

Dosya Araştırma Kütüphanesi

(File Research Library)

1) Dosya araştırma kütüphanesi nedir?

Python 3.8.5 versiyonu ile geliştirilmiş hedef platform olarak windows işletim sisteminde çalışan script uygulamalardır. Herhangi geliştirici IDE'sine (pycharm vs.) **ihtiyaç duymadan**, tıkla çalıştır mantığıyla çalışan, ms-dos uygulamaları benzeri programlardır. Bu uygulamalar, python'a ait kendi iç kütüphanesi dışında herhangi bir bağımlılığı ya da ek kütüphane ihtiyacı yoktur. Sisteminizde python 3.8.5 ya da üst sürümü yüklü olması yeterlidir. (Alt sürümüyle de çalışabilir. Test edilmemiştir.)

2) Dosya araştırma kütüphanesi kurulum notları

Python 3.8.5 ya da üst sürümü tavsiye edilmektedir.

Python kurulumda:

- Python programı kurulurken “.py” uzantılı dosyaları, python programı tarafından çalıştırarak şekilde ayarlamaktadır. Bu hususa dikkat edilmelidir. Aksi takdirde teknik bilgisi olmayan kullanıcılar programı çalıştırmada sorun yaşayabilirler.
- Python programı kurulurken adres yollarını (dosya\klasör yolları [Örneğin c:\windows vs.]) 255 karakter ile sınırlar. Kurulum sırasında bu sınırı kaldırmak mümkündür. Tavsiye edilen kurulum sırasında gelen menüden ayarlanmasıdır. Bu ayarlanırsa kullanıcı açısından iyi olur. Ancak programın çalışmasına **engel değildir**.
- Scriptler herhangi bir dış kütüphaneye **bağımlı değildir**. Python'a ait temel kütüphaneyi kullanır. Bu sebeple herhangi bir kurulum **gerekmez**.
- Klasörü indirip ya da kopyalayıp direk kullanılabilir. İlgili scriptler **lib** klasörü içindeki tarafımdan yazılmış olan kütüphaneleri kullanılır. Bu sebeple **lib** klasörü içinde herhangi bir değişiklik yapmayınız ya da silmeyiniz.
- Doğrudan klasörlerin içindeki scriptleri (*ADSEditor.py*, *FileSignature.py*, *HashIdentifier.py*, *HexEditor.py*, *JPGInspector.py*, *MicrosoftOfficeInspector.py*, *PdfInspector.py*) kullanabileceğiniz gibi kendi scriptleriniz de **lib** içindeki fonksiyonlarını kullanarak yazabilirsiniz. Ticari olmayan ve kişisel kullanımlarda kaynak belirtilmek şartıyla geçerlidir. Diğer kullanım türleri için izin verilmemiştir. İzin alınması gerekir.

3) Dosya araştırma kütüphanesi kullanım ve lisans şartları

Mühendislik , siber güvenlik gibi çeşitli alanlarda yazılım geliştiriciliği yaptım. Bir kısmı ticariydi. Büyük çoğunluğu ise ticari olmayan yazılımlardı. Basit gözüken bir yazılımı bir yerden araklamıyorsanız (çalmıyorsanız) ciddi bir arka planı oluyor. (Araştırma safhası, algoritma kurma safhası, kodlama safhası, test safhası vb.) Ciddi bir emek ve zaman harcamış oluyorsunuz.

Dosya araştırma kütüphanesi ise çok fazla zaman ve emek vermediğim ancak kütüphane kullanmadan exif, icc_profile, mpf vs. yapıları incelememe izin verecek ve dosya standartlarını ne kadar anladığımı ölçmek istediğim bir projeydi. Sıfırdan hiçbir kütüphane kullanmadan (python temel kütüphanesi hariç) bir proje yazmak istedim. Bu şekilde ortaya çıktı. (Öğrendiğiniz bir konu hakkında algoritma yazıp kodlayamıyorsanız, o konuyu öğrenememişsiniz demektir.)

Ticari olmayan ve kişisel kullanım amaçlı yazılmıştır. Scriptler çalıştırılarak kullanılabilmesi gibi kütüphane içinde bulunan **lib** klasöründeki fonksiyonlar projelere eklenerek başka projelerde kaynak belirtmek şartıyla kullanılabilir. Ticari ve/veya eğitim amacı ile kullanılmadan önce izin alınması gerekir. (Üniversite ön lisans, lisans, yüksek lisans, doktora eğitimlerinde, çalışmalarında hoca ve öğrenciler izin almadan, kaynak belirtmek şartıyla kullanılabilir.) Konuyla ilgili arkadaşlar kodları inceleyebilir (Merak her zaman kendini geliştirmenin en iyi yöntemidir. Bende başkalarının yazdığı kodları incelerim. Çözüm yöntemlerine ve kodlama biçimlerine bakarım. Bu sayede metamorfik programlama gibi yöntemler öğrendim.)

İlgili klasörde bulunan “LICENSE.txt” dosyası, programın ve kaynak kodun lisansı olup; programı kullanırken ve dağıtırken kesinlikle klasör içinde bulunmalıdır. Bu dokümanla çelişen kısımlarda esas olarak “LICENSE.txt” dosyası baz alınır. Bu doküman kullanıcılara bilgi vermek amacı taşımaktadır.

4) Dosya araştırma kütüphanesi scriptleri

Dosya araştırma kütüphanesi scriptleri 7 tanedir. Her script ayrıntılı ve ayrı başlıklarda anlatılacaktır. Kısaca temel kullanım amaçları şunlardır:

- **ADSEditor.py** : NTFS dosya sistemine ait “alternative data stream” dosyalarını yönetmeyi sağlayan script dosyadır. Ayrıca Windows işletim sistemleri internetten indirilen dosyalara, indirilen adres bilgilerini kaydetmektedir. Bu bilgileri inceleyip, silmenizi sağlar.
- **FileSignature.py** : Dosyaların hangi tür dosya olduğunu belirlemek için dosyanın ilk baytlarını okuyarak, karşılaştırır. Bu sayede dosya türünü bulur.
- **HashIdentifier.py** : Herhangi internet sitesi, dosya vb. bir yerde karşılaştırdığınız hash’ın hangi şifreleme yöntemi olduğunu belli özelliklere bakarak listeler.
- **HexEditor.py** : Her tür dosyayı okumayı sağlayan hex editördür. Buffer özelliği sayesinde boyutu ne olursa olsun her tür dosyayı rahatlıkla okur.
- **JPGInspector.py** : Jpg ve jpeg uzantılı resim dosyalarını incelemeyi sağlayan scripttir. Exif, JFIF, Icc_Profile, MPF vb diğer tagları okumayı ve incelemeyi sağlar.
- **MicrosoftOfficeInspector.py** : Yeni Office dosyalarını (docx, docm, xlsx, xslm, pptx, pptm vs.) incelemeyi sağlayan scripttir. (Eski Office dosyalarını desteklemez. doc, xls, ppt vs.)
- **PdfInspector.py** : Pdf dosyalarını incelemeyi sağlayan scripttir. Eski ve yeni pdf versiyonlarını destekler.

Dosya araştırma kütüphanesi (file research library) **1 klasör (lib) ve 7 scriptten** oluşur. Scriptlerin kullandığı fonksiyon ve classlar **lib** klasöründe bulunmaktadır.

lib	7.09.2020 19:13	Dosya klasörü	
ADSEditor.py	3.09.2020 13:08	Python File	13 KB
FileSignature.py	7.09.2020 19:54	Python File	11 KB
HashIdentifier.py	7.09.2020 14:02	Python File	8 KB
HexEditor.py	10.08.2020 12:55	Python File	6 KB
JPGInspector.py	7.09.2020 14:06	Python File	35 KB
MicrosoftOfficeInspector.py	7.09.2020 19:16	Python File	18 KB
PdfInspector.py	7.09.2020 14:22	Python File	24 KB

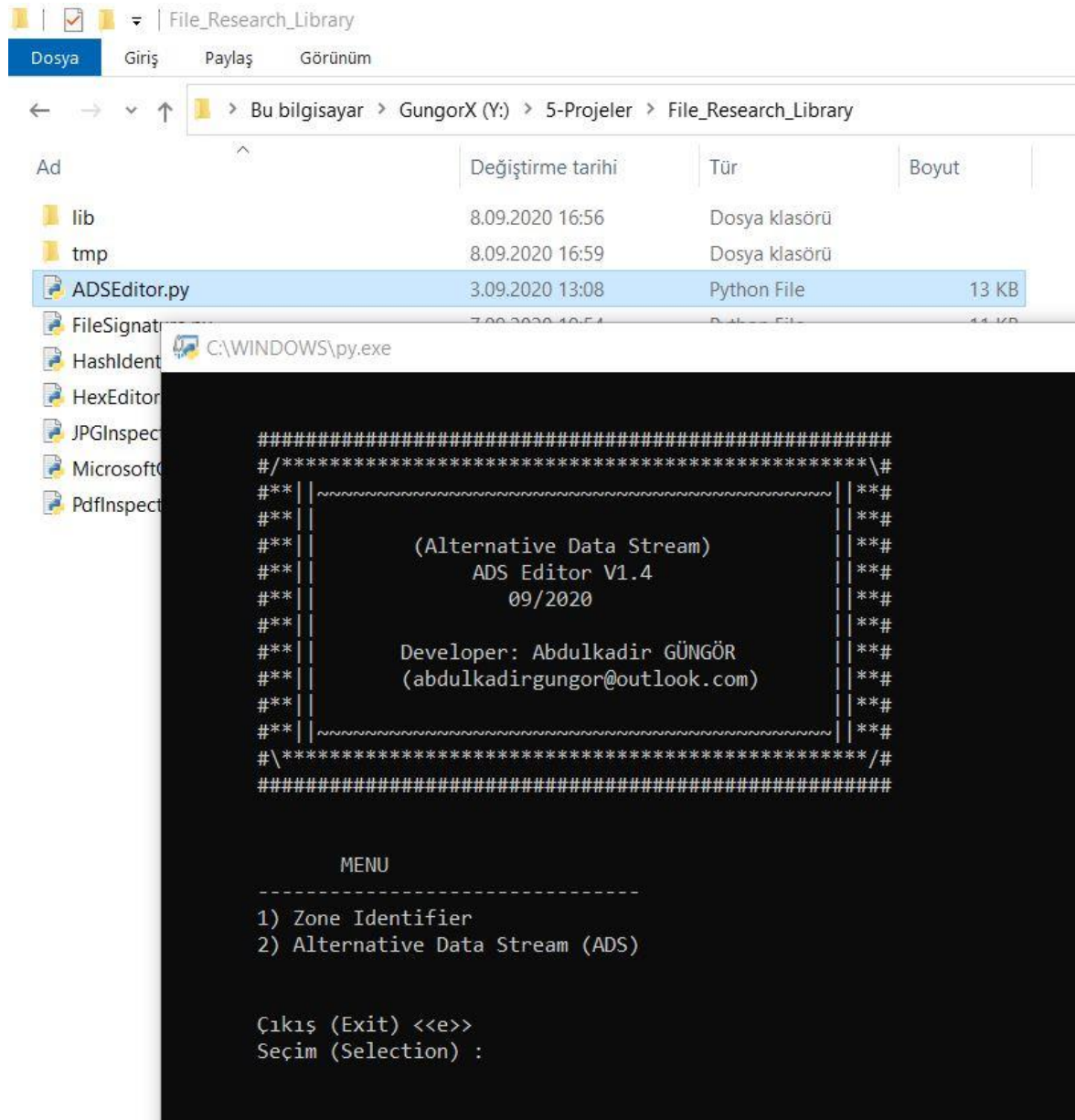
Kolaylık olması açısından incelemek istediğim klasörler ve dosyaları içine koymak için bir **tmp** klasörü oluşturuyorum. Gerekli olmamak birlikte kolaylık sağlamaktadır. Scriptlere herhangi bir dosya yolu ya da klasör yolunu girerekten işlem yapılabilir. Ancak scriptler her başlatıldığında bulunduğu klasör yolunu temel alır.

