

Funky
File
Formats

2014/12 3103

ANGE ALBERTINI



ANGE ALBERTINI

reverse engineering

VISUAL DOCUMENTATIONS

@angealbertini

ange@corkami.com

<http://www.corkami.com>



a file is:

So, this talk is about files... what are the usual files' categories?



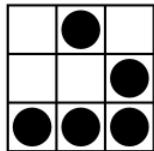
?



✓ VALID



✗ CORRUPTED



EXPLORER

It depends if you're a newbie, a user, a dev, a hacker...



BØRNING VALID

A large, stylized word 'BØRNING' is written in red, bold letters. A question mark is positioned above the letter 'O'. Below the letter 'O' is a green checkmark. To the right of 'BØRNING', the word 'VALID' is written in green, bold letters.

...but in general, valid files aren't very sexy!



VALID



CORRUPTED

However, the frontier between valid and corrupted is not straight and clear !

Here is a *valid* file...

f76f5dafdcf0818c457e6ffb50ea61a67196dcd4 *ccc.jpg

(ok, maybe not a *standard* file)



This is a JPEG picture...

```
>java -jar ccc.jpg  
Hello world! [Java]
```

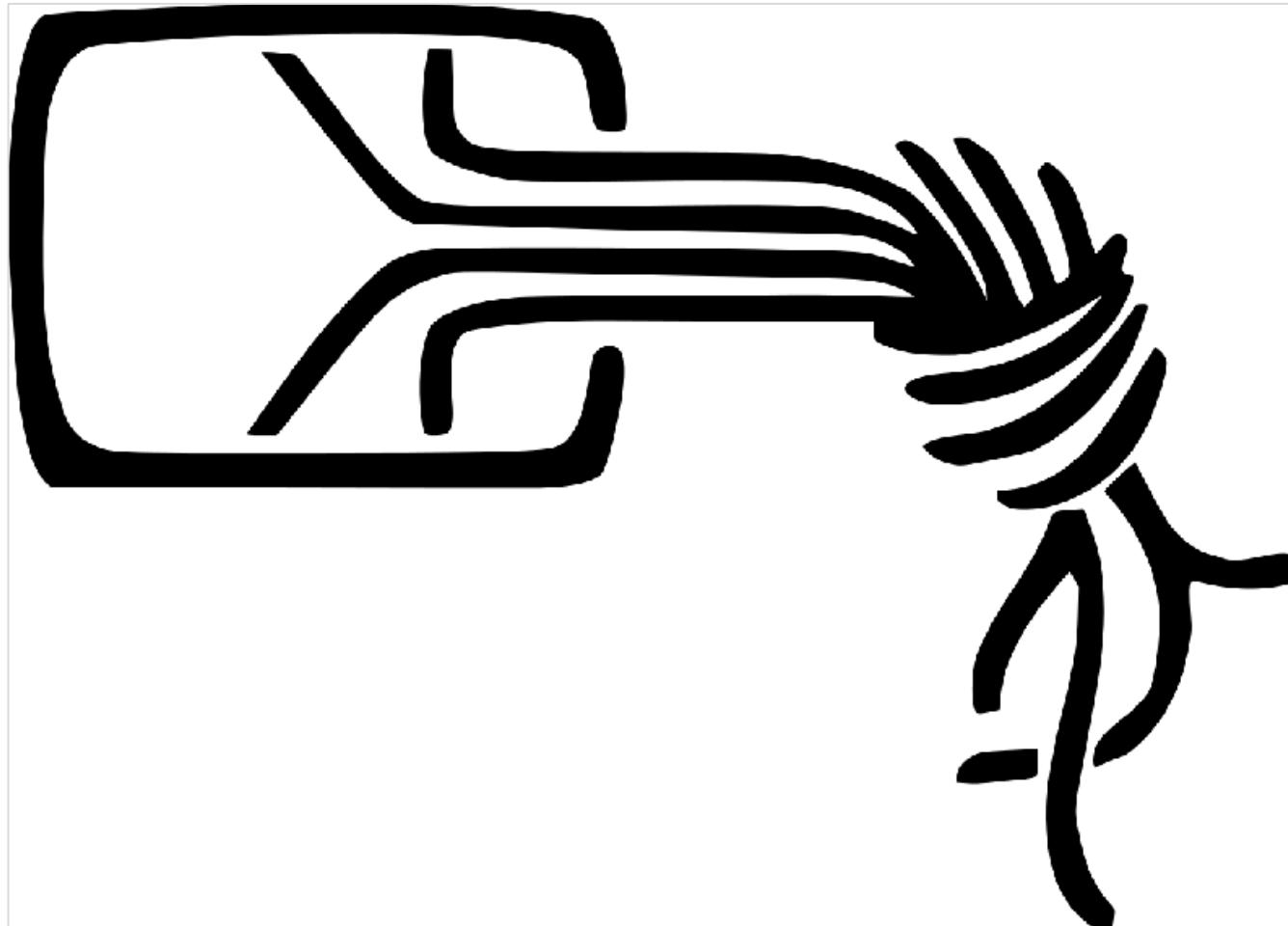
>

...that's also a Java file.

AES()



If you encrypt it with AES...



... you get a PNG picture.

3DES()



If you **decrypt** it with *Triple DES*...



...you get a PDF document.

AES_{K₂}

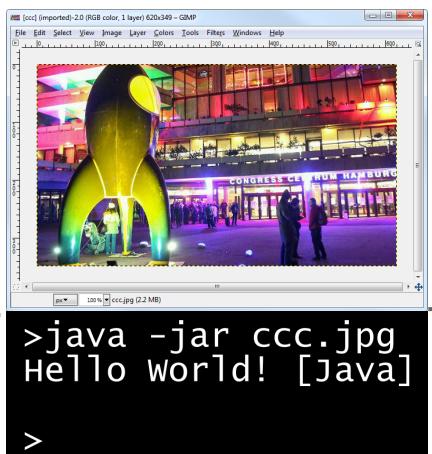


If you encrypt the original file with AES again, but with a **different key**...



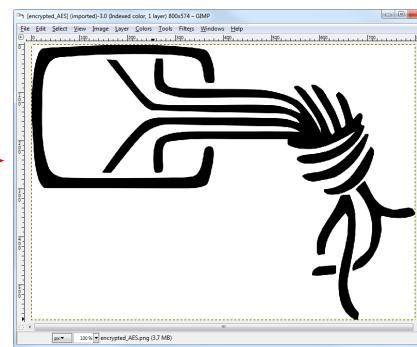
...you get a Flash Video...
..that ... oh well, nevermind, I could go on for hours...

JPG



JAR
(ZIP + CLASS)

\rightarrow
 AES_{K_1}



PNG

\rightarrow
 AES_{K_2}

PDF

\downarrow
3DES



FLV

So, as you can see, I'm just a normal guy (who likes to play with binary).

```
me@nux:~$ ./mini  
me@nux:~$ echo $?  
42
```

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
00:	7F	.E	L	.F	01	01	01									
10:	02	00	03	00	01	00	00	00	60	00	00	08	40	00	00	00
20:									34	00	20	00	01	00		
40:	01	00	00	00	00	00	00	00	00	00	00	08	00	00	00	08
50:	70	00	00	00	70	00	00	00	05	00	00	00				
60:	BB	2A	00	00	00	B8	01	00	00	00	CD	80				

ELF HEADER

IDENTIFY AS AN ELF TYPE
SPECIFY THE ARCHITECTURE

FIELDS		VALUES
e_ident		
EI_MAG	0x7F, "ELF"	
EI_CLASS, EI_DATA	1 ELFCLASS32, 1 ELFDATA2LSB	
EI_VERSION	1 EV_CURRENT	
e_type	2 ET_EXEC	
e_machine	3 EM_386	
e_version	1 EV_CURRENT	
e_entry	0x8000060	
e_phoff	0x0000040	
e_ehsize	0x0034	
e_phentsize	0x0020	
e_phnum	0001	
p_type	1 PT_LOAD	
p_offset	0	
p_vaddr	0x80000000	
p_paddr	0x80000000	
p_filesz	0x0000070	
p_memsz	0x0000070	
p_flags	5 PF_R PF_X	

PROGRAM HEADER TABLE

EXECUTION INFORMATION

CODE

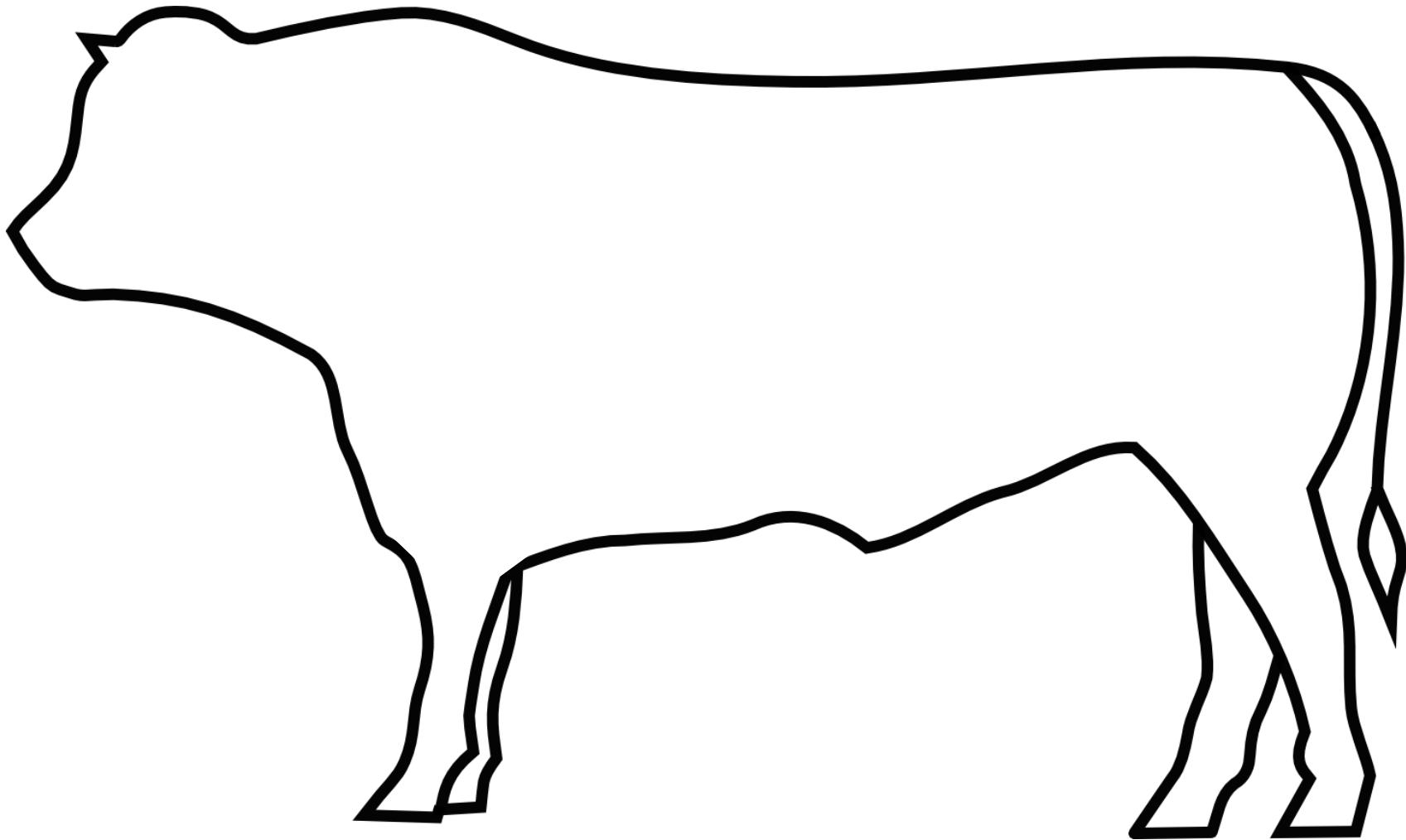
X86 ASSEMBLY

```
mov ebx, 42
mov eax, 1
int 80h
```

EQUIVALENT C CODE

```
SC_EXIT
return 42;
```

Let's talk about...

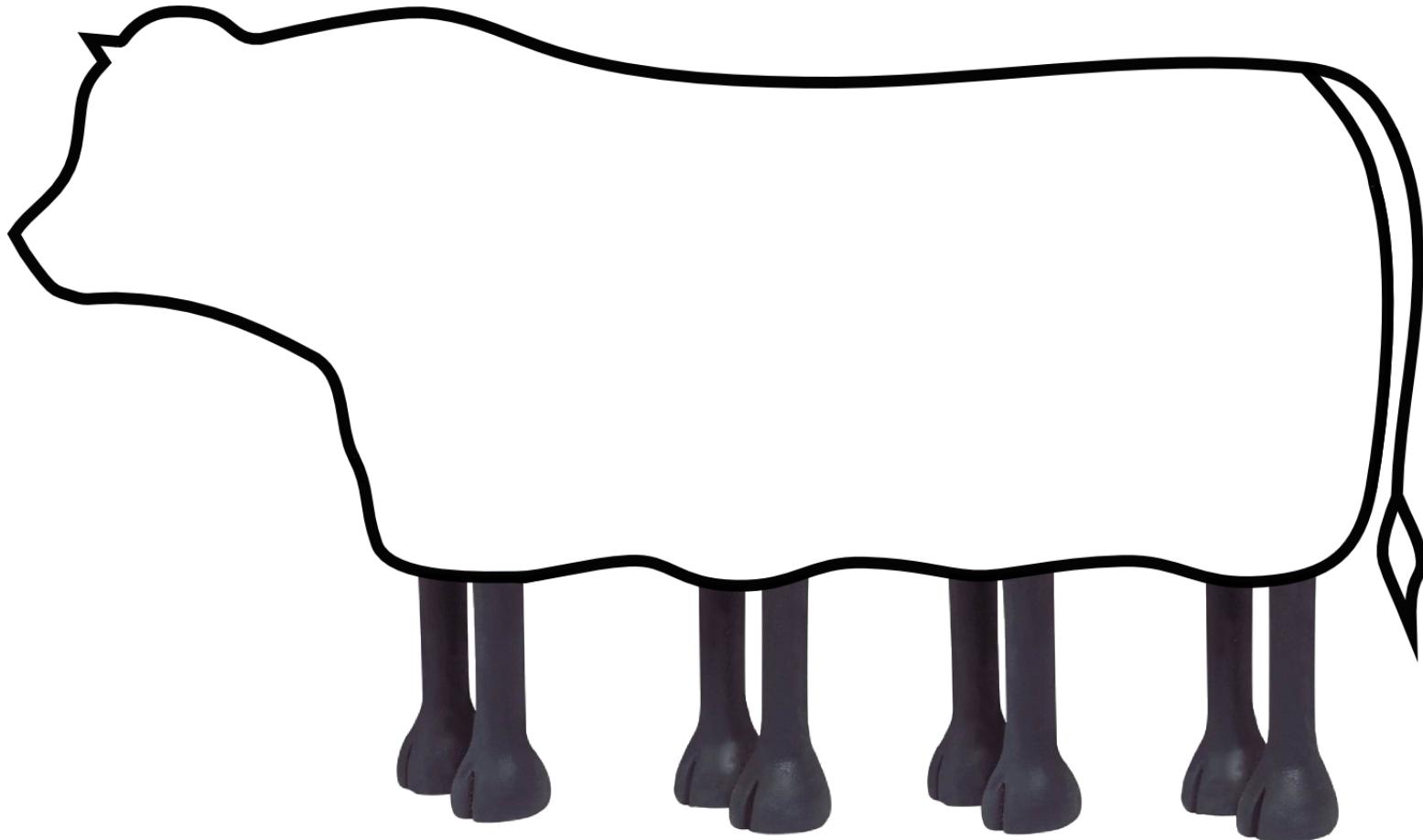


Identification

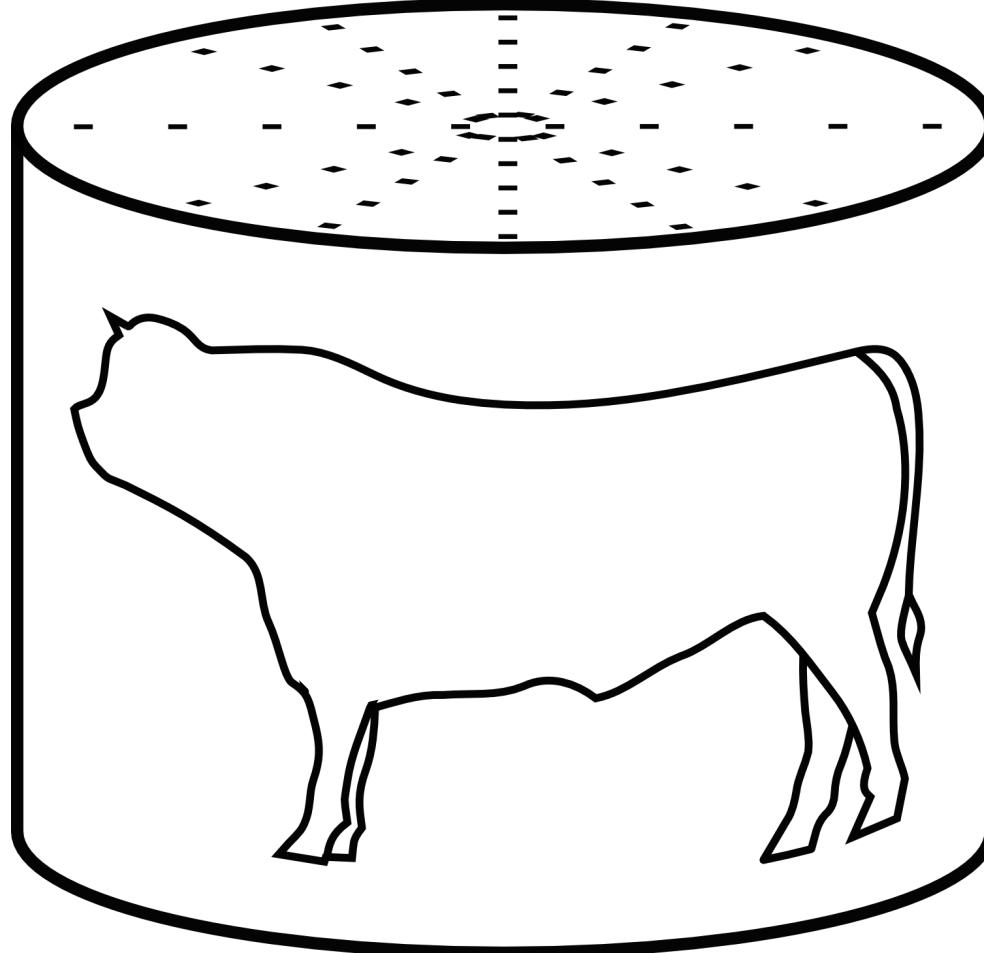
How do you identify a cow?



By its head?



By its body?



