

12th CENTRAL & EASTERN EUROPEAN
SOFTWARE ENGINEERING CONFERENCE IN RUSSIA


October 28 - 29, Moscow



Anonymity of Tor: myth and reality

Aleksandr Lazarenko

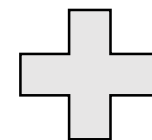
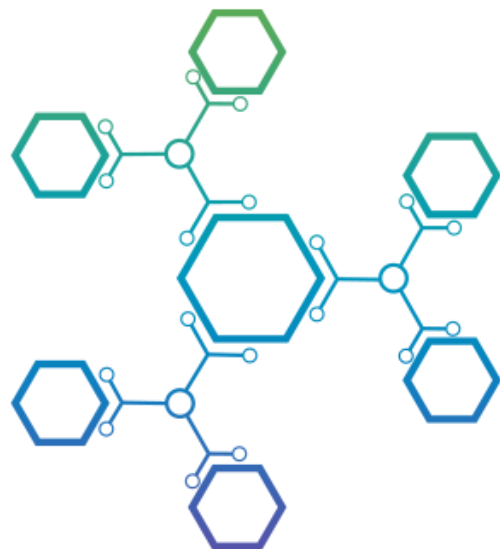
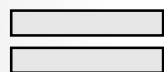
NRU HSE



What
is
Tor?