Scriptlere, dosya ve klasör yolu ya da isimlerini girerken içinde boşluk olmamasına dikkat ediniz. Aksi takdirde hatalar ile karşılaşabilirsiniz.

5) ADSEditor.py

İki ana menüden meydana gelmektedir.

- Zone Identifier
- Alternative Data Stream



1) Zone Identifier : Dosyaların internetten indirilme adresleri, windows işletim sistemi tarafından İndirilen dosyanın “alternative data stream” kısmındaki Zone.Identifier kısmına yazılmaktadır. Dosyaların “zone identifier” kısmını okuyabilir, silebilir ya da dosya olarak kaydedebilirsiniz.

İçine girildiğinde 4 menü daha açılır. Bunlar:

- Klasör yolunu değiştir. [Set Path]
- Bilgileri göster [Show info(s)]
- Bilgileri dosyaya kaydet. [Save info(s)]
- Bilgileri sil. [Delete info(s)]

C:\WINDOWS\py.exe

```
Current Directory: "Y:\5-Projeler\File_Research_Library"

-----
MENU
-----
1) Klasör yolunu değiştir. [Set path]
2) Bilgileri göster. [Show info(s)]
3) Bilgileri dosyaya kaydet. [Save info(s)]
4) Bilgileri sil. [Delete info(s)]

Çıkış (Exit) <<e>> , Geri (Back) <<b>>
Seçim (Selection) :
```

Klasör yolunu değiştir. [Set Path] : Çalışmak istenilen klasör yoludur. Default olarak scriptlerin çalıştığı klasör yolu ayarlıdır. “c:\windows” gibi adres girilebileceği gibi relative adres yolu (script içindeki “tmp” klasörü ve içindeki “pdfs” klasör yolu) “tmp\pdfs” şeklinde girilebilir.

C:\WINDOWS\py.exe

```
Klasör Yolu (Directory Path) : (Available) Y:\5-Projeler\File_Research_Library
Klasör Yolu (Directory Path) : tmp\pdfs
```

Bilgileri göster [Show info(s)] : Klasör içindeki dosyalardan okuduğu bilgileri gösterir.

C:\WINDOWS\py.exe

```
"Y:\5-Projeler\File_Research_Library\tmp\pdfs"
#####

1) "a.pdf" [Zone.Identifier]
-----
[ZoneTransfer]
ZoneId=3
ReferrerUrl=https://www.mevzuat.gov.tr/MevzuatMetin/1.5.5393.pdf
HostUrl=https://www.mevzuat.gov.tr/MevzuatMetin/1.5.5393.pdf

2) "l.pdf" [Zone.Identifier]
-----
[ZoneTransfer]
ZoneId=3
ReferrerUrl=about:client
HostUrl=about:internet

3) "w.pdf" [Zone.Identifier]
-----
[ZoneTransfer]
ZoneId=3
HostUrl=http://www.samiacar.net/about/cv.php

Devam (Continue) <<enter>>
```

Bilgileri dosyaya kaydet. [Save info(s)] : “Bilgileri göster [Show info(s)]” menüsünde gösterdiği bilgileri kaydeder. Windows’ta korumalı klasöre ya da mevcut olan dosya üzerine yazarken hatalar alabilirsiniz. Dosyayı çalışılan klasör içinde oluşturur.

d.pdf
e.pdf
f.pdf
g.pdf
kayıt.txt
l.pdf
w.pdf
x.pdf
z.pdf

C:\WINDOWS\py.exe

```
Dosya ismi (File name) : kayıt.txt

Dosyaya kaydedildi.(Save file)
Devam (Continue) <<enter>>
```



```
"Y:\5-Projeler\File_Research_Library\tmp\pdfs_yedek"  
#####
```

```
1) "a.pdf" [Zone.Identifier]  
-----
```

```
[ZoneTransfer]
```

```
ZoneId=3
```

```
ReferrerUrl=https://www.mevzuat.gov.tr/MevzuatMetin/1.5.5393.pdf
```

```
HostUrl=https://www.mevzuat.gov.tr/MevzuatMetin/1.5.5393.pdf
```

```
2) "l.pdf" [Zone.Identifier]  
-----
```

```
[ZoneTransfer]
```

```
ZoneId=3
```

```
ReferrerUrl=about:client
```

```
HostUrl=about:internet
```

```
3) "w.pdf" [Zone.Identifier]  
-----
```

```
[ZoneTransfer]
```

```
ZoneId=3
```

```
HostUrl=http://www.samiacar.net/about/cv.php
```

Bilgileri sil. [Delete info(s)] : Seçilen klasör içindeki dosyalarda bulunan zone.identifier data kısımlarını siler. Silme işlemi dosyaya ya da içindeki bilgilere zarar vermez. (Ancak hiçbir yazılımcı tarafından garanti verilmez. İşlem sorumluluğu her zaman kullanıcıya aittir.)

```
"Y:\5-Projeler\File_Research_Library\tmp\pdfs"  
#####
```

```
1) "a.pdf:Zone.Identifier" (+) [Content was deleted]
```

```
1) "a.pdf:Zone.Identifier" (+) [File was deleted]
```

```
2) "l.pdf:Zone.Identifier" (+) [Content was deleted]
```

```
2) "l.pdf:Zone.Identifier" (+) [File was deleted]
```

```
3) "w.pdf:Zone.Identifier" (+) [Content was deleted]
```

```
3) "w.pdf:Zone.Identifier" (+) [File was deleted]
```

```
Devam (Continue) <<enter>>
```

2) Alternative Data Stream (ADS) : Dosya ya da klasörlerin “alternative data stream”inde bulunan dosyaları gösterir. Bu dosyaları silebilir, dışarıya kopyalayabilir ya da dışarıdan “alternative data stream” kısmına dosya eklenebilir.

 C:\WINDOWS\py.exe

```
Current Directory: "Y:\5-Projeler\File_Research_Library"

MENU
-----
1) Klasor yolunu değiştir. [Set path]
2) ADS İşlemleri. [ADS Operations]
3) Dosyayı ADS'e kopyala. [Copy file to ADS]
4) Güncelle! [Update]

Çıkış (Exit) <<e>> , Geri (Back) <<b>>
Seçim (Selection) :
```

4 menüsü vardır.

- Klasor yolunu değiştir. [Set path]
- ADS İşlemleri. [ADS Operations]
- Dosyayı ADS'e kopyala. [Copy file to ADS]
- Güncelle! [Update]

Klasor yolunu değiştir. [Set path] : “Alternative Data Stream“ (görmek, silmek, kopyalamak amacıyla) çalışılacak klasör yolu girilir.

ADS İşlemleri. [ADS Operations] : Çalışılan klasör yolundaki içinde “alternative data stream” bulunan dosyaları listeler. Bu menüden ilgili datayı seçerek; silebilir ya da datayı dosyaya kopyalayabilirsiniz.

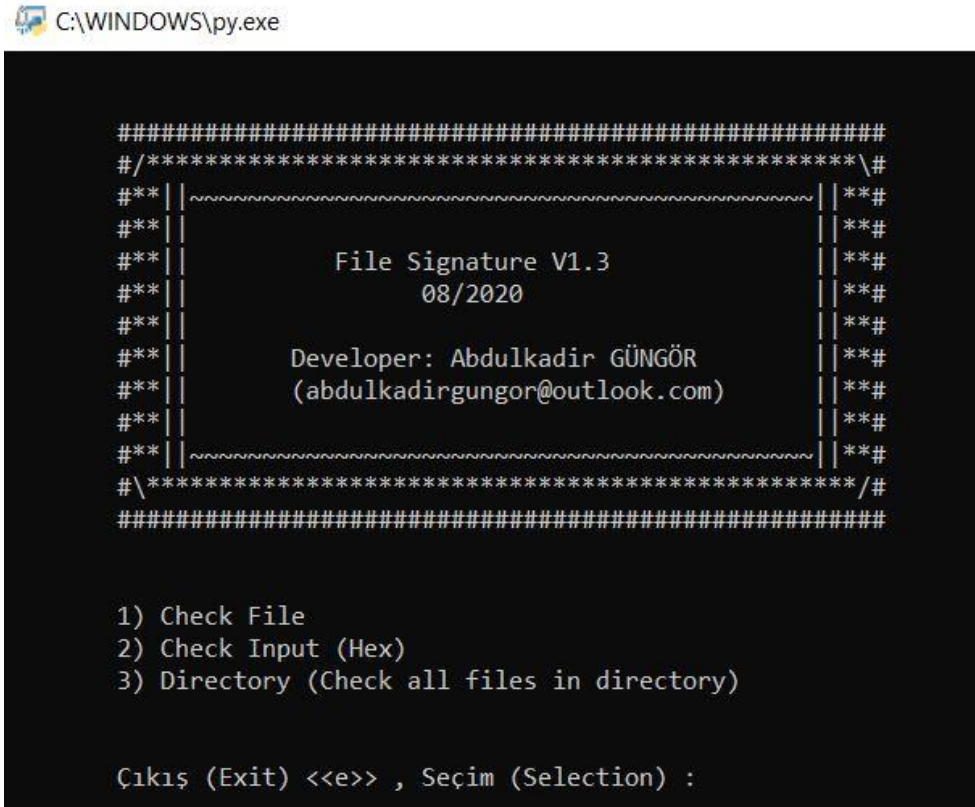
Dosyayı ADS'e kopyala. [Copy file to ADS] : Bir resim, program, text dosyasını herhangi bir dosya ya da klasörün “alternative data stream“ kısmına kopyalayabilirsiniz. Kopyalama buffer özelliği sayesinde data boyutundan bağımsız kopyalayabilirsiniz. Dosya boyutu büyüdükçe işlem süresi artmaktadır. Özellikle program kopyalarken exe, dll uzantılı dosyalar windows defender tarafından engellenebilmektedir.

Güncelle! [Update] : Klasör yolundaki dosyaları hafızaya alır. İşlem yaptıkça script kendini günceller. Ancak elle ilgili dizinde değişiklik yaptığınız zaman (dosya kopyalama ya da silme vs.) durumunda hafızayı güncellemek için kullanılır.

6) FileSignature.py

Dosyanın ilk baytlarına bakarak dosya türünü bulmaya çalışır. Bütün olasılıkları en iyi ihtimalden en düşük ihtimale doğru sıralar. Üç ana menüsü vardır.

