A Framework to Evaluate the Phishing Exposure in a Corporate Environment

Paulo Alexandre Souza Rupeiro

Mestrado em Segurança Informática

Trabalho de Projeto orientado por:
Prof. Doutor Mário João Barata Calha

2023
Acknowledgments

Firstly, I would like to thank my family for everything they have done and continue to do for me, for the education they gave me and for the unconditional love and support that always led me to become a better person and professional. A special thanks to my father Paulo, for guiding me through life and making sure I was able to get to where I am today, but also to my stepmother Christiane, brother Bernardo, sister Larissa and grandmother Maria Augusta for their support and company. I also want to give a big thank you to my grandfather Francisco for teaching me so much in his last years about life and what really matters in it. I hope you are proud up there.

Not less important, a very special thanks to my dear friends from Cascais and FCUL for their friendship and for always being there to help and contribute with their support and countless discussions. Thank you André Ginja for being such a good friend and uber, for all the memories, nicknames and music hits since 2016. Thank you Miguel Chaves for being the biggest bully but always with good intentions and motivating me to improve my health and lifestyle. Finally, thank you to João Caseirito and Diogo Avelãs for being my companions and groupmates since our undergraduate days.

I would like to express my gratitude towards Professor Mário Calha for being my counsellor, guiding me and contributing to this work’s success.

Thank you to EY’s Cybersecurity team for an amazing 3 years, especially to Ricardo Dionísio for helping and guiding me throughout this work, to Sergio Sá and Sérgio Martins for giving me the opportunity to work at EY, André Reves Pereira for being more than just the OG counsellor, Bruno Lourenço and Francesco Elisa for the countless hours of company. In addition, thank you to EY’s Risk team where I met some amazing people.

Lastly, I would like to thank myself, for not giving up and pushing through to finish this chapter in my life. As someone once said: “There are no mistakes in life, only lessons. There is no such thing as a negative experience, only opportunities to grow, learn and advance along the road of self-mastery. From struggle comes strength. Even pain can be a wonderful teacher.”
Resumo

Ao longo dos anos o *hacking* tem vindo a evoluir, assim como as suas diferentes técnicas, e o *phishing* não é um caso diferente. Este tipo de ataque, que tem como chave a base na *engenharia social* aplicada nos emails enviados para as vítimas, tem vindo a ser cada vez mais utilizado mas, com o passar do tempo, os atacantes têm também optado por uma abordagem mais direcionada em vez da, outrora mais habitual, utilização do spam. A abordagem mais clássica consiste em construir um email único com um objetivo (recolha de credenciais, disseminação de malware, etc) e enviar o mesmo para o máximo de pessoas possíveis na esperança que alguém seja ludibriado a realizar a operação que o atacante pretende. Por outro lado, a abordagem mais direcionada, que é denominada por *spear-phishing*, consiste em estudar previamente o ambiente em que irá ser realizado o ataque, e as possíveis vítimas de modo a escolher uma ou diversas para que sejam feitos ataques personalizados para cada uma desses indivíduos. Além desta técnica mais direcionada, os atacantes têm também utilizado novas abordagens acabando por ser variações do clássico phishing baseado no envio de emails: *smishing* - envio de mensagens de texto (SMS); *vishing* - chamadas telefónicas.

Com o crescimento e evolução do *hacking*, tem crescido também a preocupação mundial sobre o tema e a importância da cibersegurança para todo o tipo de negócios, sendo considerada indispensável para o dia-a-dia de todos. Dado esta informação, a sensibilização dos colaboradores é essencial. Quanto mais sensibilizados e atentos estejam, menor a probabilidade de serem ludibriados a realizar uma operação não desejada ao abrir um email ou uma mensagem de texto.

O aparecimento do vírus SARS-CoV-2 no início de 2020, foi um grande teste para todas as organizações do mundo e que elevou ainda mais a importância da cibersegurança. De um momento para o outro, toda a população mundial foi enviada para casa, o que significou que milhões de colaboradores de distintas organizações começaram a ter de trabalhar a partir das suas casas. O impacto que esta alteração teve foi enorme pois a maioria da população não tinha qualquer treino ou preparação para, de um momento para o outro, fazer uso de equipamentos da sua organização em sua casa ou até dispositivos pessoais para realizar as suas tarefas no dia-a-dia de trabalho.

O facto de que uma organização é tão segura como a sua camada mais fraca, que neste caso é o elemento humano, é algo que foi sendo revelado ao longo do tempo. Os ataques de engenharia social exploram vulnerabilidades humanas, da forma como pensamos, como processamos informação, como tomamos decisões e como gerimos as nossas emoções. Toda a população mundial está vulnerável a este tipo de ataques dado que as vulnerabilidades humanas são algo com que
todos nós já nascemos e que nunca vão desaparecer pelo simples facto de estarem associadas a algo que todos temos: cérebro, pensamento, subconsciente.

O psicólogo Daniel Kahneman explica que o ser humano tem dois modos de pensamento, o Sistema 1 (modo rápido) e o Sistema 2 (modo lento). O Sistema 1 está associado a reações emocionais automáticas, operando de forma intuitiva e quase imediata com pouco esforço ou controlo voluntário. O Sistema 2 está associado a operações de escolha, foco e racionais, operando de forma mais lenta e metódica. O que acontece então no cérebro humano é que estes dois sistemas não funcionam de forma isolada um do outro, o Sistema 1 pode mudar o foco do Sistema 2 e vice versa.

Os atacantes utilizam princípios de persuasão (Principles of Persuasion in Social Engineering - PPSE) para explorar vulnerabilidades humanas associadas ao funcionamento destes dois sistemas. O objetivo é que as vítimas achem com base no Sistema 1, de forma emocional, automática e instintiva para que sejam ludibriadas a realizar uma operação não desejada, seja fornecer informação confidencial ou acessos para, por exemplo, uma rede interna da sua organização.

Para combater o aumento deste tipo de ataques e mitigar o risco associado, as empresas necessitam de testar os seus colaboradores e sensibilizá-los/treiná-los. Para tal são necessárias ferramentas que ofereçam essas capacidades, que por norma são bastante dispendiosas, o que faz com que pequenas e médias empresas com orçamentos mais reduzidos não tenham meios para suportar os custos. Assim, estas muitas empresas acabam por optar e adoptar ferramentas de código aberto ou open-source como Gophish, Social-Engineering Toolkit ou Simple Phishing Toolkit. Apesar das mesmas oferecerem algumas funcionalidades para a realização de simulações de campanhas de phishing, existem sempre alguns passos para os quais a ferramenta não está capacitada para ajudar a realizar, seja o passo de configurar um servidor de emails, um SMTP relay ou um reverse proxy para aumentar a probabilidade de eficácia da simulação.

A Ernst & Young, agora conhecida como EY, é uma multinacional que oferece serviços de Consultoria, Auditoria, Fiscalidade e Estratégia e Transacção para capitalizar e crescer o negócio dos seus clientes. No contexto de um projeto mais alargado, um cliente da indústria IT requisitou os serviços da EY para aferir a eficácia do programa de sensibilização em cibersegurança da empresa. De realçar o facto do cliente ter um orçamento reduzido para este projeto paralelo, tendo sido requisitada a utilização de ferramentas open-source durante o trabalho.

Tendo todos estes fatores em conta, este trabalho visa estudar o quão vulnerável uma organização pode ser a este tipo de ataques, dependendo da sensibilização, preparação e treino de cada um dos seus colaboradores. A base do estudo foram duas campanhas de phishing com o objetivo de simular ataques de phishing dentro do ambiente empresarial do cliente da EY. O planeamento das campanhas tinha previamente estipulado um espaçamento temporal entre a primeira e a segunda, pretendendo realizar uma comparação entre os resultados de ambas para entender a evolução dos colaboradores, positiva ou negativa.

Durante o estudo, foi proposta uma framework constituída por seis elementos/ferramentas open-source e que permitiram que as campanhas fossem planeadas, desenhadas e realizadas. Esta framework permite que qualquer organização, caso tenha um orçamento alargado ou reduzido,
consiga e seja mais autônoma a realizar testes de segurança e sensibilização para este tipo de ataques. Adicionalmente, foi estudada a possibilidade de integrar uma ferramenta que permitisse o envio de mensagens de texto (Twilio) à framework proposta, adicionando assim a funcionalidade de realizar simulações de campanhas de smishing e ao mesmo tempo tirar partido de todas as outras funcionalidades que a framework já oferecia. Para tal foi modificado o código de programação base da ferramenta escolhida para a criação das campanhas (Gophish) para permitir a inserção de um novo dado, o número de telemóvel das vítimas, assim como criado um script na linguagem de programação Python que recolhe a informação das vítimas da ferramenta Gophish, constrói e envia a mensagem de texto, tudo via APIs de ambas as ferramentas mencionadas. De realçar que esta integração foi meramente experimental e não foi realizada nenhuma simulação no ambiente empresarial do cliente.

Por fim, foi realizada uma avaliação e comparação dos resultados obtidos ao longo do trabalho. Dado a ocorrência de alguns problemas aquando da realização da segunda campanha, esta comparação foi feita entre os resultados da primeira campanha e de duas previamente realizadas pela EY noutra cliente. Adicionalmente, foi também avaliado o trabalho em si, realçando o cumprimento do que havia sido proposto, assim como possibilidades de melhoria e trabalho futuro.

Palavras-chave: Phishing, Spear-phishing, Smishing, Engenharia social, Principles of Persuasion in Social Engineering
Over the years phishing has evolved and while the most classic approach consists of creating a single email and sending it to as many people as possible hoping that someone will be tricked into carrying out the operation that the attacker intends, the more targeted approach (spear-phishing), consists of previously studying the environment and possible victims in order to personalise attacks for each one of the victims. In addition to this more targeted technique, attackers have also used new approaches such as smishing - text messages.

Every human is vulnerable to these attacks since we are all born with “human vulnerabilities” because they are associated with something that we all have: brain, thought, subconscious. Attackers use Principles of Persuasion in Social Engineering to make victims act in an emotional and instinctive way being tricked into carrying out an unwanted operation.

During this study, a framework was proposed consisting of six open-source elements/tools that allowed campaigns to be planned, designed and carried out. Additionally, the possibility of integrating a tool that allows sending text messages (Twilio) to the proposed framework was studied, thus adding the functionality of carrying out smishing campaign simulations and at the same time taking advantage of all the other features that the framework already has. To this end, the base programming code of the tool chosen for creating the campaigns (Gophish) was modified to allow the insertion of new data, the victims’ mobile phone number, as well as a script created in the Python programming language that collects the information of victims of the Gophish tool, constructs and sends the text message, all via the APIs of both mentioned tools. It should be noted that this integration was merely experimental and no simulation was carried out in the client’s business environment.

**Keywords:** Phishing, Spear-phishing, Smishing, Social engineering, Principles of Persuasion in Social Engineering
# Contents

List of Figures xii

List of Tables xiii

List of Acronyms xiii

1 Introduction 1

1.1 Motivation 2

1.1.1 COVID-19 Pandemic 4

1.1.2 Prevention/Mitigation 4

1.1.3 Organizational Context 4

1.2 Goals 5

1.3 Contributions 5

1.4 Document structure 6

2 Concepts & Related Work 7

2.1 Social Engineering 8

2.2 Phishing 10

2.2.1 Current State - Attack 11

2.2.2 COVID-19 Pandemic 13

2.2.3 Models & Methodologies 13

2.2.4 Current State - Defense 23

2.2.5 The Human Element 26

2.2.6 Chapter overview 30

3 Proposed framework & campaigns 33

3.1 Work requirements 33

3.2 Use cases 34

3.3 Proposed framework 34

3.3.1 Framework elements 34

3.3.2 Framework design and Campaign flow 39

3.3.3 Framework Operations 42
### 3.4 Phishing Campaigns
- 3.4.1 Campaign I .................................................. 44
- 3.4.2 Campaign II .................................................. 49

### 3.5 Smishing integration in proposed framework
- 3.5.1 Twilio tool & Gophish modification .................... 55
- 3.5.2 Chapter overview ............................................ 59

### 4 Results & Discussion
- 4.1 Campaign I .................................................... 61
- 4.2 Campaign II .................................................... 64
  - 4.2.1 Two previously conducted campaigns ................... 64
- 4.3 Discussion ..................................................... 66
  - 4.3.1 Chapter overview ........................................ 68

### 5 Conclusion & Future Work
- 5.1 Conclusion ..................................................... 69
- 5.2 Future Work .................................................. 70

Bibliography .......................... 80

A Smishing integration (Gophish + Twilio) ................. 81

B Phishing study assessment/feedback questionnaire ....... 89
## List of Figures

2.1 The generic phishing attack overview, adapted from [61] .......................... 11
2.2 Lockheed Martin’s Cyber Kill Chain, adapted from [41] .......................... 13
2.3 Overview of the development of the Unified Kill Chain, adapted from [70] and comparison with Cyber Kill Chain and MITRE ATT&CK .......................... 17
2.4 Targeted phishing campaign flow related to Cyber Kill Chain’s phases, adapted from [67] ................................................................. 17
2.5 OSINT Framework .............................................................................. 19
2.6 Components of the anti-phishing ecosystem, adapted from [61] .......... 24

3.1 Framework design, connection and integration between elements (black arrows) and campaign’s high-level flow (green arrows) .......................... 40
3.2 Framework’s (presented in Section 3.3) operations .............................. 43
3.3 Phishing email with PPSE elements marked in red .............................. 46
3.4 Created landing page for data theft purposes ....................................... 47
3.5 Email preview on the target’s inbox .................................................... 48
3.6 Opened phishing email from the perspective of the victim ...................... 48
3.7 Landing page URL with proof of a valid certificate .............................. 48
3.8 Five steps required to launch a campaign in Gophish ............................ 49
3.9 Original email notifying employee of file sharing .................................. 50
3.10 Phishing email with PPSE element marked in red .............................. 51
3.11 Microsoft’s login page (original) ....................................................... 52
3.12 Cloned and adapted landing page for data theft purposes .................... 52
3.13 Email preview on the target’s inbox .................................................... 53
3.14 Opened phishing email from the perspective of the victim ...................... 54
3.15 Landing page URL with proof of a valid certificate .............................. 54
3.16 Modified Gophish “New Group” modal with user already added ............ 57
3.17 Smishing python script output ............................................................ 57
3.18 SMS Smishing message example ....................................................... 58
3.19 Gophish dashboard after smishing campaign launch ............................ 58
3.20 Gophish dashboard’s details ............................................................. 59

4.1 Results from Campaign I (presented in Section 3.4) .............................. 62
List of Tables

2.1 Information gathered in reconnaissance phase, from [64] .......................... 18
2.2 The five principles of persuasion in social engineering (PPSE), from [25] .... 28
2.3 Most used PPSE in phishing emails, from [25] ........................................ 29
2.4 Criteria to associate the most effective elements to corresponding PPSE, from [25] 30

3.1 Functional and non-functional work requirements, from both the client and EY . 33
3.2 Framework/toolkit comparison, adapted from [67]. X - full support, O - partial support, blanks - no support ................................................................. 36
3.3 MTA software comparison ................................................................. 37
3.4 Connection between requirements (presented in Section 3.1) and the constituent elements (presented in Section 3.3) of the framework that helped to meet the work’s requirements ................................................................. 42
3.5 Association between the email’s elements (Figure 3.3), the corresponding PPSE (Table 2.2) as well as the criteria (Table 2.4) ................................................. 46
3.6 DNS records table for Campaign I .................................................. 47
3.7 Association between the email’s element (Figure 3.10), the corresponding PPSE (Table 2.2) as well as the criteria (Table 2.4) ................................................. 51
3.8 DNS records table for Campaign II ................................................ 53
List of Acronyms

ACME - Automatic Certificate Management Environment
AOL - America Online
ASCII - American Standard Code for Information Interchange
AUTH - Authority
CDN - Content Delivery Network
CISO - Chief Information Security Officer
CKC - Cyber Kil Chain
CNAME - Canonical Name
CNCS - Centro Nacional de Cibersegurança
CPU - Central Processing Unit
CRC - Commitment, Reciprocation & Consistency
DDoS - Distributed Denial of Service
DIS - Distraction
DKIM - DomainKeys Identified Mail
DMARC - Domain-based Message Authentication, Reporting, and Conformance
DNS - Domain Name System
DoS - Denial of Service
ESP - Email Service Provider
EY - Ernst and Young
FBI - Federal Bureau of Investigation
FCUL - Faculdade de Ciências da Universidade de Lisboa
GUI - Graphical User Interface
HTTPS - Hypertext Transfer Protocol Secure
ID - Identity
IP - Internet Protocol
ISP - Internet Service Provider
IT - Information Technology
LSD - Liking, Similarity & Deception
MITM - Man In The Middle
MTA - Mail Transfer Agent
MX - Mail Exchange
NIST - National Institute of Standards and Technology
PPSE - Principle of Persuasion in Social Engineering
PTR - Pointer
RFID - Radio Frequency Identification
SMTP - Simple Mail Transfer Protocol
SP - Social Proof
SMS - Short Message Service
SPF - Sender Policy Framework
SSL - Secure Sockets Layer
TTPs - Tactics, Techniques and Procedures
TXT - Text
UI - User Interface
UKC - Unified Kill Chain
URL - Uniform Resource Locator
WFH - Work from home
Chapter 1

Introduction

Throughout the years, the internet has gained a lot of importance and become indispensable to everyone’s daily lives [37]. With the number of users growing exponentially over the years and the consequent need to increase/improve the associated infrastructures, concerns arise regarding the security of its users and their data [36]. A standard dictionary definition of security would be the state of being free from danger or threat, however, there is no such thing as being 100% secure on the internet.

Cybersecurity is a growing concern for businesses of all sizes as cyber criminals’ increasingly sophisticated tactics continue to disrupt organizations. In the past decade, we have seen that company’s demand for cybersecurity and cyber-risk measurements had an explosion [18]. This shows that organizations are reacting to the latest cybersecurity global incidents and investing in cybersecurity, trying to prevent and mitigate those risks.

