Bachelor's Thesis Computing Science

The potential of user experience design in countering email-based social engineering attacks

Leon Zhang
s1042092

June 26, 2023

First supervisor/assessor:
Dr. Hanna Schraffenberger
`hanna.schraffenberger@ru.nl`

Second assessor:
Dr. Ir. Erik Poll
`e.poll@cs.ru.nl`

Second supervisor:
Jorrit Geels
`jorrit.geels@ru.nl`

Radboud University

Abstract

Email is a widely used communication medium for both personal and professional uses. However, it can also be exploited for malicious intents, e.g., through social engineering attacks. In fact, a significant portion of social engineering attacks take place through email. This poses a significant risk to organisations, as they can suffer severe consequences from such attacks. In this research, we address this issue by aiming to assist organisational email users in identifying malicious emails. We have developed visual designs of identity-based digital signatures that can be incorporated in email clients. This thesis investigates how email-based social engineering attacks can be countered with user experience design. We proposed a countermeasure in the form of mockups of identity-based digital signatures. Such signatures allow the sender to sign emails with personal attributes (e.g., name, organization, employer), henceforth ID verifications. The effectiveness of these signatures were assessed in terms of their effect on email credibility. We designed three main types of identity-based digital signature banners that are presented to the user above the email: warning, irrelevant ID verifications and relevant ID verifications. These designs indicate to users whether an email has been signed or not, and contain information about the sender's identity, which we will refer to as ID verifications. If the email is digitally signed, a signature design is displayed based on the ID verifications that are included in the signature. On the other hand, if the email is not signed, a warning is shown to emphasize caution. The goal is to inform the user about the origin of the email, based on which email credibility can be assessed. The impact of these signature designs was assessed through an online survey, where participants were asked to evaluate email credibility. Our results are promising, demonstrating that the identity-based digital signature design containing relevant ID verifications increased the perceived credibility of an email. No evidence was found to conclude that the warning design significantly decreased email credibility, or that the design for irrelevant ID verifications increased email credibility. Based on these findings, we show that identity-based digital signature have the potential to be part of a solution in countering email-based social engineering attacks. We suggest conducting further research in this field, to gain a more profound understanding of the effects of identity-based digital signatures and to improve the designs.

Contents

Acknowledgements

I would like to express my gratitude and appreciation to all those who have contributed to the completion of this thesis. First and foremost, I would like to thank my supervisor, Dr. Hanna Schraffenberger, for her invaluable guidance, expertise, and support. Secondly, I would like to extend my appreciation to my second assessor, Dr. Ir. Erik Poll, for his time and support. I am also thankful to Jorrit Geels, for his feedback, time and help with the experiment and identity-based digital signature designs. Furthermore, thank you to all the participants who used their time and effort to participate in this study. Their participation has been essential to this thesis. Finally, I would like to express my appreciation to my family and friends for their support and encouragement during the entire duration of this thesis.

Thank you all.

Chapter 1

Introduction

In today's digital world, email has become an unmissable communication medium for both personal and professional uses. In 2022, there were over 333 billion emails sent and received every day around the world (Radicati, 2022). Unsurprisingly, emails can also be used by malicious actors, notably in the form of social engineering (Stine & Scholl, 2010). Social engineering refers to all techniques to talk a target into performing certain actions, such as giving away sensitive information or control over a computer system (ENISA, 2022; IBM, n.d.). Email-based social engineering attacks are a form of social engineering attacks, in which an attacker uses email as a communication medium, to manipulate these victims. All in all, humans represent the weakest link in information security due to their susceptibility to deception, manipulation, and errors (Mouton et al., 2014). Hence, it is only to be expected that social engineering attacks are becoming more common (Mitnick Security, 2022).

Social engineering is responsible for 98% of the targeted cyber attacks (Proofpoint, 2023). There are various types of social engineering attacks, take as an example: phishing and baiting (Rouse, 2022). Phishing is a type of online scam where an attacker tricks a person into performing the attacker's desired actions via communication media (Salahdine & Kaabouch, 2019); baiting achieves this by luring people using incentives (Salahdine & Kaabouch, 2019).

In this thesis, we solely focus on email-based social engineering attacks that may occur in organizations. It is estimated that 91% of the cyber attacks start with an email message, only 3% of those attacks use malware and the rest of them, 97%, rely on social engineering (Proofpoint, 2023). Successful social engineering attacks can have a devastating impact on both individuals and organizations. For example, data breaches, deployment of destructive malware or the disruption of critical systems can all result in significant losses such as damage to image, reputation and finances (Fuertes et al., 2022; Rock, 2021). But that is not all, with the rise of mobile devices being used for email communications, even more challenges are introduced (Witts, 2023). From this we pose the following initial question: How can we help the user to identify malicious emails on various types of devices by using user interface and user experience design.

There are various blogs where the relation between cyber security and user experience is discussed where most of them lead to the conclusion that user experience should not be neglected (Wagenseil, 2022; Brown, 2019; Menlo Security, 2022). This is because user experience plays a major role in emails: it defines how users perceive and react to these messages. For that reason, user experience design can be used to help users in avoiding malicious emails. Therefore, this thesis aims to counter email-based social engineering attacks on organizations by incorporating user experience in the process of proposing new countermeasures. We will examine and discuss the current state of existing measures. Since social engineering attacks are one of the biggest threats to organisations with severe consequences (HackControl, n.d.; Partida, 2020); we define the scope of this research to be organizational email accounts that are accessed on personal computers and mobile devices.

The main research question is formulated as follows: "**How to counter email-based social engineering attacks with user experience design?**"

We also introduce three sub-questions to support the main research question, these are:

1. What are the current countermeasures used by email clients and webmail services to prevent email-based social engineering attacks?

2. What are the limitations of countermeasures against email-based social engineering attacks?

3. How can countermeasures be made more effective against email-based social engineering attacks via user experience design?

The remainder of this thesis is organized as follows: Chapter 2 contains the preliminary knowledge required for this thesis. Chapter 3, provides and discusses a selection of related work useful for this thesis. In chapter 4, we present and review existing countermeasures against email-based social engineering attacks along with proposing a new way to counter email-based social engineering attacks. The scope and methodology of this research is described in chapter 5. Chapter 6 contains the results of the conducted research. Finally, in chapter 7, a discussion of results is provided along with some limitations of the research and some directions for future work. This chapter also contains a summary of the overall conclusions to this thesis, presenting key findings.

Chapter 2

Preliminaries

In this chapter we present preliminary knowledge that is required for the remainder of this thesis. First we start with some basic knowledge: the working of the internet and electronic mail (email) accompanied with their security risks. Next, we explain email security techniques that can help with countering email-based social engineering attacks. Afterwards, we explore the world of social engineering attacks by giving a brief definition of social engineering and presenting several types of email-based social engineering attacks. Lastly, we provide a definition for user experience, user interface design and mobile-first approach before we investigate visual security indicators.

2.1 An overview of the internet and electronic mail

This section starts with an explanation of several key concepts of the internet in order to achieve a better understanding of emails and email-based social engineering attacks. After that, we explain how emails work by presenting their essential protocols and security risks.

2.1.1 The internet

Today's digital communication is based on multiple standardized communication protocols. Communication protocols are a system of rules that describe how information must be formatted in order to achieve uniformity in transmitting and receiving information between entities (Australian Research Data Commons, 2022; Subedi, 2022).

The internet, a global network of computers and other devices that communicate with each other, makes use of these communication protocols to enable the functioning of the internet as we know it today. When a user sends a message over the internet, the message will be split into smaller chunks. These chunks of the original message are called packets (Cloud-

Flare, n.d.-a). Then, a technique known as packet switching is used to sent the packets across the internet. This technique allows network devices to process packets independently from each other so that a single connection does not take up an entire network. Ultimately, the recipient's device recombines these packets to view the original message.

Now we take a closer look at how computer systems communicate with each other. For this we use the five layered Transmission Control Protocol/Internet Protocol (TCP/IP) model. This model provides a framework for structuring the set of standardized communication protocols so that data can be transmitted over a network (IBM, 2023). This model consists of five different abstraction layers, namely the Application layer, the Transport layer, the Internet layer, the Data-link layer and the Physical layer. The Data-link layer and Physical layer are often combined and seen as one layer, however in this thesis we keep those separated. Each of the layers contain a set of protocols to perform their specific task, which will be further elaborated below. This can be seen in Table 2.1.

Layer name	**Protocol**
Application	SMTP, IMAP, POP3, HTTP, FTP, . . .
Transport	TCP & UDP
Internet	IP, ICMP, ARP, . . .
Data-link	Ethernet, WiFi, . . .
Physical	10BASE-T, 802.11, . . .

Table 2.1: The five layered TCP/IP model along with some of the protocols. Source: Figure designed based on Microchip (n.d.).

In order to understand what each layer exactly performs on the data, we have to bring back the term packets. Packets are constructed in such a way that for every abstraction layer a packet consist of two parts. The first part of a packet is the header and is prepended to the second part i.e. the body. The header contains protocol information that is relevant to that abstraction layer. The body contains the data and is often an 'entire packet' (header and body) for the next layer in the TCP/IP model.

So a packet at a specific layer contains all the information passed from the higher layers and is prepended with a header corresponding to the current layer. This does not hold for the Application layer since it is the highest layer, here the data is already formatted with the application headers by the application. Also, the Data-link layer does not only prepend a header to the data but it also adds a trailer/footer. The process of preserving data of the above layer while adding a new header of the current layer is known as encapsulation (Zwicky & Chapman, 1995) and supported by Figure 2.1.

TCP/IP is also known as the Internet protocol suite. This is because it

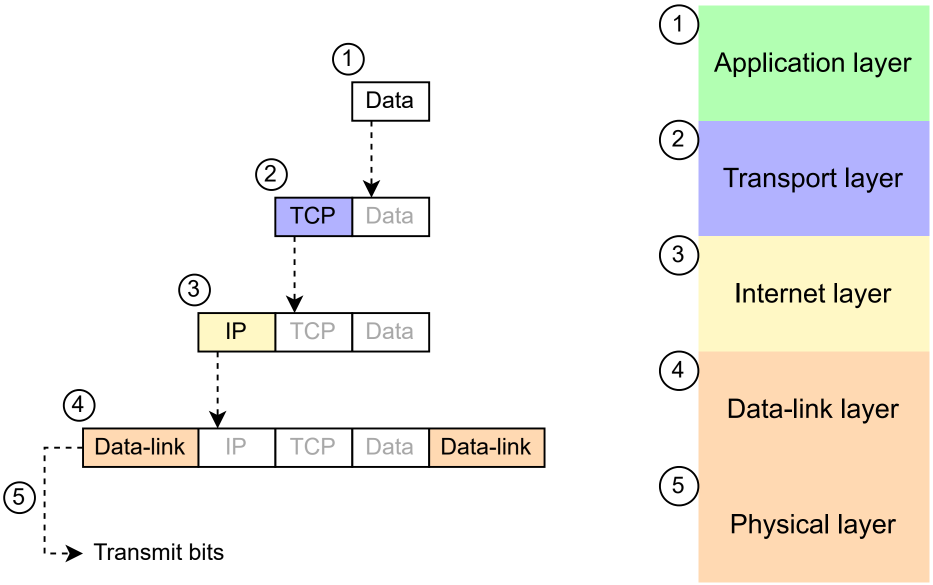


Figure 2.1: The encapsulation of data on the sending host in five steps. Source: Figure designed based on An (2015).

is simply a suite of communication protocols. If a user sends information it passes as data through all the layers in a particular order (Application to Physical layer) and when a user receives the information, it is passed through in reverse order, and can be seen in Figure 2.2. The TCP/IP model is very important for digital communication because it defines how information is processed for transmission from the sender to receiver, thus providing a standardized way for computer systems to communicate with each other.

Now we will examine each layer individually. The first and lowest layer is the Physical layer. This layer encodes and decodes data frames from the Data-link layer to bits and contains all the functions to carry this data in the form of bits to another system (Tucker, 2020).

The second layer, that is the Link layer, breaks packets received from the Internet layer into data frames. This layer enables connections between two hosts in a local network (Tucker, 2020).

The third layer, the Internet layer, provides the required functionality to deliver data packets to non-adjacent systems, and is responsible for packaging, addressing and routing the data (Alpern & Shimonski, 2010). It hosts various protocols to achieve this. Some examples of those protocols are: Internet Protocol (IP), Address Resolution Protocol (ARP) and Internet Control Message Protocol (ICMP).

The fourth layer, the transport layer, is responsible for end-to-end delivery of data from the source host to destination host and sets up communica-

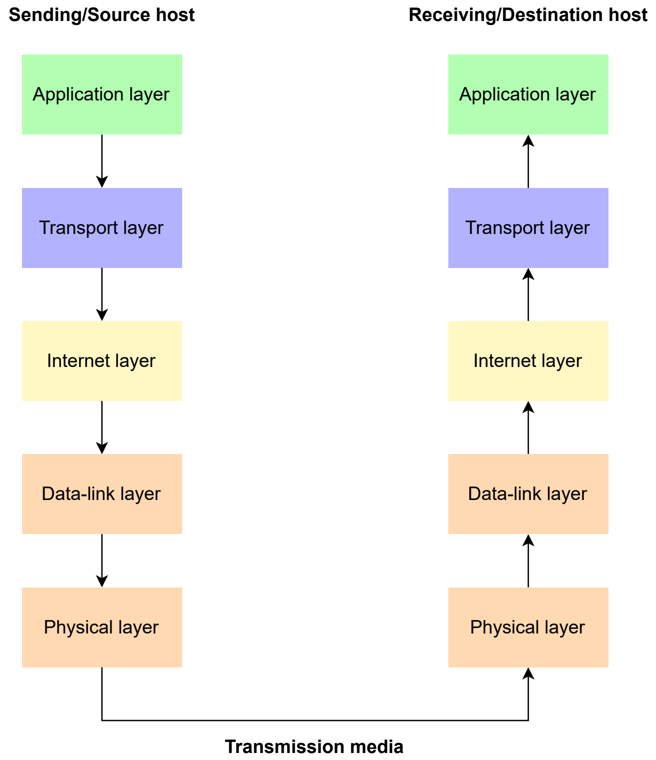


Figure 2.2: The process of sending and receiving information using TCP/IP. Source: Figure designed based on Oracle (2015).

tion between the application layer and the lower layers. The Transport layer mainly uses the following protocols: Transmission Control Protocol (TCP) and User Datagram Protocol (UDP). TCP is a connection-oriented protocol that is reliable by the fact that it retransmits lost or discarded packets, thus assuring delivery of data. It also guarantees that packets are delivered in the same order in which they were sent without errors and duplications. UDP is a connection-less protocol that is less reliable and faster than TCP since there is no error and flow control. Both TCP and UDP packets are by default not encrypted.

And at last, the highest layer named the Application layer, provides an interface between software applications and network services. This is where the protocols for sending and receiving emails reside: Simple Mail Transfer Protocol, Post Office Protocol and Internet Message Access Protocol. All three protocols communicate using TCP, SMTP is used for sending emails and the last two are used for retrieving emails. Other protocols that reside in this layer are for example the File Transfer Protocol (FTP) and Hypertext Transfer Protocol (HTTP).

Internet Protocol addresses

An Internet Protocol address (IP address) is a numerical label assigned to every device on a network. This label uniquely identifies a host and provides its location in the network. The usage of IP addresses allows the Internet Protocol to transmit messages to the correct recipient. This is done by providing a destination IP address in the IP packet.

Currently, two versions of the Internet Protocol are commonly used, namely Internet Protocol version 4 (IPv4) and Internet Protocol version 6 (IPv6). IPv4 addresses have a size of 32 bits and are written in the following format: `x.x.x.x` where `x` is a decimal value between 0 and 255, also called an octet. This limits the address space to 2^{32} addresses. IPv4 is defined in RFC 791[1].

The other version, IPv6, is defined in RFC 2460[2]. It is designed to eventually replace IPv4 as a solution to the eventual exhaustion of IPv4 addresses. These addresses have a size of 128 bits: `y:y:y:y:y:y:y:y` where `y` is a hexadecimal value between 0 and FFFF.

IP addresses come in four types: public, private, static and dynamic. A public IP address can be accessed over the internet and is used for connections outside the local network. Private IP addresses are used inside the local network. IP addresses can also be static or dynamics, indicating whether the IP address changes over time. Static IP addresses are manually configured by the host's hardware or software to be persistent. Dynamic IP addresses are assigned by the network and change over time.

Domain Name System

The Domain Name System (DNS) is an Application layer protocol that translates domain names into their corresponding IP addresses. DNS is invented to address the following problem: IP addresses can be dynamic and are notated in a format that is not meaningful to humans, therefore making them difficult to memorise. This lead to the invention of domain names and DNS. Domain names are usually intuitive and short so that they can be easily remembered. However, these domain names do have to be translated back into IP addresses, and is done by DNS. A domain name consists of one or multiple labels delimited by dots and are structured hierarchically. The right most label corresponds to the top-level domain then on the left there is the second-level domain and so on. Note that nearly all Uniform Resource Locators and email addresses contain domain names. Here, DNS is used to find the corresponding IP addresses.

Uniform Resource Locators

In order to understand how attackers deceive their targets via email, we need

[1] `https://www.ietf.org/rfc/rfc791.txt`

[2] `https://www.ietf.org/rfc/rfc2460.txt`

to know how uniform resource locators function and how they are formatted. A Uniform Resource Locator (URL) is a unique address that points to a web resource's location so that it can be retrieved. URLs consist of mandatory and optional parts in the following format:

```
scheme \://" authority \/" path \?"  parameters \#" anchor
```

Take as an example the following URL:
`https://www.example.com:80/path/to/thesis.html?key=val#chapter1`
Here "https" is the scheme. The scheme indicates the protocol that needs to be used to request the resource, and is followed by the character ":". For web pages this is usually "http" or "https" and for opening email clients and beginning an email message this is "mailto".

After the scheme we have "www.example.com:80" as the authority. An authority is optional and is separated from the scheme by the characters "//". The authority consists of the following subcomponents:

```
userinfo \@" host \:"  port
```

Where we have the following:

- An optional subcomponent `userinfo` consisting of `user \:"  password` where `password` is optional.
- The `host` subcomponent which essentially is the host name and domain name.
- An optional `port` number after the ":" character.

From this we derive that "www.example.com" is the `host` and "80" is the `port` number. This is followed by the path to the resource: "/path/to/thesis.html" and some optional parameters: "?key=val". Lastly we have an optional anchor to the resource itself: "#chapter1". The part after "#" is known as the fragment identifier. Anchors lead the user to a specific place in the resource and are never sent in the request to the server. It is not considered to be a part of the URL by RFC 1808[3].

It is also important to know for the remainder of the thesis that hyperlinks or links are text, icons or other media encapsulating a URL, because these are commonly used in malicious emails (Cofense, 2023). By interacting with links such as clicking or tapping, one navigates to the URL.

2.1.2 Electronic mail

Email is a communication medium that uses devices to deliver messages, documents, and other types of files across computer networks. The Simple

[3] `https://www.ietf.org/rfc/rfc1808.txt`

Mail Transfer Protocol (SMTP) is fundamental to email communication, and was designed to be a simple, lightweight, reliable and efficient protocol to transmit email messages. Because of this simplicity, SMTP only provides basic functionality and lacks security features. This problem is partially addressed by the introduction of various extensions to this protocol. Some of those extensions will be discussed later. We will also address two other important protocols used for email communication, namely Post Office Protocol (POP) and Internet Message Access Protocol (IMAP).

Email addresses

Email addresses are a unique string of characters that identify a mailbox to which email messages should be sent or received. An email address is made up of three parts: a local-part, "@" symbol and a domain name. Technically, one could also use an IP address instead of a domain name. Take as an example the following email address: `local-part@domain.com`. Here "local-part" denotes the local-part and is essentially the username that indicates a unique mailbox. The domain name is in this example "domain.com". SMTP uses DNS to translate domain names into the corresponding IP addresses. The IP address is then used to deliver the message to the next or final mail server.

Email message format

Email messages are structured into three components, namely the SMTP envelope, headers and the body. The SMTP envelope contains the conversation between SMTP clients and SMTP servers. A SMTP conversation consists of a SMTP client sending SMTP commands to a SMTP server where the SMTP server responds to these commands. An example conversation can be seen in Figure 2.3. We see that during a SMTP conversation the SMTP command `RCPT TO` tells the mail server the destination of the email message and the SMTP command `MAIL FROM` specifies the sender of the email message. Therefore the SMTP envelope is used to route the email message between mail transfer agents (MTA). MTAs will be explained in a later paragraph. So, the SMTP envelope is used to tell mail servers where to send the email message to, and where email clients can read the email so that it can be displayed to the user. The formatting of this envelope is specified in RFC 5321[4].

The header contains various fields that provide information about the sender and recipient such as `From`, `Date`, `CC`, `To` and `Subject`. Whereas, the body contains the message. The body and header of emails are formatted according to RFC 5322[5].

However, this approach does come with a challenge: the email addresses

[4] `https://www.ietf.org/rfc/rfc5321.txt`

[5] `https://www.ietf.org/rfc/rfc5322.txt`

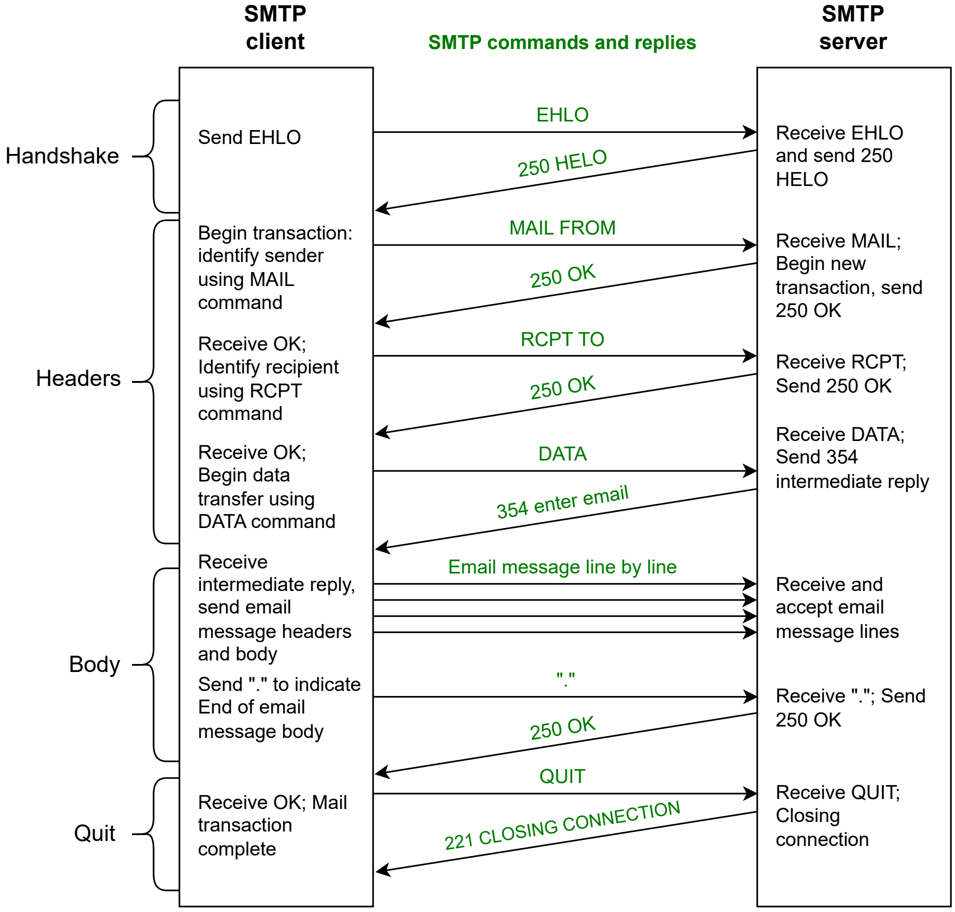

SMTP client
SMTP commands and replies
SMTP server
Handshake
Headers
Body
Quit
Send EHLO
EHLO
Receive EHLO and send 250 HELO
250 HELO
Begin transaction: identify sender using MAIL command
MAIL FROM
Receive MAIL; Begin new transaction, send 250 OK
250 OK
Receive OK; Identify recipient using RCPT command
RCPT TO
Receive RCPT; Send 250 OK
250 OK
Receive OK; Begin data transfer using DATA command
DATA
Receive DATA; Send 354 intermediate reply
354 enter email
Receive intermediate reply, send email message headers and body
Email message line by line
Receive and accept email message lines
Send "." to indicate End of email message body
"."
Receive "."; Send 250 OK
250 OK
Receive OK; Mail transaction complete
QUIT
Receive QUIT; Closing connection
221 CLOSING CONNECTION

Figure 2.3: SMTP sequence diagram when everything works correctly.
Source: Figure designed based on Elghamrawy (n.d.).

specified in the envelope `MAIL FROM` and the email header `From` can be forged. This will be discussed in section 2.4.1.

Email clients

An email client, or more formally a mail user agent (MUA), is a software program that allows users to access, send and manage email messages. Email clients can also be categorized into various types:

- Desktop email clients: Standalone email clients installed on a user's computer and are run locally.
- Mobile email clients: Email clients specifically designed for mobile devices such as tablets or smartphones.

- Webmail clients: Clients that are run remotely on server and thus can be accessed everywhere via a web browser as long as one is connected to the internet.

In this thesis, we use the term email client to address all types of email clients.

2.1.2.1 Email protocols

The aforementioned SMTP protocol is a text-based, connection-oriented Internet standard protocol for delivering email messages over a TCP connection.

A SMTP session, also known as SMTP conversation, is the communication between mail servers via SMTP commands. If the server is not the recipient's mail server, then it relays the message to another server until the recipient's mail server is reached. The specifications of this email protocol can be found in RFC 5321[6]. A sequence diagram of SMTP can be seen in Figure 2.3.

SMTP extensions

There exists various extension to the SMTP protocol, we will address two extension to SMTP: Extended Simple Mail Transfer Protocol (ESMTP) and Multipurpose Internet Mail Extensions (MIME). ESMTP allows authentication, encryption, the inclusion of email attachments and more to SMTP. MIME extends SMTP by allowing users to include email attachments such as documents, images or audio files. It also allows one to use text in other character sets than ASCII. This is done by adding five additional fields to the header of the actual email.

Internet Message Access Protocol and Post Office Protocol

IMAP and POP are Internet standard protocols over TCP and used by email clients for retrieving email messages from a mail server. The current version of IMAP i.e. IMAP4, is defined in RFC 9051[7]. POP is defined in RFC 1939[8] and the most common version is POP3.

Both protocols follow more or less the following procedure:

First a user logs in their email account in a email client. Then the email client uses IMAP or POP to communicate with the mail server. After that, the client fetches a list of available mailboxes such as "Sent items", "Inbox" and "Junk" from the mail server. The client displays these mailboxes along with the email messages. Now when a user opens an email message, the client sends a request to the mail server to obtain the content of the message.

[6] https://www.ietf.org/rfc/rfc5321.txt

[7] https://www.ietf.org/rfc/rfc9051.txt

[8] https://www.ietf.org/rfc/rfc1939.txt

Then the mail server sends this message back to the client. Upon receiving, the client displays it to the user.

The main difference between IMAP and POP is that IMAP allows users to access email messages from multiple devices, while keeping these synchronized across all devices. It also allows users to manage their email directly on the mail server. Whereas with POP3, a message is typically deleted from the mail server once the message is retrieved and downloaded from the mail server. This means that email will not be synchronized across all devices.

Sending and receiving emails
The process of sending and receiving emails is visualized in Figure 2.4. when a user composes and sends a mail via a MUA, the MUA proceeds to add email headers to the message and delivers the email using SMTP to the user's mail server. If webmail is used, then HTTP or Hypertext Transfer Protocol Secure (HTTPS) is used to deliver the composed email message to the mail server. This mail server contains a mail submission agent (MSA) and a MTA. The MSA receives the email message from the MUA and cooperates with the MTA to deliver this email message. A MTA is used for transferring email messages between computers. If the recipient's device is not hosted locally, then the MTA sends a DNS query to identify the mail server(s) of the recipient(s). Afterwards the DNS responds with a DNS mail exchange (MX) record, indicating how the email message should be routed by SMTP to the mail server(s) of the recipient(s). After obtaining this record, the email is forwarded to another MTA over the internet, eventually routing it to the final destination i.e. the recipient's mail delivery agent (MDA). Finally, the MDA delivers the email message to the recipient's mailbox.

For receiving emails, The user's MUA polls and retrieves the email from its mail server using IMAP, POP or HTTP/HTTPS if webmail is used.

2.1.2.2 Security risks of email protocols

Unfortunately, the email protocols SMTP, POP and IMAP do come with some security risks. It is important to note that we do not list all security risks but the ones that are important and related to email-based social engineering attacks.
Default IMAP and POP share more or less the same security risks that we have identified based on Ashtari (2023):

- **Lack of encryption**
  Traffic is typically unencrypted meaning that an attacker could intercept network traffic and view or temper with email messages. Also the transmission of login credentials from a client to the mail server is by default in plain text.

- **Potential weak passwords**

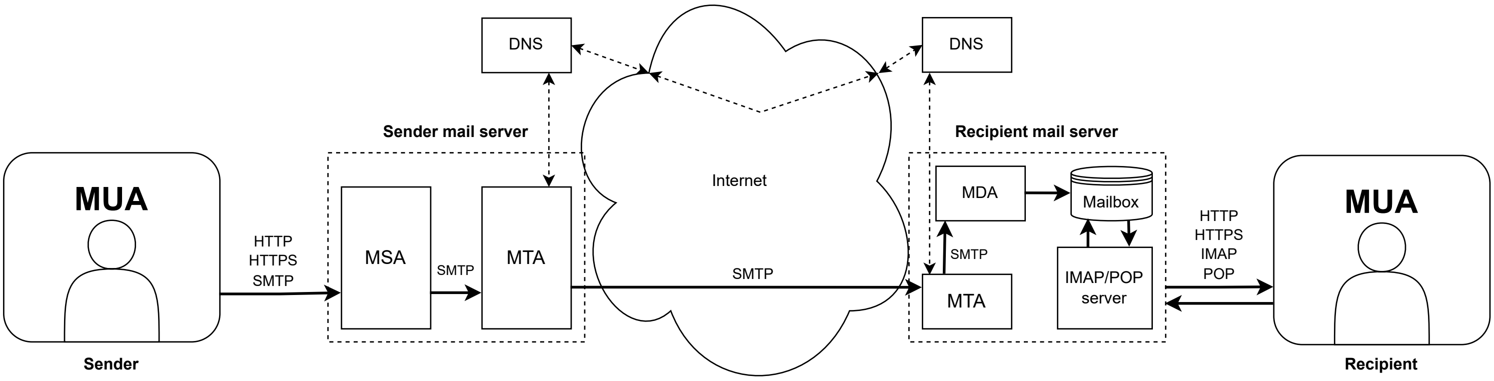


Figure 2.4: Email process flow from sender to receiver.
Source: Figure designed based on Pollock (2016).

IMAP and POP rely on username and password authentication, weak passwords could be exploited to get access to emails.

SMTP by default includes no security measures. We have identified the following security risks resulting from the design of SMTP:

- **Lack of authentication**
  SMTP by default does not have any mechanisms for the authentication of the sender of an email. This allows attackers to impersonate legitimate senders.

- **Lack of encryption**
  SMTP email messages are transmitted in plain text format over TCP, meaning that an eavesdropper with the correct tools can read and modify emails before it reaches the intended recipient.

- **Reliance on DNS**
  SMTP relies on DNS to ensure that messages are routed to their intended recipients. Therefore, attackers could manipulate DNS records to redirect email traffic to their own servers and view or tamper with the email.

2.2 Email security

In this section we start with explaining some key concepts of email security. To be precise, public-key cryptography, which is used for email encryption and email authentication. After that we examine some email encryption and email authentication protocols.

2.2.1 Public-key cryptography

Public-key cryptography (Diffie & Hellman, 1976), also known as asymmetric cryptography, is a technique that uses cryptography and a pair of keys to secure communication between two parties. Symmetric cryptography is the opposition of asymmetric cryptography and only makes use of a single key to secure communication.

In public-key cryptography, each user has a pair of keys: a public key and a corresponding private key. The public key can be shared with anyone and is used for encryption. The private key is used for decryption and should be only known to the owner, therefore it must be kept secret. If a user wants to send a message to another user securely, then this user uses the public key of the recipient to encrypt the message. Subsequently, this encrypted message is sent to the recipient. Upon receipt, the recipient uses its corresponding private key to decrypt the message. Note that this is the only key that can decrypt messages that are encrypted with the corresponding public key. This process is shown in Figure 2.5 and is known as the public-key encryption system, providing confidentiality.

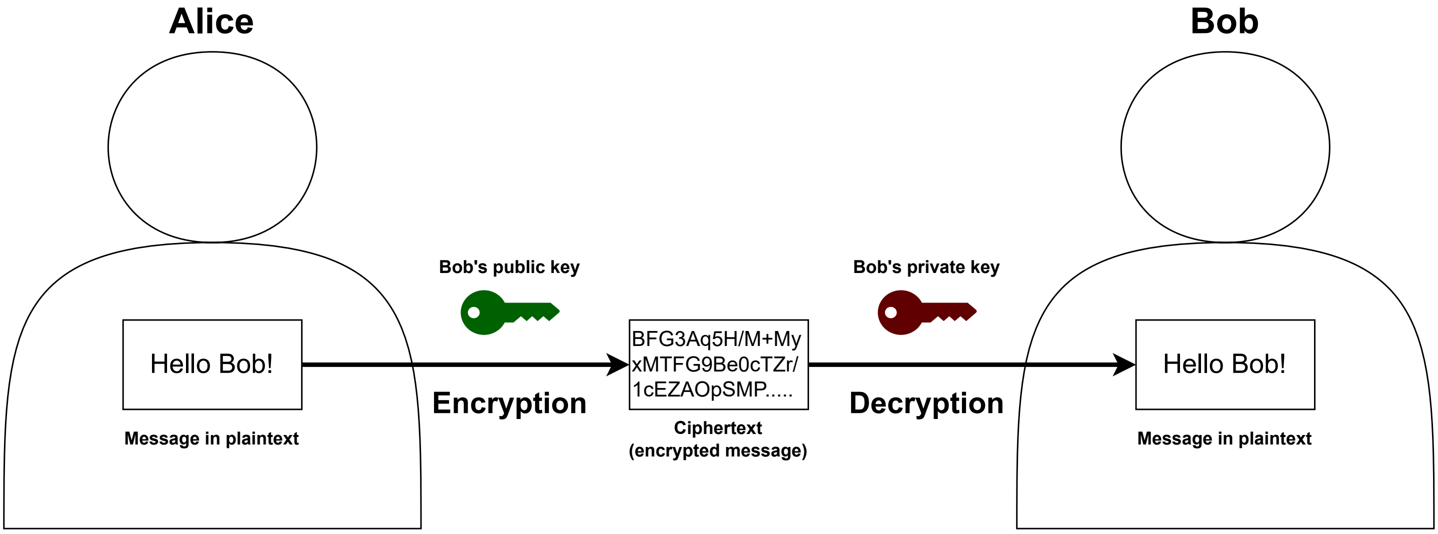


Figure 2.5: The process of public-key encryption.
Source: Figure designed based on Göthberg (2006).

It is also possible to use this pair of keys to create digital signatures. The process is as follows, the sender generates a signature using the message and the sender's private key. If one wants to verify the authenticity of the message, it then uses the corresponding public key (the sender's public key) to check whether the signature matches the message i.e. a boolean value “true” if the signature matches the message and otherwise “false”. This process can be seen in Figure 2.6. This is also known as a digital signature system, providing integrity, authenticity and non-repudiation of data.

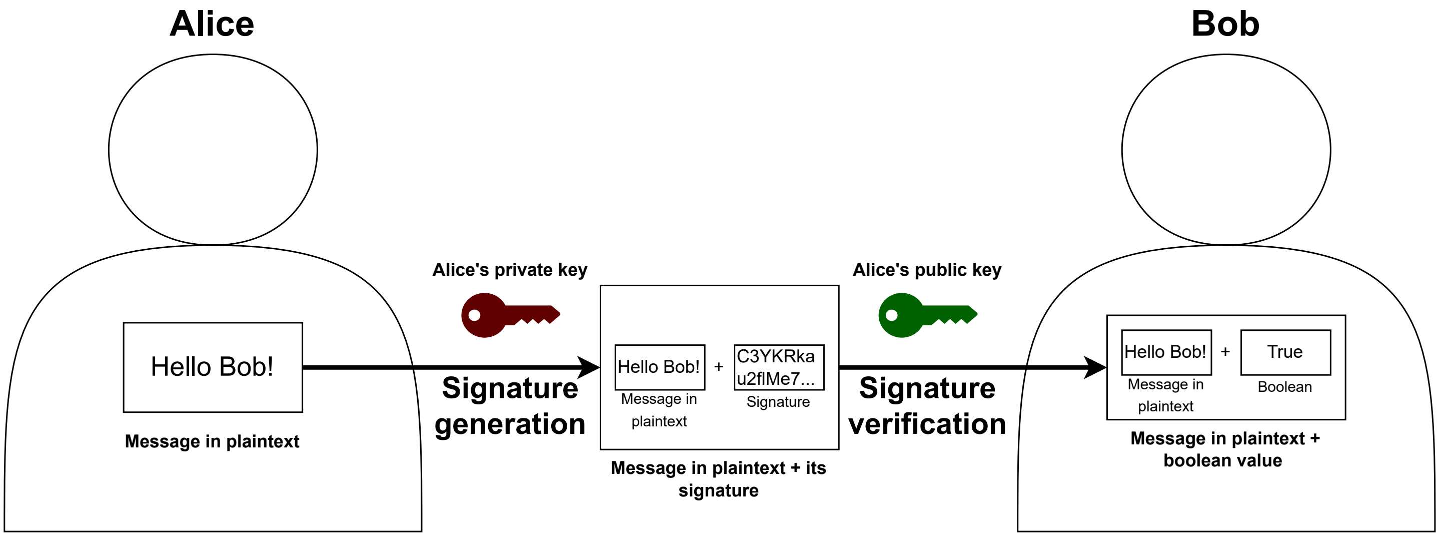


Figure 2.6: The process of public-key signing.
Source: Figure designed based on FlippyFlink (2019).

So public-key cryptography can be used to provide confidentiality, integrity, authenticity and non-repudiation of email messages. We will briefly explain these security principles in the context of emails:

- **Confidentiality:** Only authorized recipients are able to read the email message.
- **Integrity:** The email message has not been modified by an unauthorized person or process.
- **Authenticity:** The email is truly sent by the owner of the sender's email address
- **Non-repudiation:** The sender of the email cannot claim that it did not send the email.

It is worth noting that asymmetric cryptography offers more security but less performance than symmetric cryptography. This is because asymmetric cryptography makes use of a public and private key, making key management more complex than symmetric cryptography. In symmetric cryptography one uses the same key for encryption and decryption. However, it is less secure since the secret key has to be shared with recipients for decryption. In reality, both cryptography methods are used to maintain a balance between security and performance. In the following subsection we will explain how these key pairs are managed so that protocols can use public-key cryptography without having to worry about key management and the legitimacy of keys.

2.2.1.1 Public key infrastructure

Public-key cryptography relies on the usage of trusted and verified public keys, and private keys kept secret. A system for establishing trust is called a public key infrastructure (PKI). A PKI is a system for creating, storing and distributing digital certificates (Okta, 2022). Digital certificates are used to verify that a public key belongs to a specific entity. These certificates contain at minimum: some information about the public key, identity of the public key's owner and a digital signature of the one that verified the certificate's content. The PKI creates these digital certificates which bind entities to public keys and ensures that these digital certificates can be trusted.

We will discuss two approaches that makes use of digital certificates in order to achieve trust:

1. **Certificate authority (CA)**
   CA is a centralized approach in establishing trust via digital certificates. A CA is a trusted third party that issues, signs and manages digital certificates (Xolphin, n.d.). This is done using the CA's private key, meaning that trust in a user's key also depends on the trust of the authenticity of the CA's key. CAs can issue digital certificates directly to entities or, what is often the case, by authorising another CA to do so. It is a hierarchical system, meaning that there exists multiple layers of CAs where trust of a user's key relies on its superior CA. Entities that want to verify the legitimacy of a key can do this at the corresponding CA. This implementation is used by the following email encryption protocols that are discussed in this thesis: Secure / Multipurpose Internet Mail Extensions (S/MIME) and Transport Layer Security (TLS).

2. **Web of Trust (WoT)**
   WoT is a decentralized approach using a distributed system. It makes use of self-signed digital certificates and third parties to verify the certificates (MATTR, n.d.). Keys are verified among users and the authenticity of a key is based on the number of key signings. Key management and trustworthiness of keys are managed by the network of users. This approach is used by the email encryption protocol: OpenPGP.

2.2.2 Email encryption

As mentioned, SMTP is by default not secure and does not encrypt any information. This means that communication between mail servers are transmitted in plain text format and can be read or altered by eavesdroppers.

To combat this, various email encryption protocols have been introduced. Email encryption is a technique where an email, could be solely the body of the email or both header and body, is encoded in such a way that only authorized entities can read the information after decryption. In this subsection we introduce some encryption protocols that are used for email encryption.

2.2.2.1 Secure Socket Layer and Transport Layer Security

Secure Socket Layer (SSL) is the predecessor of Transport Layer Security (TLS). Both protocols use public-key cryptography and provide confidentiality and data integrity of emails. These protocols run on top of transport protocols such as TCP and UDP, encrypting Application layer traffic. For that reason these protocols are classified as Application layer protocols

The process of using SSL or TLS on top of TCP for delivering email messages is as follows: first a handshake between the MUA and MTA has to be established. If this succeeds, then the connections between MUA and MTA or MTA and MTA are encrypted. If a MUA wants to tell a MTA that it wants to make use of a secure SMTP connection via TLS or SSL, it first has to send the protocol command `STARTTLS`. The same command can be used with IMAP. POP3 has a different command called `STLS`. Note that SSL is deprecated because of various security vulnerabilities (Plesky, 2022; Möller et al., 2014).

There are two approaches in establishing a secure connection:

1. **Opportunistic SSL/TLS**
   The MUA sends the STARTTLS protocol command to the MTA to upgrade the current unsecure connection to an encrypted one. If the MTA is compatible with the SSL or TLS versions supported by the MUA and no errors occur, then a secure TLS or SSL connection will be established.

2. **Forced SSL/TLS**
   The MUA will try to establish a secure connection without asking the MTA about its compatibility. If this does not succeed then it entirely depends on what action is taken by the MUA and user. If it does succeed then a secure connection will be established.

If the sender's MTA and the recipient's MTA supports TLS, then an eavesdropper who is sniffing the traffic between MTAs cannot read the messages meaningfully. However, the email message is revealed to intermediate email relays (SMTP relays). Therefore, the email can be read and altered by these servers. Meaning that if one uses TLS or SSL it has to trust every server the email passes through. A more secure method that provides secure connections between client and client will be discussed in the next subsection.

2.2.2.2 End-to-end encryption

End-to-end encryption (E2EE) is an encryption protocol where data is encrypted and decrypted at the end points using public-key cryptography. This is different from transport-layer encryption such as TLS, since it ensures that no intermediary parties such as service providers, can decrypt and read the messages. This means that emails sent with E2EE are encrypted at the sender and decrypted at the recipient, making them not readable to any entity except the recipient. The intermediary parties can only see the encrypted data since they do not have the corresponding keys to decrypt the message.

We will briefly discuss two commonly used E2EE email encryption systems: OpenPGP and Secure/Multipurpose Internet Mail Extensions (S/MIME).

S/MIME
Secure/Multipurpose Internet Mail Extensions[9] is an email E2EE encryption protocol using public-key cryptography and CAs as a trust model. It is a security extension of the MIME protocol that is implemented using S/MIME certificates. S/MIME allows users to encrypt email messages along with their attachments using the recipient's public key. This public key can be obtained from the corresponding certificate, where the certificate can be obtained from CAs. This ensures that only the intended recipient can read the email, if its the only one that possesses the private key belonging to the associated certificate.

Additionally, S/MIME allows users to digitally sign email messages so that the recipients can validate the identity of the sender and integrity of the email message. Therefore, S/MIME provides the following security capabilities: data confidentiality, authentication, data integrity and non-repudiation. In order to use S/MIME, the MUA of both sender and recipient have to enable and support S/MIME.

OpenPGP
OpenPGP[10] is an E2EE email encryption protocol applied on MUAs. This protocol relies on the Web of Trust instead of CAs. It uses a combination of asymmetric and symmetric cryptography to encrypt and decrypt email messages. OpenPGP makes use of OpenPGP certificates. These certificates consist of a public key, information that identifies the owner of the key and a signature (one of which is self-signed). These certificates can contain more than one signature, allowing certificates (key/identification pair) to be signed by other users. This is done so that people can be, in some sense, assured that the public key belongs to the specified owner (Perrig, n.d.).

[9] `https://www.ietf.org/rfc/rfc5751.txt`
[10] `https://www.openpgp.org/`

2.2.2.3 Identity-based encryption

Identity-based encryption (IBE) is a form of public-key encryption that utilizes unique identifiers, in this context email addresses, to generate a public key (Kay, 2008; Wikipedia contributors, 2023). The process of sending emails using identity-based encryption is depicted in Figure 2.7, describing how one can encrypt email messages or verify digital signatures. They first obtain a master public key from a trusted third party. This trusted third party is also known as the private key generator (PKG; Youngblood, 2005). Then, the desired public key can be generated from the master public key and the unique identifier, which can be any kind of string. In our case, this is typically an email address. One can also retrieve its own private key from the PKG upon successful authentication. This private key is generated by the PKG using the master private key. From this process it can be seen that IBE does not need to use a PKI to distribute keys. Anyone can generate a public key for any unique identifier, as long as one has access to the master public key and the unique identifier. Also, one can retrieve its private key from the PKG if one can successfully authenticate oneself to the PKG.

After one obtains the public key of the recipient and its own private key, one can sign, encrypt and send email messages to that recipient. If a user receives an email message that is encrypted using IBE and has not been issued a private key before; it queries the PKG for its private key. This private key is given upon successful authentication. After that the user can decrypt emails that are encrypted with a public key corresponding to that user's email address (if the email address was used as the unique identifier). An advantage of IBE compared to traditional public-key cryptography is that it does not need any prior key distribution over the parties involved in an IBE conversation (Wikipedia contributors, 2023). Therefore, it reduces the complexity of the encryption process. In traditional public-key cryptography the users have to deal with publishing and distribution keys which can become complex.

Another advantage of IBE is that if there were only a finite number of users, then the secrets of the PKG can be destroyed after having issued every user with its keys (Wikipedia contributors, 2023).

2.2.3 Email authentication

SMTP does not come with any security measures, thus also no authentication methods. On top of that, encryption does not assure that information came from the specified sender.

This lead to the development of various email authentication methods. We will only discuss the most widely adopted protocols in this thesis, namely Sender Policy Framework (SPF), DomainKeys Identified Mail (DKIM) and Domain-based Message Authentication, Reporting & Confor-

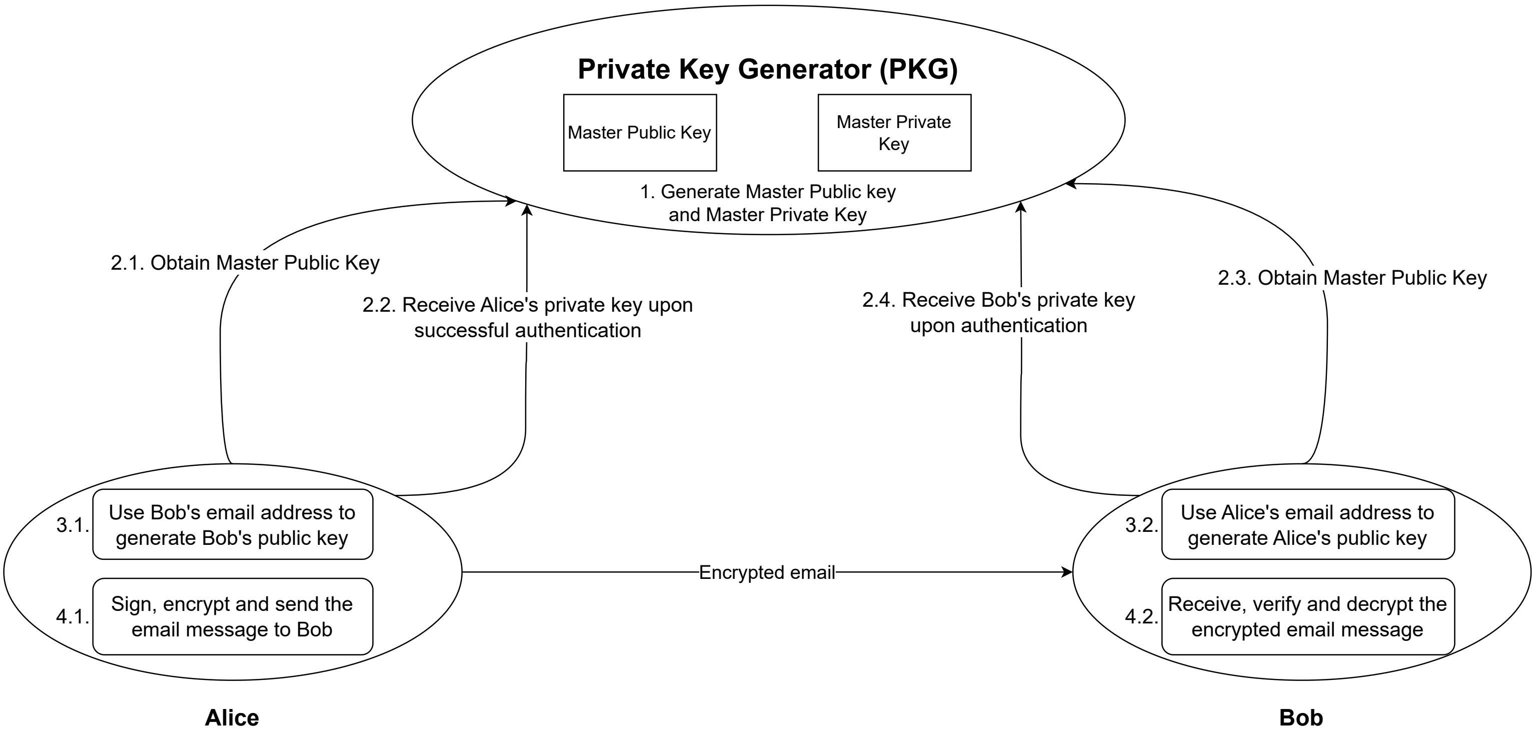

Private Key Generator (PKG)
Master Public Key
Master Private Key
1. Generate Master Public key and Master Private Key
2.1. Obtain Master Public Key
2.2. Receive Alice's private key upon successful authentication
2.4. Receive Bob's private key upon authentication
2.3. Obtain Master Public Key
3.1.
Use Bob's email address to generate Bob's public key
4.1.
Sign, encrypt and send the email message to Bob
Encrypted email
3.2.
Use Alice's email address to generate Alice's public key
4.2.
Receive, verify and decrypt the encrypted email message
Alice
Bob

Figure 2.7: Identity-based encryption steps diagram.
Source: Figure designed based on Sheffer (2009).

mance (DMARC). These protocols can work alongside SMTP to enhance email security and authentication, but do not extend the SMTP protocol. The goal of these protocols is to prevent malicious modifications in email headers and therefore it counters attacks like email spoofing, which is used by attackers to disguise malicious emails. We view these protocols as a good first step in countering email-based social engineering attacks since they prevent forged email addresses. All of these authentication protocols are designed to supplement the SMTP protocol and use DNS records that can be configured by domain administrators. First, we address SPF, DKIM and DMARC. After that, we examine BIMI which is a new email specification and the only visual method that supports recipients in identifying legitimate email messages.

It is also worth noting that the E2EE methods mentioned in a previous subsection provide email authentication. These E2EE methods provide authentication on MUAs, whereas SPF, DKIM and DMARC are often applied on MTAs. E2EE can be used in combination with SPF, DKIM and DMARC to provide more protection against attacks such as email spoofing.

2.2.3.1 Sender Policy Framework

Sender Policy Framework, defined in RFC 7208[11], determines whether a sender is permitted to send an email on behalf of a domain. This email authentication method provides a way for receiving MTAs to verify whether hosts are authorized by the domain's administrator to send emails from that domain. In order to use SPF, a domain administrator first has to publish a policy in the DNS, also referred to as an SPF record. This SPF record includes the IP addresses or domains of outbound MTAs that are allowed to send emails from that domain. When an inbound MTA receives an email, it queries the domain that is contained in the email header "Return-path" to check for a SPF record. If there is an SPF record, then this record is checked on whether it contains the IP address of the sender's MTA. The receiving MTA then handles accordingly to the specified rules in the SPF record of the sending domain. This could for example be accepting or rejecting the email. A drawback of SPF is that it does not work if an email is forwarded. The solution to this is introduced in the next subsection.

2.2.3.2 DomainKeys Identified Mail

DomainKeys Identified Mail, defined in RFC 6376[12], is a form of email authentication using public-key cryptography and a digital signature. It enables receiving MTAs to verify whether emails are authorized and sent by the owner of the domain. DKIM is used to combat email spoofing and to protect the email's content, i.e., data integrity of the email.

The process is as follows: a signing MTA creates a digital signature, known as the DKIM signature. This signature covers selected email header fields of the email header and optionally the body of the email. Then the DKIM signature is attached to the headers of that email message. This digital signature is then verified by the receivers to check whether someone modified the email headers and if included the email body. The corresponding public key to the digital signature is published in a DNS record mapped with the sender's domain name. The receiving MTA then verifies the email by first performing a DNS lookup in order to obtain the DKIM record. This record is a line of text within the DNS record containing the public key, so that the MTA can verify the signature. Usually, end users do not see DKIM signatures because it more or less functions between the sender's MTA and the recipient's MTA. The recipient's MTA checks whether an email actually came from domain that was indicated in the headers by verifying the DKIM signature. A valid signature guarantees that the parts of the email that were covered by the digital signature have not been tampered with.

[11] https://www.ietf.org/rfc/rfc7208.txt

[12] https://www.ietf.org/rfc/rfc6376.txt

2.2.3.3 Domain-based Message Authentication, Reporting & Conformance

Domain-based Message Authentication, Reporting & Conformance [13] is an email authentication, policy and reporting protocol. It is built on top of two existing email authentication protocols: SPF and DKIM. DMARC allows domain owners to publish a policy in their DNS records. This policy, also known as the DMARC policy, contains instructions to receiving MTAs on how to enforce implemented email authentication methods. It tells MTAs to send DMARC reports to the domain that is listed as the reporting email address in the DMARC record, and how email messages should be handled if an email fails authentication checks.

DMARC combats email spoofing by helping receiving MTAs authenticating incoming emails based on the instructions in the DMARC record. If an email passes all authentication checks, then this email can be viewed as trusted and will be delivered to the recipient. However, if this is not the case, then the DMARC record specifies how that email must be handled, this could for example be: quarantining, delivering or rejecting the email.

2.2.3.4 Brand Indicators for Message Identification

Brand Indicators for Message Identification (BIMI)[14] is an email specification that allows domain owners to cooperate with mailbox providers. It is used to display a brand's logo next to authenticated email messages of that brand in the recipient's email client. BIMI aims to give authenticated senders some sense of control over how their brand is represented in emails.

It is worth noting that BIMI is by itself not a new authentication protocol. It relies on current existing authentication protocols such as DMARC, SPF and DKIM for the authentication of senders. However, it can help users to identify safe and legitimate email messages by giving a visually confirmation in the form of a brand's logo. In order to make use of BIMI, the user must be using one of the participating email providers such as Gmail[15] or Yahoo[16] and the email must pass DMARC authentication checks. The structure of BIMI is as follows (Blank et al., 2022):

- Domain owners have to fully implement the DMARC mechanism and publish their brand indicators via DNS.
- Senders need a sufficiently strict DMARC policy and ensure that emails are properly authenticated.

[13] `https://dmarc.org/`

[14] `https://bimigroup.org/`

[15] `https://www.google.com/gmail/about/`

[16] `https://mail.yahoo.com/`

- Mail box providers and mail transfer agents have to confirm the authenticity of email messages using DMARC and other authentication methods if those used. Additionally, they have to check for a BIMI record. If the email message is authentic and the logo is valid then the receiving MTA ensures that the MUA can retrieve the brand indicator by adding a header.
- Mail user agents retrieve and display the brand indicator using its policy and UI.

Mailbox providers that implement BIMI also have to implement this in their MTA and MUA.

2.3 PostGuard

PostGuard[17] is a software project developed by iHub[18] at the Radboud University. The project addresses the limited adoption of email encryption by providing a free, open source and easy-to-use add-on that allows users to encrypt and share emails along with their attachments. This add-on can be installed to email clients, currently only supporting Microsoft Outlook[19] and Mozilla Thunderbird[20].

PostGuard uses identity-based encryption and end-to-end encryption for sharing files and emails. Identity-based encryption is used to address problems such as key management and recipient authentication. Additionally, it makes the process of encrypting email and files easier so that it becomes more available to the public. The user only has to authenticate itself by scanning a QR code using the authentication app Yivi, to decrypt and read PostGuard encrypted messages. Yivi is a free, open source and privacy-friendly authentication app, offering attribute-based authentication and will be explained in more details in the subsection below. File sharing and the decryption of PostGuard encrypted emails can also be done via the website if one does not have access to the PostGuard add-on.