- 1) Check file
- 2) Check input (hex)
- 3) Directory (Check all files in directory)



```
#####
#/*****\#
#**|*****|**#
#**|*****|**#
#**|          File Signature V1.3          |**#
#**|          08/2020                      |**#
#**|          Developer: Abdulkadir GÜNGÖR  |**#
#**|          (abdulkadiringor@outlook.com) |**#
#**|*****|**#
#/\*****/#
#####

1) Check File
2) Check Input (Hex)
3) Directory (Check all files in directory)

Çıkış (Exit) <<e>> , Seçim (Selection) :
```

Şekil 1) Açılış ekranı

- 1) **Check file:** Dosyadan okuma yaparak sonuçları getirir.
- 2) **Check input (hex) :** Hex editör gibi uygulama kullanıcıları (araştırmacılar) karşılaştıkları verileri elle hex türü şeklinde girerek, hangi dosya türü olduğunu kontrol edebilir.
- 3) **Directory (Check all files in directory) :** Klasör içindeki tüm dosyalar için sonuçları sırayla gösterir.

Seç C:\WINDOWS\py.exe

Kontrol edilecek dosya adını giriniz!
(Enter the file name to be checked!)

File : Y:\5-Projeler\File_Research_Library\tmp\jpgs\resimy.jpeg

The File Signatures (3)

1	Hex: 'ff d8 ff e0'	Extension: 'JFIF/JPG/JPEG'	Description: 'JPG/JPEG Image'
2	Hex: 'ff d8'	Extension: 'JPG'	Description: 'JPG Image'
3	Hex: 'ff'	Extension: 'SYS'	Description: 'Windows Executable'

Exit <<e>> Start <<enter>>

Şekil 2) "Check file" menüsü sonuçları

Abdulkadir Güngör tarafından hazırlanmıştır

7) HashIdentifier.py

Şifreleme hash'leri inceleyerek hangi şifreleme türü olduğuna dair sonuçları gösterir.

```
C:\WINDOWS\py.exe

#####
# /*****#
#**|| ****#
#**|| ****#
#**|| Hash Identifier V1.1 ****#
#**|| 09/2020 ****#
#**|| ****#
#**|| Developer: Abdulkadir GÜNGÖR ****#
#**|| (abdulkadirgungor@outlook.com) ****#
#**|| ****#
#**|| ****#
# \****#
#####

MENU
-----
1) Copy the hash from file
2) Input the hash (String)
3) Input the hash (Hex)

Çıkış (Exit) <<e>>
Seçim (Selection) :
```

Şekil 3) Açılış Ekranı

Üç ana menüden meydana gelmektedir

1) Copy the hash from file : Hash değerini dosyadan okur. Dosyayı utf-8 göre okur. (Basit “txt” dosyaları için “notepad” karakter arayı utf-8 dir.)

2) Input the hash (String) : Hash değerini doğrudan kopyalanabilir.

3) Input the hash (Hex) : Hash değeri elle hex (heximal) olarak girilebilir.

```
C:\WINDOWS\py.exe

Hash (Str) : "3F8A9C0B54671B2AA67F64570E64788BD03BE55EB26E2B2FD1023DB8393C3704"
Hash Length : 64
IsDigit : False      IsAlpha : False      IsAlphaNumeric : True
IsLower : False      IsUpper : True       IsPrintable : True

Muhtemel Sonuçlar (Possible Hashes)
[+] SHA-256
[+] Haval-256

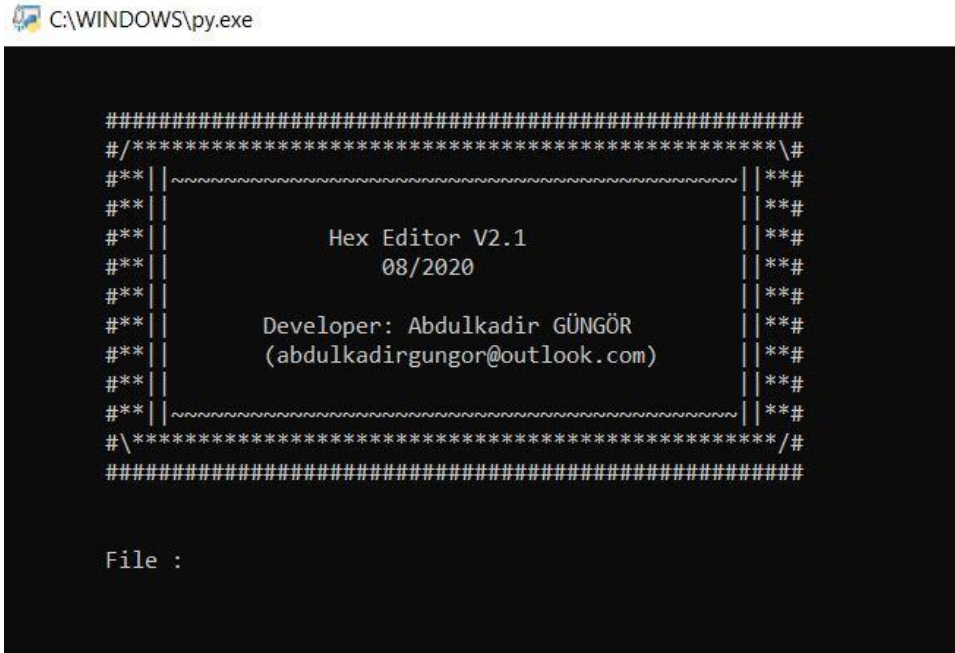
Diğer Muhtemel Sonuçlar (Other Possible Hashes)
[+] GOST R 34.11-94
[+] RipeMD-256
[+] SNEFRU-256
[+] SHA-256(HMAC)
[+] Haval-256(HMAC)
[+] RipeMD-256(HMAC)
[+] SNEFRU-256(HMAC)
[+] SHA-256(md5($pass))
[+] SHA-256(sha1($pass))

Çıkış (Exit) <<e>>, Geri (Back) <<b>>
Devam (Enter) :
```

Şekil 4) Hash Identifier Sonucu

8) HexEditor.py

Her tür dosya türünü okumak için geliştirildi. Dosya boyutu yükseldikçe çoğu dosya uygulamaları tarafında okunamayabiliyor. (Tampon) Buffer özelliği sayesinde dosya boyutundan bağımsız dosyaları rahatlıkla okuyabilir.



Dosya adı (dosya yoluyla birlikte) girildikten sonra ayarlar menüsü açılır. 5 ayar mevcuttur. Bunlar seçilerek değiştirilebilir.

- 1) **File :** (Dosya adı yoluyla beraber)
- 2) **Hex Editor Start Index:** (Standart baştan 1'den başlar. İstenirse istenilen indexten okumaya başlatılabilir.)
- 3) **Hex Editor Column Size:** (Dosya gösterimi sırasında her satırda kaç karakter olması gerektiği ayarlanabilir. Default olarak 10 karakter gösterilir.)
- 4) **Hex Editor Line Halt :** (Kaç satır okuyup ekranda gösterileceğini belirler. Girilen değer pozitif ve sıfırdan büyük olmalıdır. -1 ise dosyayı sonuna kadar okur ve gösterir. Dosya boyutu yüksek olunca ekrandaki değerlerin okunması problem olabilir. Ideal olanı bir 100 satır gibidir. Örneğin 100 değeri ayarlanırsa , 100 satıra denk gelen karakterleri okur gösterir. Sonrasında 100 'er satır olarak ilerlemeye devam eder.)
- 5) **Hex Editor Buuffer Size :** (Hex editörün tampon değeri yani buffer değeri ayarlanır. Okunacak karakter sayısına göre artırılabilir. Ayarlara göre düşük kalırsa uyarı gelir. Default olarak 1024 bayt ayarlıdır.)

C:\WINDOWS\py.exe

```
Ayarlar (Settings)
-----
1) File : 'tmp\jpgs\resim.jpg'
2) Hex Editor Start Index : 1
3) Hex Editor Column Size : 20
4) Hex Editor Line Halt : 10
5) Hex Editor Buffer Size : 1024

Selection (Continue <<Enter>> , Exit <<e>>) :
```

Şekil 5) Ayarlar Menüsü

zırlanmıştır.

C:\WINDOWS\py.exe

```
-----
|| Filename: 'tmp\jpgs\resim.jpg' - Size: 179149 (bayt) ||
-----

No:000001-000020    <Ascii>||.....ICC_PROFILE...||    <Hex>||ff d8 ff e2 0b f8 49 43 43 5f 50 52 4f 46 49 4c 45 00 01 01||
No:000021-000040    <Ascii>||.....mnrRGB ||    <Hex>||00 00 0b e8 00 00 00 00 02 00 00 00 6d 6e 74 72 52 47 42 20||
No:000041-000060    <Ascii>||XYZ .....$.acsp||    <Hex>||58 59 5a 20 07 d9 00 03 00 1b 00 15 00 24 00 1f 61 63 73 70||
No:000061-000080    <Ascii>||.....||    <Hex>||00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 01||
No:000081-000100    <Ascii>||.....-||    <Hex>||00 00 00 00 00 00 00 00 00 00 f6 d6 00 01 00 00 00 00 d3 2d||
No:000101-000120    <Ascii>||....).=.U.xB....9.||    <Hex>||00 00 00 00 29 f8 3d de af f2 55 ae 78 42 fa e4 ca 83 39 0d||
No:000121-000140    <Ascii>||.....||    <Hex>||00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00||
No:000141-000160    <Ascii>||.....desc...D||    <Hex>||00 00 00 00 00 00 00 00 00 00 00 10 64 65 73 63 00 00 01 44||
No:000161-000180    <Ascii>||...ybXYZ.....bTRC||    <Hex>||00 00 00 79 62 58 59 5a 00 00 01 c0 00 00 00 14 62 54 52 43||
No:000181-000200    <Ascii>||.....dmdd.....||    <Hex>||00 00 01 d4 00 00 08 0c 64 6d 64 64 00 00 09 e0 00 00 00 88||

%00.11 has completed
[Hex-Editor] Continue <<enter>> or quit <<q>> :
```

Şekil 6) Hex Editor

Abdulkadir Gü

Basit ama bir o kadar resim incelemesi için gerekli birçok menüye ve fonksiyona sahiptir. Başlangıçta kullanımı ve menüsü karmaşık gelebilir. Tüm menüleri anlatmak çok uzun süreceği için temel ana menüyü ve biraz tekniğine değinmek istiyorum. JPG/JPEG resim formatlarını (jpeg2000 formatı hariç) destekler. Bu resimler belli segmentlerden meydana gelir. Bazı resimler standartlara uyarken bazı resimler standartlara uymayabilir. İçinden meta verilerini silinmesi, ya da bazı üreticilerin standartlara uymaması segment bozukluklarına yol açabiliyor. Hazır uygulamalar bu verilere ulaşamayabiliyor. Hem meta verileri hem de yapısal verileri incelemek mümkündür. (Steganografi) Resim içine veri gizlemek mümkün. Ancak hem akademisyen hem de kurumsal firmalar resim içi pixellerdeki önemsiz bitlere bakarken, birçok profesyonel hacker grupları resim içlerine değil, segment yapısının içindeki padding alanı gibi (exif tag padding, icc_profile tag padding, jfif, adobe tag padding, photoshop tag padding vs.) alanlara veri gizlemesi yapıyor. (Resim içine verileri koymak için kullanılan tag ve değerlerinden değil, bu verilerin formatlanması için kullanılan padding alanlarından söz ediliyor.) Hem resimde bozulma olmuyor hem resim boyutunda bir değişiklik olmuyor, hem de steganografi gibi resim datalarının içinde yer almıyor. Bu sebepten birçok uygulama ya da incelemede farkedilmiyor ya da üzerinde durulmuyor. Steganografi tespiti için birçok yapay zeka uygulamaları dahil pek çok program vardır. Hiçbiri padding içindeki verilere tespit edemiyor çünkü programlar resim pixellerindeki datalar ile ilgilenmektedir. Bu uygulama ile çeşitli incelemeler yapılabilir. Kısaca araştırmacılar için birçok özellik sunar.