Anonymous network

Free software



The **O**nion **R**outer

Volunteer servers

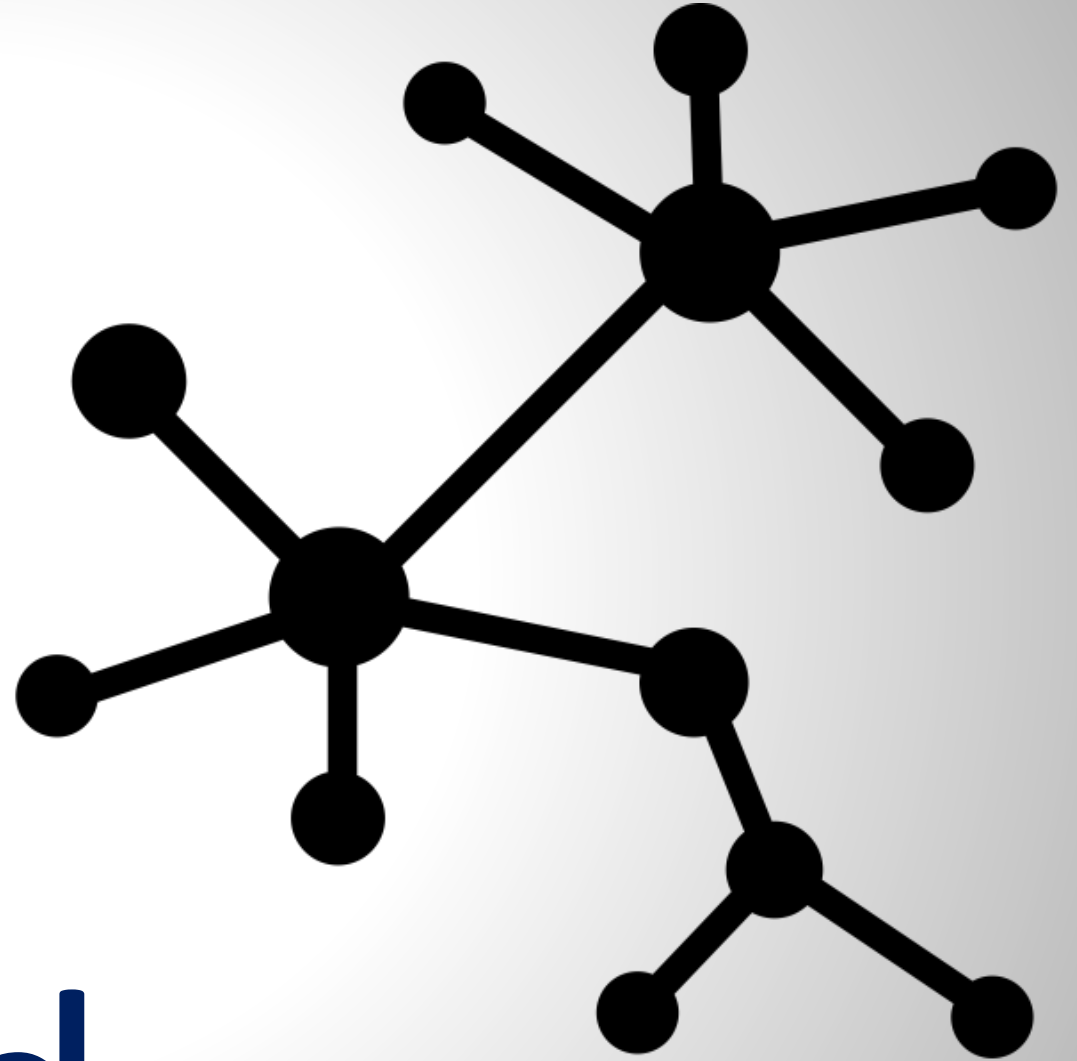
Browser
&
Messenger

Features



Tor is

distributed





Every server is VOLUNTEER

So what