By sound?

in practice...

early filetype identifier



Obvious

PE\0\0 \x7F**ELF** BPG\xFB
\x89**PNG**\x0D\x0A\x1A\x0A
dex\n035\0 RAR\x1a\7\0 BZ
GIF89a BM RIFF

Not obvious

GZip 1F 8B
JPG FF D8

Not obvious, but 133tsp34k ^_^

CAFEBABE Java / universal (old) Mach-0
DOCF11E0 Office
FEEDFACE Mach-0
FEEDFAC**F** Mach-0 (**64b**)

Egocentric

MZ (DOS header) Mark Zbikowski
PK\3\4 (ZIP) Philip Katz
BPG\xFB Fabrice Bellard

Specific logic

TIFF:
II Intel (little) endianness
MM Motorola (big) endianness
Flash:
FWS ShockWave Flash (Flat)
CWS (zlib) compressed
ZWS LZMA compressed

not enforcing signature at offset 0: ZIP, 7z, RAR, HTML
actually enforcing signature at offset 0: bzip2, GZip

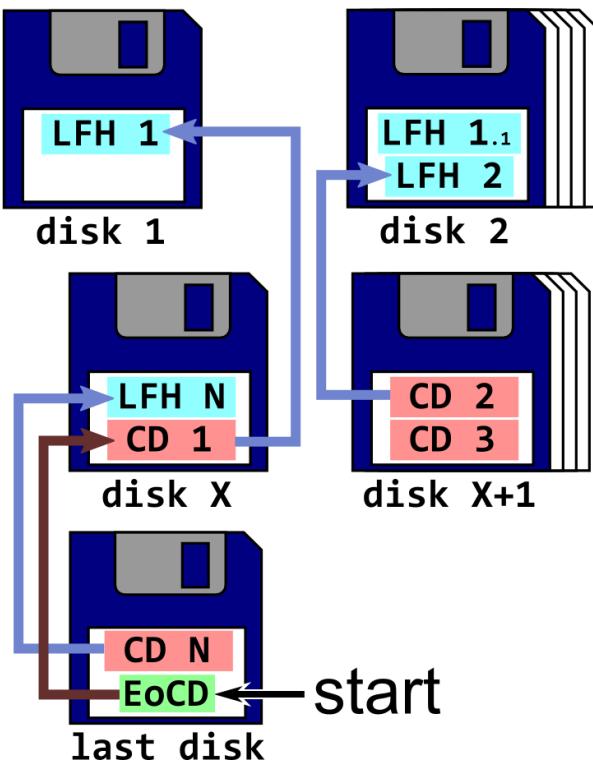
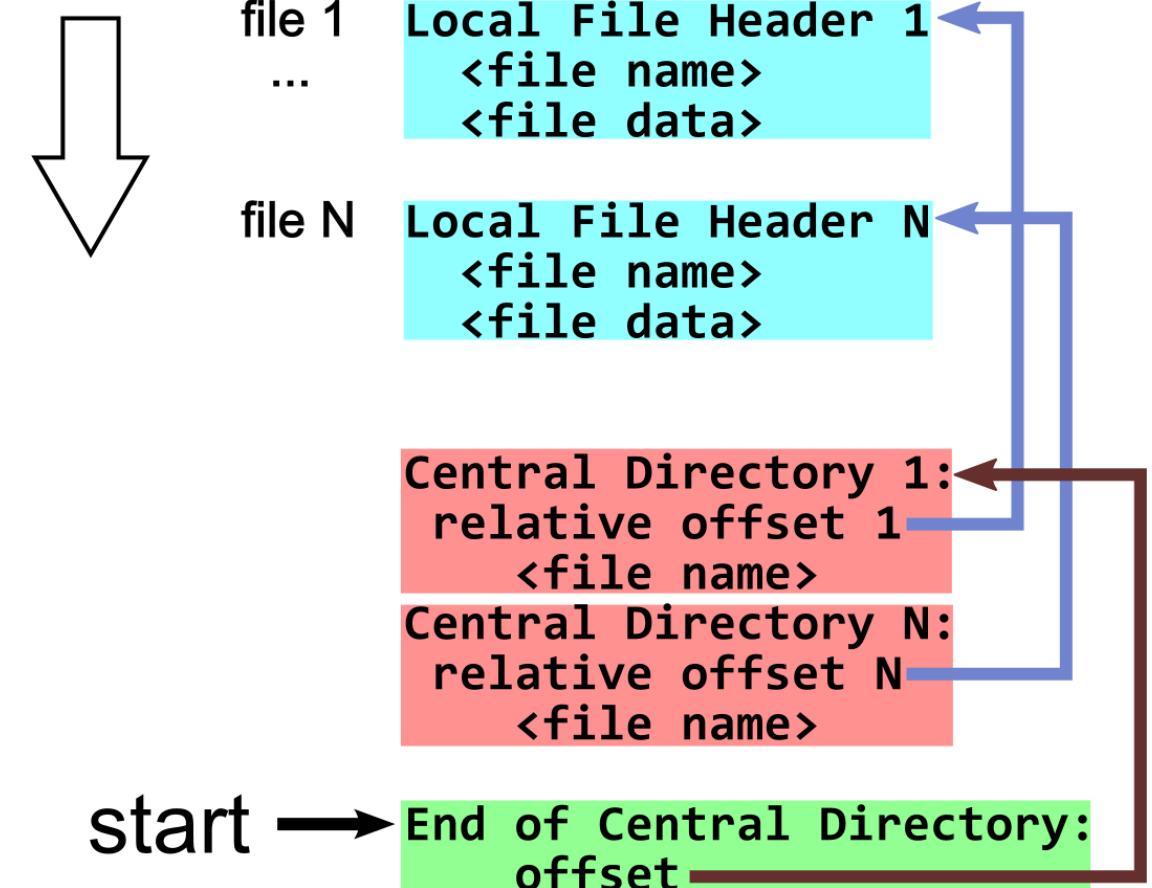
7.5.2 File Header

The first line of a PDF file shall be a *header* consisting of the 5 characters %PDF– followed by a version number of the form 1.N, where N is a digit between 0 and 7.

3.4.1, “File Header”

13. Acrobat viewers require only that the header appear somewhere within the **first 1024 bytes of the file**.

File formats not enforcing signature at offset 0
(ZIP is used in **many** formats: APK, ODT, DOCX, JAR...)



ZIP actually enforces “finishing” near the end of the file.

- TAR: Tape Archive
- Disk images: ISO, Master Boot Record
- TGA (image)
- (Console) roms

Hardware-bound formats: code/data at offset 0
'header' often (optionally) later in the memory space

a good magic signature:

- enforced at offset 0
- unique

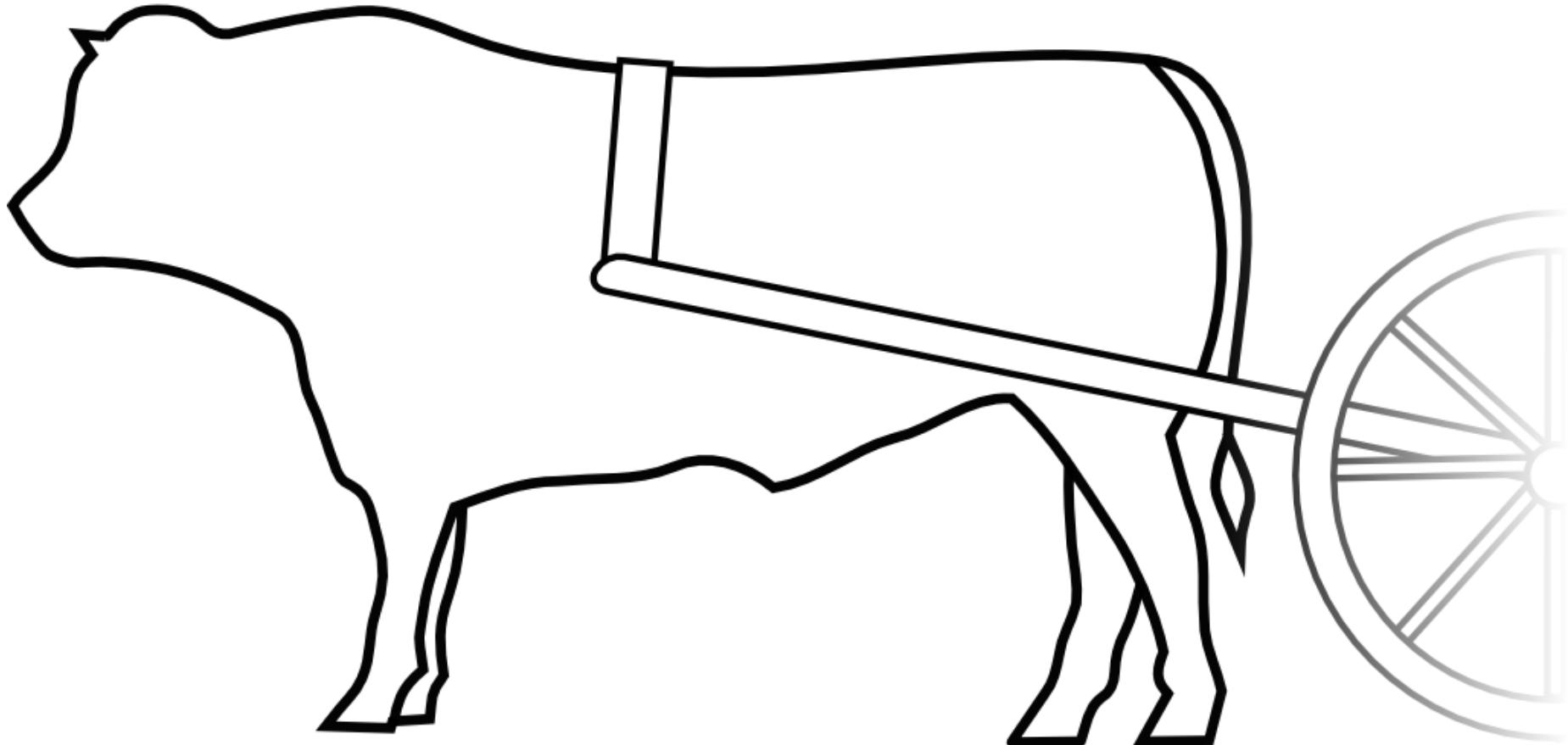
no magic ⇒ no excuse

Standard tool: checks magic,
chooses path, never returns...



Another common yet important property

(useful for abuses)

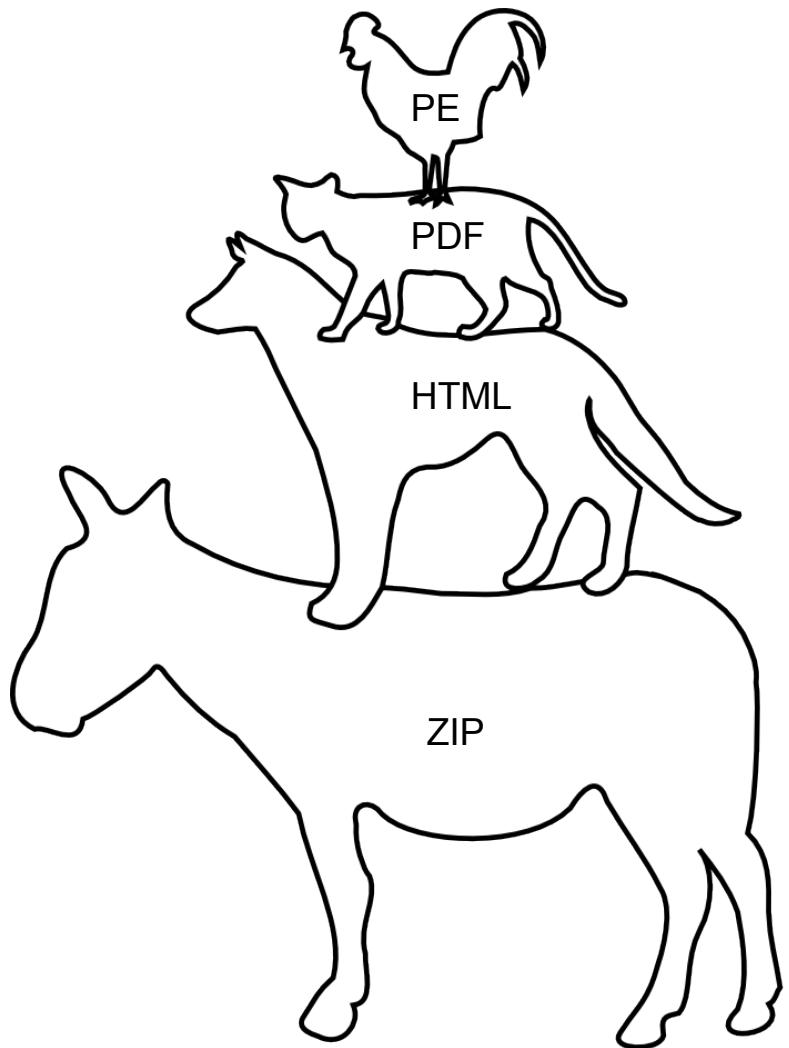


It's a complete cow (you can see its whole body), with something next:
 appending something doesn't invalidate the start.



Remember:
there's nothing to parse
after the terminator.

formats not enforced at offset 0
+ tolerating appended data
= **Polyglots by concatenation**



a JAR(JAR) || BINK polyglot
JAR = ZIP(CLASS)

```
ca: a JAR JAR BINK polyglot

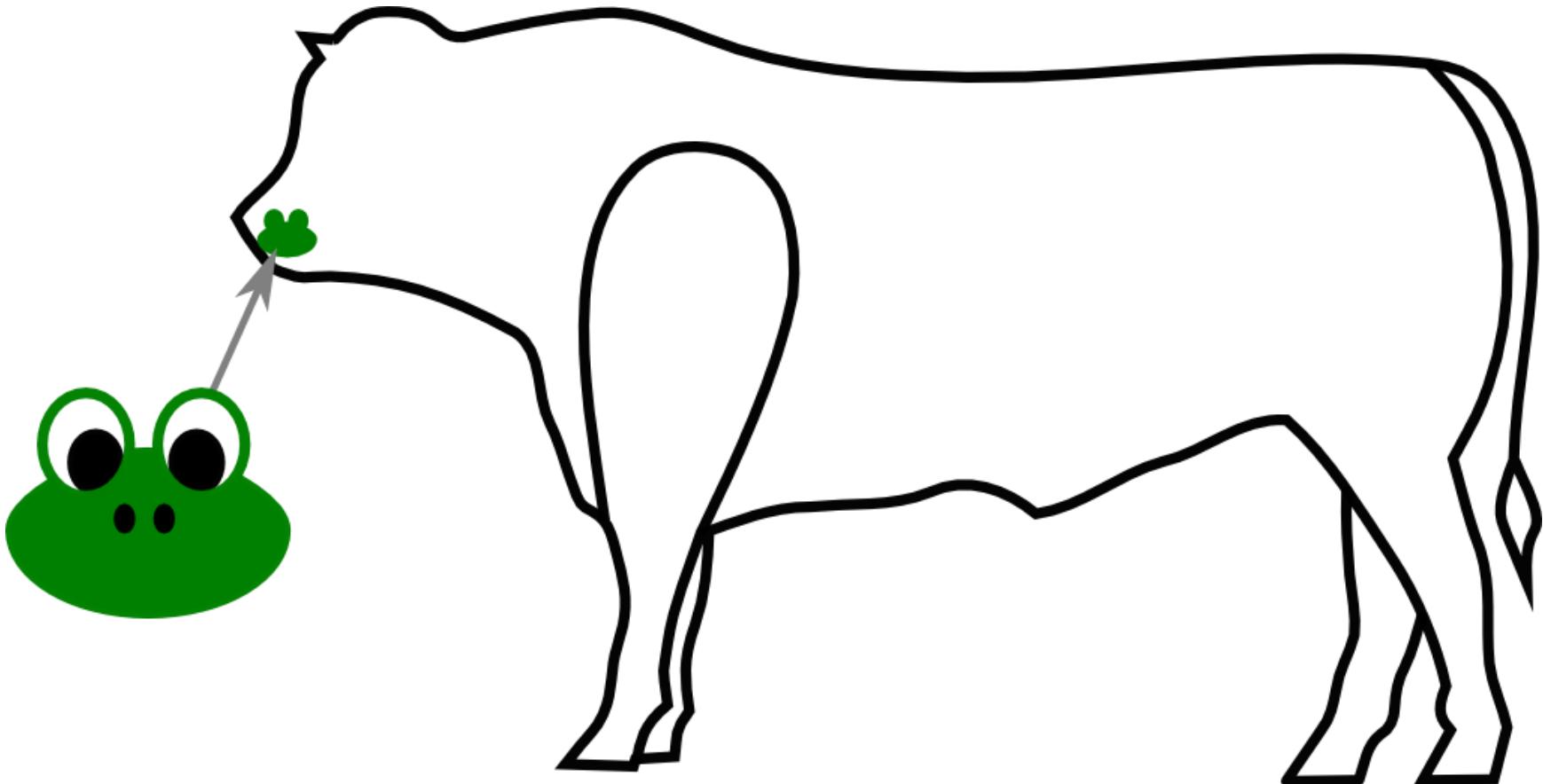
>java -jar bink.jar
Hello World!

>unzip bink.jar gungan.jar
Archive: bink.jar
warning [bink.jar]: 42732 extra bytes at beginning or within zipfile
(attempting to process anyway)
  inflating: gungan.jar

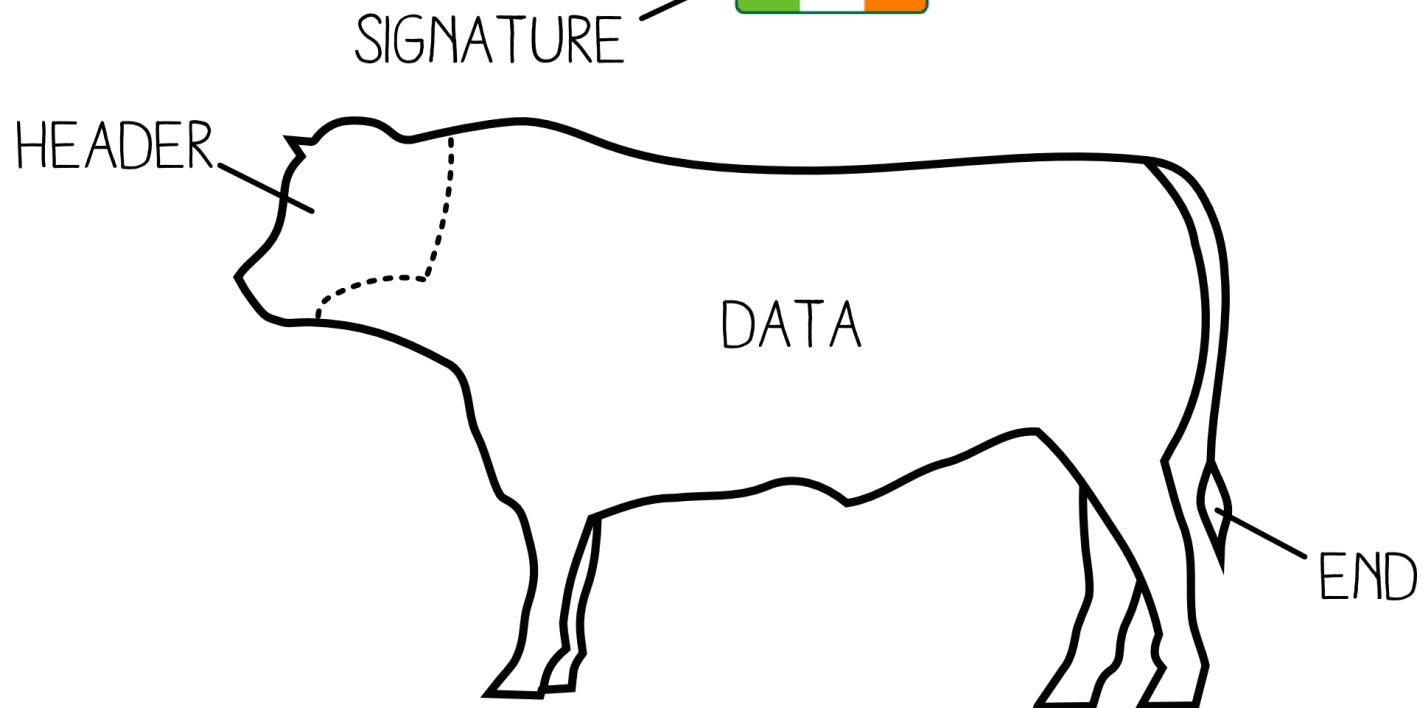
>java -jar gungan.jar
Mesa called Jar Jar Binks!
>
```