2.3.1 Yivi

Yivi[21] is a privacy-friendly identity management platform developed by Privacy by Design Foundation[22] and SIDN[23]. It allows users to disclose and

[17] https://postguard.eu/
[18] https://ihub.ru.nl/
[19] https://outlook.live.com/
[20] https://www.thunderbird.net/
[21] https://yivi.app/
[22] https://privacybydesign.foundation/
[23] https://www.sidn.nl/

prove ownership of specific personal information without revealing other information via the Yivi mobile application. These statements or properties of a person are referred to as attributes. The Yivi app can be viewed as a digital identity wallet holding the attributes. There are three parties involved in the usage of Yivi:

- **Issuers** are verified instances that provide attributes to users.
- **Verifiers** ask users to disclose ownership of certain attributes.
- **Users** are the ones that use the Yivi mobile application and reveal attributes to verifiers.

When a user wants to add an attribute to their wallet they must prove to an issuer that they own the attribute. A user can prove this through for example DigiD[24], verification codes or links.

A depiction of a Yivi session is shown in Figure 2.8. When a verifier requests a user to disclose ownership of attributes, a QR code is shown to the user. Then, the user scans the QR code so that it can give permission via the Yivi app to disclose its attributes to the verifier. After that, the verifier can use the issuer's digital signature to verify that the attributes have not been altered and were actually given to the user.

This platform is used by PostGuard as an authentication mechanism so that recipients can show that they are the intended recipient by proving that they possess attributes such as email addresses.

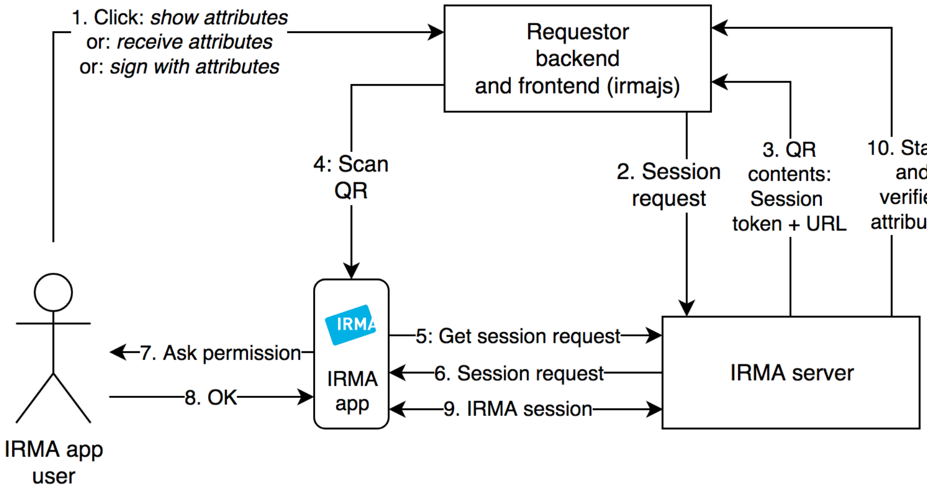


Figure 2.8: A schematic depiction of an IRMA (recently renamed to Yivi) session.
Source: Figure from IRMA (n.d.).

[24] https://www.digid.nl/

2.4 Social Engineering

In this section we will first define the term social engineering and discuss why social engineering attacks are effective. After that, we will examine various types of email-based social engineering attacks.

Social engineering is a technique that relies on the manipulation of people to obtain sensitive information, make them do certain actions or change their behavior in such a way that it benefits the attacker (ENISA, 2022; IBM, n.d.). This technique is very applicable in the digital world because of the significant usage of digital communication nowadays. Cyber attacks using social engineering can be very effective because they target the weakest link in security systems, that is humans (Mouton et al., 2014).

Mouton et al. (2014) proposed the following definition for social engineering: "The science of using social interaction as a means to persuade an individual or an organisation to comply with a specific request from an attacker where either the social interaction, the persuasion or the request involves a computer-related entity" (p. 269).

They also proposed a definition for social engineering attack: "A Social Engineering attack employs either direct communication or indirect communication, and has a social engineer, a target, a medium, a goal, one or more compliance principles and one or more techniques" (p. 269).

We use the above definitions proposed by Mouton et al. (2014) for social engineering and social engineering attacks in this thesis. We define email-based social engineering attacks based on these definitions as: "A social engineering attack that uses email as a communication medium via either direct communication or indirect communication, and has a social engineer, a target, a goal, one or more compliance principles and one or more techniques".

Targeted social engineering attacks often have a common pattern of execution with similar phases (Salahdine & Kaabouch, 2019). First, information about the target in question is gathered. Then, the attacker develops a relationship with the target. After that, the attacker exploits the obtained information about the target and executes the attack. Ultimately, exiting and leaving no traces.

Generally, social engineering attacks can be categorized into: human-based attacks and computer-based attacks. The former being carried out by interacting with the target in person. The latter utilizes devices to obtain information from targets. It is important to note that human-based attacks have their limitations since they cannot target as many people as quickly as computer-based attacks (Salahdine & Kaabouch, 2019).

These attacks have various approaches, some examples are: reverse social engineering approach, technical approach, physical approach and social approach (Salahdine & Kaabouch, 2019; Krombholz et al., 2015). The relevant approaches to this thesis will be explained very briefly, based on the work of Salahdine and Kaabouch (2019) and Krombholz et al. (2015):

- **Physical approach**
  Attacks using the physical approach are attacks where the attacker performs physical actions to obtain information about a target, a famous example is dumpster diving: searching dumpsters for valuable information.
- **Technical approach**
  Attacks that are carried out over the internet via online services or social networks gathering valuable information such as login credentials or credit card details.
- **Social approach**
  Attacks using socio-psychological techniques to deceive and manipulate the target by for example building relationships.
- **Reverse social engineering approach**
  Here an attacker attempts to make the target believe that the attacker is trustworthy with the goal that a target approaches the attacker for assistance.

In this thesis we solely focus on email-based social engineering attacks. These are computer-based, but have a human element, and mainly use a technical approach. Depending on the type of email-based social engineering attack, these are combined with either reverse social engineering, a social approach or both.

2.4.1 Email-based social engineering attacks

In this subsection we present various types of social engineering attacks that can be executed using email as a communication medium. The goal is to find similarities between the various types of attacks so that effective countermeasures can be proposed. Note that we do not consider malware such as ransomware to be a type of social engineering attack as some papers do (Salahdine & Kaabouch, 2019; Heartfield & Loukas, 2016). We define email-based social engineering attack to be the process where a target gets manipulated or deceived by interacting with the email message sent by the attacker. Hence, we do not consider the types of malware used.

2.4.1.1 Phishing

Phishing in the context of emails is a technique where an attacker sends a convincing email to one or multiple targets. Here, the email looks appearance wise similar to an email from a legitimate source and therefore, it masquerades the attacker as a trustworthy entity (McAfee, 2022). Phishing emails usually ask the target to do certain actions such as clicking a link, downloading an attachment or providing personal information such as banking details or login credentials. Links in these emails can lead the user unwarily to a fake website that is designed to look similarly to an authentic website, so that the target performs the desired actions of the attacker. This could be, for example, downloading malware (Salahdine & Kaabouch, 2019; Krombholz et al., 2015; Heartfield & Loukas, 2016). Phishing emails often contain grammatical errors and spelling mistakes and are often to be intentionally included by design. This is done to prevent detection by for example spam filters or to pick specific targets (Steinberg, 2019). Email phishing can be combined with domain spoofing, which will be explained below, to make email addresses and links within an email look more legitimate.

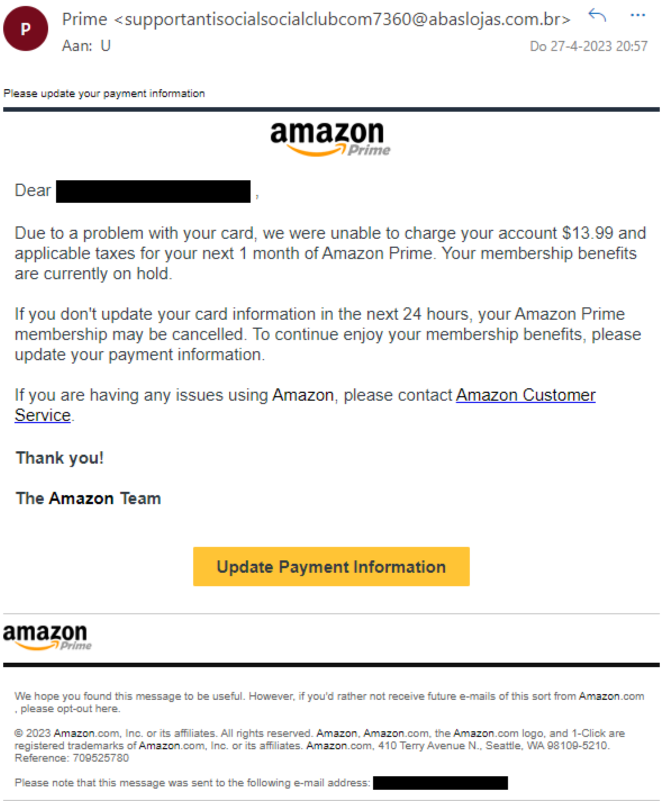
Phishing emails come in various forms and can be categorized into targeted and untargeted attacks (JCU Australia, n.d.). The former often appears as mass phishing and the latter as spear phishing. We will explain these variants in more detail below.

- **Mass phishing**
  Mass phishing is an untargeted form of phishing where an attacker sends the same email message to a large number of people, also known as mass email. The attacker hopes that one of the receivers believes that the message is legitimate and performs the attacker's desired actions e.g. providing sensitive information, opening email attachments, transferring money or visiting a malicious website via links in the email message (SALT Cyber Security, 2023). An example can be seen in Figure 2.9.

- **Spear phishing**
  Spear phishing is a targeted form of phishing where an attacker sends a customized email to a specific target, often an individual. In order to create a customized email, the attacker first obtains information about the target. These emails can include personal information or other details with the purpose that the target perceives the email as a legitimate message. The success rate of these emails are often higher than mass phishing emails (Krombholz et al., 2015). An example can be seen in Figure 2.10.

  There also exists variants of spear phishing attacks that are aimed at even more specific targets:

  - **Whaling**
    Whaling is a type of spear phishing attack that specifically targets high-profile targets (Salahdine & Kaabouch, 2019; Krombholz et al., 2015; Heartfield & Loukas, 2016).



Prime <supportantisocialsocialclubcom7360@abaslojas.com.br>

Aan: U

Do 27-4-2023 20:57

Please update your payment information

Dear [redacted] ,

Due to a problem with your card, we were unable to charge your account $13.99 and applicable taxes for your next 1 month of Amazon Prime. Your membership benefits are currently on hold.

If you don't update your card information in the next 24 hours, your Amazon Prime membership may be cancelled. To continue enjoy your membership benefits, please update your payment information.

If you are having any issues using Amazon, please contact Amazon Customer Service.

Thank you!

The Amazon Team

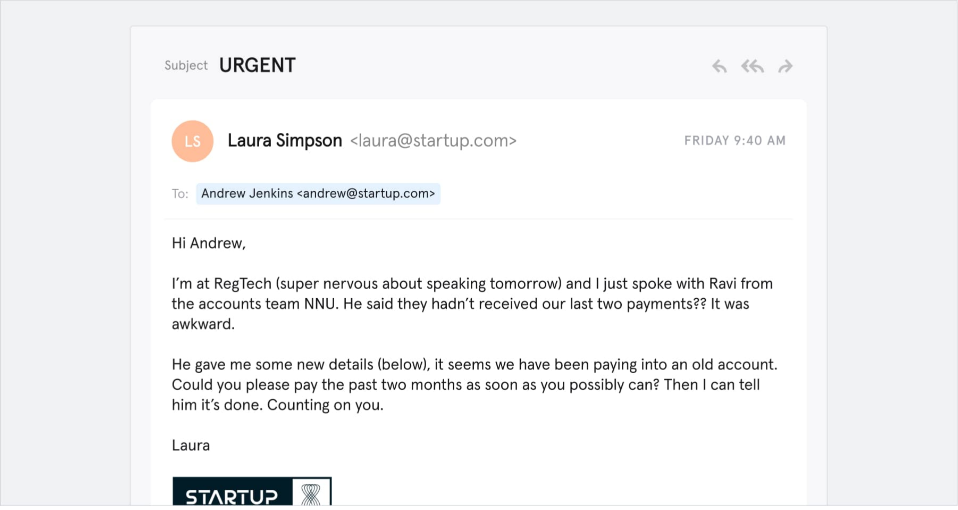
We hope you found this message to be useful. However, if you'd rather not receive future e-mails of this sort from Amazon.com , please opt-out here.

© 2023 Amazon.com, Inc. or its affiliates. All rights reserved. Amazon, Amazon.com, the Amazon.com logo, and 1-Click are registered trademarks of Amazon.com, Inc. or its affiliates. Amazon.com, 410 Terry Avenue N., Seattle, WA 98109-5210. Reference: 709525780

Please note that this message was sent to the following e-mail address: [redacted]

Figure 2.9: Mass phishing email pretending to be from Amazon. Source: Phishing email received by the author.

- **Business Email Compromise**
  Business Email Compromise (BEC) is essentially the same as whaling but here the attacker impersonates as an employee of the company to deceive high-profile targets of a company; hoping that they trust the email message and as a consequence perform the attacker's desired actions (Salahdine & Kaabouch, 2019).



Subject URGENT

Laura Simpson <laura@startup.com>

FRIDAY 9:40 AM

To: Andrew Jenkins <andrew@startup.com>

Hi Andrew,

I'm at RegTech (super nervous about speaking tomorrow) and I just spoke with Ravi from the accounts team NNU. He said they hadn't received our last two payments?? It was awkward.

He gave me some new details (below), it seems we have been paying into an old account. Could you please pay the past two months as soon as you possibly can? Then I can tell him it's done. Counting on you.

Laura

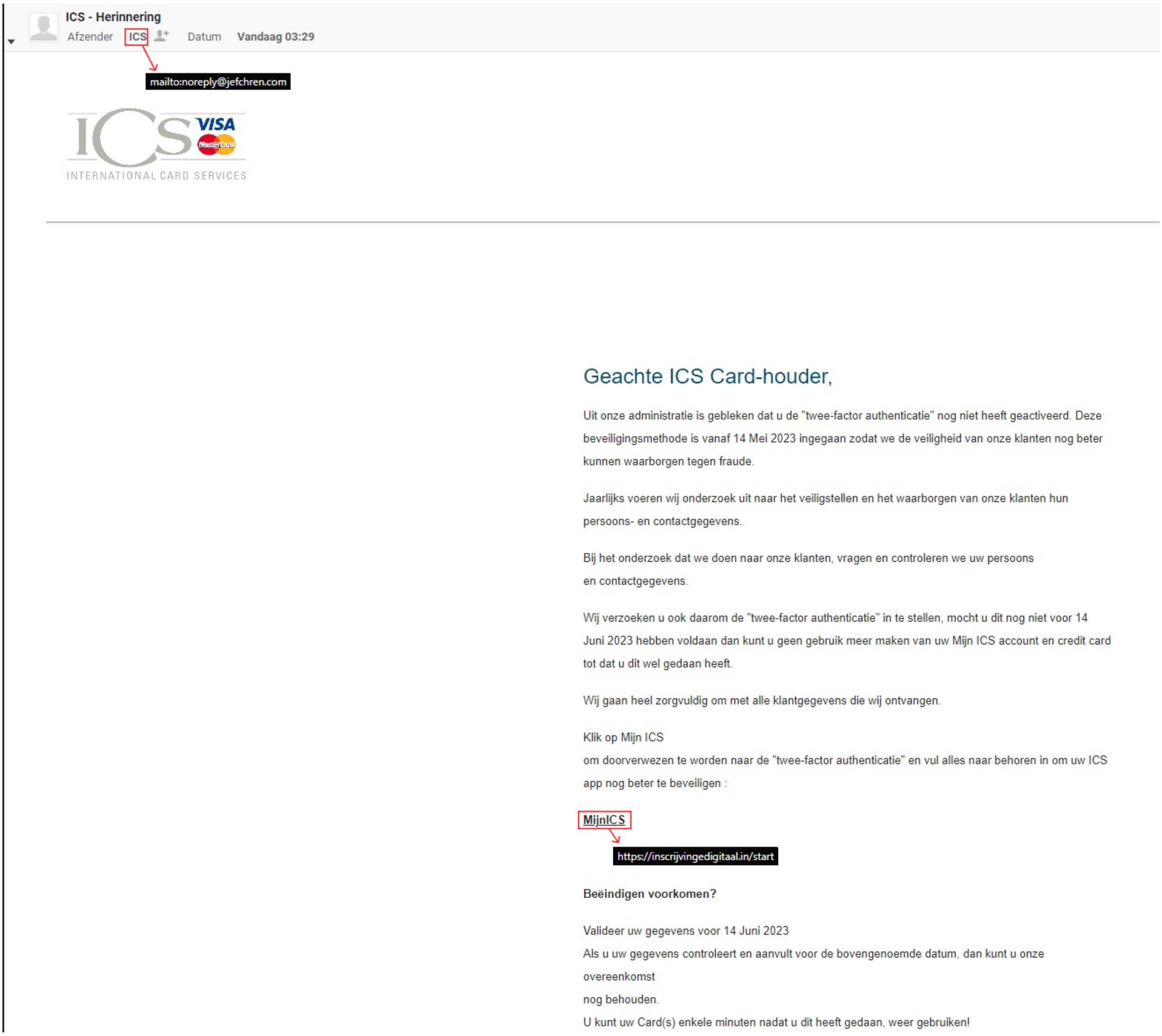
Figure 2.10: Targeted spear phishing email where the target's colleague is impersonated.
Source: Figure of Tessian (2021).

There is also another type of phishing attack technique, named clone phishing (Valimail, n.d.). This attack can be used as an untargeted or targeted attack. Here, the attacker utilizes the visual appearance of emails to manipulate and deceive victims. In this thesis, we use the terms email design and email template to refer to the visual appearance of email messages. Where we define email template as a predefined email layout consisting of either images, text or both.

Clone phishing

In clone phishing, the attacker copies an existing email template that is frequently used by a person, company or organisation and adds or modifies the email content by for example adding malicious links or attachments for malicious reasons (Eemeli, 2022). In untargeted clone phishing, the attacker clones a mail from a legitimate source and sends this email to a large number of people with the hope that one of them trusts the malicious email by familiarity. Take as an example, the clone phishing email that ended up in our organisational email account, shown in Figure 2.11.

[26] https://www.icscards.nl/



ICS - Herinnering

Afzender ICS Datum Vandaag 03:29

mailto:noreply@jefchren.com

Geachte ICS Card-houder,

Uit onze administratie is gebleken dat u de "twee-factor authenticatie" nog niet heeft geactiveerd. Deze beveiligingsmethode is vanaf 14 Mei 2023 ingegaan zodat we de veiligheid van onze klanten nog beter kunnen waarborgen tegen fraude.

Jaarlijks voeren wij onderzoek uit naar het veiligstellen en het waarborgen van onze klanten hun persoons- en contactgegevens.

Bij het onderzoek dat we doen naar onze klanten, vragen en controleren we uw persoons en contactgegevens.

Wij verzoeken u ook daarom de "twee-factor authenticatie" in te stellen, mocht u dit nog niet voor 14 Juni 2023 hebben voldaan dan kunt u geen gebruik meer maken van uw Mijn ICS account en credit card tot dat u dit wel gedaan heeft.

Wij gaan heel zorgvuldig om met alle klantgegevens die wij ontvangen.

Klik op Mijn ICS
om doorverwezen te worden naar de "twee-factor authenticatie" en vul alles naar behoren in om uw ICS app nog beter te beveiligen :

MijnICS

https://inscrijvingedigitaal.in/start

Beëindigen voorkomen?

Valideer uw gegevens voor 14 Juni 2023
Als u uw gegevens controleert en aanvult voor de bovengenoemde datum, dan kunt u onze overeenkomst
nog behouden.
U kunt uw Card(s) enkele minuten nadat u dit heeft gedaan, weer gebruiken!

Figure 2.11: A Dutch clone phishing email pretending to be from ICS[26]. The suspicious elements of the email are marked in red.
Source: Phishing email received by the author.

In the targeted variant, the attacker first has to make sure that the target already has received an email using the same design from the legitimate sender. The attacker can even send a forged email briefly after a legitimate email, so that the attacker can justify sending the forgery by for example saying that the ‘legitimate’ message contained the wrong content or that one forgot to add some attachments, text or links (Eemeli, 2022). According to Digital Check (2021), it is very easy for an attacker to create an email message or web page that looks similar to its legitimate version. The difficulty of successfully executing targeted clone phishing lies in obtaining and reading emails that are delivered to the target’s mailbox. This is one way to forge convincing malicious email (Eemeli, 2022). Therefore, clone phishing exploits the target’s feeling and trust in familiar email messages (Eemeli, 2022).

There are various techniques to make email messages look appearance wise similar to email messages from legitimate sources while avoiding detection by security solutions. A couple of those are, and will be explained below: brand impersonation with procedurally-generated graphics, text padding with invisible characters, zero-point font obfuscation and victim-specific URI (Microsoft 365 Defender Threat Intelligence Team, 2021).

- **Brand impersonation with procedurally-generated graphics**
  Here, an attacker uses procedurally-generated graphics to copy logos or branding of companies or organisations to bypass detection. This can be for example done using HTML tables (Microsoft 365 Defender Threat Intelligence Team, 2021).

- **Text padding with invisible characters**
  An attacker inserts invisible Unicode characters between words to bypass detection of automated email security analysis. Some of these Unicode characters are almost unnoticeable or not even visible to the user (Microsoft 365 Defender Threat Intelligence Team, 2021).

- **Zero-point font obfuscation**
  Zero-point font obfuscation is a technique used to bypass security solutions where one inserts irrelevant hidden words into the body of the email message. These words can be hidden by changing their font size to zero and is used to bypass spam filters or filters based on natural language processing (Microsoft 365 Defender Threat Intelligence Team, 2021).

- **Victim-specific URI**
  Victim-specific URI is a method to transmit information about a target where the information is then used to create dynamic content. First, the attacker creates a custom Uniform Resource Identifier (URI) that is a unique sequence of characters referring to a resource on the internet using the location, name or both. Then, the custom URI can pass information about the target to a malicious website of the attacker. This is done so that a malicious website seems more legitimate to the target by personalizing its content based on the target.

2.4.1.2 Domain spoofing

Earlier in the preliminaries we have addressed what domains are and their usage. In short, it is important to know that domain names are used in URLs and email addresses. Therefore, when a recipient looks at the email headers such as: `From:` and `To:`, it sees the email addresses of the recipient and sender. Additionally, links can be found in the content of an email message. These links are referred to as URLs and mainly consist of a domain name. These domain names can be spoofed or created in such a way that it resembles legitimate domain names. This is known as domain spoofing.

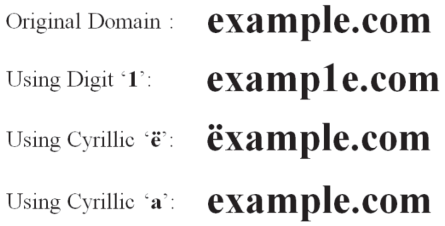
Domain spoofing is a type of attack where the attacker impersonates a legitimate or fake domain to deceive targets into believing that they are interacting with a legitimate email or URL (Bolster, 2022; CloudFlare, n.d.-b). There are various types of domain spoofing attacks:

- **Typosquatting**
  Typosquatting is a type of domain spoofing attack where an attacker uses a legitimate domain name that looks very similar to the domain name of the email address or URL that is to be impersonated, but with a slight variation (PowerDMARC, n.d.). The attacker uses misspellings or typos of the legitimate domain name, such as replacing letters with similar looking letters or numbers. It is also possible that a different top-level domain is used (Bolster, 2022). Take as an example the following legitimate URL: `https://libguides.ru.nl/`. And the following typosquat variant: `https://libguide.ru.nl/`[27].

  - **Homograph attacks**
    Homograph attacks are a sub-form of typosquatting. The difference is that one replaces characters by other characters that look similar. For example, replacing Latin characters with similar looking characters such as digits or characters from other alphabets such as Greek or Cyrillic (CloudFlare, n.d.-b). An example can be seen in Figure 2.12. Some of these letters look nearly identical to each other to an untrained eye.

- **Email spoofing**
  Email spoofing is an attack where the attacker sends an email from a fake email address that appears to be legitimate to the recipient. Spoofing email addresses is possible by the fact that default SMTP includes no security mechanisms to authenticate the identity of a sender, thus making it possible for an attacker to send an email message using an arbitrary sender email address by modifying the `MAIL FROM` field in the SMTP envelope (Bolster, 2022; CloudFlare, n.d.-b; Hu & Wang, 2018) or the email header.

[27] Note that the letter "s" is missing when compared to the legitimate URL.



Original Domain : **example.com**

Using Digit '**1**': **examp1e.com**

Using Cyrillic '**ё**': **ёxample.com**

Using Cyrillic '**а**': **example.com**

Figure 2.12: Various homograph domain names of the domain "example.com".
Source: Figure from Sawabe et al. (2019).

There are also email spoofing methods that spoof the sender's display name. We summarized methods of Dedenok (2021) and provided some examples:

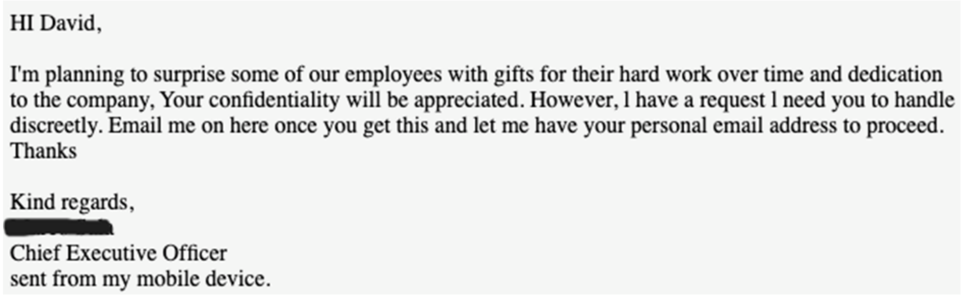
- **Display name spoofing**
  In display name spoofing, the attacker changes the name of the sender to the name of the person or company that the attacker is impersonating. This name is shown in the `From:` field in an email message and is often located before the associated email address. Note that in this type of attack, the attacker does not spoof or modify the domain name of the email addresses (Dedenok, 2021).
  - **Ghost spoofing**
    Ghost spoofing is a variant of display name spoofing where the attacker modifies the display name into the following form (name + email address). This is done on the assumption that a target only looks at the display name of an email message and not the email addresses. Then, an email may be perceived as legitimate (Dedenok, 2021).
    An example of ghost spoofing would be:
    An attacker with the email address "`fake@info.com`" impersonates a person named "John Doe" with the email address "`real@info.com`". Then, the attacker modifies its display name to: "`John Doe real@info.com`".
  - **Active directory spoofing**
    Active directory spoofing is a form of display name spoofing. Here, the attacker modifies the display name to the name of the person being impersonated and sends the mail using an email address that features this name (Dedenok, 2021).
    An example of active directory spoofing would be:
    An attacker impersonates a person named "John Doe". The attacker sends an email from an email address containing the name of the impersonated person:

"`John.Doe@info.com`". Then, if a recipient does not know the legitimate email address of John Doe and receives an email from "John Doe <`John.Doe@info.com`>" it may perceive this as legitimate, while John Doe's email address in reality is "`J.D@test.com`".

- **URL spoofing**
URL spoofing is a technique where a fake URL looks legitimate by for example typosquatting/homograph attacks (CloudFlare, n.d.-b). These malicious URLs can be included in emails such as phishing emails so that a target gets tricked into clicking the malicious URL. (CloudFlare, n.d.-b).

2.4.1.3 Pretexting

Pretexting attacks are attacks where the attacker creates fake and convincing scenarios or pretexts in order to convince the target into giving up valuable information (Salahdine & Kaabouch, 2019; Watson, 2014). Attackers use information of public websites or phone books to carry out their attack. The pretexts could vary from getting a job, helping a friend to parcel delivery (Salahdine & Kaabouch, 2019). An example can be seen in Figure 2.13.



HI David,

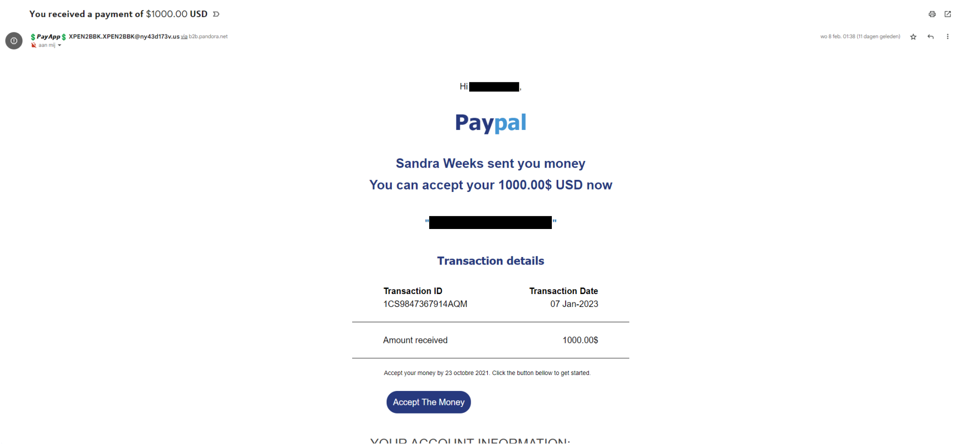
I'm planning to surprise some of our employees with gifts for their hard work over time and dedication to the company, Your confidentiality will be appreciated. However, l have a request l need you to handle discreetly. Email me on here once you get this and let me have your personal email address to proceed. Thanks

Kind regards,
[redacted]
Chief Executive Officer
sent from my mobile device.

Figure 2.13: Pretext email with a scenario concerning planning a surprise. Source: Figure from Sainio (2022).

2.4.1.4 Baiting

Baiting is a type of attack where an attacker invites a target to click a link or download an email attachment by making false promises such as offering free goods or items to the target (Salahdine & Kaabouch, 2019). An example can be seen in Figure 2.14.
The described aforementioned attacks are not the only types of social engineering attacks using email as a communication tool. However, most of the other attacks are a variation or use the same principles as the discussed attacks.



You received a payment of $1000.00 USD

$PayApp$ XPEN2BBK.XPEN2BBK@ny43d173v.us via b2b.pandora.net
aan mij

wo 8 feb. 01:38 (11 dagen geleden)

Hi [redacted],

Paypal

Sandra Weeks sent you money
You can accept your 1000.00$ USD now

"[redacted]"

Transaction details

Transaction ID	Transaction Date
1CS9847367914AQM	07 Jan-2023
Amount received	1000.00$

Accept your money by 23 octobre 2021. Click the button bellow to get started.

YOUR ACCOUNT INFORMATION:

Figure 2.14: Baiting email where money is offered.
Source: Phishing email received by the author.

2.4.1.5 Email-based social engineering attacks insights

Concludingly, email-based social engineering attacks rely on a target believing that a malicious email is legitimate i.e. coming from a person, company or organisation that the target thinks it comes from and are therefore trusted by the target. This is mainly carried out by modifying or spoofing email headers so that either display names, domain names or both, look similar to their legitimate counterparts. The email body can also be designed in such a way that it looks similar appearance wise to email messages that are sent from the person, company or organisation that is being impersonated. Furthermore, the links in the email message can be spoofed so that the target thinks that the URL or the links are legitimate, therefore trustworthy and harmless. Lastly, an email could include harmful attachments such as fake invoices or files that secretly contain malware.

Based on the aforementioned techniques and definitions, we see that the attacker typically want the victim to perform one of the actions as listed:

- **Clicking a malicious link**
  When a malicious link in an email message is clicked, it usually redirects the target to a fake malicious website that might look legitimate to the target. Here, the attacker wants the target to perform certain actions such as giving away sensitive information by entering login credentials so that these are passed to the attacker. From a personal experience, more advanced phishing websites can also ask for two-factor authentication (2FA) codes after having requested login credentials (username and password). This is done to bypass 2FA which is an extra layer of protection so that the security of accounts do not entirely depend on a username and password combination. Another possibility

is that the target unconsciously downloads and installs malware from these fake websites.

- **Opening or downloading malicious email attachments**
  By opening or downloading and running email attachments one can get their device infected with malware. Furthermore, some of these attachments can contain textual content such as fake invoices and therefore trick the user into performing actions such as transferring money.

- **Trusting textual email content**
  A target that is convinced that a malicious email is legitimate, can give sensitive information of themselves or someone else away. This can be done by replying to the malicious email or via a different communication medium. Additionally, the textual content can also induce the target into performing the attacker's desired actions such as transferring money. Take as an example the business email compromise on Facebook and Google from 2013 to 2015. The attackers impersonated a Taiwan-based hardware company and deceived Facebook and Google by sending forged invoices, contracts and letters. Here, the victims believed that the textual content of the email messages and attachments were legitimate. The victims transferred at least $100 million to the attackers (Baraniuk, 2017; Trend Micro, 2019).

2.5 Email user experience and user interface design

In this section we start with defining the terms user experience and user interface in emails along with the design approach that will be used in this thesis. After that we examine security indicators of email clients.

2.5.1 Email user experience

There is no general consensus on a definition for the term user experience (UX) (Law et al., 2009; Roto et al., 2011). According to Roto et al. (2011), the term user experience is often used as a synonym for other terms such as usability, user interface, interaction experience, general experience and much more. UX can also be used as an umbrella term incorporating the above mentioned or similar terms. In a user experience white paper (Roto et al., 2011), a definition for the term user experience was discussed by various UX researchers and practitioners. This ultimately led to the conclusion that there exists no single definition for UX that addresses all perspectives of UX.

The definition of UX provided by the International Organization for Standardization (ISO) in ISO 9241-210: "User's perceptions and responses that result from the use and/or anticipated use of a system, product or service." (Law et al., 2009; *ISO 9241-210:2019*, 2019) will be used as a starting point in this thesis. Therefore, UX designers have to a certain extent, the ability to control how a product, system or service should behave. However, they cannot control how a user feels and interacts with the product, system or service. This definition is provided with two notes: the first note indicates that responses and perceptions of users are expressed by emotions, beliefs, preferences, perceptions, comfort, behaviours, and accomplishments before, during and after using the system, product or service (*ISO 9241-210:2019*, 2019). The second note states that UX is a consequence of brand image, presentation, functionality, system performance, interactive behaviour, and assistive capabilities of a system, product or service. It also follows from the context of use and user's internal and physical state based on prior experiences, attitudes, skills, abilities and personality (*ISO 9241-210:2019*, 2019).

In this thesis and in the context of emails, we define user experience as how the user feels, behaves and experiences before, after and during interaction with email messages and the email client. There are various aspects of emails and email clients that can influence a user's experience, take as an example how the email is displayed or the functionality and accessibility of the email client. Note that these are just a few examples, many more aspects can have an effect on the user's experience.

From the first paragraph, it follows that user experience design is related to terms such as usability and user interface. This is why we consider factors such as visual design, usability and accessibility in email user experience design, as they can affect how a user interacts with emails. If we consider user experience in emails in combination with safety indicators, then we define UX as the overall experience that a user has when receiving and interacting with an email or email client including these safety indicators.

2.5.2 Email user interface

The user interface (UI) is the area where the user interacts and communicates with a system. The goal is to have an effective UI that provides an easy and intuitive experience for the user (Indeed Editorial Team, 2022). The interaction and communication between the system and the user is mainly accepting the end user's input and displaying output to the end user. UI is defined by ISO as "All components of an interactive system (software or hardware) that provide information and controls for the user to accomplish specific tasks with the interactive system." (*ISO 9241-210:2019*, 2019).

User interface design is the process of creating user interfaces where the aesthetic design of all visual elements are prioritized (Indeed Editorial Team, 2022). In our opinion, a user interface is well designed, if the interaction with the computer becomes unnoticeable i.e. a user can complete their tasks flawlessly. On top of that, we would say that a user interface should be reliable, functional and pleasant. The end user should be able to interact with the application and accomplish their goals with minimum effort. Also, accessibility plays a crucial role in UI design, and should be kept in mind during the design phase because people with low vision or other disabilities should not be neglected. One must make sure that all end users can interact with user interfaces without a lot of difficulties. Additionally, the increasing usage of devices with smaller screens supports this notion.

According to a previous study by Kurosu and Kashimura (1995), it is also important to make UI more aesthetically appealing because end-users will experience it as more usable. This is named the aesthetic–usability effect.

Important elements of an UI to the end user can be categorized into four categories (Hannah, 2021; Usability.gov, 2014):

1. **Input controls**
   Input controls enables end users to provide input to the computer.

2. **Navigation Components**
   Navigational components assist the end users with navigating through for example applications.

3. **Informational Components**
   Informational components are used to give information to end users. Some examples are: notifications, icons and tooltips.

4. **Containers**
   Containers are used to group related content, components or other groups together so that everything is organized and clear.

Six relevant important UI design principles to this thesis are mentioned below. For every principle, we considered whether they can be applied on a visual countermeasure presented in email clients.

- **Simplicity, clarity and intuitiveness**
  It is often said that good UI design is invisible (Hannah, 2021). This is because an 'invisible' user interface does not distract a user from performing its tasks (Hannah, 2021). A simple, intuitive, clear and straightforward interface can be used more quickly and effectively (Hannah, 2021). It also prevents confusion and motivates the user to keep interacting.

  Next to that, one has to keep the content and visual design focused on the essentials. This is because irrelevant information competes with relevant information and distracts the user negatively (Indeed

Editorial Team, 2022; Nielsen, 1994/2020). The UI has to support the user's goals by prioritizing relevant content and features.

- **Control** (Nielsen, 1994/2020)
  Users should feel or be in control of a product, system or service. This can be achieved by clearly giving users options on how to proceed instead of forcing an action. These should be informed to the user by using terms that they can understand. Notifications can be used to display statues of actions or to let a user know what to expect from performing an action. One has to make sure that the product, system or service clearly communicates its current status to the user using UI elements, so that the user is informed about the current and next state.

- **Error prevention** (Nielsen, 1994/2020)
  A user interface has to be designed in such a way that it minimizes user errors. A well designed UI can prevent these errors from happening by eliminating error-prone conditions or by presenting users with confirmation options before carrying out an action. Errors can be classified into two types: mistakes and slips. Slips are unconsciously made errors caused by a lack of attention. Mistakes are consciously made errors by a dissimilarity between the user's mental model and the design of the UI.

- **Match between system and the real world** (Nielsen, 1994/2020)
  User interfaces should be designed using concepts, words and phrases that the user understands. Email clients are used by the masses, and therefore jargon should be avoided. This is because unfamiliarity can cause confusion and lead to the wrong actions. Using plain language supports the user's feeling of being in control over a product, system or service and it also reduces the chance of errors.

- **Prevent insecure behaviour** (Brandon et al., 2022)
  User interfaces of security features in security-enhancing software should be distinctive from "standard" user interfaces without security features. This helps users to prevent users from unintentionally performing insecure behaviour by a lack of feedback caused by too similar looking user interfaces or insufficient status updates.

- **Explain the intended secure behaviour** (Brandon et al., 2022)
  UI of security mechanisms should support the user in performing the intended secure behaviour by providing enough information. If not, then it is possible that one does not use the security-enhancing software as intended. This can be caused by for example, a lack of knowledge or having doubts about the software and therefore result in insecure behaviour.

User interfaces are the only communication channel between a user and a computer system, making the design of user interfaces very important. An end user decides its actions based on the UI, so if the UI is not designed well enough, the user may perform undesired actions (Malisa, Luka, 2017). For that reason, user interface design plays a crucial role in informing and helping the users in identifying malicious email messages. By designing and testing user interfaces along with user experience, one can improve the UI such that the end user understands and pays attention to what is shown.

2.5.3 Mobile-first design

Mobile-first design is a design philosophy that aims to improve user experiences of mobile device users by starting the design process with these users in mind (Morales, 2021). This is done since the design of the smallest poses the most restrictions. The mobile design essentially becomes the core UX and UI design because they often only contain the essential features and can be adapted for other devices.

In this thesis, this design philosophy will be applied to ensure that email elements such as security indicators or warnings are properly optimized for mobile users, who now represent a large portion of email users (van Rijn, 2023). Therefore, when one designs countermeasures against email-based social engineering attacks, these also have to be effective in making mobile device users less susceptible. If one starts with designing countermeasures for other devices such as laptops, it might find out later that the countermeasures are ineffective for mobile devices. Or even worse, some of the designs might not even be adaptable to mobile devices due to accessibility problems. Hence, we use the mobile-first design approach because a large portion of the users access their email via their mobile phone (van Rijn, 2023).

2.5.4 Visual security information

In this thesis we use visual security information as an umbrella term to cover visual security indicators, alerts and warnings. These are used to make users more aware of potential dangers and to assist them in taking informed actions (Petelka et al., 2019). According to Stojkovski (2022), security indicators exist to inform and influence the behaviour of users by the results of a security analysis. It also assist users by reminding them of the associated consequences.

Visual security indicators are often displayed as an icon which are small images representing a concept, thought or message. Take as an example the HTTPS indicator in the address bar of one's web browser. The icon is a padlock or shield if a website has a valid SSL/TLS certificate and gives the user to a certain extent assurance of security. If a website does not have a valid certificate, then an icon is shown that represents danger, see Figure

2.15. Security indicators can be distinguished into positive and negative indicators (Hunt, 2018). Positive visual security indicators reassure the user that something is secure. Whereas, negative visual security indicators tell the user to proceed with caution. This is not the case for warnings and alerts, these are always designed to be 'negative' and are used to draw the user's attention since caution is required.

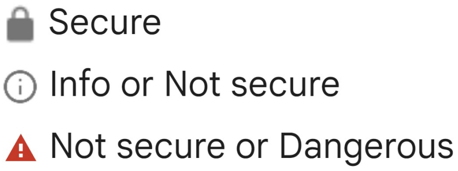


Figure 2.15: The visual security indicators of Google Chrome.
Source: Figure from Google (n.d.).

According to Hunt (2018), positive security indicators are nowadays misleading and useless. However, it is important to note that these comments were mainly about security indicators in web browsers that indicate whether a website has a valid certificate. Also, the following was said by Hunt (2018): "Positive security indicators are readily obtainable or spoofable, but nobody ever wants to show a negative indicator on a legitimate site!" (The Value of Negative Indicators, para. 3). This implies that one has to stop looking for positive visual indicators as a sign of trustworthiness. One should use the approach where exceptions are flagged with negative security indicators, and be educated on how to approach these sites with caution.

In the context of emails, visual security elements are often designed to be shown on the UI of email clients. The study of Hu and Wang (2018) examined security indicators in email clients in 2018. They found out that very few email clients had implemented visual security indicators to warn users about malicious emails. Attackers could even trigger misleading UI elements to make malicious forged emails look more legitimate.

From the work of Laughery and Wogalter (2014), we can derive that a number of design factors influence warning effectiveness in the area of attention and compliance. Namely, size, placement, color/contrast, signal word, presence of a graphic, message length and interactivity influence attention. A users' compliance can be influenced by the presence of graphics and the explicitness of content.

Chapter 3

Related work

In this chapter we compare and discuss the findings of previous related works so that interesting elements can be obtained and considered in the proposal of new effective countermeasures. We start with reviewing research about visual security indicators in a more general theme which is then followed by previous research on countering social engineering attacks on user interface level. After that, we examine papers that discuss the effectiveness of security warnings in a broader field. Lastly, we specify the points we take into consideration in coming up with a countermeasure against email-based social engineering attacks.

3.1 Visual security indicators

In this section we take a look at different types of visual security indicators and their effectiveness.

The study of Stojkovski and Lenzini (2020) examined four privacy icons by their shape and color to determine their effectiveness in conveying specific security messages in the context of a secure emailing system. They found out that these visual security indicators can actually hinder instead of helping users with detecting insecure situations. The reason for this is that certain indicators are misunderstood by users and therefore lead to insecure behaviour. Especially novice users have a difficulty with understanding what information is conveyed by the icons.

The results of the study of Dhamija et al. (2006) illustrate that security indicators are not effective for a large fraction of users; the indicators were not noticed or not understood by many participants. Some participants even mentioned that they found the padlock icon to be more legitimate if it was displayed on the web page itself than in the user interface of the web browser. This is a problem because attackers can conceal their malicious sites by including a positive security indicator on their site, since these are perceived as more legitimate than the web browser's security indicators. Additionally, some participants paid more attention to the designs of the website rather than the HTTPS indicator. The authors suggested that it is probably more important to alert users when conditions are not to be trusted, so security indicators should not only appear under trusted conditions.

Felt et al. (2016) proposed a new set of browser security indicators that are more easily understandable for users. According to the authors, security indicators face the following design constraints:

- Indicators need to be scaled down for devices with smaller screens.
- The icon shape has to be sufficient for communicating the level of risk, in order to ensure accessibility.
- The meaning of indicators have to be taught to users.

On top of that, a study by Volkamer et al. (2016) revealed that users misplace trust in HTTPS indicators.

Schechter et al. (2007) found out that positive security indicators are ignored by users and that negative indicators are perceived as more serious. From this research one could say that positive security indicators such as HTTPS and site authentication images, also known as SiteKey, deem to be ineffective because users do not notice or do not care that these are present.

Summary

Concludingly, we have seen that security indicators can be confusing for users (Stojkovski & Lenzini, 2020; Dhamija et al., 2006; Felt et al., 2016). This could result in users misusing positive security indicators by for example misplacing trust (Volkamer et al., 2016). Additionally, it does not help that most of the visual security indicators consist of a single icon, meaning that they can be interpreted differently by various users. Lastly, the ignorance of users has to be addressed (Schechter et al., 2007).

3.2 Counteracting social engineering attacks on user interface level

In this section we examine papers that give us more insight on how both social engineering attacks and email-based social engineering attacks can be prevented via user interface design. So that more effective countermeasures against email-based social engineering attacks can be proposed in this thesis.

Datta et al. (2021) wrote a paper on whether auditory representations of cyber security threat indicators could be used to warn users more effectively about cyber threats such as phishing. These auditory representations, also referred to as sonifications in the paper, can help visual impaired users and users that get overwhelmed by visual and textual security warnings notice security threats more easily. They have conducted tests with sighted participants to determine the usefulness of sonification. The user was alarmed through sonification when a URL was considered suspicious, and was determined by analysing URL heuristics in real-time. Also, when an unsafe website is visited, i.e, with an invalid or expired certificate, the sound alert is played. The results of the tests suggested that it is feasible to develop sonifications of warnings which users can understand with minimal experience. These sonifications help drawing a user's attention so that they can make a more informed decision. It is important to note that sonifications are not meant to replace visual warnings.

Similarly, Cooper et al. (2021) introduced the concept of visual alerts and warnings with haptics and audio to reduce phishing susceptibility on mobile devices. The results indicated that the combination of audio, haptic alerts and visual warnings potentially lower phishing susceptibility in emails. The work has shown that audiovisual warnings assisted participants in noticing phishing emails quicker and more easily than without audiovisual warnings. Lastly, the paper suggests that one could experiment with link analysis and hovering ability for future research of audio, visual, haptic alert and warning technology.

The work of Egelman et al. (2008) examines the effectiveness of warnings against phishing attacks. The following is recommended by Egelman et al. (2008) for improving the design of phishing warnings and alerts:

- Phishing warnings and alerts need to be designed to interrupt the task of the user. Active warnings ensures that the user's attention is switched to the warning since passive warnings do not interrupt the user's task. The effect of passive warnings is similar to providing no warnings.

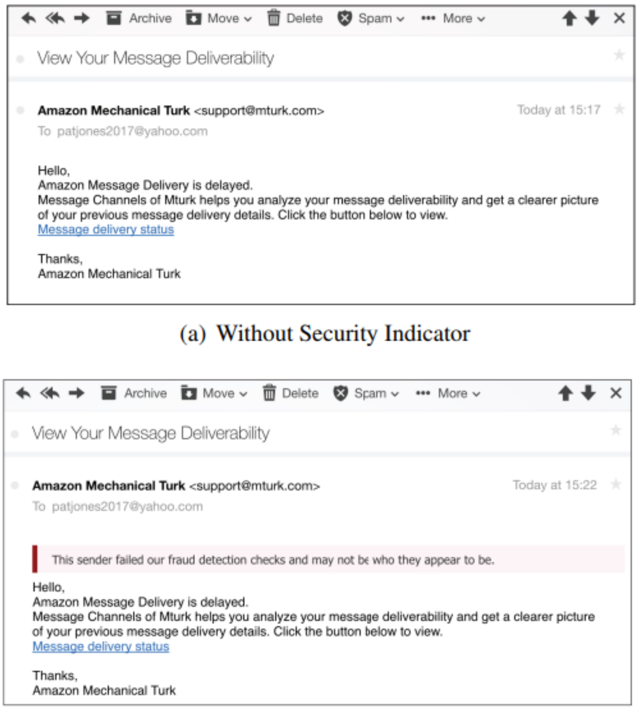
- Warnings and alerts require clear options on how to proceed instead of only displaying textual information.

- Warnings and alerts must be designed such that one can only proceed to the phishing website after reading the warning message.

- Phishing warnings and alerts need to be distinguishable from less serious warnings and only be used when there is a clear danger to prevent habituation. Introducing dynamic warning designs may help combat habituation.

- Warnings and alerts need to urge danger to the phishing site so that the user does not trust any phishing websites.

The work of Hu and Wang (2018) revealed that there exists a concerning gap between server-side spoofing detection and the protection of users. They found out that most of the email providers allowed forged emails to end up in the inboxes of users, while having the necessary protocols to detect spoofing. On top of that, a large part of the email providers do not have any warning mechanisms to inform users about forged emails, especially on mobile applications. They found out in a user study (N=488) that when a security indicator was not presented (control group), 48.9% of the control group opened and clicked the malicious URL in the spoofed email. Whereas, 37.2% clicked on the link when a security indicator was present and was consistently positive for users of different demographics. The emails that were used can be seen in Figure 3.1. Therefore, the study shows that email services with security indicators have a positive impact on reducing risky user actions when encountering phishing attacks, but cannot eliminate the risk. The paper promotes the adoption of SMTP security extensions and the development of effective security indicators for web and mobile email interfaces.



View Your Message Deliverability

Amazon Mechanical Turk <support@mturk.com> Today at 15:17
To patjones2017@yahoo.com

Hello,
Amazon Message Delivery is delayed.
Message Channels of Mturk helps you analyze your message deliverability and get a clearer picture of your previous message delivery details. Click the button below to view.
Message delivery status

Thanks,
Amazon Mechanical Turk

(a) Without Security Indicator

View Your Message Deliverability

Amazon Mechanical Turk <support@mturk.com> Today at 15:22
To patjones2017@yahoo.com

This sender failed our fraud detection checks and may not be who they appear to be.

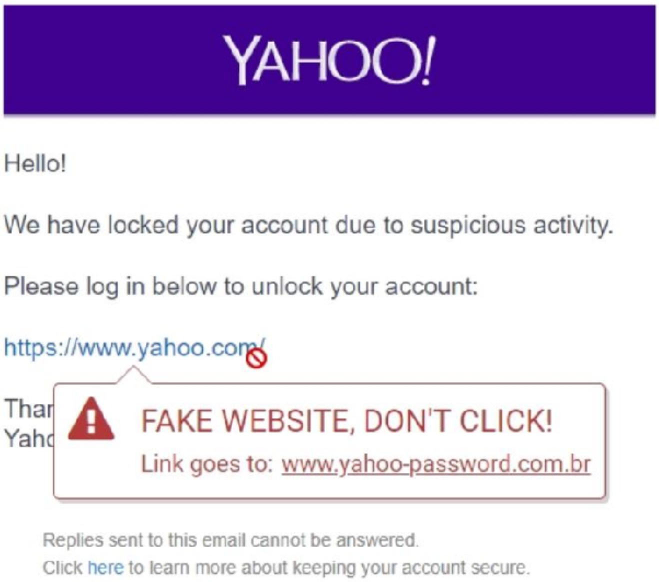
Hello,
Amazon Message Delivery is delayed.
Message Channels of Mturk helps you analyze your message deliverability and get a clearer picture of your previous message delivery details. Click the button below to view.
Message delivery status

Thanks,
Amazon Mechanical Turk

(b) With Security Indicator

Figure 3.1: The phishing email that is used in the work of Hu and Wang (2018).
Source: Figure of Hu and Wang (2018)

Petelka et al. (2019) found out that email clients have limited support in assisting a user with checking a link's URL for maliciousness before clicking. Additionally, it revealed that warnings about suspicious emails are very unspecific. The research measured the effects of moving phishing warnings closer to the suspicious link in the email (link-focused warnings). Their controlled online experiment showed that placing a warning near a link is more effective at reducing the click-through rate than a warning in the form of a banner at the top of an email. Hover warnings had little effect on click-through rates. Ultimately, the warnings that forced attention by disabling the original link were the most effective, and can be seen in Figure 3.2.



Hello!

We have locked your account due to suspicious activity.

Please log in below to unlock your account:

https://www.yahoo.com/

Than
Yaho

FAKE WEBSITE, DON'T CLICK!

Link goes to: www.yahoo-password.com.br

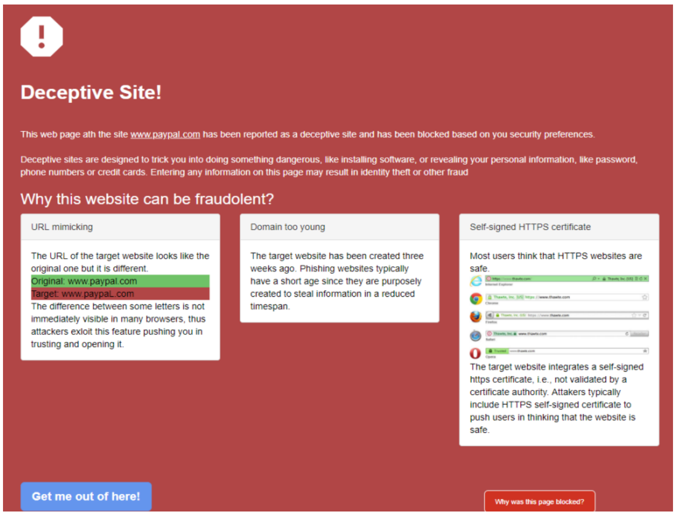
Replies sent to this email cannot be answered.
Click here to learn more about keeping your account secure.

Figure 3.2: The link-focused phishing warning that forces attention.
Source: Figure of Petelka et al. (2019).

The study of Junger et al. (2017) investigated whether priming or handing out a warning leaflet prevented users from disclosing sensitive information in a social engineering attempt. The questions were such chosen that a potential phisher could use the gathered information to send a convincing spear phishing email to the participant. An interesting finding is that the warning did not influence the degree of disclosure, in one situation the warning even worked oppositely.

According to Min (2006), the user interface level is where phishing should be solved. The paper investigates two major approaches in making user interfaces more understandable. The first approach: showing the system model to the user which used by for example anti-phishing toolbars and security indicators of browsers. This approach is not effective at preventing phishing. The second approach: telling the user's intentions to the system at data submission and letting the system check whether the actual submission corresponds to the user's intention. This approach is used by their new anti-phishing solution: Web Wallet. User studies in this work have shown that it is an effective and promising anti-phishing solution. We think that the result of the first approach is in line with a different research of Wu et al. (2006) stating that security toolbars are ineffective at combatting phishing.

The paper of Aneke et al. (2020) proposes an intelligent warning message mechanism against phishing attacks, as can be seen in Figure 3.3. However, the proposed warning message mechanism has not been tested, thus it is not known whether it is effective at countering phishing attacks. The design process could be useful to thesis in consideration with other related works.



Deceptive Site!

This web page ath the site www.paypal.com has been reported as a deceptive site and has been blocked based on you security preferences.

Deceptive sites are designed to trick you into doing something dangerous, like installing software, or revealing your personal information, like password, phone numbers or credit cards. Entering any information on this page may result in identity theft or other fraud

Why this website can be fraudolent?

URL mimicking

The URL of the target website looks like the original one but it is different.
Original: www.paypal.com
Target: www.paypaL.com
The difference between some letters is not immediately visible in many browsers, thus attackers exloit this feature pushing you in trusting and opening it.

Domain too young

The target website has been created three weeks ago. Phishing websites typically have a short age since they are purposely created to steal information in a reduced timespan.

Self-signed HTTPS certificate

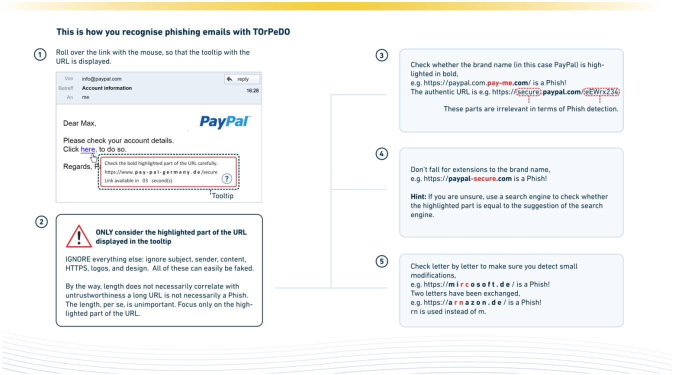
Most users think that HTTPS websites are safe.

