITY STUDY

Vragenformulier

Juli 2022
Onderzoeker: Quoc An Ha
Gerandomiseerd deelnemersnummer:
Groep B

PRE-TEST QUESTIONNAIRE

Welke van de volgende opties beschrijft uw leeftijd het beste?

☐ 18 tot 25 ☐ 26 tot 39 ☐ 40 tot 59 ☐ 60 tot 74 ☐ 75 en ouder

Met welk geslacht identificeert u zich?

☐ Vrouw ☐ Man ☐ Anders

☐ Wil ik niet zeggen

Wat is uw nationaliteit?

_____

☐ Wil ik niet zeggen

Wat is uw hoogst genoten (afgeronde) opleiding?

☐ Middelbare school ☐ MBO ☐ HBO

☐ WO Bachelor ☐ WO Master

☐ Anders, namelijk: _____

☐ Wil ik niet zeggen

Heeft u een achtergrond in de informatietechnologie of andere gerelateerde gebieden?

☐ Ja ☐ Nee

☐ Wil ik niet zeggen

Op een schaal van 1 tot 5, hoe geïnteresseerd bent u in privacy en veiligheid?

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

☐ Wil ik niet zeggen

Op een schaal van 1 tot 5, hoeveel ervaring heeft u met Windows?

1 = helemaal geen ervaring, 5 = heel veel ervaring

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

☐ Wil ik niet zeggen

Op een schaal van 1 tot 5, hoeveel ervaring heeft u met Android?

1 = helemaal geen ervaring, 5 = heel veel ervaring

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

☐ Wil ik niet zeggen

Hoe vaak heeft u gewerkt met de Outlook emailapplicatie?

1 = nooit, 2 = een aantal keer, 3 = jaarlijks, 4 = wekelijks, 5 = dagelijks

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

☐ Wil ik niet zeggen

Hoe vaak gebruikt u gemiddeld email?

☐ Een aantal keer per dag

☐ Een aantal keer per week

☐ Een aantal keer per maand

☐ Een aantal keer per jaar

☐ Nooit

Als u emails verstuurt, verstuurt u deze dan via de computer of mobiel? Meerdere vakjes mogelijk.

☐ Computer ☐ Mobiel

Welke email applicaties gebruikt u?

_____

☐ Wil ik niet zeggen

Heeft u ooit gebruik gemaakt van de IRMA app, gemaakt door de stichting Privacy by Design?

☐ Ja ☐ Nee

Heeft u ooit gebruik gemaakt van Voltage SecureMail?

☐ Ja ☐ Nee

Weet u wat de term "encryptie" inhoudt?

☐ Ja, namelijk:

_____

_____

☐ Nee

POST-VOLTAGE VRAGEN

QUESTIONNAIRE

Op een schaal van 1 tot 5, hoe soepel verliep deze taak voor u?

1 = helemaal niet soepel, 5 = heel erg soepel

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Op een schaal van 1 tot 5, hoeveel frustratie had u tijdens deze taak?

1 = helemaal geen frustratie, 2 = heel veel frustratie,

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

SYSTEM USABILITY SCALE: VOLTAGE

DIGITAL EQUIPMENT CORPORATION, 1986.

Kruist u alstublieft het vak aan wat het beste uw reactie op de uitdrukkingen beschrijft. Met “product” bedoelen wij de Voltage SecureMail software.

Uitdrukking	Helemaal mee oneens 1	2	3	4	Helemaal mee eens 5
1. Ik denk dat ik dit product frequent zou willen gebruiken.					
2. Ik vond het onnodig ingewikkeld.					
3. Ik vond het product makkelijk te gebruiken.					
4. Ik denk dat ik technische support nodig heb om het product te gebruiken.					
5. Ik vond de verschillende functies van het product goed met elkaar geïntegreerd.					
6. Ik vond dat er te veel tegenstrijdigheden in het product zaten.					
7. Ik kan me voorstellen dat de meeste mensen snel met het product overweg kunnen.					
8. Ik vond het product omslachtig in gebruik.					
9. Ik voelde me zelfverzekerd tijdens het gebruik van het product.					
10. Ik moest veel over het product leren voordat ik het goed kon gebruiken.					

Appendix E

Moderator forms

E.1 PostGuard

IBE USABILITY STUDY

Moderatorformulier

Juli 2022
Onderzoeker: Quoc An Ha
Gerandomiseerde deelnemersnummer:
Groep A

Voorbereidingen thuis

Materialen

Zorg ervoor dat al deze materialen present en beschikbaar zijn tijdens het experiment:

- ❑ Laptop #1: voor de deelnemer om het experiment op uit te voeren.
- ❑ Laptop #2: voor de onderzoeker om de psycholoog te kunnen simuleren.
- ❑ Telefoon: voor de deelnemer om IRMA op te gebruiken.
- ❑ Stekkerdoos met ten minste 5 aansluitingen.
- ❑ Verse moderator formulieren (nog niet ingevuld).
- ❑ Verse deelnemersformulieren (nog niet ingevuld).
- ❑ Toestemmingsverklaring formulier.
- ❑ Informatiebrief.
- ❑ Emailadres #1: REDACTED@REDACTED.REDACTED
- ❑ Emailadres #2: REDACTED@REDACTED.REDACTED
- ❑ Drie pennen.
- ❑ Een fles water.
- ❑ Een cadeautje voor de deelnemer.

Voorbereidingen

Wees er zeker van om de materialen als volgt van tevoren (thuis) voor te bereiden om de deelnemer niet onnodig te laten wachten:

- ❑ Laptop #1 (deelnemer):
  - ❑ Is volledig opgeladen.
  - ❑ Heeft de Outlook applicatie geïnstalleerd.
    - ❑ Is geconfigureerd met emailadres #2 voor de deelnemer.
    - ❑ Maak een nieuw profiel aan (File > Info > Account Settings > Manage Profiles > Show Profiles… > Add… > deelnemer# > Email account (vul gegevens in) > Next > Finish > Selecteer profiel voor startup > Klik Apply. Vergeet niet dit te **testen** door Outlook te herstarten.
  - ❑ Heeft de applicatie ShareX geïnstalleerd ten behoeve van de schermopname en audio opname.
    - ❑ Heeft ten minste 10GB aan schijfruimte beschikbaar voor de opnames.
  - ❑ Heeft **géén** Voltage plug-in geïnstalleerd.
  - ❑ Heeft een **lege** Download map.
  - ❑ Firefox autofill uitgeschakeld.
- ❑ Laptop #2 (onderzoeker):
  - ❑ Is volledig opgeladen.
  - ❑ Toegang tot REDACTED@REDACTED.REDACTED
  - ❑ Outlook applicatie geïnstalleerd.
  - ❑ Outlook geladen met REDACTED@REDACTED.REDACTED
  - ❑ PostGuard gekoppeld aan REDACTED@REDACTED.REDACTED
  - ❑ Voltage SecureMail geïnstalleerd.
- ❑ Telefoon:
  - ❑ Is volledig opgeladen.
  - ❑ Heeft geen pincode nodig om het scherm te openen.
  - ❑ Verwijder autosuggesties
  - ❑ Heeft **géén** IRMA geïnstalleerd.