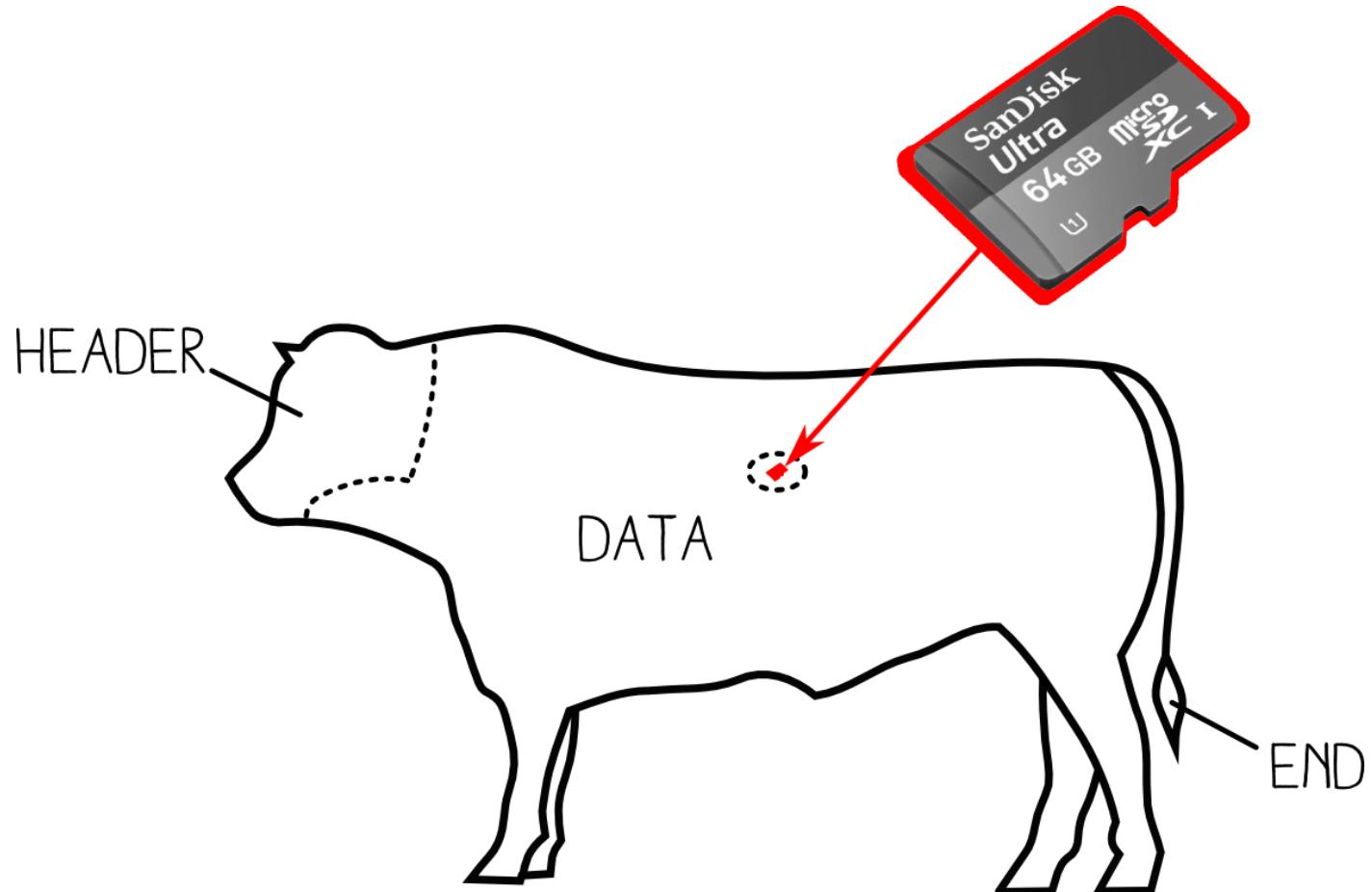
“host/parasite” polyglots



If a cow keeps a frog in its mouth, it can also speak 2 languages!
(the outer leaves space for an inner)

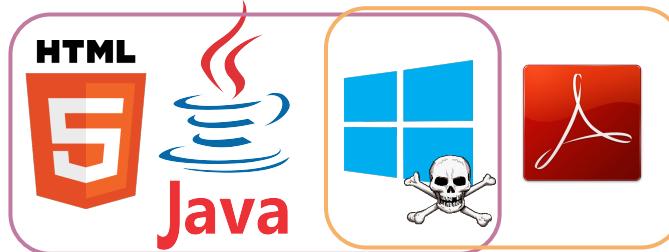


Ok, I know... here is a more realistic analogy...



...if our cow swallows a microSD, it's still a valid cow!
Even if it contains foreign data, that is tolerated by the system.

2 infection chains in one file:



```
0001. CONSTANT_Class : corkamix  
0002. CONSTANT_Utf8 : corkamix  
0003. CONSTANT_Class : java/lang/Object  
0004. CONSTANT_Utf8 : java/io/PrintStream
```

Edit CONSTANT_Utf8

```
endstreamendobj1 0 obj<</Kids[<</Parent 1 0 R/Contents[2 0 R]  
>>] /Resources<<>>> 2 0 obj<<>>>streamBT/default 80 Tf 1 0 0 1 1 715 Tm  
(CorkaMIX [PDF])Tj ETendstreamendobjtrailer <</Root<</Pages 1 0 R>>>>
```

Show References
Cancel
Save

```
0015. CONSTANT_Utf8 : CorkaMIX [Java CLASS in JAR]  
0016. CONSTANT_Methodref : class: java/io/PrintStream, name: println, descriptor: (Ljava/lang/String;)V  
0017. CONSTANT_Class : java/io/PrintStream  
0018. CONSTANT_Utf8 : java/io/PrintStream  
0019. CONSTANT_NameAndType : name: println, descriptor: (Ljava/lang/String;)V  
0020. CONSTANT_Utf8 : println  
0021. CONSTANT_Utf8 : (Ljava/lang/String;)V  
0022. CONSTANT_Utf8 : endstreamendobj1 0 obj<</Kids[<</Parent 1 0 R/Contents[2 0 R]>>] /Resources<<>>>> 2 0
```

the PDF part is stored in a Java buffer

```
>corkamix.exe  
CorkaMIX [PE]  
>java -jar corkamix.exe  
CorkaMIX [Java CLASS in JAR]
```

```
>cmp -b corkamix.exe corkamix_1b.exe  
cmp: EOF on corkamix.exe
```

```
>python corkamix_1b.exe  
CorkaMIX [python]
```

```
>copy corkamix.exe corkamix.html  
1 file(s) copied.
```

corkamix.exe - Adobe Reader

CorkaMIX [PDF]

corkamix.html

JavaScript Alert

CorkaMIX [HTML+JavaScript]

OK

	Offset	0 1 2 3 4 5 6 7	8 9 A B C D E F	Ascii	
	00000000	47 49 46 38 39 61 2F 2A	0A 00 00 FF 00 2C 00 00	GIF89a/*.....,..	<-Format data
	00000010	00 00 2F 2A 0A 00 00 02	00 3B 2A 2F 3D 31 3B 61	.../*....;*/=1;a	<-Format data - For...
	00000020	6C 65 72 74 28 22 48 65	6C 6C 6F 20 57 6F 72 6C	lert("Hello.Worl	<-Foreign data
	00000030	64 5C 6E 28 66 72 6F 6D	20 61 20 47 49 46 20 66	d\n(from.a.GIF.f	
	00000040	69 6C 65 29 22 29 3B		ile");	

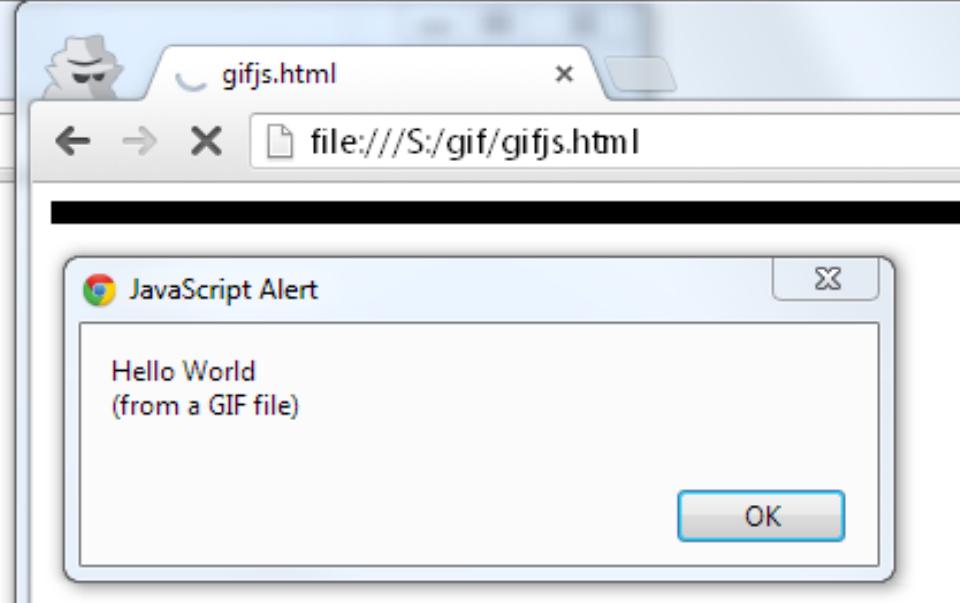
gifjs.html

```

<html><body>

<script src="gifjs.gif"></script>
</body></html>

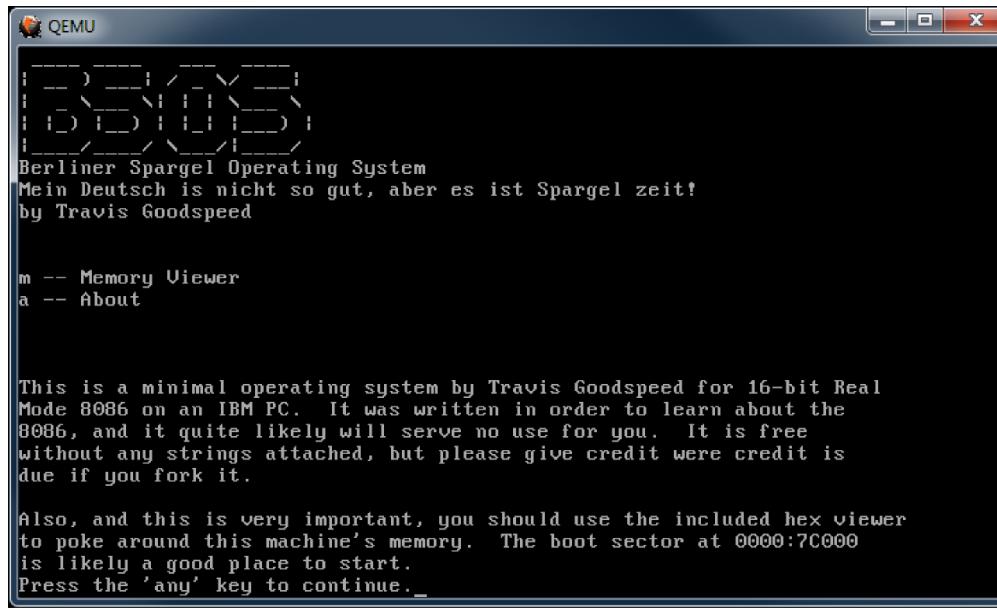
```



a JavaScript || GIF polyglot (useful for pwning - also in BMP flavor)



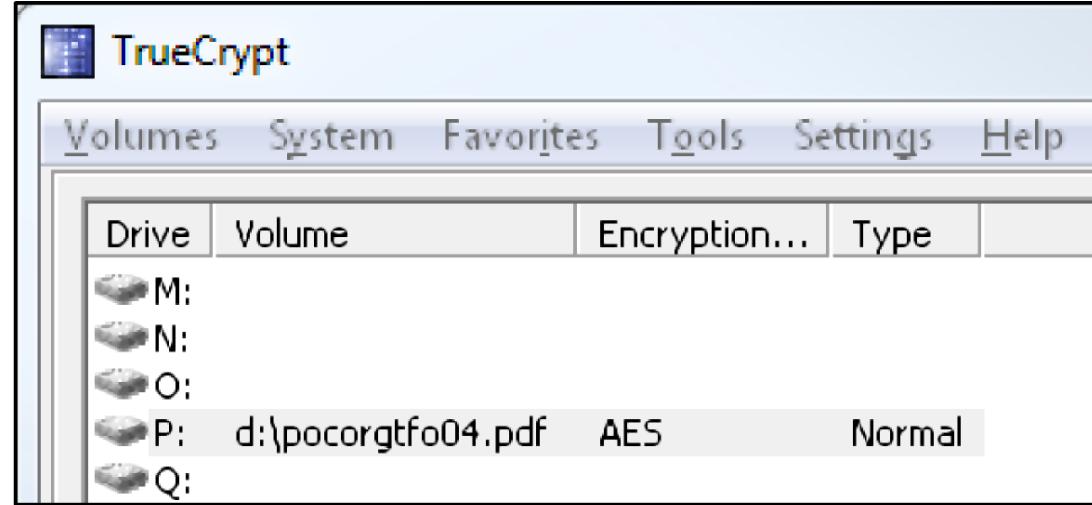
Such parasites exist already in the wild
(they just use unallocated space)



The screenshot shows a "pocorgtfo02.pdf" document viewed in Adobe Reader. The left pane shows the PDF content, which includes a cartoon illustration of a person sitting at a desk and some text. The right pane shows the file's contents as a ZIP archive:

Archive: pocorgtfo02.pdf					
warning [pocorgtfo02.pdf]: 8016414 extra bytes at beginning or within zipfile (attempting to process anyway)					
Length	EAS	ACLs	Date	Time	Name
852	0	0	12/06/13	16:25	README.txt
6794	0	0	12/06/13	16:25	coda.txt
20164	0	0	12/06/13	16:25	feeling.txt
12618	0	0	12/06/13	16:25	harrison.txt
0	0	0	12/06/13	16:25	pgpuine/
275	0	0	12/06/13	16:25	pgpuine/Makefile
1006	0	0	12/06/13	16:25	pgpuine/inflate.c
5323	0	0	12/06/13	16:25	pgpuine/quine.c
203706	0	0	12/06/13	16:25	rfc4880.txt
2046109	0	0	12/06/13	16:25	tamagotchi.zip
15565	0	0	12/06/13	16:25	theweb.txt
278598	0	0	08/05/13	13:06	pocorgtfo00.pdf
3790438	0	0	10/13/13	02:47	pocorgtfo01.pdf
6381448	0	0			-----
					13 files

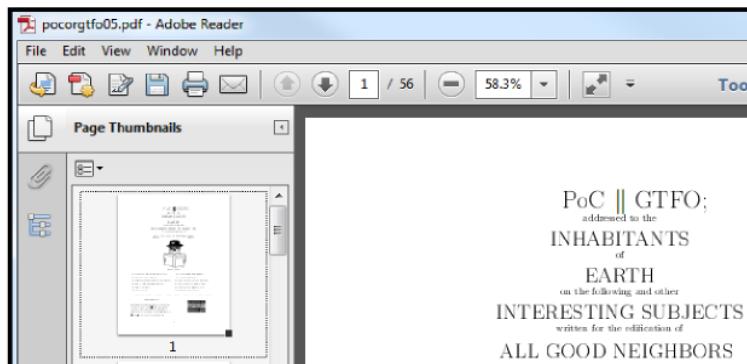
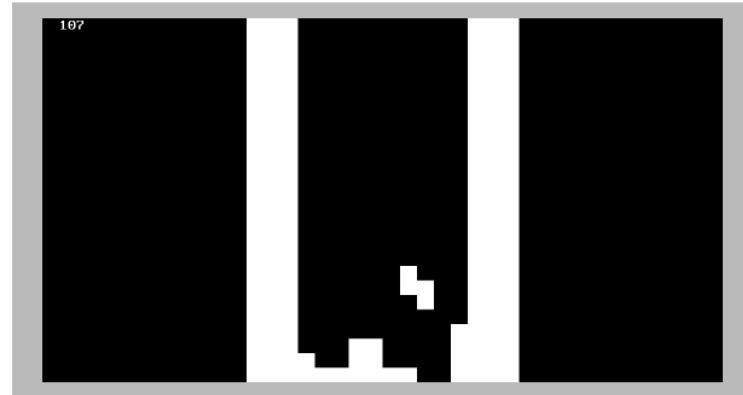
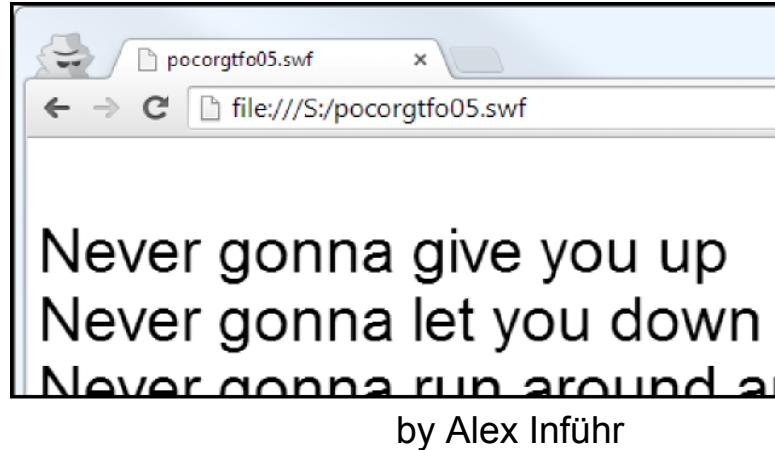
PoC||GTFO 0x2: MBR || PDF || ZIP



The Adobe Reader window shows a PDF document titled "TRACT SOCIETE SECRÈTE POC GTO". The terminal window to the right displays the contents of the ZIP file "feilles.zip.pdf".

```
Archive: pocorgtfo04.pdf
warning [pocorgtfo04.pdf]: 798586 extra bytes at beginning or within zipfile
(attempting to process anyway)
error [pocorgtfo04.pdf]: reported length of central directory is
-798586 bytes too long (Atari STZip zipfile? J.H.Holm ZIPSPLIT 1.1
zipfile?). Compensating...
Length      EAs      ACLS      Date      Time      Name
-----      --      --      --      --      --
          0      0      0 06/24/14 18:56  bin2png/
      5010      0      0 06/24/14 18:56  bin2png/bin2png.py
    18025      0      0 06/24/14 18:56  bin2png/LICENSE
     1141      0      0 06/24/14 18:56  bin2png/README.md
    140413      0      0 06/24/14 18:56  darfsteller.txt
     2841      0      0 06/24/14 18:56  gods.txt
          0      0      0 06/24/14 18:56  lenticrypt/
    36445      0      0 06/24/14 18:56  lenticrypt/lenticrypt.py
    18025      0      0 06/24/14 18:56  lenticrypt/LICENSE
       776      0      0 06/24/14 18:56  lenticrypt/README.md
      2709      0      0 06/24/14 18:56  lenticrypt/test.py
    3111965      0      0 06/24/14 18:56  pocorgtfo.png
    25986      0      0 06/24/14 18:56  thevelldt.txt
    239224      0      0 06/24/14 18:56  tsb-20140401.zip
    26750864      0      0 06/24/14 18:56  pocorgtfo03.pdf
-----      15 files
30353424      0      0
```