Gartner Insight projected that businesses would spend more than $123 billion on security in 2020 and projects that figure to grow to $170.4 billion by 2022 [10] [29]. However, cybercriminals still manage to compromise corporate data and systems with relative ease and on a regular basis [29] [40]. At the same time, although most companies today have cybersecurity as one of their concerns, it is often not high on their list of priorities.

Misguided notions can lead companies, without much knowledge in cybersecurity, to think that because they have not yet been victims of a cyberattack they do not need to invest in cybersecurity at that very moment, but instead wait for something to happen. This is also common among individuals regarding their personal data. Some people will say that since they never had their identity stolen, their email compromised, or a problem with their computer, therefore, they must be safe.

In [82], Stallings and Brown present some security concepts - Owners, Assets, Countermeasures, Risk, Threat agents, Threats - and show the relations between them. We have two sides, on the owners’ side we may have an organization or a business that owns and values some assets, which may be infrastructures, data, systems or even employees. The company’s goal is to minimize its risks and to do so they impose countermeasures. On the nefarious side, we have the adversary that wants to exploit or damage the company’s assets, so they try to discover threats which can give access to that exploitation.
One thing we have to take into account is the slight difference between cybersecurity and information security. Solms and van Niekerk [90] tell us that these two concepts are not totally analogous. They posit that cybersecurity goes beyond the boundaries of traditional information security to include not only the protection of information resources but also other assets, including the human element.

Considering this we can say that a secure company is one that actively attempts to detect and prevent incoming attacks implementing and imposing countermeasures such as backing up data and systems regularly and redundantly, training and informing staff monthly to make them aware of new threats and reinforcing best practices. However, to know how to be secure and what defences to implement, it is crucial to start by understanding the attacker’s mind. Cybersecurity expert and New York Times bestselling author, Brian Krebs, compiled a list of immutable truths about data breaches [47]. Some of them are:

- If you connect to the Internet, someone will try to hack it;
- If what you put on the Internet has value, someone will invest time and effort to steal it;
- Even if what is stolen does not have immediate value to the thief, he can easily find buyers for it.

With these truths in mind, attackers’ methods usually follow two primary courses of action, either broad or targeted. Targeting a broad range of victims is quite common and likely to have at least one victim. On the other hand, targeted attacks are less common but an approach for high-value targets. These typically do not target large networks or systems, but instead, they try to manipulate people through social engineering.

### 1.1 Motivation

Recently, there has been a significant change in social engineering attacks, with the predominance of the phishing attack type changing from massive spam emails (phishing) to targeted email phishing campaigns (spear-phishing). Furthermore, the number of smishing attacks, phishing attacks through SMS, has increased in recent times [91]. In all these attacks we can observe the usage of PPSE \(^1\) to try to deceive the potential victim into doing something they don’t want to do.

These attacks have been causing damage and a significant impact on worldwide organizations, both financially, operationally, and even at reputational levels. In their latest phishing statistics study [14], Security Boulevard shows that the trends have become more alarming as phishing attacks have become even more intelligent and creative. Here we have some of the most disturbing highlights from the report:

- 32% of confirmed data breaches involved phishing - nearly one-third of all data breaches involved phishing;

---

\(^1\)Principles of Persuasion in Social Engineering
• Overall phishing down 42% in 2019 - with this data, it may seem that the threat is going away but what is important to emphasize is that cybercriminals are targeting more their attacks, that is, opting for a quality-over-quantity approach. As they say, “Phishers are being measured, pragmatic and patient.”;

• 37.9% of untrained users fail phishing tests - this is a very alarming statistic. As KnowBe4, one of the industry’s leading cyber awareness training organizations, states, users who do not attend cyber awareness training fail phishing tests;

• $26 billion lost globally to BEC/EAC^2 Crimes - between June 2016 and July 2019, 166,346 global victims reported financial losses. In 2019 alone, these losses amounted to $1.7 Billion;

• Nearly half of data breaches are due to human error and glitches - being most of the human part due to “inadvertent insiders” who fall prey to phishing attacks. In 2019 a human error data breach averaged a cost of $3.5 Million;

• 88% of organizations reported experiencing spear-phishing attacks in 2019 - as stated before, this shows that phishers are opting for quality over quantity;

• In one week, Google blocked more than 18 million COVID-19 phishing emails daily - phishers are using COVID-19 concerns to exploit vulnerable individuals seeking information;

• According to Kaspersky [56], their anti-Phishing system was triggered 467,188,119 times. In 2019, Portugal ranked No.4 with 25.63% of users on whose computers the Anti-Phishing system was triggered out of all Kaspersky users in the country. Looking at the first two quarters of 2020 [84] [85], we have Portugal ranked No.4 on the first one and ranked No.2 on the second.

Cybercriminals are constantly making efforts to discover new tactics and techniques and changing the way they attack [84] [85] so it is essential to look beyond the numbers on these statistics. Phishing is still a huge global problem and substantially impacts businesses and organizations [14]. It is a critical threat vector against untrained and unaware employees [14], so organizations must be able to fight against it, either by developing new frameworks/methodologies to prevent, mitigate and respond to attacks or by increasing awareness of users/employees with workshops and disseminating information, trying to instil good practices and behaviours.

Additionally, with the growing number of mobile devices in the world [65], smishing attacks have become increasingly common and a method which attackers adopt. This has led to the topic being mentioned more often, raising the levels of concern about them. As users get more aware by constant suspicious emails, text messages have become a more attractive attack vector exploiting the fact that personal devices generally lack the type of security available on corporate devices. Given the proliferation of smartphones, these social engineering attacks on mobile devices are

^2Business Email Compromise/Email Account Compromise
now more prevalent [91] and consequently, lately, financial losses due to these types of attacks have been increasingly reported [23].

One of the reasons why the frequency of this type of attack has increased is due to people not being educated about and not being aware of it. Proofpoint reports that SMS-based scams increased by more than 300% in 2020 alone and that, on average, only 31% of company employees answer correctly (25% answer incorrectly and 44% don’t know) when asked what smishing is [73].

As CNCS\textsuperscript{3} state, although the volume of complaints to the authorities may be small, phishing and smishing attacks are the types of incidents most recorded by CERT.PT [52], and although they’re not a new phenomenon, the situation has been aggravated by the Covid-19 pandemic and the successive confinements.

1.1.1 COVID-19 Pandemic

The rise of the SARS-CoV-2 virus, at the beginning of 2020, further tested the level of training and best practices of the most diverse companies regarding cybersecurity. Consequently and without any warning, preparation or training, hundreds of thousands of people started working from home, some depending on their internet connection. This meant that some people started using the organization’s equipment at home or doing work-related tasks on their personal equipment. Pranggono and Arabo [71] state that many organizations were forced to revamp their physical offices and policies to enable employees to work from home without the necessary training or well-prepared arrangements, while only 38% businesses had a cybersecurity policy in place.

1.1.2 Prevention/Mitigation

Nowadays, there are several possible mitigation and prevention techniques for phishing. Some organizations invest in dedicated software (commercial or open-source solutions), in carrying out awareness campaigns, training their employees or hiring a team specialized in the subject. Something also quite common in the business world is the realization of phishing campaigns within the company itself, this being the chosen technique to carry out this work. Raising awareness and training the company’s staff are two essential mitigation techniques, as it is known that in phishing, spear-phishing or smishing attacks, the organization’s employees are the target. It is therefore imperative that these are two of the techniques that should be used by all organizations that have the objective of mitigating these types of attacks.

1.1.3 Organizational Context

This work realization is part of the Dissertação/Projeto em Segurança Informática course, which is a requirement for the Mestrado em Segurança Informática’s completion, at Faculdade de Ciências da Universidade de Lisboa (FCUL). The work was developed in an external organization, in this case a private company, Ernst & Young (EY).

\textsuperscript{3}Centro Nacional da Cibersegurança
EY is a multinational with one of the largest professional services networks in the world. It offers the expertise to capitalize and grow clients’ businesses through four service lines - assurance, consulting, strategy and transactions and tax advisory services.

In the context of a larger project, a client belonging to the IT industry requested EY’s consulting services to assess the effectiveness of the organization’s cybersecurity awareness program. For this purpose two phishing campaigns were requested by the client, a generic phishing campaign and a spear-phishing one. Later on, the possibility of carrying out smishing campaigns, to complement those previously carried out, was also discussed.

Additionally, during the project discussions, both with EY and the client, needs and requirements were raised, leading to the suggestion of developing/designing an environment/framework with the ability to integrate a selection of open-source tools/solutions that would allow performing phishing campaign simulations while collecting relevant metrics. This suggestion was made internally to EY, as an approach to the project and requirements presented by the client. Furthermore, having been motivated during the research on the topic and by the non-existence, in the analysed literature, of a solution "as open-source as possible” that by integrating different tools would obtain a better result both in terms of campaign design and deliverability.

1.2 Goals

This work aims to raise awareness for phishing, spear-phishing and smishing attacks within the corporate world. For this purpose, two phishing campaigns will be conducted, integrating the usage of PPSE in each template, with the collection of relevant data and metrics. Later on, the results will be analysed and discussed.

Specifying, this work will have the following objectives:

- Design/develop a framework that meets the need to integrate different open-source elements/tools allowing to conduct phishing campaigns and collect relevant metrics;

- Conduct two phishing campaigns with the incorporation of PPSE. Analyse and compare the campaign’s results, giving insight of: what type of phishing is the most effective; the great impact that PPSE usage in phishing campaigns has; and the level of risk to an organization, taking into account the campaigns’ success rates.

- Test the integration of a platform, able to send SMS, in the proposed framework adding the capability to conduct and collect relevant metrics of smishing campaigns.

1.3 Contributions

The contributions of the developed work are:

- A phishing simulation framework was developed and adopted by EY;

4Due to privacy reasons the client’s name will remain confidential.
• A script was developed integrating an existing solution/tool, with the ability to send SMS, into the proposed framework, to be able to launch smishing campaigns and also collect relevant metrics;

• Raising awareness for these types of attacks and for their success rate in the corporate world, with the support of the conducted campaigns’ results and received feedback.

1.4 Document structure

This document is structured as follows:

• Chapter 1 – Introduction. Presents the work’s motivation, objectives, contributions, context and document structure.

• Chapter 2 – Concepts & Related Work. Presents all the research made in the initial part of the project, as well as the concepts and the studies/works that served as a basis for it.

• Chapter 3 – Proposed framework & campaigns. Presents in detail the work, starting with its requirements and use cases, then presenting the proposed framework and finally two phishing campaign simulations. Furthermore, it also presents a tested integration of a cloud communications platform with the proposed framework to add the possibility of conducting smishing campaigns.

• Chapter 4 – Results & Discussion. Presents the results obtained throughout the project and a discussion of it.

• Chapter 5 – Conclusion & Future Work. Presents a conclusion of all work as well as possibilities for improvement and future work.
Chapter 2
Concepts & Related Work

This chapter will introduce the hacking topic, which is based on discovering and exploring vulnerabilities to exploit them later. Essential concepts such as vulnerability, attacks, and their types will be presented. Then we will make the transition to a specific type of attack, which is social engineering. This subsection will present basic and critical concepts to understand this type of attack and also what type of vulnerabilities it exploits as well as the several techniques that are used for it.

Finally, we will enter the phishing theme, one of the most used types of social engineering. In this topic, we will approach its current state and a fundamental issue in this type of attack: the human element, since it is the weakest layer of an organization and the greatest vulnerability to exploit in these cases. We will present the methodologies and related work that served as a basis for this thesis, touching on topics such as the phases of a cyber intrusion/phishing campaign and even see how the psychological part of humans impacts phishing, addressing principles of persuasion, and trying to understand which are the most effective for victims to fall into attacks.

Looking into a dictionary we will see that hacking is perceived as *the activity of using computers to get access to data in somebody else’s computer or phone system without permission* [7], and traditional hacking refers to that, attacks that focus more on exploiting technology rather than exploiting the people that use it. However, hacking consists in, the majority of times, trying to uncover vulnerabilities in a system or an organization in general.

NIST defines vulnerability as a *weakness in an information system, system security procedures, internal controls, or implementation that could be exploited or triggered by a threat source* [60], and there are two types of vulnerabilities, technical and human. In the technical realm, we have vulnerabilities related to computers, systems, software, etc. It may be a weakness in a firewall rule that lets an adversary get inside the organization’s network or software that is vulnerable to attacks that may disrupt the system. In the human realm, we have vulnerabilities that essentially derive from emotions, behaviours and instincts. An organization is as secure as its weakest element and, in information security, that is the human being.

To successfully perform an intrusion, threat actors have to attack those existent vulnerabilities

7
(Attack + Vulnerability = Intrusion). An attack is a threat that is carried out and, if successful, leads to a consequence [82]. We can separate them into two different types [82]: active attacks are an attempt to alter system resources or affect their operation. On the other hand, passive attacks are an attempt to learn or make use of information from the system that does not affect system resources.

2.1 Social Engineering

While traditional hacking focuses more on software and hardware exploitation, social engineering focuses more on the user and has them act as unwitting accomplices. It is the art of making people do things they normally would not do. These attacks exploit the social nature of human beings, the way they think and process pieces of information, to achieve an end goal. The attacker manipulates the victim, psychologically, using human reactions, people’s inattention and unawareness, people’s fear, anxiety and curiosity.

In the context of information security, social engineering is defined as the use of deception to manipulate individuals into divulging confidential or personal information [39].

The basic goals of social engineering are the same as hacking in general, to gain unauthorized access to systems or collect information to commit fraud, network intrusion, industrial espionage, identity theft or to simply disrupt the system or network. Furthermore, there are different types of social engineering attacks [48], which can be divided into groups such as:

• Physical - in this type of attack, the attacker tries to gather information by being physically present in the social engineering environment. If it is possible, they also observe the victims using their persuasion skills or some tools to collect data.

• By phone - the attacker makes contact with the victim and usually tries to collect some critical information about them. The attacker may try to gain access to the victim’s accounts or bank information by pretending to be an employee in that organization and saying that there is a security verification needed. In addition he can also try to deceive the victim into visiting a malicious website or installing a malicious piece of software.

• Aidedd - in most cases this type of social engineering attack is performed as a phishing attack. It is typically carried out by email spoofs or instant messaging and often directs users to enter personal information at a fake website. The look and feel of that website may be identical to the legitimate one and the only difference present in the URL1 of the visited website.

Every social engineering attack has different stages, which are usually: Research; Develop Trust; Exploiting Trust and Utilise Information [58] [59] [76]. However, these steps are quite similar to those performed during any other type of cyber attack [41], and there is almost a direct correlation between them:

1Uniform Resource Locator
1. Reconnaissance (Research) - here the attacker tries to collect everything which will help to perform a successful attack. They try to collect information about the people who are related to the target company (e.g., employees, employers, shareholders, etc); security measures of the company, so the threat actor can look for a way in and evasion techniques; commonly used internet browsers and its version. Knowing this they can prepare malicious websites to work with that specific browser and version; programs and their versions used inside the organization (e.g., versions of the Java runtime environment, PDF reader, Office tools, etc); and companies’ sensitive data. If the attackers know, for example, the name of the director of the human resources department, they can prepare a phishing email as if it was sent by the director.

2. Weaponization/Exploitation (Develop Trust & Exploiting Trust) - This step is the attack time, the attacker already collected all necessary information and is ready to exploit the discovered vulnerabilities. Here the nefarious actor prepares the websites to use the vulnerabilities of the browser widely used in the target company and forces the victim to visit it. They may also prepare malware, send it as an attachment to a phishing email and force the victim to open it.

3. Post-exploitation (Utilise Information) - the actions of this step are to see how far the attacker can go inside the company. Here, they try to escalate privileges to gain access to the most sensitive systems and information. Furthermore, the attackers also may create back doors for further use.

When it comes to exploiting human vulnerabilities, there are several social engineering techniques [46], some of them are:

**Shoulder Surfing**

This is the most famous of them all. Shoulder surfing is the technique used to obtain confidential data such as PINs and passwords, simply by observing the information without getting the victim’s attention, for example, by looking over the victim’s shoulder. It can be performed either at close range by directly looking over the shoulder.

This technique is most likely to be more efficient when performed in crowded places because it is easy to absorb the information without getting the victim’s attention. It is often performed in public transportation vehicles, supermarkets, shopping centres, restaurants, etc.

**Dumpster diving**

Also known as *trashing* is another popular method of social engineering. It basically consists in looking for valuable things in someone else’s waste bin. A huge amount of information can be collected through company dumpsters, e.g., company phonebooks, organizational charts, company policy manuals, meeting calendars, events, system manuals, etc. All of this information can be
used to assist a social engineering attack to gain access to the target company’s network, for example.

**Tailgating**

The common definition of tailgating is to drive to closely behind another vehicle without giving it enough security distance but as a social engineering attack, tailgating is seeking entry to a restricted area secured by unattended access control, for example, by an RFID ² card. In this technique, the attacker simply walks in behind a person who has legitimate access. Following common courtesy (social behaviour), the legitimate person usually may hold the door open for the attacker or the attacker may even ask for the victim to hold it open for him. As stated before, the human being is the weakest element and in this case, the legitimate person can fail to ask for identification or may accept an assertion that the attacker has forgotten or lost his identification or the appropriate identity token.

These methods (social engineering) were used before the internet by spies and conmen but now have a technological twist for tricking people into giving over critical information or access regardless of the approach.

### 2.2 Phishing

The term “phishing”, according to internet records, was used for the very first time on January 2nd, 1996, although the practice originated in the early 1990’s [44]. A 17-year-old hacker, known online as “Da Chronic”, created a toolkit that provided several “hacker helping” features, most notably a function for stealing passwords of AOL users, and according to its creator, containing the first recorded mention of the term “phishing”. Therefore it is believed that it all started with this program named “AOHell”.