 C:\WINDOWS\py.exe

```
#####  
#/\#####  
##|~~~~~|**#  
##|      |**#  
##|    JPG/JPEG Inspector V1.1    |**#  
##|    08/2020                     |**#  
##|      |**#  
##|Developer: Abdulkadir GÜNGÖR   |**#  
##|(abdulkadiringor@outlook.com)  |**#  
##|      |**#  
##|~~~~~|**#  
#\#####/  
#####
```

File : tmp\jpgs\resimx.JPG

Girilen resmi script bulup okuduğu zaman, 12 menü gelir.

C:\WINDOWS\py.exe

```
-----
*****  MENU  *****
-----

1) Detail 'APPn and COM' ----.
   |
2) Normal 'APPn and COM' ----+----- ***{ Application and Comment Segments }***
   |
3) Summary 'APPn and COM' ----.

4) Detail Segment ----.
   |
5) Normal Segment ----+----- ***{ All Segments (Except RST) }***
   |
6) Summary Segment ---.

7) Detail All Segment ----.
   |
8) Normal All Segment ----+----- ***{ All Segments }***
   |
9) Summary All Segment ---.

10) Application and Comment Segments --.
    |
11) All Segments (Except RST) -----+----- ***{ Use Element Index }***
    |
12) All Segment -----.

Çıkış (Exit) <<e>> , Geri (Back) <<b>>
Seçim (Selection) :
```

12 menü 4 temel gruba ayrılmıştır.

İlk grup “**Application and Comment Segments**” yapısal olmayan uygulamalar tarafından metadata bilgileri gibi çeşitli bilgilerin yazıldığı segmentlerdir. Sadece bu segmentleri gösterir. Resim’in diğer yapısal segmentlerini göstermez. Kendi içinde üçe ayrılır. “**Detail**” bu segmentler hakkında bilgi ve ilk satırlarını hex editörde getirir. Herhangi bir hata olup olmadığını incelemek için. “**Normal**” bu segmentler hakkında bilgi getirir. “**Summary**” bu segmentlerin adı ve boyutunu getirir (daha az bilgi getirir). Bu üç menüden sonra ilgili segmentler seçilip daha ayrıntılı bilgiler için yeni alt menüler açılır. Bütün segmentler hex editör ile baştan sona açılabilir. Ve hex editör ile segment açıldığı zaman yanındaki sayılar jpeg içindeki konumunun gerçek index değerlerini içerir. Hex editör dışında segmentler jfif, exif, icc_profile, vb gibi veri yapılarını tanır. Tanınan bu segmentler için ayrıca her veri yapısına özel menüler açılır. Birçok uygulamada göremeyeceğiniz bilgilere ulaşabilirsiniz.

İkinci grup “**All Segments (Except RST)**” RST segmenti hariç tüm segmentleri gösterir. Bu sayede resim içindeki resimler (thumbnail- ön izleme amaçlı küçük resimler - farklı amaçlar içinde kullanımı mümkündür.), segment tanımlamada oluşan hatalar, resmin bütün segment

yapısını en iyi görülebilecek guptur. Kendi içinde üçe ayrılır. İlk gruptaki ile aynıdır. Sadece daha fazla segment görülür. Bunun sebebi yapısal segmentlerinde gösterilmesidir. (RST Segment hariç.)

Üçüncü grup “**All Segments** ” Tüm segmentleri gösterir (RST segmenti dahil) . Özel amaçlarla kullanılabilir. Bazı resimlerde onlarca, yüzlerce RST segmenti bulunabiliyor. Bu sebepten özellikle RST segmenti incelenmek istenmiyorsa ya da özel bir amaç yoksa diğer segmentler gözden kaçabiliyor ya da çok fazla segment kullanıcıyı rahatsız edebiliyor. Kendi içinde üçe ayrılır. İlk iki gruptaki ile aynıdır. Sadece daha fazla segment bulunur. Sebebi de RST segmentleri de gösteriliyor olmasıdır.

Dördüncü grup “**Use Element Index**” diğer üç gruptan hem format olarak hem de işlem olarak farklılık içerir. İlk üç grup gösterim için aynı formatı kullanır. Ve ilgili segmentleri sıralayarak kullanıcıyı gösterir. Bu sıra numarasına göre segment seçilebilir. Dördüncü grupta segmentlerin gösterimi için farklı format kullanılır. Bu segmentleri seçmek için Element Index numaraları kullanılarak işlem yapılır. Bu dört grubun hepsinde segment seçildikten sonra aynı alt menüler açılır. Bu dördüncü grupta üç’e ayrılır. Bu üç alt menü, diğer üç grupta bulunan “**normal**” segment bilgilerini farklı şekilde gösterimini içerir.

Ana menüdeki “**5- Normal Segment ... All segments (Except RST)**” seçimine ait ekran görüntüsü aşağıdadır.

C:\WINDOWS\py.exe

<<File Name:"tmp\jpgs\resimx.jpg">>

JPG/JPEG ALL SEGMENTS (EXCEPT RST)

1)	SOI	<<2 byte>>	Start/End Hex Index:0 - 2	#Element Index:1 # Root Index:0#	
2)	APP0 ('JFIF 1.2')	<<18 byte>>	Start/End Hex Index:2 - 20	#Element Index:10 # Root Index:0#	
3)	APP1 ('Exif')	<<4320 byte>>	Start/End Hex Index:20 - 4340	#Element Index:13 # Root Index:0#	
4)	* SOI	<<2 byte>>	Start/End Hex Index:988 - 990	#Element Index:2 # Root Index:13#	
5)	* APP0 ('JFIF 1.2')	<<18 byte>>	Start/End Hex Index:990 - 1008	#Element Index:11 # Root Index:13#	
6)	* APPD ('Adobe CM')	<<14 byte>>	Start/End Hex Index:1008 - 1022	#Element Index:16 # Root Index:13#	
7)	* APPE ('Adobe')	<<16 byte>>	Start/End Hex Index:1022 - 1038	#Element Index:19 # Root Index:13#	
8)	* DQT	<<134 byte>>	Start/End Hex Index:1038 - 1172	#Element Index:28 # Root Index:13#	
9)	* SOF0	<<19 byte>>	Start/End Hex Index:1172 - 1191	#Element Index:22 # Root Index:13#	
10)	* DRI	<<4 byte>>	Start/End Hex Index:1191 - 1195	#Element Index:7 # Root Index:13#	
11)	* DHT	<<321 byte>>	Start/End Hex Index:1197 - 1518	#Element Index:25 # Root Index:13#	
12)	* SOS	<<14 byte>>	Start/End Hex Index:1518 - 1532	#Element Index:31 # Root Index:13#	
13)	* EOI	<<2 byte>>	Start/End Hex Index:4338 - 4340	#Element Index:4 # Root Index:13#	
14)	APPD ('Photoshop')	<<4818 byte>>	Start/End Hex Index:4340 - 9158	#Element Index:17 # Root Index:0#	
15)	* SOI	<<2 byte>>	Start/End Hex Index:5688 - 5690	#Element Index:3 # Root Index:17#	
16)	* APP0 ('JFIF 1.2')	<<18 byte>>	Start/End Hex Index:5690 - 5708	#Element Index:12 # Root Index:17#	
17)	* APPD ('Adobe CM')	<<14 byte>>	Start/End Hex Index:5708 - 5722	#Element Index:18 # Root Index:17#	
18)	* APPE ('Adobe')	<<16 byte>>	Start/End Hex Index:5722 - 5738	#Element Index:20 # Root Index:17#	
19)	* DQT	<<134 byte>>	Start/End Hex Index:5738 - 5872	#Element Index:29 # Root Index:17#	
20)	* SOF0	<<19 byte>>	Start/End Hex Index:5872 - 5891	#Element Index:23 # Root Index:17#	
21)	* DRI	<<4 byte>>	Start/End Hex Index:5891 - 5895	#Element Index:8 # Root Index:17#	
22)	* DHT	<<321 byte>>	Start/End Hex Index:5897 - 6218	#Element Index:26 # Root Index:17#	
23)	* SOS	<<14 byte>>	Start/End Hex Index:6218 - 6232	#Element Index:32 # Root Index:17#	
24)	* EOI	<<2 byte>>	Start/End Hex Index:9038 - 9040	#Element Index:5 # Root Index:17#	
25)	APP1 ('XMP')	<<4682 byte>>	Start/End Hex Index:9158 - 13840	#Element Index:14 # Root Index:0#	
26)	APP2 ('ICC_PROFILE')	<<3162 byte>>	Start/End Hex Index:13840 - 17002	#Element Index:15 # Root Index:0#	
27)	APPE ('Adobe')	<<16 byte>>	Start/End Hex Index:17002 - 17018	#Element Index:21 # Root Index:0#	
28)	DQT	<<134 byte>>	Start/End Hex Index:17018 - 17152	#Element Index:30 # Root Index:0#	
29)	SOF0	<<19 byte>>	Start/End Hex Index:17152 - 17171	#Element Index:24 # Root Index:0#	
30)	DRI	<<4 byte>>	Start/End Hex Index:17171 - 17175	#Element Index:9 # Root Index:0#	
31)	DHT	<<420 byte>>	Start/End Hex Index:17177 - 17597	#Element Index:27 # Root Index:0#	
32)	SOS	<<14 byte>>	Start/End Hex Index:17597 - 17611	#Element Index:33 # Root Index:0#	
33)	EOI	<<2 byte>>	Start/End Hex Index:1299450 - 1299452	#Element Index:6 # Root Index:0#	

Geri (Back) < , Seçim (Selection) :

Bu menüden segmentler sıra numarası ile seçilir.

Ana menüdeki “**11-All Segments (Except RST) ... Use Element Index**” seçimine ait ekran görüntüsü aşağıdadır.