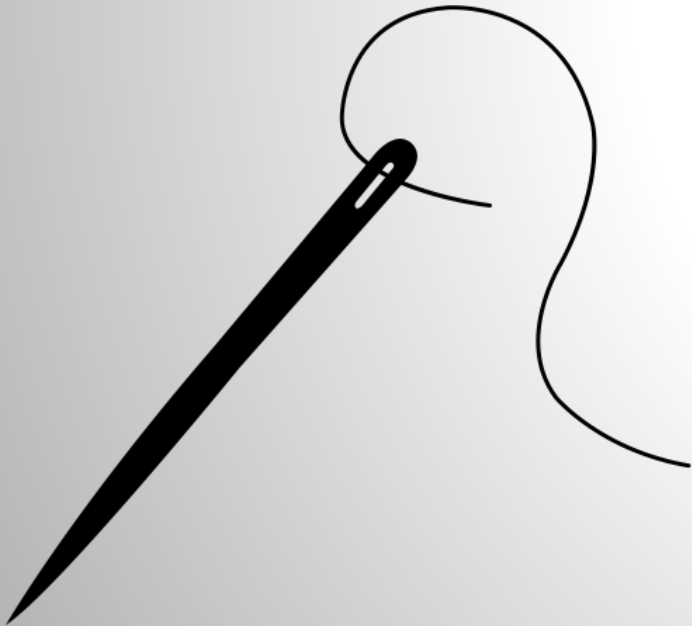
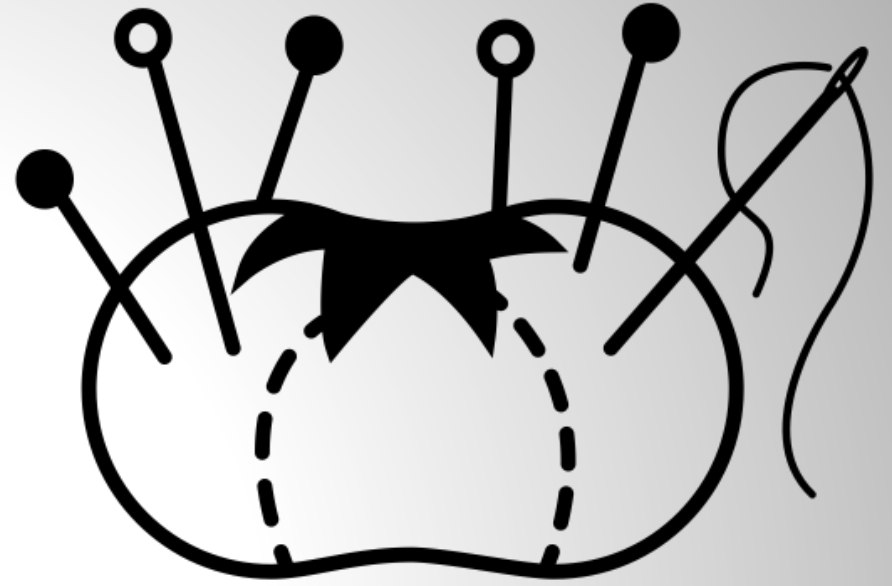
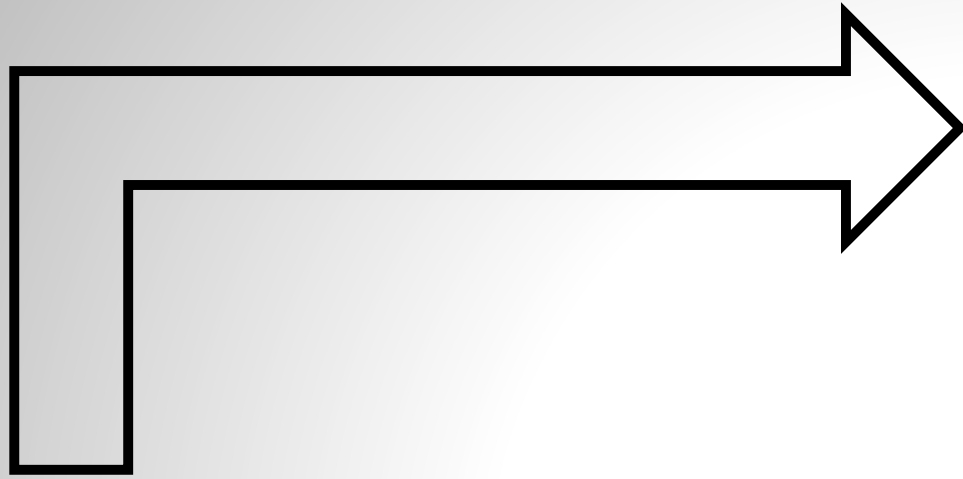


ANTI-FA
AREA
NPS
BXN



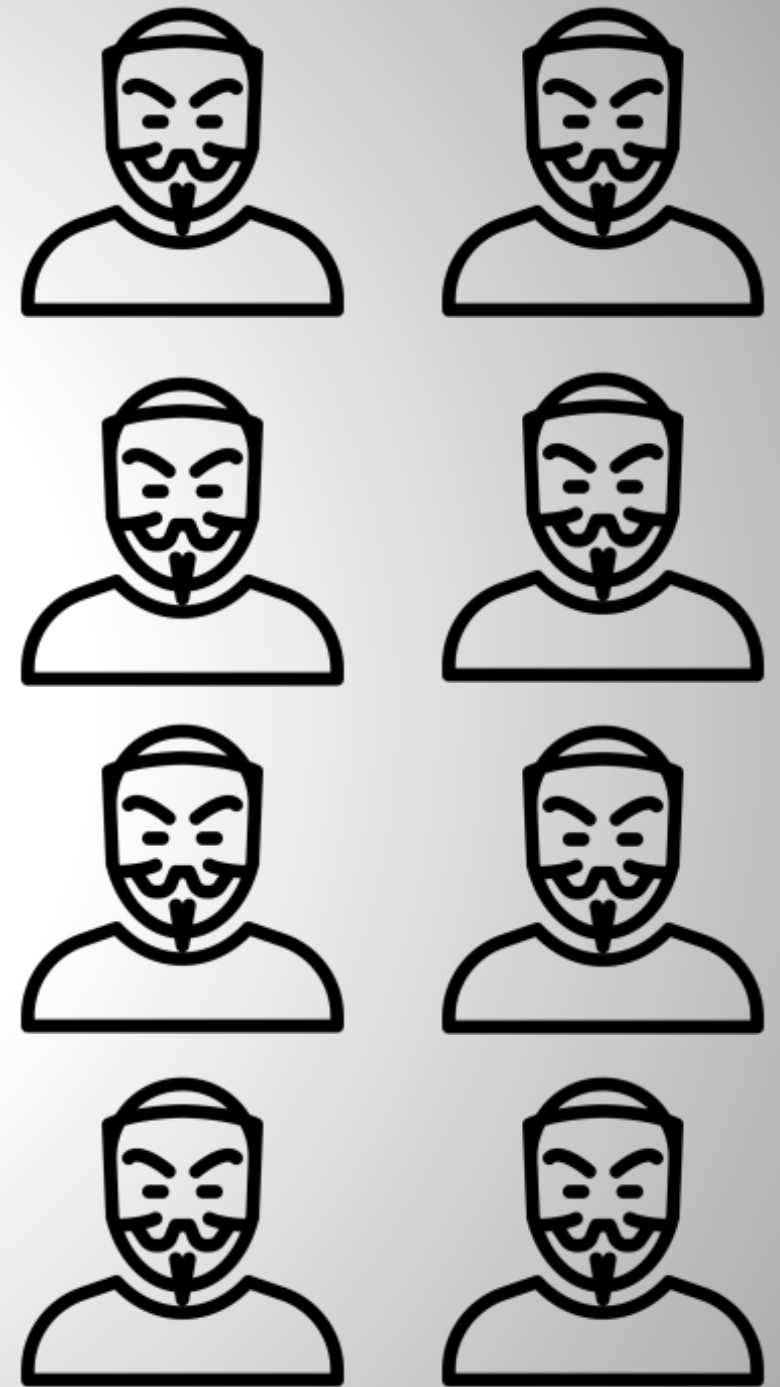
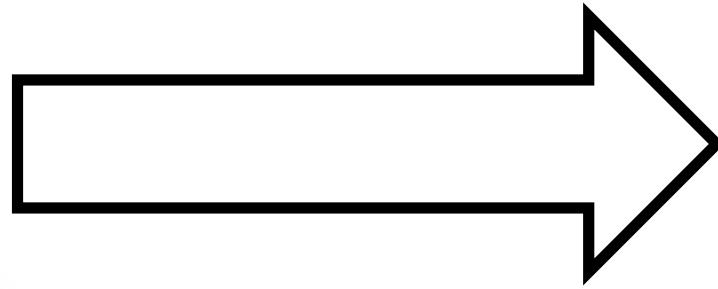
ANTI-FA
AREA