The first method “phishers” used to attack was stealing users’ passwords and using algorithms to create randomized credit card numbers [44]. When someone shut down their random credit card number generating game, phishers began sending messages to users through the AOL ³ instant messenger and email systems, while posing as the company’s employees [44]. Those messages would request them to confirm their billing information or to verify their accounts. People often fell for the deception as it was never witnessed anything like it.

Throughout the years, attackers turned their attention to online payment systems and spoofed websites [44]. The intention behind it was to send a “wormed” email to those payment system customers, which then led to a spoofed site asking, for instance, to update their credit card details or other sensitive information.

---

² Radio Frequency Identification  
³ America Online
2.2.1 Current State - Attack

The anatomy of a phishing attack consists of three major phases. First, the potential victims receive a phishing email/SMS. In the second step, the victims take the suggested action in the message. Lastly, the criminal monetizes stolen information [38]. But there are numerous types of phishing attacks and although not all are the same, they follow the same basis.

2.2.1.1 Generic Phishing

Oest et al. [61] present in Figure 2.1 the generic phishing attack overview. The classic phishing attack includes three key “moments”: the lure, the hook and the catch.

The Lure [44] consists of a phisher spamming a large number of users with an email message that typically, in a convincing way, appears to be from some legitimate institution that has a presence on the Internet. The message often uses a convincing story to encourage the user to follow a URL hyperlink encoded in the email to a website controlled by the phisher and to provide it with certain requested information. The social engineering aspect of the attack normally makes itself known in the lure, as the spam gives some legitimate sounding reason for the user to supply confidential information to the website that is hyperlinked by the spam.

The Hook [44] typically consists of a website that mimics the appearance and feel of that of a legitimate target institution. In particular, the site is designed to be as indistinguishable from the targets as possible. The purpose of the hook is for victims to be directed to it via the lure portion of the attack and for the victims to disclose confidential information to the site. Examples of the type of confidential information that is often harvested include: usernames, passwords, social-security numbers in the U.S. (or other national ID numbers), billing addresses, checking account numbers, and credit-card numbers. The hook website is generally designed both to convince the victim of its legitimacy and to encourage the victim to provide confidential information to it with as little
suspicion on the victim’s part as possible.

*The Catch* [44] is the third portion of the phishing attack, which some alternatively call the *kill*. It involves the phisher or a cashier making use of the collected information for some nefarious purpose such as fraud or identity theft.

### 2.2.1.2 Spear-phishing

Spear-phishing is a targeted attack that is always part of a wider operation with clear goals, such as compromising the security of a company, plant or institution, or stealing information, intellectual property or money [6]. Instead of it being sent to thousands of possible victims, the attacker has to make a previous analysis and choose a specific possible victim and then build a customized attack only for that specific victim. Jakobsson and Myers [44] draw a parallel from the physical world and say that most current phishing attacks can be described with the following scenario: Somebody knocks on your door and says you have a problem with your phone; if you let him in, he will repair it. A spear-phishing attack, on the other hand, can be described as somebody who first cuts your phone lines as they enter your home, waits for you to contact the phone company to ask them to come and fix the problem, and then knocks on your door and says he is from the phone company.

This type of phishing has a better success rate and that is the reason why cybercriminals are targeting their attacks, that is, opting for a quality-over-quantity approach. As stated in Section 1.1, Security Boulevard shows some alarming statistics: 37.9% of untrained users fail phishing tests and 88% of organizations reported experiencing spear-phishing attacks in 2019 [14]. This demonstrates the success rate of targeted phishing attacks, possibly to those untrained people, and that this has been a widely used attack by nefarious actors.

There are also other types, like *whale phishing*, that consist in a spear-phishing attack targeted to certain individuals with certain positions inside an organization (e.g., CEO, CISO).

### 2.2.1.3 Smishing

Smishing is a variation of email-based phishing that involves the use of text messages sent using SMS[^4], that lead the victim to access a malicious link to submit credentials, hand over confidential information (e.g. credit card number) or even download malware [51][66][91]. When conducting a malicious smishing campaign, the attacker may intend to acquire any type of information, the following being the most common in the phishing realm [51][66][91]:

- Online account credentials;
- Private information that could be used in identity theft;
- Financial data that can be used to sell on darknet markets or for online fraud.

[^4]: Short Message Service
2.2.2 COVID-19 Pandemic

With the implementation of the WFH business model, which increased attack vectors and risks to internal data, this meant that cybercriminals had a bigger attack surface to explore. As Enisa stated in their 2020 phishing report [22], it has been reported that in 1 month, between the end of February 2020 and the end of March 2020, phishing attacks involving the SARS-CoV-2 virus increased by 667%. This was a major contribution to the fact that, in 2020, phishing incidents almost doubled in frequency, from 114,702 incidents in 2019 to 241,324 incidents in 2020, according to the FBI [23]. Furthermore, while Abnormal Security [1] states that COVID-19-related scams reached their peak in the third and fourth weeks of April, in Portugal the number of phishing campaigns had a massive growth between April and June [86][87].

The whole pandemic situation also led to a cyber pandemic which caused the demand for cybersecurity services and tools to explode. As Bruno Castro states [26], Visionware detected countless cyberattacks in 2020. Although one of the big trends was stealing money [1] through malicious campaigns that directly involved companies’ leading executives, some cybercriminals steal data instead of money [26]. This is called espionage, and their goal is to steal information from public organizations to sell on the black market, which may only be noticed months or even years later. S21Sec’s data shows that phishing attacks increased by 350% globally during the pandemic [27] and also smishing attacks [52].

2.2.3 Models & Methodologies

To achieve their goals through a cyber intrusion, an adversary must follow several steps. Cybersecurity literature presents and describes these steps in several models and methodologies:

2.2.3.1 Lockheed Martin’s Cyber Kill Chain

The Cyber Kill Chain (CKC) model was introduced by Hutchins et al. from Lockheed Martin in their whitepaper [41]. The researchers state that the essence of an intrusion is that “the aggressor must develop a payload to breach a trusted boundary, establish a presence inside a trusted environment, and from that presence, take actions towards their objectives, be they moving laterally inside the environment or violating the confidentiality, integrity, or availability of a system in the environment” [41].

![Lockheed Martin’s Cyber Kill Chain](figure2.2.png)

Figure 2.2: Lockheed Martin’s Cyber Kill Chain, adapted from [41]
The Cyber Kill Chain model was adapted from the concept of U.S. military targeting doctrine that defines a Kill Chain as a “systematic process to target and engage an adversary to create desired effects”. The Cyber Kill Chain (Figure 2.2) consists of 7 distinct phases, described by Hutchins et al. [41] as follows:

1. **Reconnaissance** - Research, identification and selection of targets, often represented as crawling Internet websites such as conference proceedings and mailing lists for email addresses, social relationships, or information on specific technologies.

2. **Weaponization** - Coupling a remote access trojan with an exploit into a deliverable payload, typically by means of an automated tool (weaponizer). Increasingly, client application data files such as Adobe Portable Document Format (PDF) or Microsoft Office documents serve as the weaponized deliverable.

3. **Delivery** - Transmission of the weapon to the targeted environment. The three most prevalent delivery vectors for weaponized payloads by APT actors, as observed by the Lockheed Martin Computer Incident Response Team (LM-CIRT) for the years 2004-2010, are email attachments, websites, and USB removable media.

4. **Exploitation** - After the weapon is delivered to the victim host, exploitation triggers intruders’ code. Most often, exploitation targets an application or operating system vulnerability, but it could also simply exploit the users themselves or leverage an operating system feature that auto-executes code.

5. **Installation** - Installation of a remote access trojan or backdoor on the victim system allows the adversary to maintain persistence inside the environment.

6. **Command and Control (C2)** - Typically, compromised hosts must beacon outbound to an Internet controller server to establish a C2 channel. APT malware especially requires manual interaction rather than conducting activity automatically. Once the C2 channel is established, intruders have “hands on the keyboard” access inside the target environment.

7. **Actions on Objectives** - Only now, after progressing through the first six phases, can intruders take actions to achieve their original objectives. Typically, this objective is data exfiltration which involves collecting, encrypting and extracting information from the victim environment; violations of data integrity or availability are potential objectives as well. Alternatively, the intruders may only desire access to the initial victim box for use as a hop point to compromise additional systems and move laterally inside the network.

### 2.2.3.2 MITRE ATT&CK

MITRE started ATT&CK in 2013 to document common tactics, techniques, and procedures (TTPs) that advanced persistent threats used against Windows enterprise networks. ATT&CK is largely a knowledge base of adversarial techniques — a breakdown and classification of offensively oriented
actions that can be used against particular platforms, such as Windows [4] [83]. The framework aims to address four main issues:

1. **Adversary behaviours** - Focusing on adversary tactics and techniques allowed MITRE to develop analytics to detect possible adversary behaviours.

2. **Lifecycle models that didn’t fit** - Existing adversary lifecycle and Cyber Kill Chain concepts were too high-level to relate behaviours to defences given that the level of abstraction wasn’t useful to map TTPs to new types of sensors.

3. **Applicability to real environments** - TTPs need to be based on observed incidents to show the work is applicable to real environments.

4. **Common taxonomy** - TTPs need to be comparable across different types of adversary groups using the same terminology.

ATT&CK became one of the go-to tools, both for the adversary emulation team to plan events and for the detection team to verify their progress, as it was released to the public in May 2015.

### 2.2.3.3 Unified Kill Chain

The Unified Kill Chain (UKC) was introduced by Paul Pols in their whitepaper [70]. It states that it was developed through a hybrid research approach, combining design science with qualitative research methods, extending and combining existing models, such as Lockheed Martin’s Cyber Kill Chain and MITRE’s ATT&CK. The model describes all phases in typical and modern cyber attacks, from the attacker’s first steps to the achievement of adversarial objectives in 18 steps, providing insight into their expected ordered arrangement [70]:

1. **Reconnaissance** - Researching, identifying and selecting targets using active or passive reconnaissance.

2. **Resource Development** - Preparatory activities aimed at setting up the infrastructure required for the attack.

3. **Delivery** - Techniques resulting in the transmission of a weaponized object to the targeted environment.

4. **Social Engineering** - Techniques aimed at the manipulation of people to perform unsafe actions.

5. **Exploitation** - Techniques to exploit vulnerabilities in systems that may, amongst others, result in code execution.

6. **Persistence** - Any access, action or change to a system that gives an attacker a persistent presence on the system.
7. **Defense** - Evasion Techniques an attacker may specifically use for evading detection or avoiding other defences.

8. **Command & Control** - Techniques that allow attackers to communicate with controlled systems within a target network.

9. **Pivoting** - Tunneling traffic through a controlled system to other systems that are not directly accessible.

10. **Discovery** - Techniques that allow an attacker to gain knowledge about a system and its network environment.

11. **Privilege Escalation** - The result of techniques that provide an attacker with higher permissions on a system or network.

12. **Execution** - Techniques that result in the execution of attacker-controlled code on a local or remote system.

13. **Credential Access** - Techniques resulting in the access of, or control over, system, service or domain credentials.

14. **Lateral Movement** - Techniques that enable an adversary to horizontally access and control other remote systems.

15. **Collection** - Techniques used to identify and gather data from a target network prior to exfiltration.

16. **Exfiltration** - Techniques that result or aid in an attacker removing data from a target network.

17. **Impact** - Techniques aimed at manipulating, interrupting or destroying the target system or data.

18. **Objectives** - Socio-technical objectives of an attack that are intended to achieve a strategic goal.

### 2.2.3.4 CKC vs UKC vs MITRE ATT&CK

Comparing the 3 models/frameworks it is possible to identify, in Figure 2.3, adapted from [70], that the Unified Kill Chain is more refined and complex than the others by mapping almost the same content through more steps.
Figure 2.3: Overview of the development of the Unified Kill Chain, adapted from [70] and comparison with Cyber Kill Chain and MITRE ATT&CK

Pirocca et al. [67] also present a targeted phishing campaign flow. As Figure 2.4 shows, we have six stages, which are similar to the Cyber Kill Chain’s first three phases (in red). Although the Unified Kill Chain is a more complete framework, given the work’s objectives and scope, the followed methodology to create and conduct the two phishing campaigns will be based on the Cyber Kill Chain, its first three phases and Pirocca et al.’s targeted phishing campaign flow given that together both cover and focus more on social engineering and the key stages regarding phishing and targeted phishing attacks.

Figure 2.4: Targeted phishing campaign flow related to Cyber Kill Chain’s phases, adapted from [67]
2.2.3.5 Reconnaissance

Reconnaissance is the first phase of a cyber-attack, where the attacker collects information on the possible victims and weak points are identified. Any information gathered about the victim (individual or organization) may be crucial and revealed as a critical weakness. In Table 2.1, Sanghvi and Dahiya [64] show the type of critical information that can be obtained during this phase:

<table>
<thead>
<tr>
<th>Type</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Information</td>
<td>IP addresses, Subnet mask, Network topology, Domain names</td>
</tr>
<tr>
<td>Host Information</td>
<td>User names, Group names, Architecture type (e.g., x86 v/s SPARC), Operating system family and version, TCP and UDP services running with versions</td>
</tr>
<tr>
<td>Human Information</td>
<td>Home address, Home telephone number, Frequent hangouts, Computer knowledge, Dark secrets</td>
</tr>
<tr>
<td>Security Policies</td>
<td>Password complexity requirements, Password change frequency, Expired/disabled account retention, Physical security (e.g., locks, ID badges, etc.), Firewalls, Intrusion detection systems</td>
</tr>
</tbody>
</table>

Table 2.1: Information gathered in reconnaissance phase, from [64]

To gather this information, attackers use two methods: Active Reconnaissance and Passive Reconnaissance [64].

Active reconnaissance is commonly referred to as scanning. Scanning involves actively sending network packets to the targeted organization and is more offensive and accurate than passive reconnaissance as there is more risk that the target may be alerted of a possible attack [64].

Active reconnaissance’s objective is to gather as much network and host information as possible (table 2.1). This is often done to obtain knowledge about the environment to develop targeted payloads to compromise the organization and may be done through automated scanning or manual testing using tools like ping, traceroute, netcat, nmap (most popular), snort, tcpdump, metasploit (although primarily designed as an exploitation toolkit) or even sending a benign phishing email that is engineered to trigger a call back through externally hosted images and to disclose information such as operating system version, internal domain, hostnames [75].

Passive reconnaissance, also known as footprinting, means gathering information without ac-
actively interacting with the target to avoid alerts. It is gathering the information without alerting the victim. If the victim’s host is alerted then it drastically increases security against the attack.

Activities during footprinting include harvesting email addresses, and employee names from social media, e.g., Linkedin\(^6\). One of the most known activities to perform footprinting is *Google Hacking*, which refers to the activity of using the Google search engine and some search operators to refine results. The Google search engine is extremely agile and aggressive in its indexing, which sometimes results in information being indexed without the knowledge of the target organization.

Although Google Hacking used to dominate the reconnaissance phase, its use has been declining [34], moving a large part of this phase-out of the Google website into specific tools/services. There are numerous tools to perform passive reconnaissance (Maltego, Shodan, etc) and one of the most known frameworks to use is called OSINT Framework\(^7\). Here any person can find tool names to discover usernames, email addresses, social networks, and IP addresses, as figure 2.5 shows.

![OSINT Framework](image)

**Figure 2.5: OSINT Framework**

After gathering some information, such as an email address, other tools can help disclose used software, for instance, tool *UhOh365*\(^9\) can see if an email address is valid in Office365, disclosing if the targeted organization uses that specific software.

### 2.2.3.6 Weaponization

During weaponization one of the key elements is to build the attack infrastructure that is used to either host malicious payloads or phishing sites that are used to trick users into submitting their credentials to the attacker-controlled infrastructure.

As figure 2.4 shows, an attacker needs to go through the sending profile creation, email template creation and landing page cloning/creation. To do so, some elements are required to build its

---

\(^6\)https://github.com/initstring/linkedin2username

\(^7\)https://osintframework.com/

\(^8\)Internet Protocol

\(^9\)https://github.com/Raikia/UhOh365
infrastructure, such as [3][49]:

- Server/Virtual Private Server (VPS) - backend server to host malicious payloads or phishing websites/toolkits, mail server to handle and deliver all emails over the internet;

- Domain Name - pointing to specific IP address(es), usually the backend server’s one, so that the victims can access all landing pages/websites created. Depending on where you buy/get the domain, you have the option to create an email address, associated with the bought domain name, to send your emails.

Regarding the backend server, the attacker can use a machine he already owns or acquire a VPS from one of the many cheap providers such as Linode, Digital Ocean, Amazon Web Services, Azure and Google Cloud. Either offer VPSs as cheap as 4.85€ per month, since you don’t need a lot of resources (memory, CPU and storage) to host a phishing campaign. This backend server can be configured to also serve as a mail server, which is a computer with mail transfer agent (MTA) functions. Mail is exchanged between mail servers running special software/solutions such as Exim, Postfix and Zimbra. While using these solutions it is not unusual to encounter port 25 blocked by the ISP or hosting provider (e.g., Digital Ocean) as a way to control email spam. One way to bypass port 25 and send email to the outside world is to use SMTP relay, using port 587, which allows you to send emails without it being from your mail server directly to the destination email address [54]. This means that the email is sent from one SMTP server to another before it reaches the final recipient. For SMTP relay there are also several options, such as Mailchimp, Mailjet.

Moving to the domain name, there are also several options on where to acquire and register a domain name. Among the most known domain name registrars we have entities such as Domain.com, Namecheap and GoDaddy, where you can acquire a domain name for prices starting from around 1.83€ per year. Due to the importance of the domain name in a phishing campaign, which is demonstrated primarily by the addresses that are accessed by the victim, the acquisition of the domain name must be done after the decision about the campaign’s theme, email template and landing page. There are several tools (e.g., dnstwist) and techniques to decide what domain name to use. Usually, the attacker’s objective is to impersonate the targeted company making use of domain spoofing, which occurs when the nefarious actor appears to use a company’s domain.

