INSTITUTO SUPERIOR TÉCNICO

Departamento de Engenharia Informática

Forensics Cyber Security

MEIC / METI 2023-2024 – 1st Period

Digital Forensics Report

Diogo Pereira, N110996; **João Gonçalo Santos, N110947;** **João Matos N110846**

1 Based on your analysis of the documents, can you deduce the likely identity of the owner of this pen drive? Justify your answer with relevant findings.

Durante a nossa investigação, foram encontrados vários nomes, em diferentes contextos. Mais precisamente, nos ficheiros *Social.png, Rialva.png* e *Tagus.png*, ao executar-se o comando ***exiftool***, é possível verificar o atributo ***Author*** com o valor **César Silva Ferro**. No entanto, este trata-se de um atributo manipulável, pelo que a sua credibilidade deve ser posta em causa. Outro nome encontrado, foi, a partir do comando ***pdfinfo***, aplicado ao ficheiro *waste-of-time*, Nuno Santos, descrito na propriedade ***Author*** deste ficheiro. Uma vez mais, não consideramos este nome relevante, até porque, o facto de ser inconsistente com o *Author* Cesar Silva Ferro mencionado acima perpetua a falta de credibilidade deste. Por fim, e sendo esta a nossa análise final, achamos que a *pen-drive* pertence a Eva Rocha, fundamentado pelo facto de existir um ficheiro escondido com o *Summary Account* desta. Uma vez que se trata de um ficheiro de enorme privacidade/exclusividade, é considerado como elemento fundamental para afirmar sobre quem é o proprietário da *pen-drive*. Outro fator que motiva esta decisão, é o facto da *pen-drive* ter sido encontrada numa sala da faculdade do Instituto Superior Técnico, a qual, a indivídua Eva Rocha é funcionária (com base no *Summary Account* mencionado acima, dado o salário de 4 100€ recebido via esta instituição).

2 Were there any concealed artifacts within the provided files? If so, detail how these artifacts were embedded and your methodology to extract them.

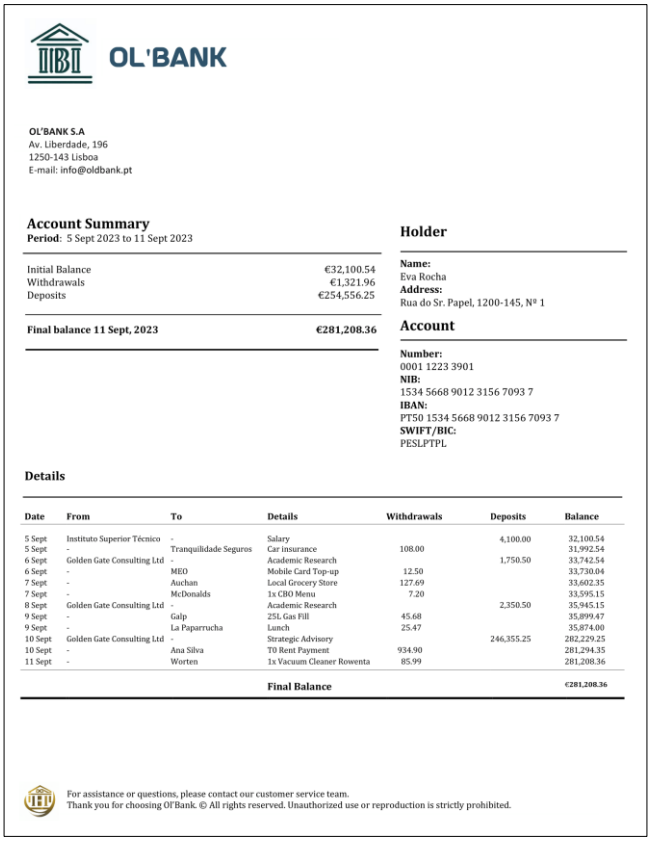
- **Ficheiro:** waste-of-time
  - Relativamente a este ficheiro, o processo de verificação realizado foi o seguinte:
    1. Executou-se o comando ***file waste-of-time***, e verificou-se que o mesmo se tratava de um ficheiro PDF.
    2. Posto isto, utilizou-se o comando ***binwalk waste-of-time***, de forma a verificar possíveis imagens escondidas. De facto, foi verificada a presença de uma imagem PNG (abaixo representada a utilização do *binwalk*).

```
DECIMAL       HEXADECIMAL     DESCRIPTION
--------------------------------------------------------------------------------
0             0x0             PDF document, version: "1.4"
552           0x228           Zlib compressed data, default compression
27753         0x6C69          Zlib compressed data, default compression
232628        0x38CB4         Zlib compressed data, default compression
264272        0x40850         Zlib compressed data, default compression
370909        0x5A8DD         PNG image, 1703 x 2203, 8-bit colormap, non-interlaced
```