PoC||GTFO 0x4: TrueCrypt || PDF || ZIP

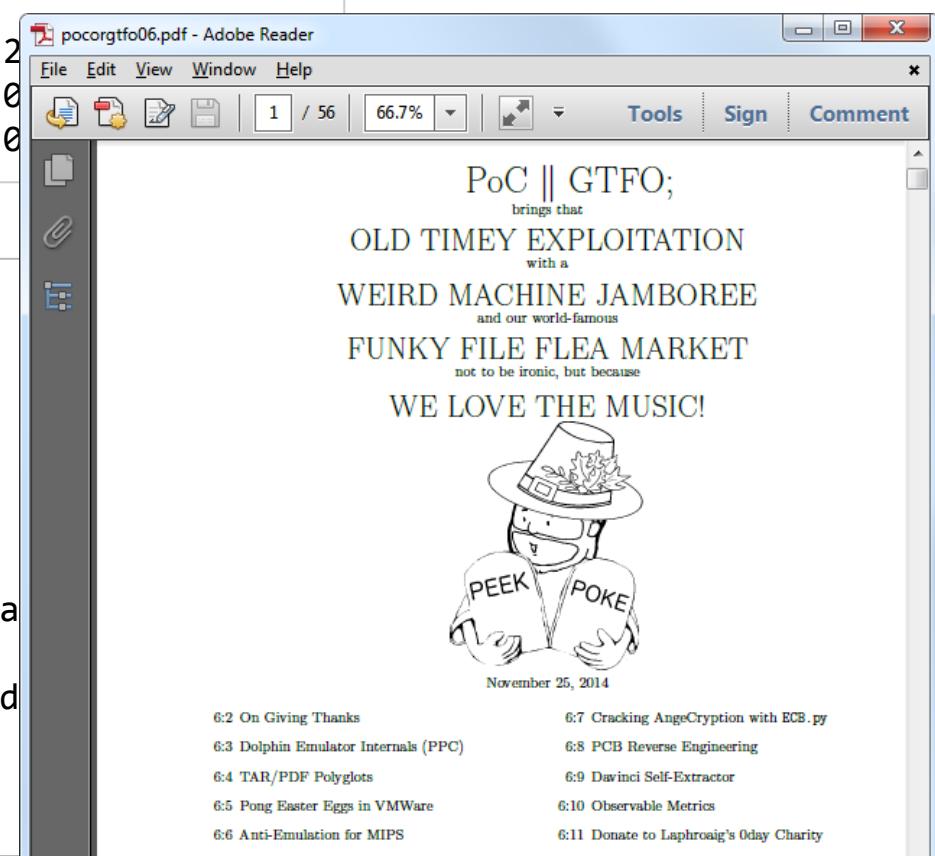


```
Archive: pocorgtfo05.pdf
warning [pocorgtfo05.pdf]:
(attempting to process an
creating: PEXternalizer/
creating: PEXternalizer/
inflating: PEXternalizer/
inflating: PEXternalizer/
inflating: PEXternalizer/
inflating: PEXternalizer/
```

PoC||GTFO 0x6: TAR || PDF || ZIP

```
$ tar -tvf pocorgtfo06.pdf
-rw-r--r-- Manul/Laphroaig    0 2014-10-06 2
-rw-r--r-- Manul/Laphroaig 525849 2014-10-0
-rw-r--r-- Manul/Laphroaig 273658 2014-10-0
```

```
$ unzip -l pocorgtfo06.pdf
Archive: pocorgtfo06.pdf
warning [pocorgtfo06.pdf]: 10672929 extra
(attempting to process anyway)
      Length      Date      Time    Name
-----  -----
        4095  11/24/2014 23:44  64k.txt
     818941  08/18/2014 23:28  acsac13_zadda
       4564  10/05/2014 00:06  burn.txt
    342232  11/24/2014 23:44  davinci.tgz.d
      3785  11/24/2014 23:44  davinci.txt
      5111  09/28/2014 21:05  declare.txt
         0  08/23/2014 19:21  ecb2/
```



unicode //

```
\u002f\u002f <html>
\u002f\u002f   <body>
\u002f\u002f     <script>
\u002f\u002f       alert('Hello World! [Javascript]');
\u002f\u002f     </script>
\u002f\u002f   </body>
\u002f\u002f </html>
```

```
public class HW
{
    public static void main(String[] args)
    {
        System.out.println("Hello World! [Java]");
    }
}
```

a Java || JavaScript polyglot (at source level)

3C	68	74	6D	6C	3E	3C	62	6F	64	79	3E	3C	73	63	72
69	70	74	3E	61	6C	65	72	74	28	27	48	65	6C	6C	6F
20	57	6F	72	6C	64	21	20	5B	4A	61	76	61	73	63	72
69	70	74	5D	27	29	3B	3C	2F	73	63	72	69	70	74	3E
3C	2F	62	6F	64	79	3E	3C	2F	68	74	6D	6C	3E	50	4B
03	04	0A	00	00	00	00	00	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	09	00	00	00	4D	45	54	41
2D	49	4E	46	2F	50	4B	03	04	0A	00	00	00	00	00	00
00	00	00	00	00	00	00	1F	00	00	00	1F	00	00	00	14
00	00	00	4D	45	54	41	2D	49	4E	46	2F	4D	41	4E	49
46	45	53	54	2E	4D	46	43	72	65	61	74	65	64	2D	42
79	3A	20	31	0D	0A	4D	61	69	6E	2D	43	6C	61	73	73
3A	20	48	57	0D	0A	50	4B	03	04	0A	00	00	00	00	00
00	00	00	00	00	00	00	00	1C	01	00	00	1C	01	00	00
00	00	00	00	CA	FE	BA	BE	00	03	00	2D	00	16	07	00
02	01	00	02	48	57	07	00	04	01	00	10	6A	61	76	61
2F	6C	61	6E	67	2F	4F	62	6A	65	63	74	01	00	04	6D
61	69	6E	01	00	16	28	5B	4C	6A	61	76	61	2F	6C	61
6E	67	2F	53	74	72	69	6E	67	3B	29	56	01	00	04	43
6F	64	65	09	00	09	00	0B	07	00	0A	01	00	10	6A	61
76	61	2F	6C	61	6E	67	2F	53	79	73	74	65	6D	0C	00
0C	00	0D	01	00	03	6F	75	74	01	00	15	4C	6A	61	76
61	2F	69	6F	2F	50	72	69	6E	74	53	74	72	65	61	6D
3B	08	00	0F	01	00	13	48	65	6C	6C	6F	20	57	6F	72
6C	64	20	21	5B	4A	61	76	61	5D	0A	00	11	00	13	07

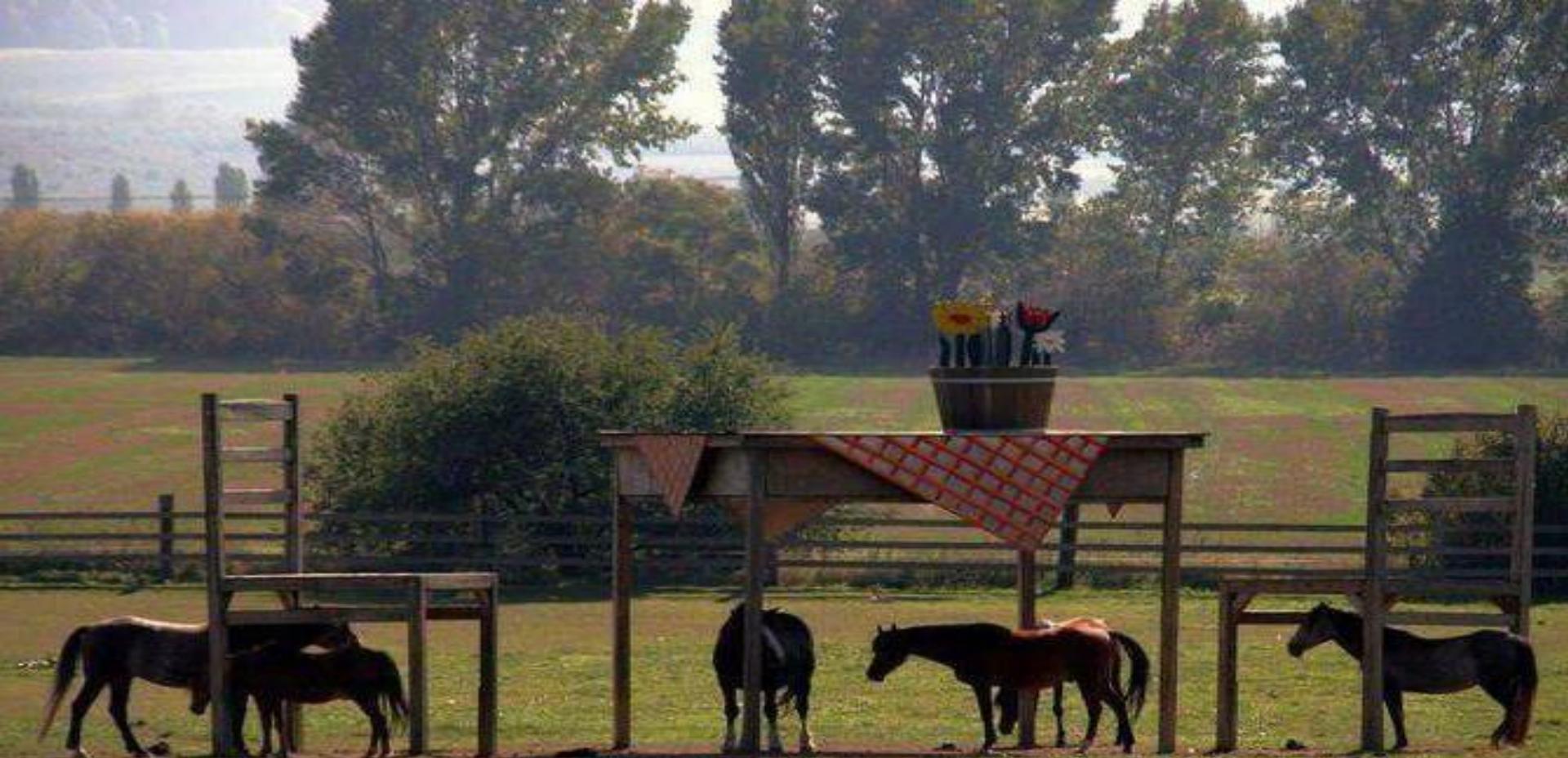
```
<html><body><script>alert('Hello  
.World!..[Javascr  
ipt]');</script>  
</body></html>PK  
.....  
.....META  
-INF/PK.....  
.....  
...META-INF/MANI  
FEST.MFCreated-B  
y:.1..Main-Class  
.HW..PK.....  
.....  
....-....  
....HW.....java  
/lang/Object...m  
ain...([Ljava/la  
ng/String;)V...C  
ode.....ja  
va/lang/System..  
....out...Ljav  
a/io/PrintStream  
;....Hello.Wor  
ld.! [Java].....
```

a Java || JavaScript polyglot (at binary level)

⇒ Java = JavaScript

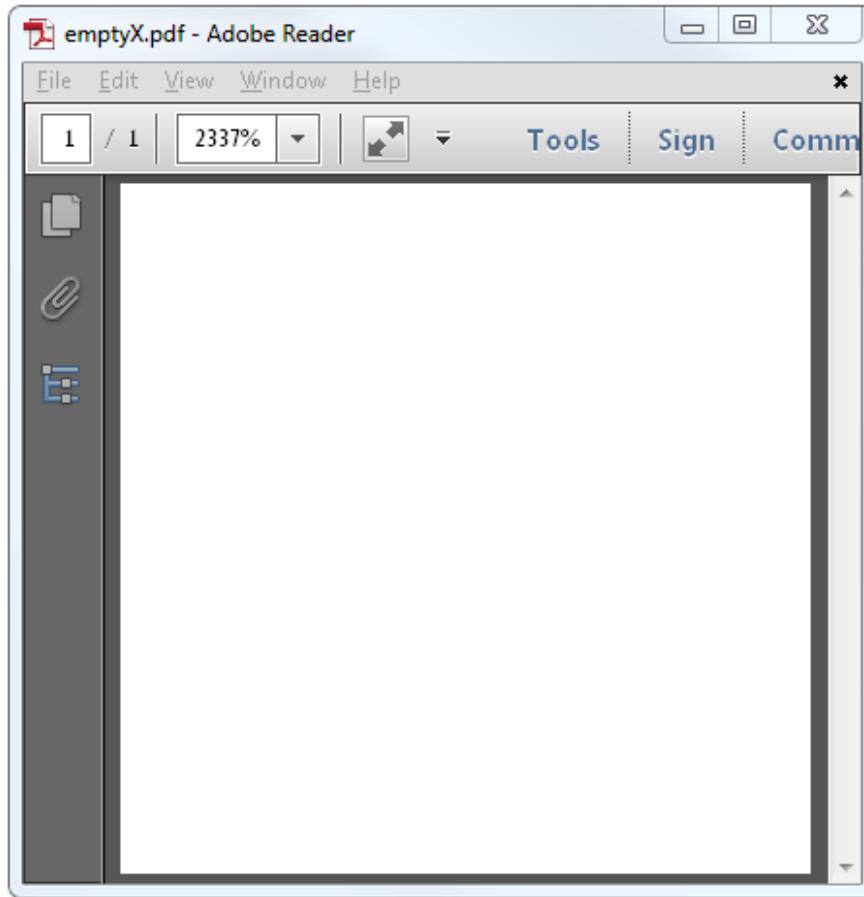
Yes, your management was right all along ;)

Extreme files bypass filters



Farmer got denied permit to build a horse shelter.
So he builds a giant table & chairs which don't need a permit.

%PDF-NULL trailer<</Root<</Pages<<>>>>



a mini PDF (Adobe-only, 36 bytes) ⇒ skipped by scanners yet valid !

```

BFF9AF2770          mov        edi, 07027AFF9 ; 'p'
B068                3mov      al, 068 ; 'h'
AA
B800102900          stosb
AB
66B8C300            stosd
AA
89D8                mov       eax, 000291000 --14
0000                add       [eax], al
0000                add       [eax], al

```

Number	Name	VirtSize	RVA	PhysSize	Offset	Flag
65524		00007000	70226000	00000000	00280200	E000000C0
65525		00007000	7022D000	00000000	00280200	E000000C0
65526		00007000	70234000	00000000	00280200	E000000C0
65527		00007000	7023B000	00000000	00280200	E000000C0
65528		00007000	70242000	00000000	00280200	E000000C0
65529		00007000	70249000	00000000	00280200	E000000C0
65530		00007000	70250000	00000000	00280200	E000000C0
65531		00007000	70257000	00000000	00280200	E000000C0
65532		00007000	7025E000	00000000	00280200	E000000C0
65533		00007000	70265000	00000000	00280200	E000000C0
65534		00007000	7026C000	00000000	00280200	E000000C0
65535		00007000	70273000	00000000	00280200	E000000C0

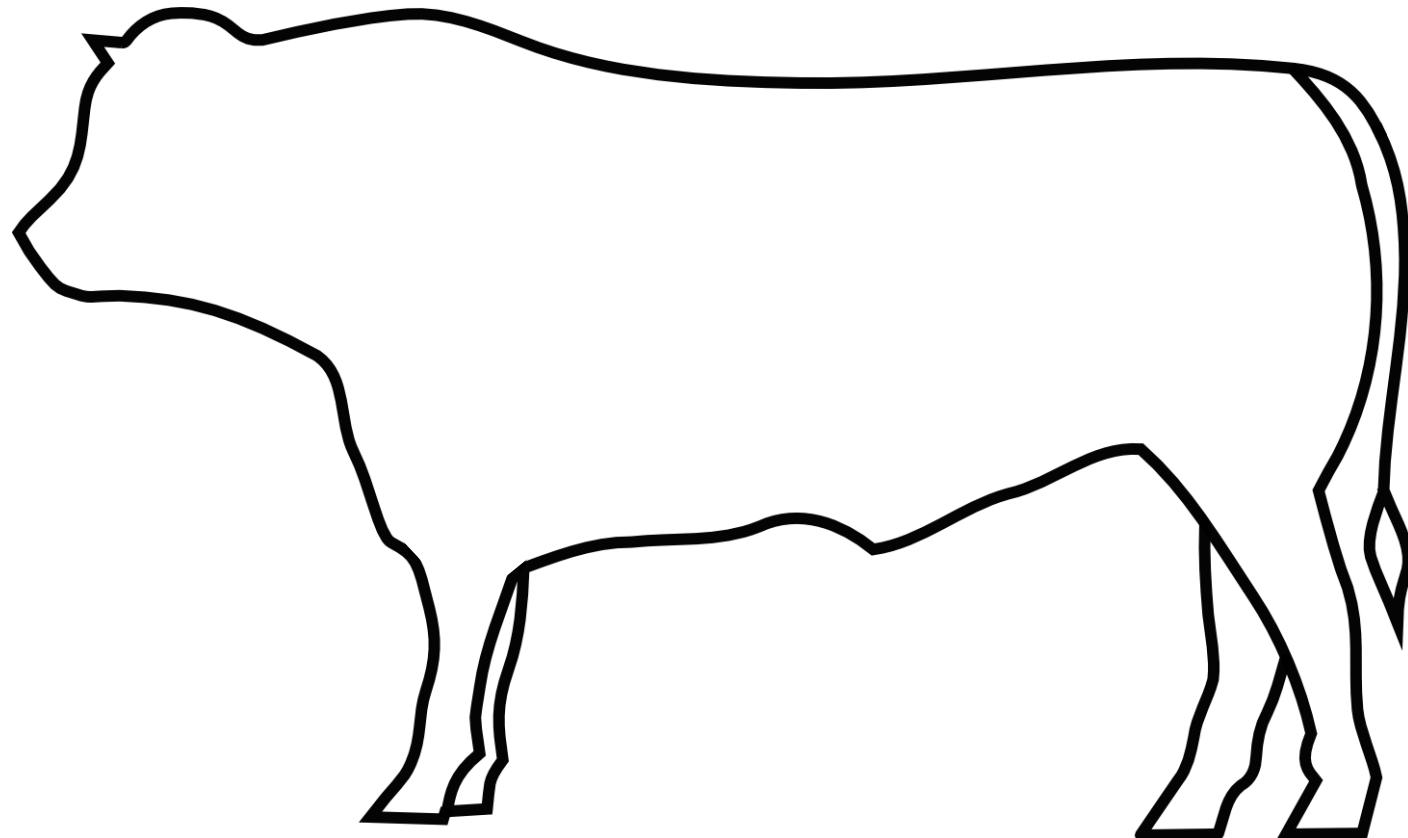
0000 add [eax], al

Windows 7 x64

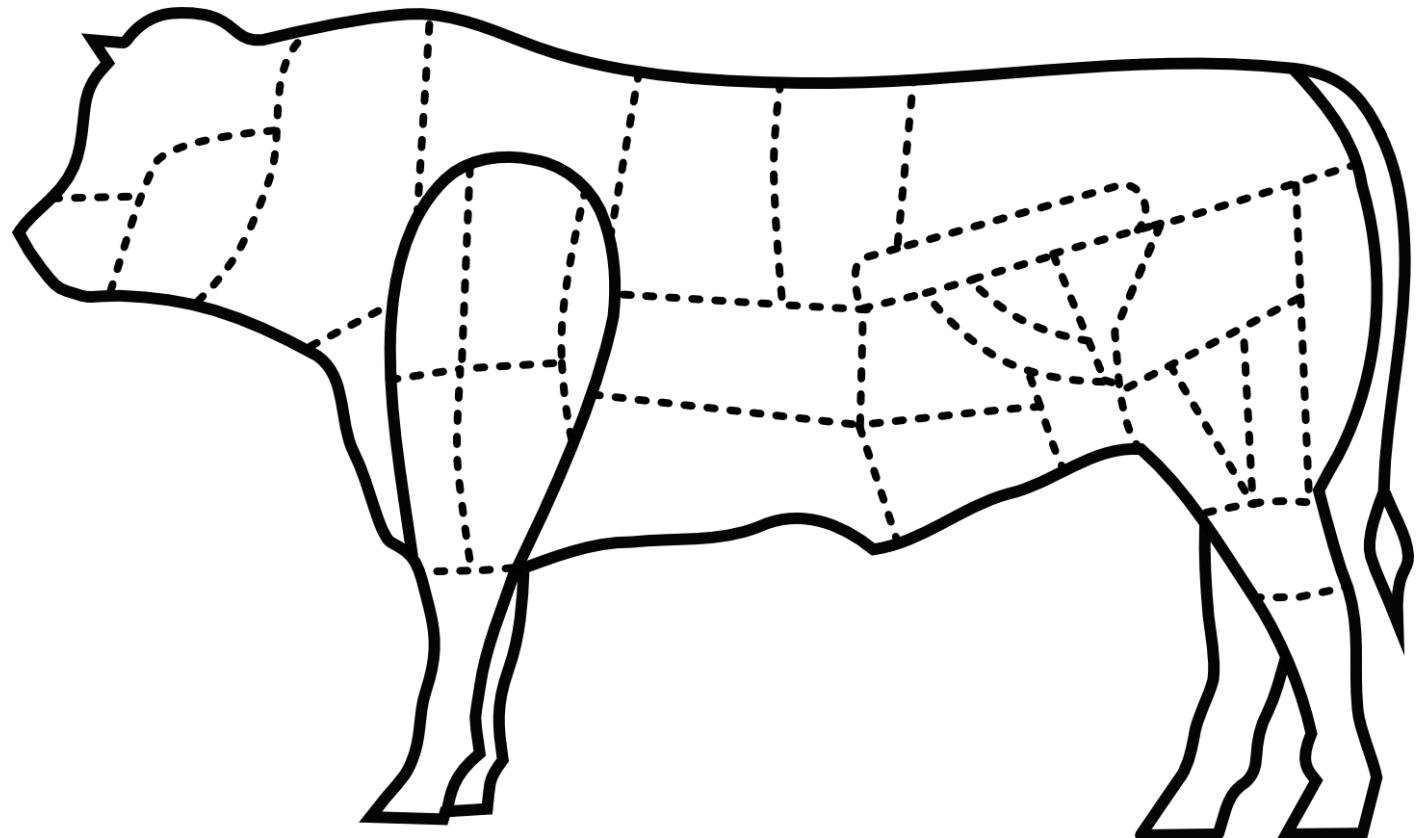
>65535sects.exe
* 65535 physically identical, virtually executed sections