Here we have some of the techniques used to select/create a domain name [88]:

- Addition - appending letters to the given domain name (e.g., googlee.com);

- Bitsquatting - registration of a domain name 1-bit different from a legitimate one (e.g., wiklpedia.com). This is a little trickier on the eyes than the ”additions” above since we can read words, based on the first and last letter, without realizing that it has some wrong letter [16];

---

10SMTP port - designed to send email through a network and to its recipient
11https://github.com/elceef/dnstwist
• Homoglyph - phishing campaigns that use this technique are often referred to as homograph attacks [55], even though they rely on the use of a homoglyph. The attack works by registering a domain name using Unicode-encoded characters that look identical to the ASCII equivalents (e.g., apple.com);

• Omission - removing letters from the original domain name (e.g., facebok.com);

• Subdomain - a period inserted at varying positions in the given domain name (e.g., wik.ipedia.com);

• Vowel-swap - vowels found in the given domain are swapped for different ones (e.g., stack-overflaw.com).

Furthermore, there are some additional elements to consider: choosing a short domain name, given that long ones are usually associated with phishing, avoiding the number usage for the same reasons and also choosing the domain extension [57].

Since this work is conducted in a corporate environment, it is of notable importance that it is possible to present the results obtained through metrics, such as number of employees who submitted sensitive data. The easiest way to do this is by using phishing simulators/toolkits, most of them also facilitate the tasks of creating email templates, landing pages, sending profiles and campaign management. There are usually two solutions for this problem: open-source phishing platforms/toolkits or commercial products [2][74]. For commercial products, we have options like KnowBe4, Barracuda PhishLine, Infosec, among others, whereas for open-source we have products such as GoPhish, Lucy, and Social-Engineer Toolkit (SET).

After acquiring the necessary infrastructure, the issue of its configuration arises. Due to the existence of measures, implemented by various organizations, to prevent and mitigate phishing attacks, elements such as a domain with a valid SSL certificate and well-configured DNS records are even more important.

Every time we interact with the web we are making use of DNS records, while most people don’t even realize they exist. The DNS is a directory lookup service that provides a mapping between the name of a host on the Internet and its numerical address [82]. When someone types in a URL such as google.com, that entry is sent to an ISP, forwarded to the DNS servers, and then directed to the proper web server using the corresponding IP address as a label. This means that if a victim tries to access the nefarious URL, embedded on the email, it has to be pointing to the server’s IP address, by using a DNS A (address) record. Furthermore, since entities like Microsoft (Outlook) and Google (Gmail) use email authentication to stop email spam and spoofing, which relies on SPF, DKIM and DMARC [15][31], being vital to a successful phishing campaign:

• Anybody can connect to Gmail’s/Outlook’s SMTP servers and send an email claiming to be anyone, so security measures were put in place to prevent this. One of these measures

---

12 Domain Name System
13 Internet Service Provider
14 verifies that email messages from a sender are legitimate and come from expected sources for that email domain
is SPF, and it allows you to define on your domain name’s DNS records the set of IPs which are authorized to send mail. So when Gmail/Outlook gets an email claiming to be from yourdomain.com, it will go to your domain’s DNS records, and look for a TXT record defining a SPF policy. If the IP from where the email came is listed there, the email gets a PASS on the SPF policy [15][31];

- DKIM lets you add a digital signature to outbound email messages in the message header. When you configure DKIM, you authorize your domain to associate, or sign, its name to an email message using cryptographic authentication. Email systems that get an email from your domain can use this digital signature to help verify whether the incoming email is legitimate. In basic, a private key encrypts the header in a domain’s outgoing email. The public key is published in the domain’s DNS records, on a TXT record, and receiving servers can use that key to decode the signature. DKIM verification helps the receiving servers confirm the mail is really coming from your domain and not someone spoofing your domain [15][31];

- DMARC works with SPF and DKIM to authenticate mail senders and ensure that destination email systems trust messages sent from your domain. Implementing DMARC with SPF and DKIM provides additional protection against spoofing and phishing emails. DMARC helps receiving mail systems determine what to do with messages sent from your domain that fail SPF or DKIM checks [15][31].

Although SPF, DKIM and DMARC are very important, they are not the only DNS records that are important so that the sent emails don’t land on the spam folder or the victim can access the website without any trouble:

- MX - indicates how email messages should be routed in accordance with the SMTP, directing email to a specific mail server. If the attacker is hosting all on the same server, it has to be pointing to it;

- PTR - provides the domain name associated with an IP address. A DNS PTR record is exactly the opposite of the A record, which provides the IP address associated with a domain name. It is then used in reverse DNS lookups: a query that starts with the IP address and looks up the domain name;

- CNAME - has the purpose of forwarding one domain/subdomain to another domain (e.g., www.google.com to google.com).

While configuring all these DNS records, there are tools to help check the “spamminess” of the sent emails, such as Mail Tester 15, that indicate a score and what you must improve so that you bypass the spam filters.

15https://www.mail-tester.com/
The attacker also has to be careful during the creation of the email templates, given that there are spam/phishing filters in organizations that look at their content. There are certain words that are blacklisted, such as *invoice*, *free* or any overuse of money-related words, on many anti-spam software.

Finally, when the victim receives the phishing email and clicks on the link, accessing the phishing domain, it is important that there is a valid SSL certificate, giving the victim a certain confidence. This is a really easy problem to solve, for instance using Let's Encrypt, making it possible for the attacker to enable HTTPS on the phishing website.

### 2.2.3.7 Delivery

Once the infrastructure is set and all artefacts have been created, the campaign can be executed. For each victim, a phishing email should be generated from the selected email template by instantiating the template based on the information on that victim. The generated emails are then sent to the corresponding victim.

With the use of phishing toolkits, this becomes much easier and all the attacker needs to do is prepare the campaign during the weaponization phase, create the sender profiles, choose the targets and then launch the campaign, without worrying about sending the emails manually.

After launching the campaign, the phishing toolkit will gather all results and present them on the dashboard according to its metrics.

### 2.2.4 Current State - Defense

Oest et al. [61] posit that phishers have extensive control over the configuration of phishing sites they deploy. From the location (URL and hosting provider) of the site to the software the site uses to display the malicious content and capture user input, and the deceptive messages distributed to victims. Each of these areas involves an extensive network of entities that are either exploited during the phishing attack, who seek to fight phishing, or who are adversaries, which gives rise to a complex anti-phishing ecosystem.

We can see the components of the anti-phishing ecosystem in Figure 2.6. In the center of the phishing attack we have the *phishers* (1), *phishing message* (2), *victim user* (3) and the *organizations being impersonated* (4). Oest et al. [61] state that without these basic components, there would exist no basis of trust between the victim user and organization and no means of exploitation by the phisher.
As stated in [61], users are prone to re-using the same credentials across different services. This means that the damage from each successful phishing attack can potentially cause a chain reaction spanning multiple organizations. Thus, the organization directly targeted by the phisher expands to a set of indirectly targeted organizations of interest to the phisher that use the same authentication scheme. Given the risk of damage that arises as a result, the organizations implement mitigation strategies consisting of their own security teams, third-party anti-phishing vendors, or law enforcement.

Mitigation strategies are fundamental for every organization if they want to be the most secure possible and this comes with also being proactive instead of only reactive. Hong [38] posits that from the end-user’s perspective, there are three strategies useful to protect organizations and needed to offer the strongest possible protection against phishing attacks. The three strategies are:

1. Make it invisible so that users do not have to do anything different. The first line of defence is to prevent phishing attacks from reaching end-users. The solutions in this space include filtering phishing emails, blocking and taking down fake websites;

2. Provide better user interfaces that either make things more obvious to users or offer additional protection. Here we have solutions like warnings, support for properly identifying websites and authentications;

3. Train end-users to recognize and avoid phishing attacks. Training is an essential part of computer security but arguably the least popular approach [38], given the inherent challenges in motivating people to be secure, as well as the fact that training does not guarantee complete protection. Nonetheless, this is a very important mitigation strategy given that the human element is the weakest layer in every organization.
2.2.4.1 Prevention/Mitigation

Before we look at existing frameworks/methodologies or systems, it is essential to realize that in the business world, depending on each organization’s context, the needs or requirements of the system/framework that the company intends to implement will be different. Nevertheless, some requirements will be similar, such as the preference for something open-source, due to the possibility of collaborative and continuous innovation, which are valid points giving the organization the ability to have control and a more flexible, configurable, and cheap system/framework.

Frameworks based on honeypots are one example. Li and Schmitz [77] posit that honeypots have been widely used, as a powerful anti-phishing tool, by security service providers and financial institutes to collect phishing emails so that they can detect them earlier and shut them down quickly. Honeypots are also used to collect information about phishers’ activities which is then used to make statistics for research and forensics. They state that there are some problems with current anti-phishing solutions based on honeypots, and propose to transform the full e-banking system itself into a honeypot equipped with honeytokens and supported by some other kinds of honeypots [77]. The use of phoneybots is perceived as an indispensable part of the created framework. Phoneybots are active honeypots running in virtual machines and mimicking real users’ behaviour to access the real e-banking system automatically, to submit honeytokens to phishers and phishing malware. Applying this to an organization would require a well-thought-out study to adapt this framework to the context of the organization and a significant investment to develop the entire system and migrate all systems to build a honeypot-based one.

CERT.PT is a service within the CNCS that coordinates the response to incidents involving Public Administration entities, critical infrastructure operators, operators of essential services and digital service providers, as well as the entire national cyberspace, including any device belonging to a network or addressing block assigned to an electronic communications operator, institution, legal or natural person with headquarters in Portuguese territory, or that is physically located in Portuguese territory [9].

Frauenstein and von Solms reported back on an organizational framework which consisted of human, organization and technology (HOT) dimensions [30]. They point three different organizational dimensions targeted by phishing: Technology Controls, Organizational Aspects and Human Factors. This framework emphasizes the human factors since each dimension has human involvement. The framework defends that the human dimension calls for effective awareness, education and training to assist in strengthening the “human firewall” and cultivating information security behaviour in the organization ideally. Of these dimensions, the human dimension is the area that phishing exposes the most and as a result, compromises the technology and organizational dimensions [30]. Thus there is a need for education to be present in all dimensions to ensure that all HOT dimensions are functioning optimally. They present some examples of things to do/implement in the organization is each dimension:

- **Human Factors**
– *Education, Training, Awareness* - understanding a need for security controls; roles and responsibilities in organization; educating third parties; identifying phishing threats; communication; reporting.

– *Organization Culture* - attitudes and behaviours.

**Organizational Aspects**

– *Organization Policies and Procedures* - human resources; recruitment; access control; website access; restrictions; compliance.


– *Risk Assessment*.

**Technology Controls**

– *Technical Controls* - anti-virus and anti-malware software; software updates; network security; web browsers; browser plug-ins; online browsing practice; email client.

### 2.2.5 The Human Element

A reliable cybersecurity approach is a combination of the proper technology systems and tools in the hands of trained staff. Businesses often invest heavily in firewalls, virus protection or other technology while taking the human component for granted. The company’s security is only as strong as their weakest layer, and like stated before, that is the human element. The workforce through neglect, incompetence or malice represents the biggest threat to any organization. A business could spend millions on complex software, but if the manager in charge uses a weak password or a pin, such as 1234, then the entire system would be worthless. Attackers are looking for the easiest way in, and usually, that approach is found through weak passwords or lazy implementation of security features on an individual level.

The organization is ultimately responsible and liable for the actions of their employees, and those who do not adequately train and prepare their staff should expect to have problems. Many of the most common forms of data loss or breaches come from avoidable errors involving employees with poor account security or who fall victim to social engineering attacks when they are tricked into giving over confidential information. These days it is easy to get carried away worrying about external threats, but cybersecurity must be done with an inside-out approach.

Phishing is particularly dangerous because it exploits the manners in which we think, make decisions, how we handle our emotions and the ways we can subconsciously be persuaded. Anyone can fall to one of these attacks.

Psychologist Daniel Kahneman posits that there is actually two modes of thinking humans have, System 1 and System 2 [45], the fast and slow mode.
2.2.5.1 System 1

Here we have the source of many automatic emotional reactions we experience. System 1 operates almost immediately and intuitively, with little voluntary control or effort. Also, it “short circuits” System 2 by making some decisions before the second system even realizes what is going on [78], e.g., braking or swerving, without thinking about it, to avoid a car in front that just stopped.

Most of the time, System 1 functions are something that everybody is born able to do: autonomic physical functions, the capacity to perceive objects, take actions to avoid the loss of life or some property [78]. It operates fast and instinctively, below what is considered the level of conscious thought, being very emotional and only useful for simple problems that require nearly zero attention, to make decisions that are considered unimportant and that can be solved easily through pattern matching.

2.2.5.2 System 2

Operations within this system are related to choice, focus, and reasoning. It is a much slower and methodical way of thinking. When someone is writing a pros and cons list for some business decision, they use the brain’s System 2. All the encapsulated processes in this system require attention and are disrupted when attention is broken or reallocated. While more focused and intentional than System 1, System 2 has some drawbacks stemming from the fact that people have limited attention resources [78]. This system is skeptical and rational but the problem is that we, usually, do not have enough time to use it everywhere.

2.2.5.3 System 1 + System 2

What we have then are two systems that do not work in isolation from each other. System 1 can change the focus of System 2 and vice versa. For example, System 1 can react to a stimulus in an environment and from then on System 2 can become aware of what is happening in that zone, just as System 2 can make System 1 process certain things, such as recognizing someone in that zone or a name in a list [78].

When interacting with System 2, System 1 generates suggestions of impressions, intuitions, intentions and emotions. In most cases, System 2 accepts these same suggestions and implements them, but when there are conflicts, System 2 can allocate the necessary attention to resolve them. System 2 also places attention when there are problems that System 1 cannot solve or find an answer to, such as something that surprises us or that is shocking, that escapes the normality we are used to [78].

System 2 is a key factor in fighting System 1’s impulses. When System 1 is about to make a mistake System 2 intervenes. This is what is called, in our everyday, of self-control. Nevertheless, System 1 and 2 work in unison and complement each other the majority of the time. They optimize performance and minimize effort, just like systems in a well programmed neural network. But like any network the brain is subject to errors [78].
2.2.5.4 Systems interacting with phishing

When we start talking about phishing, the interaction between these two systems becomes fundamental when it comes to finding the reason why most people fall for these attacks. As previously mentioned, System 1 being more impulsive and System 2 more rational and logical, the attackers exploit System 1 and expect that System 2 will not react until it is too late. The reason this is possible has to do with the fact that System 1 processes thousands of minute decisions at every moment and when it gets overloaded it can generate a short circuit in System 2 [8].

With attackers attacking the human being’s way of thinking and emotions, they want their victims to have their minds working in System 1. In essence, what attackers do with phishing attacks is abuse principles that can be processed by System 1, some of which have been identified by psychologist Robert Cialdini in what he calls “Persuasion Science” [8]. These principles include: authority, liking/similarity, distraction, reciprocation, social proof/conformity. If the attacker uses and abuses these principles it becomes easier for the victim to fall into this type of attack because the email scan is assigned to System 1, since most people in the corporate world receive hundreds of emails per day.

Information security only has been considering these techniques of persuasion very recently, although there are several studies, in other realms, on the same subject. The psychologist Cialdini identified some persuasion principles and techniques in marketing, Stajano and Wilson evidenced principles of scams in street cons, and Gragg devised psychological triggers in social engineering [25]. These three families have been working independently but researchers started to question if their principles had something in common with the others. Motivated by this question, Ferreira and Lenzini [24] studied and compared the elements of the three families and extracted five general and independent principles from which all the others seem to derive. They named these Principles of Persuasion in Social Engineering (PPSE) and summarized them as Table 2.2 shows. They state that their PPSE can be used with more confidence than the original principles because they have been composed considering all perspectives (persuasion, scam/deception, and psychology of social engineering).

<table>
<thead>
<tr>
<th>PPSE</th>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authority</td>
<td>AUTH</td>
<td>Society trains people not to question authority so they are conditioned to respond to it. People usually follow an expert or pretence of authority and do a great deal for someone they think is an authority.</td>
</tr>
<tr>
<td>Social Proof</td>
<td>SP</td>
<td>People tend to mimic what the majority of people do or seem to be doing. People let their guard and suspicion down when everyone else appears to share the same behaviours and risks. In this way, they will not be held solely responsible for their actions.</td>
</tr>
<tr>
<td>Liking, Similarity &amp; Deception</td>
<td>LSD</td>
<td>People prefer to abide to whom (they think) they know or like, or to whom they pretend to be similar to or familiar with, as well as attracted to.</td>
</tr>
<tr>
<td>Commitment, Reciprocation &amp; Consistency</td>
<td>CRC</td>
<td>People feel more confident in their decision once they commit (in public or in writing) to a specific action and need to follow it until the end. This is true whether in the workplace, or in a situation when their action is illegal. People tend to believe what others say and need, and they want to appear consistent in what they do. When they owe a favour, there is an automatic response for repaying it.</td>
</tr>
<tr>
<td>Distraction</td>
<td>DIS</td>
<td>People focus on one thing and ignore other things that may happen without them noticing; they focus attention on what they can gain, what they need, what they can lose or miss out, or if that thing will soon be unavailable, has been censored, restricted or will be more expensive later. These distractions can heighten people’s emotional state and make them forget other logical facts when making decisions.</td>
</tr>
</tbody>
</table>

Table 2.2: The five principles of persuasion in social engineering (PPSE), from [25]
In [25], it was found the most used PPSE in phishing emails and revealed which ones are commonly used together. They analysed manually 52 phishing emails and organized them, according to their goal, into three categories:

- **Data Theft** (30 emails) - emails intended to steal banking, financial, and private or confidential data from use accounts (these emails include attacks such as Man-In-The-Middle, session hijacking, and impersonation);

- **Malware** (15 emails) - emails containing attachments or links to fake websites (ultimately containing Trojan horses, virus, system reconfiguration and malicious software);

- **Fraud** (7 emails) - emails offering large sums of money and prizes (e.g., the 419 or the Nigerian scams).