Figura 1 - Output binwalk waste-of-time

3. De forma a extrair a imagem PNG, foi utilizado o comando ***foremost waste-of-time***. A partir deste, é possível verificar uma pasta ***output*** que contém os ficheiros PDF e PNG listados no ***ponto 2.*** O ficheiro PNG escondido representa o *Account Summary* da indivídua Eva Rocha no período de 5 de Setembro a 11 de Setembro de 2023, no *OL'BANK*.

4. ***SHA-256 value***:
1f70fc60e1c23b842e3278846e29c5fb66ed43c1d6ac3ccfe408f53acf005edb
00000724.png



OL'BANK

OL'BANK S.A
Av. Liberdade, 196
1250-143 Lisboa
E-mail: info@oldbank.pt

Account Summary
Period: 5 Sept 2023 to 11 Sept 2023

Initial Balance	€32,100.54
Withdrawals	€1,321.96
Deposits	€254,556.25
Final balance 11 Sept, 2023	**€281,208.36**

Holder

Name:
Eva Rocha
Address:
Rua do Sr. Papel, 1200-145, Nº 1

Account

Number:
0001 1223 3901
NIB:
1534 5668 9012 3156 7093 7
IBAN:
PT50 1534 5668 9012 3156 7093 7
SWIFT/BIC:
PESLPTPL

Details

Date	From	To	Details	Withdrawals	Deposits	Balance
5 Sept	Instituto Superior Técnico	-	Salary		4,100.00	32,100.54
5 Sept	-	Tranquilidade Seguros	Car insurance	108.00		31,992.54
6 Sept	Golden Gate Consulting Ltd	-	Academic Research		1,750.50	33,742.54
6 Sept	-	MEO	Mobile Card Top-up	12.50		33,730.04
7 Sept	-	Auchan	Local Grocery Store	127.69		33,602.35
7 Sept	-	McDonalds	1x CBO Menu	7.20		33,595.15
8 Sept	Golden Gate Consulting Ltd	-	Academic Research		2,350.50	35,945.15
9 Sept	-	Galp	25L Gas Fill	45.68		35,899.47
9 Sept	-	La Paparrucha	Lunch	25.47		35,874.00
10 Sept	Golden Gate Consulting Ltd	-	Strategic Advisory		246,355.25	282,229.25
10 Sept	-	Ana Silva	T0 Rent Payment	934.90		281,294.35
11 Sept	-	Worten	1x Vacuum Cleaner Rowenta	85.99		281,208.36
			Final Balance			**€281,208.36**

For assistance or questions, please contact our customer service team.
Thank you for choosing OlBank. © All rights reserved. Unauthorized use or reproduction is strictly prohibited.

Figura 2 - Account Summary Eva Rocha

5. **Forma de esconder os artefactos:** Uma vez que este artefacto foi extraído a partir do comando *foremost*, pode ser deduzido que o artefacto foi escondido a partir de uma alteração na estrutura do ficheiro PDF, tornando-se visível aquando executado o comando *binwalk*.

- **Ficheiro:** sporting_anthem
    - Relativamente a este ficheiro, o processo de verificação realizado foi o seguinte:
        1. Primeiramente, executou-se o comando ***file* sporting_anthem**, verificou-se que este representava um ficheiro ***RIFF (little-endian) data, WAVE audio, Microsoft PCM, 16 bit, stereo 88200 Hz***.

2. Uma vez que o ficheiro representa um áudio, ouvimo-lo de forma a analisar o seu conteúdo. Foi notado um ruído ao longo do áudio, pelo que se continuou a investigar.
3. Após pesquisa, chegou-se à conclusão de que seria necessário utilizar um programa específico de áudio de forma a analisar com maior detalhe o espectrograma deste, pois, de facto, poderia evidenciar a razão do ruído verificado. Instalou-se o programa ***Audacity*** e, a partir deste, acedeu-se a ***Options > View Spectogram > Spectogram Settings***: definiu-se o intervalo de frequência ***[10000Hz; 40000Hz]***. Posto isto, tornou-se possível verificar uma mensagem escondida com o seguinte conteúdo: '**1683461 – Look up!'.** Abaixo a representação da mensagem encontrada.