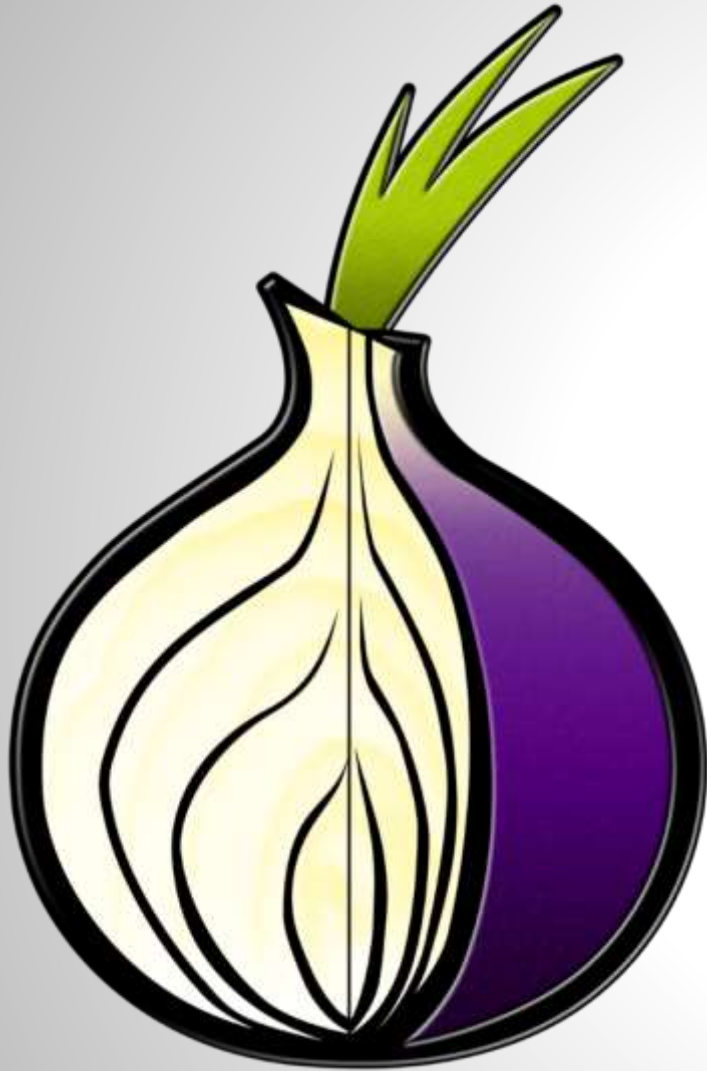
The larger the
network

The greater
the anonymity



Bio





1998

Free Haven Project

- The Onion Routing
- Ⓢ DARPA*
- 🎓 MIT



2002

DECLASSIFIED

- ◆ Launched
- ◆ Open-source

2009



Browser

- * Mozilla Firefox
- * Out-of-the box
- * Tor inside



Tor