After all the information was taken from the emails and all the necessary analysis was done, in Table 2.3, it was concluded which PPSE is the most used alone, which pair and even which triple is the most used.

<table>
<thead>
<tr>
<th>Category</th>
<th>Most used PPSE</th>
<th>Most used pair</th>
<th>Most used triplet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Theft</td>
<td>LSD</td>
<td>AUTH + LSD</td>
<td>AUTH + CRC + DIS</td>
</tr>
<tr>
<td>Malware</td>
<td>LSD</td>
<td>DIS + LSD</td>
<td>AUTH + CRC + DIS</td>
</tr>
<tr>
<td>Fraud</td>
<td>LSD</td>
<td>CRC + DIS</td>
<td>AUTH + CRC + DIS</td>
</tr>
</tbody>
</table>

Table 2.3: Most used PPSE in phishing emails, from [25]

When analysing the sample of 52 emails, Ferreira and Lenzini used certain appropriate criteria to associate the most effective elements to the corresponding PPSE. These criteria, present in Table 2.4, serve to make it easier, when analysing an email, to identify which PPSE may be being used.

All this work is fundamental, giving a great advantage to those who prepare phishing campaigns, because now they know which PPSE are more effective, whether used alone, in pairs, or triples, and know examples of elements or certain phrases/expressions to put in a phishing email.
Table 2.4: Criteria to associate the most effective elements to corresponding PPSE, from [25]

### 2.2.6 Chapter overview

In this chapter, several themes were approached to present the basics behind cyber attacks, focusing on the topic of social engineering and phishing in particular. This theme was introduced with a brief contextualization of the first attacks and how they were carried out and then moving on to phishing’s current state. Within phishing’s current state, we presented the methodologies and related work that served as a basis for this thesis, ending the chapter by addressing a key topic in this type of attack: the human element. Our brain uses two systems to process information and react to it, systems that make us all vulnerable to phishing. We also approached five persuasion principles and which are the most effective, used alone or together.
Chapter 3

Proposed framework & campaigns

This chapter will start by presenting the work’s requirements, both functional and non-functional, given its context. It will also present the proposed framework that was designed to conduct the phishing campaign simulations and its use cases, as well as each element that composes it and the reasons for its choice. Furthermore, we will present the phishing campaigns that were created, detailing all phases throughout the process.

3.1 Work requirements

As mentioned in Section 1.1.3, in a larger project’s context, a client belonging to the IT industry requested EY’s consulting services to assess the effectiveness of the company’s awareness program. Hence, there were some requirements from both the client and EY, presented in Table 3.1.

It should be noted that some requirements are consequences of others, such as the integration of different tools/solutions due to the client’s reduced budget and the consequent necessity to opt for open-source software and free solutions.

<table>
<thead>
<tr>
<th>Work Requirements</th>
<th>Functional</th>
<th>Non-functional</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1. Integrate different solutions (e.g. phishing simulator + mail server) to create and conduct phishing simulations</td>
<td>NF1. Opt for open-source/free solutions due to the client’s reduced budget</td>
<td></td>
</tr>
<tr>
<td>F2. Ability to bypass possible controls, being undetectable, as well as drawbacks caused by using open-source software so that the phishing simulations are more effective</td>
<td>NF2. Use an existing backend server (initially on-premises and later migrated to Azure)</td>
<td></td>
</tr>
<tr>
<td>F3. Ability to capture and monitor events and results to calculate/generate required metrics and present them to the client</td>
<td>NF3. Conduct two phishing simulations (generic vs spear-phishing) integrating PPSE into the email template creation, assuring email deliverability</td>
<td></td>
</tr>
<tr>
<td>F4. Required metrics: emails sent, emails opened, clicked links, submitted data and emails reported</td>
<td>NF4. Comply with the provided list of target emails</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: Functional and non-functional work requirements, from both the client and EY


3.2 Use cases

This section presents the use cases for the proposed framework in this work, aiming to demonstrate its utility within the realm of cybersecurity:

- **Security Awareness Training**: As mentioned throughout this document, security awareness training inside businesses/organizations is vital to improving cybersecurity levels. This framework allows organizations to simulate phishing attacks and collect relevant metrics from employee responses. This allows the organizations to have the necessary information to later provide training to enhance the employees’ awareness of phishing attacks.

- **Incident Response Testing**: This framework allows an organization’s security team to conduct phishing campaign simulations to assess their ability to detect/block (e.g. security tools), respond and mitigate phishing attacks.

- **Email Security Testing**: As mentioned in Section 3.1, one of the work requirements was that the framework had the "ability to bypass possible controls being undetectable (...)". This allows organizations to test and evaluate their email security policies and procedures against new mail servers, SMTP relays and reverse proxies.

These use cases highlight the versatility, accessibility and practicality of the proposed framework. It helps to improve cybersecurity measures, gather information, build custom reports and automate phishing campaign creation. Emphasising that given that the framework is fully open-source, it allows all organizations to be autonomous in their security testing, whether they have a large or reduced budget for it.

3.3 Proposed framework

To meet the work’s functional and non-functional requirements, presented in Section 3.1, a framework was proposed and developed to conduct the phishing simulations. To get to its final design, consisting of six elements/tools, we researched and compared different solutions over time. This proposal is also based on tests carried out to validate the correct integration of the different elements, as well as their ability to contribute to the phishing simulations’ good execution and performance.

3.3.1 Framework elements

1. **Backend/phishing server**

   Starting with the backend/phishing server, as mentioned in Section 3.1, it was provided by EY, being initially a physical machine and later migrated to an Azure-hosted VPS with a Linux distribution. This allowed to meet some non-functional requirements, given the ease of integration of open-source tools into the environment, as well as integration between tools.
2. Phishing framework/toolkit

Since the backend server was already fully set up and running, the first choice to be made was regarding what software to use to create and conduct phishing simulations. The most effective tools available to perform automated and large-scale phishing campaigns are phishing toolkits [13]. While planning to perform phishing simulations in a corporate environment, there are two areas to choose from: commercial products and open-source solutions. Given the non-functional requirement, presented in Section 3.1, to opt for open-source solutions due to the client’s reduced budget, this type of solutions were prioritized.

Table 3.2 compares the tools that were analysed and that are present as some of the best phishing simulators in Infosec Institute’s list [2]. In each column we can observe essential features for the evaluation/choice of a phishing simulator and to meet the requirements presented in Section 3.1. Thus, the result of the analysis performed is presented as well as the support of each solution for each feature.

The chosen features for the software analysis, were the following:

- **GUI** - The existence of a GUI makes the job of the person who will create the campaign and track its results easier, offering the possibility for a less technical employee to learn how to use the tool more easily;

- **Landing page customization** - The ability to automatically clone an original page and edit it later is also very important. This way, if the goal of the campaign is to impersonate some well-known company (e.g. Microsoft), it is possible to clone the original page to give more credibility when the victim accesses it. Furthermore, the ability to create your own landing page is also beneficial at times. If a company wants to create a new scenario for a phishing campaign to try to catch employees off guard, it is very important to have the ability to create your own landing page;

- **Email customization** - The feature of email customization is perhaps the most important one when creating a spear-phishing campaign. This allows each email to be based on different information about each target, making it seem more personal. Additionally, in another context it is also very useful to have the ability to import an original email automatically and modifying it later;

- **Event capture** - This is a necessary feature to meet the work’s functional requirements. In this case it would be necessary to be able to get information regarding opening the email, accessing the link, submitting data through the fake landing page and reporting the email. From the point of view of training employees after the simulation, it is crucial to be able to differentiate which employee submitted data so that this human vulnerability can be patched;

- **Result evaluation** - Finally, the evaluation of results is important because it means that the tool integrates the collection of metrics automatically and graphically to observe the
progress of the campaign. This feature again helps to meet the work’s functional requirements, facilitating the presentation of results to the client;

<table>
<thead>
<tr>
<th>Framework/Toolkit</th>
<th>GUI</th>
<th>Landing page customization</th>
<th>Email customization</th>
<th>Event capture</th>
<th>Result evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gophish</td>
<td>X</td>
<td>X</td>
<td>O</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>King Phisher</td>
<td>X</td>
<td>X</td>
<td>O</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Simple Phishing Toolkit</td>
<td>X</td>
<td></td>
<td>O</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Phishing Frenzy</td>
<td>X</td>
<td>X</td>
<td></td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Social-Engineering Toolkit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>SpeedPhish Framework</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>O</td>
</tr>
</tbody>
</table>

Table 3.2: Framework/toolkit comparison, adapted from [67]. X - full support, O - partial support, blanks - no support

Given each tool’s support for capturing events and evaluating campaigns, only two really presented themselves as possible solutions, so the choice ended up being between Gophish and King Phisher, with both having similar support on all areas. Although the remaining tools were not considered at the time of the final decision due to the results presented in the table above, it is also worth highlighting the results for the other fields that were analysed and identified as needing improvement. One of them is the support that exists for the tool itself, with all being open-source solutions, it is fundamental that there is a team supporting the project and trying to solve problems that appear over time, such as configurations, operation or vulnerabilities (e.g. Simple Phishing Toolkit has been “abandoned” since 2013). This field leads to the second one that was identified, which was the installation and configuration process as a whole, some times being quite complex and made even more difficult due to lack of information.

As far as the first two tools (Gophish and King Phisher) are concerned, although the end result may be alike, the whole process of installing/configuring and preparing/launching a phishing campaign in each of the two tools is quite different. This ended up being the reason for choosing Gophish over King Phisher. While King Phisher is a tool with plenty of features such as the ability to clone sites, repositories containing templates for emails and landing pages, it falls short in the installation and configuration process, with it being much more complex than Gophish’s. On the other hand, Gophish is a user-friendly tool throughout the entire usage process and offers almost the same features. As Infosec Institute states [2], Gophish’s installation is as simple as downloading and extracting a ZIP folder, the interface is simple and intuitive, and the features are thoughtfully implemented. Users are easily added, either manually or via bulk CSV importing. Email templates are easy to create and modify, creating campaigns is a straightforward process, and reports can be exported to CSV format with various levels of detail.

3. Mail server (MTA)

A mail server is simply an application or computer in a network with MTA functions and for this element there are numerous open-source MTA solutions. Although it is mandatory to choose
one so that emails can be sent, given the work’s context, our priority when choosing this element was based on finding a Linux-compatible, reliable and secure solution but also easy to install and configure. The diversity of features of each tool, besides the essentials of any MTA, were put aside during the analysis, since there would be no direct contact with the tool beyond the process of installation and integration into the framework. The tools we considered and analysed were Sendmail, Exim and Postfix. The comparison between the three softwares is presented in Table 3.3.

<table>
<thead>
<tr>
<th>MTA</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sendmail</td>
<td>- Portability</td>
<td>- Customization</td>
</tr>
<tr>
<td></td>
<td>- Flexibility</td>
<td>- Weak security mechanisms</td>
</tr>
<tr>
<td></td>
<td>- Configurability</td>
<td></td>
</tr>
<tr>
<td>Postfix</td>
<td>- Security orientation</td>
<td>- Customization</td>
</tr>
<tr>
<td></td>
<td>- Tool documentation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- High queuing operation speed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Active development</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Configurability (parameters defined in the config file)</td>
<td></td>
</tr>
<tr>
<td>Exim</td>
<td>- Support community</td>
<td>- Complexity</td>
</tr>
<tr>
<td></td>
<td>- Scalability</td>
<td>- Slow queuing operation speed</td>
</tr>
<tr>
<td></td>
<td>- Flexible configuration (variables declaration,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scripting language to process emails)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.3: MTA software comparison

- **Sendmail** - Now known as Proofpoint, is standard MTA for Unix systems. It is one of the oldest and most known but since 1996 that its market share has been dropping, from 84% to below 4% of all online public email servers on the Internet [92], having security gaps [28] reported on a daily basis;

- **Postfix** - In 1998, Wietse Venema made Postfix, that came to life in IBM research, available as an alternative to the widely-used, at the time, Sendmail program[17] [28]. Postfix provides many parameters, options and features a security-oriented architecture, while offering really good performance, clocking over 300 e-mail deliveries per second from a single Postfix instance. Furthermore, according to a recent study [21], more than 30% of the publicly reachable mail servers in the world run Postfix, so there is plenty of useful documentation available that can help users adapt much faster;

- **Exim** - A feature-rich mail server widely used (more than 60%) [21] on many Unix-like operating systems such as Linux. It also provides efficient cross-platform support, making it a perfect choice for mail servers that need to be diverse and handle a large number of requests. It has its own application language to perform complex configurations and an advanced multi-step mail processing logic which is suitable for complex cases. Moreover, the development team rolls out frequent updates and takes care of bug fixes very well. Lastly, more than 30% of the publicly reachable mail servers in the world run Postfix.

Although Sendmail was once the most widely used MTA, at the moment it loses to Postfix and
Exim in all aspects including usability, leading to the final choice between Postfix and Exim. While Exim has features that are quite useful for a larger and more complex context, it ends up having a more painful installation and configuration process. Since the goal of this element was only to configure and integrate the framework with Gophish, the tool of choice was Postfix. Basically it ends up being a tool created with goals such as reliability, security, performance, flexibility and ease-of-use [17].

4. Email service provider (SMTP relay)

After the integration and configuration of Gophish and Postfix on the VPS, a test campaign was run to check if the delivery of emails was running smoothly. This was not the case, as port 25 was blocked by the hosting provider as a way to control email spam, resulting in no emails being sent. As stated in Section 2.2.3.6, one way to bypass port 25 restrictions is to use SMTP relay.

With SMTP relay, your own mail server does not directly send email to the destination email address. Instead, there is an intermediate mail server, such as an email service provider, that sends the email for you. Your email server communicates with the relay host on port 587, then the relay host communicates with the recipient’s mail server on port 25. SMTP relay also helps to improve email deliverability and to bypass anti-spam blacklists, if the backend server’s IP address gets blacklisted for some reason, given that SMTP relay services maintain good IP reputation. The recipient’s mail server checks the relay host’s IP address against public anti-spam blacklists, instead of the backend server’s IP address.

There are several ESP that can act as a relay host and offer services for free, such as Sendinblue, Mailchimp or Mailjet. The biggest difference between the free plans of the various ESP is usually the number of emails they allow you to send per day/month. For the context of this work, the chosen solution needed to have, as mandatory requirements, the compatibility with Postfix, a reasonable number of emails allowed to send per day/month as well as the ability to be configured as a relay host.

After analysing the three solutions listed above, it was possible to see that they all offered similar features and advantages in improving email deliverability and helping to avoid emails going into the spam folder. Ultimately, the reasons that made us choose Mailjet to play the role of relay host were:

- **Reputation** - Mailjet has a very good reputation and more than 130 thousand clients, giving confidence regarding the good IP reputation;

- **Configuration** - Very easy to set up and configure additional features, such as domain authentication (SPF and DKIM records);

- **Familiarity** - The fact that we had already worked with it. As this is, again, an element that only needs to be configured and from there serves its purpose without requiring contact with it, we prioritised a solution that we had already worked with, with positive feedback.
Finally, test campaigns were rerun to check if the delivery of the emails went as expected, also analysing the email score with the help of the mail tester tool 1.

5. Domain name registrar

When it comes to the domain name registrar, there are hundreds of such services available online, and technically they each fulfill the same function. However, that doesn’t mean they’re all equal. There are a few key areas in which registrars tend to differ, such as additional hosting services, different prices for the same domains or better domain management features. Given the project’s context and the role that the domain registrar will play, the main relevant features were the price difference, a user-friendly and easy-to-use interface. Due to an existing previous experience, the choice was Namecheap, which is well known for offering unbeatable domain name prices and an excellent UI.

6. CDN 2 service provider (reverse proxy)

Before buying any domain, some tests were performed with free domains. To provide an additional sense of trust and security to the target when they open the campaign’s landing page, some tests have been made regarding the use of SSL certificates. Although there are several open certificate authorities on the market, such as ZeroSSL or Let’s Encrypt, its configuration requires a few steps, install software that uses the ACME protocol (e.g., Certbot) to demonstrate control over the domain, and after doing some tests, we came to the conclusion that the client’s filters/controls could be sending the emails to spam due to Let’s Encrypt certificates.

To facilitate this process but also add a security layer to the framework to ensure that the backend server is not IP blacklisted, it was decided to use Cloudflare, a CDN service provider, as a reverse proxy. A reverse proxy sits in front of an origin server and ensures that no client ever communicates directly with that origin server, thus protecting the original server with landing page content. Furthermore, Cloudflare also offers SSL certificates, as well as DNS record management, which will then be handled from their platform. All of this is in addition to the fact that all of these services can be used for free.

3.3.2 Framework design and Campaign flow

As detailed above, the framework consists of six different elements, these being:

1. **VPS**, where most of the weaponization phase takes place, making use of all software that was set up (Gophish, Postfix);

   1.1. **Gophish**, phishing framework/toolkit where the campaign is created and launched (email template, landing page, sending profile);
1.2. **Postfix**, mail server that manages all phishing emails that are set to be sent, by launching the campaign on Gophish, and communicating with the relay host to "pass" the emails;

2. **Mailjet**, email service provider that serves as an SMTP relay host, receiving communications from Postfix and delivering all emails to the client’s mail server. Furthermore, with Mailjet you also do additional configurations such as creating the sender profile for the emails as well as setting up SPF and DKIM records (domain authentication);

3. **Namecheap**, where the domain is purchased, which is later inserted in the Gophish tool when creating the landing page, indicating the URL that will be accessed when clicking on the link in the email;

4. **Cloudflare**, which provides the SSL certificate, to use on the campaign’s landing page, and also serves as a reverse proxy server to protect the backend server’s IP. After being generated, the certificate is integrated into the landing page via Gophish. Furthermore, when integrating the tool with the other infrastructure’s elements, the DNS records management is transferred and becomes configurable in Cloudflare.