<<File Name:"tmp\jpgs\resimx.JPG">>

JPG/JPEG ALL SEGMENTS (EXCEPT RST)

```
#####
---> Element index :1 | root :0 | SOI | ' * ' | Length :2 | (byte) --> Start Index :0 - End Index :2
---> Element index :10 | root :0 | APP0 | 'JFIF 1.2' | Length :18 | (byte) --> Start Index :2 - End Index :20
---> Element index :13 | root :0 | APP1 | 'Exif' | Length :4320 | (byte) --> Start Index :20 - End Index :4340
-----> Element index :2 | root :13 | SOI | ' * ' | Length :2 | (byte) --> Start Index :988 - End Index :990
-----> Element index :11 | root :13 | APP0 | 'JFIF 1.2' | Length :18 | (byte) --> Start Index :990 - End Index :1008
-----> Element index :16 | root :13 | APPD | 'Adobe CM' | Length :14 | (byte) --> Start Index :1008 - End Index :1022
-----> Element index :19 | root :13 | APPE | 'Adobe' | Length :16 | (byte) --> Start Index :1022 - End Index :1038
-----> Element index :28 | root :13 | DQT | ' * ' | Length :134 | (byte) --> Start Index :1038 - End Index :1172
-----> Element index :22 | root :13 | SOF0 | ' * ' | Length :19 | (byte) --> Start Index :1172 - End Index :1191
-----> Element index :7 | root :13 | DRI | ' * ' | Length :4 | (byte) --> Start Index :1191 - End Index :1195
-----> Element index :25 | root :13 | DHT | ' * ' | Length :321 | (byte) --> Start Index :1197 - End Index :1518
-----> Element index :31 | root :13 | SOS | ' * ' | Length :14 | (byte) --> Start Index :1518 - End Index :1532
-----> Element index :4 | root :13 | EOI | ' * ' | Length :2 | (byte) --> Start Index :4338 - End Index :4340
---> Element index :17 | root :0 | APPD | 'Photoshop' | Length :4818 | (byte) --> Start Index :4340 - End Index :9158
-----> Element index :3 | root :17 | SOI | ' * ' | Length :2 | (byte) --> Start Index :5688 - End Index :5690
-----> Element index :12 | root :17 | APP0 | 'JFIF 1.2' | Length :18 | (byte) --> Start Index :5690 - End Index :5708
-----> Element index :18 | root :17 | APPD | 'Adobe CM' | Length :14 | (byte) --> Start Index :5708 - End Index :5722
-----> Element index :20 | root :17 | APPE | 'Adobe' | Length :16 | (byte) --> Start Index :5722 - End Index :5738
-----> Element index :29 | root :17 | DQT | ' * ' | Length :134 | (byte) --> Start Index :5738 - End Index :5872
-----> Element index :23 | root :17 | SOF0 | ' * ' | Length :19 | (byte) --> Start Index :5872 - End Index :5891
-----> Element index :8 | root :17 | DRI | ' * ' | Length :4 | (byte) --> Start Index :5891 - End Index :5895
-----> Element index :26 | root :17 | DHT | ' * ' | Length :321 | (byte) --> Start Index :5897 - End Index :6218
-----> Element index :32 | root :17 | SOS | ' * ' | Length :14 | (byte) --> Start Index :6218 - End Index :6232
-----> Element index :5 | root :17 | EOI | ' * ' | Length :2 | (byte) --> Start Index :9038 - End Index :9040
---> Element index :14 | root :0 | APP1 | 'XMP' | Length :4682 | (byte) --> Start Index :9158 - End Index :13840
---> Element index :15 | root :0 | APP2 | 'ICC_PROFILE' | Length :3162 | (byte) --> Start Index :13840 - End Index :17002
---> Element index :21 | root :0 | APPE | 'Adobe' | Length :16 | (byte) --> Start Index :17002 - End Index :17018
---> Element index :30 | root :0 | DQT | ' * ' | Length :134 | (byte) --> Start Index :17018 - End Index :17152
---> Element index :24 | root :0 | SOF0 | ' * ' | Length :19 | (byte) --> Start Index :17152 - End Index :17171
---> Element index :9 | root :0 | DRI | ' * ' | Length :4 | (byte) --> Start Index :17171 - End Index :17175
---> Element index :27 | root :0 | DHT | ' * ' | Length :420 | (byte) --> Start Index :17177 - End Index :17597
---> Element index :33 | root :0 | SOS | ' * ' | Length :14 | (byte) --> Start Index :17597 - End Index :17611
---> Element index :6 | root :0 | EOI | ' * ' | Length :2 | (byte) --> Start Index :1299450 - End Index :1299452
```

Geri (Back) <> , Seçim (Selection -Element Index-) :

Bu menüdeki segmentler “element index“ ile seçilir.

Yukarıdaki her iki ana menü seçimini gösteren resimler segmentlerde herhangi bir hata olmayan bir resim incelemesine aittir. Dikkat edilirse root değeri 0 olan ana segmentler, root değeri 0 farklı olan segmentler, ana segmentlerin içinde bulunan alt segmentlerdir. Genellikle ilgili segmentte önizleme için küçük resimler (thumbnail) olabileceği gibi, veri gizleme, standartlara uymayan bozuk segment ya da yanlış yorumlanan baytlar olabilir. Bir segmentin başlangıç (start) ve bitiş (end) (hex editör için adresleri – onluk sistemde) değerlerine bakılarak ilgili segmentin hangi segment içinde olup olmadığı görülebilir. Bunu daha kolay görmek için root değerine bakılır. root değeri, ilgili segmenttin hangi ana segment’e ait olduğunu gösteren bir değerdir. Programda, ana segment ve alt segmentler ayırt edici bir gösterime sahiptir. Jpg/Jpeg resim segmentleri (SOI, APP1, ... DHT, EOI,) 3. kolonda gösterilir. Bu segmentler içinde tanımlanan verinin format biçimi tanındığı takdirde 4. kolonda veri tipi gösterilir. Bu kolonda tanınmayan veriler “Unknown” ile gösterilir. Eğer veri taşımayan yapısal segmentler “*” ile gösterilir. Veri taşıyan ve taşımayan tüm segmentler hex editör ile açılabilir. Veri taşıyan segmentler hem hex editör hem de özel alt menüler ile açılabilir. Hex editör açılan segmentler resim içindeki gerçek index değerlerine (konumuna) göre açılır. Sadece ilgili segment datası gösterilir.

İlgili “ **Element Index :13 - APP1 - ‘Exif’** ” segmenttin seçilince açılan menü aşağıdadır. “**2-Show <<Exif>> Header**” menü exif hakkında bulunan bilgi ve tag sayılarını verir.

C:\WINDOWS\py.exe

```
Segment Type Name : APP1
Segment Index      : 13
Segment Size       : 4320 bayt
App Name           : Exif

Menu
-----
1) Hex Editor
2) Show <<Exif>> Header
3) Show <<Exif>> '0IFD' Tag Names
4) Show <<Exif>> 'ExifIFD' Tag Names
5) Show <<Exif>> 'InterOperabilityIFD' Tag Names
6) Show <<Exif>> 'GpsIFD' Tag Names
7) Show <<Exif>> '1IFD' Tag Names
```

Geri (Back) <> , Seçim (Selection) :

Yukarıdaki menüden “2-Show <<Exif>> Header” seçilince gelen ekran.

C:\WINDOWS\py.exe

```
App Marker           : ff e1 'APP1'
App Length            : 10 de (4318 bayt)
Exif Marker           : 45 78 69 66 00 00 'Exif\x00\x00'
Exif Byte Order       : 4d 4d 'MM' 'Big Endian'
Exif Version Number   : 00 2a (42)
Exif IFD              : 00 00 00 08
```

```
Exif '0IFD' Tags Count : 9
Exif 'Exif IFD' Tags Count : 38
Exif 'Interoperability IFD' Tags Count : 0
Exif 'GPS IFD' Tags Count : 1
Exif '1IFD' Tags Count : 6
```

+

```
Exif Total Tags Count : 54
```

Devam (Continue) <<Enter>> :

Çok detaya girmeyeceğim ama ‘1IFD’ tagları (etiketleri) thumbnail (önizleme resmine) ait taglar olup diğerleri ana resme aittir. Bunu tüm segmentlerin görüldüğü ana yapıdan

anlayabileceğimiz gibi standartlarda IIFD taglarının her zaman thumbnail(önizleme resmine) işaret ettiği söylenir. (Veri gizleme vb. durumlar istisnadır.)

Enter tuşuna basıp bir üst menüye döndükten sonra bulunan tagları 3, 4, 5, 6 ve 7 menülerde seçebiliriz.

C:\WINDOWS\py.exe

```
      0IFD TAGS (9)
-----
1. Make
2. Model
3. Orientation
4. XResolution
5. YResolution
6. ResolutionUnit
7. Software
8. ModifyDate
9. YCbCrPositioning

Geri (Back) <<b>> , Seçim (Selection) :
```

C:\WINDOWS\py.exe

```
#####
Tag No      : 2
Tag (Hex)   : '01 10'
Tag Name    : Model

Data Type (Hex) : 00 02
Data Type (str) : ASCII
Data Length   : 10 bayt

(Data)
<Ascii>||NIKON D90.||    <Hex>||4e 49 4b 4f 4e 20 44 39 30 00||
#####

Devam (Continue) <<Enter>> :
```

Ya da bulunan segmentleri hex editör ile açabiliriz.

10) MicrosoftOfficeInspector.py

Yeni Microsoft Office dosya türlerini açar [docx, docm, xlsx, xlsxm, pptx, pptm vs.]. Ancak eski Office dosya türlerini **desteklemez** [doc, xls, ppt vs.].

```
C:\WINDOWS\py.exe

#####
#/\*****\#
#**| ~~~~~|**#
#**|      |**#
#**|      Microsoft Office File Inspector      |**#
#**|      (Docx, Xlsx, Pptx, Docm, Xlsm vs.)      |**#
#**|      09/2020      |**#
#**|      Developer: Abdulkadir GÜNGÖR      |**#
#**|      (abdulkadiringunor@outlook.com)      |**#
#**| ~~~~~|**#
#/\*****\#
#####

MENU
-----
1) Microsoft Office File Inspector
2) Microsoft Office File Extract
3) Microsoft Office File Package

Çıkış (Exit) <<e>>
Seçim (Selection) :
```

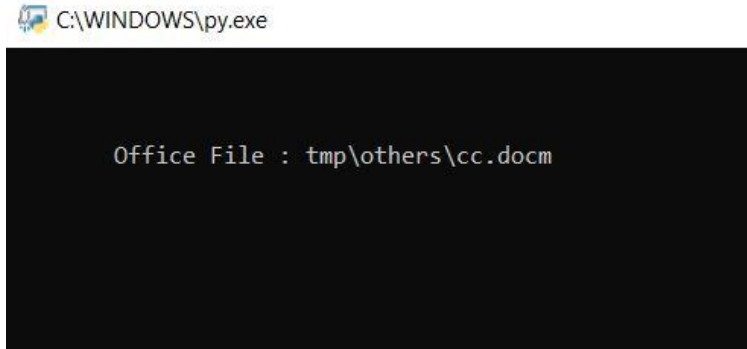
Script üç ana menüden oluşur.

1) Microsoft Office File Inspector : Microsoft Office dosyasının içinde bulunan diğer data türleri ve dosyaları incelemek, analiz etmek vb. gibi özellikler içerir. Dosyaları çıkarmadan analiz eder. Ayrıntılı olarak aşağıda tekrar değinilecektir.