- ❑ Emailadres #1 (psycholoog):
  - ❑ Voeg een draft toe voor openingsmail Voltage.
    - ❑ Ontvanger ingevuld.
    - ❑ Onderwerp ingevuld.
    - ❑ Inhoud ingevuld.
  - ❑ Voeg een draft toe voor openingsmail PostGuard.
    - ❑ Ontvanger ingevuld.
    - ❑ Onderwerp ingevuld.
    - ❑ Inhoud ingevuld.
- ❑ Emailadres #2 (deelnemer):
  - ❑ Is **niet** geregistreerd bij Voltage SecureMail.
  - ❑ Heeft een lege inbox.
  - ❑ Heeft **géén** koppeling met de Outlook PostGuard add-in.
- ❑ Deelnemersinformatie (vul deze gegevens in op de moderator- en deelnemersformulieren):
  - ❑ Bepaal het deelnemersnummer als een willekeurig nummer tussen (exclusief) 10.000 en 100.000. Controleer vervolgens of deze niet al in gebruik is genomen.
  - ❑ Bepaal of de deelnemer in groep A of groep B zit. Groep A behandelt PostGuard. Groep B behandelt Voltage. De onderzoeker moet afwisselen per deelnemer: de eerste deelnemer zit in groep A, de tweede in groep B, de derde in groep A, enzovoort.

Voorbereiding op-locatie

Deze sectie beschrijft de voorbereiding ter plekke:

- ❑ Zet laptop #1 op een tafel, doe deze aan de oplader, zet deze aan en log op Windows in:
  - ❑ Open het Outlook programma.
  - ❑ Open het ShareX programma.
- ❑ Leg de telefoon naast de laptops.
- ❑ Zet de fles water naast de laptops.
- ❑ Zet laptop #2 op een tafel, doe deze aan de oplader en log in op het bestuurssysteem.
  - ❑ Open Outlook (web) met de inbox van REDACTED@REDACTED.REDACTED
- ❑ Pak een pen en leg deze samen met de observator en moderator formulieren naast laptop #2.
- ❑ Verstuur openingsmail respectievelijk voor PostGuard of Voltage SecureMail.

Pre-test script

Allereerst wil ik u welkomen en bedanken voor uw deelname aan dit onderzoek.

Voordat we beginnen, heb ik wat informatie voor u, die ik voor zal lezen zodat ik niks mis.

Het doel van dit onderzoek is om de gebruiksvriendelijkheid van email applicaties te testen die specifiek ontwikkeld zijn voor veilige communicatie. Ik wil daarbij benadrukken dat we deze applicaties testen, en niet u. We zijn ook uiterst benieuwd naar wat u denkt dat beter zal kunnen.

Om het duidelijk voor u te houden, hebben we de belangrijkste informatie voor u op papier gezet. Ik vraag u om voor nu alleen het kopje "Introductie onderzoek" te lezen.

- Overhandig deelnemersformulier.

Om er zeker van te zijn dat het duidelijk is zal ik de informatie even herhalen. Onderbreekt u mij op elk moment wanneer u een vraag heeft.

We beginnen zo met de informatiebrief en de toestemmingsverklaring. Deze zijn nodig voordat het onderzoek plaats mag nemen. Daarna volgt een pre-test vragenlijst om vast te stellen wat uw achtergrond is en ervaringen zijn. Vervolgens introduceren wij u met het concept van hardop denken en de email applicatie Outlook, welke we beide zullen gebruiken tijdens het onderzoek.

Dan zijn we toe aan de daadwerkelijke taak voor u om uit te voeren. Nadere instructies zullen op dat moment volgen. Na de taak hebben we sluiten we af met een vragenlijst, een interview en product reactie kaarten. Dat laatste zullen we dan ook nader toelichten. In totaal zal het onderzoek ongeveer een uur duren. Ik wil nogmaals benadrukken dat u kunt stoppen wanneer u wilt zonder opgaaf van reden.

Wij zullen notities maken aan de hand van een schermopname op de computer waar u de taak op uit zult voeren. Ook nemen wij het geluid op en zullen we notities maken op pen en papier ofwel typend op de computer.

Ik wil u nu vragen de informatiebrief en toestemmingsverklaring te lezen als u dit nog niet gedaan heeft. Als u instemt, wilt u dan alstublieft de toestemmingsverklaring ondertekenen?

- Overhandig de informatiebrief en toestemmingsverklaring
- Wacht tot de deelnemer deze gelezen en ondertekent heeft.
- Onderteken het formulier zelf ook.

Heeft u tot nu toe vragen?

Oké, geweldig. Dan wil ik u nu vragen om de pre-test vragenlijst in te vullen.

- Overhandig de pre-test vragenlijst samen met een pen en wacht tot deze is ingevuld.
- Open alvast het filmpje voor het hardop denken.
- https://www.youtube.com/watch?v=BwpPIiBK0cA

Top! Dan zijn we nu bij de kennismaking met hardop denken en Outlook aangekomen. Allereerst zal ik u vertellen wat het concept van hardop denken inhoudt.

Het concept van hardop denken is zoals het klinkt, heel erg simpel. Het omvat namelijk simpelweg het letterlijk uitspreken wat u denkt. Dit proces maakt het makkelijker voor ons om uw gedachtegang te volgen, en daarmee ook hoe u de informatie u tot u neemt. Dit geeft ons inzicht in **waar** een gebruiker vast kan lopen, en **waarom** een gebruiker vastloopt.

Natuurlijk is geen enkele gedachte is fout, en daarmee dus ook enkele uitspraak. Wees dus vooral niet bang om iets verkeerds te zeggen. Het moeilijke zal zijn om u te herinneren dat u hardop moet denken, daarom zullen wij u hierop attenderen zodra u stilvalt.

Ter demonstratie zullen we nu een filmpje bekijken van een man die een appel in stukken snijdt.

- Speel het filmpje af.

Oké top, begrijpt u het concept van hardop denken nu?

- Adresseer enige vragen.

Geweldig, dan zijn we nu aangekomen tot de kennismaking met Outlook. Voor het onderzoek is het nodig dat u weet hoe u een email verzendt met dit programma. Zou u hiertoe alstublieft een test email kunnen versturen naar REDACTED@REDACTED.REDACTED terwijl u hardop denkt?