In Figure 3.1 we can see how all elements are connected as well as the high-level flow of a phishing campaign from the steps of creation/launch to the target’s data submission. All connections between elements are indicated with black arrows while the campaign’s high-level flow is indicated with green arrows. These links and flows will now be explained in more detail.

![Figure 3.1: Framework design, connection and integration between elements (black arrows) and campaign’s high-level flow (green arrows)](image)

Regarding the integration and connections between elements, Figure 3.1 shows six different links:
• Postfix & Gophish - It is required that Postfix is configured and operational to be able to: insert all information regarding the SMTP server responsible for sending all emails; and create/configure sending profiles in Gophish;

• Mailjet & Postfix - For the SMTP relay to be integrated, it is necessary to configure the service (Mailjet) and then insert the necessary information in Postfix so that all emails can be sent via Mailjet;

• Namecheap & Gophish - After purchasing the domain to be used in the phishing simulation and when creating the campaign, it is necessary to insert the domain where the landing page will be accessible to the victims. That same domain is bought via Namecheap and that information is then entered in Gophish;

• Namecheap & Cloudflare - To integrate the reverse proxy into the framework, it is required to configure Cloudflare, importing the existing website (associated with the purchased domain) and transferring the management of all DNS records;

• Mailjet & Cloudflare - For domain authentication (SPF and DKIM records) to be introduced, it is necessary to start its configuration in Mailjet and then create the records in Cloudflare. Note that this link could be with Namecheap instead of Cloudflare, depending on when the configuration is performed and where DNS records are managed: if it is performed before the reverse proxy integration then the link would be with the domain name registrar;

• Cloudflare & Gophish - Finally, for the landing page accessed to be as credible as possible, it must have an SSL certificate. This certificate is configured in Cloudflare and then imported, to a specific folder of the project, to be integrated in Gophish.

The flow of a campaign or phishing simulation has three different phases. For those conducting the phishing simulation, only actions in the first phase are required. The following ones are conducted by elements of the framework or are influenced by the targets themselves. The three phases are the following:

1. Campaign creation and launch - This entire phase takes place inside the VPS, more precisely within Gophish. To create and launch a campaign you it is required to set up/create a target group that the emails will be sent to, an email template to use, a landing page that will be visited and a sending profile so that Gophish knows which SMTP server will deliver the emails;

2. Sending and delivering all emails - From the moment the campaign is launched, Postfix “forwards” the emails to the relay host, Mailjet. This way, in the eyes of the customer’s SMTP server, they have been delivered by the Mailjet server, a globally recognised and reputable entity, increasing the deliverability of the emails. Next, the emails are delivered to the client’s SMTP server and then distributed to each target’s inbox;
3. Landing page access and data submission - The existence of the third and final phase is dependent on the targets themselves. If no one opens the email this phase ceases to exist, but otherwise, Gophish captures the different events that may exist (opening the email, accessing the link and submitting data) and if there is data submission and the option to do so has been activated, it also captures them. All this information is stored by Gophish and presented in the dashboard, where it’s possible to check each event of each victim.

After presentation of the framework design, as well as the links between each element and the flow of a campaign, it is possible to understand the role of each element and its contribution so that the work’s requirements are met. Table 3.4 summarises these links between requirements, both functional and non-functional, relevant within the context of the framework design, and the elements that contribute to their fulfilment.

<table>
<thead>
<tr>
<th>Requirements</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>Six different elements: Azure-hosted VPS, Gophish, Postfix, Mailjet, Namecheap, Cloudflare</td>
</tr>
<tr>
<td>F2</td>
<td>Mailjet, Cloudflare</td>
</tr>
<tr>
<td>F3</td>
<td>Gophish</td>
</tr>
<tr>
<td>F4</td>
<td>Gophish</td>
</tr>
<tr>
<td>NF1</td>
<td>Gophish, Postfix, Mailjet, Cloudflare</td>
</tr>
<tr>
<td>NF2</td>
<td>Azure-hosted VPS</td>
</tr>
</tbody>
</table>

Table 3.4: Connection between requirements (presented in Section 3.1) and the constituent elements (presented in Section 3.3) of the framework that helped to meet the work’s requirements

### 3.3.3 Framework Operations

This section presents the different operations that can be carried out. Figure 3.2 shows all these operations subdivided into three different phases, starting from the configuration/integration of the framework and its core elements plus Mailjet, to the preparation of the campaign itself and finally to the creation and launch of the campaign.

#### 3.3.3.1 Framework configuration

The configuration phase is about exactly that, configuring and setting up the “first”/core elements of the framework on the VPS, these being the phishing toolkit, the SMTP server and the SMTP relay. Both Gophish and Postfix need to be installed on the VPS, while Mailjet is integrated as an SMTP relay from the platform itself and later in the Postfix configuration file.

#### 3.3.3.2 Campaign preparation

The campaign preparation phase consists of all the operations required to finalise the set up of the framework but which are dependent on the purchase of the domain to be used. Once the domain chosen for the campaign has been purchased, the integration and set up of the reverse proxy,
the domain authentication in the DNS records, the SSL certificate and the sending profile can be started.

3.3.3.3 Campaign creation and launch

This whole phase takes place in Gophish, where all the operations required to create and launch a campaign are carried out, these being: Create sending profile - where the SMTP server to use is specified, as well as the email address that will be the sender; Create/import new group - where the sample of targets is created/imported; Create/import email template - where the email template that will be used during the campaign is created/imported; Create/import landing page - where the landing page is created/imported, which will be accessed when the victims click on the email’s link. Note that both admin and simple users have permissions to perform these operations. Once the campaign is running and events have been captured, it is possible to export these events as well as the overall campaign results. Furthermore, the admin user can also perform user management.

Figure 3.2: Framework’s (presented in Section 3.3) operations
3.4 Phishing Campaigns

To create and conduct the campaigns, the chosen methodology was based on Lockheed Martin’s Cyber Kill Chain model, as well as the flow of a targeted phishing campaign presented by Pirocca et al. [67], as mentioned in Section 2.2.3.4. The steps taken to conduct the phishing campaigns were:

- Reconnaissance: OSINT gathering;
- Weaponization: Sending profile creation, Email template creation, Landing page cloning/creation;
- Delivery: Campaign launch.

Despite the existence of studies in which the use of attached files was more successful [19], the embedded links technique was prioritized during the email creation phase. The integration in a corporate environment like EY’s allowed us to acquire some knowledge about controls and filters performed by several companies regarding emails. With that information, it was possible to realize that blocking emails from unknown senders is something quite common. Furthermore, one of the fundamental points during all the work developed was the integration of the PPSE in the emails.

Since the final goal of the phishing emails was data theft (employees’ credentials), we took into account the study discussed in Section 2.2.5.4 in the choice of the PPSE to be used for this purpose, as shown in Table 2.3.

Knowing that the campaigns’ launch date was dependent on the client’s prior approval, after having the infrastructure ready we immediately started thinking about context/scenarios for the email templates. By the client’s request, the first to be prepared was the spear-phishing campaign, followed by the more generic one.

3.4.1 Campaign I

3.4.1.1 Reconnaissance

The first step is to gather as much information as possible on the target, trying to identify weak points and possible victims. Given what was allowed by the client and the fact that they had previously provided a list with the target’s names and emails, the focus of this phase was on passive reconnaissance.

One technique used during this phase was Google hacking. As mentioned in Section 2.2.3.5, this technique aims at collecting information available on the internet about the client, to gather useful information for the email template creation. After some research using operators such as site, intext or related, as well as the combination of some of those operators, we found some information regarding an organization that acquired third party positions in the client’s company, consequently becoming the sole shareholder. That same information turned out to be very useful, serving as the basis for the creation and design of the first phishing campaign.
Another step we took in this phase was to use the *UhOh365* tool to verify if the provided email addresses in the list existed and were valid in Microsoft 365. Although most companies use Microsoft 365, performing this step allowed us to confirm and validate the existence of all email addresses in the list, to make sure that this would not be a reason for the campaigns not being successful.

### 3.4.1.2 Weaponization

This phase started well before the phishing campaigns were created and launched, and included all the steps for designing and configuring the infrastructure presented in Section 3.3.

The scenario created for the first campaign took into account the whole context of the COVID-19 pandemic we had been experiencing, a theme that has represented a large part of the phishing scams in recent times [22]. It was designed and created at a time when the country was still in a state of emergency but was approaching the start date of the deconfinement plan. Additionally, in this same period, we began to see an increasing concern about the mental health of the world’s population, leading the European Commission to present new strategies for health and safety at work [53]. Taking this into account, and adding the fact that the company had recently been acquired by another organization, the scenario we ended up creating was a partnership between the acquiring organization and a hotel to raffle off one employee who could enjoy a weekend in the hotel.

**Email design, PPSE integration and campaign creation**

Given that this email would forge an email sent by the acquiring company and that we did not have access to emails from it neither from the acquired company, the email design was mostly based on emails we would receive on EY’s inbox during the internship.

As mentioned in Section 2.2.5.4 and shown in Table 2.3, the most used PPSE triplet in emails with data theft purposes are AUTH, LSD and CRC. Although in Table 3.5 we can observe four distinct PPSE (AUTH, LSD, CRC and DIS), when writing the email the use of the PPSE triplet was prioritized due to the campaign’s data theft purpose. The inclusion of the fourth PPSE in Table 3.5 is due to the similarity of criteria between PPSE (Table 2.4) and the possibility that an email’s element corresponds to more than one PPSE.

Analysing the email in more detail, with Figure 3.3 and Table 3.5 we can observe the association between the email’s marked elements and the corresponding PPSE.
Chapter 3. Proposed framework & campaigns

Figure 3.3: Phishing email with PPSE elements marked in red

<table>
<thead>
<tr>
<th>Id</th>
<th>Email element</th>
<th>PPSE</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Caro(a) Paulo</td>
<td>LSD</td>
<td>Kind and respectable words directed to the user (expressing liking)</td>
</tr>
<tr>
<td>2</td>
<td>“A... compreende as incertezas de um futuro próximo, contudo a saúde e o bem estar dos colaboradores será sempre uma prioridade...”</td>
<td>LSD</td>
<td>Kind and respectable words directed to the user (expressing liking) Reference to people/colleagues from the same company (“dos colaboradores” - expressing similarity)</td>
</tr>
<tr>
<td>3</td>
<td>“Clean &amp; Safe” stamp</td>
<td>AUTH</td>
<td>Graphics and logos from well known reputable companies (to acquire the stamp it is necessary to comply with the new requirements of Turismo de Portugal, created with the help of the partnership between Direção-Geral da Saúde and NOVA Medical School - expressing deception)</td>
</tr>
<tr>
<td>4</td>
<td>“através do endereço: <a href="https://...solutions.com%E2%80%9D">https://...solutions.com”</a></td>
<td>AUTH</td>
<td>Words expressing authority (similar to the expression “click here”)</td>
</tr>
<tr>
<td>5</td>
<td>“O sorteio irá decorrer de 08/03/2021 a 12/03/2021, participe já!”</td>
<td>AUTH</td>
<td>Words expressing tight deadlines (short period to participate in the raffle)</td>
</tr>
<tr>
<td>6</td>
<td>Boa sorte</td>
<td>LSD</td>
<td>Kind and respectable words directed to the user (expressing liking)</td>
</tr>
<tr>
<td>7</td>
<td>Signature with organization’s logo, email, address, phone number, website and hotel partnership indication</td>
<td>AUTHLSD</td>
<td>Graphics, logos and contact from well known reputable companies (expressing deception)</td>
</tr>
</tbody>
</table>

Table 3.5: Association between the email’s elements (Figure 3.3), the corresponding PPSE (Table 2.2) as well as the criteria (Table 2.4)

The next step was to create the landing page. Since it would be necessary to login with company credentials to participate in the raffle and so that the partnership would look as credible as possible, a simple page was created with some elements to highlight: the background image being the hotel of the partnership, a login area (placeholder text, in the top text input box, indicates that the employee needs to log in with company credentials to enter the raffle, three logos, one of the client and two of the partnership’s organizations).
The next steps included the domain name purchase (Namecheap), reverse proxy integration (Namecheap and Cloudflare), configuration of all DNS records (Mailjet and Cloudflare), integration of the landing page’s SSL certificate (Cloudflare and Gophish) and the creation of the email sending profile (Mailjet). Since this spear-phishing campaign was created completely from scratch, it was decided that a new domain name would also be chosen, i.e. none of the techniques mentioned in Section 2.2.3.1 were used. It ended up being a junction between the name of the organization that acquired the client’s company and the word solutions, due to the fact that this organization names their services as solutions. Additionally, a subdomain with the client’s name would also be added, resulting in: \([\text{client}].[\text{organization}]\text{solutions}\). The domain was purchased, for 2.5€, with the extension “.com”, as it is the most used worldwide. The sending profile created was \(\text{sorteios}@[\text{domain}\_\text{name}]\).com. Furthermore, SPF and DKIM records were created to ensure domain authentication and improve email deliverability.

After all the steps mentioned above, we have in Table 3.6 all DNS records that were created for the first campaign.

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>([\text{organization}]\text{solutions}.com)</td>
<td>[VPS’ IP address]</td>
</tr>
<tr>
<td>CNAME</td>
<td>www</td>
<td>([\text{organization}]\text{solutions}.com)</td>
</tr>
<tr>
<td>CNAME</td>
<td>[\text{client}]</td>
<td>([\text{organization}]\text{solutions}.com)</td>
</tr>
<tr>
<td>TXT</td>
<td>mailjet._domainkey</td>
<td>k=rsa;p=(...)</td>
</tr>
<tr>
<td>TXT</td>
<td>mailjet._ownership_token_piece</td>
<td>[ownership_token_full]</td>
</tr>
<tr>
<td>TXT</td>
<td>([\text{organization}]\text{solutions}.com)</td>
<td>v=spf1 include:spf.mailjet.com(...)</td>
</tr>
</tbody>
</table>

Table 3.6: DNS records table for Campaign I

Finally, the final step was to add the SSL certificate to make the credibility of the landing page as high as possible. So, if the victim has some kind of knowledge about some steps to distinguish between a real vs. phishing page, it will not be because there is no certificate and consequent “padlock icon” that the campaign will not be successful. This step is performed in the Cloudflare tool with the creation of one digital certificate and two files ([name].cert and [name].key), inside
the Gophish folder on the VPS.

With all the elements of the infrastructure integrated and all campaign elements created, we did some tests sending the campaign’s email to verify the deliverability of the emails (also using mail tester tool to test the “spamyness” of our emails), the look of both the preview of the email in the inbox (Figure 3.5) and the email itself (Figure 3.6), verification of the entire landing page (appearance, login’s functionality) as well as the validation of the SSL certificate on it (Figure 3.7).

Figure 3.5: Email preview on the target’s inbox

Figure 3.6: Opened phishing email from the perspective of the victim

Figure 3.7: Landing page URL with proof of a valid certificate
3.4.1.3 Delivery

The last phase of this whole process is about delivering the campaign, that is, sending the emails. Hence it is necessary to perform five distinct steps that are presented in Figure 3.8. Furthermore, there is no restriction as to the order in which the first four steps have to be performed, but to perform the last one it is necessary to have previously performed the other four, with all five steps to be performed in Gophish.

![Figure 3.8: Five steps required to launch a campaign in Gophish](image)

Finally, after launching the campaign, it is possible to track all events in the dashboard, as well as the information regarding the capture of the different metrics, with a real-time update of the values for each campaign: **number of emails sent, emails opened, clicked links, submitted data, and emails reported.**

3.4.2 Campaign II

3.4.2.1 Reconnaissance

Since the second campaign was meant to be a generic phishing, everything that was performed for Campaign I (presented in Section 3.4.1.1) was useful and we did not feel the need to perform any more OSINT/information gathering.

3.4.2.2 Weaponization

Taking advantage of the fact that, in the reconnaissance phase, the customer’s use of Microsoft 365 was confirmed (Section 3.4.1.1), and also the fact that Microsoft is the world’s most impersonated brand in phishing attacks [20][68][69], we decided to go with a Microsoft email scenario. When analysing campaigns previously conducted by EY and researching which emails would be the most used in Microsoft-related phishing campaigns, since some of them are made with data theft purposes, we noticed that the focus is very much on security alerts. Although the focus is usually the same, there are several approaches, such as [20]:

- “Someone has accessed your account (...) please click here to retrieve it (...)");
- “Your inbox is full (...) please click here to update the quota (...)");
- “Someone used your password to try to log into your account. Microsoft blocked the access attempt, but we suggest that you change your password (...)");
• “We have detected suspicious activity on your account (...) please click here to change your password (...)”.

While these are some of the most used types of emails in this context, which means they probably continue to have some success rate, the likelihood that some employee has already received such an email is also considerable. Additionally, due to some time exposed to the Microsoft 365 environment at EY, and consequently to its emails, we decided to follow a different approach than the ones mentioned above. With this new approach, the goal was also to turn the campaign into a case study for EY, regarding the success rate of a new type of “impersonating Microsoft” phishing campaign, generating some knowledge for future services.

As mentioned in Section 2.2.2, as a consequence of the pandemic context, several companies had to adapt and have their employees work remotely from their own homes. For companies using Microsoft 365, this change quite possibly meant a significant increase in file sharing, between employees, in applications such as Microsoft Teams. During the period interacting with the Microsoft 365 environment we could observe that when an employee X shares a file with an employee Y, an email is always sent notifying Y that X has shared a file with him (Figure 3.9).