Messenger

2015

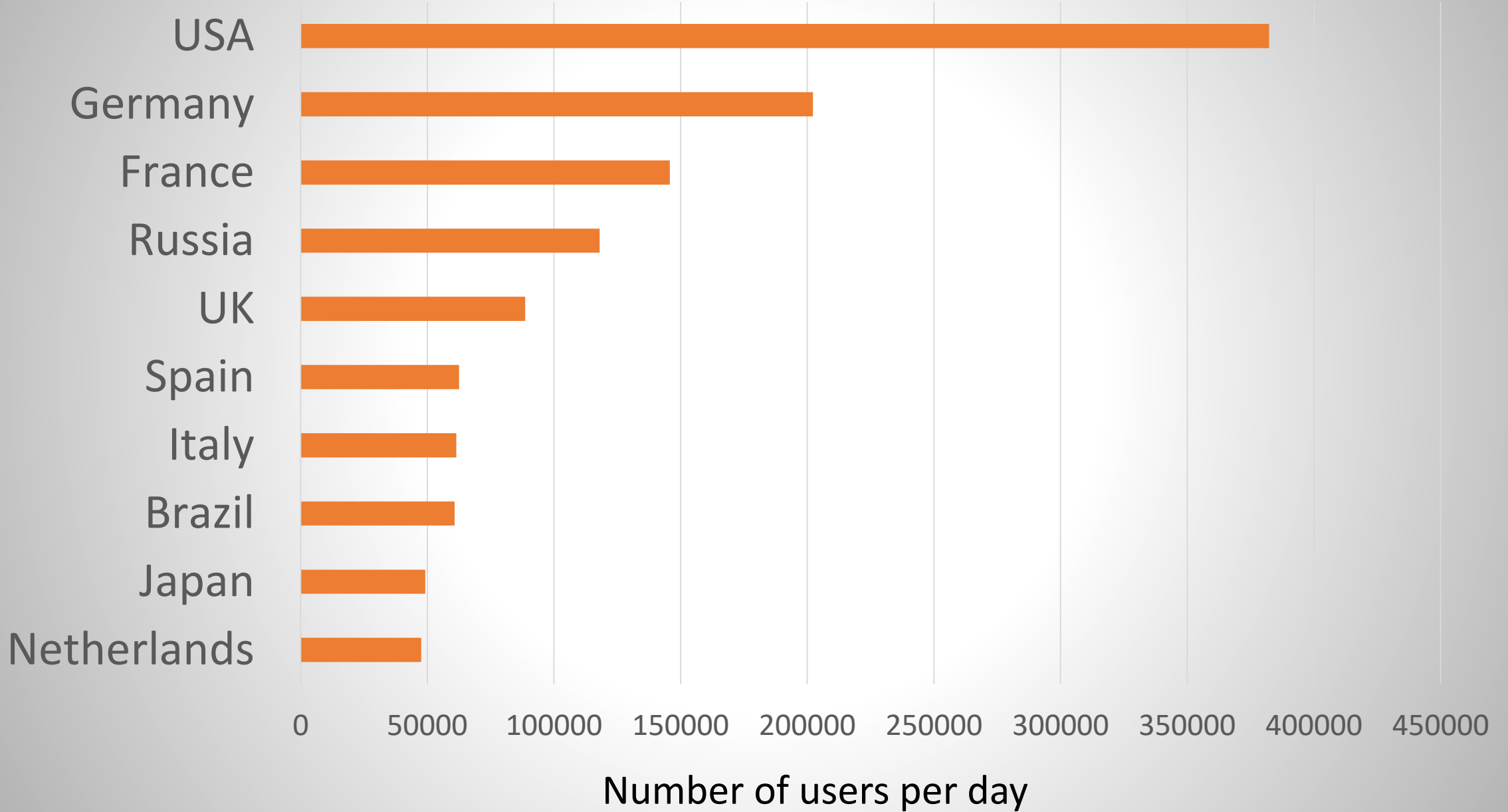
Messenger

- * Private chats
- * Anonymity
- * Tor inside

S T A T S

2 0000 0000

Users per day



Unique
Hidden
Services

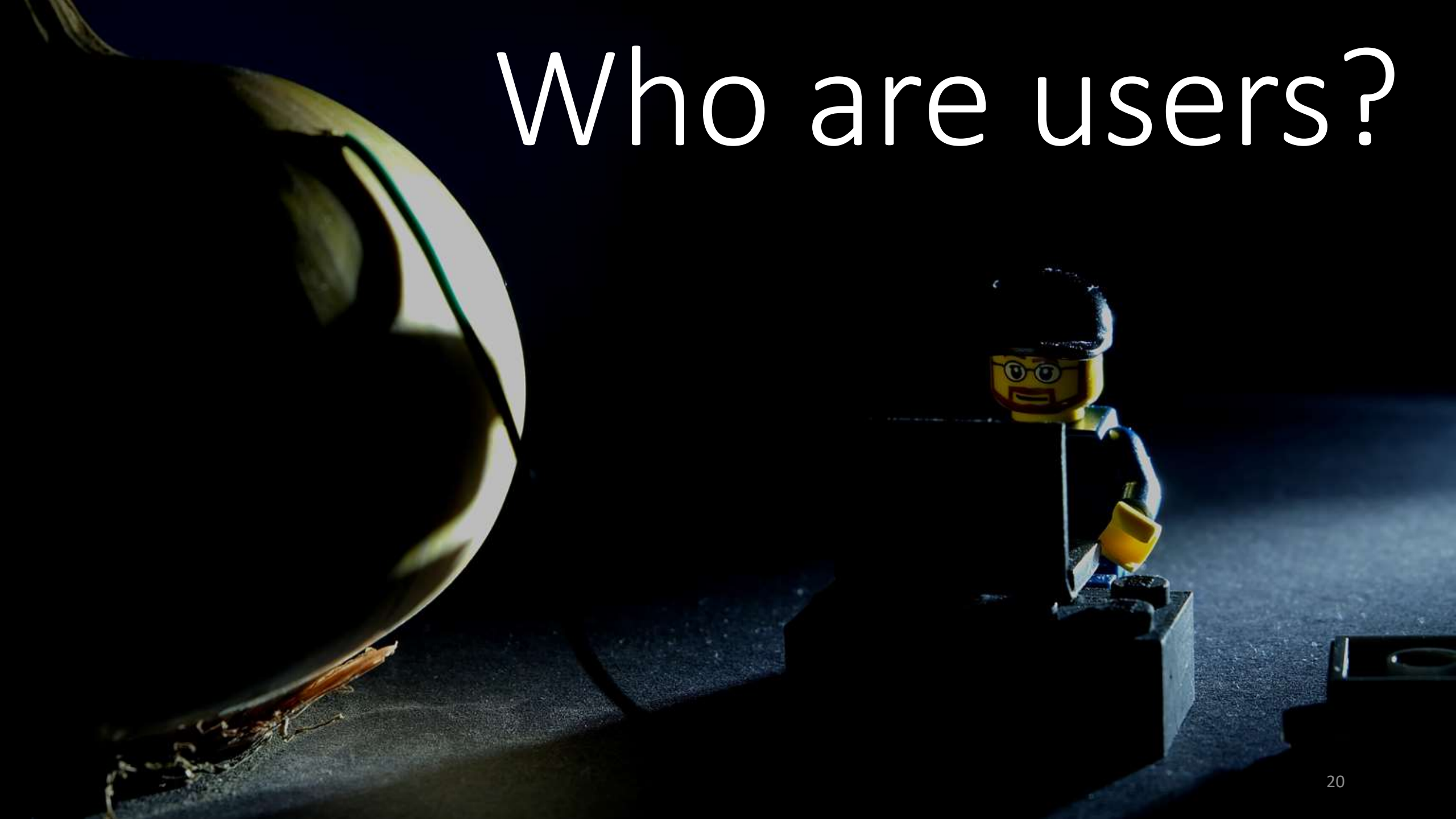
60K

7K

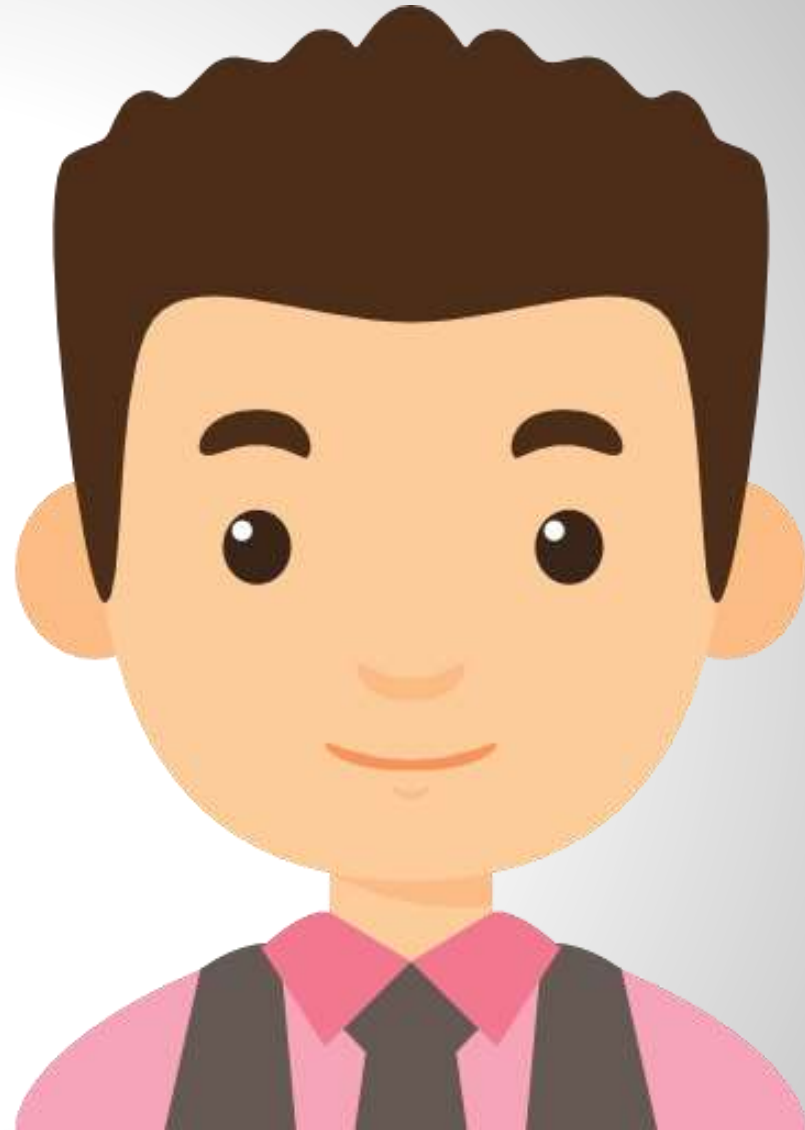
Tor

Relays

Who are users?