The target website integrates a self-signed https certificate, i.e., not validated by a certificate authority. Attakers typically include HTTPS self-signed certificate to push users in thinking that the website is safe.

Figure 3.3: Prototype of the intelligent warning message for phishing attacks.
Source: Figure of Aneke et al. (2020).

A study by Volkamer et al. (2016) examined the effectiveness of checking URLs in the address bar of a web browser to combat phishing. They have tested this with the usage of URL pruning i.e. simplifying the displayed URL. The result of the study is that one cannot rely on people checking URLs. Phish detection does increase significantly when the user is provided with a hint to check the URL and URL pruning is used. Additionally, URL pruning can be useful in countering malicious URLs on devices with limited screen space. This is because URL pruning counters attackers from creating very long URLs so that users with limited screen space cannot see the actual domain of the URL.

Volkamer et al. (2017) proposed a different anti-phishing solution against malicious URLs, named TORPEDO. TORPEDO provides tooltips that highlights the domain of the actual URL of the link. The tooltips of TORPEDO can be seen in Figure 3.4. It also delays the activation of the link, so that the user inspects the URL before they click. These tooltips assist people in judging embedded links in emails. They found out that TOR-



This is how you recognise phishing emails with TOrPeDO

1. Roll over the link with the mouse, so that the tooltip with the URL is displayed.

Von info@paypal.com
Betreff **Account information**
An me

Dear Max,

Please check your account details.
Click here, to do so.

Regards, P

Check the bold highlighted part of the URL carefully.
https://www. **pay-pal-germany.de** /secure
Link available in 03 second(s)

2. **ONLY consider the highlighted part of the URL displayed in the tooltip**

IGNORE everything else: ignore subject, sender, content, HTTPS, logos, and design. All of these can easily be faked.

By the way, length does not necessarily correlate with untrustworthiness a long URL is not necessarily a Phish. The length, per se, is unimportant. Focus only on the highlighted part of the URL.

3. Check whether the brand name (in this case PayPal) is highlighted in bold,
e.g. https://paypal.com.**pay-me.com**/ is a Phish!
The authentic URL is e.g. https://secure.**paypal.com**/eEWrx234

These parts are irrelevant in terms of Phish detection.

4. Don't fall for extensions to the brand name,
e.g. https://**paypal-secure.com** is a Phish!

Hint: If you are unsure, use a search engine to check whether the highlighted part is equal to the suggestion of the search engine.

5. Check letter by letter to make sure you detect small modifications,
e.g. https://**mircosoft.de** / is a Phish!
Two letters have been exchanged,
e.g. https://**arnazon.de** / is a Phish!
rn is used instead of m.

Figure 3.4: Just-in-time, just-in-place, trustworthy tooltips of TORPEDO as shown in the top-left corner.
Source: Figure of Volkamer et al. (2017)

PEDO performed significantly better in helping users detecting phishing mails and identifying legitimate emails than the standard banners used by Thunderbird or other webmail clients. It is worth noting that TORPEDO is an AddOn for Thunderbird, meaning that users need to be aware of its existence and install the AddOn themselves. This could limit its adoption rate.

Summary

To conclude this section, we have seen that anti-phishing toolbars are not effective at preventing phishing attacks (Min, 2006; Wu et al., 2006). In addition, multiple works (Petelka et al., 2019; Egelman et al., 2008; Junger et al., 2017) found out that warnings deem to be ineffective at preventing users from being susceptible to social engineering attacks. Although, a different work concludes that warnings do have a positive impact on reducing risky user actions (Hu & Wang, 2018). This might be true since there are types of warnings that can be successful, take as an example active warnings or link-focused warnings (Petelka et al., 2019; Datta et al., 2021; Cooper et al., 2021; Egelman et al., 2008). Also there are works showing that combatting malicious URLs in emails can help the user to avoid phishing attacks (Volkamer et al., 2016; Volkamer et al., 2017).

From the aforementioned two papers by Datta et al. (2021) and Cooper et al. (2021), it follows that introducing audio and haptics to visual email warnings and alerts can be helpful in countering e-mail based social engineering attacks. One could also experiment with hovering ability and link analysis to introduce more effective countermeasures against email-based social engineering attacks.

3.3 Effectiveness of security warnings

Here we examine works about the effectiveness of security warnings where warnings do not necessarily have any relation with social engineering attacks.

The research of Anderson et al. (2016) found out that warning messages are largely ineffective. A key contributor to them being ineffective is habituation, a decreased response to repeated warnings. In order to tackle this problem, they designed a polymorphic warning artifact which repeatedly changes its appearance so that the effects of habituation could be resisted. They found out that a polymorphic warning artifact was significantly more resistant to habituation than traditional warnings.

Ebert et al. (2022) investigates how to design more effective security warnings. Using an online experiment they examined how informational content of warning and their visual salience could affect decision making. The result of the experiment is that both factors influence decision making simultaneously. Hence, when designing security warnings, it is important to pay at least as much attention to the visual design as the informational content.

Herley (2009) states that users rejecting security advice is entirely rational from an economic perspective. Most security advice simply offer a poor cost-benefit trade-off to users and are therefore rejected.

Krol et al. (2012), conducted a study to test the effectiveness of security warnings where 81.7% of the 120 participants ignored the warning. They found out that participants ignored the warnings because of their previous experiences of security warnings being false alarms. They conclude that users need to be re-sensitised and restore trust in warnings so that security warnings can be made effective again.

Summary

In conclusion, we see that there are various papers stating that security warnings are ineffective (Anderson et al., 2016; Herley, 2009; Krol et al., 2012). However, this could be improved with the use of polymorphic warnings (Anderson et al., 2016), lowering the number of false positives (Krol et al., 2012), offering a better cost-benefit trade-off to users (Herley, 2009) or designing more effective warnings (Ebert et al., 2022).

3.4 Incorporation of related work

In order to develop a countermeasure against email-based social engineering attacks, we will incorporate insights and recommendations from related work in designing the countermeasure.

We have seen that warning designs are frequently misunderstood or ineffective. To address this issue, we aim to create a design that is both easily comprehensible and visually stands out, to grab the attention of the user. Our warning designs should clearly indicate danger and how to proceed via the visual design and textual information.

For the purpose of ensuring accessibility and minimizing the potential for misunderstandings, we made the decision to avoid using different colors or shapes to convey certain messages. And as a consequence, we refrain from designing polymorphic warnings. Additionally, we will not make use of other techniques that are used to force attention such as sonifications, haptics and active warnings. This is because we want to find out how the designs would perform without these techniques.

Chapter 4

Countermeasures against email-based social engineering attacks

In this chapter we go through the process of proposing a visual countermeasure against email-based social engineering attacks. The process involves about three steps: first we examine various current existing countermeasures against these attacks. Afterwards, we identify potential issues with the discussed existing countermeasures. Based on this, we propose a method that combines technical and visual aspects to counter these attacks in the form of digital signature designs.

4.1 Countermeasures against email-based social engineering attacks

In this section we discuss a number of existing countermeasures against email-based social engineering attacks and their potential issues. These are used as a basis for proposing a new visual countermeasure.

4.1.1 Existing countermeasures

There exists numerous anti-social engineering software solutions for organisations which are created by various companies. Most of these software are based on the same concepts but differ in implementation. Therefore, only a couple of these solutions will be examined. In particular, we will list some techniques used by two known email security software providers:

Proofpoint[1] and Microsoft 365 Defender[2].

- **URL analysis and rewriting**
  URL analysis and rewriting aim to protect the user from visiting malicious web pages through URLs that are included in email messages (Cisco, n.d.). These malicious web pages may for example distribute malware or try to trick people into handing over sensitive information.

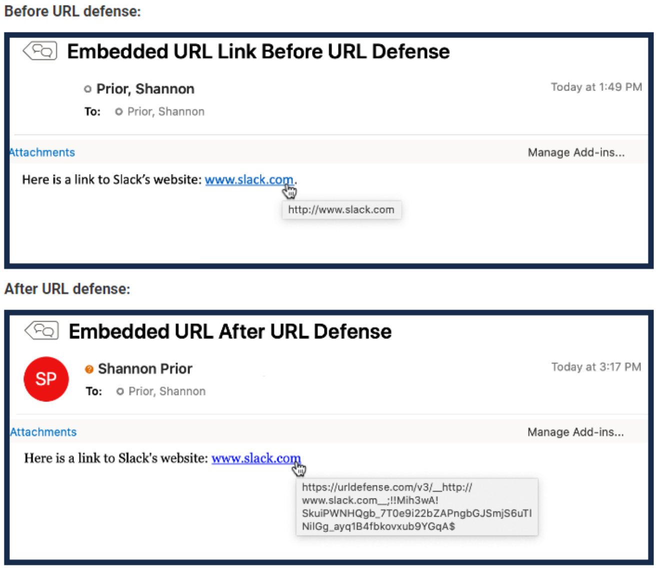
  The first method, URL analysis, performs a static or dynamic analysis to filter out, block access or neutralize suspicious URLs (S., 2020). Here, the exact implementation of the analysis depends on the vendor. An example of static analysis would be: matching suspicious URLs with other known malicious URLs contained in a database; if it matches then the URL in question is malicious. This is also known as blacklisting (Bhardwaj, 2021). With dynamic analysis, one would for example actually visit the suspicious URL via sandboxing to check whether the web page is malicious.

  Furthermore, URL rewriting modifies the URL in question so that the user is first redirected to a proxy instead of the actual web page. This is done so that the actual destination of the URL can be analysed while the user is at the proxy (Whaley, 2020). This allows URLs to be checked twice, once upon delivery and once at the time of clicking. Additionally, it can catch malicious URLs that were safe during the first scan. For example, if a link was compromised later after the first scan, then URL rewriting protects the user from visiting this malicious site by analysing it at the time of clicking. Additionally, URL rewriting can neutralize malicious URLs that use specific malicious parameters or unsecure protocols by removing or altering these parts of the URL. Scanned URLs are often rewritten or wrapped using a standard URL prefix, where the prefix is chosen by the vendor. This allows users to check whether the URL has been scanned by the security software. An example can be seen in Figure 4.1.

- **Email attachment analysis**
  Email attachment analysis provides protection against malicious email attachments (MailXaminer, n.d.), and can be done via a static or dynamic analysis (ReversingLabs, 2019; Proofpoint, n.d.-a). Proofpoint and Microsoft use dynamic analysis via sandboxing to ensure that attachments are examined before being delivered to the recipients (Proofpoint, n.d.-b; Microsoft 365 Defender, 2023a). The actions that will be taken after an analysis depend on the policy that is set

[1] https://www.proofpoint.com/us/products/email-security-and-protection/email-protection

[2] https://learn.microsoft.com/en-us/microsoft-365/security/defender/microsoft-365-defender?view=o365-worldwide



Before URL defense:

Embedded URL Link Before URL Defense

Prior, Shannon

To: Prior, Shannon

Today at 1:49 PM

Here is a link to Slack's website: www.slack.com.

http://www.slack.com

After URL defense:

Embedded URL After URL Defense

Shannon Prior

To: Prior, Shannon

Today at 3:17 PM

Here is a link to Slack's website: www.slack.com

https://urldefense.com/v3/__http://www.slack.com__;!!Mih3wA!SkuiPWNHQgb_7T0e9i22bZAPngbGJSmjS6uTINilGg_ayq1B4fbkovxub9YGqA$

Figure 4.1: Example of rewritten URL by Proofpoint.
Source: Figure of UC San Diego (2020).

up by the administrators. Also, email attachments can be checked using a form of static analysis that checks for similarity between the attachments in question and already known malicious attachments. If it is similar, then it is purged.

- **Email body content analysis**
Malicious email messages can also be identified by analysing the content of an email message's body (MailXaminer, n.d.; Egress, n.d.). Here, the body is checked for the presence of commonly used keywords or keyphrases by malicious emails. It is also possible that the email's body is examined on whether it adheres to the characteristics of email-based social engineering attacks such as urgency (Franchina et al., 2021). Lastly, the content of an email body can be checked for similarity with email bodies of already known malicious email messages.

- **Email header analysis**
Email header analysis is a technique that is used to check email header fields for authenticity (MailXaminer, n.d.; Poston, n.d.). There are various security and email authentication protocols that support this process and provide authentication results: SPF, DKIM and DMARC. These protocols were discussed in section 2.2.3. The authentication results of these protocols are shown in the email headers, so that tampered messages can be filtered out by checking whether the im-

plemented protocols have passed their checks. These three can be implemented together and are capable of working in tandem to provide stronger email authentication. As we have mentioned earlier, SPF is designed to prevent spoofing and blacklisting, DKIM is designed to protect data integrity of an email message and DMARC checks whether the headers in SPF and DKIM records match with the email message's header.

Also, as mentioned, the BIMI protocol can be used as an additional layer of authentication and verification for email messages. It requires implementation of DMARC and either SPF, DKIM or both. BIMI can help increase a recipient's trust in an email message by displaying a logo of the brand when that brand actually sent the email.

So, email header analysis in combination with email authentication protocols can be used to oppose email social engineering attacks. This is because it provides mechanisms to counter email spoofing by verifying the authenticity of email messages.

Next to that, various email clients provide some built-in protection against malicious emails. They can, for example, filter out suspicious emails or notify the user that an email message could be malicious. However, depending on the email client, the available security features may differ and may not be set up by default. Some security features that are included in Gmail and Microsoft Outlook are (Gmail, n.d.; Microsoft 365 Defender, 2023b):

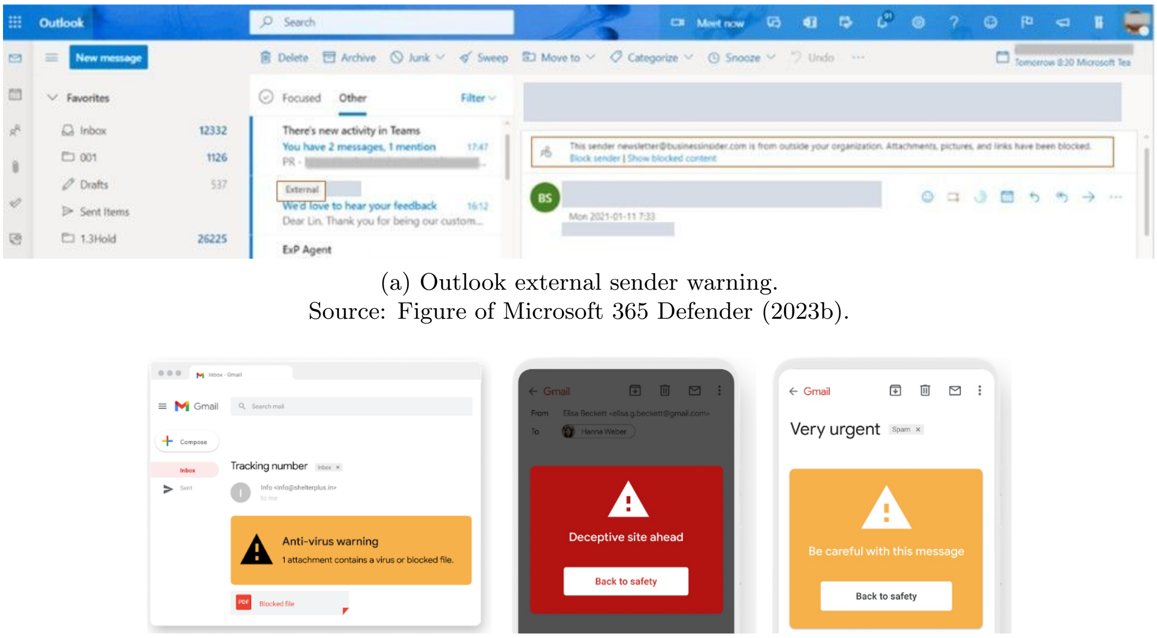
- **Spam filtering**
  Suspicious and spam email messages can be filtered out using a default or custom spam filter.

- **Malware detection**
  Links and attachments of email messages can be scanned for malware.

- **Scanning links and external images for malicious content**
  Malicious content can be identified by scanning the actual link of a masked link and by scanning external images. It is also capable of displaying a warning when the user clicks a link that is going to an untrusted domain.

- **Spoofing and unauthenticated email protection**
  It provides protection against spoofing of domains, unauthenticated emails, employee names or email messages that pretend to be from a domain.

- **Warning banners**
  When an email message is considered suspicious, a warning banner may be shown. Example warnings can be seen in Figure 4.2.



This sender newsletter@businessinsider.com is from outside your organization. Attachments, pictures, and links have been blocked.
Block sender | Show blocked content

Anti-virus warning
1 attachment contains a virus or blocked file.

Deceptive site ahead

Very urgent

Be careful with this message

(a) Outlook external sender warning.
Source: Figure of Microsoft 365 Defender (2023b).

(b) Warnings of Gmail.
Source: Figure of Gmail (n.d.).

Figure 4.2: Warning designs of Outlook and Gmail.

Additionally, there are email clients that support the installation of add-ons such as Microsoft Outlook, Mozilla Thunderbird or Gmail. Some of these add-ons can provide some protection against social engineering attacks. In the case of Gmail, it makes use of browser add-ons, since it is a webmail client. Outlook and Thunderbird add-ons can be either installed manually or via their respective add-ons page. In the related works section, we had discussed one anti-phishing add-on, named TORPEDO. TORPEDO is a Mozilla Thunderbird add-on by Volkamer et al. (2017) and is an anti-phishing solution against malicious URLs. Also, it should be noted that email encryption provides confidentiality and therefore can enhance protection against email-based social engineering attacks. For example, it can counter targeted social engineering attacks by protecting the email message from eavesdroppers, making it harder for an attacker to create convincing emails by personalizing. Lastly, email digital signatures can be used to increase email security. These will be explained below.

4.1.2 Email digital signatures

Email digital signatures can be attached to an email message to add a layer of email security (Microsoft 365 Defender, n.d., 2017). If a sender signed the contents of the email message, assurance is provided to the recipient that the email has not been altered in transit and that it genuinely originated from the mentioned sender.

These digital signatures differ from traditional signatures, i.e., a handwritten signature or some closing salutation. This is because an email message that is digitally signed can only come from the person that signed it. Whereas, with a traditional signature, anyone can copy it (Microsoft 365 Defender, n.d.).

In the case of email digital signatures of Microsoft Outlook, a digital ID is used to create a digital signature. This digital ID consists of a private key and a public key, and is based on digital certificates issued by CAs (Microsoft 365 Defender, 2017). If everything is set up correctly, one can sign an email by clicking the "Sign Message" button (Microsoft 365 Defender, n.d.).

Additionally, these signatures can be used to support email authentication, and are for example used in the aforementioned security protocols, discussed in Section 2.2.2: OpenPGP, S/MIME and DKIM. OpenPGP and S/MIME digital signatures are applied by the end user and provide authentication, message integrity and non-repudiation of origin. DKIM digital signatures are applied by the mail server that handles the email message and not by the end user.

From our observation, email digital signatures can potentially be the foundation of a solution that counters email-based social engineering attacks, since we believe that the problem is caused by a lack of authentication. Hence, we will look further into how email digital signatures can be enhanced or adjusted for the purpose of countering email-based social engineering attacks.

4.1.3 Potential issues with existing countermeasures

Unfortunately, the previously examined countermeasures are not fail proof. In this subsection, we will address several potential issues with the existing countermeasures.

URL rewriting has its flaws. If one wants check whether a URL has been rewritten, it often must hover over the link and look for the standard URL prefix. This hovering technique might be too complicated for some users (Whaley, 2020). Mobile devices or other touch screen devices do not support hover, therefore those users might neglect inspecting URLs. Additionally, URL rewriting could even be counterproductive: instead of giving a user some sense of protection, we think that the rewritten URL might scare off or discourage a user from clicking a legitimate link. Also, URL analysis and rewriting is detection based, meaning that there are false positives and not all malicious links can be detected. It might even break links immediately or after a period of time (Whaley, 2020).

We have observed that countermeasures based on the analysis of email content rely on detection. Therefore, we think that the problem is in the form of false positives and false negatives. False negatives in this context are essentially the undetected malicious emails, attachments or URLs, and can occur since attackers are capable of finding ways to bypass security solutions. False positives are legitimate emails, attachments or URLs that are wrongly flagged as malicious and can result in end user frustration. Therefore, one cannot solely rely on detection-based countermeasures.

On top of that, there exists email-based social engineering attacks that try to deceive users by using contextually relevant content. These emails do not use any kind of malicious links or attachments, making the process of detection very challenging.

Also, most of the countermeasures are based on scanning email content for maliciousness. This could lead to privacy issues. The usage of end-to-end email encryption makes countermeasures based on scanning malicious content ineffective unless it is performed at the end points where the content is decrypted. However, one has to still deal with the issue of the trade-off between security and privacy.

As far as built-in email protection is concerned, a lot of email clients (mobile, web and third-party) have not adopted security indicators for forged emails (Hu & Wang, 2018). Personally, we have encountered many malicious and suspicious email messages in our personal email account that were not flagged or filtered out and ended up in the regular inbox without any warnings, some examples can be seen in Appendix F. Also, email protection on an organisational level requires security settings such as email security rules or policies to be configured correctly so that the required email security level can be achieved. According to Paunikar (2021), the default Microsoft Office 365 email security settings are weak.

Some of the countermeasures might even provide the end user with a false sense of security. For example, one could rely too much on built-in email client protection, potentially leaving them vulnerable to undetected malicious email messages that end up in the user's inbox.

The effectiveness of email security protocols often depend on interoperability. Email authentication protocols require mail servers or email clients to be compatible with the used protocols on both ends. Take as an example S/MIME, it requires the sender's and recipient's email clients to support S/MIME in order to be used correctly. If this is not the case, then the email most likely cannot be delivered or read by the recipient which could be frustrating for both end users. The same applies to email digital signatures, both ends have to be capable of sending and receiving digitally signed messages, which can lead to usability problems since not all email providers provide the same services (Moecke & Volkamer, 2013). Another issue with email digital signatures is the limited adoption rate (Reuter et al., 2021). This is mainly caused by the complexity and cost of setting up and managing digital signatures (Garfinkel et al., 2005), an example of email digital signatures that make use of CA-based digital certificates would be Microsoft Outlook. The limited adoption rate is mainly caused by poor usability (Moecke & Volkamer, 2013; Garfinkel et al., 2005). In the context of secure email, the works of Garfinkel et al. (2005) and Moecke and Volkamer (2013) found out that the usability of secure email is impacted by a lack of awareness and education, for example: users do not understand how to use cryptography in secure email or they simply do not care. On top of that, the implementation and usage of digital signatures can be too complex for non-technical users, especially in terms of key management (Garfinkel et al., 2005).

A solution to address the issues of digital signatures might be in the form of identity-based digital signatures, since we believe that it makes the process of signing easier and less costly via identity-based cryptography. In this thesis, we opt for identity-based digital signatures, since it is, as mentioned, easier to manage than traditional cryptography. On top of that, there is an existing and ongoing project named PostGuard, discussed in section 2.3, that uses identity-based encryption and is planning to implement identity-based digital signatures. We believe that the usage of identity-based digital signatures in emails can help users to identify malicious emails and decrease the number of effective malicious email messages. Therefore, we propose a visual countermeasure that makes use of identity-based digital signatures to counter email-based social engineering attacks.

4.1.4 Proposed countermeasure

Even with the presence of various email security solutions, users can still be susceptible to social engineering attacks. This is why some solutions rather not present the end user with these security decisions, because humans are prone to make mistakes or make incorrect decisions due to a lack of knowledge. Therefore, some security solutions filter suspicious mails in mail servers before they are received by the end user. However, one can also not rely fully on that principle, since not every malicious mail can be detected. This process is even more complex concerning end-to-end encrypted email and privacy. Encryption prevents analysis of malicious content and therefore can only be used at decryption. However, the problems regarding privacy remain.

Therefore, we want to counter these attacks by introducing a new visual technique that minimizes privacy risks and relies on the revealed identity attributes by the sender. These visual elements should support the user in identifying malicious emails and indicate whether an email message can be trusted or that a user has to proceed with caution. Ideally, with the goal that these could be universally understood without requiring any special training.

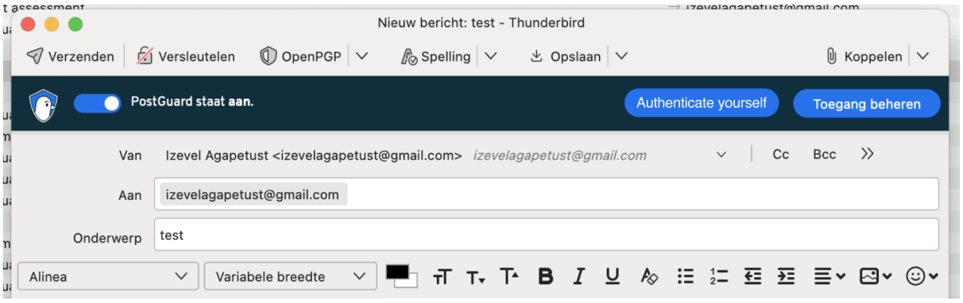
It should be noted that this visual technique does not counter email-based social engineering attacks on its own, it should be used in combination with other security solutions that can be used for identity-based end-to-end encrypted emails.

More specifically, we propose to develop a visual representation of digital signatures used for email authentication via PostGuard. This is because PostGuard allows end users to easily encrypt, send, receive and decrypt email messages based on end-to-end identity-based encryption in a privacy-friendly way. Currently, digital signatures are not supported by PostGuard but are in development. In this thesis, we assume that this feature already has been implemented, where we envision this with visual mockups. It should be noted that these designs could be adjusted to be used in other email client security software that support identity-based digital signatures. The authentication with Yivi allows us to include the presented identity attributes in our digital signature designs, and is used to present more easy-to-understand email authentication information to users. Ultimately, increasing a recipient's awareness by indicating whether and by whom an email message is digitally signed and comes from a known or unknown sender.

4.2 Identity-based digital signatures with Yivi

Our proposed countermeasure is the usage of identity-based digital signatures in emails, which in its core is nearly identical to digital signatures. The key difference is that it makes use of Yivi, so that it is easier to use (users are not bothered with key management). On top of that, we aim to support the user in identifying malicious emails by the inclusion of identity attributes of the sender, which we will further on mention as ID verifications[3]. ID verifications are the attributes that allow one to identify a unique person or organisation and is done via identity attributes and ownership attributes. An example of ID verifications for a person would be: full name, date of birth, nationality, occupation, organisation and email address. In this thesis, we envision the mockups for the email security add-on PostGuard, which was discussed in Section 2.3. One can sign and include ID verifications with PostGuard by composing an email in an email client that supports the installation of PostGuard. Specifically, to sign an email, one has to enable a slider button, see Figure 4.3. If a user wants to add ID verifications, it can do so by clicking the "Authenticate yourself" button, which will lead to a window where ID verifications can be added. This can be seen in Figure 4.4. Authentication of ID verifications is ensured by Yivi, which we have addressed in Section 2.3.1.

[3] Based on: `https://www.microsoft.com/en-us/security/blog/2023/04/12/linkedin-and-microsoft-entra-introduce-a-new-way-to-verify-your-workplace/`

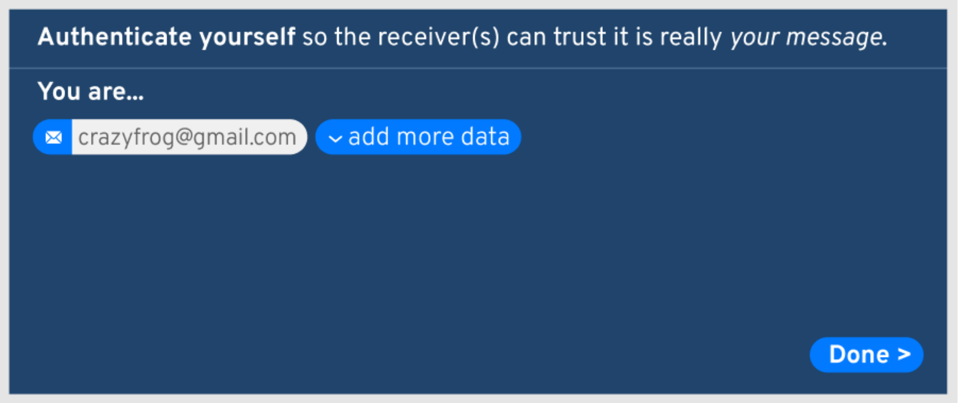


Van Izevel Agapetust <izevelagapetust@gmail.com> *izevelagapetust@gmail.com*

Aan izevelagapetust@gmail.com

Onderwerp test

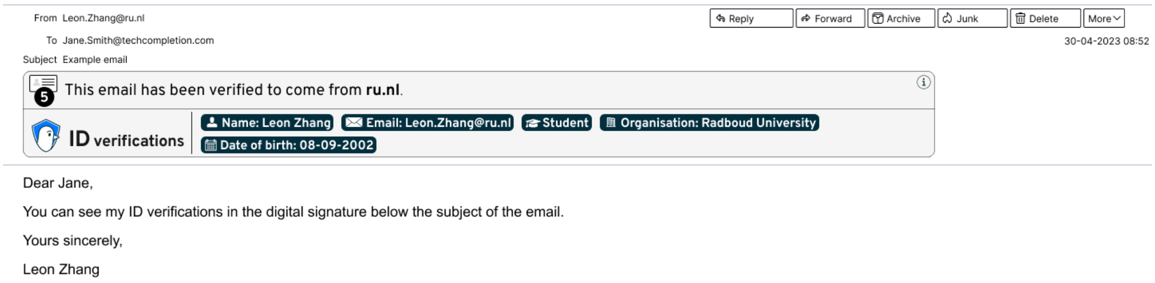
Figure 4.3: Signing emails with PostGuard.
Source: Unpublished figure of PostGuard (obtained in personal communication).



Authenticate yourself so the receiver(s) can trust it is really *your message.*

Figure 4.4: Adding ID verifications with PostGuard.
Source: Unpublished figure of PostGuard (obtained in personal communication).

Our mockups of the identity-based digital signature consist of a banner which is located within the email header block, containing the identity attributes of the sender. In this thesis, we will refer to this as the ID banner, and can be seen in Figure 4.5. The various types of identity-based digital signature designs will be discussed later in this chapter. We have opted to design a banner because the risk of social engineering attacks extends beyond just malicious links or attachments and can be present in the email message itself. Therefore, we believe that a banner would be a straightforward solution to address all three attack methods. Also, we think that when the design is properly introduced to users, it will reduce users' susceptibility to email-based social engineering attacks. Since, we aim to design the ID banner to be easily understandable. Additionally, as mentioned, Hu and Wang (2018) showed that email clients/services with security indicators in the form of banners have a positive impact on reducing risky user actions.



From Leon.Zhang@ru.nl

To Jane.Smith@techcompletion.com

30-04-2023 08:52

Subject Example email

This email has been verified to come from **ru.nl**.

ID verifications

Name: Leon Zhang | Email: Leon.Zhang@ru.nl | Student | Organisation: Radboud University | Date of birth: 08-09-2002

Dear Jane,

You can see my ID verifications in the digital signature below the subject of the email.

Yours sincerely,

Leon Zhang

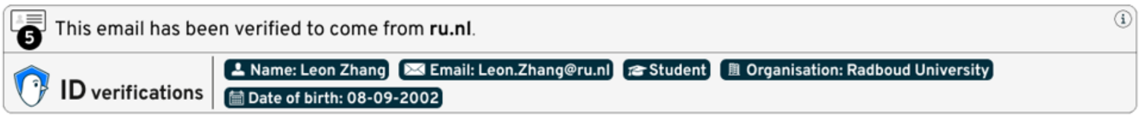
Figure 4.5: Mockup of an identity-based digital signature attached to an email.
Source: Own work.

Before we explain the various types of designs, we define the properties of an email that determine which ID banner design is used. First of all, an email can be sent from an internal sender or external sender. We define an email to be from an internal sender: if the domain of the sender's email address is equal to the domain of the recipient's email address. In all other cases, it is seen as an external sender. Secondly, an email can be signed or not signed. This corresponds to whether the email has been signed with a digital signature using PostGuard. Lastly, the included ID verifications can be viewed as either relevant or irrelevant.

In this thesis, we defined relevant and irrelevant ID verifications based on the goal of staying privacy-friendly by, e.g., minimizing ID verifications, while providing enough information to the recipient so that they can determine whether the sender is known to them. Therefore, relevant and irrelevant ID verifications are defined using the following definitions: For external senders, a combination of ID verifications is deemed to be irrelevant if it misses at least one of the following attributes:

- Email address.
- (For organisations) The name and domain name of the organisation.
- (For persons) The name of the person and the name of the organisation that they are part of (in this thesis we do not address personal email accounts).

If these are included, then the ID verifications are relevant. The detection of relevant or irrelevant ID verifications can be done automatically, because the initial distinction of the sender being a person or organisation can be made using the included ID verifications. After that, the verifications can be checked on the above criteria. For internal senders, emails that are signed are always legitimate, regardless of the given verifications. Thus, these verifications are always seen as relevant. An example of Relevant ID verifications can be seen in Figure 4.6 We assume that one cannot give ID verifications


5 This email has been verified to come from **ru.nl**.

ID verifications

Name: Leon Zhang | Email: Leon.Zhang@ru.nl | Student | Organisation: Radboud University | Date of birth: 08-09-2002

Figure 4.6: Example Relevant ID verifications of the author.
Source: Own work.

that do not match with the sender's email address and sender's name. If one want to include multiple email addresses, it must at least include the email address that was used to send the email.

If an external sender has not signed the email, then the email is viewed as malicious (depends on the adoption rate of identity-based digital signatures). In this thesis, we assume that we are studying the effectiveness of the designs in an ideal situation, i.e, high adoption rate.

If one wants to impersonate an external sender, then it must create an organisation with a similar legal name which is subject to laws and regulations. Impersonation should be a lot more difficult to execute. Also, a person cannot sign with the name of the person they impersonate unless they happen to have the same name.

Although PostGuard cannot be used on mobile devices, we do start with creating the designs for mobile devices. This is done to ensure that the designs are compatible for all devices. Then, from these designs, we create the designs for larger screens such as desktops. We have created four identity-based digital signature designs. Each ID banner state has a corresponding digital signature design. These were first created for mobile devices and converted later to fit larger screens.

Based on whether the sender's email address is internal or external, whether it is signed or not, and whether relevant ID verifications or irrelevant ID verifications are included, we determine what information is displayed on the ID banner. We define the following states for the banner, supported by Figure 4.7:

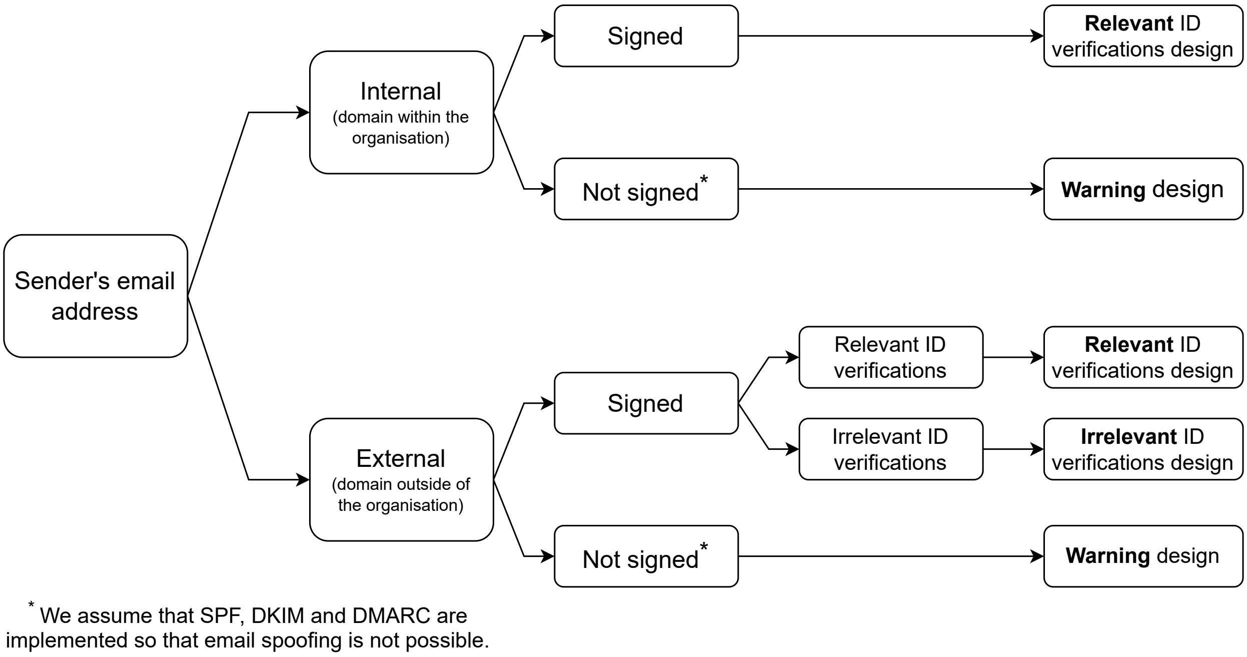
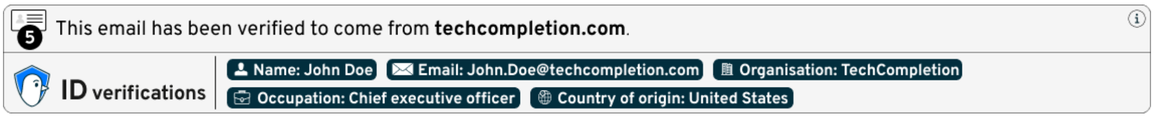

Sender's email address
Internal
(domain within the organisation)
External
(domain outside of the organisation)
Signed
Not signed*
Signed
Not signed*
Relevant ID verifications
Irrelevant ID verifications
Relevant ID verifications design
Warning design
Relevant ID verifications design
Irrelevant ID verifications design
Warning design
* We assume that SPF, DKIM and DMARC are implemented so that email spoofing is not possible.

Figure 4.7: A flowchart on how the state of the ID banner is determined. Source: Own work.

- **Relevant ID verifications design:** This design is used for an email message with an internal or external sender that has signed the email and included relevant ID verifications. The designs can be seen in Figure 4.8.

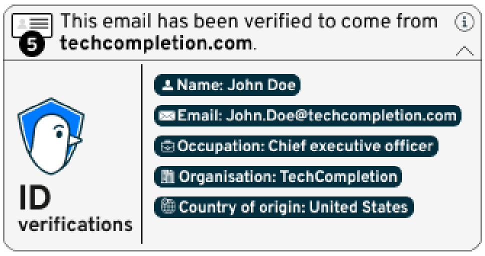


This email has been verified to come from **techcompletion.com**.

ID verifications

Name: John Doe
Email: John.Doe@techcompletion.com
Organisation: TechCompletion
Occupation: Chief executive officer
Country of origin: United States

(a) Relevant ID verifications design for large screens.



This email has been verified to come from **techcompletion.com**.

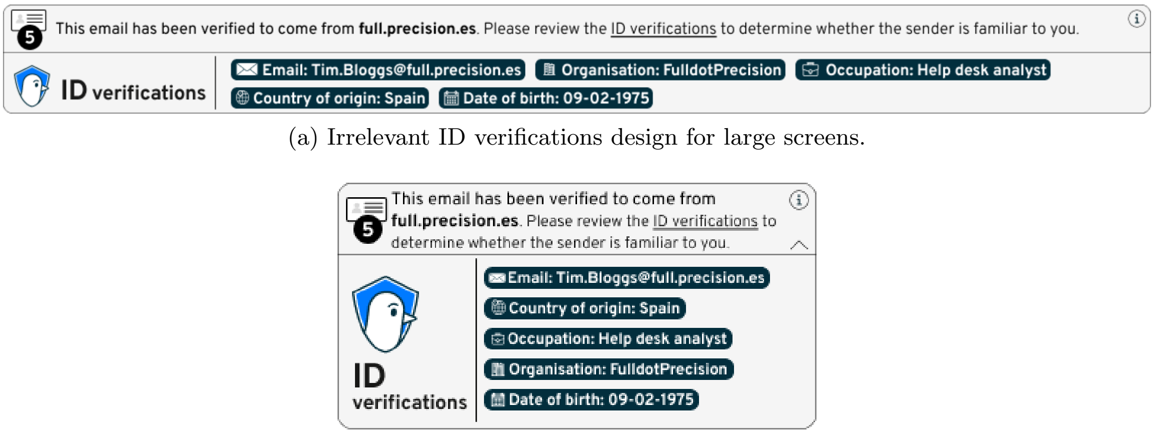
ID verifications

Name: John Doe
Email: John.Doe@techcompletion.com
Occupation: Chief executive officer
Organisation: TechCompletion
Country of origin: United States

(b) Relevant ID verifications design for small screens.

Figure 4.8: Relevant ID verifications designs for large and small screens. Source: Own work.

- **Irrelevant ID verifications design:** This design is either used for an email message with an external sender that has signed the email and included irrelevant ID verifications. The designs can be seen in Figure 4.9, where the ID verifications are irrelevant because the “Name” ID verification is missing.



5 This email has been verified to come from **full.precision.es**. Please review the ID verifications to determine whether the sender is familiar to you.

ID verifications

Email: Tim.Bloggs@full.precision.es | Organisation: FulldotPrecision | Occupation: Help desk analyst | Country of origin: Spain | Date of birth: 09-02-1975

(a) Irrelevant ID verifications design for large screens.

5 This email has been verified to come from **full.precision.es**. Please review the ID verifications to determine whether the sender is familiar to you.

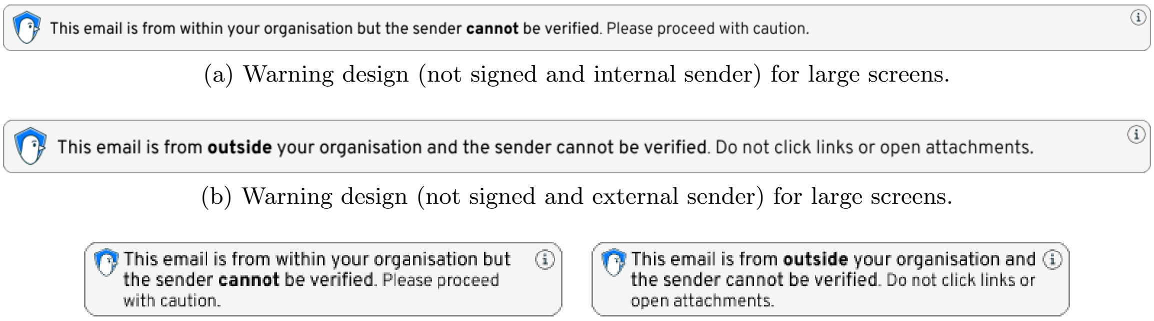
ID verifications

Email: Tim.Bloggs@full.precision.es
Country of origin: Spain
Occupation: Help desk analyst
Organisation: FulldotPrecision
Date of birth: 09-02-1975

(b) Irrelevant ID verifications design for small screens.

Figure 4.9: Irrelevant ID verifications designs for large and small screens.
Source: Own work.

- **Warning designs:** This design is used if the email came from an internal or external sender and has not been signed. The designs for an email from an internal or external sender that has not been signed can be seen in Figure 4.10.



This email is from within your organisation but the sender **cannot** be verified. Please proceed with caution.

(a) Warning design (not signed and internal sender) for large screens.

This email is from **outside** your organisation and the sender cannot be verified. Do not click links or open attachments.

(b) Warning design (not signed and external sender) for large screens.

This email is from within your organisation but the sender **cannot** be verified. Please proceed with caution.

This email is from **outside** your organisation and the sender cannot be verified. Do not click links or open attachments.

(c) Warning design (not signed and internal sender) for small screens.

(d) Warning design (not signed and external sender) for small screens.

Figure 4.10: Warning designs for large and small screens.
Source: Own work.

Note that these states should support the user in identifying malicious emails, but should not be used on its own to determine maliciousness of emails. The purpose of the banner is to provide some information of the sender so that the user can for example make a decision based on the ID verifications, email header and body.

The following elements were incorporated in the designs:

- Mobile designs have an expand/collapse button since the designs occupy more vertical screen space due to width constraints.
- Each ID verification has its name and icon displayed to avoid misunderstandings.
- If an email is not digitally signed, then no ID verifications will be shown. This is displayed very apparently in the design.
- If an email contains ID verifications, then the number of ID verifications is displayed by the icon in the upper-left corner of the design.
- The designs for both mobile and desktop are identical, therefore one does not have to get used to a different design when switching devices.

Chapter 5

Methodology

In this chapter we elaborate on the design of the online experiment that is used in this thesis to measure the effectiveness of the proposed digital signature designs in countering email-based social engineering attacks.

5.1 Online experiment design

In this section, we will explain the research design that was used in this thesis along with a clarification behind the reasoning of choices that were made; to measure the effect of including our visual identity-based digital signatures designs in email messages. In other words, we measure the perceived credibility of an email message with and without the proposed designs. First, we start with explaining how the experiment was designed. Then we explain the general set-up and procedure along with some information about the group of participants.

For this experiment, we use a repeated measures, also known as within-subjects, design. This defines how test participants are assigned to multiple conditions in a single study (Budiu, 2018). In repeated measures design, the participants remain the same for every condition, therefore every participant produces one result for every condition in the study (Budiu, 2018; Field & Hole, 2002). The conditions in this study are digital signature designs attached to an email message, broken down into two scenarios. The first scenario has seven conditions and the second scenario contains six conditions.

We have chosen for a repeated measures design because it is in terms of time, effort and participant numbers more viable to run than between-subjects design. It requires less participants and has greater statistical power than between-subjects design when access is limited to a few participants (Field & Hole, 2002). However, repeated measures design also comes with a drawback, namely carry-over effects. These are the effects that influence a participant's behaviour in later presented conditions and are caused by, for example, the order in which the conditions appear or by learning and transferring information across conditions.

To counteract these effects, one uses randomisation or counterbalancing. Randomisation is defined as presenting conditions in a random order to the participant, while counterbalancing uses a systematic variation of the order of conditions (Field & Hole, 2002; Bhandari, 2022). Counterbalancing can be sometimes more favourable, this is because the researcher can ensure that every defined sequence is presented to an even portion of the sample group (Field & Hole, 2002; Bhandari, 2022). Ideally, each condition appears equally often in each position of the sequence, to balance out the order effect on the results and is also known as complete counterbalancing. However, this is less feasible for experiments with numerous conditions. In this thesis, we have 13 conditions in total: 7 for the first scenario and 6 for the second scenario. Therefore, we use randomisation instead of partial or complete counterbalancing due to the number of sequences needed and the difficulty of implementing this in the software (LimeSurvey) used for constructing the survey.

We also try to keep the experiment length short, so that the experiment does not become taxing or tedious for the participants. Otherwise, the results may suffer from a decline of motivation. The scenarios and conditions will be explained in more details in the next section 5.1.1.

5.1.1 Conditions

We have opted to use the mobile-first approach in designing the identity-based digital signatures. Therefore, all the designs have been first created for mobile devices. The designs for larger screens followed from the mobile designs. In this experiment, we have chosen to only measure the digital signature designs for larger screens to minimize the survey length. Hence, the conditions are based on the designs for larger screens. The conditions that were used in the experiment are included in Appendix A. Note that we have included an equal number of ID verifications to the conditions that are very similar to each other but differ in the aspect of being malicious or legitimate. Take as an example, a signed legitimate email with 3 irrelevant ID verifications and a malicious signed email with 7 irrelevant ID verifications. Here, a participants might judge the credibility of the email based on the number of ID verifications, making our measurement less accurate. Hence, to measure the effects of the visual designs more accurately, we have chosen for an equal number of ID verifications. This makes it more likely that a participant's behaviour is not influenced by the number of ID verifications but solely by the design of the signatures.

Now we will look at the two different scenarios of the experiment. The first scenario contains emails that are from an internal and external sender, whereas the second scenario only contains emails from an external sender. This is done so that the effects of the designs can be examined more accurately by testing them in multiple situations.

5.1.1.1 Scenario 1: TechCompletion

The participants of the experiment were first presented with the scenario about a fictional company, named TechCompletion. The scenario description is as follows:
You are Jane Smith, a warehouse manager of the fictive company TechCompletion (`techcompletion.com`) located in the United States.

The chief executive officer (CEO) of TechCompletion, John Doe, made sure that all employees, including you, installed the email add-on PostGuard. PostGuard allows users to securely communicate via email and provides information about the identity of an email's sender, in the form of ID verifications. Examples of ID verifications are: full name, email address, date of birth, nationality, occupation and many more.

In this scenario we test the following conditions, supported by Table 5.1, which shows for each condition: the origin of the email (internal or external), the authenticity of the email (legitimate or malicious) and the type of identity-based digital signature design that is displayed.

Condition	Sender origin	Authenticity	Design
A	Internal	Legitimate	None (Control condition)
B	External	Legitimate	None
C	External	Malicious	Warning (external)
D	External	Malicious	Irrelevant ID verifications
E	Internal	Legitimate	Warning (not signed)
F	External	Legitimate	Relevant ID verifications
G	Internal	Legitimate	Relevant ID verifications

Table 5.1: Conditions of scenario 1 (TechCompletion).

Condition A and B serve as control conditions and are used to see how users perceive the credibility of the email without the inclusion of any identity-based digital signature designs, so that effect of the designs can be investigated. The design of the email messages shown in these condition are discussed in subsection 5.2.1.

The visual representations of the conditions are shown in Appendix A and the email message used can be seen in subsection 5.2.1. In Table 5.2, one can see that we defined all signed email messages from an internal sender to have relevant ID verifications, regardless of which ID verifications are included. This is because of the unlikeliness that an attacker compromised both email account and authentication account (Yivi). Hence, the condition with a signed email and irrelevant ID verifications from an internal sender does not exist.

Internal sender	External sender	
Condition G	Condition F	Signed - relevant ID verifications
	Condition D	Signed - irrelevant ID verifications
Condition E	Condition C	Not signed
Condition A	Condition B	None (no design)

Table 5.2: The possible conditions of scenario 1. We have crossed out the table cells that under our assumption are unlikely to occur.

5.1.1.2 Scenario 2: MijnOverheid

The second scenario of the experiment differs from the first scenario by excluding the designs used for internal senders. This is because the email message is from MijnOverheid, the digital platform facilitating interactions with Dutch authorities, and is an external sender. We included this scenario so that the effects of the designs for external senders can be examined more comprehensively. The scenario is explained as follows:

You are still Jane Smith, a warehouse manager of TechCompletion. You have received some emails from MijnOverheid (`mijn.overheid.nl`) which is the digital platform for your dealings with Dutch authorities.

You also still have PostGuard installed on the applications you use for sending and receiving emails.

You will now see six email messages.

In this scenario we test the conditions seen in Table 5.3:

Condition	Sender origin	Authenticity	Design
I	External	Legitimate	None
II	External	Malicious	None
III	External	Malicious	Warning (external)
IV	External	Malicious	Irrelevant ID verifications
V	External	Legitimate	Irrelevant ID verifications
VI	External	Legitimate	Relevant ID verifications

Table 5.3: Conditions of scenario 2 (MijnOverheid).

Here, condition I and II serve as control conditions and are used to see how users perceive an email's credibility when no identity-based digital signature design is shown. This is done so that effect of the designs can be investigated. The design of the email messages shown in these condition are discussed in subsection 5.2.1.

The visual representations of the conditions are shown in A and the email messages used are shown in subsection 5.2.1. We do not test the condition (external sender, malicious email, signed - relevant ID verifications), as can be seen in Table 5.4, because of the increased difficulty for an attacker to impersonate someone, since they have to create an organisation with their malicious email address. We also assume that the adoption rate of identity-based digital signatures is high, meaning that not signed external emails are likely to be malicious.

External sender

Malicious email	Legitimate email	
	Condition VI	Signed - relevant ID verifications
Condition IV	Condition V	Signed - irrelevant ID verifications
Condition III		Not signed
Condition II	Condition I	None (no design)

Table 5.4: The possible conditions of scenario 2: MijnOverheid. We have crossed out the table cells that under our assumption are unlikely to occur.

5.1.2 Introduction of visual digital signatures

We included the identity-based digital signatures in email messages which we made ourselves. For scenario 1: TechCompletion, we have chosen to use two email messages, one from an internal sender and one from an external sender. The email message from the external sender can be interpreted as malicious and legitimate depending on the given sender's email address and digital signature.

For scenario 2: MijnOverheid, we have chosen to use one email message that can be interpreted as malicious and legitimate depending on other information such as email addresses. Conditions II, III and IV are malicious and conditions I, V and VI are legitimate. This scenario is included in the experiment to more specifically examine the effect of identity-based digital signature designs in email messages coming from external senders.

5.2 Apparatus and materials

In this section we clarify the design of our experiment and the materials used in which the aforementioned decisions are incorporated.

We used a interface design software tool named Figma[1] to design the email messages and the identity-based digital signatures. The survey tool LimeSurvey[2] was used to create and let participants fill in the survey.

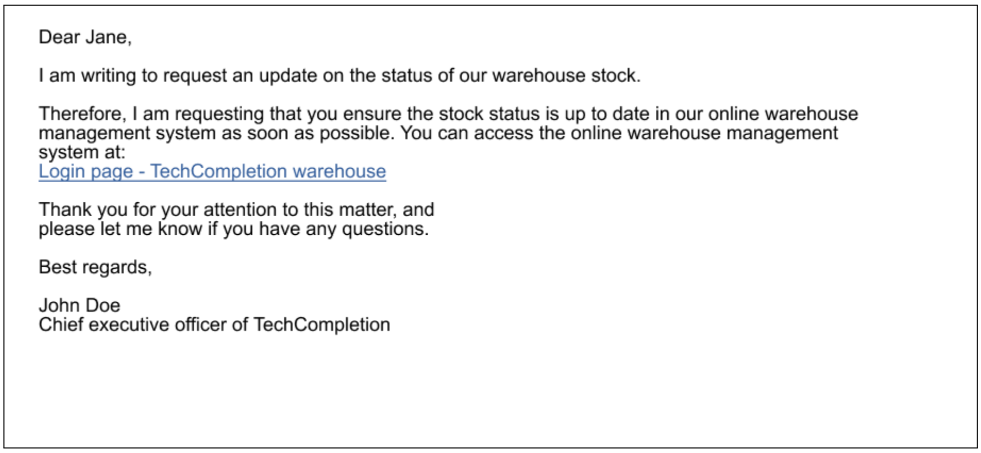
We will now address the email messages used for the experiments on which the identity-based digital signature designs are attached.

5.2.1 The email messages

5.2.1.1 Email messages of scenario 1: TechCompletion

We have chosen to create an email which is sent by an internal sender and an email message sent by an external sender. The external sender's message could be malicious or legitimate depending on the sender's email address and digital signature.

The malicious form of the email message corresponds to a business email compromise attack. We have chosen for a BEC attack since these are on the rise (Microsoft Security, 2022). For conditions A, B, F and G the email message is legitimate. For conditions C, D and E the email message is malicious. The participant can identify whether the email is malicious by inspecting the email header and the state of the digital signature design. The body of the email message can be seen in Figure 5.1 and 5.2 respectively. We have chosen to minimize the differences between the two email messages so that the effect of including digital signatures can be measured more accurately.



Dear Jane,

I am writing to request an update on the status of our warehouse stock.

Therefore, I am requesting that you ensure the stock status is up to date in our online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and
please let me know if you have any questions.

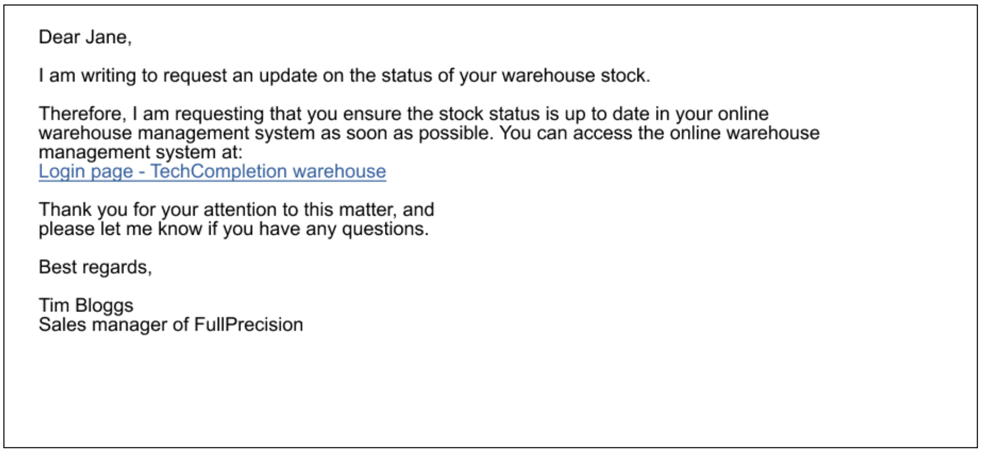
Best regards,

John Doe
Chief executive officer of TechCompletion

Figure 5.1: The email message's body from an internal sender of scenario 1. Source: Own work.

[1] https://www.figma.com/

[2] https://www.limesurvey.org/en/



Dear Jane,

I am writing to request an update on the status of your warehouse stock.

Therefore, I am requesting that you ensure the stock status is up to date in your online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and
please let me know if you have any questions.

Best regards,

Tim Bloggs
Sales manager of FullPrecision

Figure 5.2: The email message's body from an external sender of scenario 1.
Source: Own work.