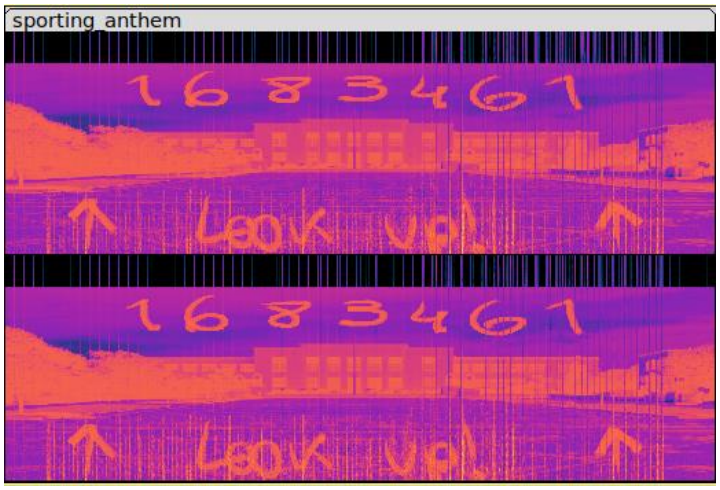


Figura 3 - Mensagem escondida no ficheiro sporting_anthem

4. ***SHA-256 value*** do ficheiro txt com o código descoberto na ***Figura 3***: 3c3eb82d92c24fe88da1d900665c1463cafc58ef7c9b59c735ba19f4b35990a6 message_sporting_anthem.txt

5. **Forma de esconder os artefactos:** A partir de um editor de imagem, foi criada a mensagem acima mencionada. Esta foi renderizada como áudio e, também, guardada como ficheiro áudio. Depois, este foi adicionado, paralelamente, ao áudio fornecido, explicando o ruído existente.

- **Ficheiros:** Social.png, Rialva.png, Tagus.png
    - Relativamente a estes ficheiros, o processo de verificação realizado foi o seguinte:
        1. Verificou-se, a partir do comando ***file***, que de facto os ficheiros mencionados são, efetivamente, ficheiros PNG.
        2. Consequentemente, executou-se, para todos os ficheiros mencionados, o comando ***binwalk***, no entanto, nenhuma informação relevante foi retirada.

3. Posto isto, de forma a verificar a *meta data* dos ficheiros mencionados, executou-se o comando ***exiftool***. É possível verificar, para todos os ficheiros, que

   - ***Tittle:*** *Food Review*,
   - ***Author:*** *Cesar Silva Ferro*.

   E, relativamente aos ficheiros ***Social.png*** e ***Tagus.png***, estes contêm

   - ***Description*****:** http://www.pdf-tools.com

   No entanto, foi o atributo ***Web Statement*** que achámos ser mais relevante, dada a sua existência e tamanho. Após alguma análise, foi possível verificar que este se tratava de texto na forma de base64.

4. Inicialmente, verificou-se, individualmente, se o texto em base64 poderia significar algo relevante. Apenas quando se tentou transformar o texto numa imagem foi possível obter resultados relevantes. Mais precisamente, o atributo *Web Statement* da imagem ***Social.png*** resultou num ficheiro PNG representativo de uma planta da área adjacente à **Casa da Moeda**, mais precisamente, referente ao Piso -1 desta, ou parte.