Just
people

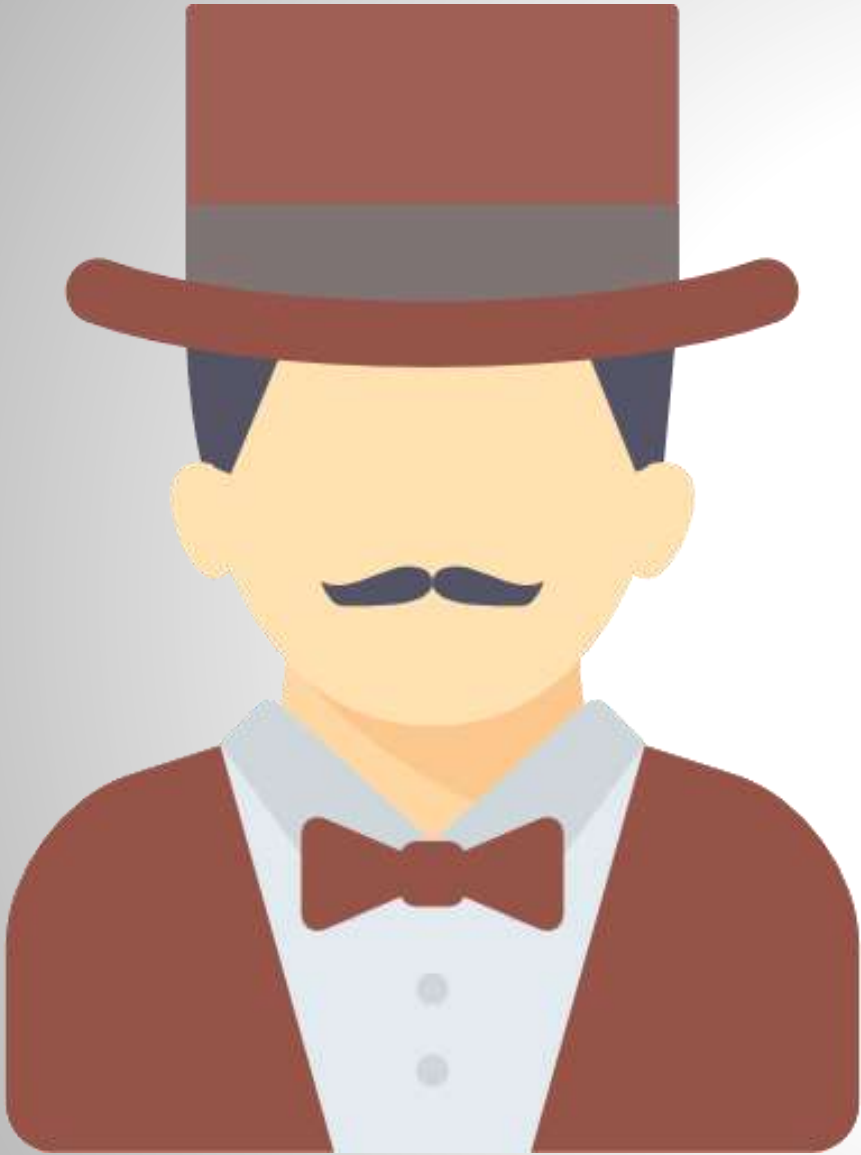




Journalists & Bloggers

Police
&
friends





Business

Military





IT

pros

Crime



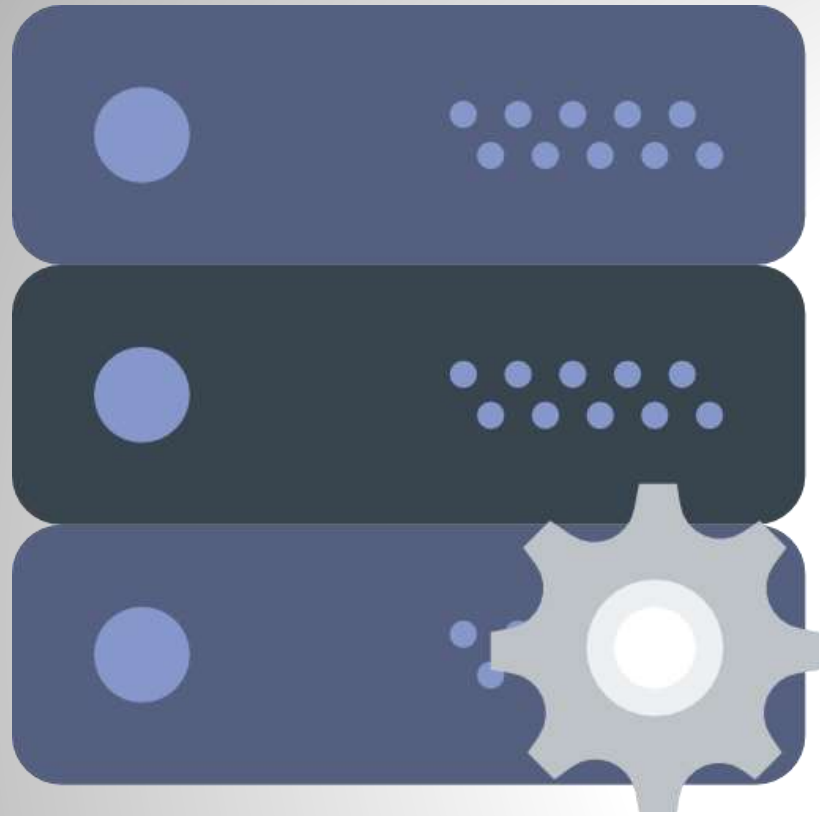
WHY

DEEP WEB?



Because

HIDDEN Services!



Anonymous
server

2004

.....
Anonymity for Servers
.....



Only for Tor




.onion

Inaccessible
On the
Internet



WikiLeaks:

<http://suw74isz7wqzpmgu.onion>



How
does it
work?

User



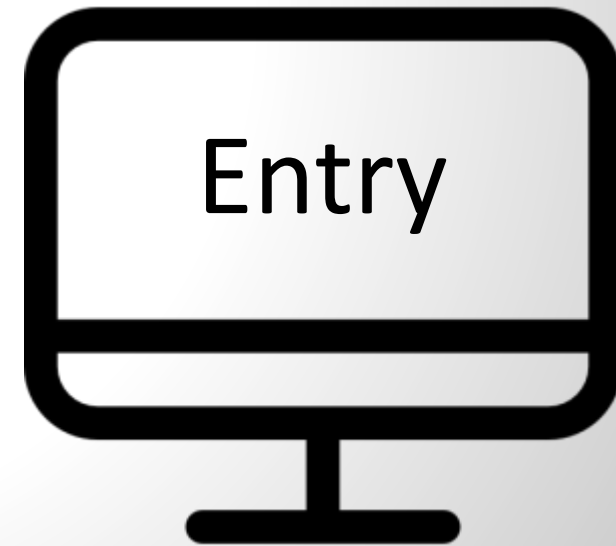
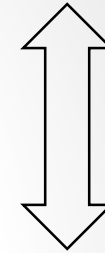
Tor Client

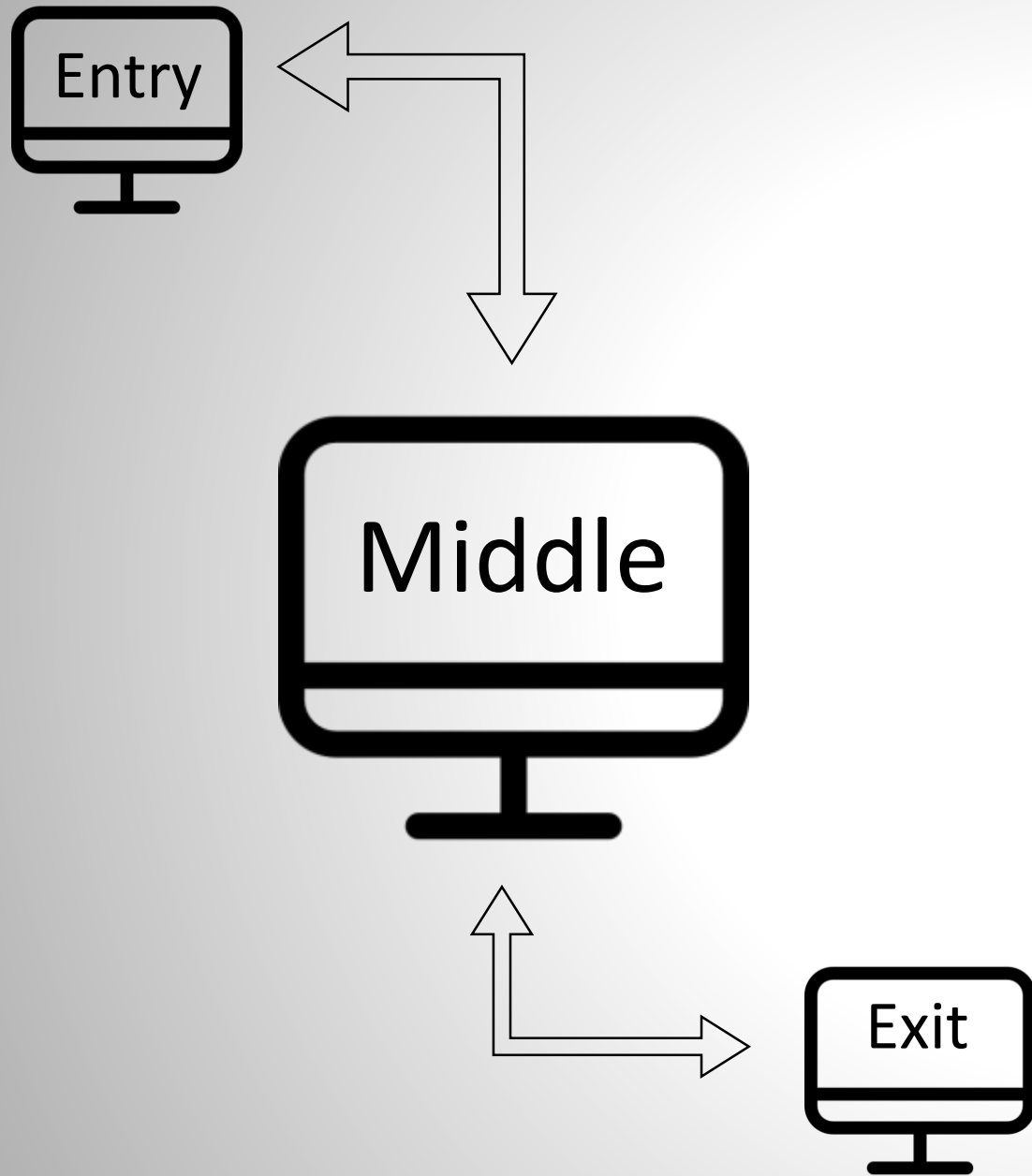
- * Connects with Tor
- * Has installed soft
- * Any PC