![Figure 3.9: Original email notifying employee of file sharing](image)

**Email design, PPSE integration and campaign creation**

By having access to the original emails and the messages’ HTML in EY’s inbox, it was possible to create an almost identical copy (Figure 3.10). It should be noted that the differences in font that can be observed between the emails is only caused by the emails having been opened in different programs (e.g., one in the browser and one in the application itself).
Chapter 3. Proposed framework & campaigns

Figure 3.10: Phishing email with PPSE element marked in red

<table>
<thead>
<tr>
<th>Element</th>
<th>PPSE</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsoft logo + “Declaração de Privacidade”</td>
<td>AUTH</td>
<td>Graphics and logos from well known reputable companies (expressing deception)</td>
</tr>
<tr>
<td></td>
<td>LSD</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.7: Association between the email’s element (Figure 3.10), the corresponding PPSE (Table 2.2) as well as the criteria (Table 2.4)

Analysing the email in terms of PPSE usage, as can be seen in Figure 3.10 and Table 3.7, there is only one element that deserves attention: the Microsoft logo itself along with the “Privacy Statement” just below. This element appears as if replacing a “normal” email signature, coming from a human sender who intends to identify himself. However, given the existing criteria, the use of this element covers two distinct PPSE (AUTH and LSD) that make up the most commonly used PPSE pair in emails for data theft purposes (Table 2.3). However, other elements are worth mentioning, even though they do not qualify to be considered PPSE. These elements are:

1. The name “Rodrigo Oliveira” was chosen randomly, so as not to be a name known within the client’s company and to generate some kind of curiosity in the possible victims, with the possibility of being a new employee;

2. The supposed shared file, “Top 10 National Cybersecurity Companies”, since this name was chosen to generate curiosity among employees and, taking into account the context of the sector in which the company provides services.

Regarding the landing page creation, both the campaign’s goal and the email’s context were taken into account. The goal was data theft and the context of the email suggested that when the victim clicked on “Abrir” it would open a file, since the employee would probably already be logged in. Additionally, when we analysed the Microsoft 365 login page, we could see that it was done in two steps: entering the email and then the password. From the moment the email is
inserted, we are redirected to a similar page that asks us for the password, but with curiosity, the email is “hidden” at the top, as shown in Figure 3.11.

Figure 3.11: Microsoft’s login page (original)

Using Gophish’s “Import Site” feature, we were able to import the Microsoft 365 login page and then make a few surgical modifications to it (Figure 3.12). The most relevant was the incorporation of a reference to an existing variable, `{{.Email}}`, when creating emails/landing pages in Gophish, which allows us to insert that variable into the landing page and have different content displayed, depending on the victim. That is, to increase the credibility of the landing page and the campaign as a whole, the moment the victim clicked on the email link, the login page would be opened with the indication of his/her own email and only requiring the insertion of the corresponding password.

Figure 3.12: Cloned and adapted landing page for data theft purposes

In the next steps, for the domain acquiring we made use of one of the techniques indicated in Section 2.2.3.1 when creating it, more precisely the addition technique, where the letter c was added to Microsoft’s domain, resulting in `microscoft`. When checking domain availability at Namecheap, we found most of the most used extensions all purchased, so we tried to look for other solutions. Given the wide use of the word `live` by Microsoft on several pages, it was cho-
sen to be the domain’s extension, resulting in the purchase of the domain \texttt{microsoft.live}, for 2.5€. The sending profile created was \texttt{noreply@microsoft.live}. SPF and DKIM records were created to ensure domain authentication and improve email deliverability. After all the steps mentioned above, we have in Table 3.8 all DNS records that were created for Campaign II.

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>\texttt{microsoft.live}</td>
<td>[VPS’ IP address]</td>
</tr>
<tr>
<td>CNAME</td>
<td>\texttt{www}</td>
<td>\texttt{microsoft.live}</td>
</tr>
<tr>
<td>CNAME</td>
<td>\texttt{login}</td>
<td>\texttt{microsoft.live}</td>
</tr>
<tr>
<td>TXT</td>
<td>\texttt{mailjet._domainkey}</td>
<td>k=rsa;p=(...)</td>
</tr>
<tr>
<td>TXT</td>
<td>\texttt{mailjet._ownership_token_piece}</td>
<td>[ownership_token_full]</td>
</tr>
<tr>
<td>TXT</td>
<td>\texttt{microsoft.live}</td>
<td>v=spf1 include:spf.mailjet.com(...)</td>
</tr>
</tbody>
</table>

Table 3.8: DNS records table for Campaign II

The last element to add was the SSL certificate, done the same way as mentioned in Campaign I.

Finally, with all elements created, configured and integrated, some tests were made to check the deliverability of the emails, the look of both the preview of the email in the inbox (Figure 3.13) and the email itself (Figure 3.14), verification of the landing page as well as the validation of the SSL certificate on it (Figure 3.15).

![Microsoft]( Rodrigol Oliveira partilhou um documento consigo Rodrigol Oliveira partilhou um documento consigo)

Figure 3.13: Email preview on the target’s inbox
3.4.2.3 Delivery

Since this step does not have any difference from what was performed in Campaign I, every relevant information is mentioned in Section 3.4.1.3.

3.5 Smishing integration in proposed framework

This section will present the integration of a text messaging tool (Twilio) into the designed framework, aiming to add the ability to conduct smishing campaigns and collect relevant metrics, in the same way as would be done in a phishing campaign. This integration is only possible by making a few changes to the Gophish tool, as this is the core tool for developing the campaigns and collecting metrics, with every other tool being linked to it.

It is important to emphasise that this section will only present the integration and what makes it possible and not any result since it has not been put into practice in environments other than test environments. After the 2 phishing campaigns had been carried out, a proposal to the client was made to carry out at least 1 smishing campaign, but the client’s response was negative.
3.5.1 Twilio tool & Gophish modification

3.5.1.1 Twilio

Twilio is a cloud communications platform allowing the integration of communication features into applications, sending and text messages (SMS) being one of them, through communication API. One of the key advantages of Twilio is its flexibility and scalability having more than 50,000 companies as clients and over 2 million developers as users. Furthermore, it was the chosen tool given previous experience with it and its API as well as its easiness to configure/use.

There are 3 steps to be able to send SMS messages in Twilio, them being [89]:

1. Sign up for/login to your Twilio account;
2. Acquire a phone number with SMS capabilities;
3. Send an SMS message via Twilio’s REST API.

It should be noted that, although every step to configure the framework was free of charge, except the domain purchase mentioned in Sections 3.4.1.2 and 3.4.2.2, that is not the case with the Twilio integration, or not forever. It is a pay-as-you-go type of pricing and while performing step 1 and creating a new account, the user may receive a certain amount of credits to spend while on a trial period. When this period ends, Twilio will no longer cover the usage costs. Both steps 2 and 3 have costs associated where, for instance, acquiring a Portuguese phone number has a monthly fee of 13.74€ versus a monthly fee of 1.05€ for an American one and sending SMS messages has its cost starting at 0.0072€ per message. Furthermore, all testing was made during the trial period so no operation was charged for this work.

3.5.1.2 Gophish

The Gophish phishing framework is written in the Go programming language, and its codebase is organized into different packages and directories following a typical Go project structure [32][33]:

- auth: Package auth implements the authentication in use for Gophish.
- config: Package config implements a global configuration to be used with Gophish (e.g. settings for the web server, email configurations, and other customizable options).
- context: Package context provides the ability to store request-scoped values on an http.Request instance.
- controllers: Package controllers is responsible for setting up the routing and controllers (http.Handlers) for Gophish.
- db: Gophish uses a database to store campaign data and results. This directory typically contains database models and code for interacting with the database.
• logger: Package logger contains a wrapper around logrus to support a centralized logging config.

• middleware: Package middleware is responsible for the definition/implementation of middleware functionality.

• models: Package models implements the types and structs needed in Gophish.

• static: Package static is responsible for storing static assets such as JavaScript, CSS, and images used in the web interface.

• templates: This directory typically contains HTML templates for rendering web pages in the application’s user interface.

• util: Package util provides misc utility functions for Gophish.

• webhook: Package webhook contains the functionality for handling outcoming webhooks.

• worker: Package worker contains the functionality for the background worker process.

After reading Gophish’s documentation and analysing every package and file, the directories and relevant files to the integration were identified [32][33]. Each file modification is presented in Appendix A where it can be observed that there were variables/attributes named “phone number” added to allow Gophish to ask for, receive and save the victim’s phone number (insert the phone number in groups for each target, show phone number in dashboards, details, results and events):

• Package models
  – File group.go listed in Appendix A, Listing A.1
  – File campaign.go listed in Appendix A, Listing A.2

• static
  – group.js (/static/js/src/app) listed in Appendix A, Listing A.4
  – campaign_results.js (/static/js/src/app) listed in Appendix A, Listing A.7

• templates
  – group.html listed in Appendix A, Listing A.3

We can observe the difference, in the “New Group Modal” after the modification, presented in Figure 3.16.
After modifying every file, a Python script was created to collect information from a built campaign in Gophish (e.g. victim’s name and phone number) through Gophish’s API to build the SMS message and send it through Twilio’s API. Python was the chosen programming language given that Gophish has a Python API and Twilio’s API also supports Python. The developed script is listed in Appendix A, Listing A.8. Furthermore, both the configuration file and the message file example mentioned in the Python script are also listed in Appendix A, Listing A.9 and Listing A.10, respectively.

The script execution output and the received SMS message are presented below in Figure 3.17 and Figure 3.18.
In Gophish we can observe that after the “victim” clicks on the malicious link and submits data, it automatically collects that information updating the main dashboard and details, presented in Figure 3.19 and Figure 3.20, proving that the integration between the developed framework and Twilio is possible and that the pilot test was successful.
3.5.2 Chapter overview

This chapter discussed the work done throughout the project, starting by identifying requirements/restrictions taking into account the project’s context, followed by a discussion presenting the proposed infrastructure for the project. Finally, the entire process that led to the creation of both campaigns was detailed, addressing the psychological nuances and the PPSE used in the emails. Additionally, the process required to create a campaign in the Gophish tool was described.

Finally, this chapter also presented the integration between the proposed framework and Twilio, a cloud communications platform that allows sending SMS messages via REST API. This integration was made through the creation of a Python script, making use of both Gophish’s and Twilio’s API. Furthermore, it was shown that the integration is possible. The pilot test was a success with the SMS message being delivered and Gophish being able to collect the relevant metrics.
Chapter 4

Results & Discussion

This chapter will present and discuss the results of Campaign I (presented in Section 3.4), as well as an explanation of what happened during Campaign II and what led to no results. Given this same situation, we will discuss the results expectation for Campaign II, taking into account similar campaigns previously conducted by EY and their results. Furthermore, before the results presentation and the discussion, a brief introduction will be made giving context on what metrics were collected and what data was used as a benchmark when it comes to phishing campaign results. Finally, we go through every work requirement, one by one, discussing if it was met or not.

As mentioned at the end of the previous chapter, since the Gophish phishing framework captures all events during a campaign, it was possible to collect the following metrics: number of emails sent, emails opened, clicked links, submitted data and emails reported. Furthermore, to understand the performance level of the client’s employees, the results obtained were compared with other data, such as those presented in Proofpoint’s annual phishing report [73], which is the industry benchmark report on the phishing topic and is based on millions of data points from multiple continents and from across 19 industry verticals, or those provided by Lance Spitzner from SANS Institute [81].

4.1 Campaign I

During the preparation for campaign I, the client provided a list of 95 targets, to whom spear-phishing emails were sent indicating the possibility of entering a raffle to win a weekend at a hotel.

The existing expectation for this campaign’s results was somewhat reserved due to the lack of information about the employees’ awareness regarding this subject, despite working for a company in the IT industry, and given the fact that this was the first campaign to be carried out with this client. Furthermore, given that it was a spear-phishing campaign different from the usual ones (e.g., invoices, human resources issues) and that greater care was taken in the choice and correct use of the PPSE to exploit the psychological vulnerabilities of the employees, we expected that
some employees would submit their credentials.

In Figure 4.1 we present the results obtained throughout the campaign, where it is possible to see that they exceeded the initial expectations and that it was a very successful campaign in the various metrics.

![Figure 4.1: Results from Campaign I (presented in Section 3.4)](image)

From general or industry-specific IT benchmarks [73][81], we were able to make some comparisons with the results obtained and draw some conclusions. Lance Spitzner, from SANS, state that typically organizations get a click rate of 25-30% the first time they do a phishing simulation, but also that this number may vary greatly depending on how targeted the phishing template is [81]. Although this was a spear-phishing campaign, with a phishing template targeted at a specific client, it resulted in a click rate of more than 45%, which translates into an increment of more than 15% of the indicated range. Furthermore, we can see a failure rate close to 32%, which may be the most alarming statistic out of these results. In Proofpoint’s annual report [73], it is indicated that the average failure rate both globally and specifically for the IT industry stands at 11%. The results obtained translate a failure rate almost 3 times higher than the average, which is a very positive result from the point of view of the campaign’s effectiveness and possibly of the use and chosen combination of PPSE in it. In addition to those indicated above, we also observed some more alarming statistics, which are indicated below the arrows in Figure 4.1:

- Close to 50% of the emails sent were opened;
- More than 90% of “openers” clicked the link;
- Close to 70% of “clickers” submitted data (company credentials);
- 0% report rate.

The last metric reflects possibly the most concerning statistic in these results, from the client’s perspective. A 0% report rate means that none of the employees reported the email to their IT team or help desk. This may be the most important metric when running a phishing simulation within an organization, for the simple fact that it represents the likelihood that the responsible teams, upon being informed, will intervene in an attempt to respond to the incident and mitigate the risk (e.g., blocking the malicious email domain) before another employee opens the email.
Halevi et al. [35] and Oliveira et al. [63] state in their studies that there is a gender-based difference in the susceptibility to phishing, with women being more susceptible to phishing and more likely to respond to a spear-phishing email. Given these studies, in an attempt to understand if the same would have happened in this campaign, all the events that occurred during the campaign were analysed and a distribution of them by gender was made, presented in Figure 4.2.

Looking at the “Total” column we can see a large disparity when it comes to the amount of employees of each gender in the sample, this could lead to a misinterpretation of the graph where male employees were the most susceptible to phishing. However, by analysing the data in percentages, we were able to observe that the statements in [35] and [63] are observed in this case as well:

- **Men**
  - 49.3% open rate
  - 43% click rate
  - 28.4% failure rate

- **Women**
  - 50% open rate
  - 50% click rate
  - 39.3% failure rate

Although the results of the first two metrics (email opened and clicked link) are quite similar or closer to each other, it is in the submission data where the biggest difference lies, which is in
line with the statement that women are more likely to respond to a spear-phishing email, translated into an increase of almost 11%.

Just like the mentioned studies in previous paragraphs, there are numerous others that attempt to make a relationship between susceptibility to phishing and genders, age groups, departments in the company. This was also one of the goals of the work, as well as later making a relationship between this data and that of human behavior, i.e. what would have caused that employee to fail the phishing simulation and how would this be related to gender, age group, department and awareness classification. To this end, the questionnaire present in Appendix B was created, in which each employee would be asked to fill in some information about themselves (e.g., gender, age group), as well as rate or give feedback on each of the campaigns and indicating what emotion/reaction they had when seeing the email for the first time (correlation with PPSE and psychological aspect of phishing). Unfortunately, since we did not get a response regarding the client’s approval/authorization for sharing the questionnaire with the employees who received the emails, it was not possible to proceed with these analyses.

4.2 Campaign II

As far as campaign II is concerned, it was not possible to collect any results because the emails ended up in the employees’ spam folder. Initially after getting the confirmation and authorization from the client to start the campaign on a certain date, all the necessary preparation was done (e.g., domain purchase, DNS records configuration). However, after everything was operational, the customer decided to postpone the phishing simulation’s launch date, subsequently giving notice of the new date at a very short notice. This incident and the period until the new date led to the previously purchased domain being blocked so it was necessary to prepare the campaign one more time. Having the new domain purchase been made with such short notice (less than 7 days) led to a significantly lower score regarding the “spammyness” of the emails (Mail Tester tool). This whole situation consequently led to emails being delivered to the employees’ spam folder, resulting in no results.

4.2.1 Two previously conducted campaigns

This section will then discuss the expectations that existed for this campaign based on a comparison with the emails and results of two other campaigns that also attempted to impersonate Microsoft, previously conducted by EY with other clients. The two campaigns used as comparison are present in Figures 4.3 and 4.4, with a security alert/suspicious activity being the theme of both.

Regarding the first campaign, it was thought of and defined as a generic phishing campaign with a lower degree of complexity, and the template presented in Figure 4.3 was created in which the link to change the password referred to a landing page similar to the one presented in Figure 3.12. It is then possible to verify the lower degree of complexity, due to the simplicity of the email.
and the lack of similarity to a real Microsoft email informing that there was an improper access on the person's account.

The results collected from this campaign were as follows:

- Email Sent - 521;
- Email Opened - 431 (close to 83% open rate);
- Clicked Link - 59 (over 11% click rate);
- Submitted Data - 12 (over 2% failure rate);

![Microsoft](image)

Figure 4.3: First campaign previously conducted by EY, trying to impersonate Microsoft, similarly to Campaign II (presented in Section 3.4)

On the other hand, the second campaign was also thought of and defined as a generic phishing campaign but with a slightly higher degree of complexity. Although the context of the campaign is the same as the previous one, it is possible to see a greater effort in trying to make the template (Figure 4.4) look very similar to an official Microsoft email.

The results collected from this campaign were as follows:

- Email Sent - 250;
- Email Opened - 45 (18% open rate);
- Clicked Link - 18 (over 7% click rate);
- Submitted Data - 10 (4% failure rate);
Figure 4.4: Second campaign previously conducted by EY, trying to impersonate Microsoft, similarly to Campaign II (presented in Section 3.4)

From the analysis of these two samples, both the templates and the results, a positive expectation was generated for what could be the end result of the second phishing simulation. Compared to the two templates, the one presented in Figure 3.4 is also based on an attempt to impersonate Microsoft, but it differs in that it is a new template when it comes to phishing emails, i.e., most of the times that an email impersonating Microsoft is received there is a high probability that it has the security alerts/suspicious activity as a theme. Additionally, the fact that it was very similar to an official Microsoft email that regularly appears in employees’ inboxes, especially at a time when working from home is prevalent, as well as the possible curiosity generated in employees when reading that someone had shared a file with them, gave reason to believe that it would have been a more successful campaign than the two presented above.