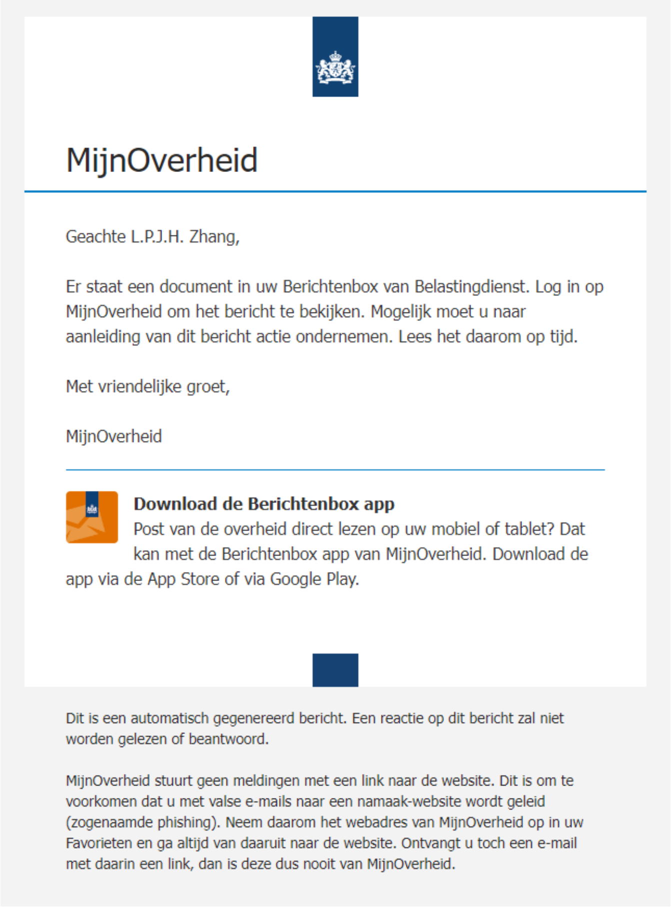
5.2.1.2 Email messages of scenario 2: MijnOverheid

The most common form of social engineering attack is phishing via email (Proofpoint, 2023; Dyer, 2023). Since we have used a BEC attack in the first scenario, we will use a less targeted attack in this scenario. Specifically, a mass phishing email along with its legitimate version. The content of this email has been designed to serve as a first phase of a mass phishing attack. It essentially checks whether one replies to the email so that suitable targets can be acquired.

Also, we imitated the design of a legitimate email message of MijnOverheid[3] and used the content of this message to create the email messages of scenario 2. The legitimate email message can be seen in Figure 5.3. In particular, we imitated the "Download the Berichtenbox app" and partially the footer of the legitimate email message. The other parts were designed in such a way to be suitable for the experiment.

We have chosen to create a MijnOverheid mass phishing email because most of the participants were located in the Netherlands and are likely to be familiar with these email messages. In this scenario, we have chosen to use the same content for both legitimate and malicious email messages so that the effect of including digital signatures can be measured more accurately. The created email message for scenario 2 can be seen in Figure 5.4.

[3] https://www.netherlandsworldwide.nl/mijnoverheid-abroad/what-is-mijnoverheid



MijnOverheid

Geachte L.P.J.H. Zhang,

Er staat een document in uw Berichtenbox van Belastingdienst. Log in op MijnOverheid om het bericht te bekijken. Mogelijk moet u naar aanleiding van dit bericht actie ondernemen. Lees het daarom op tijd.

Met vriendelijke groet,

MijnOverheid

Download de Berichtenbox app
Post van de overheid direct lezen op uw mobiel of tablet? Dat kan met de Berichtenbox app van MijnOverheid. Download de app via de App Store of via Google Play.

Dit is een automatisch gegenereerd bericht. Een reactie op dit bericht zal niet worden gelezen of beantwoord.

MijnOverheid stuurt geen meldingen met een link naar de website. Dit is om te voorkomen dat u met valse e-mails naar een namaak-website wordt geleid (zogenaamde phishing). Neem daarom het webadres van MijnOverheid op in uw Favorieten en ga altijd van daaruit naar de website. Ontvangt u toch een e-mail met daarin een link, dan is deze dus nooit van MijnOverheid.

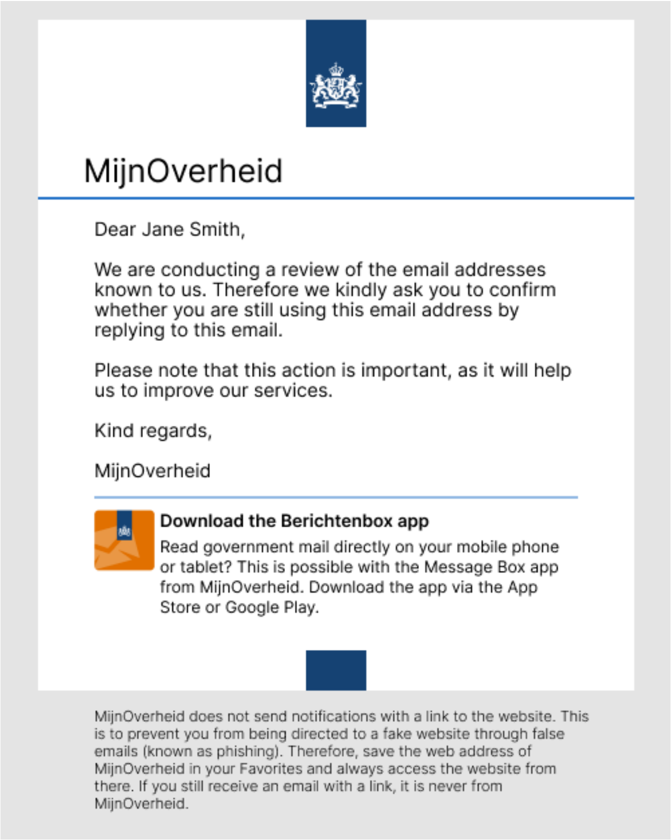
Figure 5.3: A legitimate email sent by MijnOverheid.
Source: Email received by the author.

5.2.2 Online survey and procedure

Our independent variable was the identity-based digital signature design attached to an email message, consisting of seven manipulations for the first scenario and six manipulations for the second scenario.

We have the following main research question: “**How to counter email-based social engineering attacks with user experience design?**”. However, now that we have proposed a countermeasure in the form of identity-based digital signatures. We answer the question: “**What is the effect of identity-based digital signatures on email credibility?**”. Where email credibility is split up into the following variables: credibility of the sender, credibility of the email message, email sender certainty, email interaction and email interaction comfort.

Therefore, we have the following dependent variables in this research: email sender credibility, email message credibility, email sender certainty,



MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

MijnOverheid

Download the Berichtenbox app

Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

Figure 5.4: The email message's body of scenario 2 where the identity-based digital signature designs are attached on.
Source: Own work.

email interaction and email interaction comfort. These variables encapsulates the effectiveness of the created designs in countering email-based social engineering attacks. To measure these dependent variables we make use of a self-report online survey. This is used to measure the beliefs or feelings of a participant about the credibility of each email message and its sender with or without an identity-based digital signature design.

5.2.2.1 Survey

Now we will examine the survey that is conducted to measure the dependent variables. We have chosen to conduct a survey because it involves actual participants and provides us with insights from various perspectives. We chose to conduct this survey via LimeSurvey which is hosted on the servers of Radboud University. In this way, participants could use their own devices for filling in the survey, while being in their usual environment.

First, every participant in the study was presented with an information letter which provided information about the research. If the participant agreed to participate in the experiment, it could do so via the consent form. After agreeing to this, they had to fill in a few questions related to their demographic background, and is used to determine how representative the group of participants were. This was followed by some questions about their email literacy and activity e.g. number of emails sent and received.

Subsequently, since we used a repeated measures design, every participant saw a total of 13 conditions. The order of condition appearance was assigned via randomisation for each scenario. Before the participants saw the conditions and the respective questions, they were first introduced to the add-on PostGuard and the idea of an identity-based digital signature. This is because in reality, one should not introduce these kind of changes without notifying users. However, we did strive to make the designs self-explanatory. If we had decided to not explain the designs beforehand, then participants might get confused or overwhelmed. This likely has an effect on their behaviour and therefore the quality of results. So, we expect to measure the effectiveness of these designs more accurately when it is tested with somewhat informed participants. Also, as we have seen in the Related Works chapter, many studies have shown that visual security indicators are not understood or misunderstood by users (Stojkovski & Lenzini, 2020; Dhamija et al., 2006; Felt et al., 2016). Our chosen approach can support users in interpreting the designs. After that, all the participants were first presented with scenario 1 and its conditions along with the questions about the credibility of the email message and sender. This was paired with some scenario specific email interaction questions e.g. would you reply to the email? or would you click the link in the email? We have chosen to let the participants see the conditions while answering questions about it, without a time limit. This is done because a user can often view email messages at their own pace. However, in reality, people just spend nine seconds, on average, looking at email messages (Nanji, 2023). But since we do not want to encourage users to quickly glance through emails, we have given them unlimited time. After scenario 1, the participants were shown the conditions and questions of scenario 2.

Additionally, we incorporated attention checks that assessed the participant's attentiveness to the questions. Unfortunately, one of the attention checks included in the experiment was phrased unclearly. To be more specific, the question: "Which organisations were mentioned in the emails you have seen in the first scenario?" failed to indicate that two answers needed to be selected. As a consequence, a significant number of participants only selected one organisation. Hence, we made the decision to exclude this question from the experiment.

After the attention checks, some questions were asked about one's experiences and attitudes with emails, to determine whether these had an influence on their decision making. Afterwards, we asked the participants whether they would use the PostGuard add-on and we gave the possibility to leave feedback about the experiment. At the end of the experiment, we gave the participants some information about the study. In Figure 5.5, a flowchart of the experiment can be seen. The entire online survey is included in Appendix E.

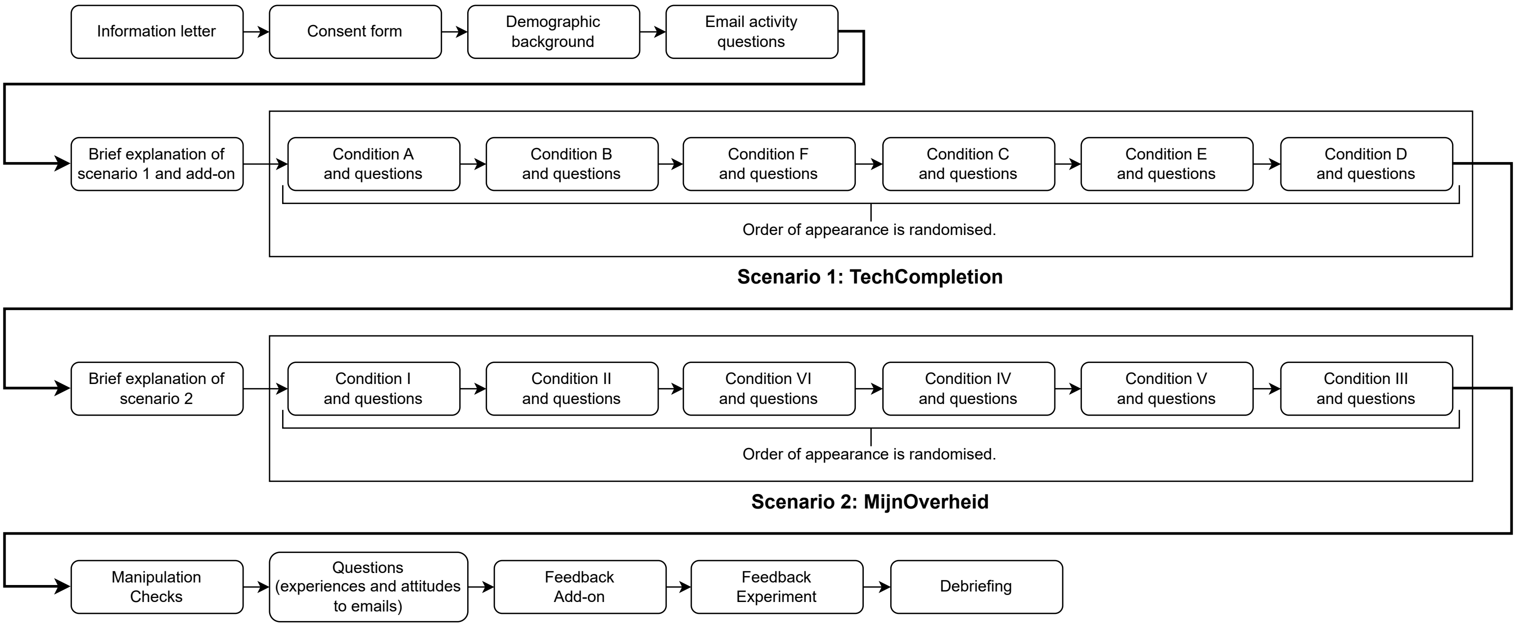

Information letter
Consent form
Demographic background
Email activity questions
Brief explanation of scenario 1 and add-on
Condition A and questions
Condition B and questions
Condition F and questions
Condition C and questions
Condition E and questions
Condition D and questions
Order of appearance is randomised.
Scenario 1: TechCompletion
Brief explanation of scenario 2
Condition I and questions
Condition II and questions
Condition VI and questions
Condition IV and questions
Condition V and questions
Condition III and questions
Order of appearance is randomised.
Scenario 2: MijnOverheid
Manipulation Checks
Questions (experiences and attitudes to emails)
Feedback Add-on
Feedback Experiment
Debriefing

Figure 5.5: Flowchart of the experiment.
Source: Own work.

5.2.2.2 Hypotheses

In this part of the thesis, we will list our hypotheses regarding the dependent variables before analysis is performed on the responses of the experiment. We do this per design, where there are three types of identity-based digital signature designs. The designs can be seen in Section 4.2.

The warning designs i.e. the designs used for emails that are not signed decreases all five dependent variables

Although, we have seen works in the related works section stating that warnings can be ineffective, we do still expect that all five dependent variables decrease. This is because of the lack of warnings in email clients Hu and Wang (2018), therefore due to unfamiliarity we believe that these warning designs will grab the attention of users. On top of that, Hu and Wang (2018) found out that email warning banners have a positive impact on reducing risky user actions. Hence, we expect that all five dependent variables will decrease.

Relevant ID verifications design increases all five dependent variables
If we examine other works studying the presence of ID verifications more generally, thus not in the context of emails, we see that it increases the users' perceived credibility and trustworthiness of the source (Li & Liang, 2022; Zloteanu et al., 2018; Morris et al., 2012), this is also supported by the following posts of Cohen (2022) and Kolaja (2021). Hence, we think that all five dependent variables will increase.

Irrelevant ID verifications design decrease all five dependent variables
Irrelevant ID verifications can occur in both legitimate and malicious emails. However, we ideally want legitimate emails to have relevant ID verifications. Therefore, we have chosen to create the design to influence the user to be cautious by including a warning message. We believe that it results in a slight decrease of all five dependent variables.

Note that all of these hypotheses are also based on the expectation that participants of the experiment will focus more on finding the details and look more carefully at the ID verifications than in real life scenarios. And is caused by them being aware that they are participating in a email credibility experiment.

5.2.2.3 Measurement scales

For measuring the credibility of the email message and the sender, we use Likert scales so that participants can choose the best option that corresponds to how they feel about a presented question or statement. We have chosen to not use Visual-Analog Rating (VAS) scales, which is a rating scale in the form of a slider. This is because various studies have shown that the Likert scale was preferred by evaluators because of its simplicity and usability; both type of scales produce mostly similar results (Kuhlmann et al., 2017; Dourado et al., 2021; Buskirk, 2015; van Laerhoven et al., 2004).

The following questions were asked in both scenarios:

- **Email interaction:**
  In scenario 1, the participant was asked a binary yes/no question on whether they would click the link in the email message, followed by an optional open-ended question asking which aspects of the email message contributed to their response.

  In scenario 2, the participant was asked a binary yes/no question on whether they would send a reply to the email message, followed by an optional open-ended question asking which aspects of the email message contributed to their response.

- **Email interaction comfort:**
  The questions of email interaction were followed by a 7-point Likert

question on how comfortable they would feel if they were to perform the asked action depending on the scenario (replying or clicking a link).

- **Email message credibility:**
The participant had to judge on how believable, accurate and professional they found the content of the email message to be using a 7-point Likert scale ranging from “Describes very poorly” to “Describes very well” with three items (Appelman & Sundar, 2016).

- **Email sender credibility:**
Here, a participant had to judge on how trustworthy, credible and reputable they perceive the sender of the email message using a 7-point Likert scale ranging from “Not at all” to “Extremely” with three items (Metzger et al., 2020; Flanagin & Metzger, 2000).

- **Email sender certainty:**
Participants were asked how certain they were that the email originated from the mentioned sender using a 7-point Likert scale ranging from “Not at all” to “Extremely”. Here, the participants also had to specify which sections of the email message influenced their decision via an optional open-ended question.

5.2.2.4 Participants

We conducted a trial involving three participants to gather feedback on the experiment, minimizing errors and improving the flow prior to public release. Some important feedback points were:

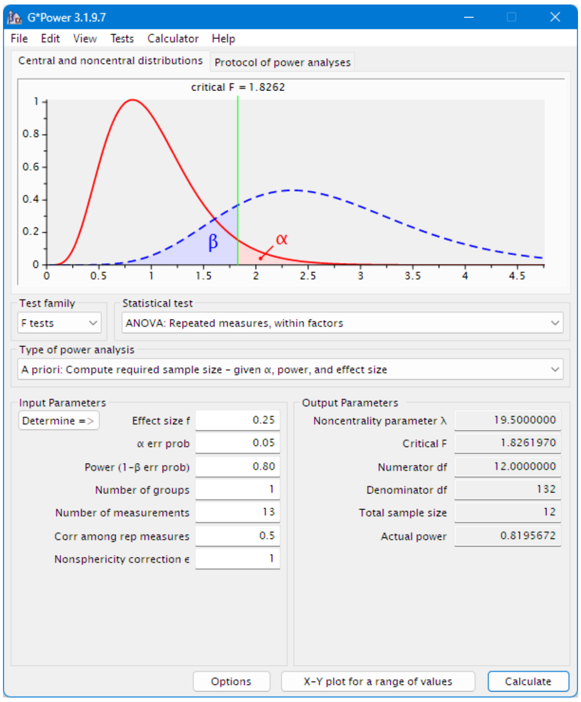
- Survey was lengthy, therefore some questions were dismissed, some were made optional and large blocks of text were made more concise to reduce reading time.

- Some questions had to be rephrased due to ambiguity and needed more introduction.

- Email message of scenario 2 needed some adjustments to match with the questions asked.

The results of these participants were not used and therefore deleted. We solely conducted this trial to obtain feedback that is used to improve the experiment. For the actual experiment, we acquired participants through LinkedIn and the computing science mailing list of Radboud University. Additionally, we asked via LinkedIn whether people could repost the survey to their LinkedIn network. This resulted in 21 complete responses without applying the attention checks. As stated earlier, one of the attention checks included in the experiment was phrased unclearly. This check was excluded from the experiment. Participants that answered the attention checks wrongly or indicated that they did not know or have forgotten what the experiment was about, were left out. This resulted in 20 complete responses.

The minimum number of required participants was calculated with the software tool GPower (Faul et al., 2007; Faul et al., 2009). A screenshot of the inputted values can be seen in Figure 5.6. Here we can see that the required total sample size was 12. Therefore, with 20 responses we have passed the threshold.



Test family: F tests

Statistical test: ANOVA: Repeated measures, within factors

Type of power analysis: A priori: Compute required sample size – given α, power, and effect size

Input Parameters		Output Parameters	
Effect size f	0.25	Noncentrality parameter λ	19.5000000
α err prob	0.05	Critical F	1.8261970
Power (1-β err prob)	0.80	Numerator df	12.0000000
Number of groups	1	Denominator df	132
Number of measurements	13	Total sample size	12
Corr among rep measures	0.5	Actual power	0.8195672
Nonsphericity correction ϵ	1		

Determine =>

Figure 5.6: G*Power calculation for obtaining the required sample size. Source: Own work.

The survey took on average approximately 1858 seconds, which is about 31 minutes (SD ≈ 1215, min = 504.66s, max = 5551.2s) where the median was around 1670 seconds.

The participants were in the age group 18-24 (80%) or 25-34 (20%). and the majority had completed a Bachelor's degree (50%) or some high school diploma or equivalent (45%). A large part of the participants read on average 1-3 email message per day (40%) or 4-6 (30%). Where the majority sent on average 0 emails per day (50%), followed by 1-3 emails (40%).

In Appendix C, one can find the complete descriptive statistics of the participants.

Chapter 6

Results

In this chapter, the responses obtained from the aforementioned conducted experiment are examined and analysed to answer our research question, namely "**What is the effect of identity-based digital signatures on email credibility?**". This was done by performing analyses separate on the two scenarios of the experiment. First, we performed statistical tests for the responses of scenario 1 and consequently, we performed the same procedure for the responses of scenario 2. The statistical tests were conducted using the guides of Laerd Statistics: Cochran's Q test (Laerd Statistics, 2018), Friedman test (Laerd Statistics, 2015) and Shapiro-Wilk's test (Laerd Statistics, n.d.). Before we conducted the statistical tests, we first checked whether the assumptions held.

6.1 Results of scenario 1 (TechCompletion)

Before we present the results, we will briefly mention the conditions of scenario 1. Scenario 1 consisted of 7 conditions. The origin and legitimacy of the email plus the identity-based digital signature design that was presented in each condition can be seen in Table 6.1.

Condition	Sender origin	Authenticity	Design
A	Internal	Legitimate	None
B	External	Legitimate	None
C	External	Malicious	Warning (external)
D	External	Malicious	Irrelevant ID verifications
E	Internal	Legitimate	Warning (not signed)
F	External	Legitimate	Relevant ID verifications
G	Internal	Legitimate	Relevant ID verifications

Table 6.1: Conditions of scenario 1 (TechCompletion).

To find out what effects the identity-based digital signature designs had on all five dependent variables, statistical tests were conducted. We first addressed the dependent variable email interaction. This is followed by statistical analyses on email interaction comfort, email message credibility, email sender credibility and lastly, email sender certainty.

Email interaction

This question addressed whether the user would click the link in the email and log in. Overall users indicated that they would click the links in 4/40 malicious emails and click the links of 50/100 legitimate emails. The distribution of the responses for this question can be seen in Figure 6.1.

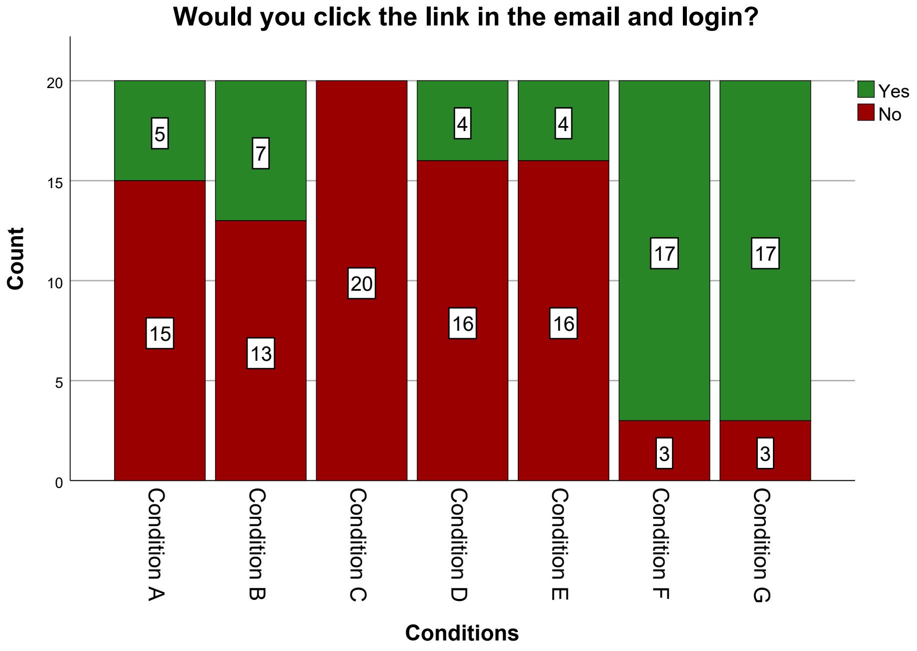


Figure 6.1: Scenario 1: Distribution of email interaction (N=20).

As the plot suggests, click-through rates differ per condition. This was confirmed by Cochran's Q test, $\chi^2(6)= 63.82$, $p < .001$. Which specific conditions differed was investigated using a post-hoc analysis, namely pairwise comparisons using Dunn's procedure with a Bonferroni correction for multiple comparisons. We found out that F and G (Relevant ID verifications design) differed significantly from all other conditions (all $p < .05$), while not differing from each other ($p = 1.00$), and is in line with what we had observed.

Email interaction comfort

Email interaction comfort was measured using a 7-point Likert scale ranging from 1 to 7 (higher values indicate higher credibility; $M = 3.51$, $SD = 2.29$), along with a qualitative question where one could write down their reasoning behind the provided answer. The distribution of the answers can be seen in Figure 6.2.

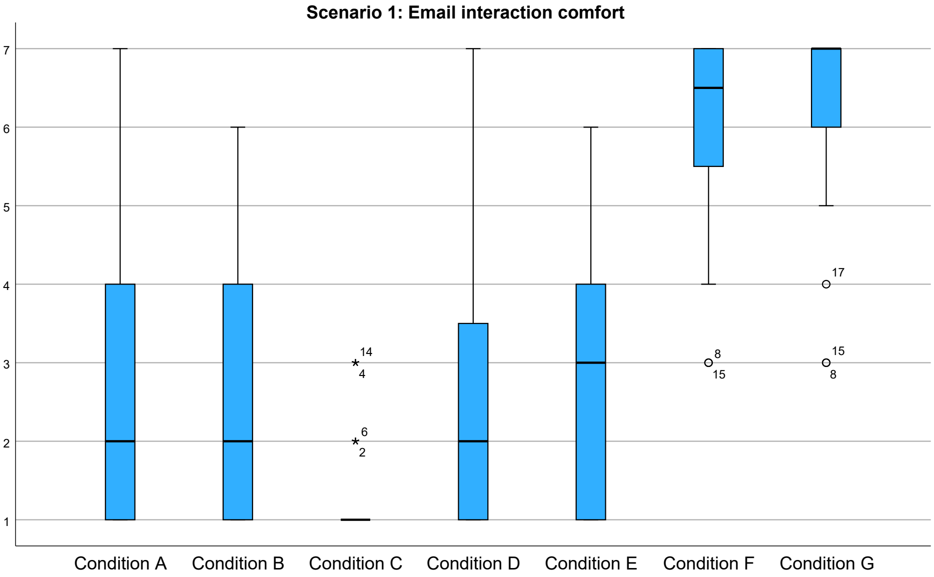


Figure 6.2: Scenario 1: Distribution of answers to email interaction comfort question (N=20).

We see that conditions F and G ($M = 6.08$, $SD = 1.33$) differ from conditions A-E ($M = 2.48$, $SD = 1.71$). This was confirmed by Friedman's test, $\chi^2(6)= 77.11$, $p < .001$, where Shapiro-Wilk's test revealed that the data was not normally distributed, $p < .05$. Which specific conditions differed was investigated using pairwise comparisons with a Bonferroni correction for multiple comparisons. Here, we also had that conditions F and G (relevant ID verifications) statistically significantly differed from all other conditions (all $p < .05$), while not differing from each other ($p = 1.00$).

After that, we performed a thematic analysis for the qualitative question concerning email interaction comfort (Mortensen, 2021). We found out that all the cases could be categorized in the following themes with some examples of the responses:

- **Sender identity verification and confirmation**
  "ID has been verified so I assume everything is fine" and "The ID is not verified".

- **Sender characteristics**
  "The ID verifications seem to match the info i have. It is verified to come from my company, country of origin checks out etc".
- **Suspicious email content**
  "The email seems reputable. But why would they need to provide me a link to the website of the company that I work at?"
- **Sender trust**
  "Confirmed to be from my CEO".

Here, the most important theme was "Sender identity verification and confirmation". The majority of the cases stated that their comfort in interacting with the emails was influenced by the absence or presence of the identity-based digital signature or the information given by the digital signature. The analysis also showed that email interaction comfort was influenced by the sender's email address or the content of the email message. A more detailed version of the analysis can be found in Appendix D.1.

Email message credibility

Email message credibility was measured using a 7-point Likert scale with 3 items ranging from 1 to 7 (higher values indicate higher credibility; $M = 4.73$, $SD = 1.63$). We have converted this into one score for each participant by taking the average of the scores of the three items. The distribution of the answers can be seen in Figure 6.3.

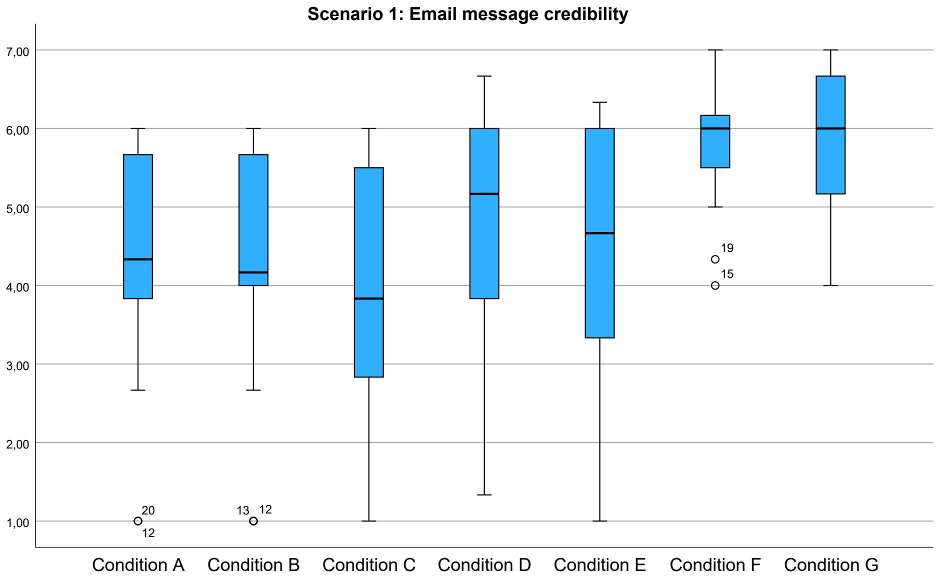


Figure 6.3: Scenario 1: Distribution of email message credibility (N=20).

We see that condition G ($M = 5.87,\ SD = 0.94$) differs from conditions A-E ($M = 4.28,\ SD = 1.64$) and that condition F ($M = 5.88,\ SD = 0.80$) differs from condition C ($M = 3.85,\ SD = 1.70$).

With Shapiro-Wilk's test ($p < .05$), we assessed that the data of email message credibility was not normally distributed. Our observations were confirmed by Friedman's test, $\chi^2(6)= 49.58,\ p < .001$. Which specific conditions differed was investigated using pairwise comparisons with a Bonferroni correction for multiple comparisons. Here, we have seen that condition G (relevant ID verifications) statistically significantly differed from conditions A-E (all $p < .05$). On top of that, condition F significantly differed from condition C ($p < .05$).

Email sender credibility

Email sender credibility was measured using a 7-point Likert scale with 3 items ranging from 1 to 7 (higher values indicate higher credibility; $M = 4.00,\ SD = 1.94$). We have converted this into one score for each participant by taking the average of the scores of the three items. The distribution of the answers can be seen in Figure 6.4.

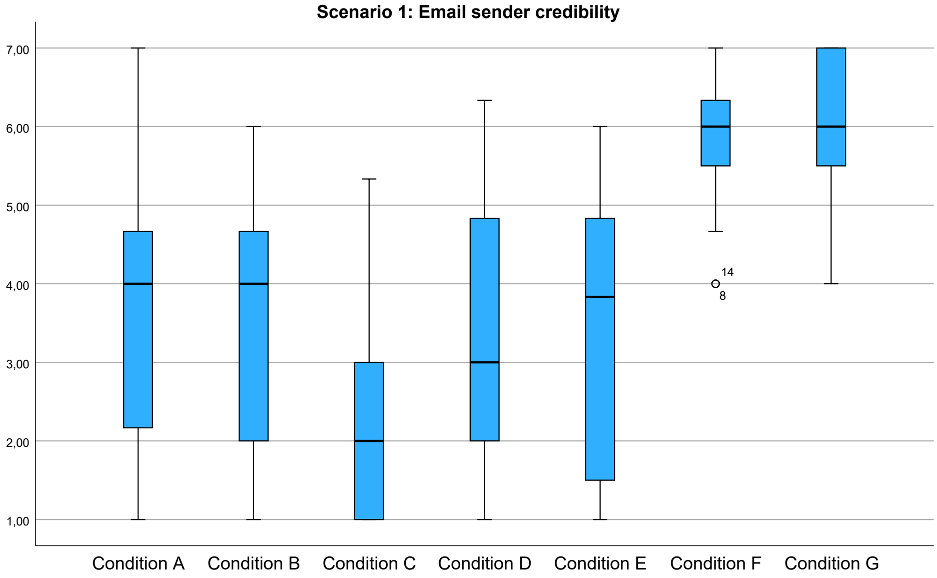


Figure 6.4: Scenario 1: Distribution of email sender credibility (N=20).

We see that conditions F and G ($M = 5.88,\ SD = 0.84$) differ from conditions A-E ($M = 3.24,\ SD = 1.73$). This was confirmed by Friedman's test, $\chi^2(6)= 72.29,\ p < .001$, where Shapiro-Wilk's test revealed that the data was not normally distributed, $p < .05$. Which specific conditions differed was investigated using pairwise comparisons with a Bonferroni correction for multiple comparisons. Here, we also had that conditions F and G (relevant ID verifications) statistically significantly differed from all other conditions (all $p < .05$), while not differing from each other ($p = 1.00$).

Email sender certainty

Email sender certainty was measured using a 3-point scale ranging from 1 to 3 (No, certainly/I do not know/Yes, certainly; $M = 2.14$, $SD = 0.74$) along with a qualitative question that evaluated the reasoning behind the provided answer. The distribution can be seen in Figure 6.5.

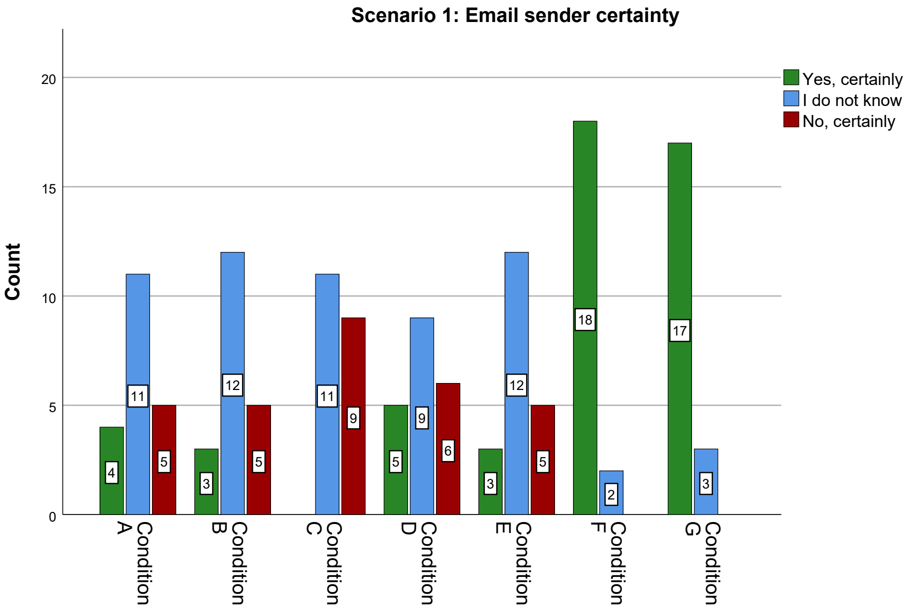


Figure 6.5: Scenario 1: Distribution of email sender certainty (N=20).

We see that conditions F and G ($M = 2.88$, $SD = 0.33$) differ from conditions A-E ($M = 1.85$, $SD = 0.66$). This was confirmed by Friedman's test, $\chi^2(6)= 67.70$, $p < .001$, where Shapiro-Wilk's test revealed that the data was not normally distributed, $p < .05$. Which specific conditions differed was investigated using pairwise comparisons with a Bonferroni correction for multiple comparisons. Here, we also had that conditions F and G (relevant ID verifications) statistically significantly differed from all other conditions (all $p < .05$), while not differing from each other ($p = 1.00$).

The thematic analysis for the qualitative question resulted in the following main themes that influence email sender certainty with some examples of the responses:

- **Influence of identity-based digital signature design**
  "The CEO would not send an email such as this, and there is no postguard id verification" and "ID verifications and email address corresponds to the scenario".

- **Sender characteristics**
  "This person has been verified to be John Doe, who is the CEO of the company I work for. I would consider him trustworthy".

- **Email content and context**
  "He doesn't have the ID confirmation label. Also, why is someone outside of my company informing me which website I need to do my job? I should know that myself better than he does".

- **Sender familiarity and trust**
  "The email corresponds to the sender".

- **Need for external confirmation**
  "Yes, it is verified, but why would the CEO send such an email? I would either ask them in person or call them by phone to clarify that".

Here, the most important theme was "Influence of identity-based digital signature design". The majority stated that their confidence in the email being sent by the mentioned sender was mainly influenced by the absence or presence of the identity-based digital signature in combination with irrelevant or relevant ID verifications. The sender's email address and the content of the email message were also important factors. A more detailed version of the analysis can be seen in Appendix D.2.

6.2 Results of scenario 2 (MijnOverheid)

We will first name all the conditions of scenario 2. Scenario 2 consisted of 6 conditions. The origin and legitimacy of the email plus the identity-based digital signature design that was presented in each condition can be seen in Table 6.2.

Condition	Sender origin	Authenticity	Design
I	External	Legitimate	None
II	External	Malicious	None
III	External	Malicious	Warning (external)
IV	External	Malicious	Irrelevant ID verifications
V	External	Legitimate	Irrelevant ID verifications
VI	External	Legitimate	Relevant ID verifications

Table 6.2: Conditions of scenario 2 (MijnOverheid).

Email interaction

This question addressed whether the user would click the link in the email and log in. Overall users indicated that they would reply to 14/60 malicious emails and reply to 30/60 legitimate emails. The distribution of the responses for this question can be seen in Figure 6.6.

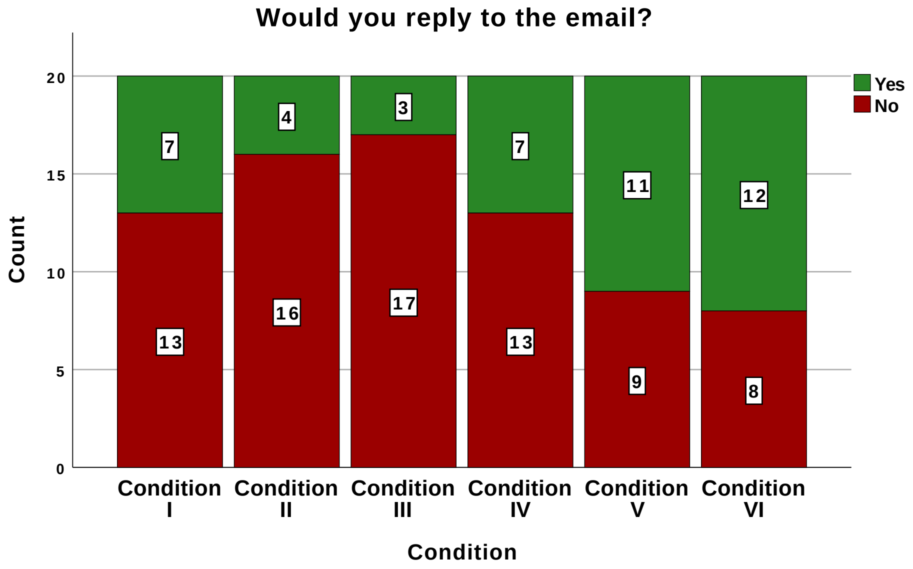


Figure 6.6: Scenario 2: Distribution of email sender credibility (N=20).

As the plot suggests, click-through rates differ per condition. This was confirmed by Cochran's Q test, $\chi^2(5)$= 23.33, p < .001. Which specific conditions differed was investigated using a post-hoc analysis, namely pairwise comparisons using Dunn's procedure with a Bonferroni correction for multiple comparisons. We found out that condition V (Irrelevant ID verifications design) and VI (Relevant ID verifications design) differed significantly from conditions II and III (all $p < .05$), while not differing from each other ($p = 1.00$).

Email interaction comfort

Email interaction comfort was measured using a 7-point Likert scale ranging from 1 to 7 (higher values indicate higher credibility; $M = 3.49$, $SD = 2.19$), along with a qualitative question where one could write down their reasoning behind the provided answer. The distribution of the answers can be seen in Figure 6.7.

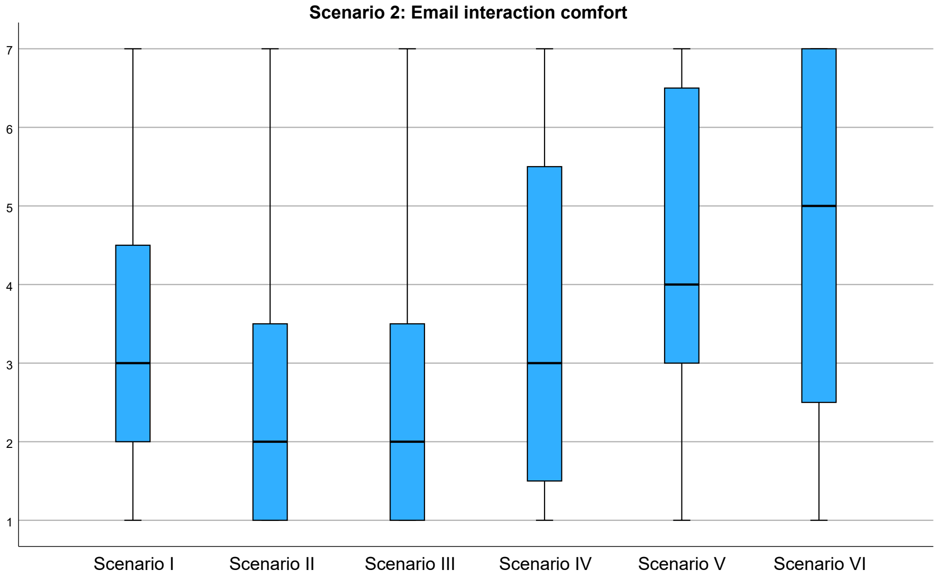


Figure 6.7: Scenario 2: Distribution of answers to email interaction comfort question (N=20).

We see that conditions V and VI (M = 4.45, SD = 2.22) differ from conditions II and III (M = 2.50, SD = 1.89). This was confirmed by Friedman's test, $\chi^2(5)$= 44.48, $p < .001$, where Shapiro-Wilk's test revealed that the data was not normally distributed, $p < .05$. Which specific conditions differed was investigated using pairwise comparisons with a Bonferroni correction for multiple comparisons. Here, we also had that condition V (Irrelevant ID verifications design) and VI (Relevant ID verifications design) differed significantly from conditions II and III (all $p < .05$), while not differing from each other ($p = 1.00$).

After that, we performed a thematic analysis for the qualitative question concerning email interaction comfort (Mortensen, 2021). We found out that all the cases could be categorized in the following main themes with some examples of the participants' responses:

- **Email consequences**
  "No real harm can be done by replying to a email".

- **Verification and legitimacy**
  "Confirmed to be from mijn overheid" and "The email address domain matches the original one from the government. The verifications also match.".

- **Suspicious origin and content**
  "This is generally not how MijnOverheid behaves, especially stressing

that this action is important seems odd. I would log in to MijnOverheid or google".

, Here, the most important theme was "Verification and legitimacy". The majority stated that their comfort in interacting with emails was influenced by the absence or presence of the identity-based digital signature or the information given by the digital signature. A large part also mentioned that the email content was unusual based on their experiences and knowledge about emails of MijnOverheid and that this affected their comfort in interacting with the email. On top of that, a part of the participants stated that there were no serious consequences to replying to an email. A more detailed version of the analysis can be found in Appendix D.3.

Email message credibility

Email message credibility was measured using a 7-point Likert scale with 3 items ranging from 1 to 7 (higher values indicate higher credibility; $M = 4.41,\ SD = 1.76$). We have converted this into one score for each participant by taking the average of the scores of the three items. The distribution of the answers can be seen in Figure 6.8.

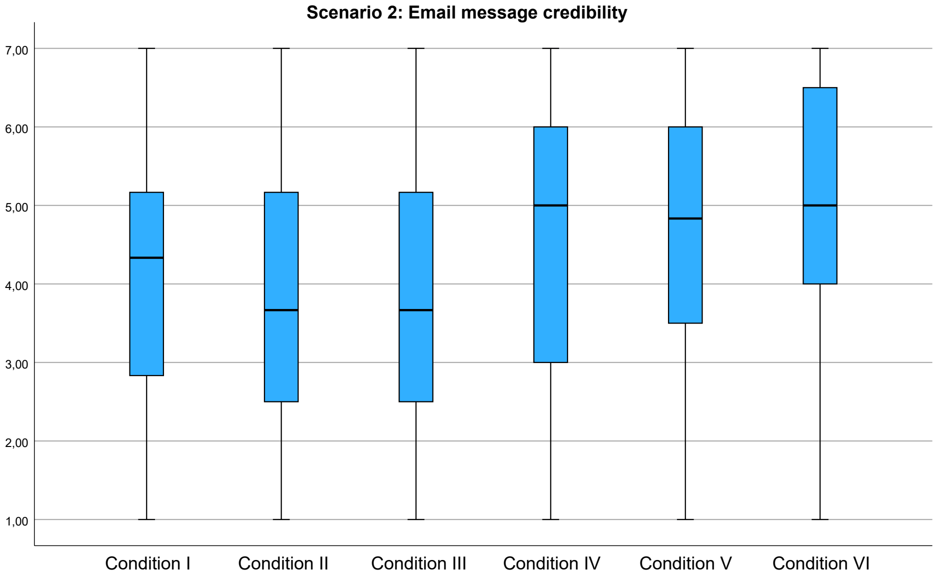


Figure 6.8: Scenario 2: Distribution of email message credibility (N=20).

We see that condition VI ($M = 5.13,\ SD = 1.59$) differs from conditions III ($M = 3.88,\ SD = 1.93$).

With Shapiro-Wilk's test ($p < .05$), we assessed that the data of email message credibility was not normally distributed. Our observations were confirmed by Friedman's test, $\chi^2(5)= 21.51,\ p < .001$. Which specific conditions differed was investigated using pairwise comparisons with a Bonferroni correction for multiple comparisons. Here, we have seen that condition VI (relevant ID verifications) statistically significantly differed from condition III ($p < .05$).

Email sender credibility

Email sender credibility was measured using a 7-point Likert scale with 3 items ranging from 1 to 7 (higher values indicate higher credibility; $M = 3.84,\ SD = 1.96$). We have converted this into one score for each participant by taking the average of the scores of the three items. The distribution of the answers can be seen in Figure 6.9.

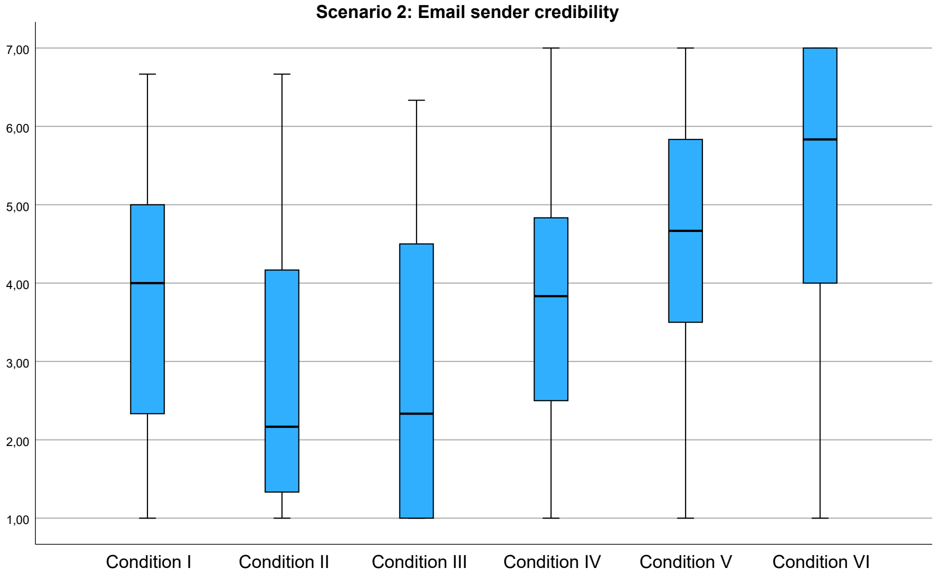


Figure 6.9: Scenario 2: Distribution of email sender credibility (N=20).

We see that conditions V and VI ($M = 4.87,\ SD = 1.87$) differ from conditions II and III ($M = 2.87,\ SD = 1.79$). This was confirmed by Friedman's test, $\chi^2(5)= 44.71,\ p < .001$, where Shapiro-Wilk's test revealed that the data was not normally distributed, $p < .05$. Which specific conditions differed was investigated using pairwise comparisons with a Bonferroni correction for multiple comparisons. Here, we also had that condition V (Irrelevant ID verifications design) and VI (Relevant ID verifications design) differed significantly from conditions II and III (all $p < .05$), while not differing from each other ($p = 1.00$).

Email sender certainty

Email sender certainty was measured using a 3-point scale ranging from 1 to 3 (No, certainly/I do not know/Yes, certainly; $M = 1.98$, $SD = 0.75$) along with a qualitative question that evaluated the reasoning behind the provided answer. The distribution can be seen in Figure 6.10.

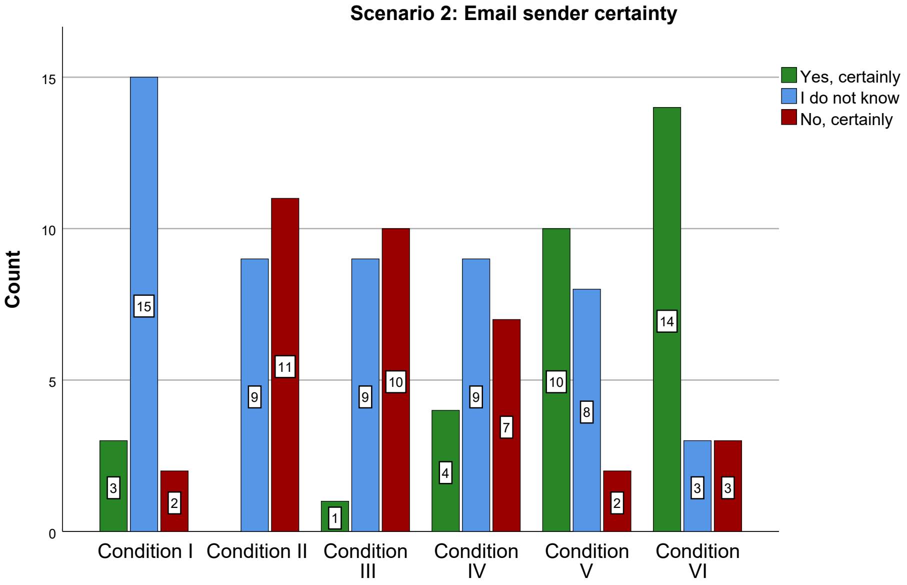


Figure 6.10: Scenario 2: Distribution of email sender certainty (N=20).

We see that conditions V and VI ($M = 2.48$, $SD = 0.72$) differ from conditions II and III ($M = 1.50$, $SD = 0.55$). This was confirmed by Friedman's test, $\chi^2(5)= 49.45$, $p < .001$, where Shapiro-Wilk's test revealed that the data was not normally distributed, $p < .05$. Which specific conditions differed was investigated using pairwise comparisons with a Bonferroni correction for multiple comparisons. Here, we also had that condition V (Irrelevant ID verifications design) and VI (Relevant ID verifications design) differed significantly from conditions II and III (all $p < .05$), while not differing from each other ($p = 1.00$).

The thematic analysis for the qualitative question resulted in the following main themes that influence email sender certainty with some examples of the responses:

- **Verification concerns**
  "You can see from the verifications that the email originates from the government, and also which domain it has and the email address, and all of this matches" and "Sender seems alright, but no verification banner".

- **Suspicious email domain and content** "MijnOverheid doesn't ask you to send emails or click links".

Here, the most important theme was "Verification concerns". The majority stated that their certainty of the email being sent by the mentioned sender was influenced by the sender's email address and the content of the email message. The identity-based digital signature and ID verifications also played a role, but the content of the email raised a lot of doubts. A more detailed version of the analysis can be seen in Appendix D.4.

6.3 Experiences and attitudes with emails

Experiences

The majority of the participants (75%) have received a social engineering message before. Therefore, it highlights the importance of assisting users in identifying malicious emails so that harmful consequences can be prevented.

Attitudes and advice

We have also asked a question about what advise a participant would give to their colleagues about opening links or attachments in emails. We have summarized these responses into the following points:

- Pay attention to the sender, email subject, email content, style of writing, spelling, grammar and email address and check for suspiciousness. Contact the IT department or organisation when you are uncertain about the legitimacy of the email or confirm the sender's identity by asking the person who is being impersonated or coworkers.

- Be wary of opening links and attachments in emails, especially if the sender is not known to you or if an email seems suspicious e.g. asking you to log in or handing over some information. Only click URLs when you are absolutely certain that you know the sender.

- Stay informed about the latest phishing and cyber-attack tactics, and educate yourself on how to recognize and avoid them.

- Never click links or open attachments unless you were expecting an email with a link or attachment and when you a certain that it came from a sender you know. Never automatically open attachments.

- Think about the risks. What is exactly being asked of you and what are consequences of performing that particular action.

6.4 Expected add-on use

Furthermore, we also asked the participants whether they would make use of the identity-based digital signature feature of PostGuard. From our 20 participants 60% said “Yes”, 5% said “No” and 35% said “Maybe”. From these numbers, it is clear that the majority of participants would use this feature and that a large part is still hesitant.

6.5 Feedback about the experiment

This was the last qualitative question of the survey, and allowed the participants to leave feedback about the experiment. The majority of answers were either empty or indicated that they did not encounter any issues during the experiment. Some participants mentioned that the survey had no indication of progress, and that the questions per email were quite long.

Chapter 7

Discussion and conclusions

In this chapter, we will interpret the results of our online survey to understand the impact of the designs on email credibility. Essentially, to find out whether identity-based digital signatures can be used to counter email-based social engineering attacks. After that, we put these results into the context of earlier related research. Furthermore, we discuss the limitations of our research and provide some directions for future research in this field. Lastly, we present the conclusions of this thesis.

7.1 Discussion

From the results of both qualitative and quantitative questions, we have seen that the sender's email address and the content of the email played a role in determining the credibility of an email. Hence, we found out that these are not neglected in the process of judging an email's credibility, even when the identity-based digital signature design was present. However, the participants might have been more focused and cautious than usual, since they were participating in an experiment about malicious emails.

We found out that the relevant ID verifications design had an effect on email credibility that was in line with our hypothesis i.e *Relevant ID verifications design increases all five dependent variables.* For the conditions of scenario 1 ("TechCompletion") where relevant ID verifications were present (conditions F and G), we have seen an significant increase in email interaction, email interaction comfort, email message credibility, email sender credibility and email sender certainty compared to the control conditions. This was also the case for scenario 2 ("MijnOverheid"). Here, we also had an increase in all dependent variables. The increase was proportionally less significant compared to scenario 1. This was likely caused by the email of scenario 2, which was seen by the majority as suspicious, even when the email came from the legitimate email address. The relevant ID verifications gave the participants the confirmation that the email's sender and message were legitimate, therefore we can conclude that this type of design worked as intended.

For the warning designs of scenario 1 (condition C and E) and scenario 2 (condition III), we had an decrease in all five dependent variables when compared to the control conditions, except for control condition II of scenario 2. This control condition scored even lower in most of the dependent variables than condition III i.e. the malicious email with the warning design. For the warning designs, we see that in most cases a decrease in the dependent variables, which is in line with our hypothesis: *The warning designs i.e. the designs used for emails that are not signed decreases all five dependent variables.* However, the differences were not statistically significant, therefore we do not have enough evidence to conclude that the decrease in dependent variables was caused by the warning. We have seen that it decreases email credibility in both scenarios, but not statistically significantly when compared to the control conditions. We could say that the design likely does work as intended, but needs to be tested in different conditions.

Regarding the irrelevant ID verifications design of scenario 1 (condition D) and scenario 2 (condition IV and V): for scenario 1, we either had a decrease or no change in the dependent variables except for email message credibility when compared to the control conditions. For scenario 2, we saw a slight increase in most of the dependent variables when compared to the control conditions. However, none of these differences were statistically significant. Thus we do not have enough evidence to support or reject the hypothesis: *Irrelevant ID verifications design decreases all five dependent variables.* We can conclude that the design did not completely work as intended, as we saw a decrease and an increase in email credibility, depending on the presented condition. Ideally, we want the irrelevant ID verifications design to consistently decrease email credibility, and as observed this was not the case.

Now that we have addressed our hypotheses, we will continue with discussing the designs in relation with the results.