a 64K sections PE (all executed) ⇒ crashes many softwares, evades scanning

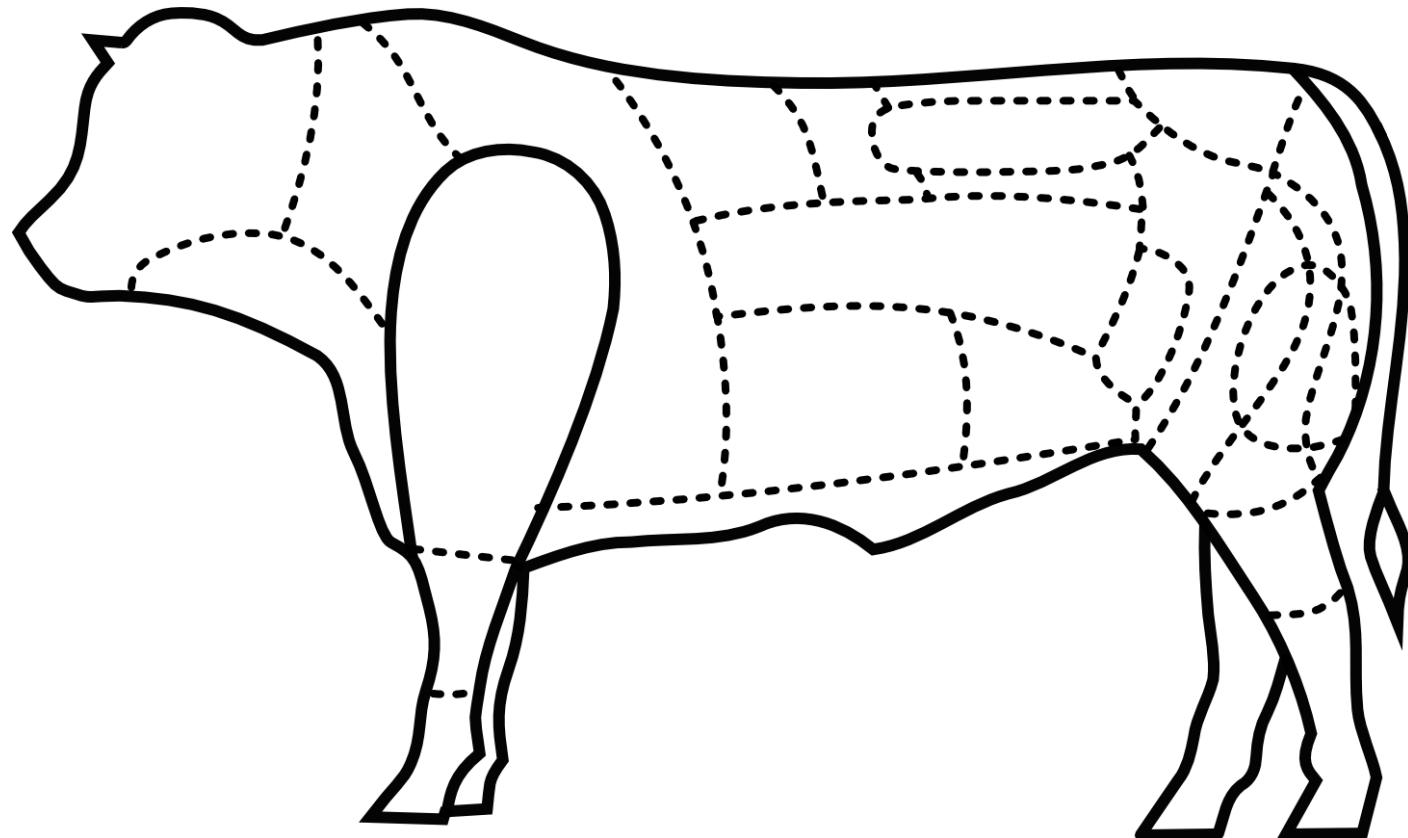
Parsing



This is how a **user** sees a cow.



This is how a **dev** sees a cow...



This is how **another** dev sees a cow !
(this one: brazilian beef cut - previous: french beef cut)

Same data, different parsers

it would have been too easy ;)

commented line

% trailer <</Root ...>>



trailer <</Root ...>>

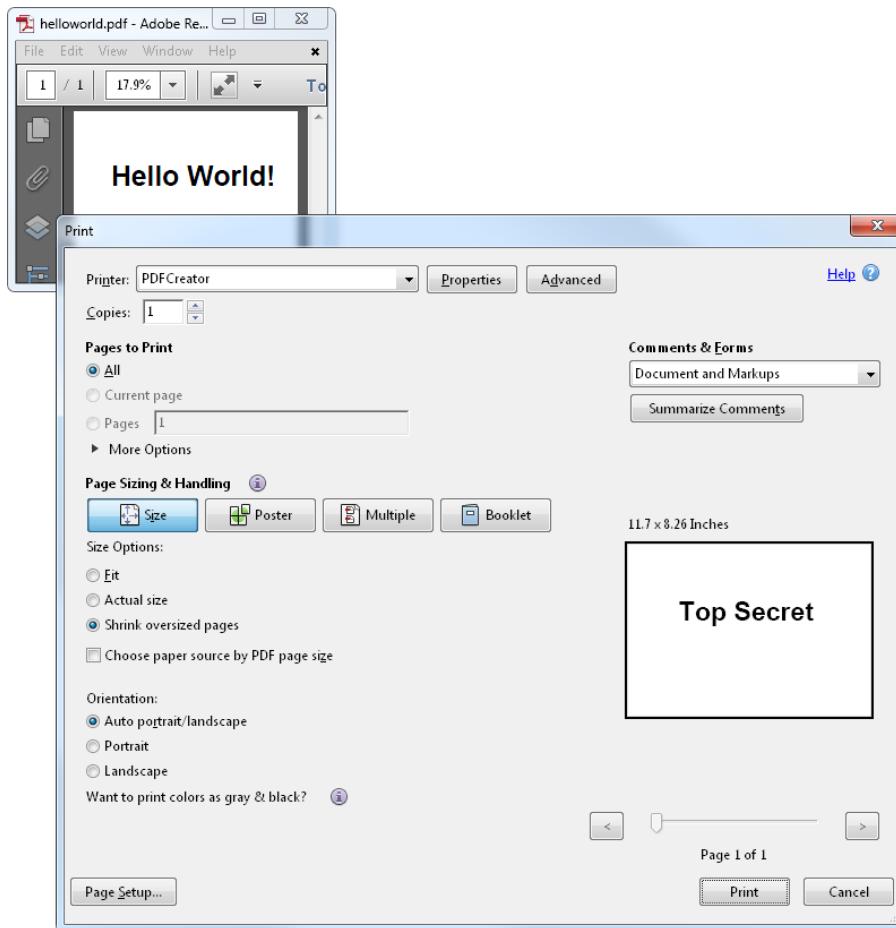


missing trailer keyword

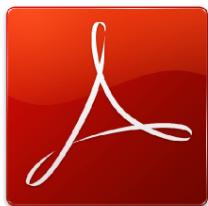
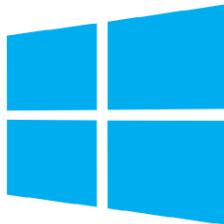
<</Root ...>>



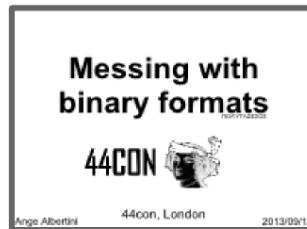
a schizophrenic PDF: 3 different trailers, seen by 3 different readers



a schizophrenic PDF (screen ⇄ printer)



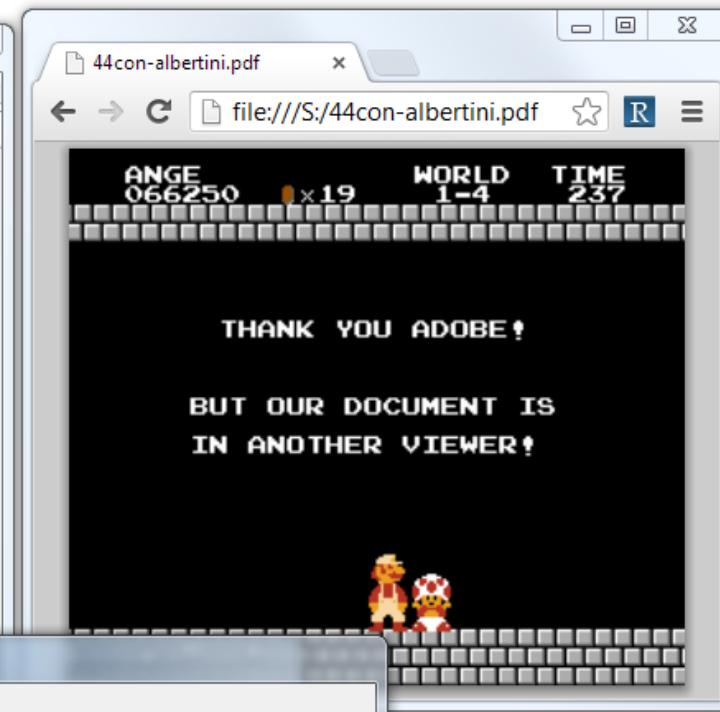
HTML



PDF viewer

PDF slides

a (generated) PDF || PE || JAR [JAVA+ZIP] || HTML polyglot...



...which is also a schizophrenic PDF

```
$ du -h stringme  
141      stringme
```

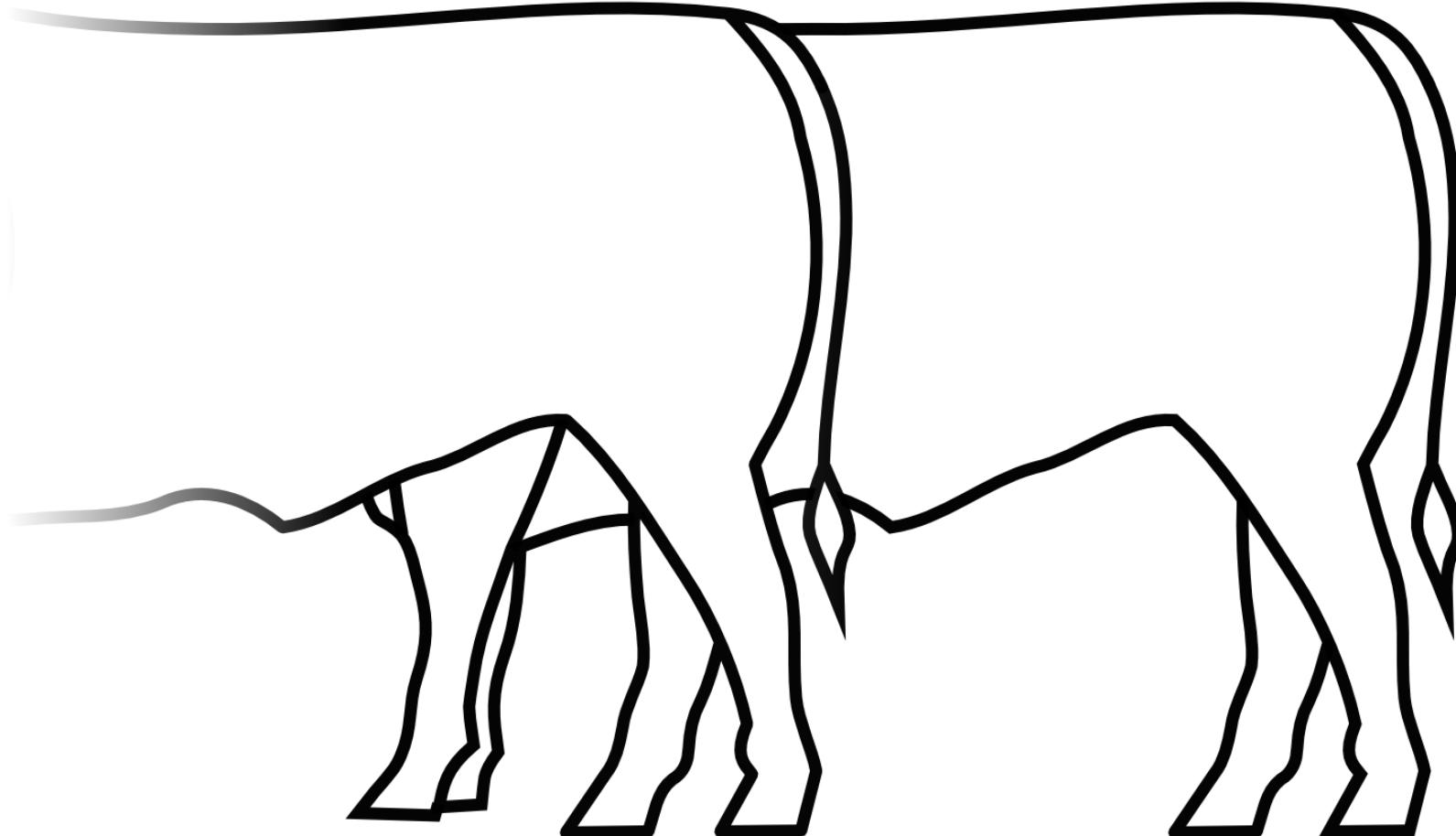
```
$ strings stringme  
Segmentation fault (core dumped)
```

Extra problem: parsers can be present in unexpected places

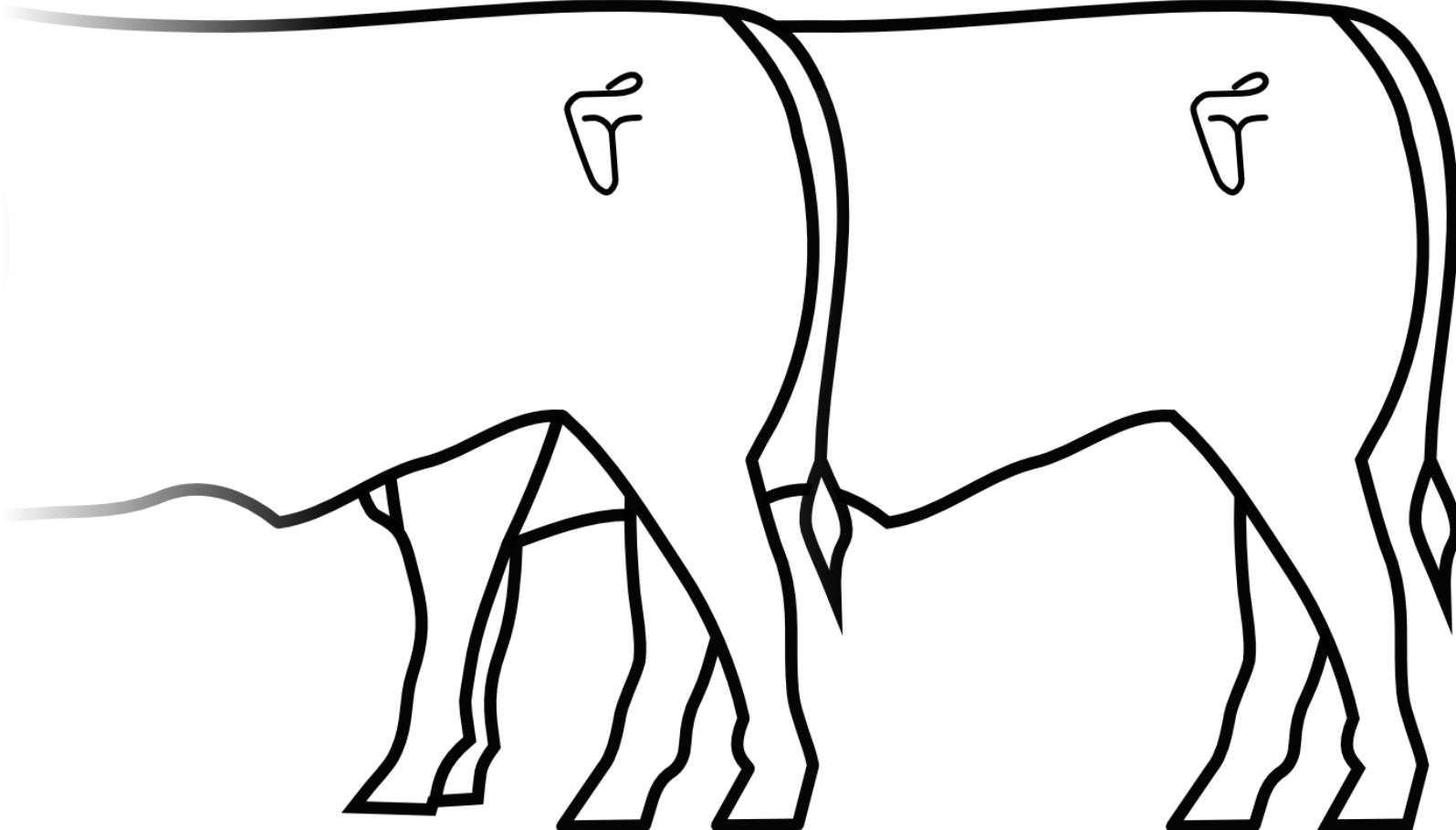
<http://lcamtuf.blogspot.de/2014/10/psa-dont-run-strings-on-untrusted-files.html> (CVE-2014-8485)

metadata

Who's the owner?



A hidden cow just looks like another cow...



... so cattle is branded.