4.3 Discussion

After Campaign I was conducted and after analysing its results, having offered a first information regarding the awareness and susceptibility of the client’s employees to phishing, this same information also influenced the expectation generated. On the one hand the difference in technique used in each of the campaigns (spear-phishing vs. generic phishing), would always influence the difference in results, since spear-phishing campaigns are usually more effective. Due to the personalization for the specific customer and the greater focus on the use of PPSE throughout the email, it was expected that Campaign II would be less successful and that its results would show
some discrepancy compared to Campaign I. On the other hand, some of the metrics collected during Campaign I raised some concerns regarding the preparation and training of the client’s employees, such as the high failure rate and the 0% reporting rate. These same metrics led us to believe that despite being less successful than Campaign I, Campaign II would still have some data submitted.

As mentioned before, we were not able to collect responses to the questionnaire created with the aim of understanding the relationship between employees failing in the phishing campaign and their age, department, level of awareness, as well as the quality of each campaign, be it the presentation of the email in general, its text, credibility, what type of emotion/reaction was provoked in the victim when reading each email. Although we were not able to receive formal feedback in the form of questionnaire’s answers or a written and signed document by the client, we received verbal feedback from the client’s representative mentioning that both campaign’s templates were very well designed and built, which we believe is directly related to the campaign’s success.

Regarding the work requirements presented in Section 3.1 we will go through one by one and discuss if it was met:

**Functional**

1. This requirement was met by proposing a framework consisting of 6 different tools with the ability to create and conduct phishing simulations.

2. This requirement was met up to a certain point by proposing a framework with an SMTP relay, a reverse proxy and also SPF and DKIM records configured, helping to improve email deliverability and to bypass anti-spam controls. The second conducted phishing campaign did not bypass anti-spam controls due to events outside our control.

3.&4. Both these requirements were met by choosing Gophish as the phishing toolkit element in the proposed framework, as the toolkit offers the ability to not only create and conduct phishing campaigns but also collect relevant metrics and monitor every event (emails sent, emails opened, clicked links, submitted data, emails reported).

**Non-functional**

1. This requirement was met by proposing a framework consisting of only open-source/free solutions whenever possible, that is, it was not possible to acquire a domain for each campaign without having an associated cost.

2. This requirement was met by configuring both the phishing toolkit and the mail server on an EY’s existing backend server that initially was located on-premises and later migrated to a virtual machine in Microsoft Azure.
3. This requirement was met up to a certain point, both phishing campaign simulations were created and sent, although the email deliverability was not assured for the second one due to events outside our control.

4. This requirement was met by sending the phishing emails only to the employees present on the previously made available list (95 and 91 employees for the first and second campaigns, respectively).

4.3.1 Chapter overview

This chapter presents the results collected throughout Campaign I, using different metrics, as well as a discussion of them in comparison to phishing benchmarks around the world and in the client’s specific industry. A breakdown of the results by gender is also made in an attempt to verify claims in some studies that there is a difference in phishing susceptibility between the two genders. Next, Campaign II and the situation that led to it having no results is discussed. Despite this occurrence, the expectation that existed for the results of the campaign is discussed when based on two similar campaigns previously conducted by EY, as well as the results of Campaign I. Finally, each of the work requirements was discussed to clarify whether it was fulfilled or not.
Chapter 5

Conclusion & Future Work

This chapter presents an overview and summary of all the developed work and its main conclusions as well as possibilities for future work that could contribute to its improvement.

5.1 Conclusion

The work described in this document was developed as part of the Dissertação/Projeto em Segurança Informática course, which is a requirement for the Mestrado em Segurança Informática’s completion, at FCUL. Furthermore, it was developed during an internship at an external organization, EY. An EY’s client requested its service to assess the effectiveness of their cybersecurity awareness program and this ended up being the project that the work was based on. To do this, the client required two phishing campaign simulations, highlighting that they had a reduced budget for it and that they wanted to opt for open-source solutions. It was defined with the client that the first campaign would be a spear-phishing one, while the second would be a more generic (spam) one.

Given this context, the goals of the work were to design/develop a framework that met the client’s requirements to opt for open-source solutions and create an environment able to conduct phishing campaign simulations and collect relevant metrics. Another goal was to take advantage of the proposed framework and conduct two phishing campaign simulations with the incorporation of PPSE in each campaign’s email template. Then, analyse and compare both campaign’s results. Finally, another objective was to test a cloud communication platform’s integration in the proposed framework to allow smishing campaigns to be carried out.

The first phase of this project was dedicated to researching relevant literature on the topics of hacking, social engineering attacks, phishing, smishing, its relation with psychology and the human brain, principles of persuasion in social engineering, models and methodologies that mapped out a nefarious actor’s different steps to perform a cyberattack and also a study on these steps and its nuances.

The second phase was dedicated to the proposed framework’s development, based on a previous study of what kind of tools should be a part of it and what were the best ones not only by itself but integrated with others as well. Different phishing frameworks/toolkits, mail server softwares, email service providers, domain registrars and CDN service providers were analysed and
The third phase consisted of using the proposed framework to design, create and conduct phishing campaign simulations. A methodology based on the CKC and a targeted phishing campaign flow, that consisted on three steps was followed. These three steps are: Reconnaissance, Weaponization and Delivery. Furthermore, after both campaigns, its results were analysed and compared with industry benchmarks resulting in some alarming conclusions that can not be generalized given the work’s scope: security awareness is still a big issue even for IT industry organizations; targeted phishing is more effective than generic phishing; women are more likely to respond to a spear-phishing email. Finally, it was also tested the cloud communication platform integration in the proposed framework with success.

Regarding the work’s goals, the main three mentioned in Section 1.2 were achieved although some minor ones were not. As mentioned in Section 3.4.2, the second campaign could not be carried out due to some timing issues associated with the client. This meant that it was not possible to truly compare the first and second campaigns and draw conclusions about them. To try to tackle this issue, we ended up comparing the first campaign’s results with the results of two similar generic phishing campaign simulations conducted by EY with other clients.

Overall, the result of this work can be seen as positive as its contributions were a proposed phishing simulation framework adopted by EY that is still used, a proof of concept was performed regarding the ability to conduct smishing campaigns with the proposed framework and a new tool integration, raise awareness for phishing/spear-phishing/smishing attacks in by exposing alarming results and conclusions. Furthermore, the positive feedback that was received from the client verbally, as mentioned in 4.3, regarding the campaign’s template design, text and credibility.

5.2 Future Work

Regarding future work to be developed, the proposed framework must be subject to a continuous improvement process, that is, open-source tools and their features must be monitored to decide whether it makes sense to change any tool in a specific element of the framework.

Given that it was not possible to carry out the assessment correctly during the work, the same concept should be carried out in a new project with the following planning involving an employee training activity: study the organization and its environment, with the information collected, prepare attack scenarios, conducting the first campaign, receiving results and analyzing them, based on the results adapting a workshop to raise employee awareness on the topic, subsequently and after a period carrying out a second campaign. It should be noted that after a collaborator fails in one of the campaigns, they would immediately have access to a note or a short video explaining what had just happened and highlighting some points that they should have paid attention to to detect the fraudulent email. By following this planning it is possible to draw real conclusions about the company’s maturity level and its awareness program, as well as the evolution of employees from the first to the second campaign. Additionally, with an action immediately following the employee’s undesired action in the email (e.g. clicking on the link, submitting information,
downloading a file), be it a note or video, the employee would see the topic as an attempt to teach and help instead of feeling like the company is trying to catch you making a mistake.
Bibliography


[27] Rui da Rocha Ferreira. RDDoS: O renascer do ciberataque que explora o medo e que já faz vítimas em Portugal. EXAME INFORMÁTICA, 2021. https://visao.sapo.pt/exameinformatica/noticias-ei/mercados/2021-03-23-rddos-ciberataque-que-explora-o-medo-e-que-ja-faz-vitimas-em-portugal/?fbclid=IwAR2NM9qVzI4GDDAd7nuQM1_0dqlrX7UzYy8brMGJsz7XHxEWYxa2A9nl6E.


[90] Rossouw von Solms and Johan van Niekerk. From information security to cyber security.  


https://mailtrap.io/blog/postfix-sendmail-exim/.
Appendix A

Smishing integration (Gophish + Twilio)

Listing A.1: group.go

Listing A.2: campaign.go
Appendix A. Smishing integration (Gophish + Twilio)

Listing A.3: group.html

```html
<!--- Modal -->

<table id="targetsTable" class="table table-hover table-striped table-condensed">
  <thead>
    <tr>
      <th>First Name</th>
      <th>Last Name</th>
      <th>Email</th>
      <th>Phone Number</th>
      <th>Position</th>
      <th class="no-sort"></th>
    </tr>
  </thead>
  <tbody>
    <!-- Add table rows here -->
  </tbody>
</table>

function save(id) {
  $.each($("#targetsTable").DataTable().rows().data(), function (i, target) {
    targets.push(
      {first_name: unescapeHtml(target[0]),
       last_name: unescapeHtml(target[1]),
       email: unescapeHtml(target[2]),
       phone_number: unescapeHtml(target[3]),
       position: unescapeHtml(target[4])
    })
  })
}

function edit(id) {
  $.each(group.targets, function (i, record) {
    targetRows.push([escapeHtml(record.first_name),
                     escapeHtml(record.last_name),
                     escapeHtml(record.email),
                     escapeHtml(record.phone_number),
                     escapeHtml(record.position),
                     '<span style="cursor:pointer;"><i class="fa fa-trash-o"></i></span>'
    ])
  });

  done: function (e, data) {
    $.each(data.result, function (i, record) {
      addTarget(
        record.first_name,
        record.last_name,
        record.email,
```
function addTarget(firstNameInput, lastNameInput, emailInput, phoneInput, positionInput) {
// Create new data row.
  var email = escapeHtml(emailInput).toLowerCase();
  var newRow = [
    escapeHtml(firstNameInput),
    escapeHtml(lastNameInput),
    email,
    escapeHtml(phoneInput),
    escapeHtml(positionInput),
    '<span style="cursor:pointer;"><i class="fa fa-trash-o"></i></span>'
  ];
}

$(document).ready(function () {
  addTarget(
    $('#firstName').val(),
    $('#lastName').val(),
    $('#email').val(),
    $('#phoneNumber').val(),
    $('#position').val());
});
Appendix A. Smishing integration (Gophish + Twilio)

Listing A.5: result.go

```
func AddEvent(e, r.CampaignId) {
    return e, nil
}
(...)
```

Listing A.6: campaign_results.html

```
function renderTimeline(data) {
    record = {
        "id": data[0],
        "first_name": data[2],
        "last_name": data[3],
        "email": data[4],
        "phone_number": data[5],
        "position": data[6],
        "status": data[7],
        "reported": data[8],
        "send_date": data[9]
    }
    results = '<div class="timeline col-sm-12 well well-lg">
        <h6>Timeline for ' + escapeHtml(record.first_name) + ' ' + escapeHtml(record.last_name) + '</h6>
        <span class="subtitle">Email: ' + escapeHtml(record.email) + '<br>Phone Number: ' + escapeHtml(record.phone_number) + '<br>Result ID: ' + escapeHtml(record.id) + '</span>
    </div>
    <div class="timeline-graph col-sm-6">
        (...)
    </div>
    (...)
}
```

function load() {
```
Appendix A. Smishing integration (Gophish + Twilio)

Listing A.7: campaign_results.js

```javascript
import configparser
import requests
from twilio.rest import Client  # Twilio API

CONFIG = 'config.ini'

# Returns targets information from Gophish’s campaign
def getTargetsFromGophish(api_key, campaign_id, gophish_admin_server_ip):
    print('[*] Getting campaign details from Gophish...')
    results = requests.get('https://' + gophish_admin_server_ip + ':3333/api/campaigns/' + str(campaign_id) + '/results/?api_key=' + api_key, verify=False).json()['results']
    targets = []
    for result in results:
        victim = {}
        victim['rid'] = result['id']
        victim['first_name'] = result['first_name']
        victim['last_name'] = result['last_name']
        victim['phone_number'] = result['phone_number']
        victim['status'] = result['status']
        victim['reported'] = moment(result.send_date).format('MMMM Do YYYY, h:mm:ss a')
        targets.append(victim)
    return targets

# Replaces, with each victim’s data, the Gophish template references in the message text
def setCustomSms(account_sid, auth_token, msg_service_id, message_body, targets, link):
    client = Client(account_sid, auth_token)
    for victim in targets:
        custom_message = message_body.replace('{{.FIRST_NAME}}', victim['first_name'])
        custom_message = custom_message.replace('{{.LAST_NAME}}', victim['last_name'])
        if link[-1] != '/':
            link = link + '/'
        result = client.messages.create(to=victim['phone_number'], from_='Gophish', body=custom_message, link=link, media_url=link, message_type='sms', status_callback='https://' + gophish_admin_server_ip + ':3333/api/campaigns/' + str(campaign_id) + '/results/report?api_key=' + api_key)
```

Listing A.7: campaign_results.js
custom_message = custom_message.replace('{{.URL}}', link + '?rid=' + victim['rid'])
print('[+] Name: ' + victim['first_name'] + ' ' + victim['last_name'] + ' | Phone number: ' + victim['phone_number'])
twilioSendSms(client, msg_service_id, custom_message, victim)
print('[*] Launching smishing campaign...')

# Launch smishing campaign sending all SMSes
def twilioSendSms(client, msg_service_sid, message_body, target):
    message = client.messages.create(messaging_service_sid=msg_service_sid, body=message_body, to=target['phone_number'])
    print(message.sid)

def main():
    config = configparser.ConfigParser()
    config.read(CONFIG)

    # read values from GOPHISH section
    gophish_api_key = config.get('GOPHISH', 'GOPHISH_API_KEY')
    gophish_admin_server_ip = config.get('GOPHISH', 'GOPHISH_ADMIN_SERVER_IP')
    gophish_landing_page_url = config.get('GOPHISH', 'GOPHISH_LANDING_PAGE_URL')
    gophish_campaign_id = config.get('GOPHISH', 'GOPHISH_CAMPAIGN_ID')

    # read values from TWILIO section
    twilio_account_sid = config.get('TWILIO', 'TWILIO_ACCOUNT_SID')
    twilio_auth_token = config.get('TWILIO', 'TWILIO_AUTH_TOKEN')
    twilio_msg_service_id = config.get('TWILIO', 'TWILIO_MSG_SERVICE_SID')
    twilio_msg_filepath = config.get('TWILIO', 'MSG_FILEPATH')

    targets = getTargetsFromGophish(gophish_api_key, gophish_campaign_id, gophish_admin_server_ip)
    message_body = open(twilio_msg_filepath, 'r').read()
    setCustomSms(twilio_account_sid, twilio_auth_token, twilio_msg_service_id, message_body, targets, gophish_landing_page_url)

if __name__ == '__main__':
    main()
Caro {{.FIRST_NAME}} {{.LAST_NAME}},

Queremos confirmar a sua inscrição no sorteio que a [Empresa] está a organizar.

Se ainda não se inscreveu, tem até amanhã para o fazer a partir deste link: {{.URL}}.

Obrigado e boa sorte!

[Empresa]

---

Listing A.10: Message file example (message.txt), mentioned in A.9
Appendix B

Phishing study assessment/feedback questionnaire

Phishing study questionnaire

This anonymous questionnaire addresses the assessment of the phishing study based on the conducted phishing campaigns. It consists of 3 sections with a total of 16 questions that focus on gathering some information and evaluating both campaign's quality and effectiveness.

The results of this questionnaire are anonymous, as mentioned above, and will be used in the master’s thesis of the student Paulo Rugeiro, student of the M.Sc. in Cybersecurity at Faculdade de Ciências da Universidade de Lisboa.

This questionnaire takes about 5 minutes to complete.

*Required

Employee Information

In this section you are asked to fill in some information about yourself.

1. Gender *

Mark only one oval.

☐ Male
☐ Female
☐ Other: 

2. Age *

Mark only one oval.

☐ 18-25
☐ 26-40
☐ 41-60
☐ >60

3. Department *

Your department in the company

Figure B.1: Part I of the questionnaire
Appendix B. Phishing study assessment/feedback questionnaire

Figure B.2: Part II of the questionnaire
8. Human behaviour *
   What were your emotions/reactions when you first saw the phishing email?

   *Tick all that apply.*
   - [ ] Fear
   - [ ] Duty/Obligation
   - [ ] Curiosity
   - [ ] Trust
   - Other: ___

9. Best aspects *
   Example: trustworthy elements

   __________________________
   __________________________
   __________________________
   __________________________
   __________________________

10. Needed Improvements *
    Example: phishing email concept/templates

    __________________________
    __________________________
    __________________________
    __________________________
    __________________________

**Phishing campaign feedback (Hotel weekend giveaway)**

This section's objective is to gather feedback about the conducted phishing campaign regarding the hotel weekend giveaway.

Figure B.3: Part III of the questionnaire
11. **Overall campaign quality/credibility**
   Phishing email, landing page
   
   *Mark only one oval.*
   
   1  2  3  4  5  
   Very low  □  □  □  □  □  Very high  

12. **Campaign’s appearance quality/credibility**
   Phishing email, landing page
   
   *Mark only one oval.*
   
   1  2  3  4  5  
   Very low  □  □  □  □  □  Very high  

13. **Campaign’s grammar quality**
   Phishing email, landing page
   
   *Mark only one oval.*
   
   1  2  3  4  5  
   Very low  □  □  □  □  □  Very high  

14. **Human behaviour**
   What were your emotions/reactions when you first saw the phishing email?
   
   *Tick all that apply.*
   
   [ ] Fear
   [ ] Duty/Obligation
   [ ] Curiosity
   [ ] Trust
   Other: □  

---

Figure B.4: Part IV of the questionnaire
15. **Best aspects** *
   Example: trustworthy elements

16. **Needed improvements** *
   Example: phishing email concept/template

Figure B.5: Part V of the questionnaire