Relay

Entry guard

- * Speaks with Client
- * Encrypts data
- * Retranslates data





Relay

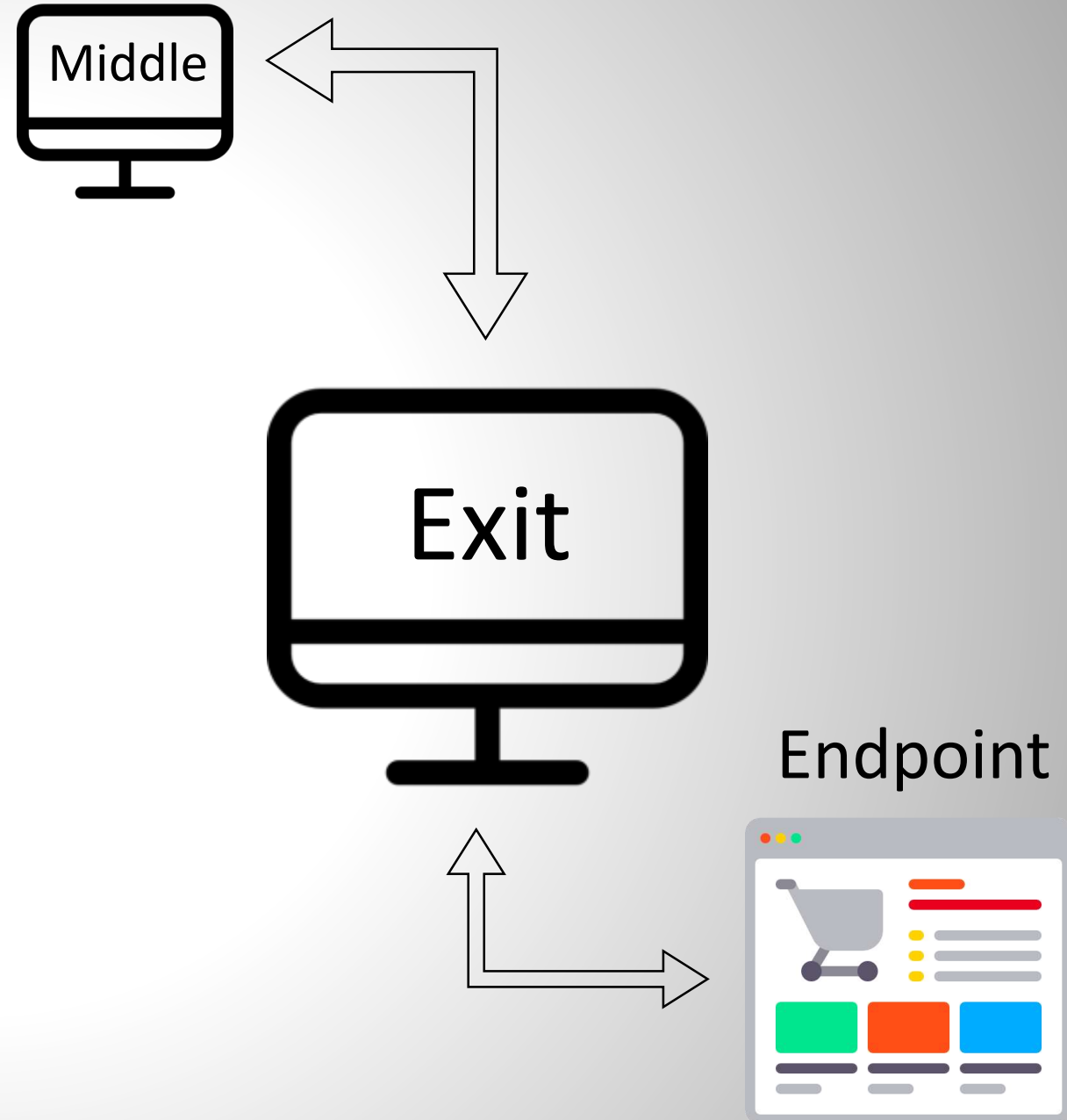
Middle

- * Speaks with Entry
- * Encrypts data
- * Speaks with Exit

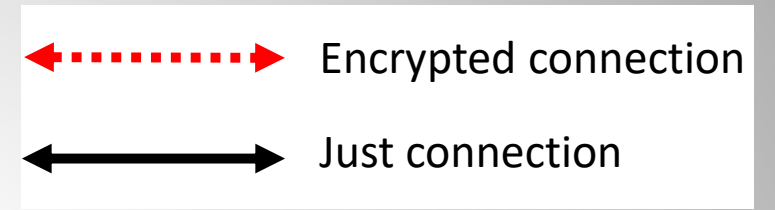
Relay

Exit

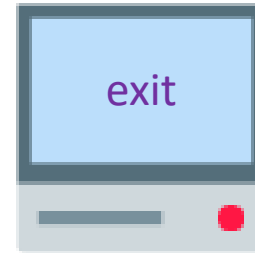
- * Speaks with Middle
- * Encrypts data
- * Speaks with Endpoint