- Begeleid de gebruiker met het versturen van een email terwijl ze hardop denken.

Top, dan zijn we nu aangekomen tot de daadwerkelijke taak. In dit onderzoek vragen we u om u in te leven in een persoon die email contact heeft met hun psycholoog. Daarbij speel ikzelf de rol van de psycholoog. We vragen u om alles zo zelfstandig mogelijk uit te voeren. In het uiterste geval kunt u wel vragen stellen, maar juist omdat we ook benieuwd zijn naar wat gebruikers doen als ze vastlopen, kan het zijn dat ik niet altijd zal antwoorden. U mag wel gewoon het internet zelf gebruiken alsof u thuis op uw computer zit. Ik wil u bij deze taak eraan herinneren om hardop te denken, zoals we net besproken hebben. Verder zullen we op bepaalde momenten u onderbreken om wat vragen te stellen. Heeft u vragen hierover tot nu toe?

Oké top, dan stel ik voor dat we nu samen de introductie tot het scenario lezen en het gegevensoverzicht doornemen.

- Wijs de deelnemer op het informatieblad op het deelnemersformulier.

Introductie scenario

Het is een onstuimige periode wat ertoe heeft geleid dat u sinds kort bij een psycholoog, Jane Doe, loopt. Tijdens jullie laatste gesprek heeft mevrouw Doe benadrukt dat u haar altijd kunt emailen voor acute problemen. Om haar veilig vertrouwelijke informatie te kunnen emailen, moeten emails echter wel versleuteld zijn. Ze heeft beloofd om u een email te sturen vanuit haar emailadres "REDACTED@REDACTED.REDACTED", die uitlegt hoe dit werkt.

Taak instructie

Lees de email van de psycholoog en handel naar eigen inzicht. We vragen u om tijdens het onderzoek zo veel mogelijk hardop te denken, zodat we inzicht in uw gedachteproces krijgen. De onderzoeker zal u mogelijk op bepaalde punten onderbreken om wat vragen te stellen. U bent klaar zodra de psycholoog u een email heeft gestuurd met dat het proces goed opgezet is.

Gegevensoverzicht

Hieronder staat informatie die u mogelijk nodig heeft.

Heeft u vragen hierover?

Oké, dan hebben we nu alle benodigde informatie behandelt. Bent u klaar om het onderzoek te starten?

- Verstuur de mail vanuit de psycholoog.
- Begin schermopname via ShareX.
- Volg het script verder via de interviews. Onderbreek de deelnemer na elke taak! Maak aantekeningen. Script voor onderbreken: “Goed gedaan! Ik wil u graag nu even kort onderbreken.”

HULP PROCEDURE

Bij de volgende scenario's is het van belang om de meest realistische omgeving te creëren. Dat wil zeggen met zo min mogelijk hulp van de onderzoeker. Van de hulp die gegeven wordt, is het cruciaal dat deze consistent is voor alle deelnemers, zodat de data tussen de deelnemers vergeleken kan worden. Hieronder wordt beschreven in welk geval hulp gegeven kan worden, en op wat voor manier.

POSTGUARD

Situatie	Hulp
De deelnemer loopt vast op het QR-code scherm omdat ze nog geen IRMA gedownload hebben.	Geef geen hulp. De informatie is wel te vinden, maar ze moeten terug gaan naar de decryption instructions van PostGuard.
De deelnemer begrijpt niet dat ze hun emailattribuut moeten laden.	Geef ze pas hulp als ze de informatie op http://irma.app hebben geopend en gelezen. Vertel ze: "Om PostGuard gebruiken, moet je IRMA app geladen zijn met het emailadres attribuut.
Deelnemer kan de "Get add-ins" knop niet vinden in Outlook	Wijs ze op het icoontje na 5 minuten.
De deelnemer kan de "Send encrypted email" of "Decrypt Email" knop niet vinden.	Wijs ze na 5 minuten op de het icoontje (op de toolbar, of verschuilt onder de submenu met drie punten).
De deelnemer verstuurt een onversleutelde email.	Reageer met het account van de psycholoog met de opmerking dat die onversleuteld was en met het verzoek om het opnieuw te proberen.

Scenario takenoverzicht

PostGuard

1) Ontsleutelen aan de hand van de browser.
2) Add-in/plug-in installeren en verzenden versleutelde email.
3) Ontvangen en ontsleutelen versleutelde email.

Instructie onderzoeker

- Help de deelnemer zo min mogelijk, als de deelnemer echt vastloopt, geef dan hulp aan de hand van het opgestelde protocol.
- Herinner om de deelnemer te stimuleren om hardop te denken zodra ze een tijdje stilvallen.
- Let goed op wanneer de taken voltooid zijn en onderbreek de deelnemer.
    - Stel de interview vragen voor die taak.
    - Neem altijd de vrijheid om over dingen door te vragen die je gehoord hebt tijdens het hardop denken.

Scenario: PostGuard

INTERVIEW

Herinner: Neem altijd de vrijheid over dingen door te vragen die je gehoord hebt tijdens het hardop denken.

Taak 1

U heeft nu IRMA opgezet en aan de hand van de browser de mail ontsleuteld. De volgende vragen gaan daar over:

- Hoe verliep dit proces en waarom?

  _____

- Leek de instructie email van PostGuard betrouwbaar voor u?

  _____

- Waren deze instructies duidelijk voor u? Wat was er duidelijk of onduidelijk?

  _____

- Waarvoor diende u de QR-code te scannen met uw telefoon?

  _____

- Kunt u in uw eigen woorden uitleggen wat het doel van de IRMA applicatie is?

  _____

- Heeft u problemen ervaren tijdens het ontsleutelen? Zo ja, wat waren deze?

  _____

- Wat zou volgens u verbeterd kunnen worden om de gebruiksvriendelijkheid te verhogen?

  _____

- In het geval dat u hulp nodig had, was u zelf doorgegaan als u niet deelnam aan een onderzoek?

  _____

- Heeft u verder nog opmerkingen die u met ons wilt delen?

  _____

We zijn nu klaar met de vragen. U kunt nu verdergaan. U heeft de email van uw psycholoog zojuist ontsleuteld en bent klaar om deze te lezen.

Taak 2

U heeft nu de PostGuard add-in geïnstalleerd en een versleutelde email verstuurd. De volgende vragen gaan daar over:

- Hoe verliep dit proces en waarom?

  _____

- Vond u de installatie instructies van de add-in duidelijk en goed te volgen? Zo nee, wat precies?

  _____

- Heeft u problemen ervaren terwijl u de add-in installeerde? Zo ja, wat waren deze?