But brandings can be faked!

or “patched” into another symbol
⇒ attribution is hard

... and in a pure PoC||GTFO fashion,
@munin forged a branding iron !



an *encrypted* file is not always “*encrypted*”
⇒ `encrypt(file)` is not always “random”

`encrypt(file)` can be *valid*

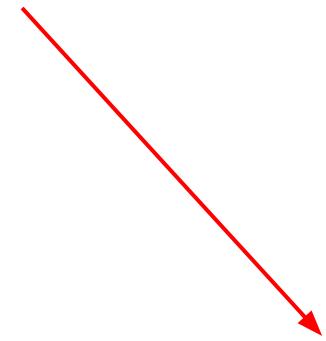
.D.A.T.A.[.1.2.3.4.5.6.7.8.9.A.B
.C.D.E.F.] .E.N.D

?

.T.E.X.T0A.t.h.i.s. .i.s. .a. .t
.e.x.t0A

We want to encrypt a **DATA** file to a **TEXT** file.
DATA tolerates appended data after it's **END** marker
TEXT accepts /* */ comments chunk (think 'parasite in a host')

.D.A.T.A.[.1.2.3.4.5.6.7.8.9.A.B
.C.D.E.F.].E.N.D



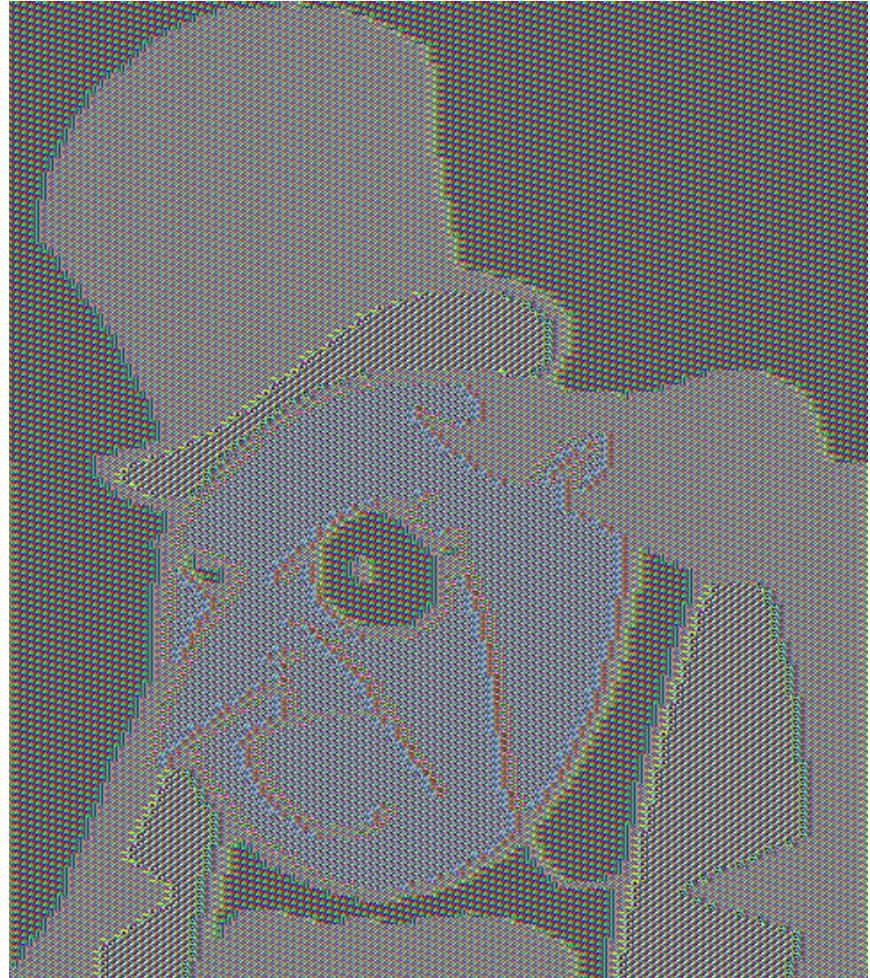
<random>

if we encrypt, we get random result. we can't control AES output & input together.

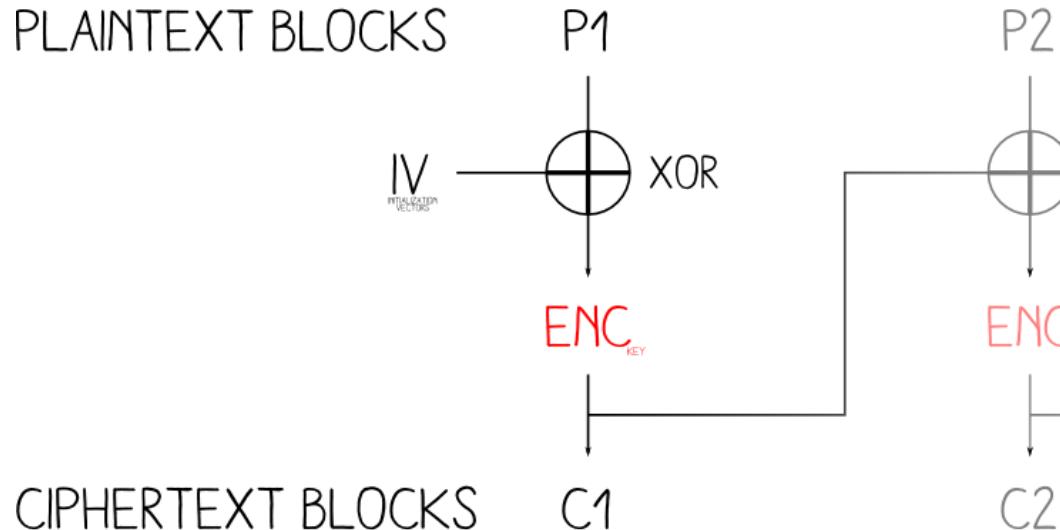
AES works with blocks

File encryption applies AES via a mode of operation

Electronic Code Book:
penguin = bad



CIPHER BLOCK CHAINING



choose the IV to control
both first blocks (P1 & C1)

$$C1 = \text{ENC}_{\text{KEY}}(P1 \wedge IV)$$

$$\text{DEC}_{\text{KEY}}(C1) = P1 \wedge IV$$

$$IV = \text{DEC}_{\text{KEY}}(C1) \wedge P1$$

.D.A.T.A.[.1.2.3.4.5.6.7.8.9.A.B
.C.D.E.F.].E.N.D

+IV1

.T.E.X.T <something we control>
<random rest>

Encrypt with pure AES, then determine IV to control the output block

```
.D.A.T.A.[ .1.2.3.4.5.6.7.8.9.A.B  
.C.D.E.F. ].E.N.D
```

+IV2

```
.T.E.X.T./.*  
<ignored random rest>
```

We can't control the rest of the garbage... so let's put a comment start in the first block

```
.D.A.T.A.[ .1.2.3.4.5.6.7.8.9.A.B  
.C.D.E.F. ].E.N.D
```

```
.T.E.X.T./.*  
<ignored random rest>  
.*./0A.t.h.i.s. .i.s. .a. .t  
.e.x.t0A
```

If we close the comment and append the target file's data in the encrypted file.
then this file is valid and equivalent to our initial target.

```
.D.A.T.A.[ .1.2.3.4.5.6.7.8.9.A.B  
.C.D.E.F. ].E.N.D  
<pre-decrypted ignored random>
```

+IV2

```
.T.E.X.T./.*  
<ignored random rest>  
.*./0A.t.h.i.s. .i.s. .a. .t  
.e.x.t0A
```

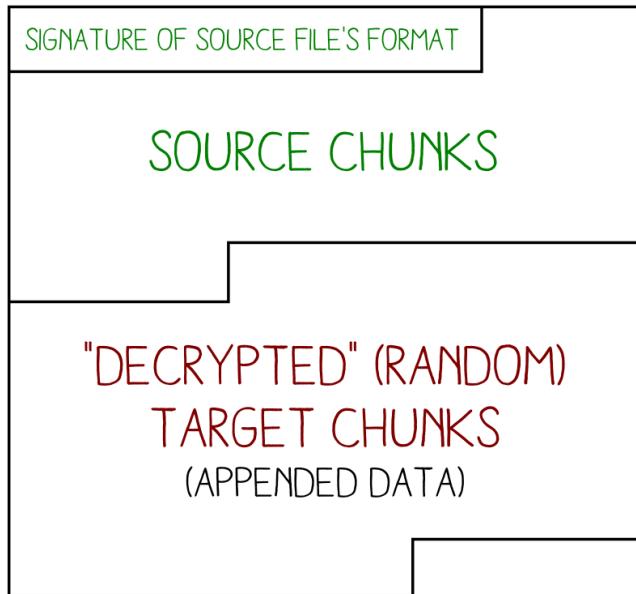
...then we decrypt that file: we get the original source file,
with some random data, that will be ignored since it's appended data.

.D.A.T.A.[.1.2.3.4.5.6.7.8.9.A.B
.C.D.E.F.].E.N.D
<pre-decrypted ignored random>

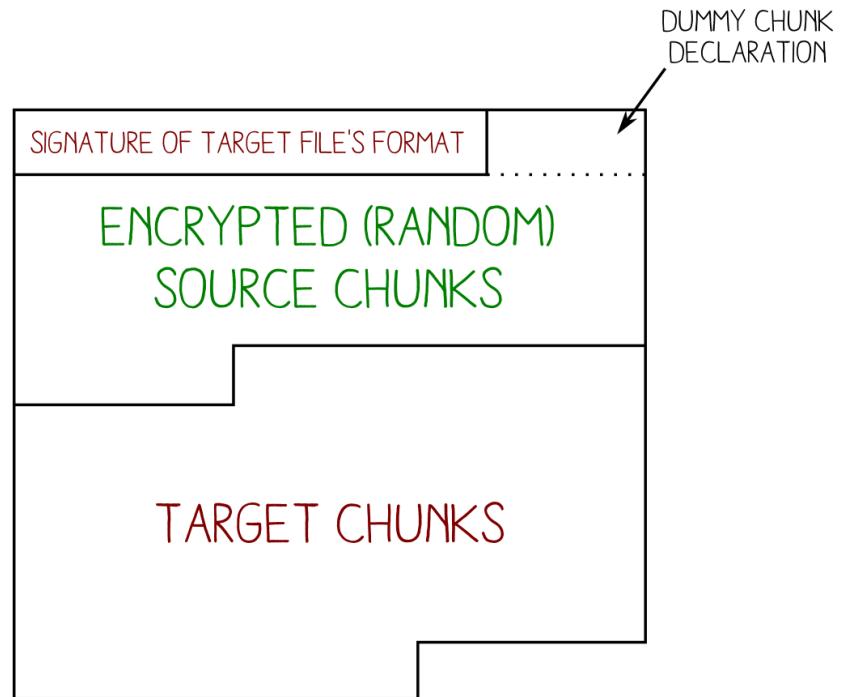
+IV2

.T.E.X.T./.*
<ignored random rest>
.*./0A.t.h.i.s. .i.s. .a. .t
.e.x.t0A

Since AES CBC only depends on **previous** blocks,
this DATA file will indeed encrypt to a TEXT file.



BEFORE ENCRYPTION



AFTER ENCRYPTION

```
00: 4441 5441 5b31 3233 3435 3637 3839 4142  DATA[123456789AB
10: 4344 4546 5d45 4e44 0000 0000 0000 0000  CDEF]END.....
20: f6fe 17cf 0802 7449 58de cdf2 f9c4 45ce  .....tIX.....E.
30: 2e8e 6996 5854 824c c09c 1b7d 4898 a29e  ..i.XT.L...}H...
```

```
openssl enc -aes-128-cbc -nopad
-K `echo OurEncryptionKey|xxd -p`
-iv A37A69F13417F5AB3CC4A1546B97FD76
```

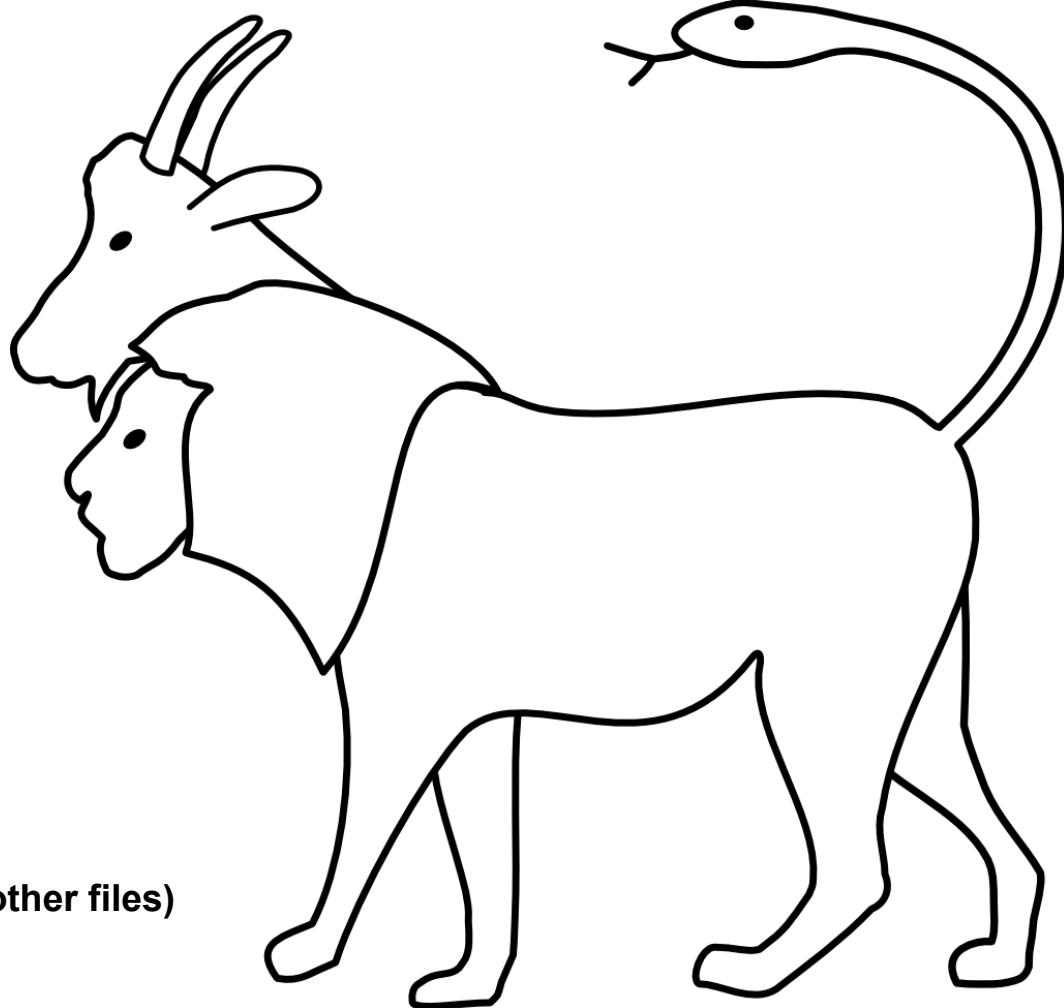


```
00: 5445 5854 2f2a 0000 0000 0000 0000 0000  TEXT/*.....
10: 3f81 11a9 2540 ded5 096a 83c9 f191 d8bb  ?...%@...j.....
20: 2a2f 0a74 6869 7320 6973 2061 2074 6578  */.this is a tex
30: 740a 454e 4400 0000 0000 0000 0000 0000  t.END.....
```

You can even try it at home :)

Chimera

(if you skip identified bodies, you'll miss other files)



zipjpg.pdf - WinRAR

File Commands Tools Favorites Options Help

zipjpg.pdf - SFX ZIP archive, unpacked size 69,782 bytes

Name	Size	Pac...	Type	CRC32
..			Folder	
corkami.jpg	69,782	69,782	JPEG Image	2A142635

Test finished

No errors found during test operation

OK

Selected 69,782 bytes in 1 file Total 69,782 bytes in 1 file

endstream
endobj

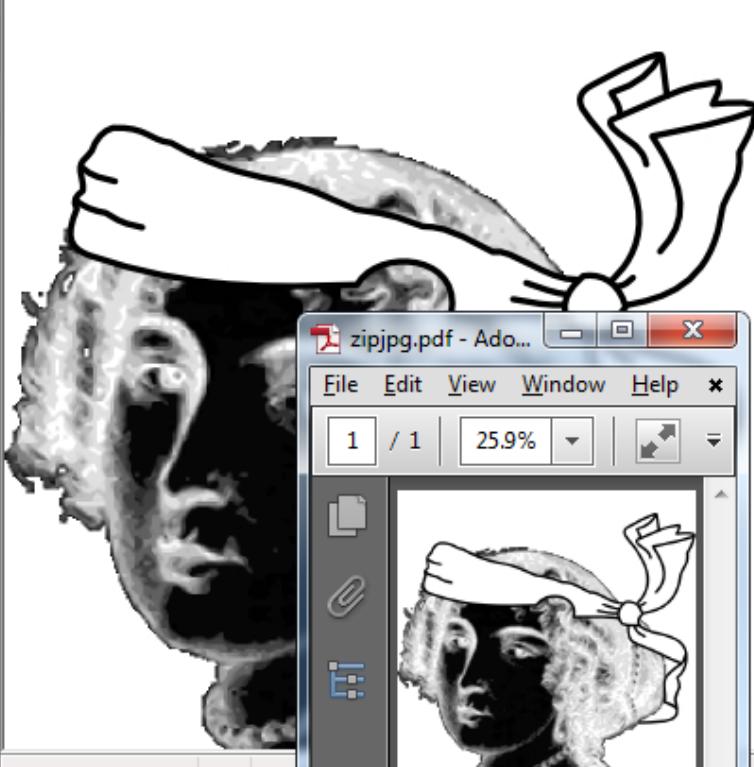
xref
0 1
000000000000 65535 f
000000000010 00000 n

trailer
<</Root 1 0 R>>

startxref
70488
%%EOF
ÿÙ

zipjpg.pdf - IrfanView

File Edit Image Options View Help



zipjpg.pdf - Ado...