The design for irrelevant ID verification was quite challenging in terms of conveying the correct message to the user and assisting the user to make the correct decision. This is because the design had to be applicable on malicious emails as well as legitimate ones. This was supported by the results of the qualitative questions. Participants mentioned that one still had to look carefully at the ID verifications to determine whether the sender was legitimate. It was also mentioned that they could see how irrelevant ID verifications could be used by a malicious actor to trick targets into believing that it originated from a legitimate sender. However, our identity-based digital signature designs do make the process of deceiving users much harder, since one cannot forge ID verifications. If identity-based digital signatures are highly adopted, then irrelevant ID verifications would likely be the weakest aspect of the design in countering email-based social engineering attacks, since this is were the design could be misused to manipulate victims. Furthermore, persons can accidentally or unknowingly include irrelevant ID verifications in their legitimate email. From our results, we have seen that legitimate emails with irrelevant ID verifications were seen as less credible and trustworthy than emails with relevant ID verifications. Malicious emails with irrelevant ID verifications were seen as more credible in scenario 2 and less credible in scenario 1. Therefore, additional research is required in finding an optimal way for users to include ID verifications so that user errors are minimized and sufficient information is provided to the users. In other words, the design used in combination with irrelevant ID verification requires more research so that an effective design can be created.

The control conditions were perceived by the majority as malicious or suspicious, even when the email message originated from a legitimate email address. However, there were participants that mentioned that a legitimate email address increased the credibility of the emails of the control conditions. Also, we have seen that the content of the email had an influence on the perceived credibility of an email. Especially in the control conditions of scenario 2 (condition I and II). Most of the participants experienced these emails as suspicious, as could be seen in the statistics of email sender certainty. Condition randomisation might have played a role in these results, since it is possible that some participants felt uncomfortable after first seeing a condition with an identity-based digital signature design, and then a condition without the design. The results underscore the significance of attaining a high adoption rate, as the absence of the design on legitimate emails could lead recipients to perceive them as malicious.

On top of that, the warning design used for not signed emails from a legitimate sender was perceived as untrustworthy. This is likely to be caused by the warning message that the design conveys, and is in line with the work of Levine (2014). This paper states that people as a default start with the belief that others' communication is honest, but that certain events can trigger them to abandon this state (trust), in this case, the warning message.

Regarding the advice that participants gave about opening links and attachment in emails, all of these were very reasonable. Overall, we have not noticed any misconceptions about email security. An interesting remark is that a participant answered the qualitative question regarding email certainty with the following: “I would overlook it [information about the sender] quicker, because there is no tool at all.”. This shows that our design likely grabs the attention of the participants, since it makes information about the sender more apparent. Therefore, our visual designs might be a promising start.

Lastly, we have seen from the responses of the qualitative questions that the majority of participants would follow the advice that the design presented. With these results, we have examined identity-based digital signatures could be implemented in PostGuard and its effectiveness in countering email-based social engineering attacks. We have seen that the relevant ID verifications design was effective, therefore a similar or identical design could be used for when this feature is implemented in PostGuard. We also have shown that the irrelevant ID verifications and warning designs require more attention, so that the desired results can be achieved. We suggest developers of email clients or email security software to pay more attention to the design of irrelevant ID verifications, since we expect that this design is going to be used by social engineers.

7.1.1 Relation to previous research

As we had seen in Section 3.2 and Section 3.3, passive warnings were considered ineffective due to not being able to grab the attention of the user, habituation or ignorance. We do not have enough evidence to conclude that our warning designs were ineffective or effective, but we have seen that the designs did successfully force the participant's attention to the ID verifications. This might be explained by the fact that participants were briefly introduced to identity-based digital signatures, and might have encouraged them to look for the ID banner and verifications.

We believe that when the proposed warning designs are tested in more scenarios, the outcomes will be in line with the research of Hu and Wang (2018), where they found warnings to have a positive impact in reducing risky user actions. This is because we have observed email interaction to decrease compared to the control conditions when the warning design was present. Take as an example: the warning design for not signed legitimate emails shown in scenario 1 (condition E). If we specifically look at email interaction for this condition, we see that the majority would not interact with the email, even though it is legitimate. This is very likely to be caused by the warning message. However, this was not a statistically significant difference. We think that this is caused by the control conditions being perceived as less credible than we initially thought they would be.

7.1.2 Limitations of the study

We had chosen to first introduce the concept of identity-based digital signatures to the participant before they saw them. This means that we do not know how these designs would perform when they are shown to uninformed participants.

Another limitation was the content of the email messages used in the conditions of scenario 2. A legitimate email of MijnOverheid would probably not ask for a reply to the email, and would not be written in English. A better option would have been to choose a different fictional or real organisation for the emails of scenario 2. Also, the legitimate version of the

MijnOverheid email looked a bit questionable since MijnOverheid would in reality never send an email similar to the ones we have created. This is because we had to keep everything the same for the research and a legitimate email would not contain the same content as a malicious mail, and vice versa.

Additionally, there was a minor mistake in condition IV of scenario 2. Here, the ID banner mentioned mijn.overheid.com instead of mijnoverheid.com. It was mentioned in the qualitative questions, and therefore might have affected the results.

Also, we did not take the first manipulation question into consideration when filtering the responses. This is because the question was not formulated correctly, making it ambiguous. We did not specify that exactly two organisations had to be selected, resulting in a large part of the participants selecting only one organisation.

The duration of the online survey was relatively lengthy for the participants. We saw that a lot of partial responses and some feedback regarding the survey length. This might have affected the results.

On top of that, our sample was not very representative. Most of the participants were in the age group 18-24 (80%) where no participants were older than 34 years old.

Also, it should be noted that we designed the mockups of identity-based digital signatures for the email security add-on PostGuard. We believe that email client add-ons can face adoption rate problems, because the usage of email client add-ons is entirely optional and might go unnoticed.

In the questions about the participants' experiences with social engineering attacks, we have not asked whether any of the participants experienced negative consequences or harm as a result of these messages. So, we cannot conclude whether past experiences had an influence on a participant's behaviour. Also, it is worth mentioning that we kept this question a bit more general since the message could be from other communication media than email.

Regarding the identity-based signature design, one participant disagreed on how this would be a security boost, because one still has to look very carefully in order to make sure that the emails are authentic.

There was also feedback on the emails of scenario 2 (MijnOverheid). There was some uncertainty about the possibility of English emails from MijnOverheid. Which made the emails of MijnOverheid uncomfortable for them. It also indicates to a certain extent that the email message of scenario 2 did not seem very realistic. Some clarification might have made this more clear.

The results of scenario 2 could have been affected by a lack of motivation or concentration, since it appeared after scenario 1 which had seven conditions. Additionally, the content of the email likely had an effect on the results of scenario 2, because some participants were familiar with MijnOverheid. These participants used their own experiences and knowledge to determine that the email asked an unusual action i.e. replying to the email which MijnOverheid normally would not ask. Also, the email was written in English which is unusual for an email from MijnOverheid.

Lastly, after the experiment was conducted, we found out that the randomisation sequence of the conditions could not be retrieved for each of the participants. Therefore, no analysis could be performed on the effect of different sequences on the dependent variables or results.

7.1.3 Direction for future research

A lot of research can still be done on identity-based digital signatures for emails. First of all, we have created mobile designs of these signatures in this thesis. These have not been tested due to keeping the survey length as short as possible. These mobile designs could be examined in future work.

Additionally, the emails presented in the conditions of this experiment were a static image. One could research the effects of identity-based digital signatures in a more interactive environment to perceive more accurate and realistic results, since one can interact with email messages. This could be tested in combination with features that we have seen in related works, such as haptics or sound alerts.

Lastly, the proposed designs in this research can be improved upon, take as an example the message of the ID banner. A participant mentioned that the message of the ID banner for signed emails, "this email has been verified to come from . . . [email domain]", did not indicate that the email originated from the person or organisation. Therefore, further research is required on finding the optimal message for the designs. Moreover, the design should help the user in identifying malicious emails without requiring too much effort. In our case, participants still had to look very carefully on whether an email was not malicious. One could for example make more distinct designs for irrelevant and relevant ID verifications. Also, the design for not signed legitimate emails can be improved so that the majority of legitimate emails will not be perceived as not credible if the adoption rate is low.

7.2 Conclusion

While identity-based digital signature may not be the sole solution to counter email-based social engineering attacks, this research has set an important step in assisting users in assessing the credibility of email messages. Our findings indicate that relevant ID verifications enhance email credibility by assisting the recipient in confirming the email's sender. Furthermore, we found out that the effects of warning designs and irrelevant ID verifications require further research. The warning designs appear to have the potential to decrease email credibility. On top of that, this research highlights that users evaluate email credibility not only by considering the design of the identity-based digital signature, but also by email content and the sender's email address. These findings are important for the development of email identity-based digital signatures and will further assist us in combating email-based social engineering attacks.

Bibliography

Radicati, S. (2022). Email statistics report, 2022-2026. *The Radicati Group, Inc.* https://www.radicati.com/wp/wp-content/uploads/2022/11/Email-Statistics-Report-2022-2026-Executive-Summary.pdf

Stine, K., & Scholl, M. (2010). E-mail security. an overview of threats and safeguards. *J AHIMA*, *81*(4), 28–30, quiz 31.

ENISA. (2022). What is social engineering? https://www.enisa.europa.eu/topics/incident-response/glossary/what-is-social-engineering

IBM. (n.d.). What is social engineering? https://www.ibm.com/topics/social-engineering

Mouton, F., Leenen, L., Malan, M. M., & Venter, H. S. (2014). Towards an ontological model defining the social engineering domain. In K. Kimppa, D. Whitehouse, T. Kuusela, & J. Phahlamohlaka (Eds.), *Ict and society* (pp. 266–279). Springer Berlin Heidelberg.

Mitnick Security. (2022). Are social engineering attacks on the rise? https://www.mitnicksecurity.com/blog/are-social-engineering-attacks-on-the-rise

Proofpoint. (2023). What is social engineering? - definition, types & more. https://www.proofpoint.com/us/threat-reference/social-engineering

Rouse, G. (2022). Common types of social engineering attacks (2022). https://www.datto.com/blog/common-types-of-social-engineering-attacks

Salahdine, F., & Kaabouch, N. (2019). Social engineering attacks: A survey. *Future Internet*, *11*(4). https://doi.org/10.3390/fi11040089

Fuertes, W., Arévalo, D., Castro, J., Ron, M., Estrada, C., Andrade, R., Peña, F., & Benavides, E. (2022). Impact of social engineering attacks: A literature review. https://doi.org/10.1007/978-981-16-4884-7_3

Rock, T. (2021). Understanding the real danger of social engineering. https://invenioit.com/security/danger-of-social-engineering/

Witts, J. (2023). Mobile email security: How we can stay secure using smartphones for email. https://expertinsights.com/insights/mobile-email-security-how-we-can-stay-secure-using-smartphones-for-email/

Wagenseil, P. (2022). Security vs. user experience: Finding the balance. https://www.scmagazine.com/resource/identity-and-access/security-vs-user-experience-finding-the-balance

Brown, J. (2019). The art of balancing user experience and security. https://usabilitygeek.com/user-experience-and-security/

Menlo Security. (2022). Security vs. user experience. https://www.menlosecurity.com/blog/security-vs-user-experience-87-say-user-experience-is-what-counts/

HackControl. (n.d.). What is the impact of social engineering attacks? https://hackcontrol.org/cases/what-is-the-impact-of-social-engineering-attacks/

Partida, D. (2020). Social engineering cyberattacks and how they're impacting businesses. https://www.securityinfowatch.com/cybersecurity/article/21203580/social-engineering-cyberattacks-and-how-theyre-impacting-businesses

Australian Research Data Commons. (2022). Standardised communications protocols. https://ardc.edu.au/resource/standardised-communications-protocols/

Subedi, H. (2022). Basics of computer networking: Communication protocols. https://www.itjones.com/blogs/basics-of-computer-networking-communication-protocols

CloudFlare. (n.d.-a). What is a packet? https://www.cloudflare.com/learning/network-layer/what-is-a-packet/

IBM. (2023). Tcp/ip protocols. https://www.ibm.com/docs/en/aix/7.2?topic=protocol-tcpip-protocols

Microchip. (n.d.). Tcp/ip five layer software model terminology reference. https://microchipdeveloper.com/local--files/tcpip:tcp-ip-five-layer-model/layer%5C_terminology.JPG

Zwicky, E. D., & Chapman, D. B. (1995). *Building internet firewalls* (1995th ed.). O'Reilly Media.

An, S. (2015). Data encapsulation terminology. https://notes.shichao.io/icnd1/figure_2-11.png

Oracle. (2015). How a packet travels through the tcp/ip stack. https://docs.oracle.com/cd/E18752_01/html/816-4554/figures/ipov.fig88.png

Tucker, C. (2020). The osi model – the 7 layers of networking explained in plain english. https://www.freecodecamp.org/news/osi-model-networking-layers-explained-in-plain-english/

Alpern, N. J., & Shimonski, R. J. (2010). Chapter 5 - the osi model and networking protocols. In N. J. Alpern & R. J. Shimonski (Eds.), *Eleventh hour network+* (pp. 73–88). Syngress. https://doi.org/10.1016/B978-1-59749-428-1.00006-0

Cofense. (2023). Urls 4x more likely than phishing attachments to reach users. https://cofense.com/blog/urls-4x-more-likely-than-phishing-attachments-to-reach-users/

Elghamrawy, K. (n.d.). The smtp protocol. https://d6x8u9i2.rocketcdn.me/blog/wp-content/uploads/2017/11/SMTP-sequence-diagram.png

Pollock, W. (2016). How mail works on the internet. http://wpollock.com/AUnix2/EmailDiagram.png

Ashtari, H. (2023). Imap vs. pop3: 4 leading differences you should know. https://www.spiceworks.com/tech/tech-general/articles/imap-vs-pop3/

Diffie, W., & Hellman, M. (1976). New directions in cryptography. *IEEE Transactions on Information Theory*, *22*(6), 644–654. https://doi.org/10.1109/TIT.1976.1055638

Göthberg, D. (2006). Public key encryption. https://commons.wikimedia.org/wiki/File:Public_key_encryption.svg

FlippyFlink. (2019). Private key signing. https://upload.wikimedia.org/wikipedia/commons/7/78/Private_key_signing.svg

Okta. (2022). What is public key infrastructure (pki); how does it work? https://www.okta.com/identity-101/public-key-infrastructure/

Xolphin. (n.d.). Certificate authority (ca). https://www.xolphin.com/support/Terminology/Certificate%5C_Authority%5C_(CA)

MATTR. (n.d.). Web of trust 101. https://learn.mattr.global/docs/concepts/web-of-trust-101

Plesky, E. (2022). Tls vs ssl: What is the right choice? https://www.plesk.com/blog/various/tls-vs-ssl-what-is-the-right-choice/

Möller, B., Duong, T., & Kotowicz, K. (2014). This poodle bites: Exploiting the ssl 3.0 fallback. https://www.openssl.org/~bodo/ssl-poodle.pdf

Perrig, A. (n.d.). How pgp works. https://users.ece.cmu.edu/~adrian/630-f04/PGP-intro.html

Kay, R. (2008). Identity-based encryption. https://www.computerworld.com/article/2551479/identity-based-encryption.html

Wikipedia contributors. (2023). Identity-based encryption. https://en.wikipedia.org/w/index.php?title=Identity-based_encryption&oldid=1149227008

Youngblood, C. (2005). An introduction to identity-based cryptography. https://courses.cs.washington.edu/courses/csep590/06wi/finalprojects/youngblood_csep590tu_final_paper.pdf

Sheffer, Y. (2009). Identity based encryption steps. https://commons.wikimedia.org/wiki/File:Identity_Based_Encryption_Steps.png

Blank, S., Goldstein, P., Loder, T., Zink, T., Bradshaw, M., & Brotman, A. (2022). *Brand Indicators for Message Identification (BIMI)* (Internet-Draft draft-brand-indicators-for-message-identification-02) [Work in Progress]. Internet Engineering Task Force. Internet Engineering Task Force. https://datatracker.ietf.org/doc/draft-brand-indicators-for-message-identification/02/

IRMA. (n.d.). Irma session flow. https://irma.app/docs/assets/irmaflow.png

Krombholz, K., Hobel, H., Huber, M., & Weippl, E. (2015). Advanced social engineering attacks [Special Issue on Security of Information and Networks]. *Journal of Information Security and Applications*, *22*, 113–122. https://doi.org/10.1016/j.jisa.2014.09.005

Heartfield, R., & Loukas, G. (2016). A taxonomy of attacks and a survey of defence mechanisms for semantic social engineering attacks. *ACM Computing Surveys*, *48*. https://doi.org/10.1145/2835375

McAfee. (2022). Phishing email examples: How to recognize a phishing email. https://www.mcafee.com/learn/phishing-email-examples-how-to-recognize-a-phishing-email/

Steinberg, J. (2019). Why scammers make spelling and grammar "mistakes". https://josephsteinberg.com/why-scammers-make-spelling-and-grammar-mistakes/

JCU Australia. (n.d.). How to avoid phishing attacks. https://www.jcu.edu.au/information-and-communications-technology/secure-it/how-to-avoid-phishing-attacks

SALT Cyber Security. (2023). Phishing. https://www.linkedin.com/pulse/phishing-salt-cyber-security/

Tessian. (2021). Phishing vs. spear phishing examples. https://www.tessian.com/wp-content/uploads/2019/12/WebsiteBlog-What-Is-Spear-Phishing_-Example-2.jpg.jpg

Valimail. (n.d.). What is clone phishing: How it works, examples & defenses. https://www.valimail.com/guide-to-phishing/clone-phishing/

Eemeli. (2022). Clone phishing. https://www.hoxhunt.com/blog/clone-phishing

Digital Check. (2021). Two simple rules that can spot nearly every email phishing scam. https://www.digitalcheck.com/how-to-spot-phishing-scams/

Microsoft 365 Defender Threat Intelligence Team. (2021). Trend-spotting email techniques: How modern phishing emails hide in plain sight. *Microsoft*. https://www.microsoft.com/en-us/security/blog/2021/08/18/trend-spotting-email-techniques-how-modern-phishing-emails-hide-in-plain-sight/

Bolster. (2022). What is domain spoofing? https://bolster.ai/blog/what-is-domain-spoofing

CloudFlare. (n.d.-b). What is domain spoofing? | website and email spoofing. https://www.cloudflare.com/learning/ssl/what-is-domain-spoofing/

PowerDMARC. (n.d.). What is typosquatting in cybersecurity. https://powerdmarc.com/what-is-typosquatting/

Sawabe, Y., Chiba, D., Akiyama, M., & Goto, S. (2019). Detection method of homograph internationalized domain names with ocr. *Journal of Information Processing*, *27*, 536–544. https://doi.org/10.2197/ipsjjip.27.536

Hu, H., & Wang, G. (2018). End-to-End measurements of email spoofing attacks. *27th USENIX Security Symposium (USENIX Security 18)*, 1095–1112. https://www.usenix.org/conference/usenixsecurity18/presentation/hu

Dedenok, R. (2021). Email spoofing: How attackers impersonate legitimate senders. *Securelist.* https://securelist.com/email-spoofing-types/102703/

Watson, G. (2014). Chapter 3 - the techniques of manipulation. In G. Watson, A. Mason, & R. Ackroyd (Eds.), *Social engineering penetration testing* (pp. 39–63). Syngress. https://doi.org/10.1016/B978-0-12-420124-8.00003-X

Sainio, R. (2022). Scenario 5: I'm planning a surprise. https://www.hoxhunt.com/blog/pretexting-is-a-simple-and-effective-phishing-attack

Baraniuk, C. (2017). Google and facebook duped in huge 'scam'. *BBC.* https://www.bbc.com/news/technology-39744007

Trend Micro. (2019). Google and facebook fraudster pleads guilty to $100 million scam. https://www.trendmicro.com/vinfo/fr/security/news/cybercrime-and-digital-threats/google-and-facebook-fraudster-pleads-guilty-to-100-million-scam

Law, E. L.-C., Roto, V., Hassenzahl, M., Vermeeren, A. P., & Kort, J. (2009). Understanding, scoping and defining user experience: A survey approach. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 719–728. https://doi.org/10.1145/1518701.1518813

Roto, V., Law, E.-C., Vermeeren, A., & Hoonhout, J. (2011). *User experience white paper: Bringing clarity to the concept of user experience.* s.n.

Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems approach (Standard ISO 9241-210:2019). (2019). International Organization for Standardization. Geneva, CH. https://www.iso.org/standard/62711.html

Indeed Editorial Team. (2022). What is user interface (ui)? https://www.indeed.com/career-advice/career-development/user-interface

Kurosu, M., & Kashimura, K. (1995). Apparent usability vs. inherent usability: Experimental analysis on the determinants of the apparent usability. *Conference Companion on Human Factors in Computing Systems*, 292–293. https://doi.org/10.1145/223355.223680

Hannah, J. (2021). What is a user interface & what are the key elements? https://careerfoundry.com/en/blog/ui-design/what-is-a-user-interface/

Usability.gov. (2014). User interface design basics. https://www.usability.gov/what-and-why/user-interface-design.html

Nielsen, J. (2020). 10 usability heuristics for user interface design. *NN/g Nielsen Norman Group.* https://www.nngroup.com/articles/ten-usability-heuristics/

Brandon, M., Schraffenberger, H. K., Sluis-Thiescheffer, W., van der Geest, T., Ostkamp, D., & Jacobs, B. (2022). Design principles for actual security. *Adjunct Proceedings of the 2022 Nordic Human-Computer Interaction Conference*. https://doi.org/10.1145/3547522.3547684

Malisa, Luka. (2017). *Security of user interfaces: Attacks and countermeasures* (Doctoral dissertation). ETH Zurich. https://doi.org/10.3929/ETHZ-B-000217453

Morales, J. (2021). Mobile first design strategy: The when, why and how. https://xd.adobe.com/ideas/process/ui-design/what-is-mobile-first-design/

van Rijn, J. (2023). The ultimate mobile email statistics overview. https://www.emailmonday.com/mobile-email-usage-statistics/

Petelka, J., Zou, Y., & Schaub, F. (2019). Put your warning where your link is: Improving and evaluating email phishing warnings. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3290605.3300748

Stojkovski, B. (2022). *User experience design for cybersecurity & privacy: Addressing user misperceptions of system security and privacy* (Doctoral dissertation). University of Luxembourg.

Google. (n.d.). Check if a site's connection is secure. https://support.google.com/chrome/answer/95617?hl=en

Hunt, T. (2018). The decreasing usefulness of positive visual security indicators (and the importance of negative ones). https://www.troyhunt.com/the-decreasing-usefulness-of-positive-visual-security-indicators-and-the-importance-of-negative-ones/

Laughery, K., & Wogalter, M. (2014). A three-stage model summarizes product warning and environmental sign research. *Safety Science - SAF SCI*, *61*. https://doi.org/10.1016/j.ssci.2011.02.012

Stojkovski, B., & Lenzini, G. (2020). Evaluating ambiguity of privacy indicators in a secure email app. *Italian Conference on Cybersecurity*.

Dhamija, R., Tygar, J. D., & Hearst, M. (2006). Why phishing works. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 581–590. https://doi.org/10.1145/1124772.1124861

Felt, A. P., Reeder, R. W., Ainslie, A., Harris, H., Walker, M., Thompson, C., Acer, M. E., Morant, E., & Consolvo, S. (2016). Rethinking connection security indicators, 1–13.

Volkamer, M., Renaud, K., & Gerber, P. (2016). Spot the phish by checking the pruned url. *Information and Computer Security*, *24*, 372–385. https://doi.org/10.1108/ICS-07-2015-0032

Schechter, S. E., Dhamija, R., Ozment, A., & Fischer, I. (2007). The emperor's new security indicators. *2007 IEEE Symposium on Security and Privacy (SP '07)*, 51–65. https://doi.org/10.1109/SP.2007.35

Datta, P., Namin, A. S., Jones, K. S., & Hewett, R. (2021). Warning users about cyber threats through sounds. *SN Applied Sciences*, *3*(7). https://doi.org/10.1007/s42452-021-04703-4

Cooper, M., Levy, Y., Wang, L., & Dringus, L. (2021). Heads-up! an alert and warning system for phishing emails. *Organizational Cybersecurity Journal: Practice, Process and People*, *ahead-of-print*. https://doi.org/10.1108/OCJ-03-2021-0006

Egelman, S., Cranor, L. F., & Hong, J. (2008). You've Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings. https://doi.org/10.1184/R1/6626570.v1

Junger, M., Montoya, L., & Overink, F. (2017). Priming and warnings are not effective to prevent social engineering attacks. *Computers in human behavior*, *66*, 75–87. https://doi.org/10.1016/j.chb.2016.09.012

Min, W. (2006). Fighting phishing at the user interface.

Wu, M., Miller, R. C., & Garfinkel, S. L. (2006). Do security toolbars actually prevent phishing attacks? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 601–610. https://doi.org/10.1145/1124772.1124863

Aneke, J., Ardito, C., & Desolda, G. (2020). Designing an intelligent user interface for preventing phishing attacks. In J. Abdelnour Nocera, A. Parmaxi, M. Winckler, F. Loizides, C. Ardito, G. Bhutkar, & P. Dannenmann (Eds.), *Beyond interactions* (pp. 97–106). Springer International Publishing.

Volkamer, M., Renaud, K., Reinheimer, B., & Kunz, A. (2017). User experiences of torpedo: Tooltip-powered phishing email detection. *Computers & Security*, *71*, 100–113. https://doi.org/10.1016/j.cose.2017.02.004

Anderson, B. B., Vance, A., Kirwan, C. B., Jenkins, J. L., & Eargle, D. (2016). From warning to wallpaper: Why the brain habituates to security warnings and what can be done about it. *Journal of Management Information Systems*, *33*(3), 713–743. https://doi.org/10.1080/07421222.2016.1243947

Ebert, N., Ackermann, K. A., & Bearth, A. (2022). When information security depends on font size: How the saliency of warnings affects protection behavior. *Journal of Risk Research*, *26*(3), 233–255. https://doi.org/10.1080/13669877.2022.2142952

Herley, C. (2009). *So long, and no thanks for the externalities: The rational rejection of security advice by users* (NSPW, tech. rep. MSR-TR-2009-46) [NSPW]. Association for Computing Machinery, Inc. https://www.microsoft.com/en-us/research/publication/so-long-and-no-thanks-for-the-externalities-the-rational-rejection-of-security-advice-by-users/

Krol, K., Moroz, M., & Sasse, M. A. (2012). Don't work. can't work? why it's time to rethink security warnings. *2012 7th International Conference* *on Risks and Security of Internet and Systems (CRiSIS)*. https://doi.org/10.1109/crisis.2012.6378951

Cisco. (n.d.). Url rewriting and analysis (using outbreak filters). https://docs.ces.cisco.com/docs/url-rewriting-and-analysis

S., H. (2020). Url analysis: How to determine maliciousness. https://hustlelead.medium.com/url-analysis-how-to-determine-maliciousness-f630b4e51b9e

Bhardwaj, P. (2021). What is url blacklist? https://www.tutorialspoint.com/what-is-url-blacklist

Whaley, E. (2020). Security for emails: Rethinking url rewriting. *Vade*. https://www.vadesecure.com/en/blog/rethinking-url-rewriting-in-email-security

UC San Diego. (2020). Example of rewritten url when embedded. https://blink.ucsd.edu/technology/email/security/url-defense.html#Example-of-Rewritten-URL-When-E

MailXaminer. (n.d.). Email content analysis software. https://www.mailxaminer.com/email-content-analysis.html

ReversingLabs. (2019). Automated static analysis vs. dynamic analysis - better together? https://www.reversinglabs.com/blog/automated-static-analyis-vs.dynamic-analysis

Proofpoint. (n.d.-a). Malicious email attachments. https://www.proofpoint.com/us/threat-reference/malicious-email-attachments

Proofpoint. (n.d.-b). Targeted attack protection. https://www.proofpoint.com/us/products/advanced-threat-protection/targeted-attack-protection

Microsoft 365 Defender. (2023a). Safe attachments in microsoft defender for office 365. *Microsoft*. https://learn.microsoft.com/en-us/microsoft-365/security/office-365-security/safe-attachments-about?view=o365-worldwide

Egress. (n.d.). The quick guide to email content analysis. https://www.egress.com/blog/email-encryption/email-content-analysis

Franchina, L., Ferracci, S., & Palmaro, F. (2021). Detecting phishing e-mails using text mining and features analysis.

Poston, H. (n.d.). How to scan email headers for phishing and malicious content. https://resources.infosecinstitute.com/topic/how-to-scan-email-headers-for-phishing-and-malicious-content/

Gmail. (n.d.). Email that keeps your private information safe. https://safety.google/gmail/

Microsoft 365 Defender. (2023b). Microsoft defender for office 365 service description. *Microsoft*. https://learn.microsoft.com/en-us/office365/servicedescriptions/office-365-advanced-threat-protection-service-description

Microsoft 365 Defender. (n.d.). Secure messages by using a digital signature. *Microsoft*. https://support.microsoft.com/en-us/office/secure-messages-by-using-a-digital-signature-549ca2f1-a68f-4366-85fa-b3f4b5856fc6

Microsoft 365 Defender. (2017). What is a digital signature? *Microsoft*. https://learn.microsoft.com/en-us/previous-versions/office/office-2007-resource-kit/cc545901(v=office.12)?redirectedfrom=MSDN

Paunikar, C. (2021). Microsoft office 365 email security defaults are bad, so fix them! *Fractional CISO*. https://fractionalciso.com/microsoft-office-365-email-security-defaults-are-bad-so-fix-them/

Moecke, C., & Volkamer, M. (2013). Usable secure email communications - criteria and evaluation of existing approaches. *Information Management and Computer Security*, *21*, 41–52. https://doi.org/10.1108/09685221311314419

Reuter, A., Abdelmaksoud, A., Boudaoud, K., & Winckler, M. (2021). Usability of end-to-end encryption in e-mail communication. *Frontiers in Big Data*, *4*. https://doi.org/10.3389/fdata.2021.568284

Garfinkel, S., Margrave, D., Schiller, J., Nordlander, E., & Miller, R. (2005). How to make secure email easier to use. *Proc. SIGCHI conference on Human factors in computing systems*, 701–710. https://doi.org/10.1145/1054972.1055069

Budiu, R. (2018). Between-subjects vs. within-subjects study design. *NN/g Nielsen Norman Group*. https://www.nngroup.com/articles/between-within-subjects/

Field, A., & Hole, G. J. (2002). *How to design and report experiments*. SAGE Publications.

Bhandari, P. (2022). Within-subjects design | explanation, approaches, examples. https://www.scribbr.com/methodology/within-subjects-design/

Microsoft Security. (2022). What is business email compromise (bec)? https://www.microsoft.com/en-us/security/business/security-101/what-is-business-email-compromise-bec

Dyer, J. (2023). How social engineering attacks work (with examples). *Egress*. https://www.egress.com/blog/phishing/how-social-engineering-attacks-work-with-examples

Nanji, A. (2023). How much time do people typically spend looking at an email? *MarketingProfs*. https://www.marketingprofs.com/charts/2023/48537/how-much-time-do-people-typically-spend-looking-at-an-email

Li, J., & Liang, X. (2022). Reviewers' identity cues in online product reviews and consumers' purchase intention. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.784173

Zloteanu, M., Harvey, N., Tuckett, D., & Livan, G. (2018). Digital identity: The effect of trust and reputation information on user judgement in the sharing economy (J. A. Aimone, Ed.). *PLOS ONE*, *13*(12), e0209071. https://doi.org/10.1371/journal.pone.0209071

Morris, M. R., Counts, S., Roseway, A., Hoff, A., & Schwarz, J. (2012). Tweeting is believing? *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. https://doi.org/10.1145/2145204.2145274

Cohen, Z. (2022). How digital identities add a layer of trust to the internet. https://www.trulioo.com/blog/identity-verification/layer-of-trust

Kolaja, C. O. (2021). Why identity verification is critical for establishing credibility and maintaining integrity on social media platforms. https://www.linkedin.com/pulse/why-identity-verification-critical-establishing-media-oconnor-kolaja/

Kuhlmann, T., Dantlgraber, M., & Reips, U.-D. (2017). Investigating measurement equivalence of visual analogue scales and likert-type scales in internet-based personality questionnaires. *Behavior Research Methods*, *49*(6), 2173–2181. https://doi.org/10.3758/s13428-016-0850-x

Dourado, G. B., Volpato, G. H., de Almeida-Pedrin, R. R., Oltramari, P. V. P., Fernandes, T. M. F., & de Castro Ferreira Conti, A. C. (2021). Likert scale vs visual analog scale for assessing facial pleasantness. *American Journal of Orthodontics and Dentofacial Orthopedics*, *160*(6), 844–852. https://doi.org/10.1016/j.ajodo.2020.05.024

Buskirk, T. D. (2015). Are sliders too slick for surveys? an experiment comparing slider and radio button scales for smartphone, tablet and computer based surveys. *methods, data*, No 2 (2015). https://doi.org/10.12758/MDA.2015.013

van Laerhoven, H., van der Zaag-Loonen, H., & Derkx, B. (2004). A comparison of likert scale and visual analogue scales as response options in children's questionnaires. *Acta Paediatrica*, *93*(6), 830–835. https://doi.org/10.1111/j.1651-2227.2004.tb03026.x

Appelman, A., & Sundar, S. (2016). Measuring message credibility: Construction and validation of an exclusive scale. *Journalism and Mass Communication Quarterly*, *93*(1), 59–79. https://doi.org/10.1177/1077699015606057

Metzger, M. J., Hartsell, E. H., & Flanagin, A. J. (2020). Cognitive dissonance or credibility? a comparison of two theoretical explanations for selective exposure to partisan news. *Communication Research*, *47*(1), 3–28. https://doi.org/10.1177/0093650215613136

Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of internet information credibility. *Journalism & Mass Communication Quarterly*, *77*(3), 515–540. https://doi.org/10.1177/107769900007700304

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using g*power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Laerd Statistics. (2018). Cochran's q test using spss statistics. https://statistics.laerd.com/

Laerd Statistics. (2015). Friedman test using spss statistics. https://statistics.laerd.com/

Laerd Statistics. (n.d.). Testing for normality using spss statistics. https://statistics.laerd.com/spss-tutorials/testing-for-normality-using-spss-statistics.php

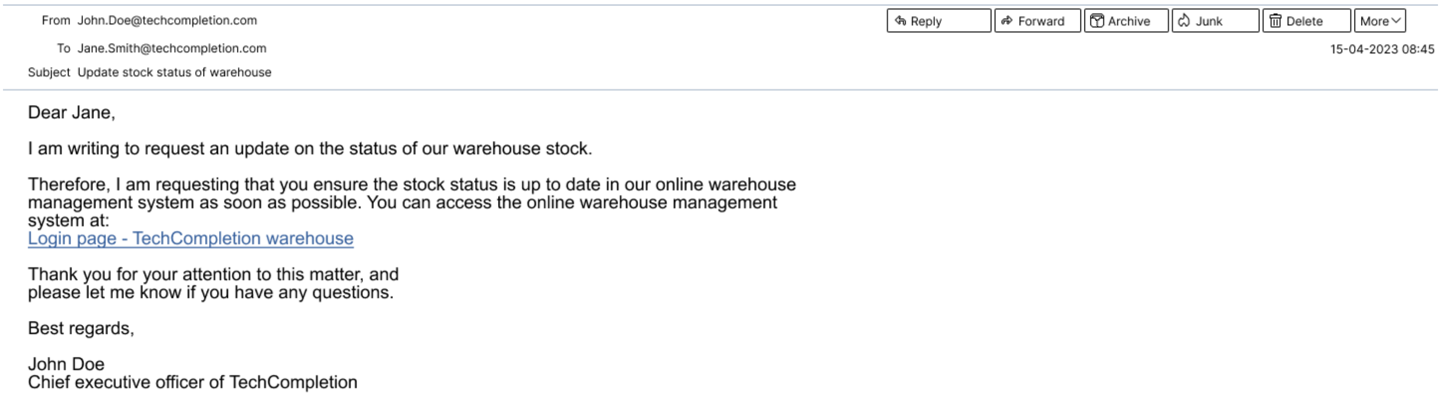
Mortensen, D. H. (2021). How to do a thematic analysis of user interviews. https://www.interaction-design.org/literature/article/how-to-do-a-thematic-analysis-of-user-interviews

Levine, T. R. (2014). Truth-default theory (tdt): A theory of human deception and deception detection [p. 386, Table 3. point 7.]. *Journal of Language and Social Psychology*, *33*(4), 378–392. https://doi.org/10.1177/0261927X14535916

Appendix A

Conditions

A.1 Scenario 1: TechCompletion



From John.Doe@techcompletion.com

To Jane.Smith@techcompletion.com

Subject Update stock status of warehouse

15-04-2023 08:45

Dear Jane,

I am writing to request an update on the status of our warehouse stock.

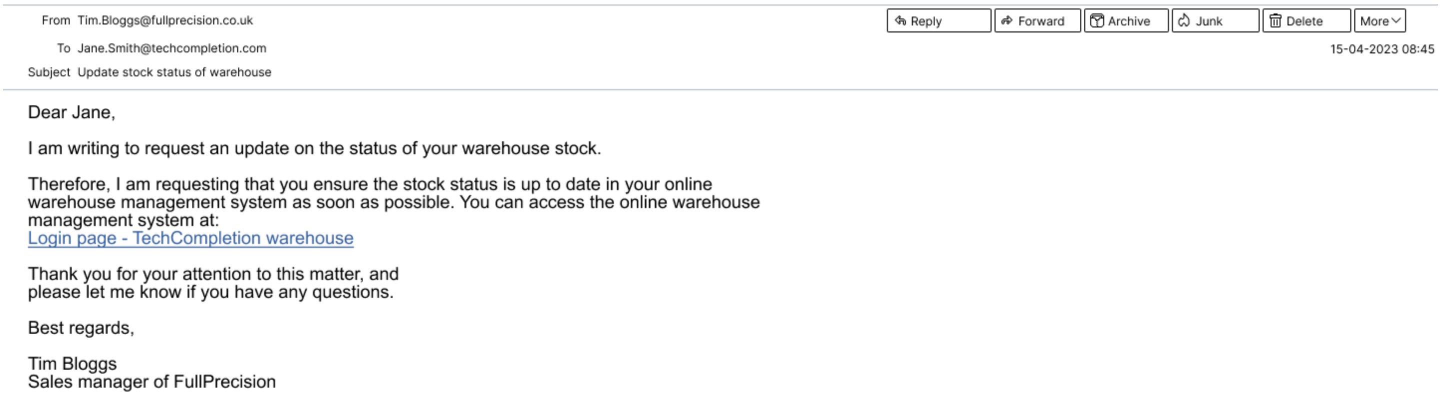
Therefore, I am requesting that you ensure the stock status is up to date in our online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and
please let me know if you have any questions.

Best regards,

John Doe
Chief executive officer of TechCompletion

Figure A.1: Condition A of scenario 1 - internal sender - none.
5



From Tim.Bloggs@fullprecision.co.uk

To Jane.Smith@techcompletion.com

15-04-2023 08:45

Subject Update stock status of warehouse

Dear Jane,

I am writing to request an update on the status of your warehouse stock.

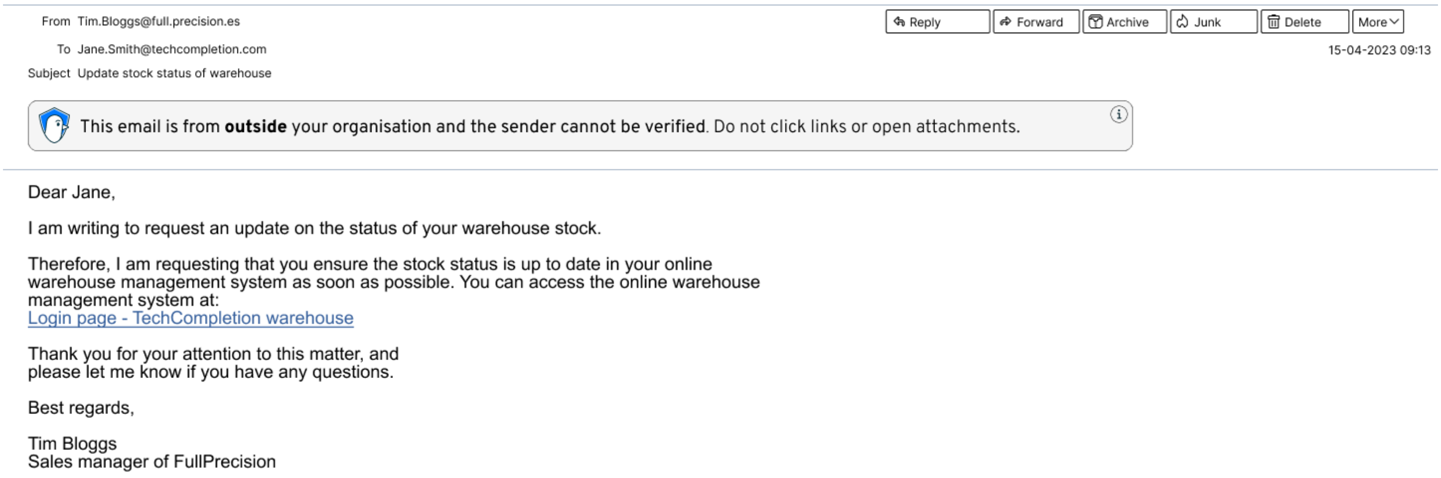
Therefore, I am requesting that you ensure the stock status is up to date in your online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and
please let me know if you have any questions.

Best regards,

Tim Bloggs
Sales manager of FullPrecision

Figure A.2: Condition B of scenario 1 external sender - none.



From Tim.Bloggs@full.precision.es

To Jane.Smith@techcompletion.com

15-04-2023 09:13

Subject Update stock status of warehouse

This email is from **outside** your organisation and the sender cannot be verified. Do not click links or open attachments.

Dear Jane,

I am writing to request an update on the status of your warehouse stock.

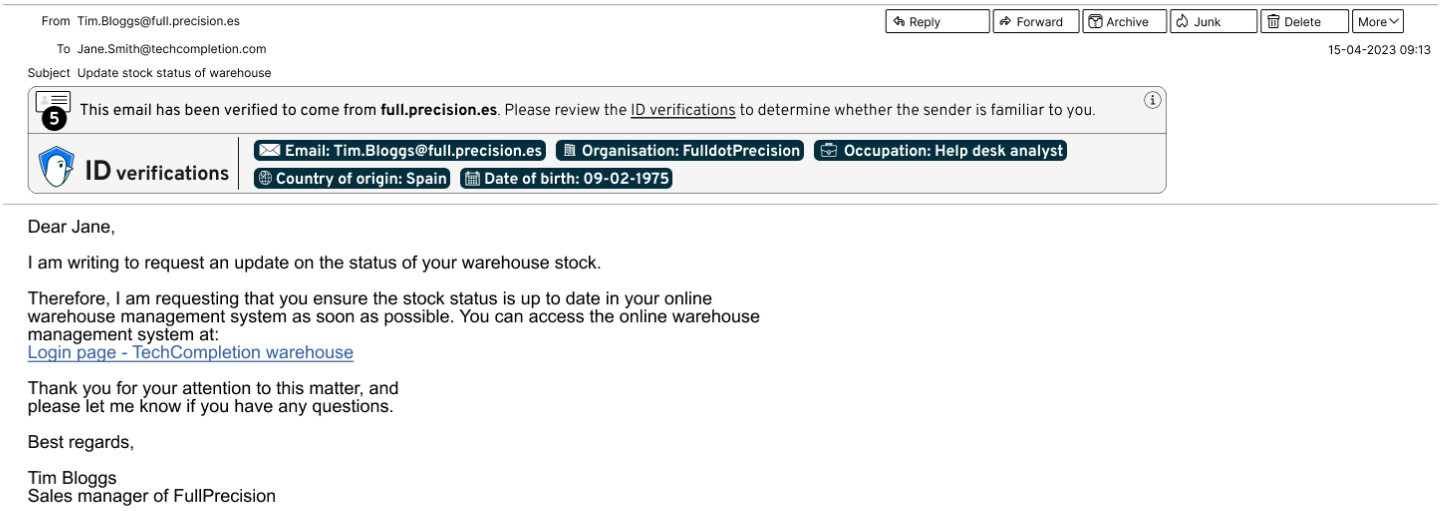
Therefore, I am requesting that you ensure the stock status is up to date in your online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and
please let me know if you have any questions.

Best regards,

Tim Bloggs
Sales manager of FullPrecision

Figure A.3: Condition C of scenario 1 - external sender - not signed.



From Tim.Bloggs@full.precision.es

To Jane.Smith@techcompletion.com

15-04-2023 09:13

Subject Update stock status of warehouse

This email has been verified to come from **full.precision.es**. Please review the ID verifications to determine whether the sender is familiar to you.

Email: Tim.Bloggs@full.precision.es | Organisation: FulldotPrecision | Occupation: Help desk analyst | Country of origin: Spain | Date of birth: 09-02-1975

Dear Jane,

I am writing to request an update on the status of your warehouse stock.

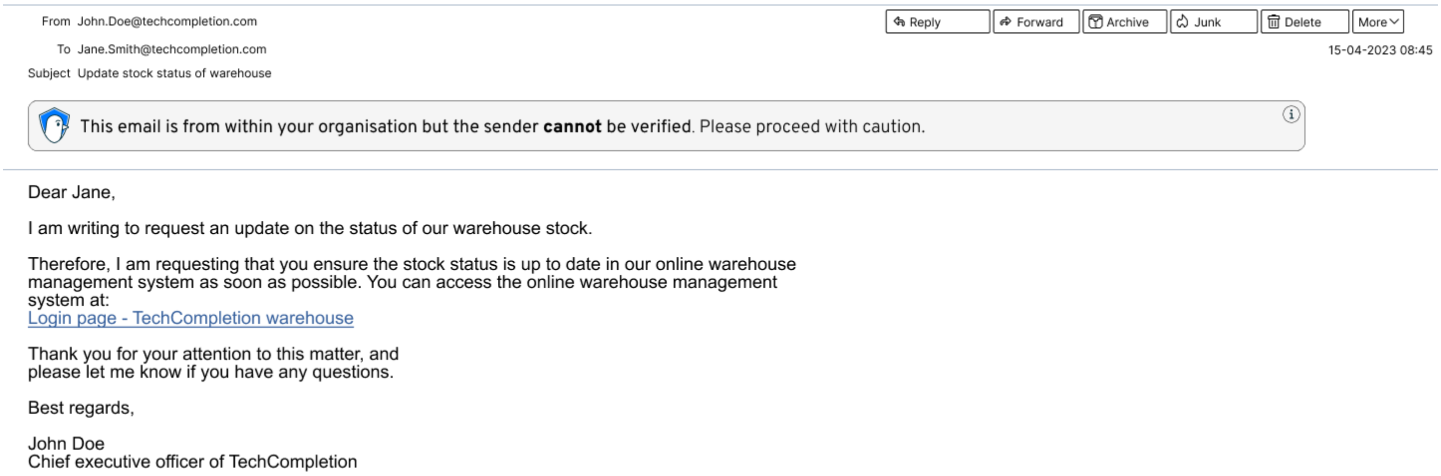
Therefore, I am requesting that you ensure the stock status is up to date in your online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and please let me know if you have any questions.

Best regards,

Tim Bloggs
Sales manager of FullPrecision

Figure A.4: Condition D of scenario 1 - external sender - signed (irrelevant).



From John.Doe@techcompletion.com

To Jane.Smith@techcompletion.com

15-04-2023 08:45

Subject Update stock status of warehouse

This email is from within your organisation but the sender **cannot** be verified. Please proceed with caution.

Dear Jane,

I am writing to request an update on the status of our warehouse stock.

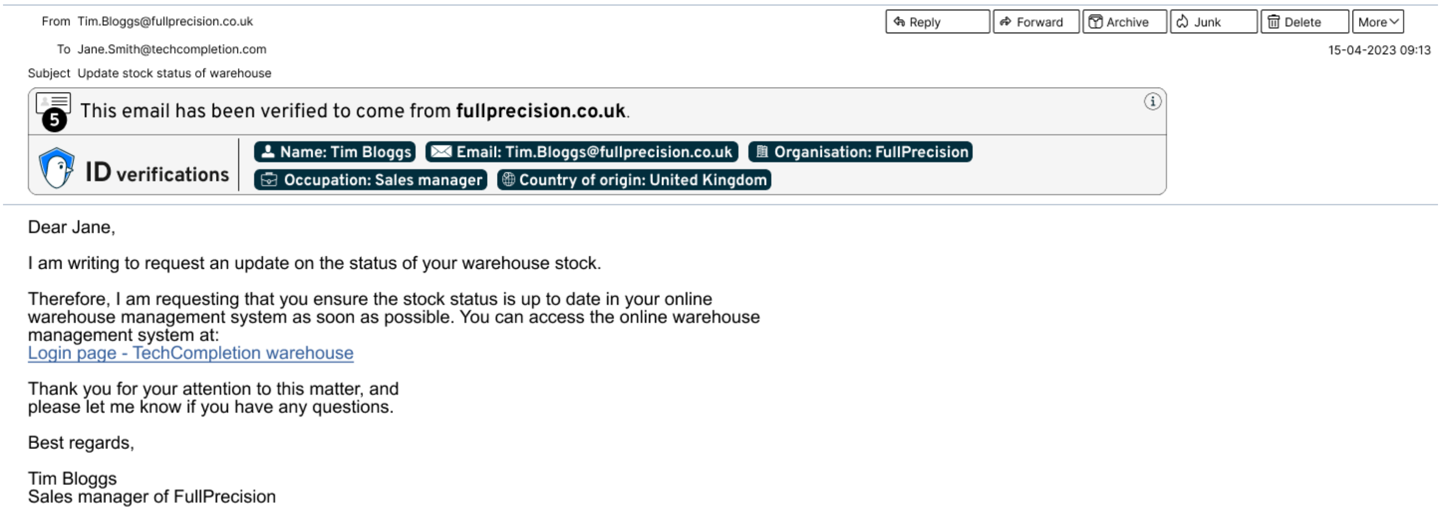
Therefore, I am requesting that you ensure the stock status is up to date in our online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and please let me know if you have any questions.

Best regards,

John Doe
Chief executive officer of TechCompletion

Figure A.5: Condition E of scenario 1 - internal sender - not signed.



From Tim.Bloggs@fullprecision.co.uk

To Jane.Smith@techcompletion.com

15-04-2023 09:13

Subject Update stock status of warehouse

This email has been verified to come from **fullprecision.co.uk**.

ID verifications

Name: Tim Bloggs | Email: Tim.Bloggs@fullprecision.co.uk | Organisation: FullPrecision | Occupation: Sales manager | Country of origin: United Kingdom

Dear Jane,

I am writing to request an update on the status of your warehouse stock.

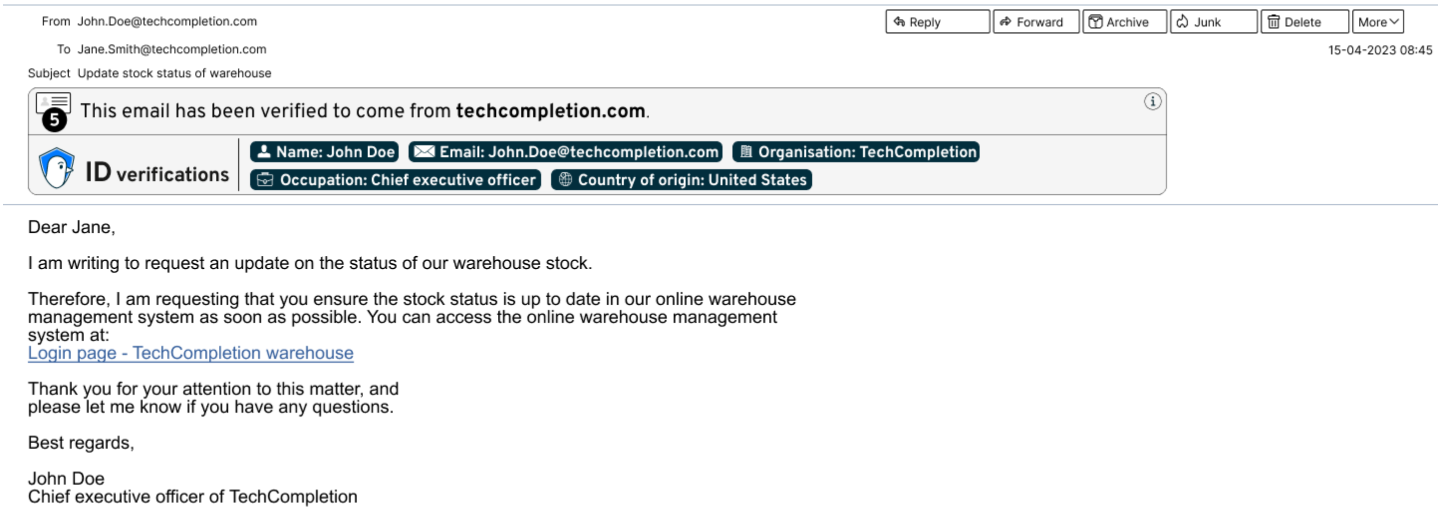
Therefore, I am requesting that you ensure the stock status is up to date in your online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and please let me know if you have any questions.

Best regards,

Tim Bloggs
Sales manager of FullPrecision

Figure A.6: Condition F of scenario 1 - external sender - signed (relevant).



From John.Doe@techcompletion.com

To Jane.Smith@techcompletion.com

15-04-2023 08:45

Subject Update stock status of warehouse

This email has been verified to come from **techcompletion.com**.

ID verifications

Name: John Doe | Email: John.Doe@techcompletion.com | Organisation: TechCompletion | Occupation: Chief executive officer | Country of origin: United States

Dear Jane,

I am writing to request an update on the status of our warehouse stock.

Therefore, I am requesting that you ensure the stock status is up to date in our online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

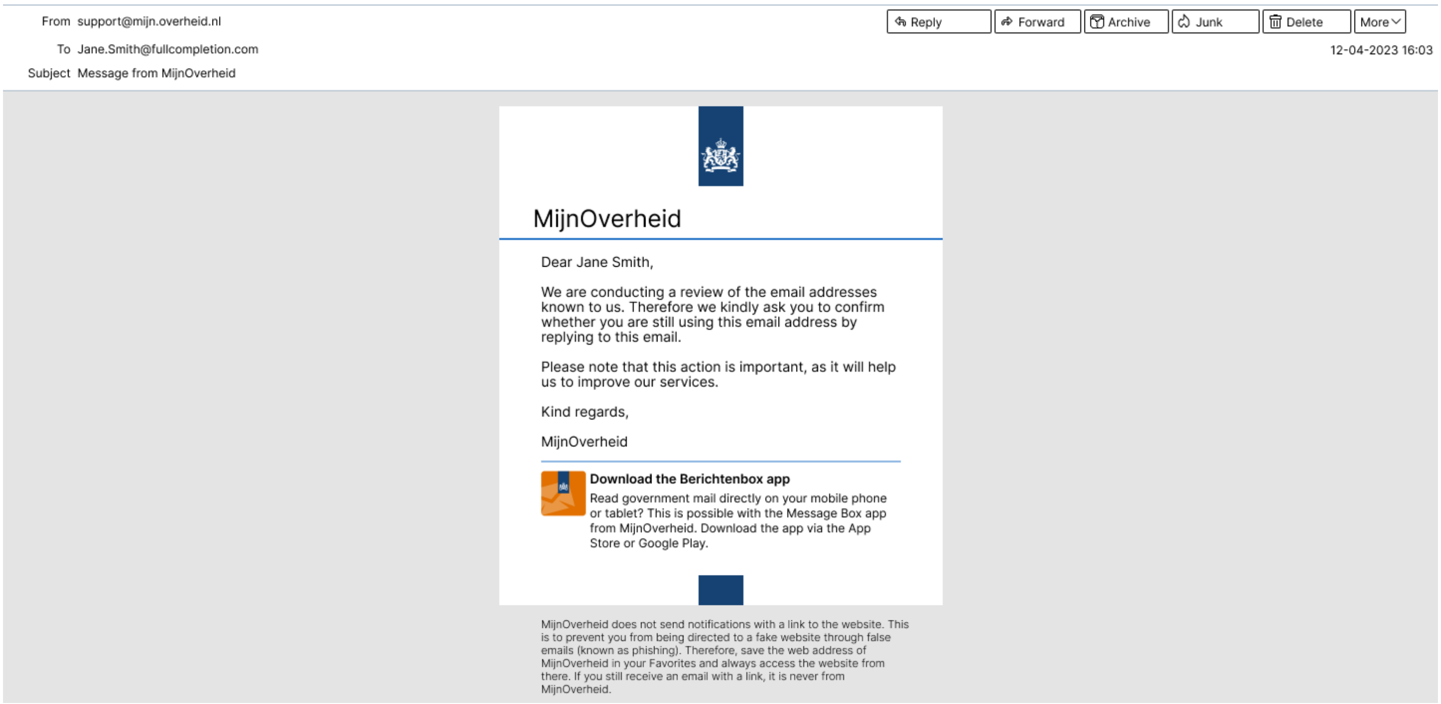
Thank you for your attention to this matter, and please let me know if you have any questions.

Best regards,

John Doe
Chief executive officer of TechCompletion

Figure A.7: Condition G of scenario 1 - internal sender - signed.