  _____

- Wat vond u specifiek van het proces om de add-in te installeren?

  _____

- Wat zou volgens u verbeterd kunnen worden om de gebruiksvriendelijkheid te verhogen?

  _____

- Heeft u problemen ervaren terwijl u de versleutelde email verstuurde? Zo ja, wat waren deze?

  _____

- Was het duidelijk voor u hoe u een versleutelde email verstuurd tegenover een niet versleutelde email?

  _____

- Wat gebeurt er als u op de “normale” verzendknop had geklikt?

  _____

- Kunnen de ontwikkelaars van PostGuard uw email lezen?

  _____

- Kunnen de ontwikkelaars van PostGuard emails als u verzenden?

  _____

- In het geval dat u hulp nodig had, was u zelf doorgegaan als u niet deelnam aan een onderzoek?

  _____

We zijn nu klaar met de vragen. U kunt nu verdergaan. U was aan het wachten op de reactie van uw psycholoog.

Taak 3

U heeft nu de reactie van uw psycholoog ontsleuteld en gelezen. De volgende vragen gaan daar over:

- Wat verliep prettig of gemakkelijk tijdens het ontsleutelen van de email?

  _____

- Heeft u problemen ervaren terwijl u de email ontsleutelde? Zo ja, wat waren deze?

  _____

- Waarom moest u de QR-code scannen?

  _____

- Hoe herkent u dat een email versleuteld verzonden is?

  _____

- In het geval dat u hulp nodig had, was u zelf doorgegaan als u niet deelnam aan een onderzoek?

  _____

We zijn nu klaar met dit onderdeel van het onderzoek. Als volgt gaan we nu de post-test procedure uitvoeren, wat bestaat uit een vragenlijst, interview en product reaction cards.

Post-test PostGuard

Hier nu allereerst de vragenlijst.

- Overhandig de vragenlijst van het deelnemers vragenformulier.

Top, dankuwel! Dan wil ik u nu vragen om de system usability scale in te vullen, weer een vragenlijst.

- Overhandig de SUS vragenlijst.

Top, dankuwel! Dan nu het interview over het product.

- Vond u PostGuard fijn om te gebruiken? Zo ja, wat was er fijn aan? Zo nee, wat was er niet fijn aan?

  _____

- Zou u uit kunnen leggen hoe PostGuard werkt?

  _____

- Is PostGuard te vertrouwen denkt u? Waarom wel/niet?

  _____

- Zou u PostGuard gebruiken in uw dagelijks leven? Waarom wel/niet?

  _____

- Zou u PostGuard aanraden aan anderen? Zo ja, met wat voor reden? Zo nee, waarom niet?

  _____

- Ziet u zelf een praktische toepassing voor versleutelde emails in uw dagelijks leven? Zo ja, waarvoor? En wat zou dan de toegevoegde waarde zijn?

  _____

- Is er iets wat u zou veranderen aan hoe PostGuard nu werkt?

  _____

We zijn nu klaar met het interview. Als laatst hebben we nog de product reaction cards. Geef me alstublieft een moment om deze op te zetten.

PRODUCT REACTIE KAARTEN

- Verspreid willekeurig de product reactie kaarten.

Hier voor u heb ik 108 kaarten verspreid. Ik wil u vragen om degene te kiezen die naar uw mening het beste PostGuard/Voltage SecureMail beschrijven. U kunt er zoveel kiezen als dat u wilt.

- Wacht tot ze een selectie gemaakt hebben.
- Als er meer of minder dan exact vijf kaarten gekozen zijn, vraag ze dan om een selectie van vijf te maken die ze het best passend vinden.
- Vraag ze de kaarten te overhandigen.
- Noteer de gekozen kaarten hieronder en vraag ze hun keuze toe te lichten.

1. _____

   _____

   _____

2. _____

   _____

   _____

3. _____

   _____

   _____

4. _____

   _____

   _____

5. _____

   _____

   _____

Hartelijk bedankt, we zijn nu klaar! Heeft u nog enige vragen of opmerkingen?

_____

_____

_____

Nogmaals bedankt voor uw deelname!

APPENDIX

Email sjablonen

Openings email PostGuard

Beste Alex Smith,

Zoals beloofd stuur ik u hierbij een email met informatie over het versleutelen van emails.

U zei Outlook te gebruiken, wat toevallig goed werkt met de software waar wij op de praktijk mee werken, genaamd PostGuard. Informatie over PostGuard is hier (https://www.postguard.eu/) te vinden.

Als u deze mail ontsleuteld heeft via uw browser, dan moet u nog de add-in installeren voor Outlook, zodat dat binnen in het programma kan. U kunt hier (https://www.postguard.eu/install_instructions.html) de handleiding vinden.

Ik stel voor dat we dit gelijk testen, zodat u mij direct kunt mailen wanneer dat nodig is. Zou u mij alstublieft een versleutelde email kunnen sturen aan de hand van de add-in met daarin de datum van onze volgende afspraak? Ik zal die email versleuteld beantwoorden met het tijdstip, die u controleert. Zo garanderen we dat u ook versleutelde emails kunt ontvangen, alsmede verzenden.

Met vriendelijke groet,
Jane Doe

Openings email Voltage

Beste Alex Smith,

Zoals beloofd stuur ik u hierbij een email met informatie over het versleutelen van emails.

U zei Outlook te gebruiken, wat toevallig goed werkt met de software waar wij op de praktijk mee werken, genaamd Voltage. Informatie over Voltage is hier (https://www.microfocus.com/en-us/cyberres/data-privacy-protection/secure-mail) te vinden.

Als u deze mail ontsleuteld heeft via uw browser, dan moet u nog de plug-in installeren voor Outlook, zodat dat binnen in het programma kan. U kunt hier (https://voltage.com/products/email-security/micro-focus-securemail-cloud/quickstart/#show_client) de handleiding vinden.

Ik stel voor dat we dit gelijk testen, zodat u mij direct kunt mailen wanneer dat nodig is. Zou u mij alstublieft een versleutelde email kunnen sturen aan de hand van de plug-in met daarin de datum van onze volgende afspraak? Ik zal die email versleuteld beantwoorden met het tijdstip, die u controleert. Zo garanderen we dat u ook versleutelde emails kunt ontvangen, alsmede verzenden.

Met vriendelijke groet,
Jane Doe

Tijdstip antwoord

Beste Alex Smith,

Dankuwel voor de datum, ik heb het gecontroleerd en het klopt. Door de versleuteling kan ik erop vertrouwen dat dit emailadres ook echt bij u hoort en dat u deze mail geschreven heeft.

Het tijdstip van onze volgende afspraak is om 14:30.