File Edit View Window Help

1 / 1 25.9%

400 x 400 x 24 BPP 1/1 100

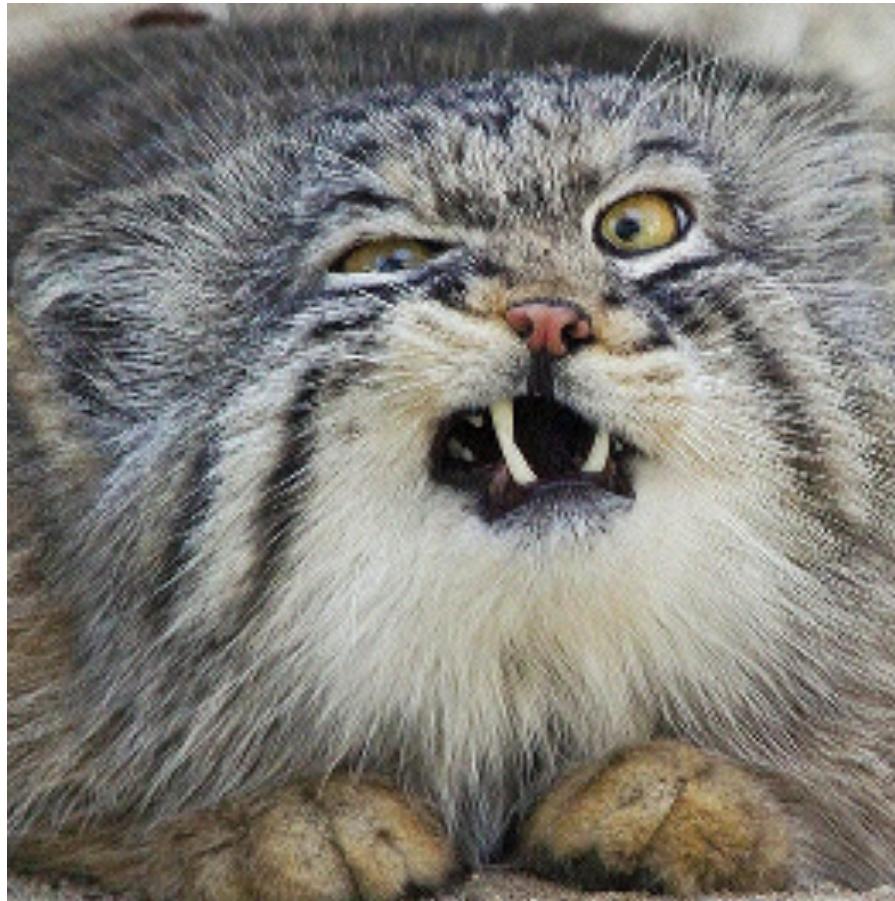
a JPEG || ZIP || PDF Chimera

OFFSETS	CONTENT	JPEG	PDF	ZIP
00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F				
00000: FF D8 00 E0 00 10 .J .F .I .F 00 01 01 01 00 48 00 48 00 00	MAGIC & HEADER			
14: FF FE 02 1E	COMMENT SEGMENT START (LENGTH)			
18: %PDF-1.4		PDF HEADER & DOCUMENT		
1 0 obj ...				
140: 20 0 obj <</Length 69786>> stream		DUMMY OBJECT START		
168: .P .K 03 04			LOCAL FILE HEADER START	
181: 00 9b			FILE NAME LENGTH	
186: endstream endobj		DUMMY OBJECT END	LFHS FILENAME (ABUSED)	
5 0 obj <<Width 400 ... stream			IMAGE OBJECT START	
221: FF D8 00 E0 00 10 .J .F .I .F 00 01 01 01 00 48 00 48 00 00	(END OF COMMENT)	IMAGE HEADER	STORED FILE DATA	
235: FF DB 00 43 ...	IMAGE DATA (IDQT)	--	--	
112B5: FF D9	END OF IMAGE	--	--	
112B7: FF FE 00 E6	SEGMENT COMMENT START (NOT STRICTLY REQ.)			
112bc: endstream endobj		END OF IMAGE OBJECT		
24 0 obj stream ...		DUMMY OBJECT START		
112de: 01 02	.P .K		CENTRAL DIRECTORY	
1130c: corkami.jpg			FILENAME (CORRECT)	
11317: .P .K 05 06			END OF CENTRAL DIR	
1132b: 75 00			LENGTH OF COMMENT	
1132e: endstream endobj		END OF DUMMY OBJECT	ARCHIVE COMMENT	
xref ...		XREF, TRAILER		
1139a: %%EOF %		END OF FILE LINE COMMENT		
113a1: FF D9	END OF IMAGE MARKER	(END OF LINE)	(END OF COMMENT)	

image data

a chimera defeats sequential parsing with optimization

a ***Picture of Cat***
(BMP ! uncompressed ! OMG)



BMP let us define bit masks for each color:

32 bits: 0000000000000000rrrrrggggggbbbb (no alpha)

⇒ 16 bits of free space!

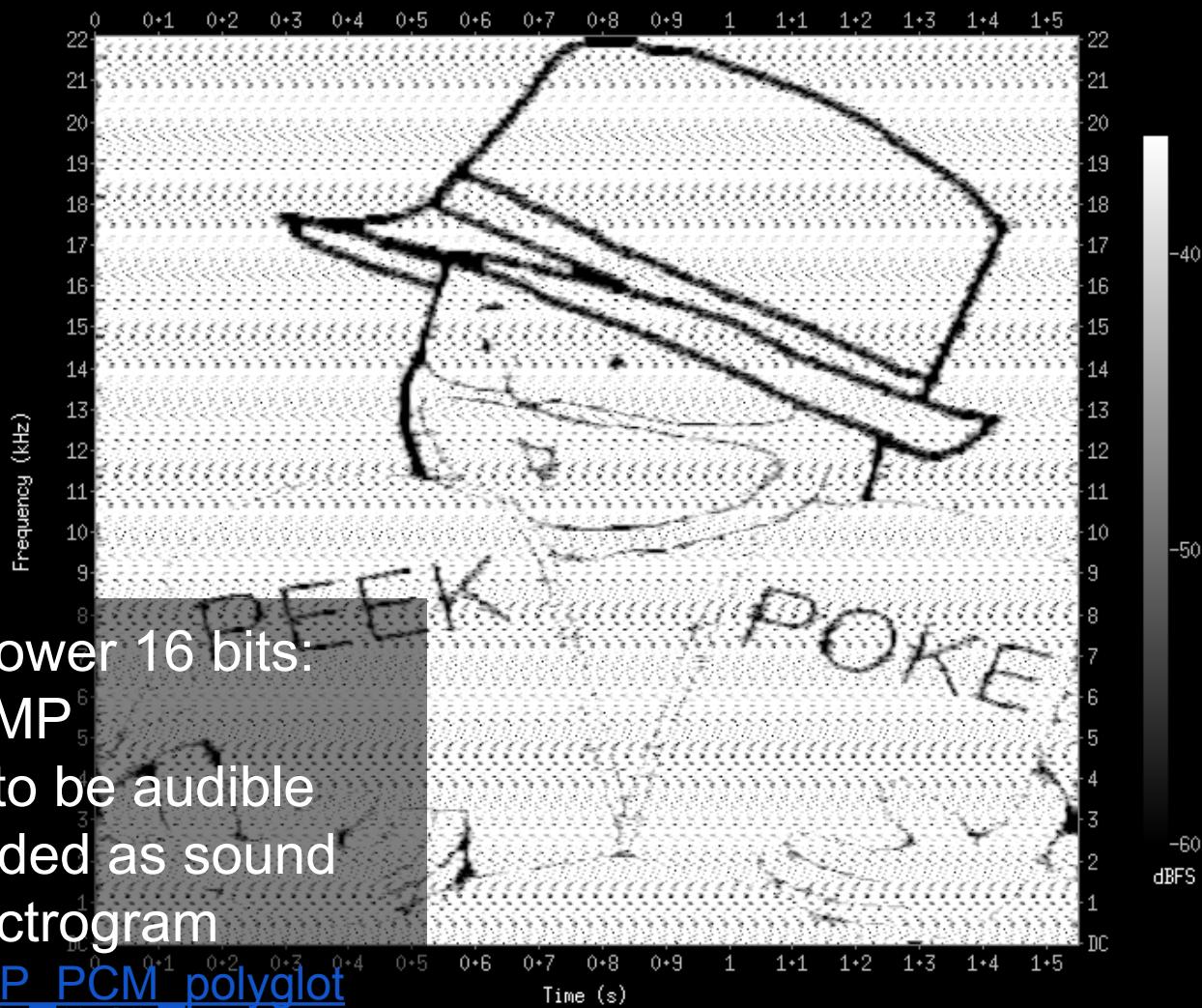
let's *play* the picture!

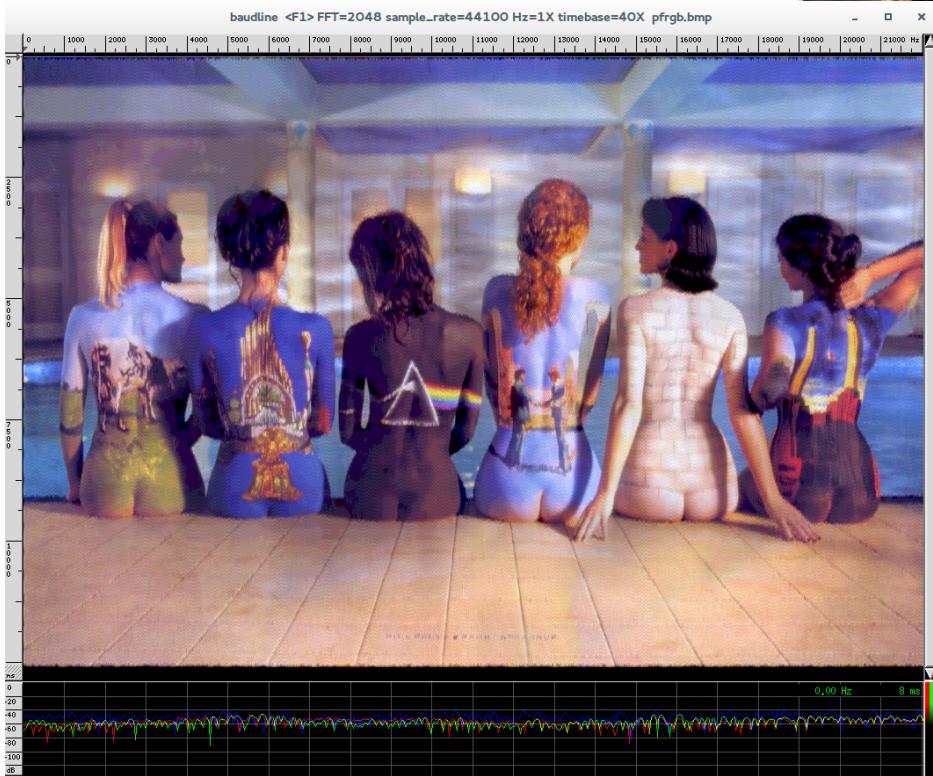
no, seriously :)

Consider the BMP
as RAW 32b PCM

1. store **sound** in the lower 16 bits:
sound ignored by BMP
image data too low to be audible
2. store a picture encoded as sound
 - o viewable as spectrogram

http://wiki.yobi.be/wiki/BMP_PCM_polyglot

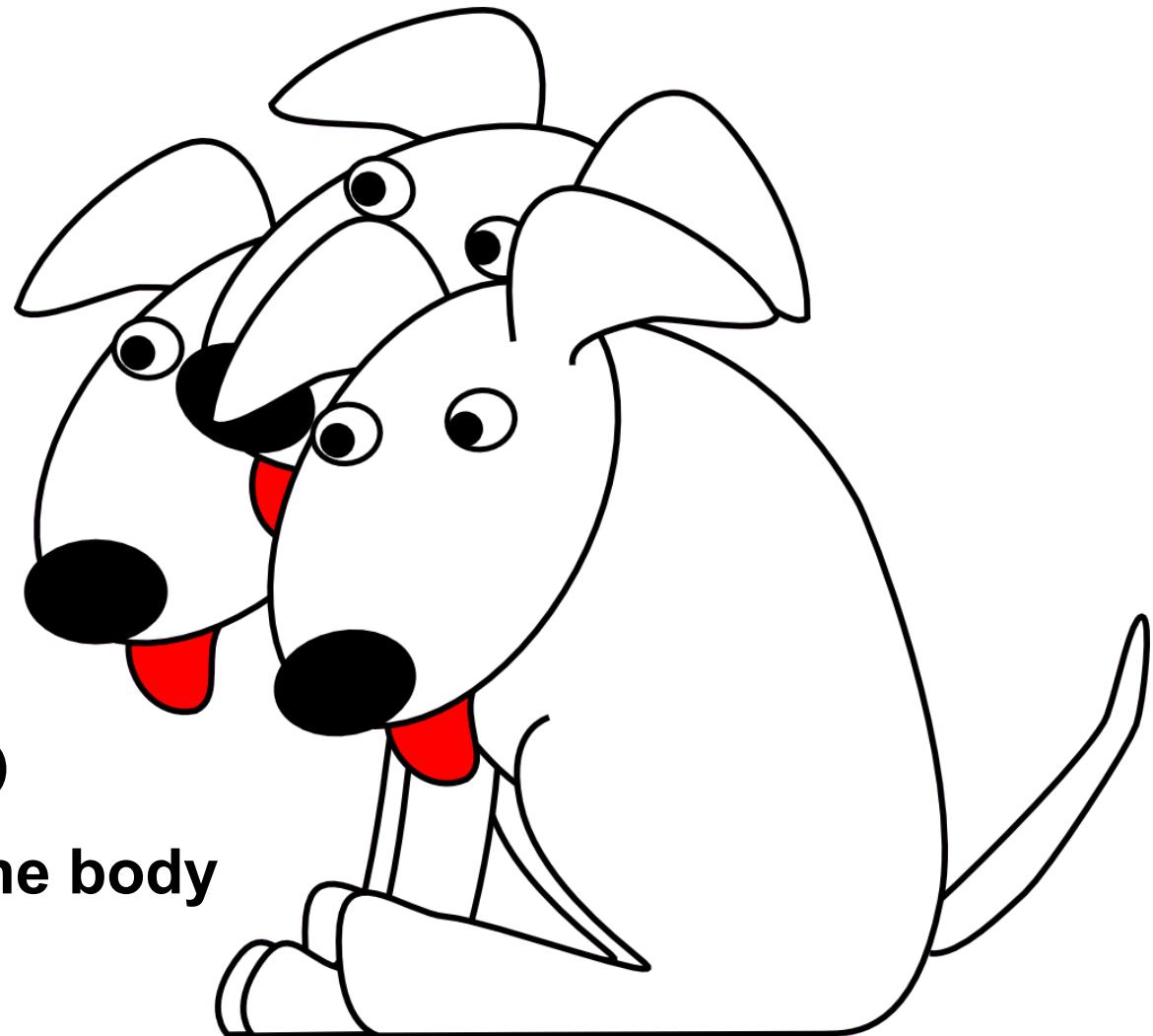




an RGB BMP || raw (3-channel spectrogram) polyglot by @doegox

Cerbero

same type of heads, one body



an RGB picture...

RGB picture data = bytes triplets for **R**, **G**, **B** colors

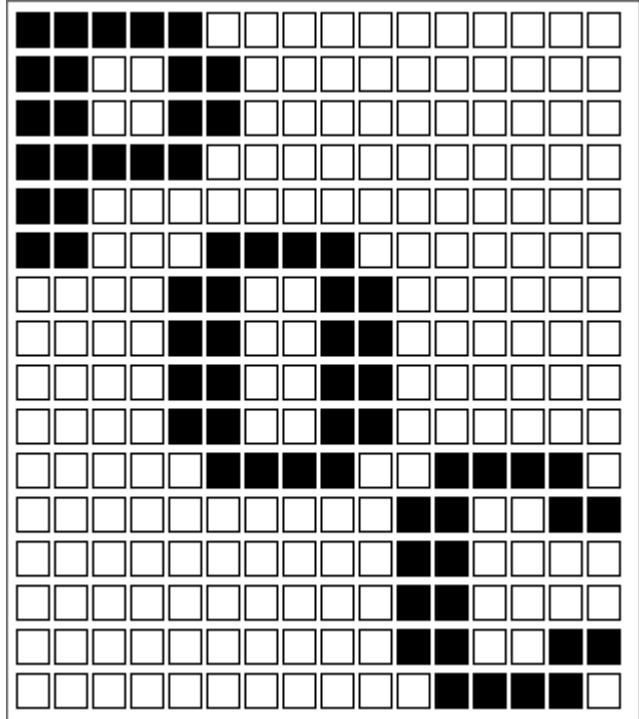


...with an unused palette

palette picture data = each byte is an index in the palette

in theory, it could be used:

For colour types 2 and 6 (truecolour and truecolour with alpha), the **PLTE** chunk is optional. If present, it provides a suggested set of colours (from 1 to 256) to which the truecolour image can be quantized if it cannot be displayed directly. It is, however, recommended that the **sPLT** chunk be used for this purpose, rather than the **PLTE** chunk.



How to make a pic-ception

adjust each RGB value to the closest palette index

⇒ store a ***second*** picture with the ***same*** data....

(original idea by @reversity)

We get another picture of
the same type from the
same data!

BTW, that's a barcode inception:
a DataMatrix barcode inside a QRCode, both valid
<https://www.iseclab.org/people/atrox/qrinception.pdf>



Malicious Hashing: Eve's Variant of SHA-1

Ange Albertini¹, Jean-Philippe Aumasson², Maria Eichlseder³,
Florian Mendel³, and Martin Schläffer³

This is the actual SHA-1 with only 4 of its 5 constants modified
This doesn't give a collision in the actual SHA-1

μΣ 9 ►m♦. . ú

, ç n 8ÿâø¥2%! j j ð

òC▼ Yfç sà≈ÿ T L R

σ 9 Tÿm♦. L ú

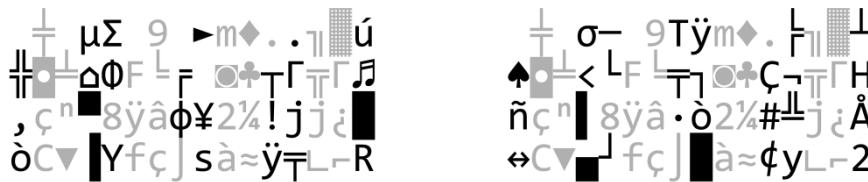
ñç n 8ÿâø. ò2%# j ð Å

↔C▼ fç J à≈çy L 2

2 colliding blocks: mostly random and unpredictable

At most three consecutive bytes without a difference.

Typically, in every dword, only the middle two bytes have no differences.



JPEG signature

Chunk marker

- ff e5 in block 1

- ff e6 in block 2

Chunk length

- c4 00 in block 1

- e4 00 in block 2

00000: ff d8 ff e? ?4 00 39 54 ?? 6d 04 2e ?? b7 b2 ??
?? 08 cf ?? ?? 46 d4 ?? ?? 0a 05 ?? ?? cb e2 ?? (contains no 0xff)
?? 87 fc ?? 38 98 83 ?? ?? 32 ac ?? ?? 6a a8 ??
?? 43 1f ?? ?? 66 87 f5 ?? 85 f7 ?? ?? 1c a9 ??