A.2 Scenario 2: MijnOverheid



From support@mijn.overheid.nl

To Jane.Smith@fullcompletion.com

Subject Message from MijnOverheid

12-04-2023 16:03

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

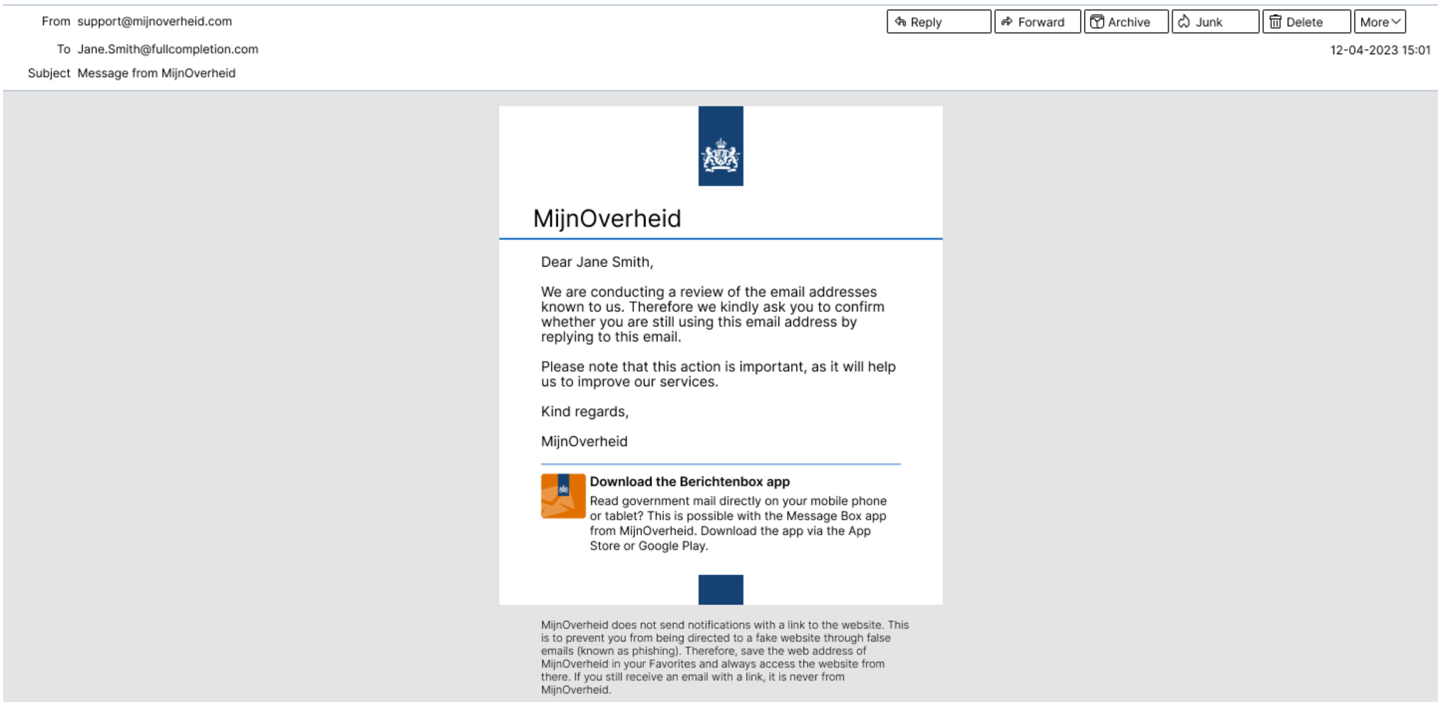
MijnOverheid

Download the Berichtenbox app

Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

Figure A.8: Condition I of scenario 2 - legitimate email - none



From support@mijnoverheid.com

To Jane.Smith@fullcompletion.com

Subject Message from MijnOverheid

12-04-2023 15:01

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

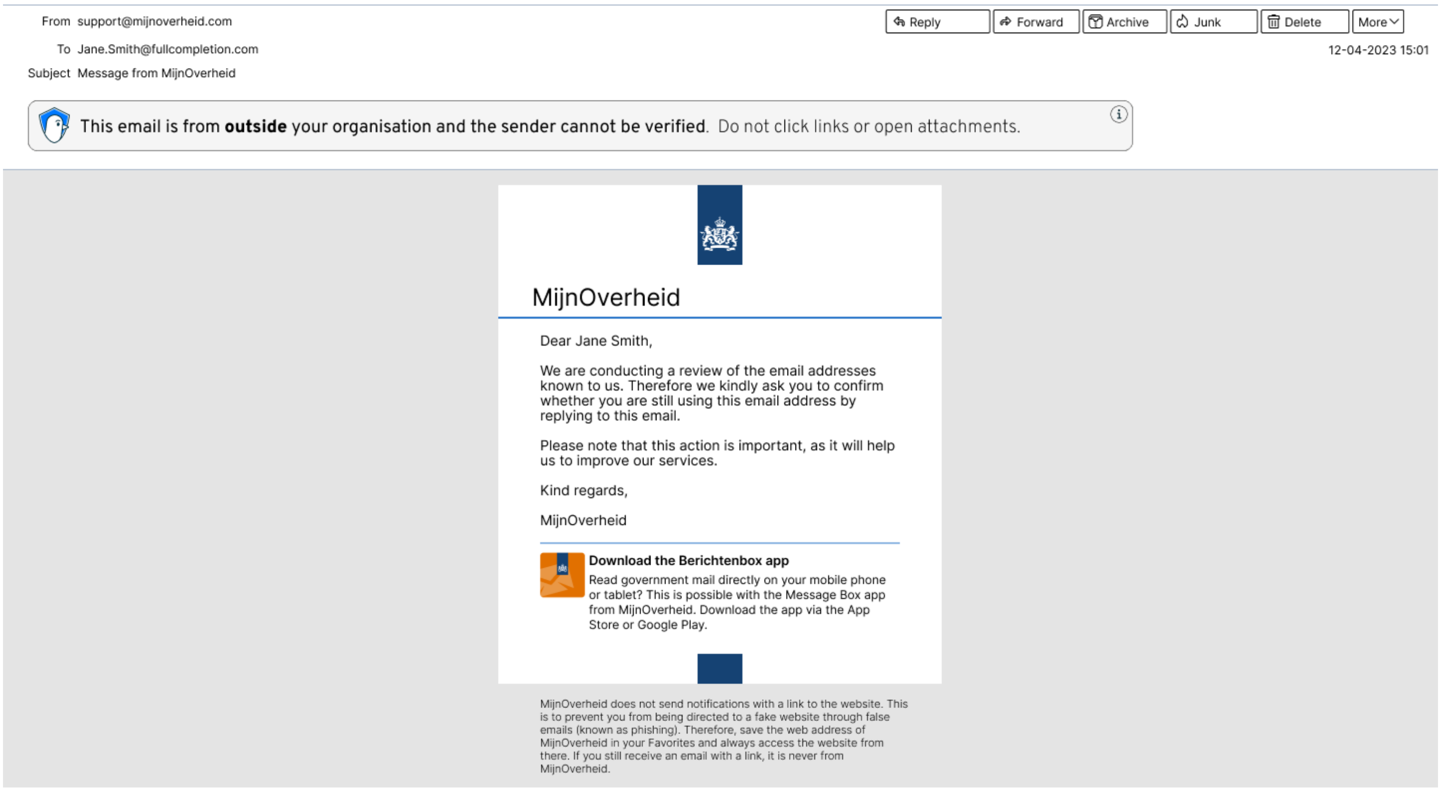
MijnOverheid

Download the Berichtenbox app

Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

Figure A.9: Condition II of scenario 2 - malicious email - none (no design)



From support@mijnoverheid.com

To Jane.Smith@fullcompletion.com

Subject Message from MijnOverheid

12-04-2023 15:01

This email is from **outside** your organisation and the sender cannot be verified. Do not click links or open attachments.

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

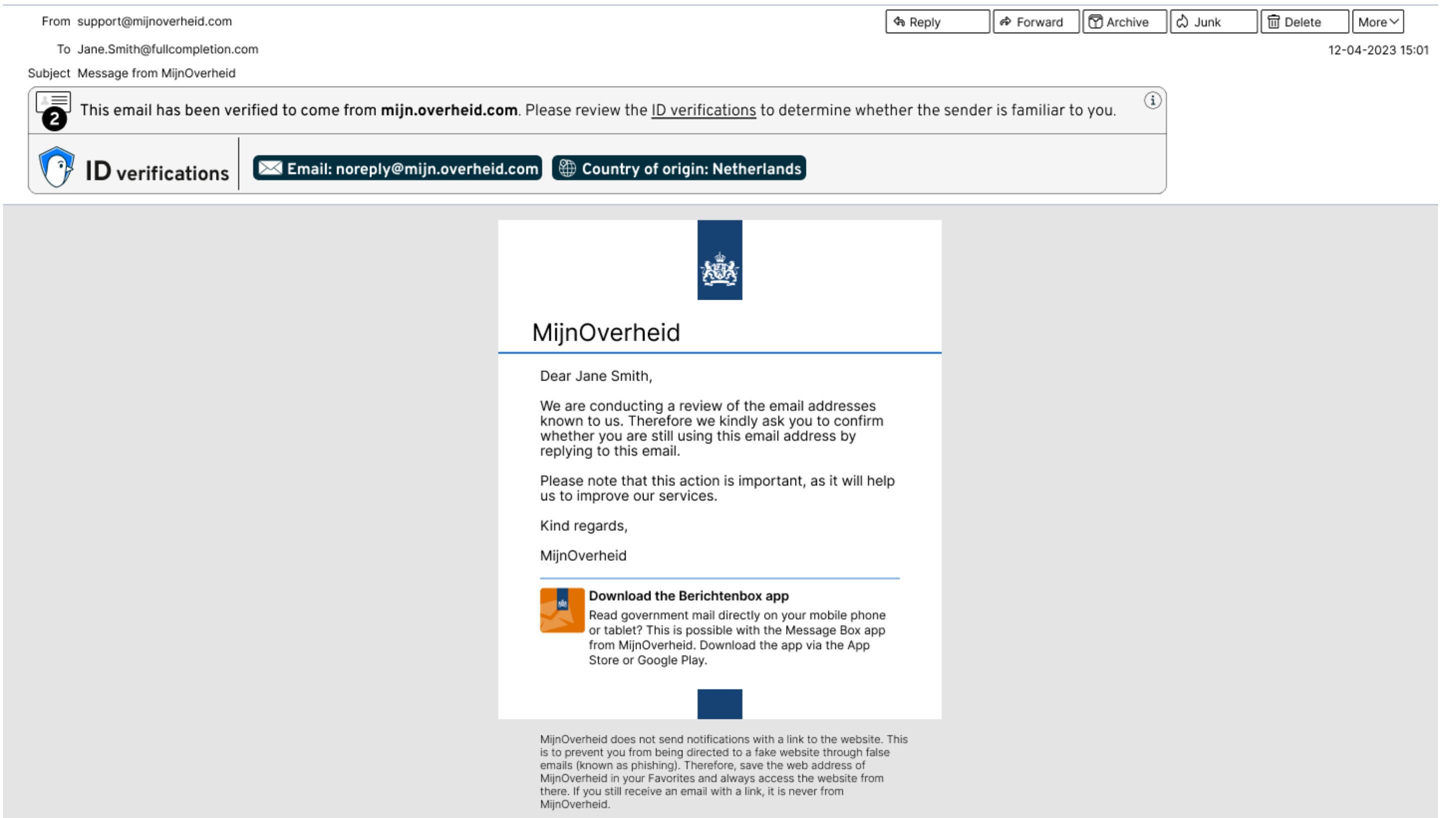
MijnOverheid

Download the Berichtenbox app

Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

Figure A.10: Condition III of scenario 2 malicious email - not signed



From support@mijnoverheid.com

To Jane.Smith@fullcompletion.com

Subject Message from MijnOverheid

12-04-2023 15:01

2 This email has been verified to come from **mijn.overheid.com**. Please review the ID verifications to determine whether the sender is familiar to you.

Email: noreply@mijn.overheid.com

Country of origin: Netherlands

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

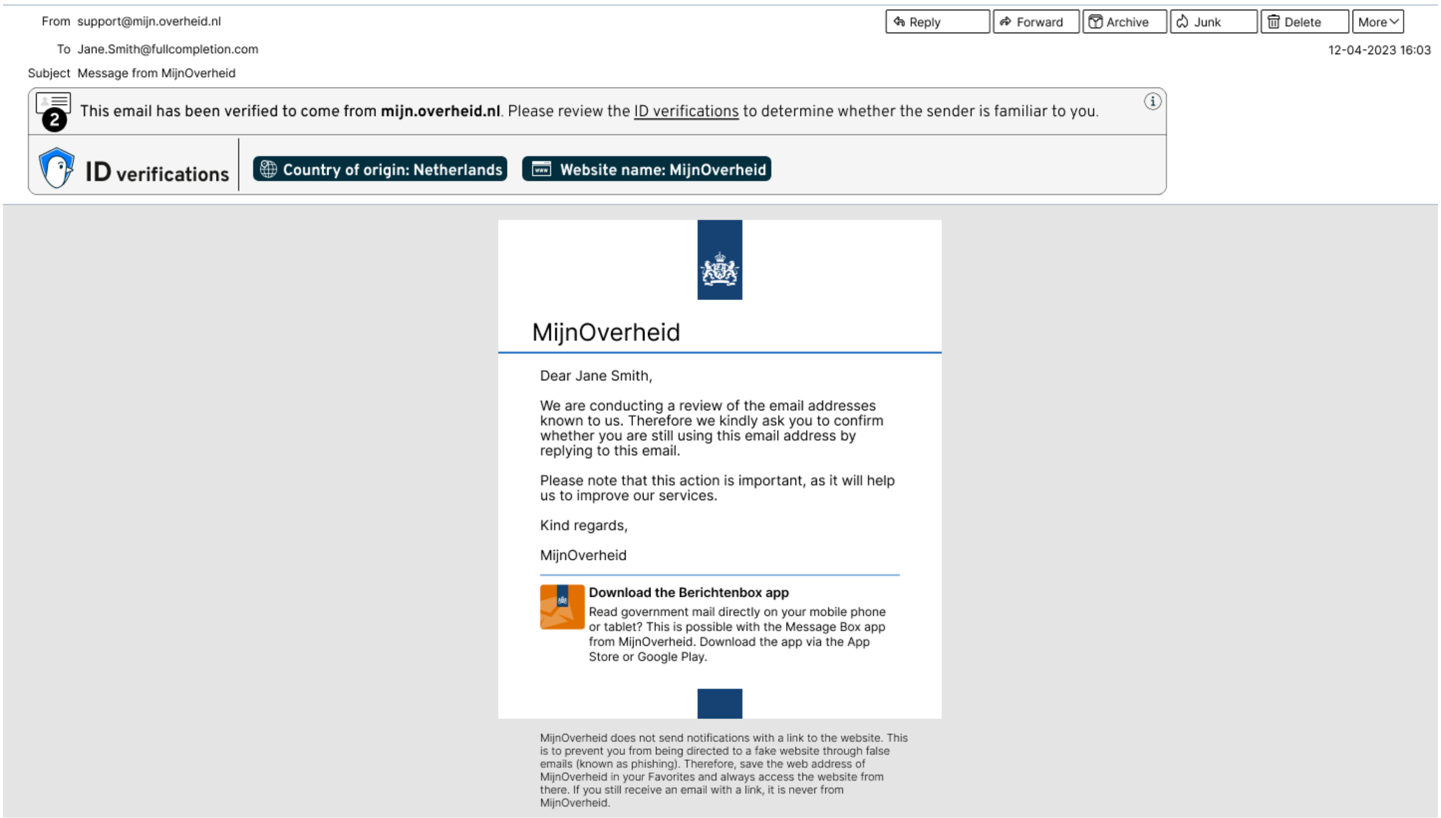
MijnOverheid

Download the Berichtenbox app

Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

Figure A.11: Condition IV of scenario 2 - malicious email - signed (irrelevant)



From support@mijn.overheid.nl

To Jane.Smith@fullcompletion.com

Subject Message from MijnOverheid

12-04-2023 16:03

This email has been verified to come from **mijn.overheid.nl**. Please review the ID verifications to determine whether the sender is familiar to you.

ID verifications **Country of origin: Netherlands** **Website name: MijnOverheid**

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

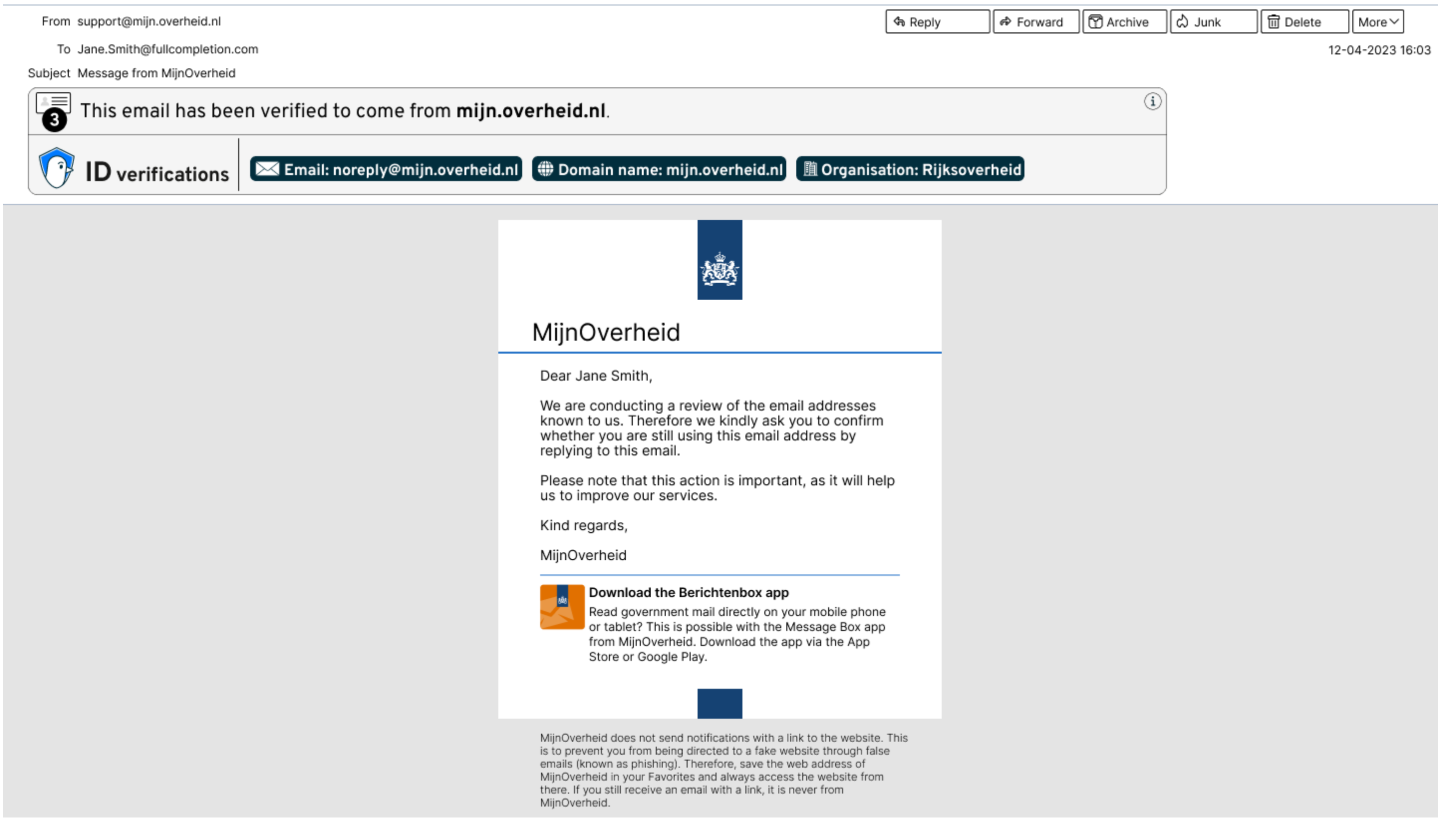
MijnOverheid

Download the Berichtenbox app

Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

Figure A.12: Condition V of scenario 2 - legitimate email - signed (irrelevant)



From support@mijn.overheid.nl

To Jane.Smith@fullcompletion.com

12-04-2023 16:03

Subject Message from MijnOverheid

This email has been verified to come from **mijn.overheid.nl**.

Email: noreply@mijn.overheid.nl **Domain name: mijn.overheid.nl** **Organisation: Rijksoverheid**

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

MijnOverheid

Download the Berichtenbox app

Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

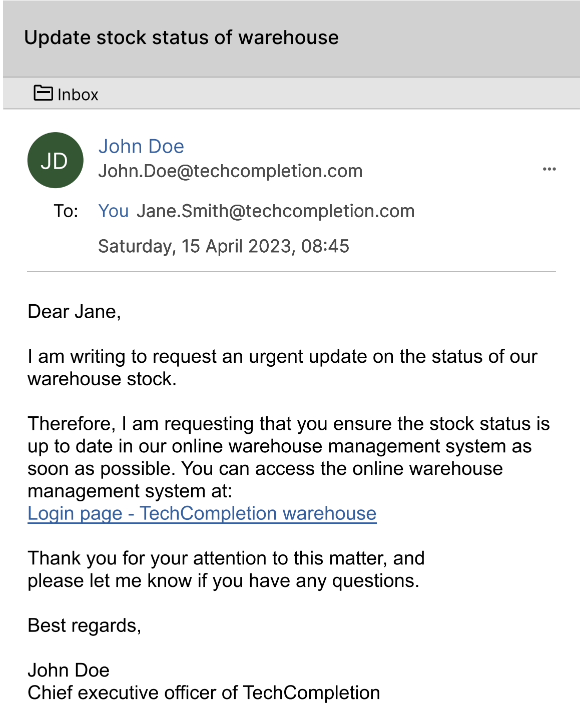
Figure A.13: Condition VI of scenario 2 - legitimate email - signed (relevant)

Appendix B

Mobile conditions

Below are the conditions containing the designs for mobile devices. These have not been tested in this thesis.

B.1 Scenario 1: TechCompletion



Update stock status of warehouse

John Doe
John.Doe@techcompletion.com

To: You Jane.Smith@techcompletion.com

Saturday, 15 April 2023, 08:45

Dear Jane,

I am writing to request an urgent update on the status of our warehouse stock.

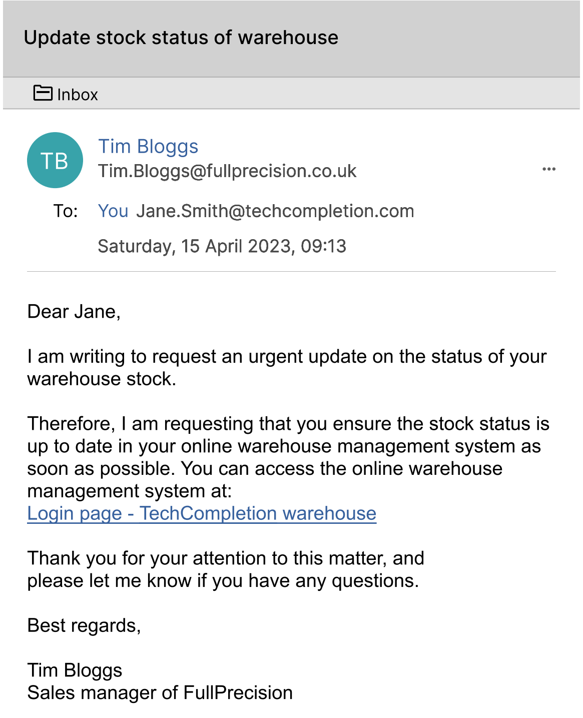
Therefore, I am requesting that you ensure the stock status is up to date in our online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and please let me know if you have any questions.

Best regards,

John Doe
Chief executive officer of TechCompletion

(a) Condition A - internal sender - none



Update stock status of warehouse

Tim Bloggs
Tim.Bloggs@fullprecision.co.uk

To: You Jane.Smith@techcompletion.com

Saturday, 15 April 2023, 09:13

Dear Jane,

I am writing to request an urgent update on the status of your warehouse stock.

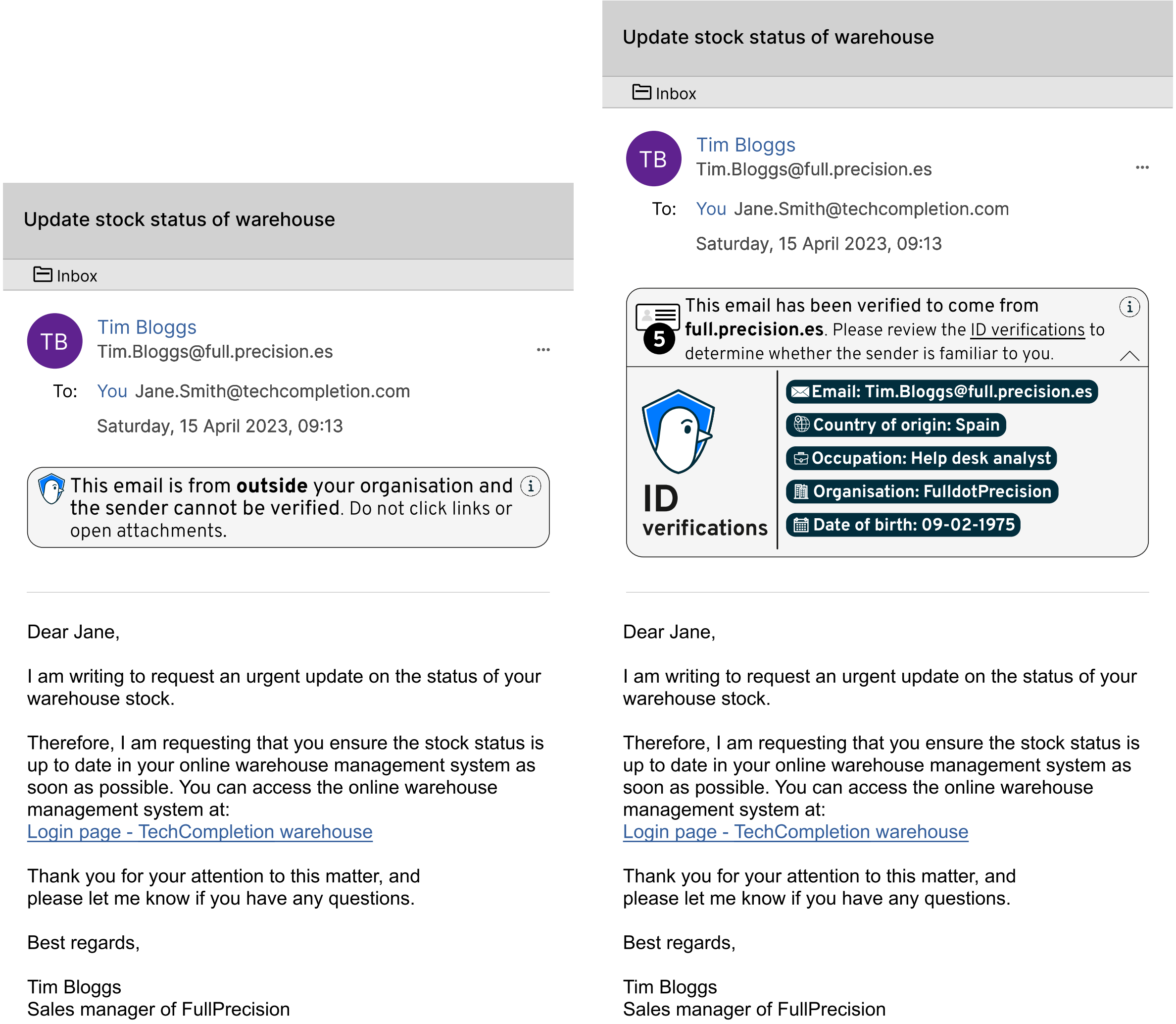
Therefore, I am requesting that you ensure the stock status is up to date in your online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and please let me know if you have any questions.

Best regards,

Tim Bloggs
Sales manager of FullPrecision

(b) Condition B - external sender - none



Update stock status of warehouse

Inbox

TB Tim Bloggs
Tim.Bloggs@full.precision.es

To: You Jane.Smith@techcompletion.com

Saturday, 15 April 2023, 09:13

This email is from **outside** your organisation and the sender cannot be verified. Do not click links or open attachments.

Dear Jane,

I am writing to request an urgent update on the status of your warehouse stock.

Therefore, I am requesting that you ensure the stock status is up to date in your online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and please let me know if you have any questions.

Best regards,

Tim Bloggs
Sales manager of FullPrecision

(a) Condition C - external sender - not signed

Update stock status of warehouse

Inbox

TB Tim Bloggs
Tim.Bloggs@full.precision.es

To: You Jane.Smith@techcompletion.com

Saturday, 15 April 2023, 09:13

5 This email has been verified to come from **full.precision.es**. Please review the ID verifications to determine whether the sender is familiar to you.

ID verifications

- **Email: Tim.Bloggs@full.precision.es**
- **Country of origin: Spain**
- **Occupation: Help desk analyst**
- **Organisation: FulldotPrecision**
- **Date of birth: 09-02-1975**

Dear Jane,

I am writing to request an urgent update on the status of your warehouse stock.

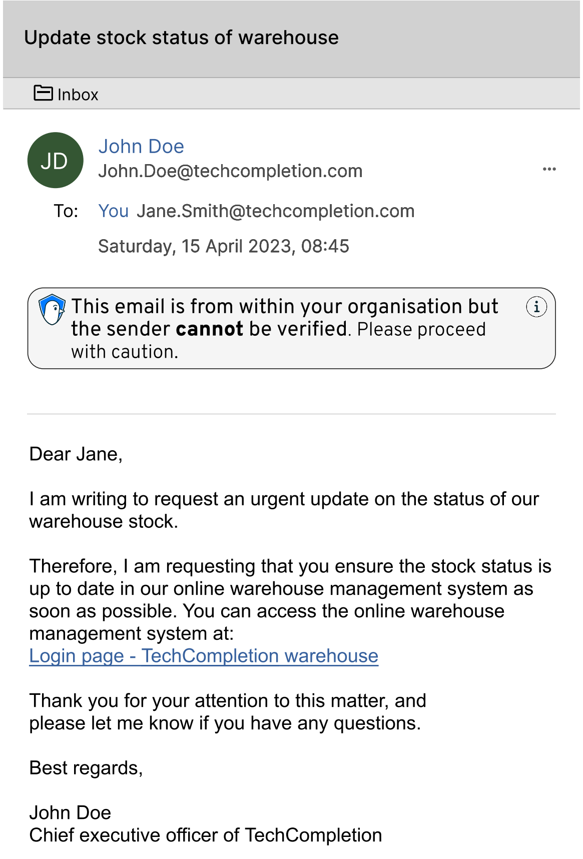
Therefore, I am requesting that you ensure the stock status is up to date in your online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and please let me know if you have any questions.

Best regards,

Tim Bloggs
Sales manager of FullPrecision

(b) Condition D - external sender - signed (irrelevant)



Update stock status of warehouse

Inbox

JD John Doe
John.Doe@techcompletion.com

To: You Jane.Smith@techcompletion.com

Saturday, 15 April 2023, 08:45

This email is from within your organisation but the sender **cannot** be verified. Please proceed with caution.

Dear Jane,

I am writing to request an urgent update on the status of our warehouse stock.

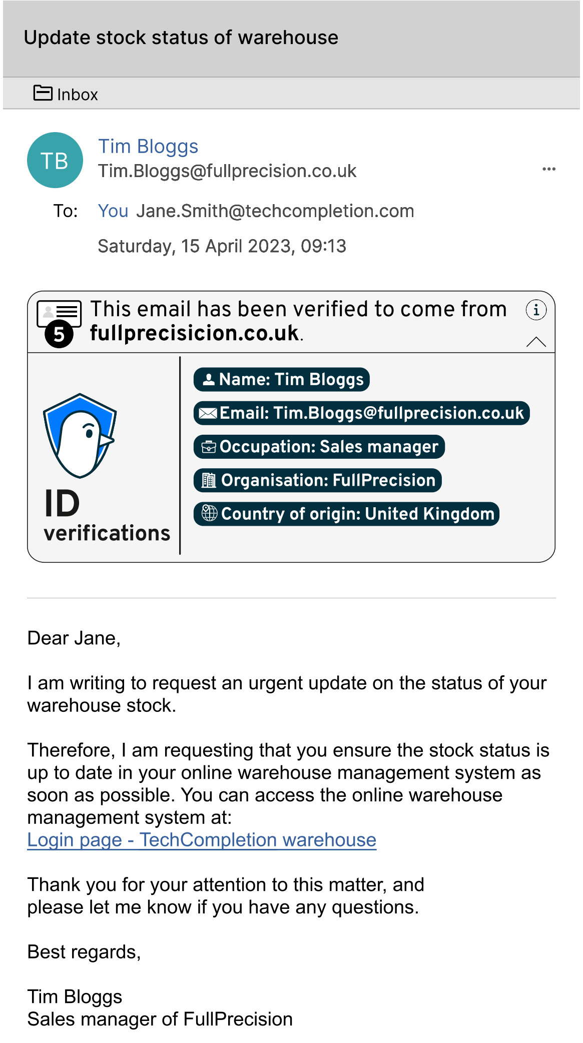
Therefore, I am requesting that you ensure the stock status is up to date in our online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and please let me know if you have any questions.

Best regards,

John Doe
Chief executive officer of TechCompletion

(a) Condition E - internal sender - not signed



Update stock status of warehouse

Inbox

TB Tim Bloggs
Tim.Bloggs@fullprecision.co.uk

To: You Jane.Smith@techcompletion.com

Saturday, 15 April 2023, 09:13

5 This email has been verified to come from **fullprecisicion.co.uk**.

ID verifications

- Name: Tim Bloggs
- Email: Tim.Bloggs@fullprecision.co.uk
- Occupation: Sales manager
- Organisation: FullPrecision
- Country of origin: United Kingdom

Dear Jane,

I am writing to request an urgent update on the status of your warehouse stock.

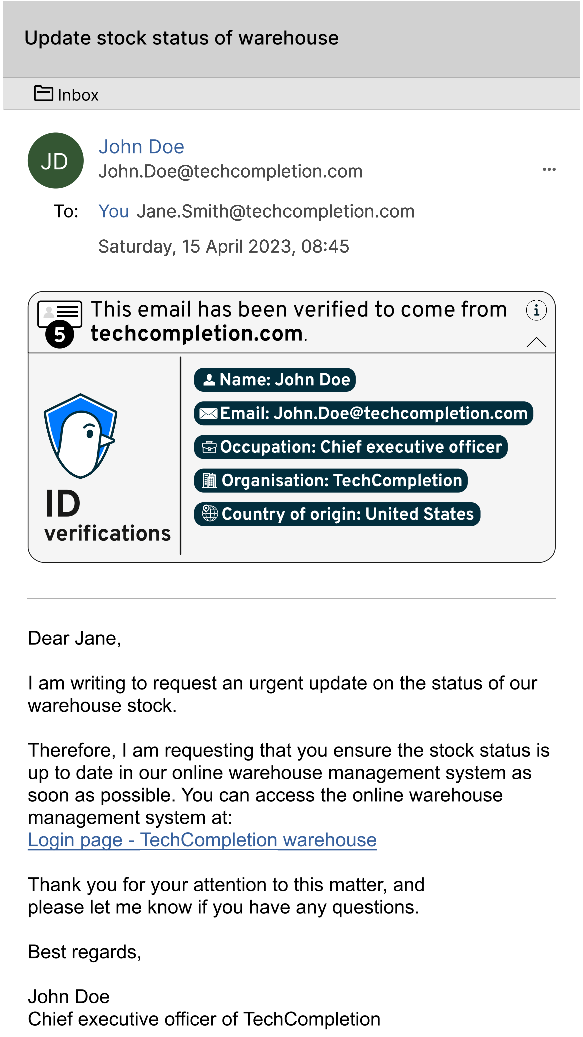
Therefore, I am requesting that you ensure the stock status is up to date in your online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and please let me know if you have any questions.

Best regards,

Tim Bloggs
Sales manager of FullPrecision

(b) Condition F - external sender - signed (relevant)


Update stock status of warehouse

John Doe
John.Doe@techcompletion.com

To: You Jane.Smith@techcompletion.com

Saturday, 15 April 2023, 08:45

This email has been verified to come from **techcompletion.com**.

- Name: John Doe
- Email: John.Doe@techcompletion.com
- Occupation: Chief executive officer
- Organisation: TechCompletion
- Country of origin: United States

Dear Jane,

I am writing to request an urgent update on the status of our warehouse stock.

Therefore, I am requesting that you ensure the stock status is up to date in our online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

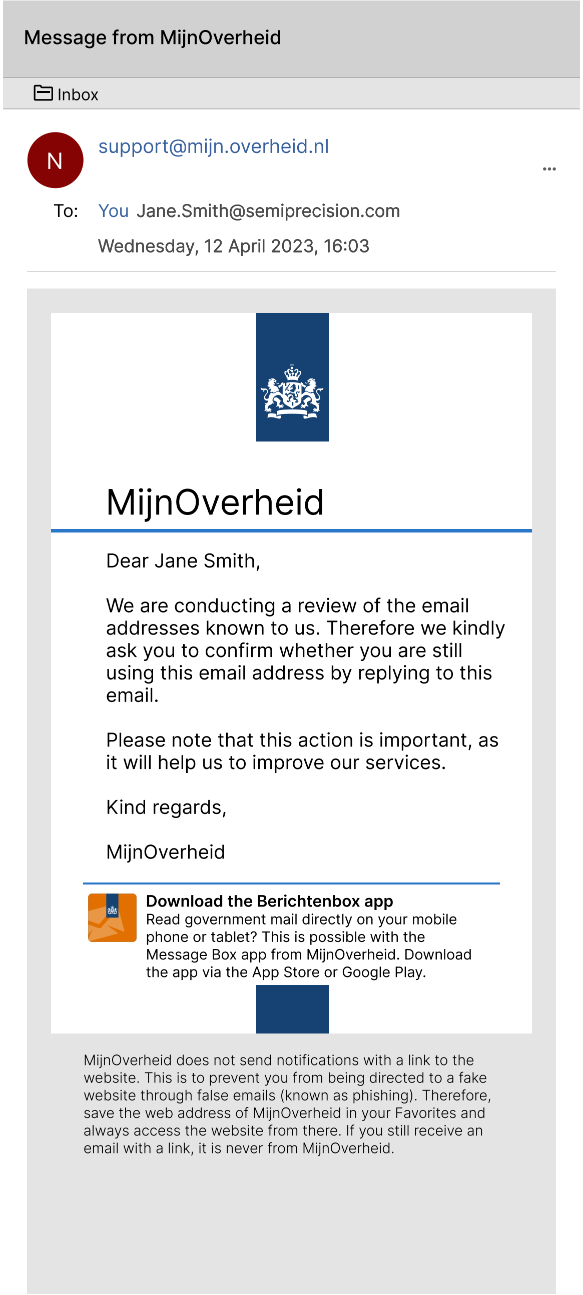
Thank you for your attention to this matter, and please let me know if you have any questions.

Best regards,

John Doe
Chief executive officer of TechCompletion

(a) Condition G - internal sender - signed

B.2 Scenario 2: MijnOverheid



Message from MijnOverheid

Inbox

support@mijn.overheid.nl

To: You Jane.Smith@semiprecision.com

Wednesday, 12 April 2023, 16:03

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

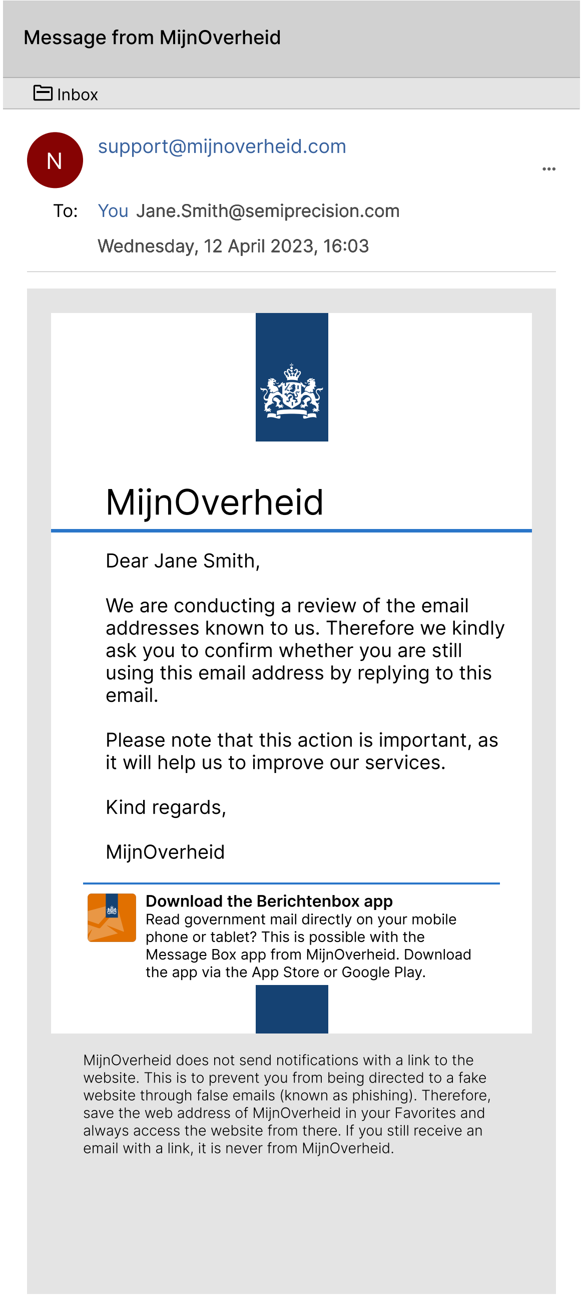
Kind regards,

MijnOverheid

Download the Berichtenbox app
Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

(a) Condition I - legitimate email - none



Message from MijnOverheid

Inbox

support@mijnoverheid.com

To: You Jane.Smith@semiprecision.com

Wednesday, 12 April 2023, 16:03

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

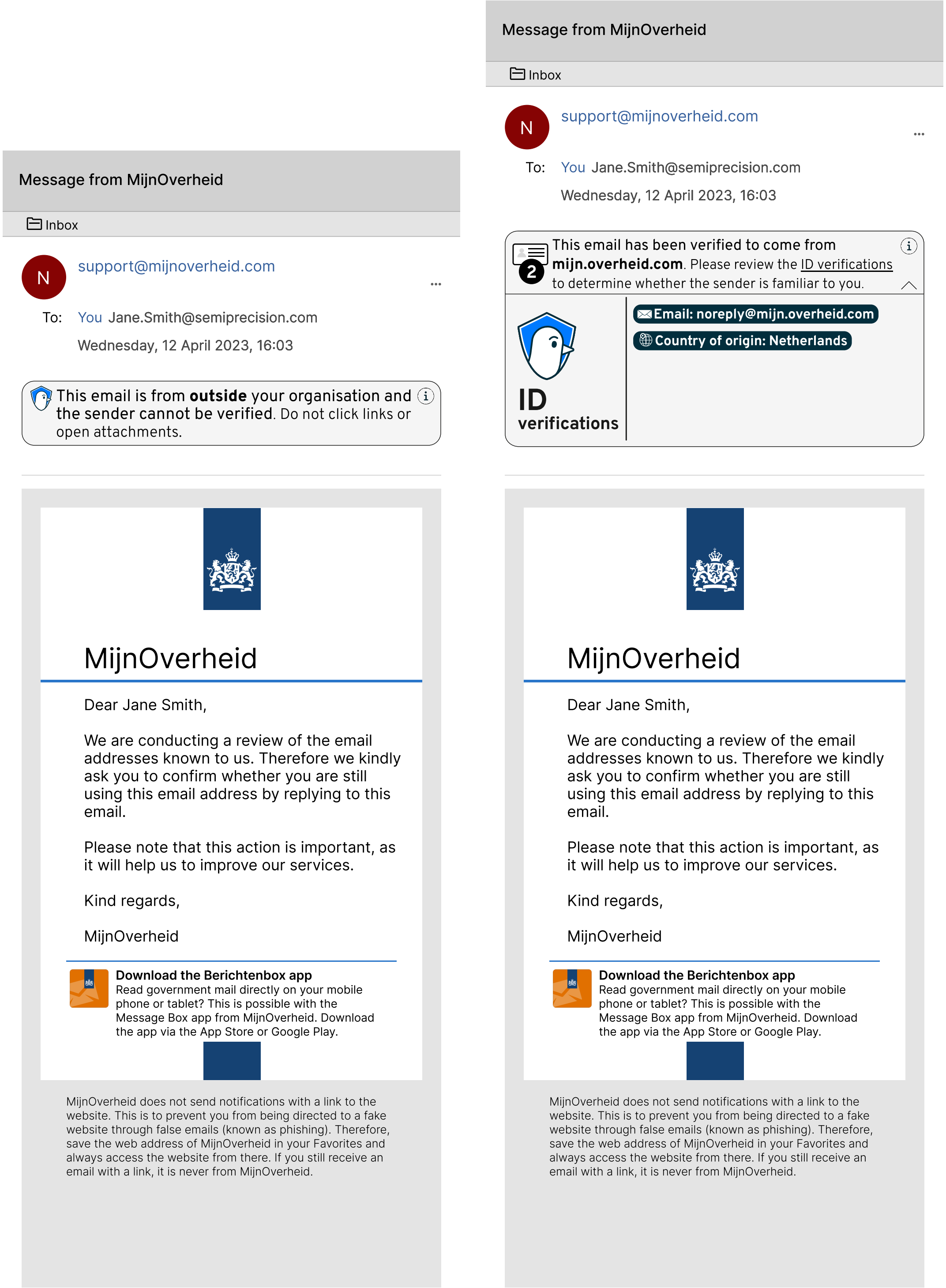
Kind regards,

MijnOverheid

Download the Berichtenbox app
Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

(b) Condition II - malicious email - none



Message from MijnOverheid

Inbox

support@mijnoverheid.com

To: You Jane.Smith@semiprecision.com

Wednesday, 12 April 2023, 16:03

This email is from **outside** your organisation and the sender cannot be verified. Do not click links or open attachments.

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

MijnOverheid

Download the Berichtenbox app
Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

(a) Condition III - malicious email - not signed

Message from MijnOverheid

Inbox

support@mijnoverheid.com

To: You Jane.Smith@semiprecision.com

Wednesday, 12 April 2023, 16:03

This email has been verified to come from **mijn.overheid.com**. Please review the ID verifications to determine whether the sender is familiar to you.

ID verifications

Email: noreply@mijn.overheid.com

Country of origin: Netherlands

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

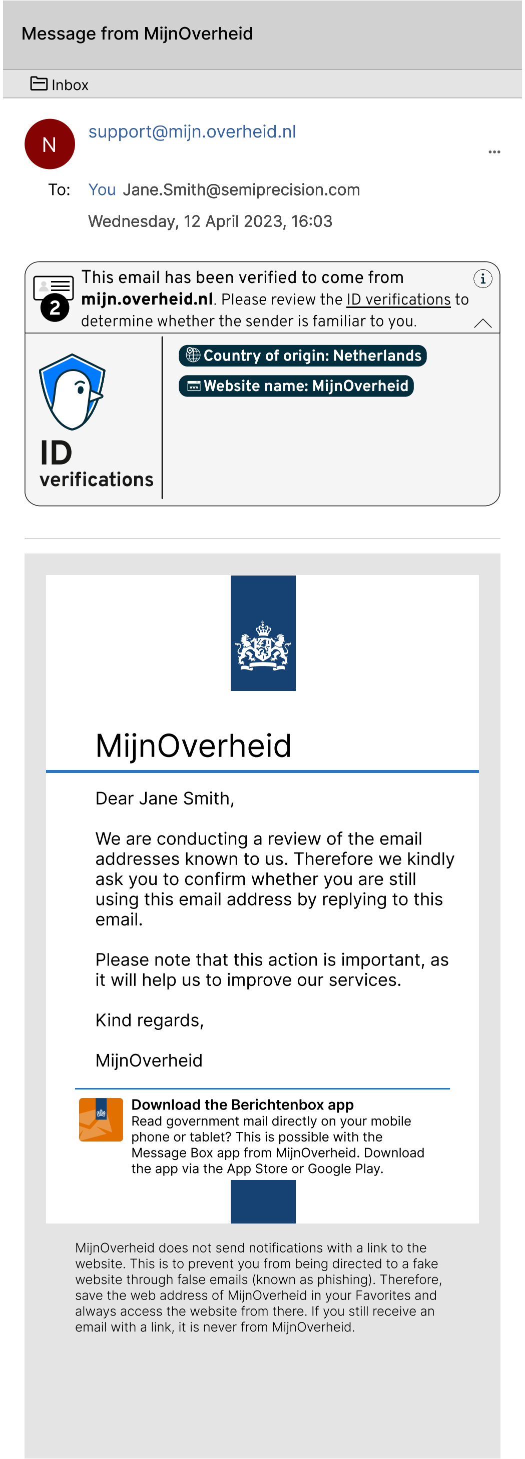
Kind regards,

MijnOverheid

Download the Berichtenbox app
Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

(b) Condition IV - malicious email - signed (irrelevant)



Message from MijnOverheid

support@mijn.overheid.nl

To: You Jane.Smith@semiprecision.com

Wednesday, 12 April 2023, 16:03

This email has been verified to come from **mijn.overheid.nl**. Please review the ID verifications to determine whether the sender is familiar to you.

Country of origin: Netherlands

Website name: MijnOverheid

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

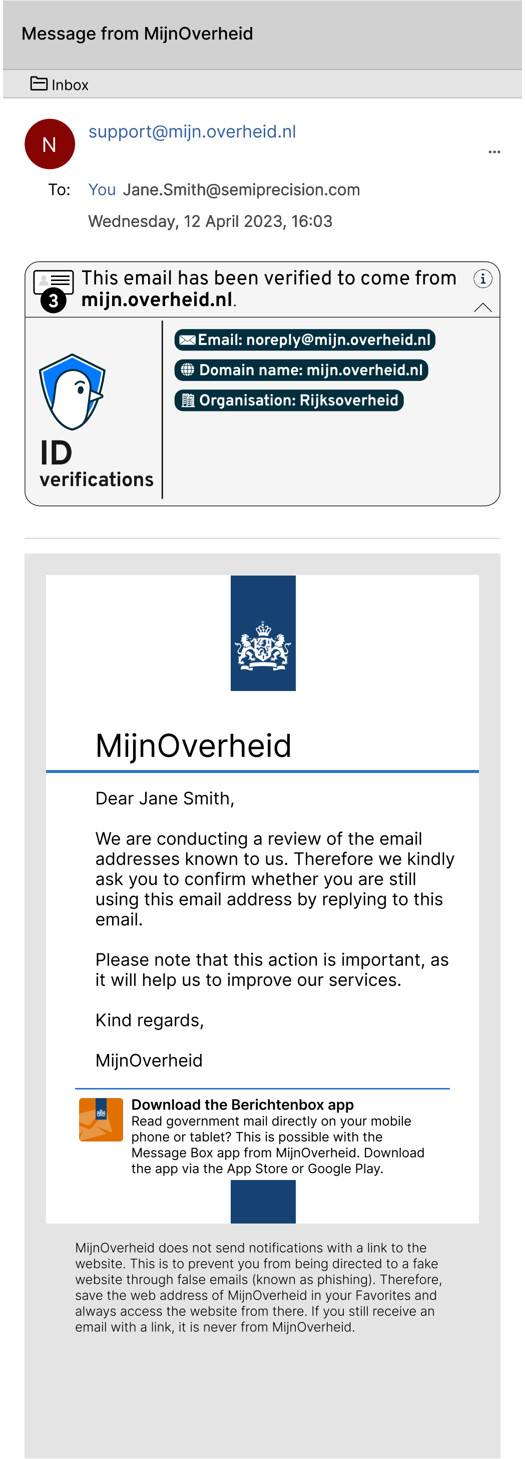
Kind regards,

MijnOverheid

Download the Berichtenbox app
Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

(a) Condition V - legitimate email - signed (irrelevant)



Message from MijnOverheid

support@mijn.overheid.nl

To: You Jane.Smith@semiprecision.com

Wednesday, 12 April 2023, 16:03

This email has been verified to come from **mijn.overheid.nl**.

Email: noreply@mijn.overheid.nl

Domain name: mijn.overheid.nl

Organisation: Rijksoverheid

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

MijnOverheid

Download the Berichtenbox app
Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

(b) Condition VI - legitimate email - signed (relevant)

Appendix C

Demographics

Here are the complete descriptive statistics regarding the demographic background of our participants.