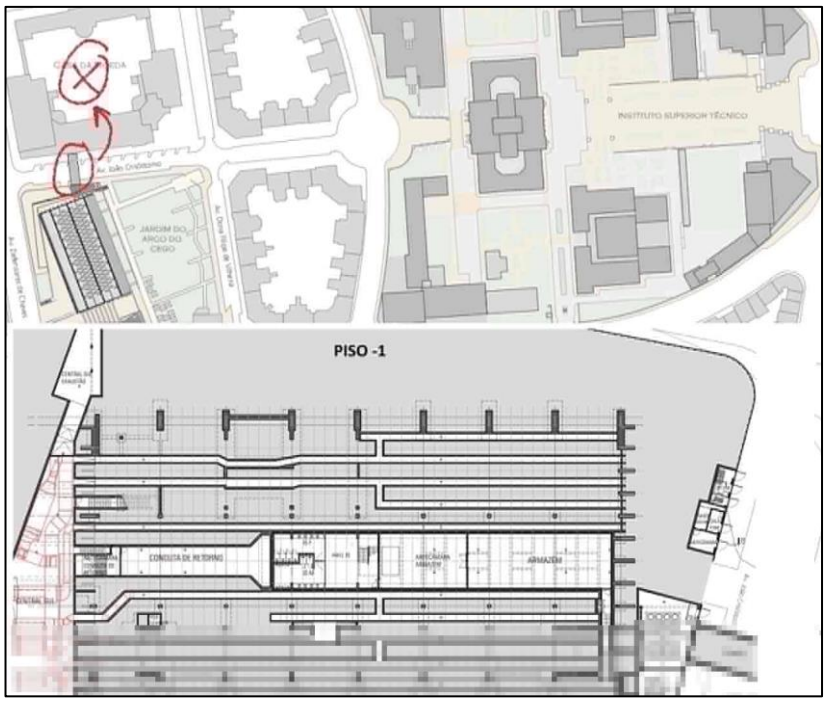


Figura 4 - Imagem obtida após descofidicação do Web Statement referente ao ficheiro Social.png

5. ***SHA-256 value*** do ficheiro referente à Figura 4: 8d5117114353f5399dbad61534902c1eee97505da767167ed85a8c14d67d1656 evidence1.png (*evidencia intermédia, por razões relativas ao tamanho máximo do ficheiro de submissão aceite, este foi removido do zip, mas mantido no relatório)

6. No entanto, é possível verificar uma distorção na parte inferior da imagem, pelo que se seguiu com a investigação. Uma vez que os três ficheiros PNG

mencionados contém o atributo *Web Statement*, considerou-se a existência de um possível padrão. Inicialmente, guardou-se o texto referente aos três atributos em ficheiros de texto, separadamente. Após tal, utilizou-se o comando ***cat***, de forma a concatenar o conteúdo dos ficheiros. Primeiramente, por ordem alfabética (R(ialva) - S(ocial) - T(agus)), uma vez que se trata de uma sequência exata existente no abecedário, e, de seguida, descodificar o texto na forma de base64 para imagem, mas sem sucesso. Consequentemente, executou-se o mesmo processo, mas agora, tentou-se ordenar pela ordem de classificação dos pratos (conteúdo relativo à imagem), tornando-se assim possível extrair uma imagem sem distorção. Abaixo a sua representação.

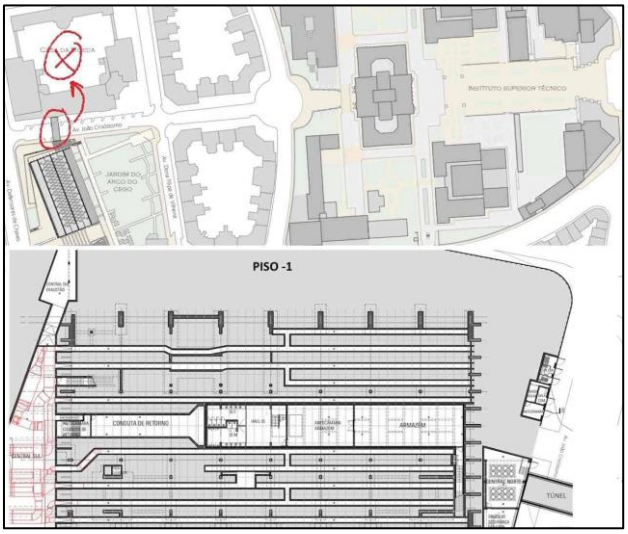


Figura 5 - Imagem extraída a partir da concatenação dos atributos Web Statement relativos às imagens Social.png, Rialva.png e Tagus.png

7. ***SHA-256 value*** do ficheiro referente à Figura 5:
   ad962cbd8f1d558d6e3cb8a46e88793500c078a029682b3ead703d1baf9ffa84
   evidence2.png

8. **Forma de esconder os artefactos:** Foi, inicialmente, realizada a conversão da imagem para texto na forma de base64. Este foi dividido em três e, separadamente, cada uma das três partes do texto, foi definida como atributo do *Web Statement* na meta-data das imagens, separadamente.

- **Ficheiro:** Cool_stuff.mp4
  - Relativamente a este ficheiro, o processo de verificação realizado foi o seguinte:
    1. Verificou-se que, de facto, se tratava de um ficheiro mp4.
    2. Analisou-se o conteúdo do vídeo e, reparou-se que a imagem relativa a Sua Excelência, o Senhor Presidente da República Portuguesa Professor Doutor

Marcelo Rebelo de Sousa, se encontrava diferente da, por nós pesquisada, original. Com isto, retirou-se o *frame* relativo a esta imagem, no caso, o último com o comando ***ffmpeg -sseof -3 -i Cool_stuff.mp4 -update 1 -q:v 1 last.png***.