We hebben nu het proces omtrent beveiligde emails versturen helemaal doorlopen; we zijn klaar!

Met vriendelijke groet,
Jane Doe

E.2 Voltage

IBE USABILITY STUDY

Moderatorformulier

Juli 2022
Onderzoeker: Quoc An Ha
Gerandomiseerde deelnemersnummer:
Groep B

Voorbereidingen thuis

Materialen

Zorg ervoor dat al deze materialen present en beschikbaar zijn tijdens het experiment:

- ❑ Laptop #1: voor de deelnemer om het experiment op uit te voeren.
- ❑ Laptop #2: voor de onderzoeker om de psycholoog te kunnen simuleren.
- ❑ Telefoon: voor de deelnemer om IRMA op te gebruiken.
- ❑ Stekkerdoos met ten minste 5 aansluitingen.
- ❑ Verse moderator formulieren (nog niet ingevuld).
- ❑ Verse deelnemersformulieren (nog niet ingevuld).
- ❑ Toestemmingsverklaring formulier.
- ❑ Informatiebrief.
- ❑ Emailadres #1: REDACTED@REDACTED.REDACTED
- ❑ Emailadres #2: REDACTED@REDACTED.REDACTED
- ❑ Drie pennen.
- ❑ Een fles water.
- ❑ Een cadeautje voor de deelnemer.

Voorbereidingen

Wees er zeker van om de materialen als volgt van tevoren (thuis) voor te bereiden om de deelnemer niet onnodig te laten wachten:

- ❑ Laptop #1 (deelnemer):
  - ❑ Is volledig opgeladen.
  - ❑ Heeft de Outlook applicatie geïnstalleerd.
    - ❑ Is geconfigureerd met emailadres #2 voor de deelnemer.
    - ❑ Maak een nieuw profiel aan (File > Info > Account Settings > Manage Profiles > Show Profiles… > Add… > deelnemer# > Email account (vul gegevens in) > Next > Finish > Selecteer profiel voor startup > Klik Apply. Vergeet niet dit te **testen** door Outlook te herstarten.
  - ❑ Heeft de applicatie ShareX geïnstalleerd ten behoeve van de schermopname en audio opname.
    - ❑ Heeft ten minste 10GB aan schijfruimte beschikbaar voor de opnames.
  - ❑ Heeft **géén** Voltage plug-in geïnstalleerd.
  - ❑ Heeft een **lege** Download map.
  - ❑ Firefox autofill uitgeschakeld.
- ❑ Laptop #2 (onderzoeker):
  - ❑ Is volledig opgeladen.
  - ❑ Toegang tot REDACTED@REDACTED.REDACTED
  - ❑ Outlook applicatie geïnstalleerd.
  - ❑ Outlook geladen met REDACTED@REDACTED.REDACTED
  - ❑ PostGuard gekoppeld aan REDACTED@REDACTED.REDACTED
  - ❑ Voltage SecureMail geïnstalleerd.
- ❑ Telefoon:
  - ❑ Is volledig opgeladen.
  - ❑ Heeft geen pincode nodig om het scherm te openen.
  - ❑ Verwijder autosuggesties
  - ❑ Heeft **géén** IRMA geïnstalleerd.

- ❑ Emailadres #1 (psycholoog):
  - ❑ Voeg een draft toe voor openingsmail Voltage.
    - ❑ Ontvanger ingevuld.
    - ❑ Onderwerp ingevuld.
    - ❑ Inhoud ingevuld.
  - ❑ Voeg een draft toe voor openingsmail PostGuard.
    - ❑ Ontvanger ingevuld.
    - ❑ Onderwerp ingevuld.
    - ❑ Inhoud ingevuld.
- ❑ Emailadres #2 (deelnemer):
  - ❑ Is **niet** geregistreerd bij Voltage SecureMail.
  - ❑ Heeft een lege inbox.
  - ❑ Heeft **géén** koppeling met de Outlook PostGuard add-in.
- ❑ Deelnemersinformatie (vul deze gegevens in op de moderator- en deelnemersformulieren):
  - ❑ Bepaal het deelnemersnummer als een willekeurig nummer tussen (exclusief) 10.000 en 100.000. Controleer vervolgens of deze niet al in gebruik is genomen.
  - ❑ Bepaal of de deelnemer in groep A of groep B zit. Groep A behandelt PostGuard. Groep B behandelt Voltage. De onderzoeker moet afwisselen per deelnemer: de eerste deelnemer zit in groep A, de tweede in groep B, de derde in groep A, enzovoort.

Voorbereiding op-locatie

Deze sectie beschrijft de voorbereiding ter plekke:

- ❑ Zet laptop #1 op een tafel, doe deze aan de oplader, zet deze aan en log op Windows in:
  - ❑ Open het Outlook programma.
  - ❑ Open het ShareX programma.
- ❑ Leg de telefoon naast de laptops.
- ❑ Zet de fles water naast de laptops.
- ❑ Zet laptop #2 op een tafel, doe deze aan de oplader en log in op het bestuurssysteem.
  - ❑ Open Outlook (web) met de inbox van REDACTED@REDACTED.REDACTED
- ❑ Pak een pen en leg deze samen met de observator en moderator formulieren naast laptop #2.
- ❑ Verstuur openingsmail respectievelijk voor PostGuard of Voltage SecureMail.

Pre-test script

Allereerst wil ik u welkomen en bedanken voor uw deelname aan dit onderzoek.

Voordat we beginnen, heb ik wat informatie voor u, die ik voor zal lezen zodat ik niks mis.

Het doel van dit onderzoek is om de gebruiksvriendelijkheid van email applicaties te testen die specifiek ontwikkeld zijn voor veilige communicatie. Ik wil daarbij benadrukken dat we deze applicaties testen, en niet u. We zijn ook uiterst benieuwd naar wat u denkt dat beter zal kunnen.

Om het duidelijk voor u te houden, hebben we de belangrijkste informatie voor u op papier gezet. Ik vraag u om voor nu alleen het kopje "Introductie onderzoek" te lezen.

- Overhandig deelnemersformulier.

Om er zeker van te zijn dat het duidelijk is zal ik de informatie even herhalen. Onderbreekt u mij op elk moment wanneer u een vraag heeft.

We beginnen zo met de informatiebrief en de toestemmingsverklaring. Deze zijn nodig voordat het onderzoek plaats mag nemen. Daarna volgt een pre-test vragenlijst om vast te stellen wat uw achtergrond is en ervaringen zijn. Vervolgens introduceren wij u met het concept van hardop denken en de email applicatie Outlook, welke we beide zullen gebruiken tijdens het onderzoek.