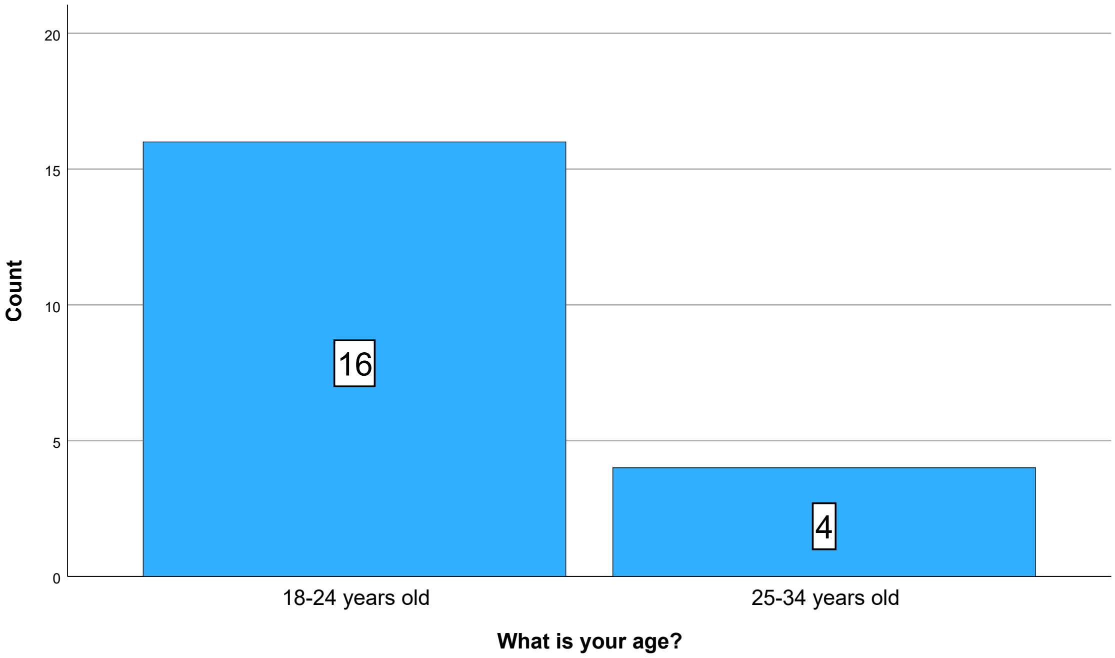


Figure C.1: Distribution of age groups (N = 20)

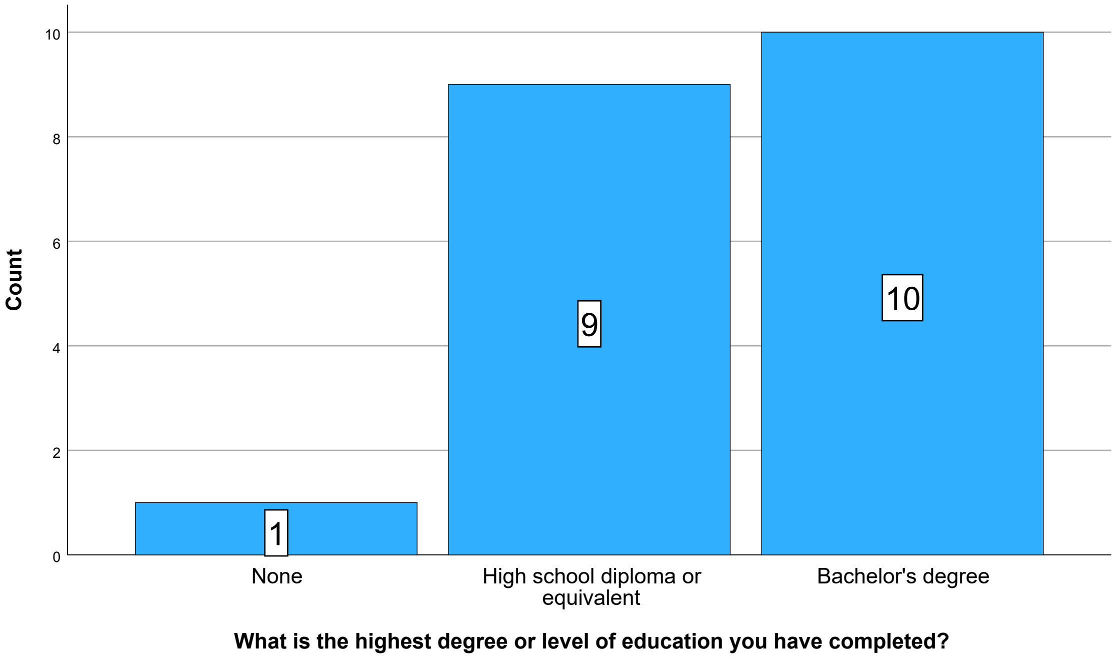


Figure C.2: Distribution of education (N = 20)

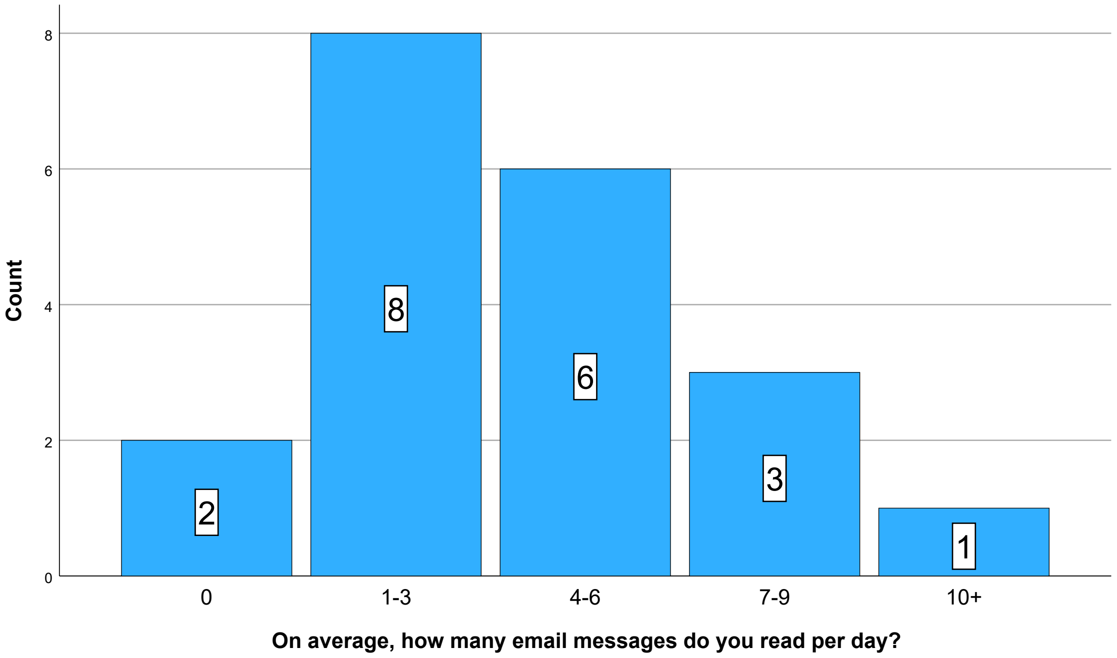


Figure C.3: Distribution of the average emails received per day (N = 20)

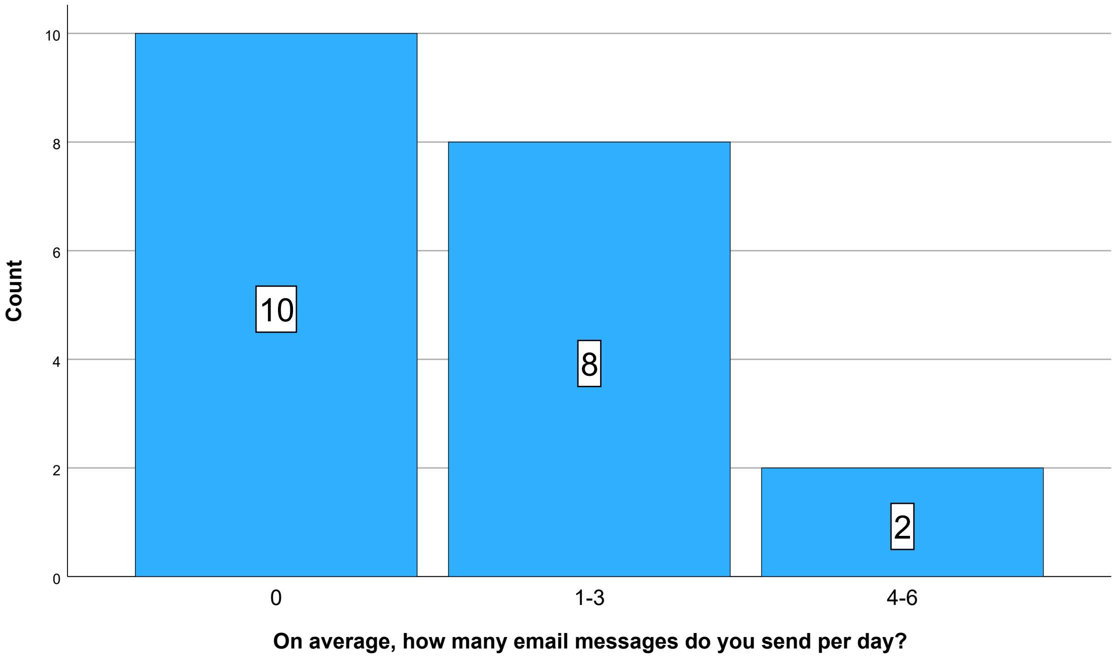


Figure C.4: Distribution of the average emails sent per day (N = 20)

Appendix D

Thematic analyses

D.1 Scenario 1: Email interaction comfort

Sender identity verification and confirmation

- When the design showed a warning and mismatching ID verifications, suspicion was raised. When the design was absent, emails were perceived as less credible. Participants mentioned in both cases their reluctance to click links or open attachments.
- The presence of matching ID verifications and the corresponding identity-based digital signature design raise trust and comfort in email interaction.
- When the sender's occupation, organisation name or domain name did not match their genuine counterparts, doubts about the authenticity of the email were raised.

Sender characteristics

- Participants expressed comfort when the email originated from within the organisation and when the ID verifications corresponded to the sender's identity.
- Participants found emails from known colleagues, when the request was reasonable and the email address seemed legitimate, to be trustworthy.
- When the sender's ID verifications did not match such as a mismatching occupation, concerns were raised about the legitimacy of the email.

Suspicious email content

- Participants raised concerns about emails that contained direct links to login pages, especially when assumed that the recipient already knew where to log in.

- The request in the email is viewed as suspicious because according to some participants it normally would not be requested by the CEO. This also holds when it was asked by an external party.
- The full link address could not be seen, which decreased participants' comfort in clicking the link.

Sender trust

- Emails that were expected or came from a known sender are more trustworthy.
- Emails that are well-written from a legitimate email address with a reasonable request are considered trustworthy.

D.2 Scenario 1: Email sender certainty

Influence of identity-based digital signature design

- Doubts were raised since the message on the banner of the identity-based digital signature did not specify that the email came from that organisation or person. It only indicated that the email came from the mentioned domain.
- When the design was absent, it raised doubts about the email originating from the sender. When it was present with relevant ID verifications, it increased participants' trust in the sender and email.

Sender characteristics

- Emails that originated from within the organisation were seen as more credible and trustworthy.
- Emails that came from an external sender raised concerns and doubts about their authenticity.
- When the email domain did not match with the organisation's domain, it raised doubts about the legitimacy of the email and sender.

Email content and context

- Suspicion was raised by the participants, when the sender's occupation in ID verifications did not match the content of the email.
- The request in the email was viewed as suspicious by some participants because it normally would not be requested by the CEO.
- The email included a link to the login page which raised questions about the authenticity of the email.

Sender familiarity and trust

- Participants expressed higher trust when the email came from a familiar sender with relevant ID verifications.
- Recognition of the domain or email address lead to increased trust in the email.

Need for external confirmation

- Some participants mentioned to seek external confirmation of the sender's identity, by contacting the sender directly via other communication media or contacting the company.

D.3 Scenario 2: Email interaction comfort

Verification and legitimacy

- Concerns were expressed about the legitimacy of the email when the identity-based digital signature was absent, or that the design indicated that the sender was not verified.
- Mismatching ID verifications undermined the trust of the email.
- When the email was verified, participants indicated that it was from a genuine sender.

Suspicious origin and content

- Some participants mentioned that the email was from outside their organisation, and in combination with the sender not being verified, the email was viewed as untrustworthy.
- MijnOverheid typically does not ask for replies or clicking on links and does not write emails in English. Therefore, the email looked suspicious.
- Participants view the email as untrustworthy if the sender's email address does not match the genuine email address of MijnOverheid.
- Doubts were still expressed after confirming the sender to be verified, since certain aspects of the email were found suspicious.

Email consequences

- Participants believed that no harm can be done by replying to the email, which made interaction less uncomfortable.
- Some participants were skeptic about the importance of the requested action, decreasing email interaction.
- A lack of interest or motivation was expressed by participants, because they had no interest in responding to the email to improve the service.

D.4 Scenario 2: Email sender certainty

Verification concerns

- Absence of ID verifications or the identity-based digital signature decreased trust and credibility. The same applies for when the design communicated that the sender was not verified. Participants expressed that they were uncertain about the origin of the email.
- When the identity-based digital signature communicated that the sender was verified and had matching ID verifications, the email was perceived as more credible and authentic.

Suspicious email domain and content

- When the email domain did not match the genuine domain, it was perceived to be certainly not from the mentioned sender.
- The content of the email, especially the request and language, was perceived as very suspicious by the participants that were familiar with MijnOverheid. Resulting in the perception that it certainly not came from MijnOverheid.
- Some participants were very uncertain about the authenticity of the email. This was caused by the email content and ID verifications.

Appendix E

Questionnaire

Here you can see the entire questionnaire as exported from Lime survey. Visually the exported questionnaire looks a bit off due to the margins.

User experience and security research of email digital signatures

There are 121 questions in this survey.

Welcome!

INFORMATION ABOUT THE EXPERIMENT

User experience and security research of visual digital signature designs in emails

Attention: *This survey has to be filled in using a* ***laptop*** *or* ***computer****. Devices with* ***small*** *screens are* ***not*** *supported.*

Introduction

This research is part of the Bachelor's Thesis of the Computing Science program at Radboud University. If you want to participate, we will ask you to sign a consent form. Before you decide whether or not to take part, we will give you information about the study. Please take time to read the following information carefully. If something is not clear, or you would like more information, please ask me via *leon.zhang@ru.nl*.

Outline and aim of the research

In this research we want to *measure how users perceive the credibility of an email message when it is presented with information about the identity of the email's sender.*

What is expected of you?

In this user experience and security research you will first have to answer a few questions about your background and email activity. After that, you will get to see some screenshots of email messages where you will be asked to answer a few questions about it. Participation in this research will take about 15 to 20 minutes.

Voluntary participation

Your participation in this research is voluntary. This means that you can withdraw your participation and consent at any time during the research, without giving a reason.

What data is collected?

We will collect data on your personal background: age and education, as well as the answers to the survey. Additionally, the timestamps of when a question is answered will be collected.

What will happen to my data?

The research data we collect during this study will be used for my bachelor's thesis. The anonymized research data might be accessible publicly as part of my thesis. The data might be used in a presentation where outcomes of my research is described. Personal data collected remain confidential. When we share data with others, these data cannot be traced back to you.

All research and personal data are safely stored following the Radboud University guidelines.

More information?

If you have any questions or would like more information about the research, please contact me using the contact information at the bottom of this letter.

Should you have any complaints regarding this research, please contact me.
You can also file a complaint with my supervisor: hanna.schraffenberger@ru.nl.

Consent form

If you want to participate in this research, we ask you to sign the consent form. With this written consent, you declare that you have understood the information we have provided and consent to participate in this research.

Please note the following:
The survey has to be completed in one sitting. You cannot save your progress and resume later.

Kind regards,

Leon Zhang (leon.zhang@ru.nl)

CONSENT FORM

for participation in the scientific study: U*ser experience and security research of visual digital signature designs in emails.*

Statement of participant

The aim of the research has been outlined to me.

- I am at least 18 years of age.
- I was given the opportunity to ask questions regarding the research study.
- I participate voluntarily in the research study.
- I understand that I can stop at any point during the research study, should I wish to do so.
- I understand how the data of the research study will be stored and how they will be used.
- I consent to participating in the research study as described in the information letter.

*

❗ Choose one of the following answers
Please choose **only one** of the following:

○ I agree to participate.

○ I do not want to participate.

Demographical background

What is your age? *

Please choose **only one** of the following:

○ 18-24 years old

○ 25-34 years old

○ 35-44 years old

○ 45-54 years old

○ Over 55

What is the highest degree or level of education you have completed? *

Please choose **only one** of the following:

○ None

○ High school diploma or equivalent

○ Bachelor's degree

○ Master's degree

○ Doctorate (e.g. PhD, EdD)

○ Other

Email Literacy/Activity

Email activity

You will now have to answer some questions about your email activity.

On average, how many email messages do you **read** per **day?**

*

Please choose **only one** of the following:

○ 0

○ 1-3

○ 4-6

○ 7-9

○ 10+

On average, how many email messages do you **send** per **day**?

*

Please choose **only one** of the following:

○ 0

○ 1-3

○ 4-6

○ 7-9

○ 10+

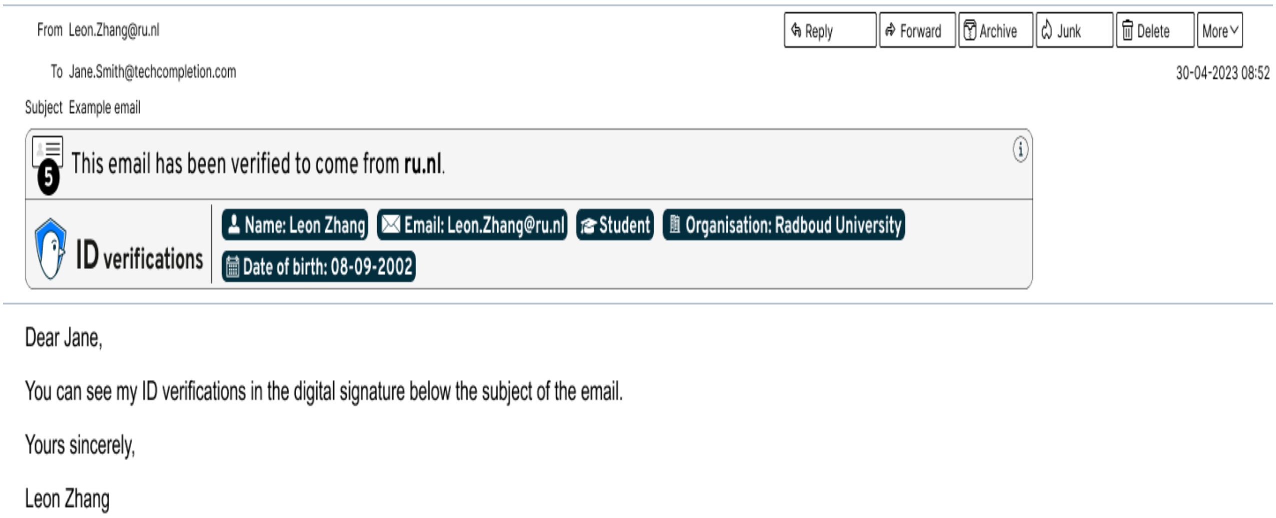
Scenario 1: TechCompletion - introduction

Welcome to the first scenario of this research.

Please read the information below carefully.
You are Jane Smith, a warehouse manager of the fictive company TechCompletion (`techcompletion.com`) located in the United States.

The chief executive officer (CEO) of TechCompletion, John Doe, made sure that all employees, including you, installed the email add-on **PostGuard**.
PostGuard allows users to securely communicate via email and provides information about the identity of an email's sender, in the form of ID verifications.
Examples of ID verifications are: full name, email address, date of birth, nationality, occupation and many more.

If you were to receive a PostGuard signed email from me, it would look like the image below.
It shows the ID verifications I have included, allowing you to confirm that the email was indeed sent by me. You can only include ID verifications that you have proven to possess. Therefore, fake ID verifications cannot be included.



From Leon.Zhang@ru.nl

To Jane.Smith@techcompletion.com

30-04-2023 08:52

Subject Example email

5 This email has been verified to come from **ru.nl**.

ID verifications

Name: Leon Zhang | Email: Leon.Zhang@ru.nl | Student | Organisation: Radboud University | Date of birth: 08-09-2002

Dear Jane,

You can see my ID verifications in the digital signature below the subject of the email.

Yours sincerely,

Leon Zhang

You will now be shown 7 emails you have received.

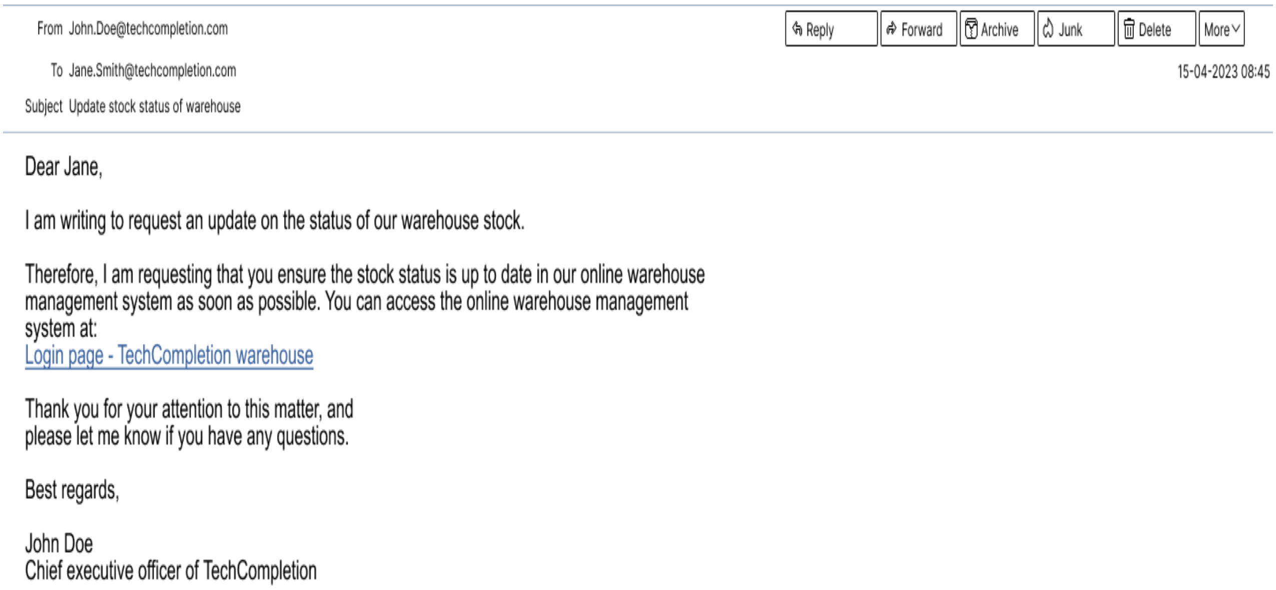
□

Scenario 1: TechCompletion - condition A

Brief description of scenario 1:

You are Jane Smith, a warehouse manager of TechCompletion (`techcompletion.com`) located in the United States.
You communicate with colleagues within the company and outside the company such as FullPrecision (`fullprecision.co.uk`).

Given this scenario, consider the following email:



From John.Doe@techcompletion.com

To Jane.Smith@techcompletion.com

Subject Update stock status of warehouse

15-04-2023 08:45

Dear Jane,

I am writing to request an update on the status of our warehouse stock.

Therefore, I am requesting that you ensure the stock status is up to date in our online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and
please let me know if you have any questions.

Best regards,

John Doe
Chief executive officer of TechCompletion

Would you click the link in the email and login? *

Please choose **only one** of the following:

○ Yes

○ No

How comfortable would you feel clicking the link in the email and login? *

Please choose the appropriate response for each item:

	Very uncomfortable						**Very comfortabl**
	○	○	○	○	○	○	○

And why? What aspects of the email message contributed to your response?

Please write your answer here:

How well do the following adjectives describe the **content of the email message** you just read?

*

Please choose the appropriate response for each item:

	Describes very poorly						Describes very well
Professional	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○

Please indicate for each statement how much it applies to you. I found the **sender of the email** ... *

Please choose the appropriate response for each item:

	Not at all						Extremely
Trustworthy	○	○	○	○	○	○	○
Credible	○	○	○	○	○	○	○
Reputable	○	○	○	○	○	○	○

Did this email come from your colleague **John Doe**?

*

Please choose the appropriate response for each item:

	No, certainly	I do not know	Yes, certainly
	○	○	○

And why? What aspects of the email message contributed to your response?

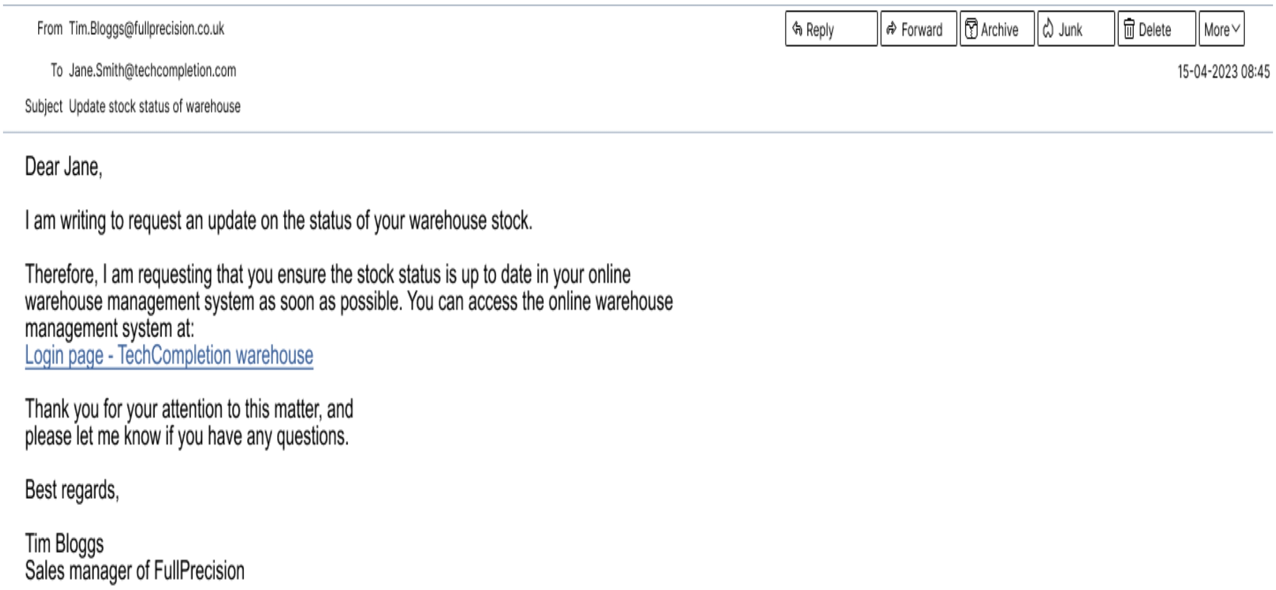
Please write your answer here:

Scenario 1: TechCompletion - condition B

Brief description of scenario 1:

You are Jane Smith, a warehouse manager of TechCompletion (`techcompletion.com`) located in the United States.
You communicate with colleagues within the company and outside the company such as FullPrecision (`fullprecision.co.uk`).

Given this scenario, consider the following email:



From Tim.Bloggs@fullprecision.co.uk

To Jane.Smith@techcompletion.com

Subject Update stock status of warehouse

15-04-2023 08:45

Dear Jane,

I am writing to request an update on the status of your warehouse stock.

Therefore, I am requesting that you ensure the stock status is up to date in your online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and
please let me know if you have any questions.

Best regards,

Tim Bloggs
Sales manager of FullPrecision

Would you click the link in the email and login? *

Please choose **only one** of the following:

○ Yes

○ No

How comfortable would you feel clicking the link in the email and login? *

Please choose the appropriate response for each item:

	Very uncomfortable						Very comfortabl
	○	○	○	○	○	○	○

And why? What aspects of the email message contributed to your response?

Please write your answer here:

How well do the following adjectives describe the **content of the email message** you just read?

*

Please choose the appropriate response for each item:

	Describes very poorly						**Describes very well**
Professional	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○

Please indicate for each statement how much it applies to you. I found the **sender of the email** ... *

Please choose the appropriate response for each item:

	Not at all						**Extremely**
Trustworthy	○	○	○	○	○	○	○
Credible	○	○	○	○	○	○	○
Reputable	○	○	○	○	○	○	○

Did this email come from your contact **Tim Bloggs**?

*

Please choose the appropriate response for each item:

	No, certainly	**I do not know**	**Yes, certainly**
	○	○	○

And why? What aspects of the email message contributed to your response?

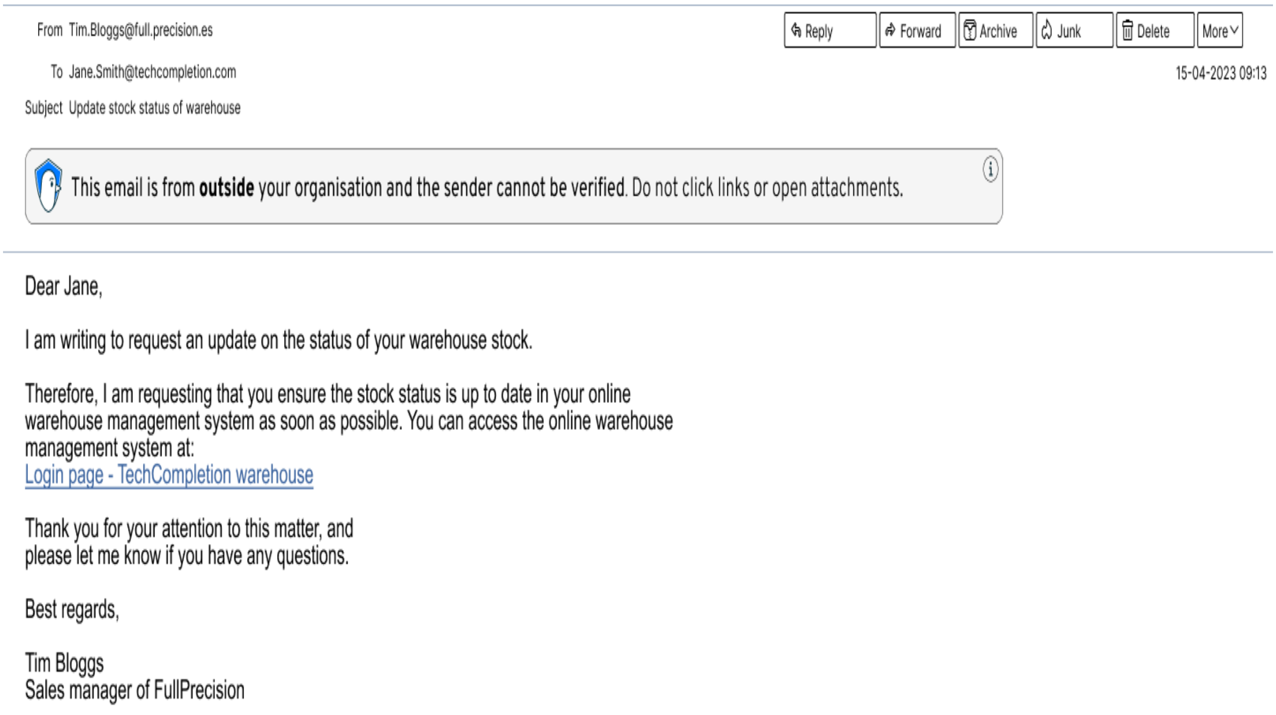
Please write your answer here:

Scenario 1: TechCompletion - condition C

Brief description of scenario 1:

You are Jane Smith, a warehouse manager of TechCompletion (`techcompletion.com`) located in the United States.
You communicate with colleagues within the company and outside the company such as FullPrecision (`fullprecision.co.uk`).

Given this scenario, consider the following email:



From Tim.Bloggs@full.precision.es

To Jane.Smith@techcompletion.com

15-04-2023 09:13

Subject Update stock status of warehouse

This email is from **outside** your organisation and the sender cannot be verified. Do not click links or open attachments.

Dear Jane,

I am writing to request an update on the status of your warehouse stock.

Therefore, I am requesting that you ensure the stock status is up to date in your online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and please let me know if you have any questions.

Best regards,

Tim Bloggs
Sales manager of FullPrecision

Would you click the link in the email and login? *

Please choose **only one** of the following:

○ Yes

○ No

How comfortable would you feel clicking the link in the email and login? *

Please choose the appropriate response for each item:

	Very uncomfortable						Very comfortabl
	○	○	○	○	○	○	○

And why? What aspects of the email message contributed to your response?

Please write your answer here:

How well do the following adjectives describe the **content of the email message** you just read?

*

Please choose the appropriate response for each item:

	Describes very poorly						Describes very well
Professional	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○

Please indicate for each statement how much it applies to you. I found the **sender of the email** ... *

Please choose the appropriate response for each item:

	Not at all						Extremely
Trustworthy	○	○	○	○	○	○	○
Credible	○	○	○	○	○	○	○
Reputable	○	○	○	○	○	○	○

Did this email come from your contact **Tim Bloggs**?

*

Please choose the appropriate response for each item:

	No, certainly	I do know know	Yes, certainly
	○	○	○

And why? What aspects of the email message contributed to your response?

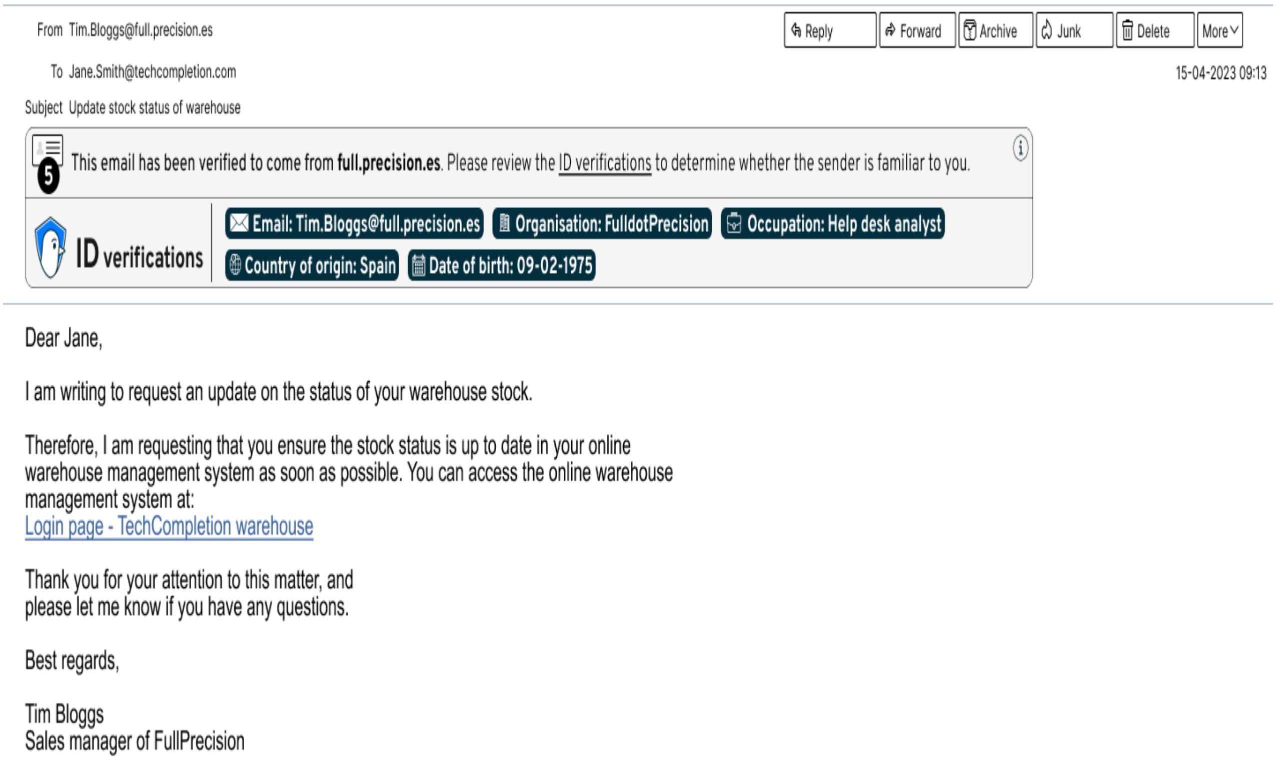
Please write your answer here:

Scenario 1: TechCompletion - condition D

Brief description of scenario 1:

You are Jane Smith, a warehouse manager of TechCompletion (`techcompletion.com`) located in the United States.
You communicate with colleagues within the company and outside the company such as FullPrecision (`fullprecision.co.uk`).

Given this scenario, consider the following email:



From Tim.Bloggs@full.precision.es

To Jane.Smith@techcompletion.com

15-04-2023 09:13

Subject Update stock status of warehouse

5 This email has been verified to come from **full.precision.es**. Please review the ID verifications to determine whether the sender is familiar to you.

ID verifications

Email: Tim.Bloggs@full.precision.es | Organisation: FulldotPrecision | Occupation: Help desk analyst | Country of origin: Spain | Date of birth: 09-02-1975

Dear Jane,

I am writing to request an update on the status of your warehouse stock.

Therefore, I am requesting that you ensure the stock status is up to date in your online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and
please let me know if you have any questions.

Best regards,

Tim Bloggs
Sales manager of FullPrecision

Would you click the link in the email and login? *

Please choose **only one** of the following:

○ Yes

○ No

How comfortable would you feel clicking the link in the email and login? *

Please choose the appropriate response for each item:

	Very uncomfortable						Very comfortabl
	○	○	○	○	○	○	○

And why? What aspects of the email message contributed to your response?

Please write your answer here:

How well do the following adjectives describe the **content of the email message** you just read?

*

Please choose the appropriate response for each item:

	Describes very poorly						Describes very well
Professional	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○

Please indicate for each statement how much it applies to you. I found the **sender of the email** ... *

Please choose the appropriate response for each item:

	Not at all						Extremely
Trustworthy	○	○	○	○	○	○	○
Credible	○	○	○	○	○	○	○
Reputable	○	○	○	○	○	○	○

Did this email come from your contact **Tim Bloggs**?

*

Please choose the appropriate response for each item:

	No, certainly	I do not know	Yes, certainly
	○	○	○

And why? What aspects of the email message contributed to your response?

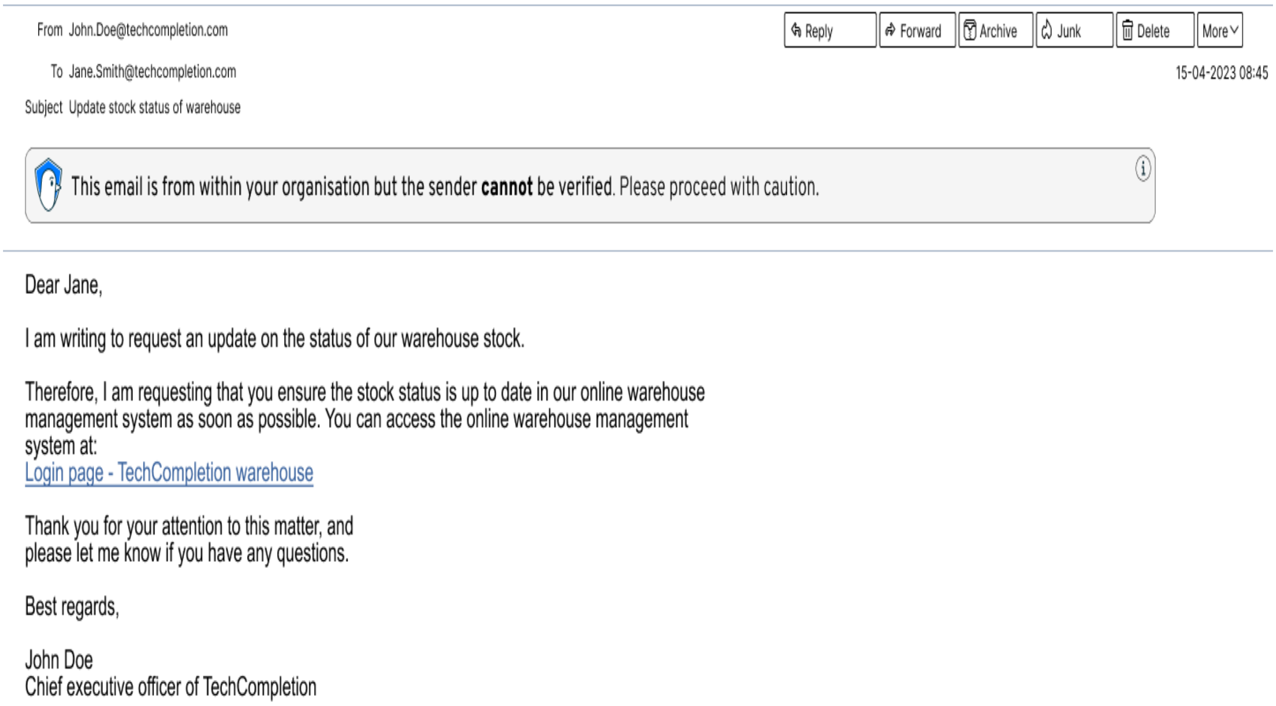
Please write your answer here:

Scenario 1: TechCompletion - condition E

Brief description of scenario 1:

You are Jane Smith, a warehouse manager of TechCompletion (`techcompletion.com`) located in the United States.
You communicate with colleagues within the company and outside the company such as FullPrecision (`fullprecision.co.uk`).

Given this scenario, consider the following email:



From John.Doe@techcompletion.com

To Jane.Smith@techcompletion.com

Subject Update stock status of warehouse

15-04-2023 08:45

This email is from within your organisation but the sender **cannot** be verified. Please proceed with caution.

Dear Jane,

I am writing to request an update on the status of our warehouse stock.

Therefore, I am requesting that you ensure the stock status is up to date in our online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and
please let me know if you have any questions.

Best regards,

John Doe
Chief executive officer of TechCompletion

Would you click the link in the email and login? *

Please choose **only one** of the following:

○ Yes

○ No

How comfortable would you feel clicking the link in the email and login? *

Please choose the appropriate response for each item:

	Very uncomfortable						**Very comfortabl**
	○	○	○	○	○	○	○

And why? What aspects of the email message contributed to your response?

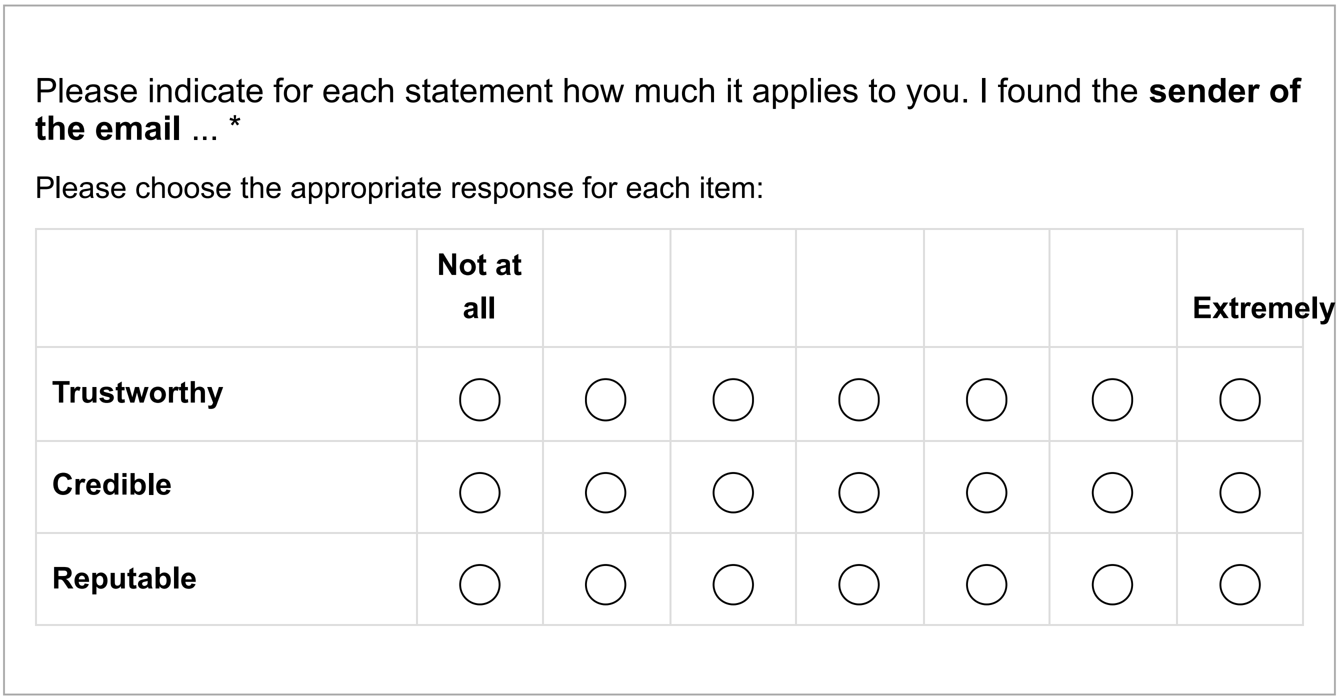
Please write your answer here:

How well do the following adjectives describe the **content of the email message** you just read?

*

Please choose the appropriate response for each item:

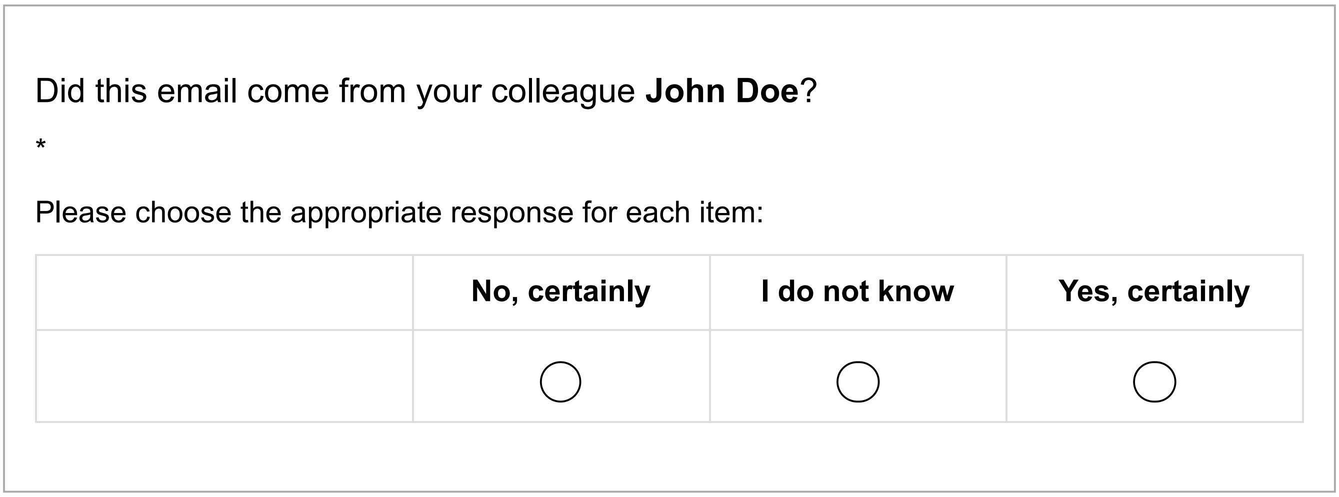
	Describes very poorly						**Describes very well**
Professional	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○



Please indicate for each statement how much it applies to you. I found the **sender of the email** ... *

Please choose the appropriate response for each item:

	Not at all						**Extremely**
Trustworthy	○	○	○	○	○	○	○
Credible	○	○	○	○	○	○	○
Reputable	○	○	○	○	○	○	○

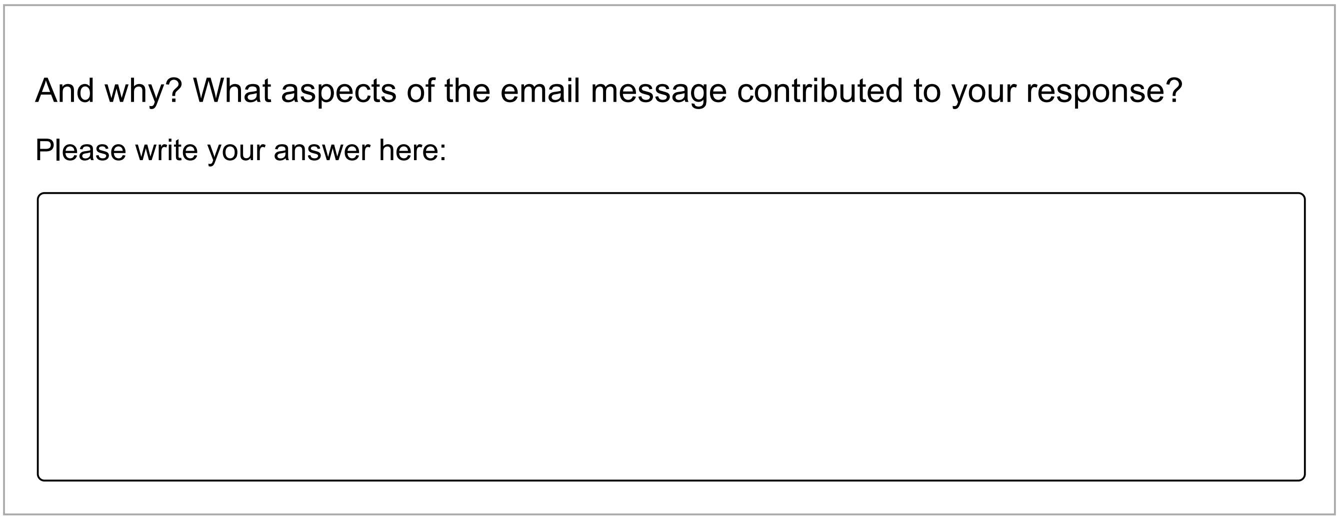


Did this email come from your colleague **John Doe**?

*

Please choose the appropriate response for each item:

	No, certainly	**I do not know**	**Yes, certainly**
	○	○	○



And why? What aspects of the email message contributed to your response?

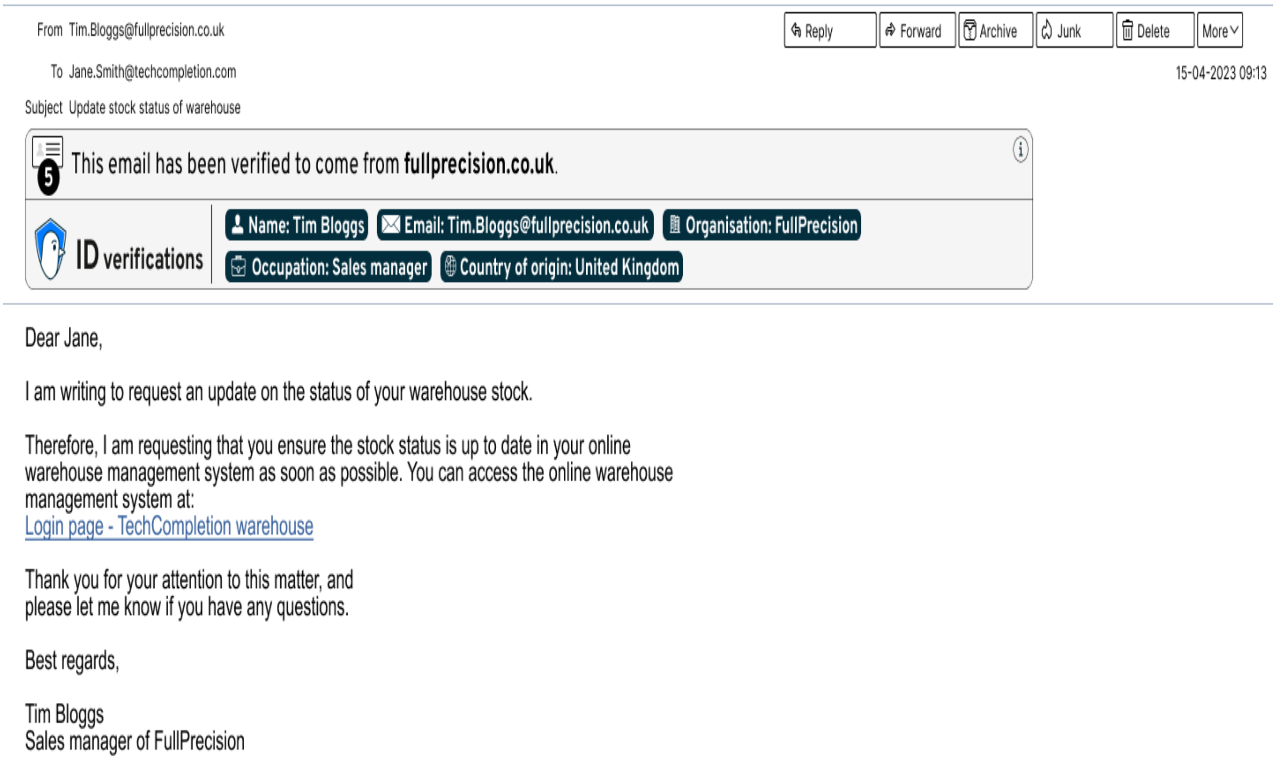
Please write your answer here:

Scenario 1: TechCompletion - condition F

Brief description of scenario 1:

You are Jane Smith, a warehouse manager of TechCompletion (`techcompletion.com`) located in the United States.
You communicate with colleagues within the company and outside the company such as FullPrecision (`fullprecision.co.uk`).

Given this scenario, consider the following email:



From Tim.Bloggs@fullprecision.co.uk

To Jane.Smith@techcompletion.com

15-04-2023 09:13

Subject Update stock status of warehouse

This email has been verified to come from **fullprecision.co.uk**.

Name: Tim Bloggs | Email: Tim.Bloggs@fullprecision.co.uk | Organisation: FullPrecision | Occupation: Sales manager | Country of origin: United Kingdom

Dear Jane,

I am writing to request an update on the status of your warehouse stock.

Therefore, I am requesting that you ensure the stock status is up to date in your online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and
please let me know if you have any questions.

Best regards,

Tim Bloggs
Sales manager of FullPrecision

Would you click the link in the email and login? *

Please choose **only one** of the following:

○ Yes

○ No

How comfortable would you feel clicking the link in the email and login? *

Please choose the appropriate response for each item:

	Very uncomfortable					**Very comfortabl**
	○	○	○	○	○	○	○

And why? What aspects of the email message contributed to your response?

Please write your answer here:

How well do the following adjectives describe the **content of the email message** you just read?

*

Please choose the appropriate response for each item:

	Describes very poorly						**Describes very well**
Professional	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○

Please indicate for each statement how much it applies to you. I found the **sender of the email** ... *

Please choose the appropriate response for each item:

	Not at all						Extremely
Trustworthy	○	○	○	○	○	○	○
Credible	○	○	○	○	○	○	○
Reputable	○	○	○	○	○	○	○

Did this email come from your contact **Tim Bloggs**?

*

Please choose the appropriate response for each item:

	No, certainly	I do not know	Yes, certainly
	○	○	○

And why? What aspects of the email message contributed to your response?

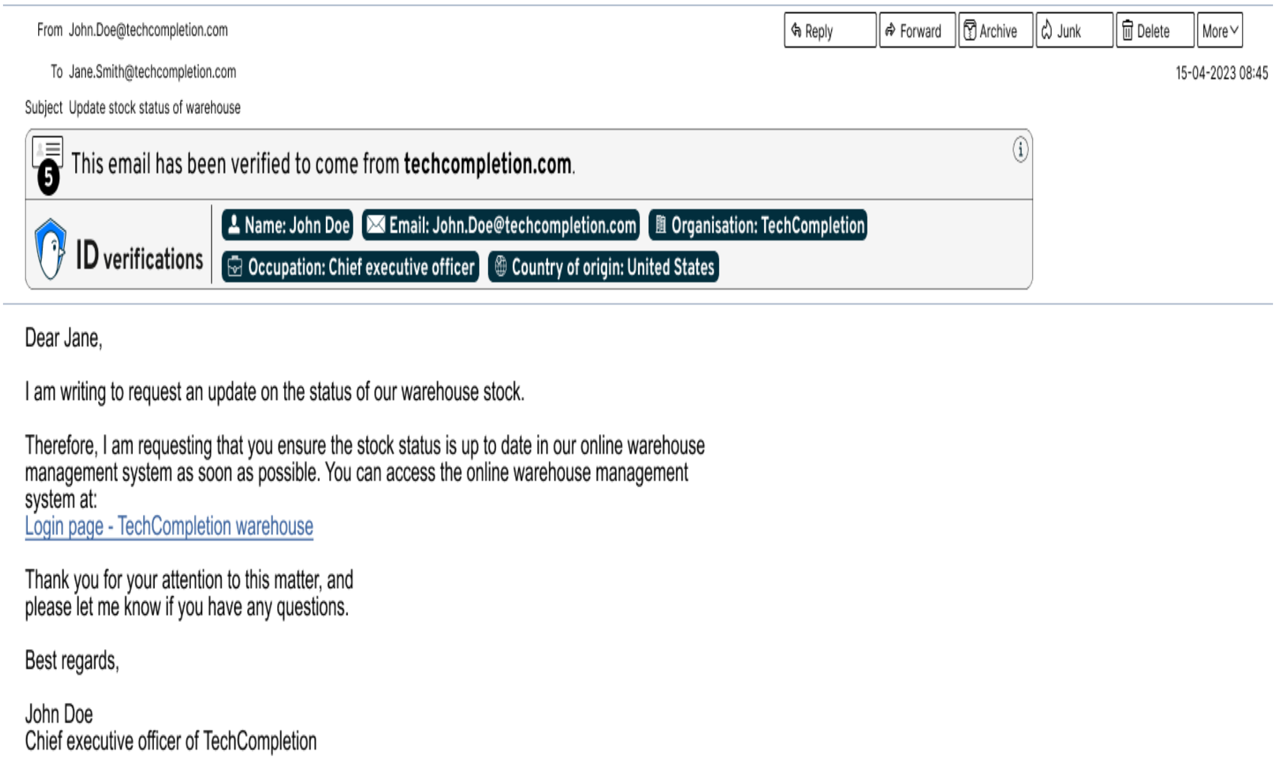
Please write your answer here:

Scenario 1: TechCompletion - condition G

Brief description of scenario 1:

You are Jane Smith, a warehouse manager of TechCompletion (`techcompletion.com`) located in the United States.
You communicate with colleagues within the company and outside the company such as FullPrecision (`fullprecision.co.uk`).

Given this scenario, consider the following email:



From John.Doe@techcompletion.com

To Jane.Smith@techcompletion.com

15-04-2023 08:45

Subject Update stock status of warehouse

5 This email has been verified to come from **techcompletion.com**.

ID verifications: Name: John Doe | Email: John.Doe@techcompletion.com | Organisation: TechCompletion | Occupation: Chief executive officer | Country of origin: United States

Dear Jane,

I am writing to request an update on the status of our warehouse stock.

Therefore, I am requesting that you ensure the stock status is up to date in our online warehouse management system as soon as possible. You can access the online warehouse management system at:
Login page - TechCompletion warehouse

Thank you for your attention to this matter, and please let me know if you have any questions.

Best regards,

John Doe
Chief executive officer of TechCompletion

Would you click the link in the email and login? *

Please choose **only one** of the following:

○ Yes

○ No

How comfortable would you feel clicking the link in the email and login? *

Please choose the appropriate response for each item:

	Very uncomfortable						Very comfortabl
	○	○	○	○	○	○	○

And why? What aspects of the email message contributed to your response?

Please write your answer here:

How well do the following adjectives describe the **content of the email message** you just read?

*

Please choose the appropriate response for each item:

	Describes very poorly						Describes very well
Professional	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○

Please indicate for each statement how much it applies to you. I found the **sender of the email** ... *

Please choose the appropriate response for each item:

	Not at all						Extremely
Trustworthy	○	○	○	○	○	○	○
Credible	○	○	○	○	○	○	○
Reputable	○	○	○	○	○	○	○

Did this email come from your colleague **John Doe**?

*

Please choose the appropriate response for each item:

	No, certainly	I do not know	Yes, certainly
	○	○	○

And why? What aspects of the email message contributed to your response?

Please write your answer here:

Scenario 2: MijnOverheid - introduction

Welcome to the second scenario of this research.

Please read the information below carefully.
You are still Jane Smith, a warehouse manager of TechCompletion.
You have received some emails from MijnOverheid (`mijn.overheid.nl`) which is the digital platform for your dealings with Dutch authorities.