3. De forma a analisar melhor a imagem, recorreu-se ao website https://stegonline.georgeom.net/ e, via funcionalidade deste, confirmou-se que a imagem estaria adulterada. Abaixo a imagem representativa.

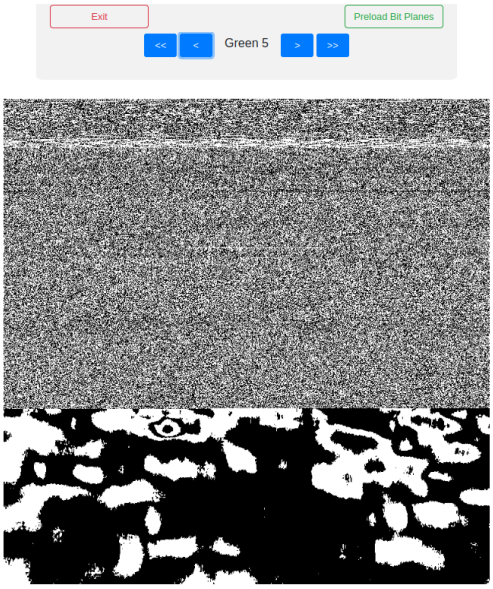


Figura 6 - Bitmap verde do 5º bit menos significativo

4. Uma vez confirmada a existência de algo suspeito, executaram-se os comandos ***binwalk*** e ***foremost***, mas sem sucesso. Assim sendo, prosseguiu-se com a investigação de deteção de esteganografia. De forma a encontrar possíveis ficheiros escondidos, utilizou-se o comando ***zsteg --all last.png*** (sendo ***last.png*** a imagem extraída no **ponto 2.**), resultando, em parte, no output abaixo representado.

Figura 7 - Output após utilização da ferramenta zsteg

5. Como se verifica acima, é possível notar a existência de ficheiro JPEG embebido na imagem em análise. De forma a extraí-la, utilizou-se o comando ***zsteg -E b6,g,lsb,xy last.png > evidence.jpeg.*** Com sucesso, verificou-se que o ficheiro JPEG escondido representa detalhes sobre o túnel que dá acesso à Casa da Moeda.

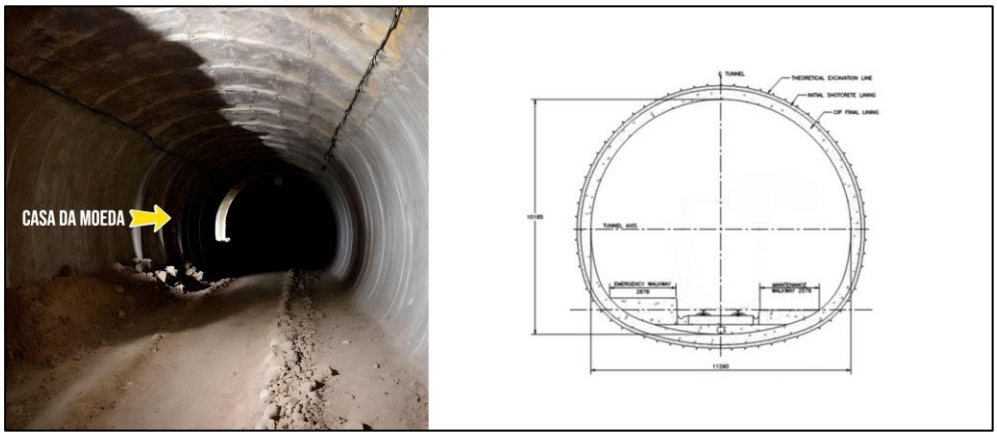


Figura 8 - Imagem extraída a partir do último frame do vídeo

6. ***SHA-256 value*** do ficheiro referente à Figura 8:
fa76deb62acc6dfb2fa0ea3d344d49e37711510d5efd160cca31363c6eb01d55
evidence.jpeg

7. **Forma de esconder os artefactos:** Foram alterados os bytes da componente verde da imagem original, de forma que os seis bits menos significativos correspondam a seis bits da imagem escondida, progressivamente, até ao fim desta.