Dan zijn we toe aan de daadwerkelijke taak voor u om uit te voeren. Nadere instructies zullen op dat moment volgen. Na de taak hebben we sluiten we af met een vragenlijst, een interview en product reactie kaarten. Dat laatste zullen we dan ook nader toelichten. In totaal zal het onderzoek ongeveer een uur duren. Ik wil nogmaals benadrukken dat u kunt stoppen wanneer u wilt zonder opgaaf van reden.

Wij zullen notities maken aan de hand van een schermopname op de computer waar u de taak op uit zult voeren. Ook nemen wij het geluid op en zullen we notities maken op pen en papier ofwel typend op de computer.

Ik wil u nu vragen de informatiebrief en toestemmingsverklaring te lezen als u dit nog niet gedaan heeft. Als u instemt, wilt u dan alstublieft de toestemmingsverklaring ondertekenen?

- Overhandig de informatiebrief en toestemmingsverklaring
- Wacht tot de deelnemer deze gelezen en ondertekent heeft.
- Onderteken het formulier zelf ook.

Heeft u tot nu toe vragen?

Oké, geweldig. Dan wil ik u nu vragen om de pre-test vragenlijst in te vullen.

- Overhandig de pre-test vragenlijst samen met een pen en wacht tot deze is ingevuld.
- Open alvast het filmpje voor het hardop denken.
- https://www.youtube.com/watch?v=BwpPIiBK0cA

Top! Dan zijn we nu bij de kennismaking met hardop denken en Outlook aangekomen. Allereerst zal ik u vertellen wat het concept van hardop denken inhoudt.

Het concept van hardop denken is zoals het klinkt, heel erg simpel. Het omvat namelijk simpelweg het letterlijk uitspreken wat u denkt. Dit proces maakt het makkelijker voor ons om uw gedachtegang te volgen, en daarmee ook hoe u de informatie u tot u neemt. Dit geeft ons inzicht in **waar** een gebruiker vast kan lopen, en **waarom** een gebruiker vastloopt.

Natuurlijk is geen enkele gedachte is fout, en daarmee dus ook enkele uitspraak. Wees dus vooral niet bang om iets verkeerds te zeggen. Het moeilijke zal zijn om u te herinneren dat u hardop moet denken, daarom zullen wij u hierop attenderen zodra u stilvalt.

Ter demonstratie zullen we nu een filmpje bekijken van een man die een appel in stukken snijdt.

- Speel het filmpje af.

Oké top, begrijpt u het concept van hardop denken nu?

- Adresseer enige vragen.

Geweldig, dan zijn we nu aangekomen tot de kennismaking met Outlook. Voor het onderzoek is het nodig dat u weet hoe u een email verzendt met dit programma. Zou u hiertoe alstublieft een test email kunnen versturen naar REDACTED@REDACTED.REDACTED terwijl u hardop denkt?

- Begeleid de gebruiker met het versturen van een email terwijl ze hardop denken.

Top, dan zijn we nu aangekomen tot de daadwerkelijke taak. In dit onderzoek vragen we u om u in te leven in een persoon die email contact heeft met hun psycholoog. Daarbij speel ikzelf de rol van de psycholoog. We vragen u om alles zo zelfstandig mogelijk uit te voeren. In het uiterste geval kunt u wel vragen stellen, maar juist omdat we ook benieuwd zijn naar wat gebruikers doen als ze vastlopen, kan het zijn dat ik niet altijd zal antwoorden. U mag wel gewoon het internet zelf gebruiken alsof u thuis op uw computer zit. Ik wil u bij deze taak eraan herinneren om hardop te denken, zoals we net besproken hebben. Verder zullen we op bepaalde momenten u onderbreken om wat vragen te stellen. Heeft u vragen hierover tot nu toe?

Oké top, dan stel ik voor dat we nu samen de introductie tot het scenario lezen en het gegevensoverzicht doornemen.

- Wijs de deelnemer op het informatieblad op het deelnemersformulier.

Introductie scenario

Het is een onstuimige periode wat ertoe heeft geleid dat u sinds kort bij een psycholoog, Jane Doe, loopt. Tijdens jullie laatste gesprek heeft mevrouw Doe benadrukt dat u haar altijd kunt emailen voor acute problemen. Om haar veilig vertrouwelijke informatie te kunnen emailen, moeten emails echter wel versleuteld zijn. Ze heeft beloofd om u een email te sturen vanuit haar emailadres "REDACTED@REDACTED.REDACTED", die uitlegt hoe dit werkt.

Taak instructie

Lees de email van de psycholoog en handel naar eigen inzicht. We vragen u om tijdens het onderzoek zo veel mogelijk hardop te denken, zodat we inzicht in uw gedachteproces krijgen. De onderzoeker zal u mogelijk op bepaalde punten onderbreken om wat vragen te stellen. U bent klaar zodra de psycholoog u een email heeft gestuurd met dat het proces goed opgezet is.

Gegevensoverzicht

Hieronder staat informatie die u mogelijk nodig heeft.

Heeft u vragen hierover?

Oké, dan hebben we nu alle benodigde informatie behandelt. Bent u klaar om het onderzoek te starten?

- Verstuur de mail vanuit de psycholoog.
- Begin schermopname via ShareX.
- Volg het script verder via de interviews. Onderbreek de deelnemer na elke taak! Maak aantekeningen. Script voor onderbreken: “Goed gedaan! Ik wil u graag nu even kort onderbreken.”

HULP PROCEDURE

Bij de volgende scenario's is het van belang om de meest realistische omgeving te creëren. Dat wil zeggen met zo min mogelijk hulp van de onderzoeker. Van de hulp die gegeven wordt, is het cruciaal dat deze consistent is voor alle deelnemers, zodat de data tussen de deelnemers vergeleken kan worden. Hieronder wordt beschreven in welk geval hulp gegeven kan worden, en op wat voor manier.

VOLTAGE

De deelnemer weet niet welke bit-versie van de plug-in nodig is.	Informeer ze gelijk dat de computer 64-bit is zodra ze dit vragen.
De deelnemer vraagt waar ze de Voltage plug-in moeten installeren.	Informeer ze gelijk met dat de standaard waarde voldoet.
De eerste keer openen van Outlook gaf een error met dat Voltage nog aan het opstarten was. Als er geen extra tab genaamd "Voltage" naast "Help" staat, dan moet Outlook herstart worden. Als deelnemer dit niet doet, dan kunnen ze de knop (Send Secure) niet vinden omdat die er niet is.	• Als ze desondanks een onversleutelde mail sturen, reageer dan als de psycholoog met dat die onversleuteld is en of ze het opnieuw willen proberen. • Als ze na 5 minuten onderzoeken het nog steeds niet gelukt is. Wijs ze op dat ze Outlook moeten herstarten.
De deelnemer verstuurd een onversleutelde email.	Reageer met het account van de psycholoog met de opmerking dat die onversleuteld was en met het verzoek om het opnieuw te proberen.