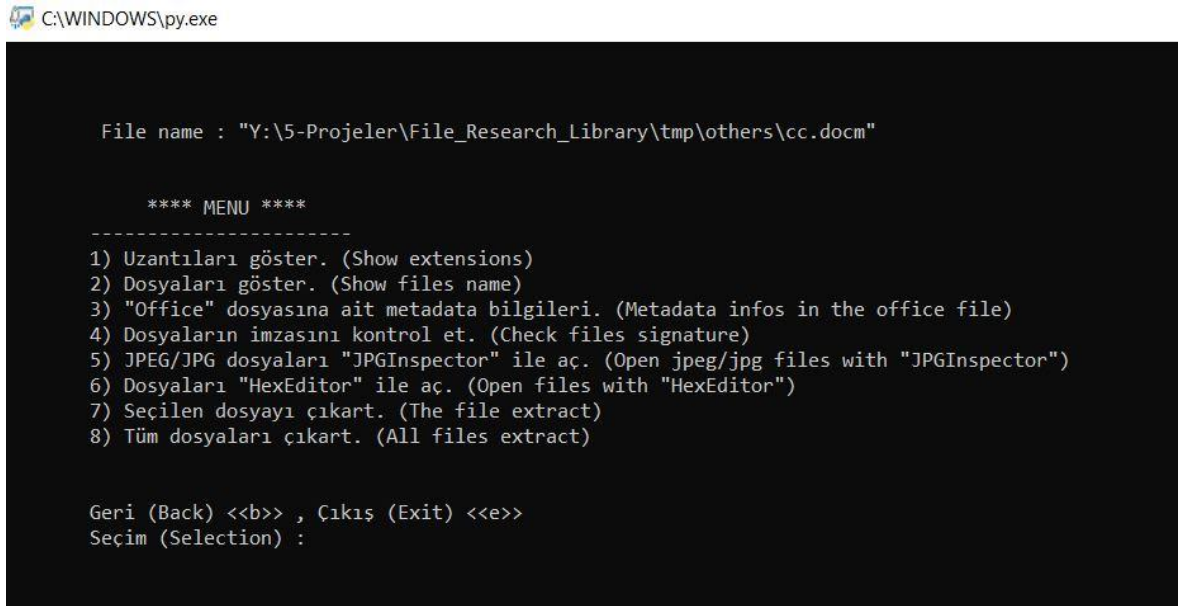
2) Microsoft Office File Extract : Microsoft Office dosyasının içinde bulunan diğer data türleri ve dosyaları dışarıya çıkarır.

3) Microsoft Office File Package : Microsoft Office dosyasının içinden çıkarılan data türlerini ve dosyaları tekrar Office dosyası haline getirmek için kullanılır.

“1) Microsoft Office File Inspector” seçildikten sonra Office dosyasının yolu ve adı girilir.



Dosya girildikten sonra 8 ana menü gösterilir. Bunlar:



1) Uzantıları göster. (Show extensions) : Dosya içerisinde bulunan dosyaların uzantılarını gösterir.

C:\WINDOWS\py.exe

```
File Name : "Y:\5-Projeler\File_Research_Library\tmp\others\cc.docm"

Bulunan Uzantılar
-----

[+] XML [10]
[+] RELS [3]
[+] BIN [1]

Devam (Continue) <<enter>> :
```

2) Dosyaları göster. (Show files name) : Dosyaları uzantılarına göre gruplandırarak tüm dosya ve dataları gösterir. Microsoft dosyasına hangi işletim sisteminde dosya ve dataların eklendiğini görebilirsiniz. Zararlı içerikler genelde Linux'ta hazırlanır (Metasploit gibi.). (Windows işletim sisteminde de hazırlanabilir ancak daha nadirdir.)

C:\WINDOWS\py.exe

```
File Name : "Y:\5-Projeler\File_Research_Library\tmp\others\cc.docm"

[-] XML [10]
|----- [Content_Types].xml           Windows      1453 Bayts
|----- docProps/app.xml              Windows      709 Bayts
|----- docProps/core.xml             Windows      739 Bayts
|----- word/document.xml             Windows      2703 Bayts
|----- word/fontTable.xml            Windows      1567 Bayts
|----- word/settings.xml             Windows      3152 Bayts
|----- word/styles.xml               Windows      29370 Bayts
|----- word/theme/theme1.xml         Windows      6799 Bayts
|----- word/vbaData.xml              Windows      2369 Bayts
|----- word/webSettings.xml          Windows      803 Bayts

[-] RELS [3]
|----- _rels/.rels                   Windows      590 Bayts
|----- word/_rels/document.xml.rels  Windows      939 Bayts
|----- word/_rels/vbaProject.bin.rels Windows      277 Bayts

[-] BIN [1]
|----- word/vbaProject.bin            Windows      7680 Bayts

Devam (Continue) <<enter>> :
```

3) "Office" dosyasına ait metadata bilgileri. (Metadata infos in the office file) : İki alt menü gelir. İlki metadata bilgilerini gösterir. Diğer metadata bilgilerini siler. Burada metadata bilgisi silinen Office dosyasıdır. İçerdiği resim vb dataların metadata bilgileri **silinmez**. Sadece ana dosyanın metadata bilgisini siler.

C:\WINDOWS\py.exe

File name : "Y:\5-Projeler\File_Research_Library\tmp\others\cc.docm"

**** MENU ****

-
- 1) Metadata bilgilerini göster. (Show metadata infos)
 - 2) Metadata bilgilerini sil. (Delete metadata infos)

Geri (Back) <> , Çıkış (Exit) <<e>>
Seçim (Selection) :

C:\WINDOWS\py.exe

File name : "Y:\5-Projeler\File_Research_Library\tmp\others\cc.docm"

Metadata Bilgileri (Metadata Infos)

title	: -
subject	: -
creator	: Gungor
keywords	: -
description	: -
lastModifiedBy	: Gungor
revision	: 2
created	: 2020-09-05T15:08:00Z
modified	: 2020-09-05T15:08:00Z

Devam (Continue) << enter >> :

4) Dosyaların imzasını kontrol et. (Check files signature) : Office dosyasının içindeki dosyaları çıkarmadan dosya türlerini kontrol eder.

C:\WINDOWS\py.exe

File name : "Y:\5-Projeler\File_Research_Library\tmp\others\cc.docm"

[-] [Content_Types].xml

----- Hex : "3c 3f 78 6d 6c 20 76 65 72 73 69 6f 6e 3d"	Extension : "MANIFEST"	Description : "Windows Visual Stylesheet"
----- Hex : "3c"	Extension : "ASX"	Description : "Advanced Stream Redirector"
----- Hex : "3c"	Extension : "XDR"	Description : "BizTalk XML-Data Reduced Schema"

[-] _rels/.rels

----- Hex : "3c 3f 78 6d 6c 20 76 65 72 73 69 6f 6e 3d"	Extension : "MANIFEST"	Description : "Windows Visual Stylesheet"
----- Hex : "3c"	Extension : "ASX"	Description : "Advanced Stream Redirector"
----- Hex : "3c"	Extension : "XDR"	Description : "BizTalk XML-Data Reduced Schema"

[-] word/document.xml

----- Hex : "3c 3f 78 6d 6c 20 76 65 72 73 69 6f 6e 3d"	Extension : "MANIFEST"	Description : "Windows Visual Stylesheet"
----- Hex : "3c"	Extension : "ASX"	Description : "Advanced Stream Redirector"
----- Hex : "3c"	Extension : "XDR"	Description : "BizTalk XML-Data Reduced Schema"

[-] word/_rels/document.xml.rels

----- Hex : "3c 3f 78 6d 6c 20 76 65 72 73 69 6f 6e 3d"	Extension : "MANIFEST"	Description : "Windows Visual Stylesheet"
----- Hex : "3c"	Extension : "ASX"	Description : "Advanced Stream Redirector"
----- Hex : "3c"	Extension : "XDR"	Description : "BizTalk XML-Data Reduced Schema"

[-] word/vbaProject.bin

----- Hex : "d0 cf 11 e0 a1 b1 1a e1"	Extension : "AC"	Description : "CaseWare Working Papers"
----- Hex : "d0 cf 11 e0 a1 b1 1a e1"	Extension : "ADP"	Description : "Access Project File"
----- Hex : "d0 cf 11 e0 a1 b1 1a e1"	Extension : "APR"	Description : "Lotus (IBM Approach 97 File)"
----- Hex : "d0 cf 11 e0 a1 b1 1a e1"	Extension : "DB"	Description : "MSWorks Database File"
----- Hex : "d0 cf 11 e0 a1 b1 1a e1"	Extension : "doc"	Description : "Microsoft Office Document"
----- Hex : "d0 cf 11 e0 a1 b1 1a e1"	Extension : "DOT"	Description : "Microsoft Office Document"
----- Hex : "d0 cf 11 e0 a1 b1 1a e1"	Extension : "MSC"	Description : "Microsoft Common Console Document"
----- Hex : "d0 cf 11 e0 a1 b1 1a e1"	Extension : "MSI"	Description : "Microsoft Installer Package"
----- Hex : "d0 cf 11 e0 a1 b1 1a e1"	Extension : "MTW"	Description : "Minitab Data File"
----- Hex : "d0 cf 11 e0 a1 b1 1a e1"	Extension : "OPT"	Description : "Developer Studio File Options File"
----- Hex : "d0 cf 11 e0 a1 b1 1a e1"	Extension : "PPS"	Description : "Microsoft Office Document"
----- Hex : "d0 cf 11 e0 a1 b1 1a e1"	Extension : "PPT"	Description : "Microsoft Office Document"

5) JPEG/JPG dosyaları "JPGInspector" ile aç. (Open jpeg/jpg files with "JPGInspector") : Office dosyasının içindeki resimleri çıkarmadan JPGInspector uygulaması ile açar. Bu uygulama "9) JPGInspector.py" başlığı altında açıklanmıştır.

C:\WINDOWS\py.exe

```
File name : "Y:\5-Projeler\File_Research_Library\tmp\others\tez.docx"

[-] JPEG
|----- 1) word/media/image10.jpeg           Windows           26592 Bayts
|----- 2) word/media/image11.jpeg           Windows           40531 Bayts
|----- 3) word/media/image12.jpeg           Windows           42251 Bayts
|----- 4) word/media/image14.jpeg           Windows           14973 Bayts
|----- 5) word/media/image15.jpeg           Windows           17516 Bayts
|----- 6) word/media/image16.jpeg           Windows           18034 Bayts
|----- 7) word/media/image17.jpeg           Windows           83395 Bayts
|----- 8) word/media/image18.jpeg           Windows           94176 Bayts
|----- 9) word/media/image19.jpeg           Windows           86647 Bayts
|----- 10) word/media/image20.jpeg          Windows           86533 Bayts
|----- 11) word/media/image4.jpeg           Windows           100861 Bayts
|----- 12) word/media/image5.jpeg           Windows           71878 Bayts
|----- 13) word/media/image8.jpeg           Windows           87571 Bayts
|----- 14) word/media/image9.jpeg           Windows           98553 Bayts

[-] JPG
|----- 15) word/media/image21.JPG           Windows           215555 Bayts

Geri (Back) <<b>> , Çıkış (Exit) <<e>>
Seçim (Selection) :
```

6) Dosyaları "HexEditor" ile aç. (Open files with "HexEditor") : Office dosyasının içindeki dosyaları çıkarmadan HexEditor uygulaması ile açar. Bu uygulama "8) HexEditor.py" başlığı altında açıklanmıştır. 4 alt menüye sahiptir.

C:\WINDOWS\py.exe

```
File name : "Y:\5-Projeler\File_Research_Library\tmp\others\tez.docx"

**** MENU ****
-----
1) Dosyanın ilk satırlarını "hex editor"de aç.(Show first line of the file with "hex editor")
2) Dosyayı "hex editor"de aç. (Open file with "hex editor")
3) Tüm dosyaların ilk satırlarını "hex editor"de aç. (Show first line of the files with "hex editor")
4) Tüm dosyaları sırayla "hex editor"de aç. (Open files in order with "hex editor" )

Geri (Back) <<b>> , Çıkış (Exit) <<e>>
Seçim (Selection) :
```


7) Seçilen dosyayı çıkart. (The file extract) : Office dosyasının içindeki istenilen dosyayı incelemek için çıkarabilirsiniz.