- **Ficheiro:** report.docx
    - Relativamente a este ficheiro, o processo de verificação realizado foi o seguinte:
        1. Primeiramente, executou-se o comando ***file report.docx***, e verificou-se que este representa *ASCII text.*
        2. Em segunda instância, utilizou-se o comando ***ghex report.docx*** e verificou-se que o ficheiro é composto apenas por caracteres referentes ao alfabeto hexadecimal. O que, consequentemente, levou à conversão deste para a sua forma original em binário a partir do comando ***xxd -r -p report.docx > extracted_report.bin***.
        3. Uma vez extraído o ficheiro ***extracted_report.bin***, verificou-se que este terminava com os caracteres '==', sendo tal um indicador de estarmos perante um texto na forma de base64. Assim sendo, utilizou-se o comando

base64 -d extracted_report.bin > decoded.bin, conseguindo assim extrair o ficheiro **decoded.bin**.

4. De forma a verificar o tipo de ficheiro extraído, executou-se o comando ***file*** **decoded.bin** e concluiu-se ser um ficheiro do tipo **Z4 compressed data**.
5. Uma vez conhecido o tipo de ficheiro, prosseguiu-se ao *unzip* deste, a partir do comando ***lz4 -d decoded.bin evidence_zip***.
6. ***SHA-256 value*** do ficheiro referente ao ficheiro LZ4 extraído:
   98aedee8f8ad66f8a54295520a7b42291ec03fe9496722dc767756c4aa6b3cfd
   evidence_zip
7. Após tentativa de extração, esta anunciava que o ficheiro zip extraído estaria protegido por uma password. Uma vez que nos foi fornecido o ficheiro **BdC_on_the_beat**, sendo este composto por enumeras palavras, foi desde logo suspeitado que este poderia ser utilizado como *wordlist*, de forma a descobrir a password através da ferramenta ***john***. Mais precisamente, criou-se um ficheiro txt, com o conteúdo do ficheiro **BdC_on_the_beat,** mas, agora, separando cada palavra por *line breaks.* Com o objetivo final definido (descobrir a palavra-passe), foi utilizado o comando ***zip2john evidence_zip.zip > hash***, de forma a obter o *hash* da password do zip em análise e, consequentemente, utilizado o comando ***john --wordlist=wordlist.txt hash*** (sendo wordlist.txt o ficheiro composto pelo conteúdo do ficheiro **BdC_on_the_beat**, mas com cada palavra separada por *line break* e hash o ficheiro relativo ao output do comando *zip2john* acima descrito), de forma a descobrir efetivamente qual a password correta. Posto isto, foi possível verificar que a password que dá acesso aos ficheiros contidos no zip em análise é `(Three-time-champion)`.
8. Após ter sido concedido acesso aos ficheiros que compõem o zip em análise, foi verificado, a partir do comando ***ghex corrupted.pdf***, que o ficheiro não se tratava de um PDF, mas sim de um possível ficheiro MP4, dado o respetivo cabeçalho observado. No entanto, após alteração da extensão do ficheiro de *.pdf* para *.mp4*, este continuava corrompido. Depois de uma análise mais minuciosa, verificou-se que o cabeçalho não estava em conformidade com um cabeçalho correto de ficheiro MP4.

Figura 9 - Cabeçalho ficheiro corrupted.pdf

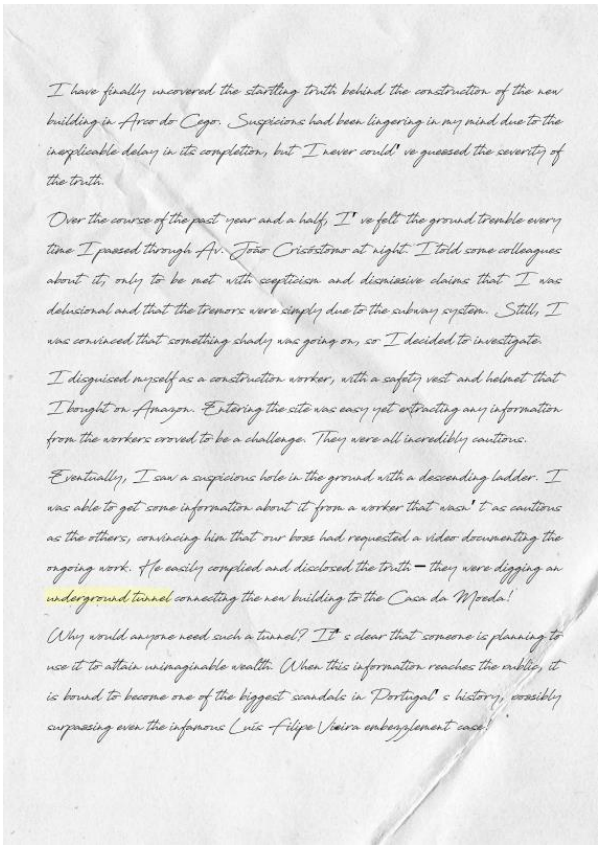
Mais precisamente, faltariam quatro caracteres ao cabeçalho na parte relativa ao ***'isom'***, isto porque, comparando com um cabeçalho MP4 válido, é possível verificar que o valor correto deste, é, de facto, '***ftypisom'***. A partir da ferramenta ***ghex*** foi possível realizar esta alteração, obtendo assim um ficheiro MP4 válido.