Scenario takenoverzicht

Voltage

1) Ontsleutelen aan de hand van de browser.
2) Add-in/plug-in installeren en verzenden versleutelde email.
3) Ontvangen en ontsleutelen versleutelde email.

Instructie onderzoeker

- Help de deelnemer zo min mogelijk, als de deelnemer echt vastloopt, geef dan hulp aan de hand van het opgestelde protocol.
- Herinner om de deelnemer te stimuleren om hardop te denken zodra ze een tijdje stilvallen.
- Let goed op wanneer de taken voltooid zijn en onderbreek de deelnemer.
    - Stel de interview vragen voor die taak.
    - Neem altijd de vrijheid om over dingen door te vragen die je gehoord hebt tijdens het hardop denken.

Voltage SecureMail

INTERVIEW

Herinner: Neem altijd de vrijheid over dingen door te vragen die je gehoord hebt tijdens de think-aloud.

Taak 1

U heeft nu een account voor Voltage SecureMail aangemaakt en aan de hand van de browser uw mail ontsleuteld. De volgende vragen gaan hierover:

- Hoe verliep dit proces en waarom?

  _____

- Leek de instructie email van Voltage SecureMail betrouwbaar voor u?

  _____

- Waren deze instructies duidelijk voor u? Wat was er duidelijk of onduidelijk?

  _____

- Heeft u problemen ervaren terwijl u de email ontsleutelde? Zo ja, wat waren deze?

  _____

- Wat zou volgens u verbeterd kunnen worden om de gebruiksvriendelijkheid te verhogen?

  _____

- Waarom moest u een account aanmaken?

  _____

- Waarom denkt u dat u nog een keer een email moest openen en daar van de link te volgen?

  _____

- In het geval dat u hulp nodig had, was u zelf doorgegaan als u niet deelnam aan een onderzoek?

  _____

We zijn nu klaar met de vragen. U kunt nu verdergaan. U heeft de email van uw psycholoog zojuist ontsleuteld en bent klaar om deze te lezen.

Taak 2

U heeft nu de Voltage plug-in geïnstalleerd en een versleutelde email verstuurd. De volgende vragen gaan daar over:

- Hoe verliep dit proces en waarom?

  _____

- Vond u de installatie instructies van de plug-in duidelijk en goed te volgen? Zo nee, wat precies?

  _____

- Heeft u problemen ervaren terwijl u de plug-in installeerde? Zo ja, wat waren deze?

  _____

- Wat vond u specifiek van het proces om de plug-in te installeren?

  _____

- Wat zou volgens u verbeterd kunnen worden om de gebruiksvriendelijkheid te verhogen?

  _____

- Heeft u problemen ervaren terwijl u de versleutelde email verstuurde? Zo ja, wat waren deze?

  _____

- Was het duidelijk voor u hoe u een versleutelde email verstuurd tegenover een niet versleutelde email?

  _____

- Wat gebeurt er als u op de “normale” verzendknop had geklikt?

  _____

- Kunnen de ontwikkelaars van Voltage SecureMail uw email lezen?

  _____

- Kunnen de ontwikkelaars van Voltage SecureMail emails als u verzenden?

  _____

- In het geval dat u hulp nodig had, was u zelf doorgegaan als u niet deelnam aan een onderzoek?

  _____

We zijn nu klaar met de vragen. U kunt nu verdergaan. U was aan het wachten op de reactie van uw psycholoog.

Taak 3

U heeft nu de reactie van uw psycholoog ontsleuteld en gelezen. De volgende vragen gaan daar over:

- Wat verliep prettig of gemakkelijk tijdens het ontsleutelen van de email?

  _____

- Hoe herkent u dat een email versleuteld verzonden was?

  _____

- In het geval dat u hulp nodig had, was u zelf doorgegaan als u niet deelnam aan een onderzoek?

  _____

We zijn nu klaar met dit onderdeel van het onderzoek. Als volgt gaan we nu de post-test procedure uitvoeren, wat bestaat uit een vragenlijst, interview en product reaction cards.

Post-test Voltage SecureMail

Hier nu allereerst de vragenlijst.

- Overhandig de vragenlijst van het deelnemers vragenformulier.

Top, dankuwel! Dan wil ik u nu vragen om de system usability scale in te vullen, weer een vragenlijst.

- Overhandig de SUS vragenlijst.

Top, dankuwel! Dan nu het interview over het product.

- Vond u Voltage fijn om te gebruiken? Zo ja, wat was er fijn aan? Zo nee, wat was er niet fijn aan?

  _____

- Zou u uit kunnen leggen hoe Voltage werkt?

  _____

- Is Voltage te vertrouwen denkt u? Waarom wel/niet?

  _____

- Zou u Voltage gebruiken in uw dagelijks leven? Waarom wel/niet?

  _____

- Zou u Voltage aanraden aan anderen? Zo ja, met wat voor reden? Zo nee, waarom niet?

  _____

- Ziet u zelf een praktische toepassing voor versleutelde emails in uw dagelijks leven? Zo ja, waarvoor? En wat zou dan de toegevoegde waarde zijn?

  _____

- Is er iets wat u zou veranderen aan hoe Voltage SecureMail nu werkt?

  _____

We zijn nu klaar met het interview. Als laatst hebben we nog de product reaction cards. Geef me alstublieft een moment om deze op te zetten.

PRODUCT REACTIE KAARTEN

- Verspreid willekeurig de product reactie kaarten.

Hier voor u heb ik 108 kaarten verspreid. Ik wil u vragen om degene te kiezen die naar uw mening het beste PostGuard/Voltage SecureMail beschrijven. U kunt er zoveel kiezen als dat u wilt.

- Wacht tot ze een selectie gemaakt hebben.
- Als er meer of minder dan exact vijf kaarten gekozen zijn, vraag ze dan om een selectie van vijf te maken die ze het best passend vinden.
- Vraag ze de kaarten te overhandigen.
- Noteer de gekozen kaarten hieronder en vraag ze hun keuze toe te lichten.

1. _____

   _____

   _____

2. _____

   _____

   _____

3. _____

   _____

   _____

4. _____

   _____

   _____

5. _____

   _____

   _____

Hartelijk bedankt, we zijn nu klaar! Heeft u nog enige vragen of opmerkingen?

_____

_____

_____

Nogmaals bedankt voor uw deelname!