C:\WINDOWS\py.exe

```
File name : "Y:\5-Projeler\File_Research_Library\tmp\others\tez.docx"

[-] tez.docx [57]
|----- (1) [Content_Types].xml          XML      Windows      3321 Bayts
|----- (2) word/document.xml           XML      Windows      592693 Bayts
|----- (3) word/footnotes.xml           XML      Windows      3140 Bayts
|----- (4) word/endnotes.xml            XML      Windows      3134 Bayts
|----- (5) word/header1.xml             XML      Windows      3588 Bayts
|----- (6) word/header2.xml             XML      Windows      3749 Bayts
|----- (7) word/header3.xml             XML      Windows      3588 Bayts
|----- (8) word/header4.xml             XML      Windows      3797 Bayts
|----- (9) word/header5.xml             XML      Windows      3628 Bayts
|----- (10) word/header6.xml            XML      Windows      3588 Bayts
|----- (11) word/header7.xml            XML      Windows      3588 Bayts
|----- (12) word/header8.xml            XML      Windows      4212 Bayts
|----- (13) word/header9.xml            XML      Windows      3588 Bayts
|----- (14) word/theme/theme1.xml       XML      Windows      6797 Bayts
|----- (15) word/settings.xml           XML      Windows      35365 Bayts
|----- (16) word/webSettings.xml        XML      Windows      205234 Bayts
|----- (17) word/fontTable.xml          XML      Windows      2806 Bayts
|----- (18) customXml/item1.xml         XML      Windows      26940 Bayts
|----- (19) docProps/core.xml           XML      Windows      789 Bayts
|----- (20) customXml/itemProps1.xml    XML      Windows      341 Bayts
|----- (21) docProps/custom.xml         XML      Windows      361 Bayts
|----- (22) word/numbering.xml          XML      Windows      85683 Bayts
|----- (23) word/styles.xml             XML      Windows      42412 Bayts
|----- (24) docProps/app.xml            XML      Windows      1039 Bayts
|----- (25) _rels/.rels                 RELS     Windows      737 Bayts
|----- (26) word/_rels/document.xml.rels RELS     Windows      5203 Bayts
|----- (27) word/_rels/header6.xml.rels RELS     Windows      289 Bayts
```

8) Tüm dosyaları çıkart. (All files extract) : Office dosyasının içindeki dosya ve dataları bir klasör içinde çıkarır. Çıkardığı dizin office dosyasının bulunduğu dizindir.

C:\WINDOWS\py.exe

```
Office file name : "Y:\5-Projeler\File_Research_Library\tmp\ornek\tez.docx"
Extract Directory : (Default) "Y:\5-Projeler\File_Research_Library\tmp\ornek\[DOCX]_tez.docx"

Default <<enter>> , New Extract Directory :
İşlem başarılı. (Operation successful)


Devam (Continue) <<enter>> :
```

bu bilgisayar > GungorX (Y:) > 5-Projeler > File_Research_Library > tmp > ornek

Ad	Değiştirme tarihi	Tür	Boyut
[DOCX]_tez.docx	9.09.2020 18:14	Dosya klasörü	
tez.docx	8.09.2020 16:46	Microsoft Word Be...	1.468 KB

11) PdfInspector.py

Basit ama güçlü araçlardan bir tanesi de “PdfInspector.py” programıdır.

 C:\WINDOWS\py.exe

```
#####
#/\*****\#
#**|| ****#
#**|| ****#
#**|| Pdf Inspector V2.3 ****#
#**|| 08/2020 ****#
#**|| ****#
#**|| Developer: Abdulkadir GÜNGÖR ****#
#**|| (abdulkadiringunor@outlook.com) ****#
#**|| ****#
#**|| ****#
#/\*****/#
#####

Çıkış (Exit) <<e>>
File Name : tmp\pdfs\x.pdf
```

İlgili pdf yolunu girdikten sonra 12 ana menü açılır. Menü gelmeden önce pdf yüklenmesi ve parse edilme süreci vardır. Dosya boyutuna göre değişiklik gösterebilir.

 C:\WINDOWS\py.exe

```
-----
MENU
-----
1) Pdf Summary
2) Check Pdf Tags
3) * Show Object (Normal - use Obj Header)
4) Show Object (Stream Summary)
5) Show Object (Normal)
6) Show Object (Hex)
7) Show Stream (Normal)
8) Show Stream (Hex)
9) Write Stream
10) Show Xref (Hex)
11) Show Trailer (Hex)
12) Show StartXref (Hex)

Çıkış (Exit) <<e>> , Geri (Back) <<b>>
Seçim (Selection) :
```

1) **Pdf Summary** : Parse edilmiş pdf tag değerleri ile ilgili bilgileri gösterir.

C:\WINDOWS\py.exe

```
File Name : tmp\pdfs\c2.pdf
File Size : 1113784 bayts

Pdf Header (1. Marker)      : %
Pdf Header (Version)       : PDF-1.6
Pdf Header (1. Split) Hex  : "0d"
Pdf Header (2. Marker)     : %
Pdf Header (Signature) Hex : "e2 e3 cf d3"
Pdf Header (2. Split) Hex  : "0d"

Pdf <<obj-endobj>> count    : 160
Pdf <<stream-endstream>> count : 127

Pdf <<xref>> count          : 0
Pdf <<trailer>> count       : 0
Pdf <<startxref>> count     : 1
Pdf <<%EOF>> count          : 1

Devam (Continue) <<enter>> :
```

2) **Check Pdf Tags** : Çeşitli pdf taglarını (53 adet tag) bulur. İlgili tagın numarasını girerek tagın içinde geçtiği (obj-endobj) objectler görüntülenebilir.

C:\WINDOWS\py.exe

No	Tag	Obj Count	Obj Tag Count	Pdf Tag Count
1	"/Page"	25	25	25
2	"/Encrypt"	0	0	0
3	"/AuthEvent"	0	0	0
4	"/ObjStm"	0	0	0
5	"/JS"	1	2	2
6	"/JavaScript"	1	2	2
7	"/AA"	1	1	1
8	"/XML"	1	1	1
9	"/Action"	1	1	1
10	"/OpenAction"	0	0	0
11	"/AcroForm"	1	1	1
12	"/AcroField"	0	0	0
13	"/FDF"	0	0	0
14	"/XForm"	0	0	0
15	"/Form"	6	9	9
16	"/XFA"	0	0	0
17	"/Launch"	0	0	0
18	"/JBIG2Decode"	0	0	0
19	"/ASCIITextDecode"	0	0	0
20	"/RichMedia"	0	0	0
21	"/FileAttachment"	0	0	0
22	"/Type"	88	121	121
23	"/FileSpec"	0	0	0
24	"/EmbeddedFile"	0	0	0

Görüldüğü gibi 5 kolondan oluşur. 3. kolonda geçen “Obj Count” , ilgili tagın kaç adet (obj-endobj) object’in içinde geçtiğini gösterir. 4. kolon “Obj Tag Count” , (obj-endobj) object içindeki tagın toplamda kaç kere geçtiği gösterir. Biraz karışık geliş olabilir. Daha öz bir biçimde örneğin, ‘/Type’ tagı 88 object içinde 121 kere geçtiğini gösterir. (3. ve 4 .kolon bu işe yarar). 5. kolon ise “pdf tag count” ilgili tagın pdf içinde toplam kaç kere geçtiğini gösterir. Bu sayede ilgili (obj-endobj) object parse edilirken hata olup olmadığı ya da bazı taglar (obj-endobj) object dışında da bulunabiliyor, bu tagları görmek için. “\Encrpyt” tagı obj dışında trailer bloğunda da bulunabiliyor. Toplam 53 tag var. Bu taglar birbirleriyle kontrol edilerek kullanılır. Hem oluşan hataları görmek hem de analiz açısından önemlidir. Örneğin pdf içindeki xmp veri datası bulunmak istenirse içinde hem “\xml” hem de “\Metadata” değerlerine bakmak gerekir. Bu iki tag, xmp veri datası olduğuna işaret eder. İlgili tagların bulunduğu (obj-endobj) objectler görülmek istenirse tag numarası girilerek ilgili (obj-endobj) object’ler incelenebilir. İlgili tag içeren tüm (obj-endobj) object’ler sırayla gösterilir. İnceleme açısından kolaylık sağlar.

C:\WINDOWS\py.exe

```
<<Obj No : 0>> -      Obj header : "1 0 obj"
-----
<<
  /Outlines 1174 0 R
  /Pages 2 0 R
  /AA
    <<
      /WC
        <<
          /S
          /JavaScript
          /JS (app.alert\('Merhaba D nya'\));)
        >>
      /WP
        <<
          /S
          /JavaScript
          /JS (URI \http:/
            /www.mynet.com\))
        >>
    >>
  /Metadata 1146 0 R
  /ViewerPreferences 1147 0 R
  /AcroForm 1151 0 R
  /Type
  /Catalog
  /Lang (tr-TR)
  /MarkInfo
    <<
      /Marked true
    >>
  >>
-----
```

“/Javascript” tagı seçildikten sonra sırayla tagı içeren objectler yukardaki resimdeki gibi incelemek üzere kullanıcıya gösterilir.

3) * **Show Object (Normal - use Obj Header)** : İlgili programdaki tüm (obj-enobj) objectler ve (stream-endstream) streamler bu program tarafından numara verilir (obj no, stream no). Bu numara sıfırdan başlar. Ana menüdeki diğer seçenekler de bu numaralar kullanılır. Ancak bu seçenekte istisnai olarak “obj no” değil “obj header” kullanılır. Çünkü tag incelemelerinde (obj-enobj) object içindeki taglar başka (obj-enobj) objectlere referans eder. Bu referansı “obj header” değerine göre yapar. Bu objectleri görüntülemek için kullanılır. Örneğin yukarıdaki resimde sağ üst taraftaki **obj header : “1 0 obj”** yazar. Bu “Adobe Acrobat” ın pdf yorumlarken kullandığı referans numarasıdır. Örneğin bu object nesnesini bu seçenekten çağırmak için ilk sayıyı yani 1 girmek yeterlidir.

C:\WINDOWS\py.exe

```
<<obj-endobj>> 135 adet bulundu.
```

```
Geri (Back) <<b>>, Çıkış (Exit) <<e>>
```

```
Obj Header (First No) : 1
```

C:\WINDOWS\py.exe

```
<<Obj No : 0>> -
```

```
Obj header : "1 0 obj"
```

```
<<
  /Type
  /Catalog
  /Pages 2 0 R
  /Lang(en-US)
  /StructTreeRoot 119 0 R
  /MarkInfo
    <<
      /Marked true
    >>
  >>
```

```
>>
```

```
Devam (Continue) <<enter>> :
```


4) Show Object (Stream Summary) : Ana menüdeki 2, 3, 5 seçeneklerinde (obj-enobj) objectler stream dataları ile beraber gösterilir. Bu stream dataları bazen resim, gibi çeşitli data değerleri içerdiği için sayfada aşırı gereksiz bilgiyle dolabiliyor. Stream dataları program tarafından analiz ediliyor. Şıkıştırılmış olanlar açılıp gösteriliyor. Stream bilgilerinin ilk satırı yalnızca 20 baytı gösterilir. Bu sayede stream datasında boğulmadan rahatça objectler incelenebilir. Daha önce belirtildiği gibi bu seçenekte objectler “object no” ile çağrılır.

```
C:\WINDOWS\py.exe

<<Obj No : 3>> -      Obj header : "4 0 obj"
-----
<<
  /Filter
  /FlateDecode
  /Length 748
>>
/*** STREAM (No:0) ***
-----

Stream No          : 0
Stream Start-End Index : 705-1472
Stream Length      : 2551 bytes
Stream Data Type    : Zlib
Zlib Stream Marker  : "78 9c ad 56"

Stream Data (Only first 20 bytes)
No:000001-000020    <Ascii>|| /P <</MCID 0>> BDC ||      <Hex>||20 2f 50 20 3c 3c 2f 4d 43 49 44 20 30 3e 3e 20 42 44 43 20||

Devam (Continue) <<enter>> :
```

5) Show Object (Normal) : Objectler obj no ile çağrılır (obj no sıfırdan başlar). Objecti ve içindeki tüm stream datasını gösterir. Bu getirilen stream datasını analiz edilmiş stream datasıdır.