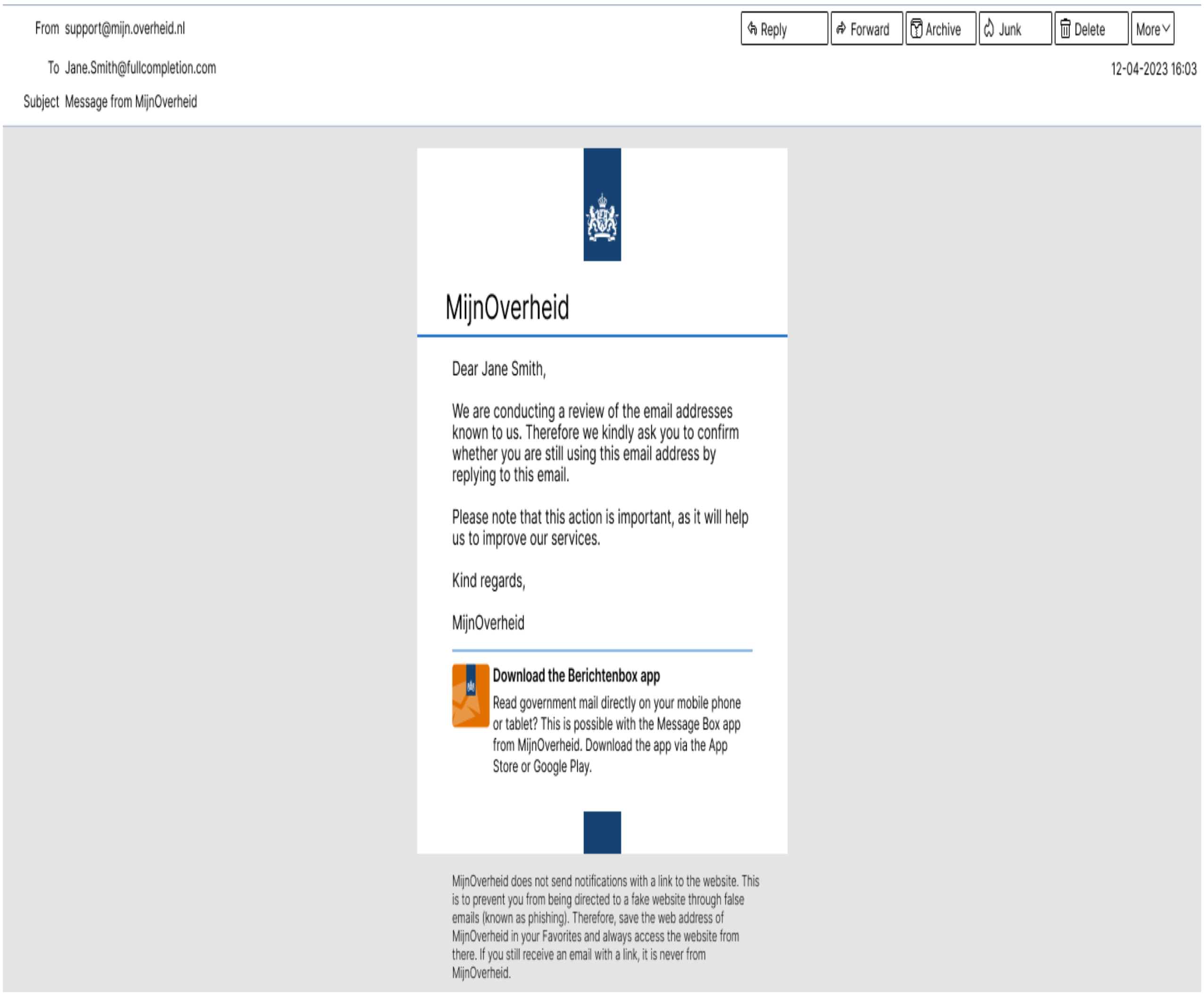
You also still have PostGuard installed on the applications you use for sending and receiving emails.

You will now see six email messages.

Scenario 2: MijnOverheid - condition I

Brief description of scenario 2:

You are still Jane Smith, a warehouse manager of TechCompletion.
You have received the following email:



From support@mijn.overheid.nl
To Jane.Smith@fullcompletion.com
Subject Message from MijnOverheid
12-04-2023 16:03

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

MijnOverheid

Download the Berichtenbox app
Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

Would you reply to the email? *

Please choose **only one** of the following:

○ Yes

○ No

How comfortable would you feel replying to the email? *

Please choose the appropriate response for each item:

	Very uncomfortable						**Very comfortabl**
	○	○	○	○	○	○	○

And why? What aspects of the email message contributed to your response?

Please write your answer here:

How well do the following adjectives describe the **content of the email message** you just read?

*

Please choose the appropriate response for each item:

	Describes very poorly						**Describes very well**
Professional	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○

Please indicate for each statement how much it applies to you. I found the **sender of the email** ... *

Please choose the appropriate response for each item:

	Not at all						**Extremely**
Trustworthy	○	○	○	○	○	○	○
Credible	○	○	○	○	○	○	○
Reputable	○	○	○	○	○	○	○

Did this email come from **MijnOverheid**?

*

Please choose the appropriate response for each item:

	No, certainly	**I do not know**	**Yes, certainly**
	○	○	○

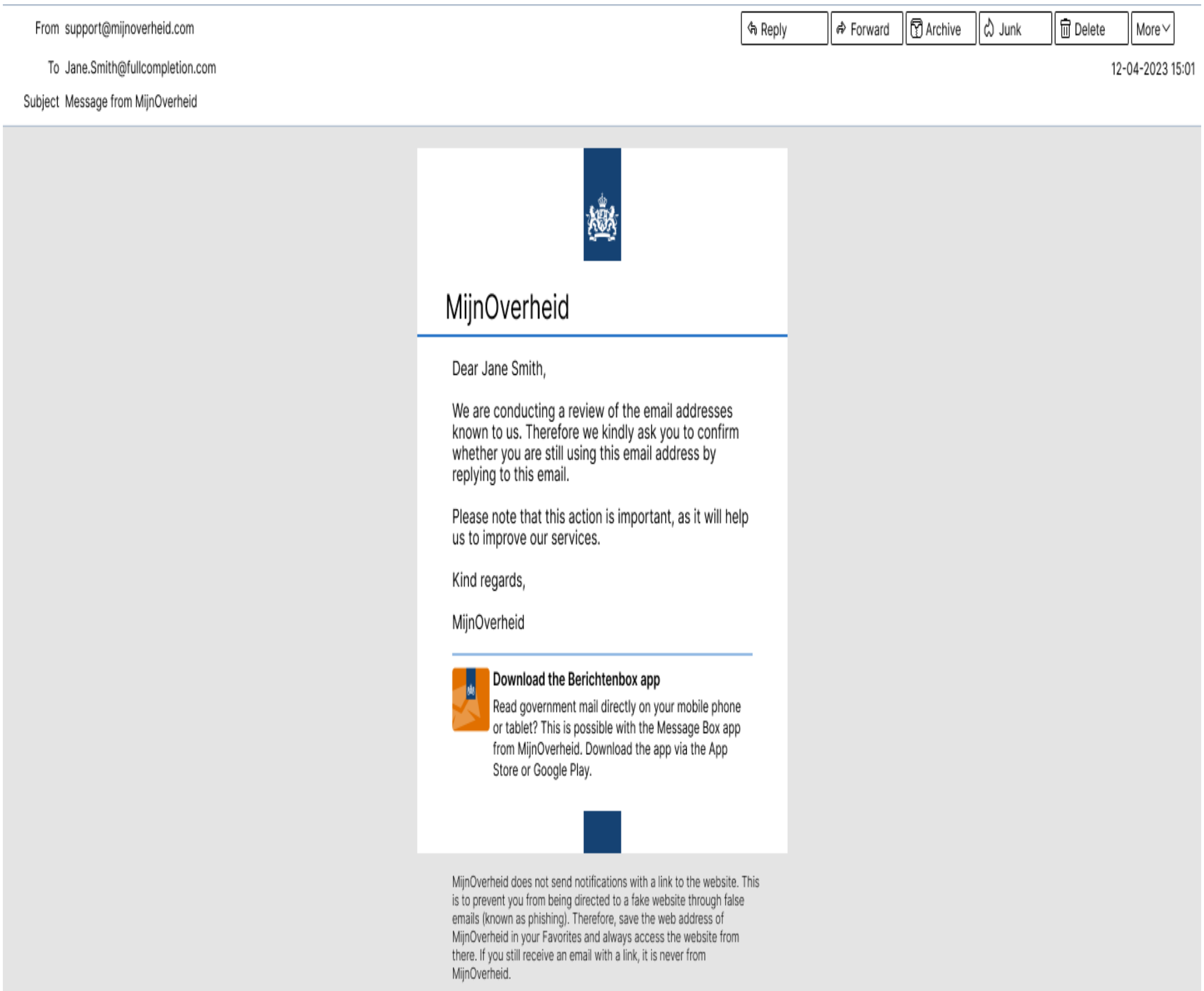
And why? What aspects of the email message contributed to your response?

Please write your answer here:

Scenario 2: MijnOverheid - condition II

Brief description of scenario 2:

You are still Jane Smith, a warehouse manager of TechCompletion.
You have received the following email:



From support@mijnoverheid.com

To Jane.Smith@fullcompletion.com

Subject Message from MijnOverheid

12-04-2023 15:01

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

MijnOverheid

Download the Berichtenbox app

Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

Would you reply to the email? *

Please choose **only one** of the following:

○ Yes

○ No

How comfortable would you feel replying to the email? *

Please choose the appropriate response for each item:

	Very uncomfortable						**Very comfortabl**
	○	○	○	○	○	○	○

And why? What aspects of the email message contributed to your response?

Please write your answer here:

How well do the following adjectives describe the **content of the email message** you just read?

*

Please choose the appropriate response for each item:

	Describes very poorly					Describes very well
Professional	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○

Please indicate for each statement how much it applies to you. I found the **sender of the email** ... *

Please choose the appropriate response for each item:

	Not at all						Extremely
Trustworthy	○	○	○	○	○	○	○
Credible	○	○	○	○	○	○	○
Reputable	○	○	○	○	○	○	○

Did this email come from **MijnOverheid**?

*

Please choose the appropriate response for each item:

	No, certainly	I do not know	Yes, certainly
	○	○	○

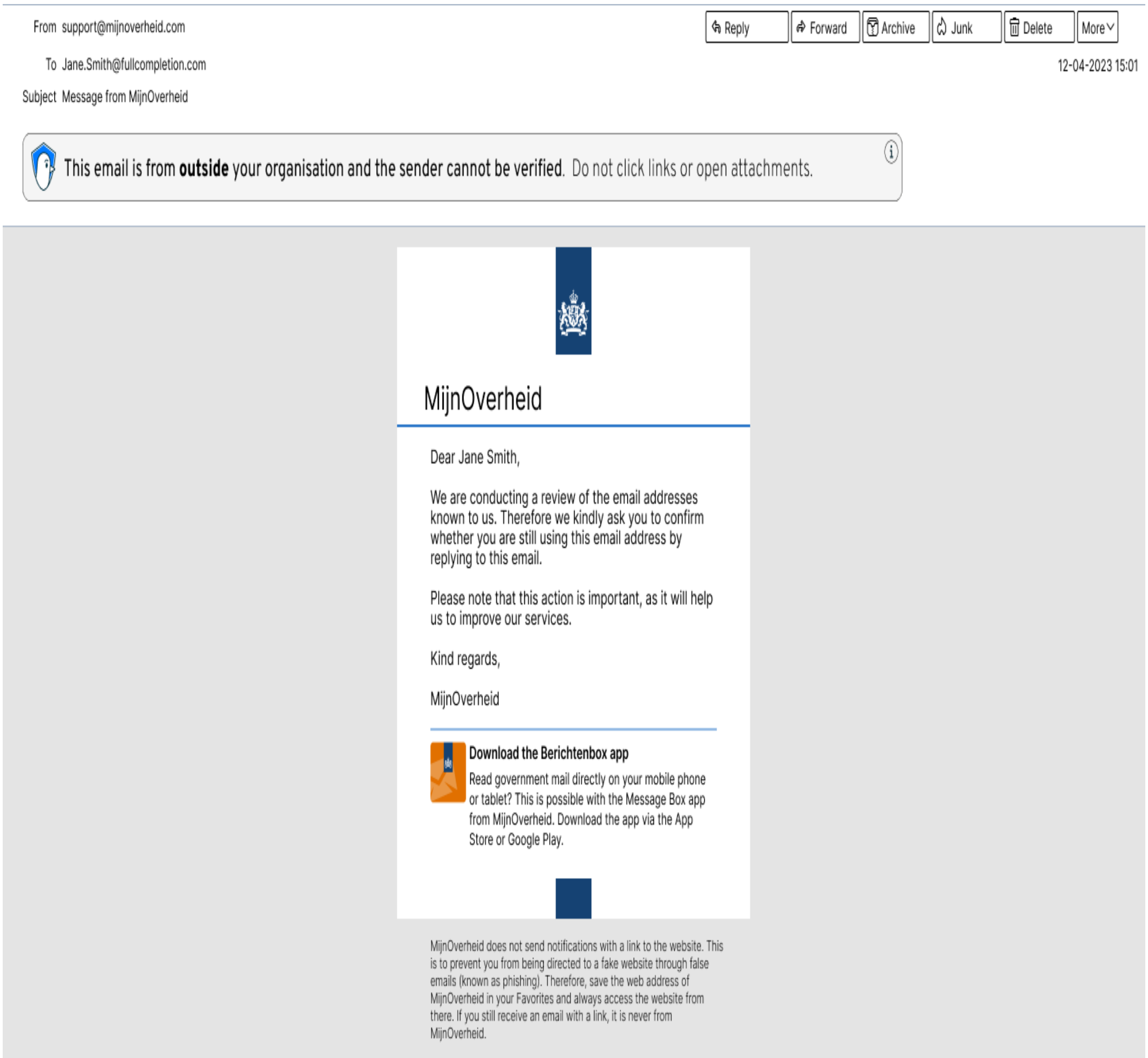
And why? What aspects of the email message contributed to your response?

Please write your answer here:

Scenario 2: MijnOverheid - condition III

Brief description of scenario 2:

You are still Jane Smith, a warehouse manager of TechCompletion.
You have received the following email:



From support@mijnoverheid.com

To Jane.Smith@fullcompletion.com

Subject Message from MijnOverheid

12-04-2023 15:01

This email is from **outside** your organisation and the sender cannot be verified. Do not click links or open attachments.

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

MijnOverheid

Download the Berichtenbox app

Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

Would you reply to the email? *

Please choose **only one** of the following:

○ Yes

○ No

How comfortable would you feel replying to the email? *

Please choose the appropriate response for each item:

	Very uncomfortable						Very comfortabl
	○	○	○	○	○	○	○

And why? What aspects of the email message contributed to your response?

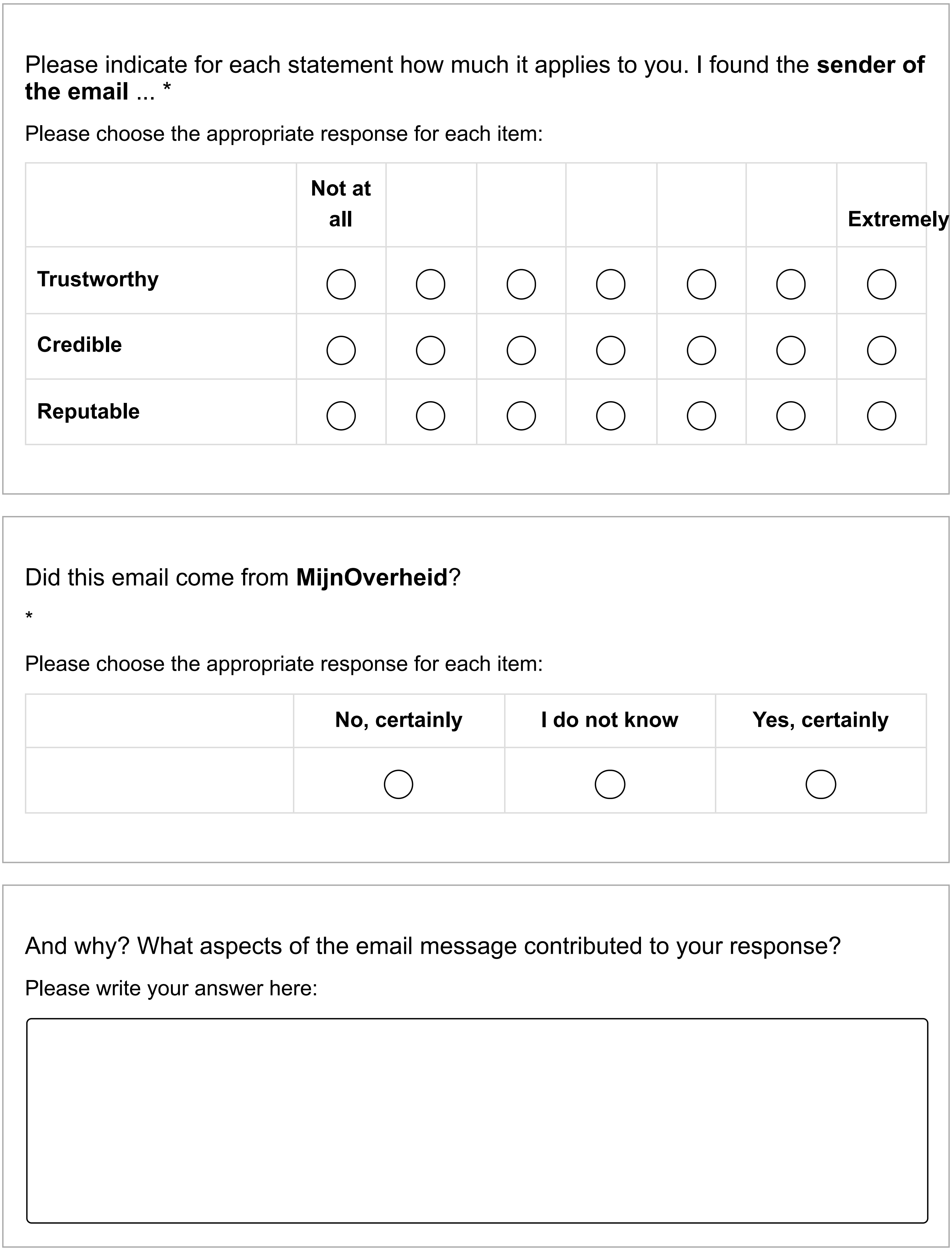
Please write your answer here:

How well do the following adjectives describe the **content of the email message** you just read?

*

Please choose the appropriate response for each item:

	Describes very poorly						Describes very well
Professional	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○



Please indicate for each statement how much it applies to you. I found the **sender of the email** ... *

Please choose the appropriate response for each item:

	Not at all						Extremely
Trustworthy	○	○	○	○	○	○	○
Credible	○	○	○	○	○	○	○
Reputable	○	○	○	○	○	○	○

Did this email come from **MijnOverheid**?

*

Please choose the appropriate response for each item:

	No, certainly	I do not know	Yes, certainly
	○	○	○

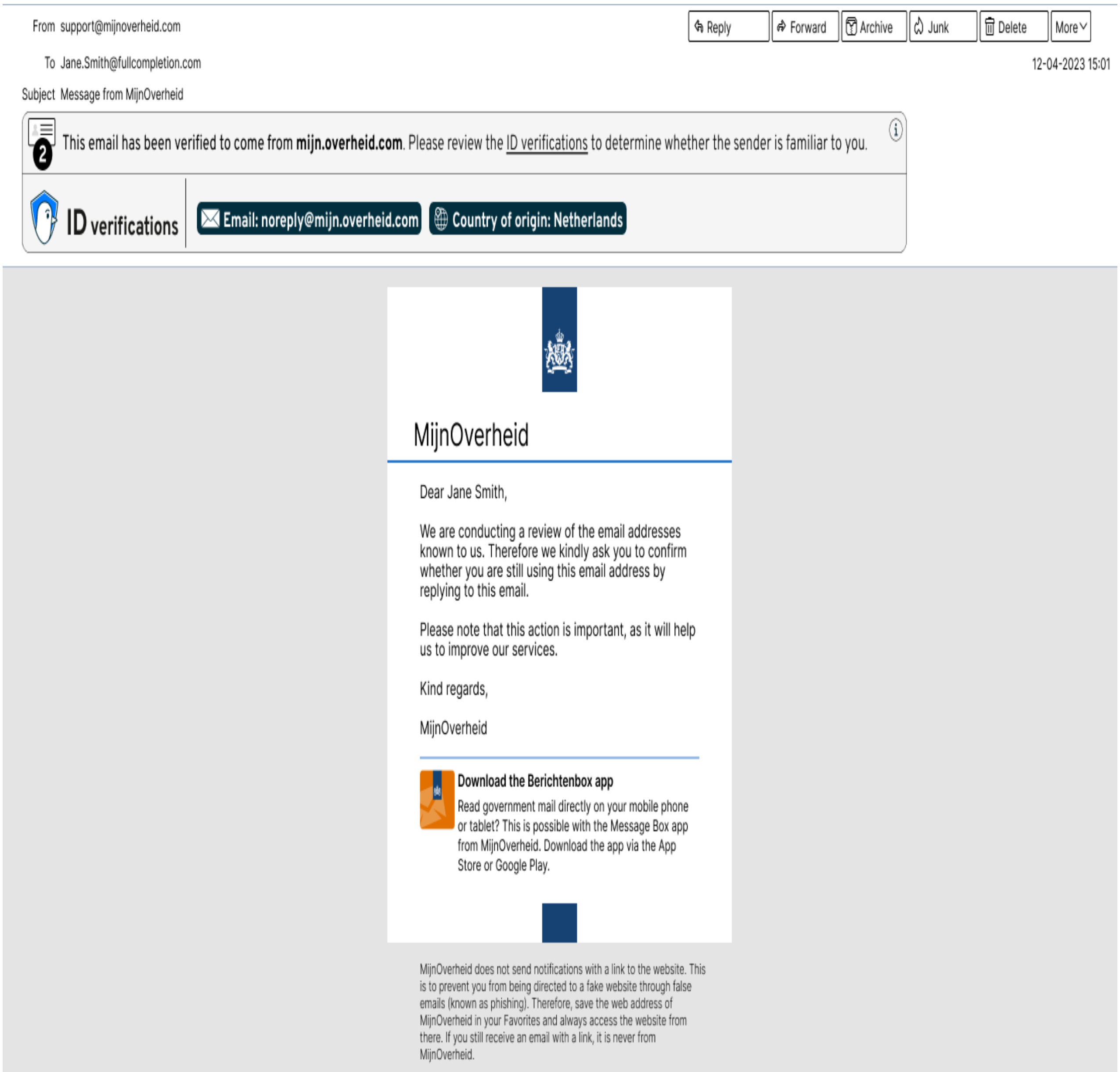
And why? What aspects of the email message contributed to your response?

Please write your answer here:

Scenario 2: MijnOverheid - condition IV

Brief description of scenario 2:

You are still Jane Smith, a warehouse manager of TechCompletion.
You have received the following email:



From support@mijnoverheid.com

To Jane.Smith@fullcompletion.com

12-04-2023 15:01

Subject Message from MijnOverheid

This email has been verified to come from **mijn.overheid.com**. Please review the ID verifications to determine whether the sender is familiar to you.

ID verifications

Email: noreply@mijn.overheid.com

Country of origin: Netherlands

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

MijnOverheid

Download the Berichtenbox app

Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

Would you reply to the email? *

Please choose **only one** of the following:

○ Yes

○ No

How comfortable would you feel replying to the email? *

Please choose the appropriate response for each item:

	Very uncomfortable						**Very comfortabl**
	○	○	○	○	○	○	○

And why? What aspects of the email message contributed to your response?

Please write your answer here:

How well do the following adjectives describe the **content of the email message** you just read?

*

Please choose the appropriate response for each item:

	Describes very poorly						**Describes very well**
Professional	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○

Please indicate for each statement how much it applies to you. I found the **sender of the email** ... *

Please choose the appropriate response for each item:

	Not at all						Extremely
Trustworthy	○	○	○	○	○	○	○
Credible	○	○	○	○	○	○	○
Reputable	○	○	○	○	○	○	○

Did this email come from **MijnOverheid**?

*

Please choose the appropriate response for each item:

	No, certainly	I do not know	Yes, certainly
	○	○	○

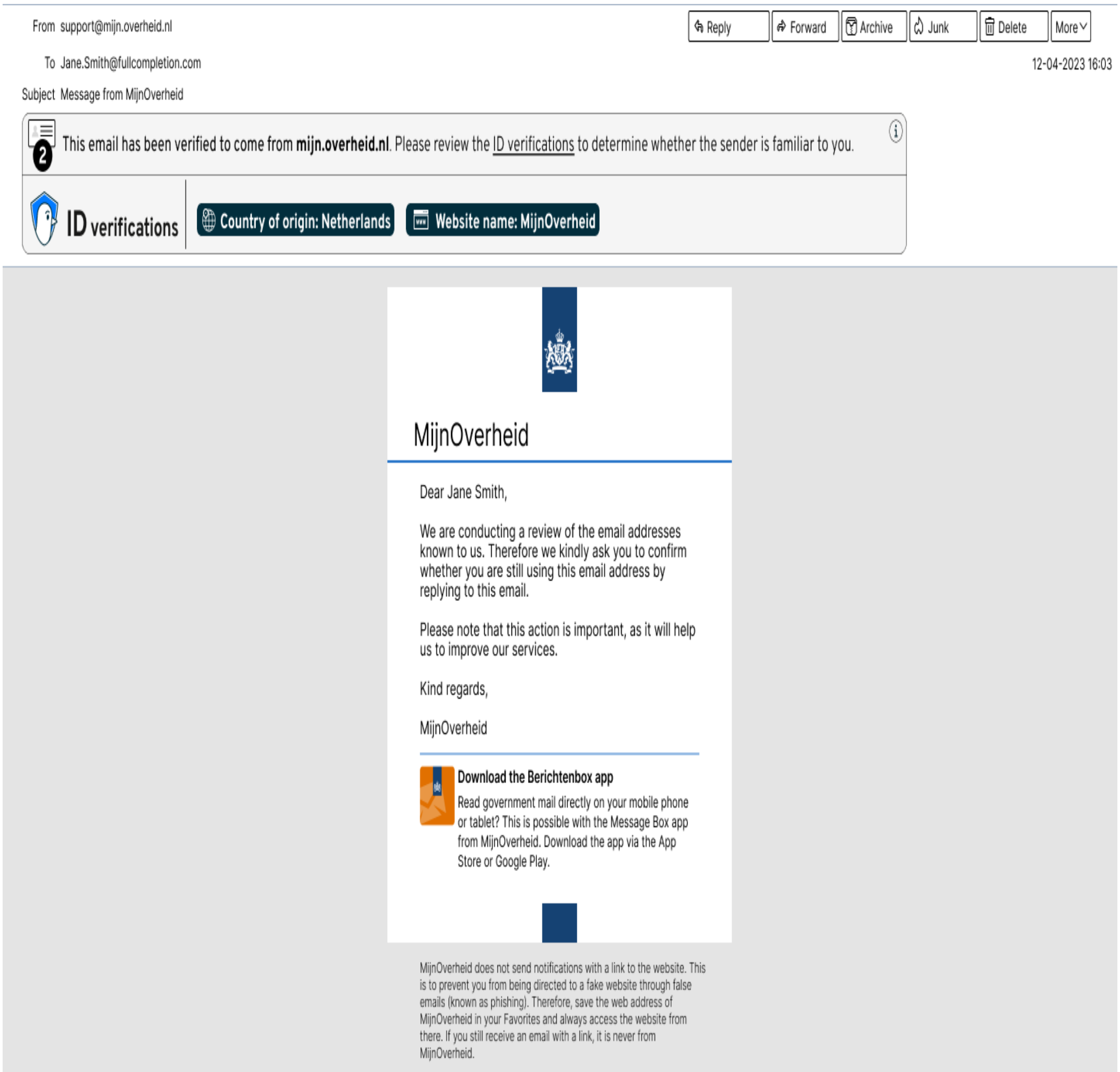
And why? What aspects of the email message contributed to your response?

Please write your answer here:

Scenario 2: MijnOverheid - condition V

Brief description of scenario 2:

You are still Jane Smith, a warehouse manager of TechCompletion.
You have received the following email:



From support@mijn.overheid.nl

To Jane.Smith@fullcompletion.com

12-04-2023 16:03

Subject Message from MijnOverheid

2 This email has been verified to come from **mijn.overheid.nl**. Please review the ID verifications to determine whether the sender is familiar to you.

ID verifications | Country of origin: Netherlands | Website name: MijnOverheid

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

MijnOverheid

Download the Berichtenbox app

Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

Would you reply to the email?

*

Please choose **only one** of the following:

○ Yes

○ No

How comfortable would you feel replying to the email? *

Please choose the appropriate response for each item:

	Very uncomfortable						Very comfortabl
	○	○	○	○	○	○	○

And why? What aspects of the email message contributed to your response?

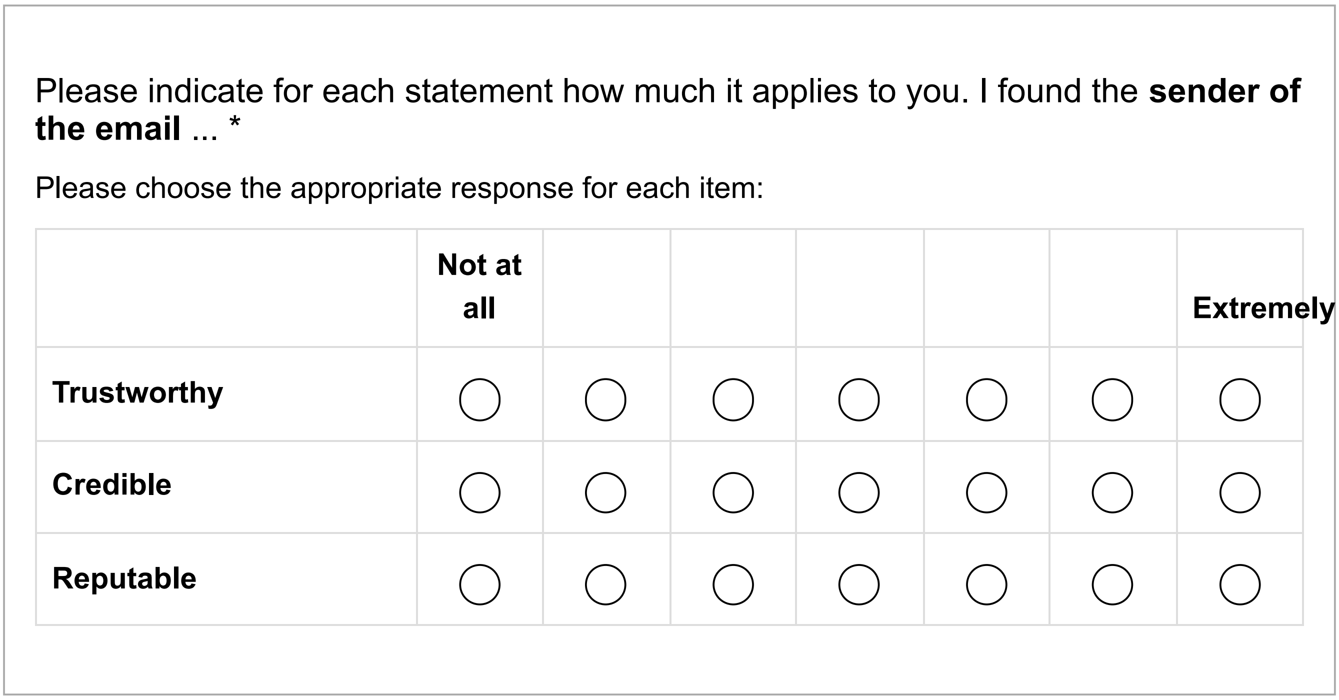
Please write your answer here:

How well do the following adjectives describe the **content of the email message** you just read?

*

Please choose the appropriate response for each item:

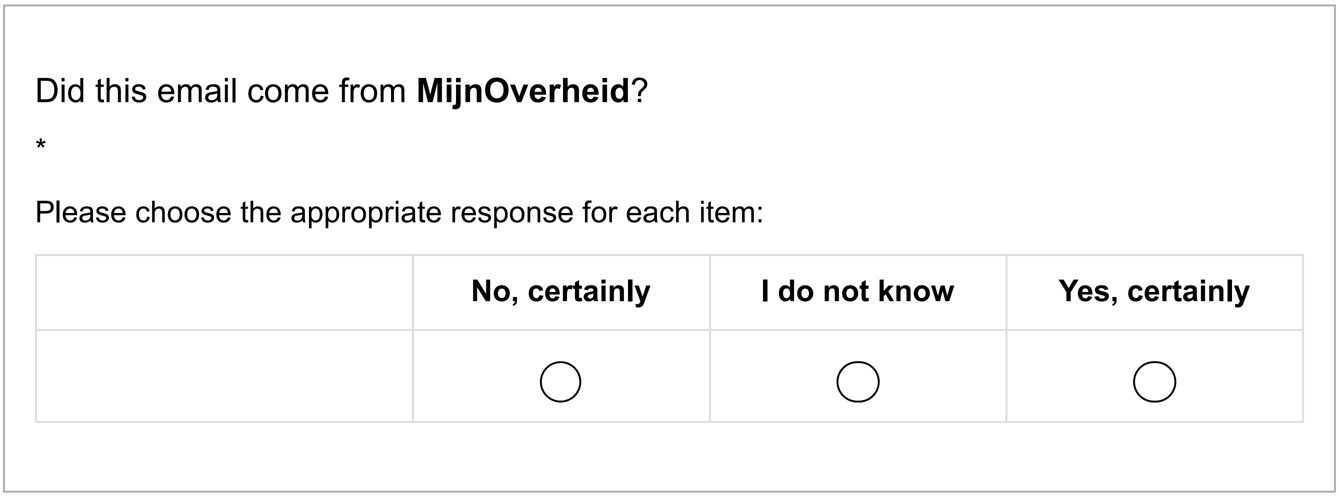
	Describes very poorly						Describes very well
Professional	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○



Please indicate for each statement how much it applies to you. I found the **sender of the email** ... *

Please choose the appropriate response for each item:

	Not at all						Extremely
Trustworthy	○	○	○	○	○	○	○
Credible	○	○	○	○	○	○	○
Reputable	○	○	○	○	○	○	○

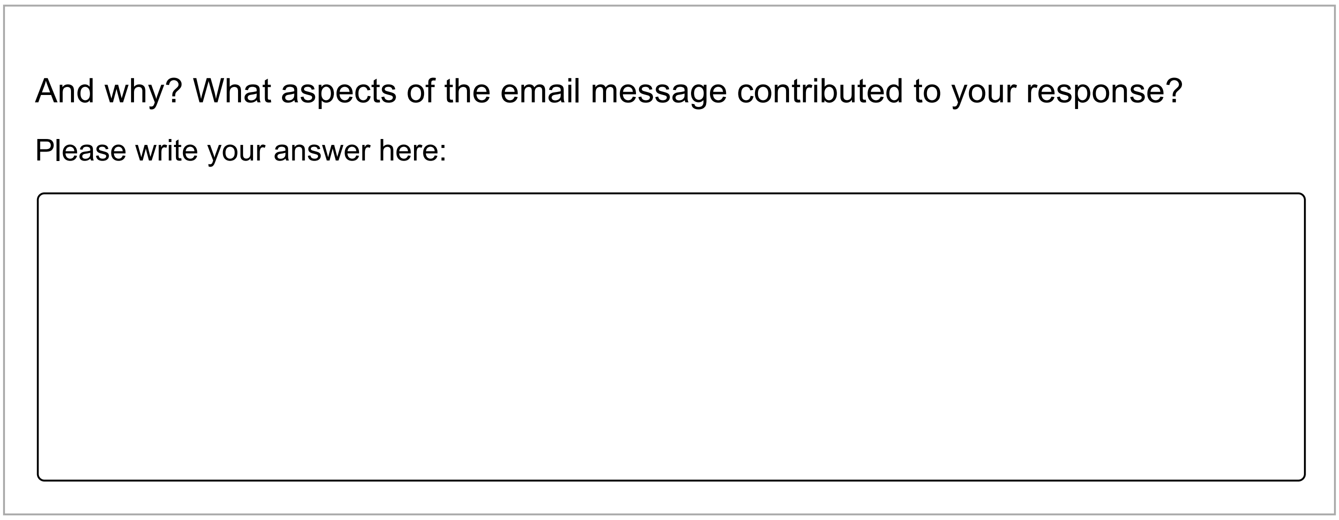


Did this email come from **MijnOverheid**?

*

Please choose the appropriate response for each item:

	No, certainly	I do not know	Yes, certainly
	○	○	○



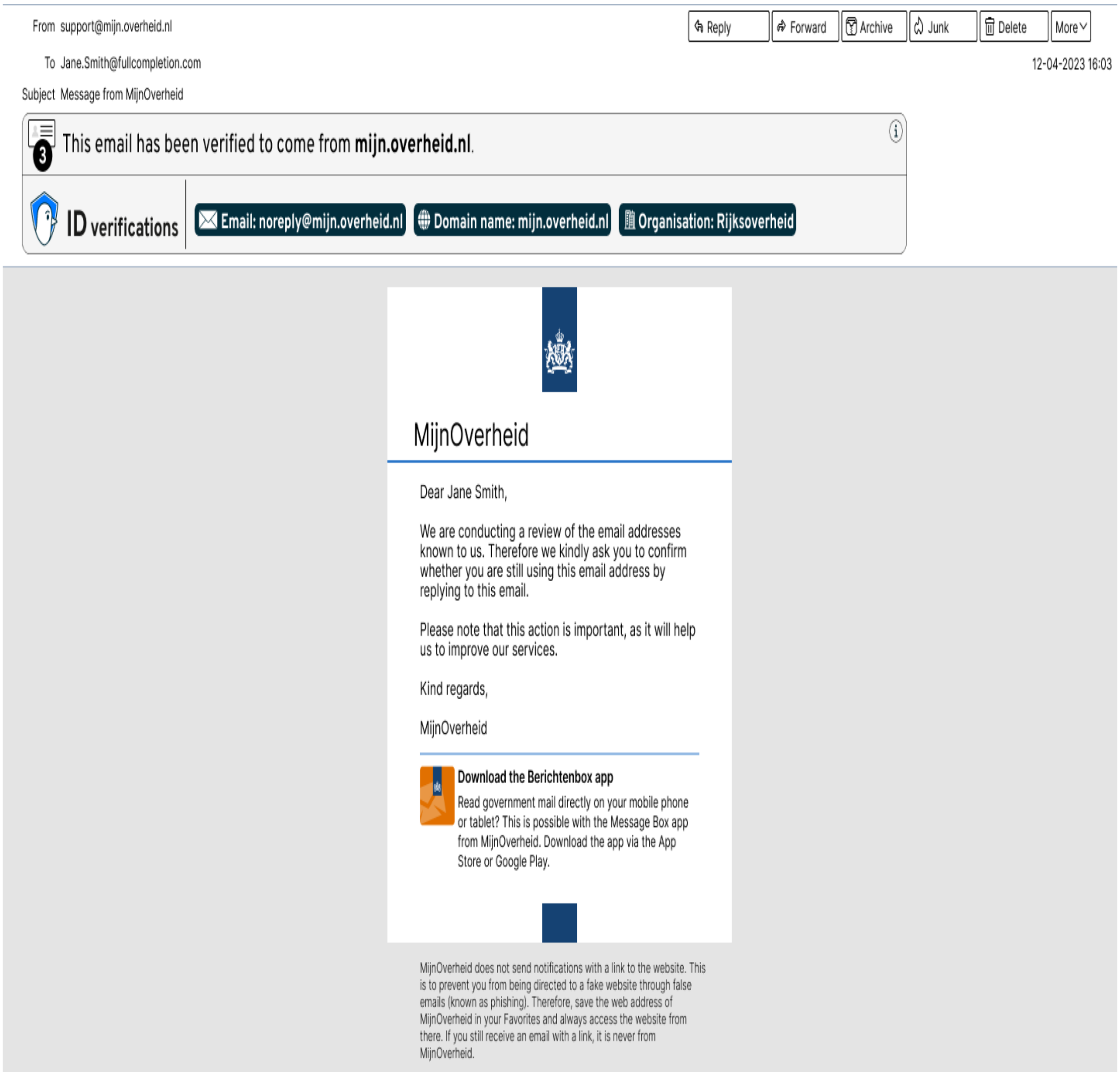
And why? What aspects of the email message contributed to your response?

Please write your answer here:

Scenario 2: MijnOverheid - condition VI

Brief description of scenario 2:

You are still Jane Smith, a warehouse manager of TechCompletion.
You have received the following email:



From support@mijn.overheid.nl

To Jane.Smith@fullcompletion.com

12-04-2023 16:03

Subject Message from MijnOverheid

This email has been verified to come from **mijn.overheid.nl**.

ID verifications: Email: noreply@mijn.overheid.nl | Domain name: mijn.overheid.nl | Organisation: Rijksoverheid

MijnOverheid

Dear Jane Smith,

We are conducting a review of the email addresses known to us. Therefore we kindly ask you to confirm whether you are still using this email address by replying to this email.

Please note that this action is important, as it will help us to improve our services.

Kind regards,

MijnOverheid

Download the Berichtenbox app

Read government mail directly on your mobile phone or tablet? This is possible with the Message Box app from MijnOverheid. Download the app via the App Store or Google Play.

MijnOverheid does not send notifications with a link to the website. This is to prevent you from being directed to a fake website through false emails (known as phishing). Therefore, save the web address of MijnOverheid in your Favorites and always access the website from there. If you still receive an email with a link, it is never from MijnOverheid.

Would you reply to the email?

*

Please choose **only one** of the following:

○ Yes

○ No

How comfortable would you feel replying to the email? *

Please choose the appropriate response for each item:

	Very uncomfortable						Very comfortabl
	○	○	○	○	○	○	○

And why? What aspects of the email message contributed to your response?

Please write your answer here:

How well do the following adjectives describe the **content of the email message** you just read?

*

Please choose the appropriate response for each item:

	Describes very poorly						Describes very well
Professional	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○

Please indicate for each statement how much it applies to you. I found the **sender of the email** ... *

Please choose the appropriate response for each item:

	Not at all						Extremely
Trustworthy	○	○	○	○	○	○	○
Credible	○	○	○	○	○	○	○
Reputable	○	○	○	○	○	○	○

Did this email come from **MijnOverheid**?

*

Please choose the appropriate response for each item:

	No, certainly	I do not know	Yes, certainly
	○	○	○

And why? What aspects of the email message contributed to your response?

Please write your answer here:

Manipulation checks

Which **organisations** were mentioned in the emails you have seen in the **first** scenario? *

Please choose **all** that apply:

□ TechCompletion

□ FullPrecision

□ Google

□ Microsoft

□ I do not remember

Which organisation was mentioned in the email you have seen in the **second** scenario?

*

Please choose **only one** of the following:

○ Radboud University

○ MijnOverheid

○ Google

○ Microsoft

○ I do not remember

Did you pay attention when filling in the questions? *

Please choose **only one** of the following:

○ Yes

○ No

Experience and attitudes with emails

Have you ever **received** a message via email, SMS, phone call or other communication media where someone **pretended** to be a person or organisation that you are familiar with, and requested you to provide personal information or carry out an action? *

Please choose **only one** of the following:

○ Yes

○ No

○ I do not know

What advise would you give to colleagues about opening links and attachments in emails?

Please write your answer here:

Feedback Add-on

If your email application (Outlook, Gmail, ...) supported the feature of displaying ID verifications of the email's sender, would you carefully look at the sender's ID verifications and include your own ID verifications when sending an email?

*

Please choose **only one** of the following:

○ Yes

○ No

○ Maybe

Feedback Experiment

Did you encounter any issues during the experiment or is there something about the experiment that can be improved?

Please write your answer here:

Debriefing and final consent

Experiment debrief:

The aim of this experiment is to investigate whether our visual designs of digital signatures help users in identifying malicious email messages that pretend to be from a known sender but in reality are not. This is done so that effective designs of digital signatures can be created and included in the email add-on **PostGuard** (*https://postguard.nl (https://postguard.nl)*).
These digital signatures can be used by a sender of an email to show the recipient that it really is the sender by including ID verifications.
All the emails that were presented in this experiment are fictional.

If you are interested in receiving the final thesis about this research, or more information about the presented emails, please contact me via leon.zhang@ru.nl. For questions or concerns, please feel free to contact my supervisor Hanna Schraffenberger via hanna.schraffenberger@ru.nl or me via leon.zhang@ru.nl.

Thank you for your time!

□

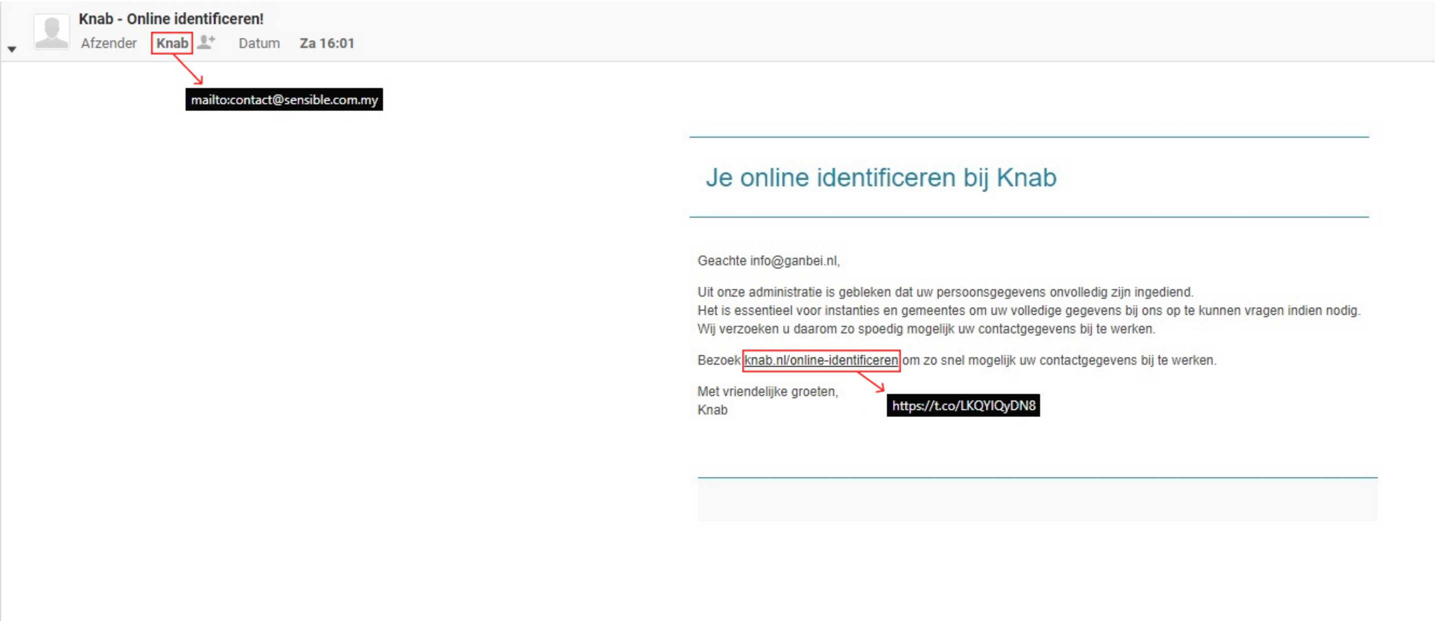
15.05.2023 – 12:38

Submit your survey.
Thank you for completing this survey.

Appendix F

Example emails

Here are some examples of malicious emails that ended up in our regular inbox without any warnings.



Knab - Online identificeren!

Afzender Knab Datum Za 16:01

mailto:contact@sensible.com.my

Je online identificeren bij Knab

Geachte info@ganbei.nl,

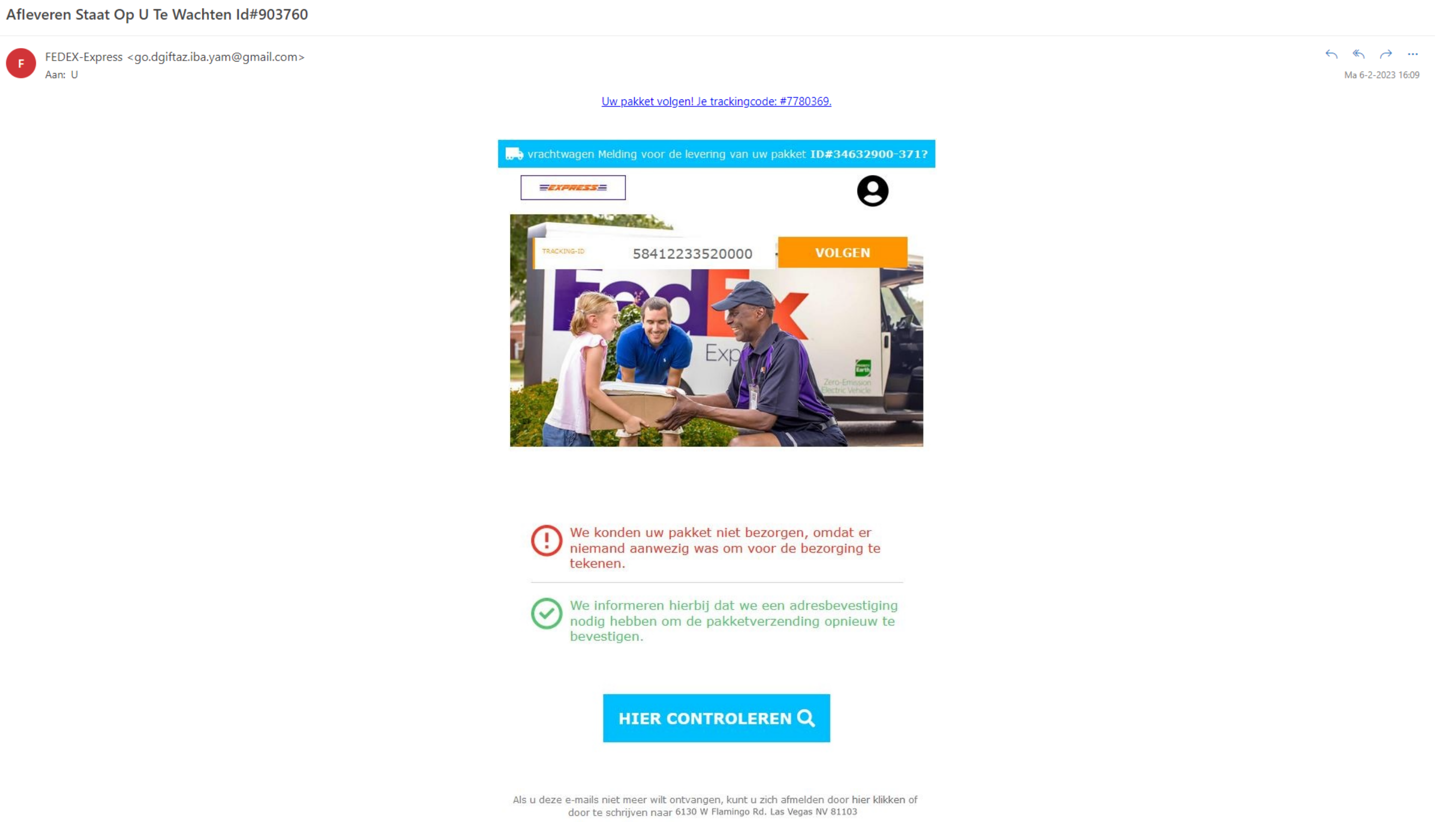
Uit onze administratie is gebleken dat uw persoonsgegevens onvolledig zijn ingediend.
Het is essentieel voor instanties en gemeentes om uw volledige gegevens bij ons op te kunnen vragen indien nodig.
Wij verzoeken u daarom zo spoedig mogelijk uw contactgegevens bij te werken.

Bezoek knab.nl/online-identificeren om zo snel mogelijk uw contactgegevens bij te werken.

https://t.co/LKQYIQyDN8

Met vriendelijke groeten,
Knab

Figure F.1: Knab phishing email in our organisational email account.



Afleveren Staat Op U Te Wachten Id#903760

FEDEX-Express <go.dgiftaz.iba.yam@gmail.com>
Aan: U

Ma 6-2-2023 16:09

Uw pakket volgen! Je trackingcode: #7780369.

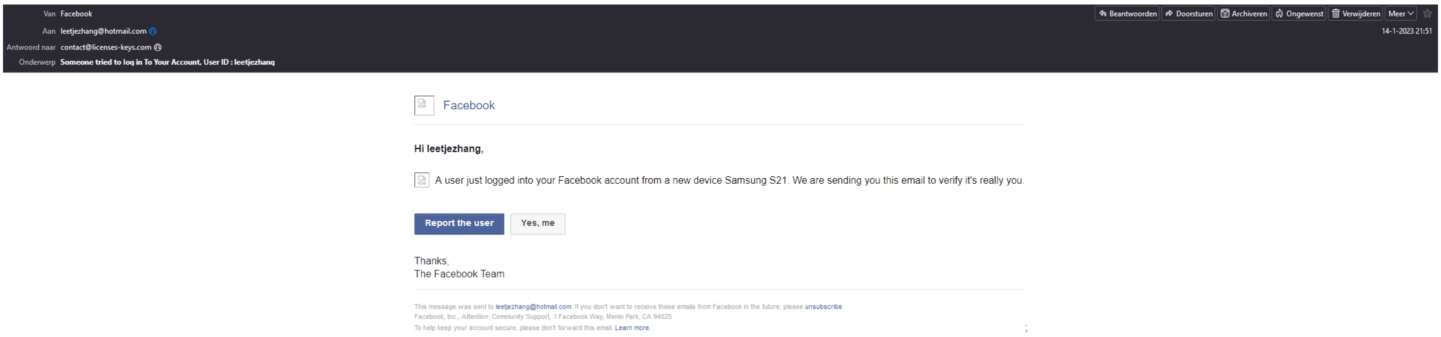
vrachtwagen Melding voor de levering van uw pakket **ID#34632900-371?**

We konden uw pakket niet bezorgen, omdat er niemand aanwezig was om voor de bezorging te tekenen.

We informeren hierbij dat we een adresbevestiging nodig hebben om de pakketverzending opnieuw te bevestigen.

Als u deze e-mails niet meer wilt ontvangen, kunt u zich afmelden door hier klikken of door te schrijven naar 6130 W Flamingo Rd. Las Vegas NV 81103

Figure F.2: Fake FedEx email.



Van Facebook
Aan leetjezhang@hotmail.com
Antwoord naar contact@licenses-keys.com
Onderwerp **Someone tried to loq in To Your Account, User ID : leetjezhanq**

14-1-2023 21:51

Facebook

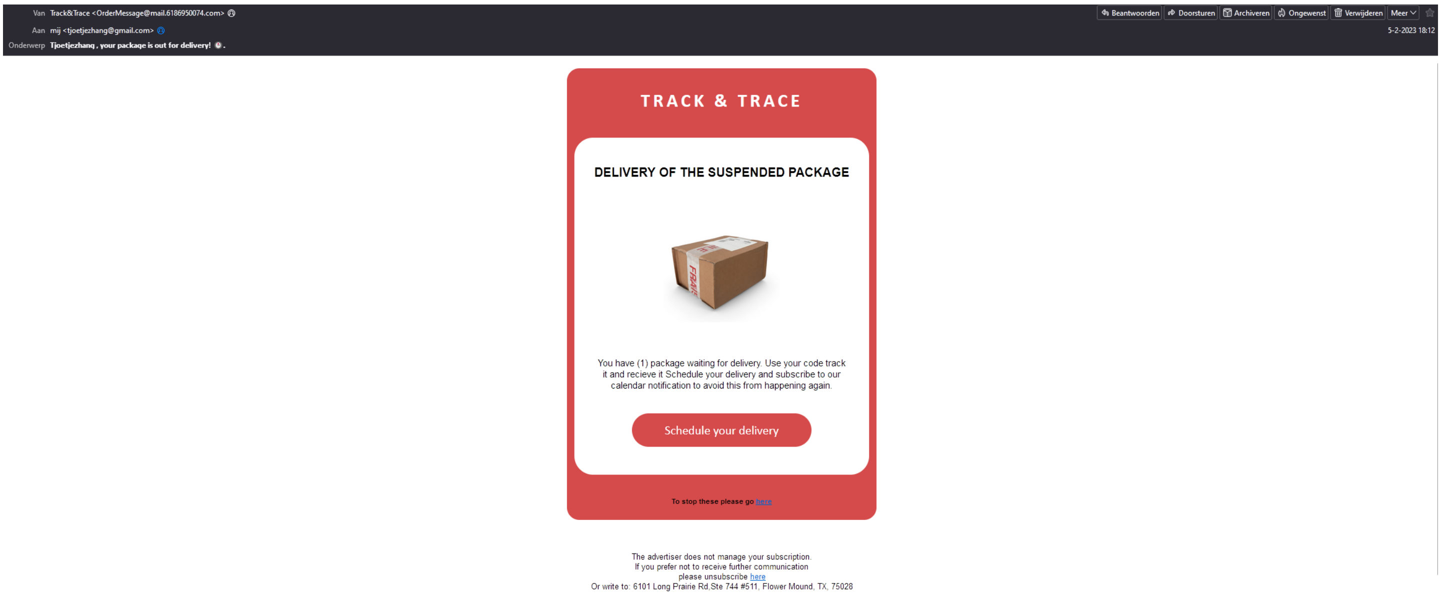
Hi leetjezhang,

A user just logged into your Facebook account from a new device Samsung S21. We are sending you this email to verify it's really you.

Thanks,
The Facebook Team

This message was sent to leetjezhang@hotmail.com. If you don't want to receive these emails from Facebook in the future, please unsubscribe.
Facebook, Inc., Attention: Community Support, 1 Facebook Way, Menlo Park, CA 94025
To help keep your account secure, please don't forward this email. Learn more.

Figure F.3: Fake Facebook email.



Van Track&Trace <OrderMessage@mail.6186950074.com>

Aan mij <tjoetjezhang@gmail.com>

Onderwerp **Tjoetjezhang , your packaqe is out for delivery! .**

5-2-2023 18:12

TRACK & TRACE

DELIVERY OF THE SUSPENDED PACKAGE

You have (1) package waiting for delivery. Use your code track it and recieve it Schedule your delivery and subscribe to our calendar notification to avoid this from happening again.

To stop these please go here

The advertiser does not manage your subscription.
If you prefer not to receive further communication
please unsubscribe here
Or write to: 6101 Long Prairie Rd,Ste 744 #511, Flower Mound, TX, 75028

Figure F.4: Fake delivery email