APPENDIX

Email sjablonen

Openings email PostGuard

Beste Alex Smith,

Zoals beloofd stuur ik u hierbij een email met informatie over het versleutelen van emails.

U zei Outlook te gebruiken, wat toevallig goed werkt met de software waar wij op de praktijk mee werken, genaamd PostGuard. Informatie over PostGuard is hier (https://www.postguard.eu/) te vinden.

Als u deze mail ontsleuteld heeft via uw browser, dan moet u nog de add-in installeren voor Outlook, zodat dat binnen in het programma kan. U kunt hier (https://www.postguard.eu/install_instructions.html) de handleiding vinden.

Ik stel voor dat we dit gelijk testen, zodat u mij direct kunt mailen wanneer dat nodig is. Zou u mij alstublieft een versleutelde email kunnen sturen aan de hand van de add-in met daarin de datum van onze volgende afspraak? Ik zal die email versleuteld beantwoorden met het tijdstip, die u controleert. Zo garanderen we dat u ook versleutelde emails kunt ontvangen, alsmede verzenden.

Met vriendelijke groet,
Jane Doe

Openings email Voltage

Beste Alex Smith,

Zoals beloofd stuur ik u hierbij een email met informatie over het versleutelen van emails.

U zei Outlook te gebruiken, wat toevallig goed werkt met de software waar wij op de praktijk mee werken, genaamd Voltage. Informatie over Voltage is hier (https://www.microfocus.com/en-us/cyberres/data-privacy-protection/secure-mail) te vinden.

Als u deze mail ontsleuteld heeft via uw browser, dan moet u nog de plug-in installeren voor Outlook, zodat dat binnen in het programma kan. U kunt hier (https://voltage.com/products/email-security/micro-focus-securemail-cloud/quickstart/#show_client) de handleiding vinden.

Ik stel voor dat we dit gelijk testen, zodat u mij direct kunt mailen wanneer dat nodig is. Zou u mij alstublieft een versleutelde email kunnen sturen aan de hand van de plug-in met daarin de datum van onze volgende afspraak? Ik zal die email versleuteld beantwoorden met het tijdstip, die u controleert. Zo garanderen we dat u ook versleutelde emails kunt ontvangen, alsmede verzenden.

Met vriendelijke groet,
Jane Doe

Tijdstip antwoord

Beste Alex Smith,

Dankuwel voor de datum, ik heb het gecontroleerd en het klopt. Door de versleuteling kan ik erop vertrouwen dat dit emailadres ook echt bij u hoort en dat u deze mail geschreven heeft.

Het tijdstip van onze volgende afspraak is om 14:30.

We hebben nu het proces omtrent beveiligde emails versturen helemaal doorlopen; we zijn klaar!

Met vriendelijke groet,
Jane Doe

Appendix F

Product reaction cards

Simplistic *Simplistisch*	Inviting *Uitnodigend*	Clean *Schoon*	Irrelevant *Irrelevant*	Patronizing *Neerbuigend*
Not Valuable *Niet waardevol*	Approachable *Benaderbaar*	Dated *Verouderd*	Valuable *Waardevol*	Consistent *Consistent*
Boring *Saai*	Effortless *Moeiteloos*	Comprehensive *Uitgebreid*	Stable *Stabiel*	Easy to use *Makkelijk te gebruiken*
Motivating *Motiverend*	Compelling *Overtuigend*	Overbearing *Aanmatigend*	Disconnected *Onsamenhangend*	Satisfying *Bevredigend*
Organized *Georganiseerd*	Fragile *Breekbaar*	Accessible *Toegankelijk*	Confusing *Verwarrend*	Useful *Nuttig*
Fresh *Fris*	Creative *Creatief*	Relevant *Relevant*	Impressive *Indrukwekkend*	Ordinary *Gewoon*
Energetic *Energiek*	Not Secure *Niet veilig*	Low Maintenance *Weinig onderhoud*	Stimulating *Stimulerend*	Enthusiastic *Enthousiast*
Empowering *Mogelijkheiden gevend*	Unconventional *Onconventioneel*	Controllable *Te controleren*	Exceptional *Uitzonderlijk*	Predictable *Voorspelbaar*
Desirable *Wenselijk*	Comfortable *Comfortabel*	Impersonal *Onpersoonlijk*	Business-like *Zakelijk*	Convenient *Handig*
Effective *Effectief*	Difficult *Moeilijk*	Frustrating *Frustrerend*	Clear *Duidelijk*	Gets in the way *Staat in de weg*
Powerful *Krachtig*	Customizable *Aanpasbaar*	Hard to Use *Moeilijk te gebruiken*	Fast *Snel*	Stressful *Stressvol*
Time-Saving *Tijdbesparend*	Connected *Samenhangend*	Compatible *Bijpassend*	Calm *Rustig*	Undesirable *Ongewenst*

Attractive *Aantrekkelijk*	Efficient *Efficiënt*	Poor quality *Slechte kwaliteit*	Inconsistent *Inconsistent*	Uncontrollable *Onbestuurbaar*
Familiar *Vertrouwd*	Overwhelming *Overweldigend*	Unpredictable *Onvoorspelbaar*	Complex *Complex*	Confident *Zelfverzekerd*
Unrefined *Onafgewerkt*	Rigid *Stijf*	Engaging *Innemend*	Annoying *Vervelend*	Busy *Druk*
Expected *Verwacht*	Sterile *Steriel*	Advanced *Gevorderd*	Essential *Essentieel*	Straight Forward *Evident*
Unapproachable *Onbenaderbaar*	Distracting *Afleidend*	Meaningful *Betekenisvol*	Trustworthy *Betrouwbaar*	Old *Oud*
Intuitive *Intuïtief*	Cutting edge *Baanbrekend*	Integrated Geïntegreerd	Unattractive *Onaantrekkelijk*	Intimidating *Intimiderend*
Time-consuming *Tijdrovend*	Secure *Veilig*	Ineffective *Ineffectief*	Helpful *Behulpzaam*	Too Technical *Te technisch*
Optimistic *Optimistisch*	Personal *Persoonlijk*	Exciting *Opwindend*	Professional *Professioneel*	High quality *Hoge kwaliteit*
Disruptive *Storend*	Collaborative *Samenwerkend*	Fun *Leuk*	Entertaining *Vermakelijk*	Flexible *Flexibel*
Inspiring *Inspirerend*	Slow *Langzaam*	Appealing *Aansprekend*	Understandable *Begrijpelijk*	Incomprehensible *Onbegrijpelijk*
Dull *Saai*	Responsive *Responsief*	Reliable *Degelijk*	Sophisticated *Geavanceerde*	Innovative *Innovatief*

Novel *Nieuw*	Usable *Bruikbaar*	Friendly *Vriendelijk*		