6) Show Object (Hex) : Objectleri ve içindeki stream datalarını ham olarak gösterir.

5 menüden elde edilen ekran görüntüsü ile 6 menüden elde edilen ekran görüntüsü aşağıda gösterilmiştir. (İki menüde görüntülenen object aynıdır.)


```
C:\WINDOWS\py.exe

<<Obj No : 3>> -      Obj header : "4 0 obj"
-----
<<
/Filter
/FlateDecode
/Length 748
>>
/*** STREAM (No:0) ***
-----

Stream No      : 0
Stream Start-End Index : 705-1472
Stream Length  : 2551 bayts
Stream Data Type : Zlib
Zlib Stream Marker : "78 9c ad 56"
-----
|| Stream - Size: 2551 (bayt) ||
-----
No:000001-000020 <Ascii>| /P <</MCID 0>> BDC || <Hex>| 20 2f 50 20 3c 3c 2f 4d 43 49 44 20 30 3e 3e 20 42 44 43 20 ||
No:000021-000040 <Ascii>| /GS5 gs...1 g..575.65 || <Hex>| 2f 47 53 35 20 67 73 0d 0a 31 20 67 0d 0a 35 37 35 2e 36 35 ||
No:000041-000060 <Ascii>| 101.84 m..457.99 53 || <Hex>| 20 31 30 31 2e 38 34 20 6d 0d 0a 34 35 37 2e 39 39 20 35 33 ||
No:000061-000080 <Ascii>| .671 345.58 39.44 25 || <Hex>| 2e 36 37 31 20 33 34 35 2e 35 38 20 33 39 2e 34 34 20 32 35 ||
No:000081-000100 <Ascii>| 2.65 39.44 c..202.44 || <Hex>| 32 2e 36 35 20 33 39 2e 34 34 20 63 0d 0a 32 30 32 2e 34 34 ||
No:000101-000120 <Ascii>| 39.44 157.47 43.819 || <Hex>| 20 33 39 2e 34 34 20 31 35 37 2e 34 37 20 34 33 2e 38 31 39 ||
No:000121-000140 <Ascii>| 120.75 49.292 c..17 || <Hex>| 20 31 32 30 2e 37 35 20 34 39 2e 32 39 32 20 63 0d 0a 31 37 ||
No:000141-000160 <Ascii>| 0.21 38.345 234.66 2 || <Hex>| 30 2e 32 31 20 33 38 2e 33 34 35 20 32 33 34 2e 36 36 20 32 ||
```

Şekil 7) 5 menüden elde edilen ekran görüntüsü

```
C:\WINDOWS\py.exe

-----
|| Obj (index:3) - Size: 820 (bayt) ||
-----
No:000660-000679 <Ascii>| 4 0 obj.<</Filter/F || <Hex>| 34 20 30 20 6f 62 6a 0d 0a 3c 3c 2f 46 69 6c 74 65 72 2f 46 ||
No:000680-000699 <Ascii>| lateDecode/Length 74 || <Hex>| 6c 61 74 65 44 65 63 6f 64 65 2f 4c 65 6e 67 74 68 20 37 34 ||
No:000700-000719 <Ascii>| 8>>..stream..x..V.O. || <Hex>| 38 3e 3e 0d 0a 73 74 72 65 61 6d 0d 0a 78 9c ad 56 df 4f 1a ||
No:000720-000739 <Ascii>| A.~'....;.[v.f.%... || <Hex>| 41 10 7e 27 e1 7f 98 c7 3b 93 5b 76 f6 66 7f 25 c6 b4 88 da ||
No:000740-000759 <Ascii>| 65.J....'A.QH...w.8 || <Hex>| 36 35 b1 4a d2 87 a6 0f 27 41 a4 51 48 91 d6 f8 df 77 e6 38 ||
No:000760-000779 <Ascii>| ....@H.f....o..... || <Hex>| cb a1 e0 fa 40 48 e0 66 99 fd f6 9b 6f be 1d 80 de 05 1c 1e ||
No:000780-000799 <Ascii>| ...?.@.A.p...+...n. || <Hex>| f6 ce 8f 3f 0f 40 1f 1d 41 7f 70 0c bd b3 2b 0b 93 87 6e 07 ||
No:000800-000819 <Ascii>| a..Xo....U ..v.z.#.R || <Hex>| 61 d2 ed 58 6f 95 b3 80 1a 55 20 b8 ef 76 c8 7a 15 23 d8 52 ||
No:000820-000839 <Ascii>| 9.P.U6@....k$w..... || <Hex>| 39 8f 50 92 55 36 40 19 15 11 18 6b 24 77 15 8c ba 1d a3 8d ||
No:000840-000859 <Ascii>| <.b.}.J.0.....D#.. || <Hex>| 3c ad 62 e4 7d e4 81 4a 15 30 02 1a ad bc 05 8a ca 44 23 b9 ||
No:000860-000879 <Ascii>| ..2...C.)I9.&2..R... || <Hex>| e8 b5 32 8c 17 14 43 82 29 49 39 07 26 32 b8 87 52 07 15 fe ||
No:000880-000899 <Ascii>| G...\4=P.5.V.$h.f.L || <Hex>| 47 9c 5c 86 a8 34 3d 2f 50 d0 35 b2 56 b6 24 68 08 7b a3 4c ||
```

Şekil 8) 6 menüden elde edilen ekran görüntüsü

7) Show Stream (Normal) : Object içerisinde sadece stream datasını görüntülemek için kullanılır. Burada işlenmiş stream dataları görüntülenir. Streamler “stream no” göre çağrılır. (Stream no sıralamaya 0 başlar.)

8) Show Stream (Hex) : Object içerisindeki stream datasını hex editörde görüntüler. Stream datasının ham hali gösterilir. Streamler “stream no” göre çağrılır. (Stream no sıralamaya 0 başlar.)

9) Write Stream : Stream datalarını dosya olarak kaydetmeye izin verir. Bu sayede pdf içindeki resimleri, ek dosyaları (zararlı script, vb.), sıkıştırılmış dosyaları vb dışarıya çıkarmak için kullanılır. Yazılan data dosyaları, scriptlerin bulunduğu klasörün içine yazılır.

C:\WINDOWS\py.exe

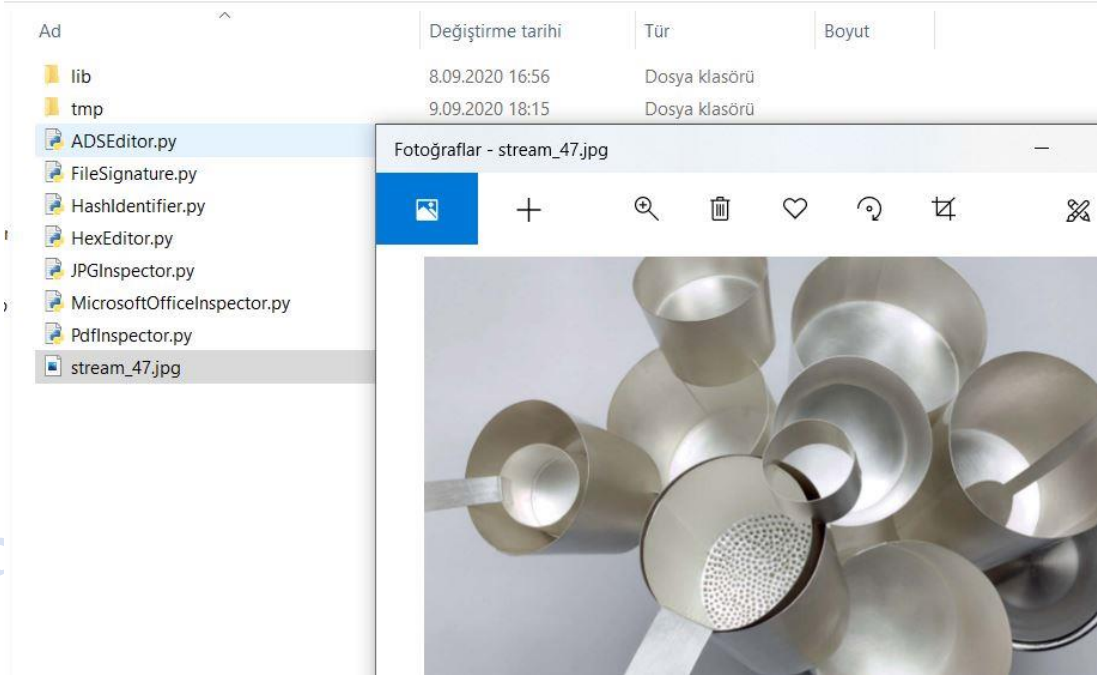
```
<<stream-endstream>> 127 adet bulundu.  
No: 0 - 126
```

```
Geri (Back) <<b>>, Çıkış (Exit) <<e>>  
(Seçim) Selection No : 47
```

C:\WINDOWS\py.exe

```
Dafult file name 'stream_47' <<enter>>  
New File Name : stream_47.jpg
```

u bilgisayar > GungorX (Y:) > 5-Projeler > File_Research_Library



Pdf içindeki 47 nolu stream datasının seçilerek dataya stream_47.jpg ismi verilerek dosya olarak çıkarılmıştır. Stream'ı dosya olarak çıkarıldıktan sonra dosya türü bilinmiyorsa "Filesignature.py" kullanılarak dosya türü belirlenebilir. Ben özellikle jpg datasını içeren bir stream seçtim.

Pdf'in sonunda xref, trailer, startxref, eof tagları sırayla bulunur. Ancak bazı pdf dosyalarında tagların bazıları eksik olabilir. Bu pdf oluşturduğunuz programla ilgilidir. Bu tag sayısının 1 den fazla olması bu pdf'e sonradan ekleme yapıldığını gösterir. Bu bilgi zararlı tespitinde kullanılan ufak püf noktalarından birisidir. (Ancak her ekleme yapılan pdf zararlı değildir.) İlgili tagları hex editörde görüntülenmek için 11, 12, 13 seçenekleri kullanılır. Genelde geçerli obj numaraları , /encrypt gibi bazı tag ların ve bilgilerin bulunduğu bölümdür. (Bazı pdf uygulamaları, kişisel bilgileri [metadata bilgileri] standart dışı olarak pdf sonuna gömebiliyor.)

10) Show Xref (Hex) : Pdf sonunda yer alan xref taglarını içeren bölümü hex editörde gösterir.

11) Show Trailer (Hex) : Pdf sonunda yer alan trailer taglarını içeren bölümü hex editörde gösterir.

12) Show StartXref (Hex) : Pdf sonunda yer alan startxref taglarını içeren bölümü hex editörde gösterir.

Abdulkadir Güngör tarafından hazırlanmıştır.