9. ***SHA-256 value*** do ficheiro referente ao ficheiro MP4 extraído:
   8f9a03d13221bf0477f8c7d178bcfbdd94c3ef22df05dcee05cc3962c82ff11c
   corrupted.mp4

10. ***SHA-256 value*** do ficheiro referente à imagem do bolo:
    8112eada7a480d85ef8b4c43010bda2192dd37dbaca91abafde25006d7397d7c
    grandmas_cake.png
11. ***SHA-256 value*** do ficheiro referente ao ficheiro txt com a receita:
    a4b09747a56a8a3b2670f205541498699fe4157b807e8327bd91182a6bfaf649
    grandmas_recipe.txt
12. ***SHA-256 value*** do ficheiro referente ao jpeg do tesouro:
    96e434b503d68822a03ad2886e243dbbd6d6723b661e29d4025a54b681d5ec2e
    my_fortune.jpeg
13. **Forma de esconder os artefactos:** Foi alterado o cabeçalho do ficheiro MP4, bem como a sua extensão. Todos os ficheiros associados foram comprimidos num ficheiro ZIP, protegido por uma *password*. A *password* foi escondida no ficheiro BdC_on_the_beat. O ficheiro ZIP foi, inicialmente, codificado em base64 e, posteriormente, foi calculado o hexadecimal deste. Este valor foi guardado num ficheiro ASCII, cuja extensão foi alterada para ***.docx***.

- **Ficheiro:** logo.png
    - Relativamente a este ficheiro, o processo de verificação realizado foi o seguinte:
        1. Verificou-se discrepâncias na cor de fundo do ficheiro.
        2. Utilizou-se uma ferramenta online para verificar as possíveis diferenças existentes nos valores RGB da imagem, e concluiu-se que estes diferiam em até 5 bits (https://imagecolorpicker.com/).
        3. Uma vez retirada a conclusão do **ponto 2.**, utilizou-se outra ferramenta online para fazer a análise esteganográfica e, consequentemente, extrair os 5 bits menos significativos (https://stegonline.georgeom.net/upload). Com isto, foi possível extrair um ficheiro PDF, mas corrompido.
        4. Reparou-se que tanto a parte azul, como a parte final (retângulo a branco no inferior da imagem) não tinham nada escondido, por tal, desenvolveu-se um programa python (***extract_info.py***) com o objetivo de extrair os 5 bits menos significativos, mas ignorando a parte azul (componentes red == 0 && green == 159 && blue == 227) e o retângulo no inferior da imagem a branco, a partir da coordenada (30, 1372) em direção à direita e baixo, obtida através de uma ferramenta online (https://pt.pixspy.com/).

5. Após executado o programa realizado, foi possível extrair a sexta prova escondida. Abaixo o ficheiro encontrado.



I have finally uncovered the startling truth behind the construction of the new building in Arco do Cego. Suspicions had been lingering in my mind due to the inexplicable delay in its completion, but I never could've guessed the severity of the truth.

Over the course of the past year and a half, I've felt the ground tremble every time I passed through Av. João Crisóstomo at night. I told some colleagues about it, only to be met with scepticism and dismissive claims that I was delusional and that the tremors were simply due to the subway system. Still, I was convinced that something shady was going on, so I decided to investigate.

I disguised myself as a construction worker, with a safety vest and helmet that I bought on Amazon. Entering the site was easy yet extracting any information from the workers proved to be a challenge. They were all incredibly cautious.

Eventually, I saw a suspicious hole in the ground with a descending ladder. I was able to get some information about it from a worker that wasn't as cautious as the others, convincing him that our boss had requested a video documenting the ongoing work. He easily complied and disclosed the truth — they were digging an underground tunnel connecting the new building to the Casa da Moeda!