0c404: ff fe b5 e9 <COMment chunk covering Image 1>

0e404: ff e0 <start of Image 1>

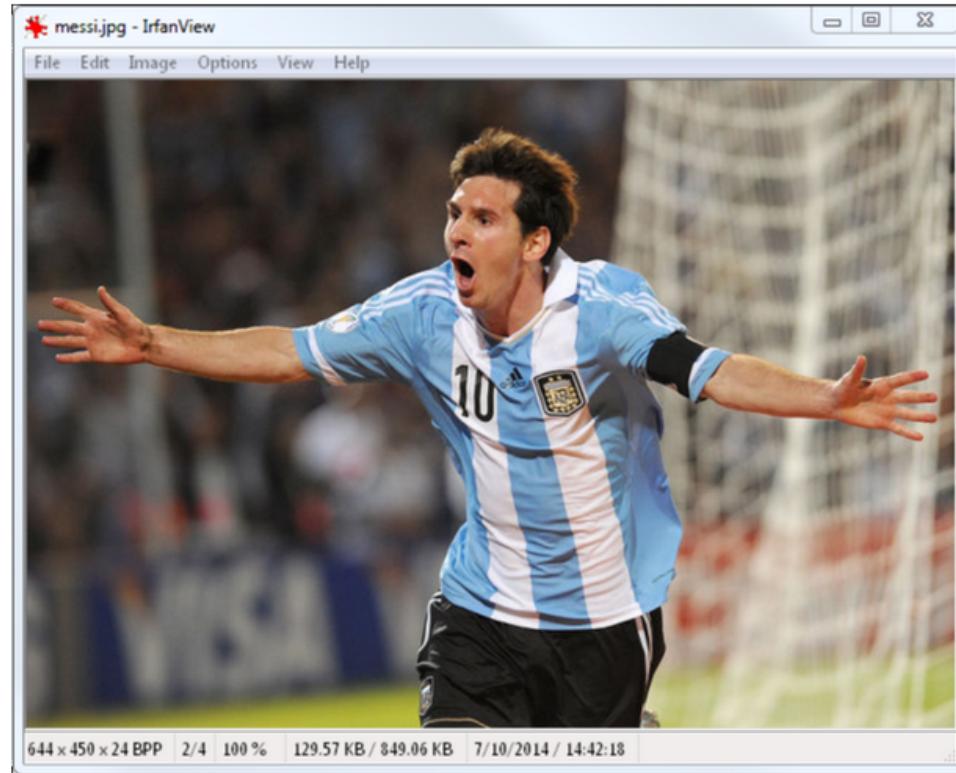
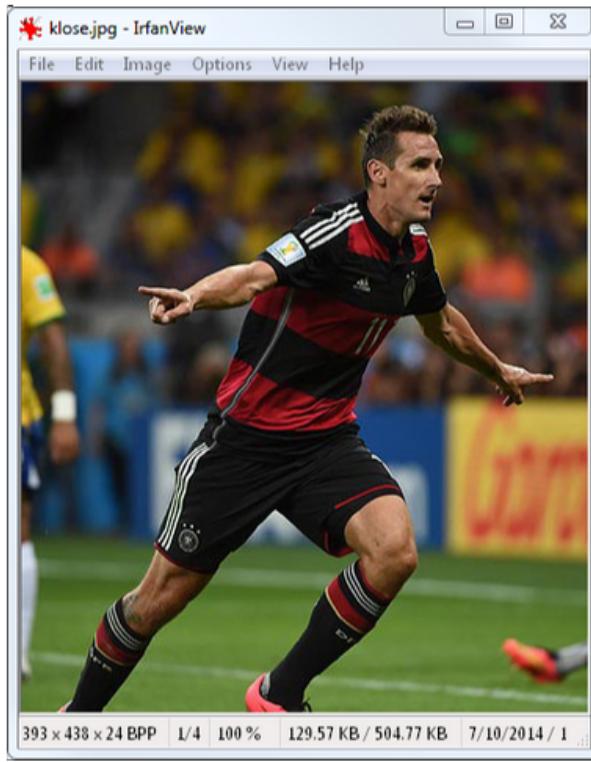
...
ff d9

<end of Image 1> <end of comment>

179ed: ff e0 <start of Image 2>

1b0d7: ff d9 <end of Image 2>

Abusing JPEG's multiple unused APPx (FF Ex) markers



```
>crypto_hash *.jpg
fb1847ac1342acb9c52c30f4b477997938a4a0a *klose.jpg
fb1847ac1342acb9c52c30f4b477997938a4a0a *messi.jpg
```

Much better! (images chosen at random)



shmbrar0.mbr



shmbrar0.sh



shmbrar0.rar

} identical



shmbrar1.mbr



shmbrar1.sh



shmbrar1.rar

} identical

{ collision }

{ collision }

{ collision }

a polyglot collision (multiple use for a single backdoor)

the PWNIE AWARDS



Pwnie award... for the best song! err... what is it pwning exactly ?

The SSL Smiley Song

Dashing through the cloud
On a ten gigabit link
One packet in a crowd
Falls into the data sink!
Draw a smiley face
On the diagram
Suck up data, leave no trace
It's all for Uncle Sam!

SSL terminators at the datacenter
Just gotta get on the other side
Just gotta break and enter!
No need to hack that server rack
Just gotta tap that fiber
Download all the private data
Win the war on CYBER!

Even songs should also have a nice PoC
(never forget to load your PDFs in your favorite NES emulator)



GENESIS DOES WHAT NINTENDON'T.

ARCADE GAMES:



Super Monaco GP™



Michael Jackson's Moonwalker™

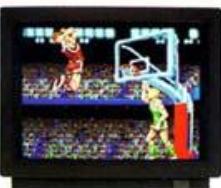


E-SWAT™

SPORTS GAMES:



Joe Montana Football™



Pat Riley Basketball™



James "Buster" Douglas
Knockout Boxing™

ADVENTURE GAMES:



The Sword of Vermilion™



Columns™



Dynamite Duke™

STRATEGY GAMES:

ACTION GAMES:

Get the hottest new video games going. Arcade, sports, adventure, strategy and action hits available only on the 16-bit Genesis System by Sega.®

Today's latest blockbuster arcade hits like Super Monaco GP.™ Climb into the cockpit of the world's fastest Grand Prix machines as you race wheel to wheel through the streets at over two-hundred miles per hour. Or take on the evil villain Mr. Big in Michael Jackson's Moonwalker™ as you use dance-kicks, hat-tricks and finally transform into a powerful robot that does it all. Or become a Cybercop in E-SWAT™ and clean up the city besieged by mad terrorists.

Get ready for the most action-packed sports games ever. In Joe Montana Football,™ check out the defense, make the call, fake a pass and scramble for a touchdown. Or force your opponent to move inside your left hook and nail him with an uppercut that puts him on the mat in James "Buster" Douglas Knockout Boxing.™ Or in Pat Riley Basketball,™ get the ball with seven seconds left in the game, drive the length of the court, slam-dunk and draw the foul which you make to break the tie.

In The Sword of Vermilion,™ make your way through 14 towns and 14 mazes in this adventure thriller where encounters with the evil demons are played in real time on the hand controller. And dazzle your friends with your skills on the puzzle game Columns.™ Or become the ultimate commando warrior in Dynamite Duke™ as you blast the enemy from an over-your-shoulder first person view.

There's only one true 16-bit system and it's got the hottest video game hits going. You can only play these on Genesis by Sega. Genesis does what Nintendo™





A Super NES & Megadrive rom
(and PDF at the same time)

snes_md.pdf - Adobe Reader

File Edit View Window Help

GENESIS DOES WHAT NINTENDON'T.

Get the hottest new video games going. Arcade, sports, adventure, strategy and action hits available only on the 16-bit Genesis System by Sega®. Today's latest blockbuster arcade hits like Super Monaco GP™ climb into the cockpit of the world's fastest Grand Prix machines as you race wheel to wheel through the streets at over two-hundred miles per hour. Or take on the evil villain Mr. Big in Michael Jackson's Moonwalker™ as you use dance-kicks, hat-tricks and finally transform into a powerful robot that does it all. Or become a Cybercop in E-SWAT™ and clean up the city besieged by mad terrorists. Get ready for the most action-packed sports games ever. In Joe Montana Football™ check out the defense, make the call, fake a pass and scramble for a

ARCADE GAMES:

Super Monaco GP™ Michael Jackson's Moonwalker™ E-SWAT™

SNES - snes md

File Emulation View Config Tools SNES Help

Genesis - snes md

Score 72 High Score 72

4 8 8
16
2

GET TO THE 2048 TILE!
TOUCHING EQUAL TILES MERGE INTO ONE!

CONTROLS: DPAD - MOVE GAME FIELD
C - RESTART GAME

Save slots 1 2 3 4 5 6 7 8 9 0 🔍 🔥

Save slots 1 2 3 4 5 6 7 8 9 0 🔍 🔥 Genplus-gx

The image shows a composite of several windows. At the top is a PDF viewer window titled 'snes_md.pdf - Adobe Reader' displaying an advertisement for the Sega Genesis system. Below it are two windows for SNES/Megadrive emulation. The left window, titled 'SNES - snes md', shows a pink screen with a small Mario character. The right window, titled 'Genesis - snes md', shows a 4x4 grid of numbers (4, 8, 8, 16, 2) from the game 2048, with text at the bottom about reaching the 2048 tile. Both emulators have their respective menu bars and toolbars visible.

Conclusion

Ange's recipes :)

Never forget to:

- open your PDFs in a hex editor
- open your pictures in a sound player
- run your documents in a console emulator
- encrypt/decrypt with any cipher
- double-check what you printed

Security advice:

DON'T *

It's easy to blame others - new insecure paths appear everyday

Research advice:

DO *

PoC||GTFO ! stop the marketing! cheap blamers \leftrightarrow blatant marketers?

F.F.F. conclusion

- many abuses of the specs
 - specs often are wrong or misleading
- few parsers, even fewer dissectors
- standard tools evolve the wrong way
 - try to repair ‘corrupted’ file outside the specs
 - standard and recovery mode

For technical details, check my previous talks.

ACK

@doegox @pdfkungfoo @veorq @reversity
@travisgoodspeed @sergeybratus qkumba
@internot @gynvael @munin
@solardiz @0xabadidea @ashutoshmehra
lytron @JacobTorrey @thicenl
...and anybody who gave me feedback!

Bonus

after the talk, we tried some PoCs on professional (very expensive!) forensic softwares:

- polyglot files
 - a single file format found + no warning whatsoever
- schizophrenic files:
 - no warning yet different tabs of the same software showing different content :D

BIG FAIL - yet we trust them for court cases ?

00000

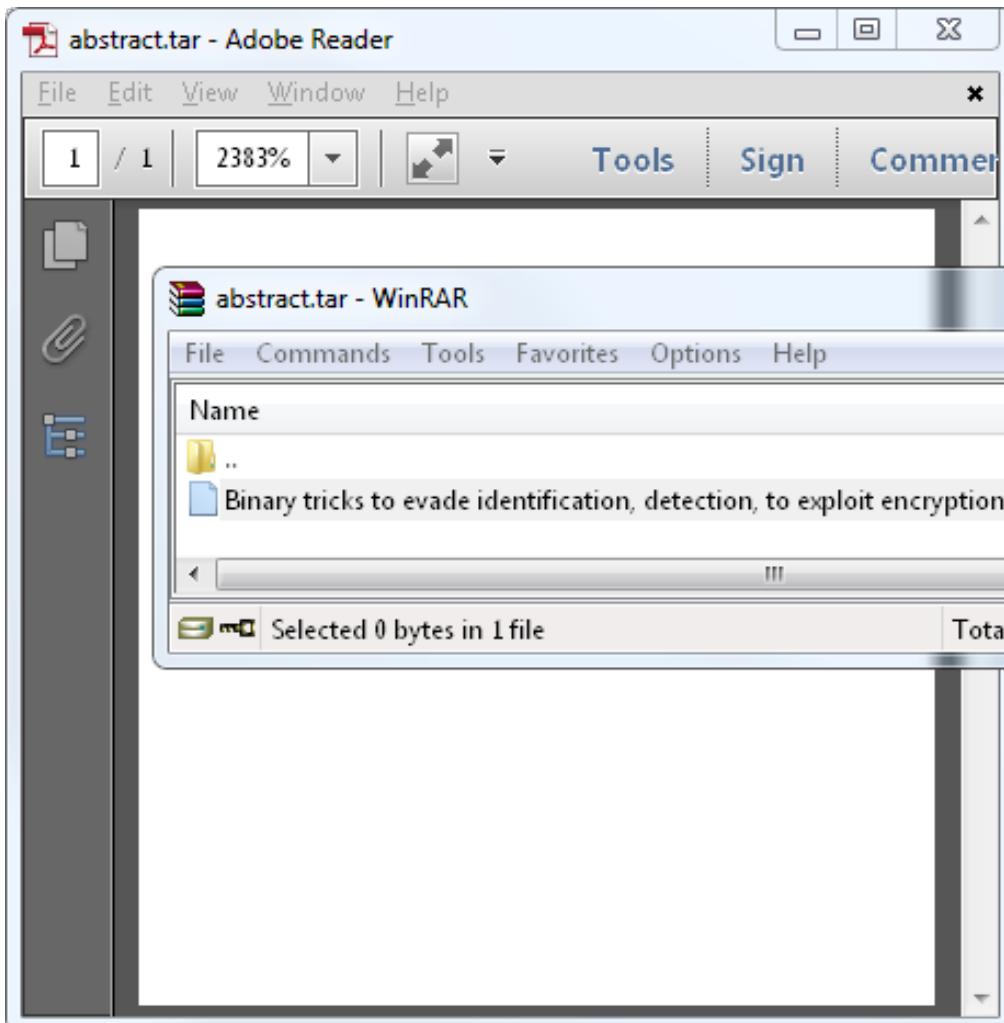
RYU

40

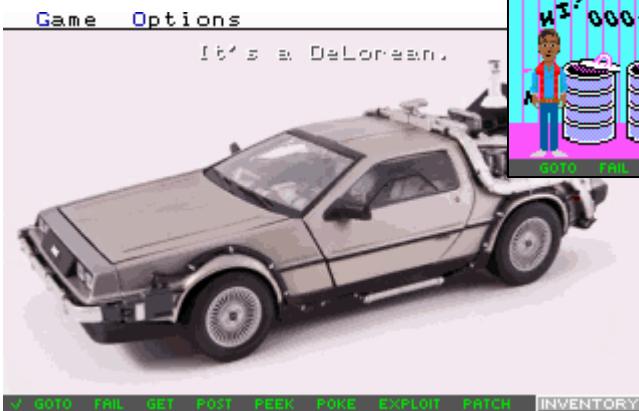
B O N U S

S T A G E





The initial abstract of this talk: ASCII-only, PDF/TAR polyglot



Solar Designer made a great keynote - that's actually a real game to play!
But one have to load and play through the game - not so accessible!
<http://openwall.com/presentations/ZeroNights2014-Is-Infosec-A-Game/>

A game by Solar Designer (@solardiz)
for ZeroNights 2014 (Moscow, Russia)
written in 1994-95 ("code"), 2014 ("data")
(includes pre-1994 library code and fonts)

<http://www.openwall.com/zn2014>

PDF/ZIP by Ange Albertini (@angealbertini)

a PDF:

- containing the game as ZIP
- hand-written
 - with walkthrough's screenshots
(in original resolution)
 - a lightweight title
 - while maintaining compatibility

a good way to distribute as a single file!

```
$ unzip -t ZeroNights2014-Is-Infosec-A-Game.pdf
Archive: ZeroNights2014-Is-Infosec-A-Game.pdf
warning [ZeroNights2014-Is-Infosec-A-Game.pdf]: 6381506 extra bytes
(attempting to process anyway)
      testing: ZN14GAME/                           OK
      testing: ZN14GAME/COMMON/                     OK
      ...
      testing: ZN14GAME/COMMON/TEST/                OK
      testing: ZN14GAME/COMMON/TEST/TEST/            OK
```

Quine

prints its own source

```
ZJPPPPPPPPPPPPPPPPPP  
PPPPPPPPPPPPPPPPPPPP  
PPPPPPPPPPPPPPPPPPPP  
PPPPPPPPPPPPPPPPXL  
;this is a type-  
able PE quine P  
E:E;a working P  
E file, made ent  
irely in assembl  
y, with no need
```

```
C:\WINDOWS\system32\cmd.exe  
;this is a type-able PE quine PE:  
;a working PE file, made entirely  
ded, which it displays on executio  
;you can do it manually via 'type'  
;Ange Albertini, BSD Licence, 2011  
IMAGEBASE equ 400000h
```

Windows XP

D>quine.exe

D>type quine.exe | more

X^L

```
;this is a type-able PE quine PE:  
;a working PE file, made entirely in assembly, with n  
ded, which it displays on execution, via 'tuning' its
```

a PE quine (in assembler, no linker)

Most quines aren't very sexy

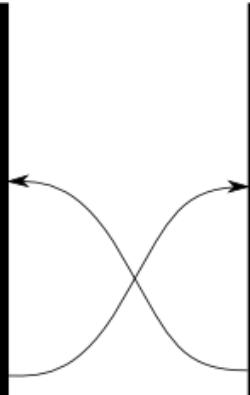
Using a compiler is cheap :p

Quine Relay

A prints B's source

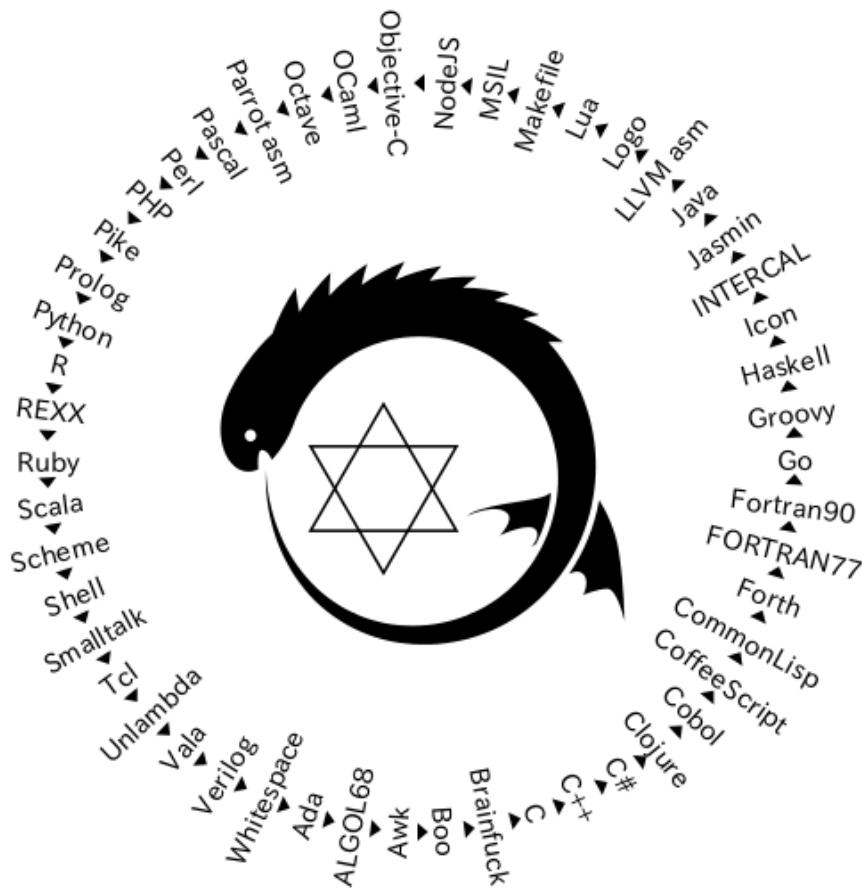
B prints A's source

```
>ver  
Microsoft Windows [Version 6.1.7601]  
  
>shasum relay.exe  
c46307a2faec73902bc70e0d7e89a2f412935eb9 *relay.exe  
>relay.exe > relay.asm  
>yasm -o relay relay.asm  
>shasum relay  
1f6594a24e593e32b490c83d4112c9ca7237a553 *relay
```



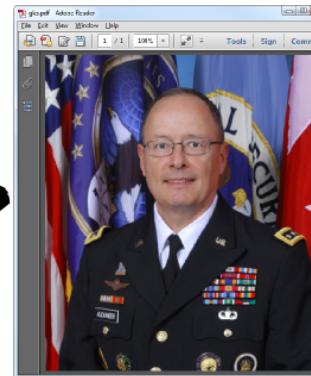
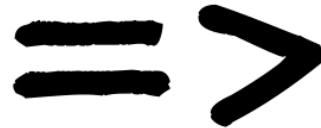
```
dev@nux:~$ uname  
Linux  
  
dev@nux:~$ sha1sum relay  
1f6594a24e593e32b490c83d4112c9ca7237a553 relay  
dev@nux:~$ ./relay > relay.asm  
dev@nux:~$ yasm -o relay.exe relay.asm  
dev@nux:~$ sha1sum relay.exe  
c46307a2faec73902bc70e0d7e89a2f412935eb9 relay.exe
```

a PE \Leftrightarrow ELF quine relay
(no linker)

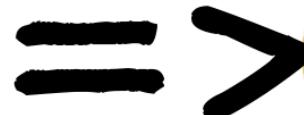


a 50-languages quine relay

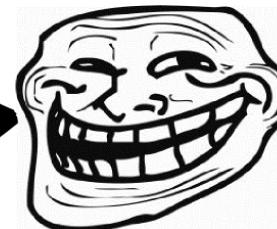
<https://github.com/mame/quine-relay>



RSA
SECURITY



Adobe®



other AngeCryption PoCs (PDF, PNG, JPG)

Schizophrenics

different contents with different tools



Fraternal twins

hash collision

```
> m_sha1sum.exe *
10382a6d3c949408d7cafaaf6d110a9e23230416 *0
10382a6d3c949408d7cafaaf6d110a9e23230416 *1
```



0



1



Booting from Floppy... MBR
good!

Booting from Floppy...
evil!

.0.sh
good.

shell
script

.1.sh
evil.

Polyglots

multiple file formats

A bit of everything

@angealbertini

corkami.com

Funky
File
Formats