Why would anyone need such a tunnel? It's clear that someone is planning to use it to attain unimaginable wealth. When this information reaches the public, it is bound to become one of the biggest scandals in Portugal's history, possibly surpassing even the infamous Luís Filipe Vieira embezzlement case!

Figura 10 - Ficheiro extraído do logo.png

6. ***SHA-256 value*** do ficheiro referente à Figura 10:
a6a005c97a758f04220fe161c269a2ac229d7fe3b6fbfc88e40d6d8675be8fd6 logo.pdf

7. **Forma de esconder os artefactos:** Foram alterados os bytes das componentes vermelha, verde e azul da imagem original, de forma que os cinco bits menos significativos correspondam a cinco bits do ficheiro PDF escondido, progressivamente, até ao fim deste. Com a particularidade de ignorar os pixéis com o valor RGB inicial (0, 159, 227), correspondente à cor azul presente no logotipo.

3 Based on the secrets you recovered, is there any indication that the pen drive was intended to spread malware or present a specific security threat? If there's no direct evidence of malicious intent, how would you interpret the data? Formulate a hypothesis regarding their purpose and justify it using the content of the recovered secrets.

Após obtenção dos seis artefactos escondidos e, consequentemente, análise dos mesmos, é possível afirmar que não foram detetados quaisquer indícios de *malware*. Relativamente a potenciais *security threats*, a nossa análise sugere que, para a faculdade em si, não existe, pelo menos de forma direta, uma ameaça associada. Mais precisamente, em nenhum dos ficheiros (visíveis e escondidos) foi encontrado código executável, ou qualquer tipo de diminuição na performance dos recursos da máquina (a responsável pela análise forense). Para além disso, não foram encontrados processos anormais a serem executados pela máquina. De forma geral, não há indícios de que exista um mal direcionado à faculdade, pelo menos no contexto do enunciado. Contudo, apesar de não existir uma evidência direta de intenção maliciosa perante a faculdade, é de realçar que foram encontrados vários ficheiros escondidos que, de certa forma, podem, ou não, sugerir alguma atividade suspeita. A nossa visão para este caso é que, a indivídua Eva Rocha, dona da *pen-drive*, está, de forma autónoma, a investigar a construção de um túnel na obra corrente do Jardim do Arco do Cego, mais precisamente, de um túnel que, a partir deste, dá acesso direto ao piso -1 da adjacente a esta área, Casa da Moeda. O fundamento desta teoria advém da ligação dos artefactos encontrados, mais concretamente, o facto de existir uma imagem da planta referente ao piso -1 da Casa da Moeda, do túnel que dá acesso à Casa da Moeda (a partir da obra corrente a ser realizada no Jardim do Arco do Cego), um vídeo explicativo e, possivelmente, confirmativo do que a obra se trata de facto. E, por fim, na nossa opinião, o ponto crucial, uma imagem referente a uma folha de papel escrita a relatar toda a operação realizada contendo a confirmação de que existe uma suspeita sobre os motivos que levaram à construção deste túnel. Em suma, devido à individua Eva Rocha manifestar desconfiança perante o real motivo do ruído existente à noite naquela zona, que, supostamente, seria apenas o ruído normal da circulação do metro, esta decide iniciar uma investigação em torno de tal. Até este momento, a individua percebe que, de facto, o ruído advém da construção do túnel mencionado acima e que, na sua visão, existe um motivo lucrativo por de trás desta construção, dizendo até que tal, a ser anunciado, seria o maior escândalo de Portugal. Uma vez que se trata de uma investigação de cunho pessoal, perante um tema sensível, assumimos que a dona da *pen-drive*, decidiu por bem, ocultar os ficheiros referentes à investigação e, também, documentos privados.

4 Given your discoveries, what would be your recommendations for the subsequent course of action? Advise Mr. Golias Matos on how best to proceed with this investigation.

Dado o descoberto, a nossa recomendação será, como dito no enunciado, devolver a *pen-drive* ao proprietário, uma vez que não foi encontrado qualquer indício de *malware* ou *security threat*. No entanto, aconselharíamos a que o Sr. Golias falasse com o proprietário da *pen-drive*, de forma a perpetuar a ideia de que realizar uma investigação deste nível a cunho pessoal, não deva ser uma ação a realizar por um individuo comum não qualificado para tal. Relativamente às provas, o Sr. Golias teria de perceber com o proprietário a admissibilidade destas, de modo a perceber se as mesmas poderão ser entregues às autoridades devidas, para futura análise.