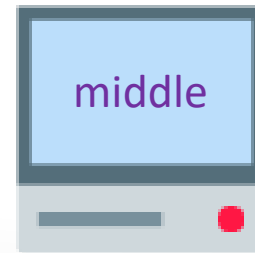
Default circuit



Tor Client



Endpoint



Step #1

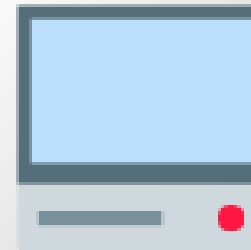
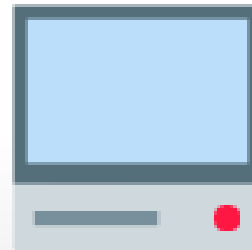
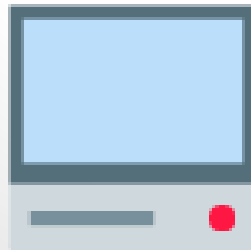
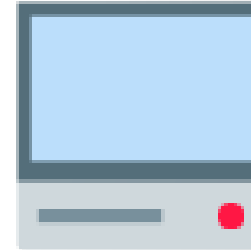
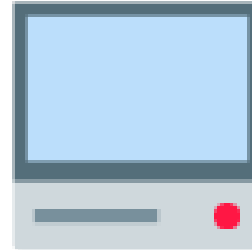
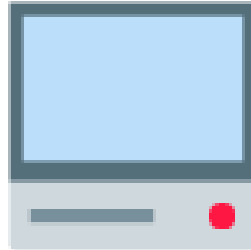
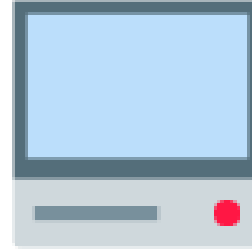
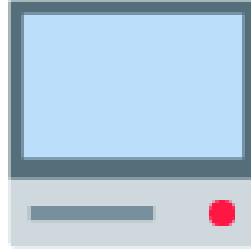


Tor Client



Directory server

Client receives the list of all Tor nodes from directory server

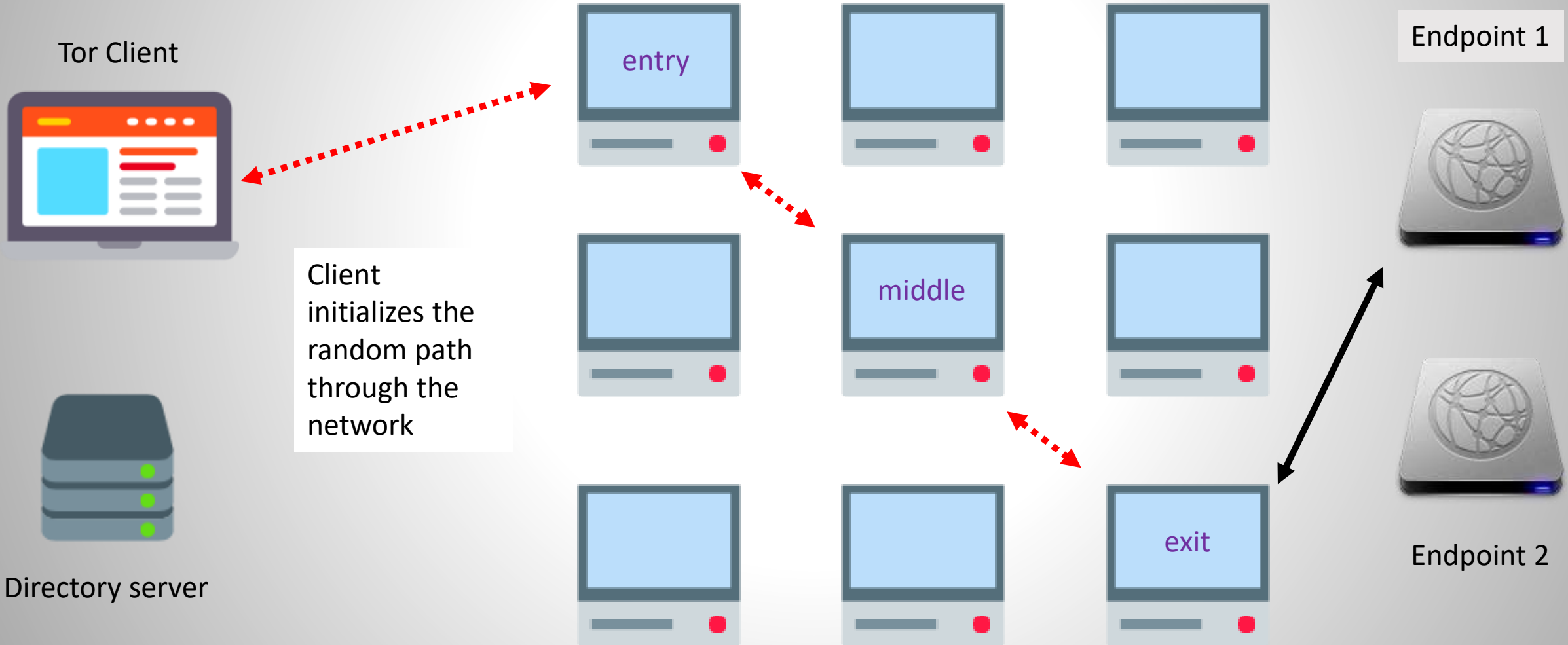
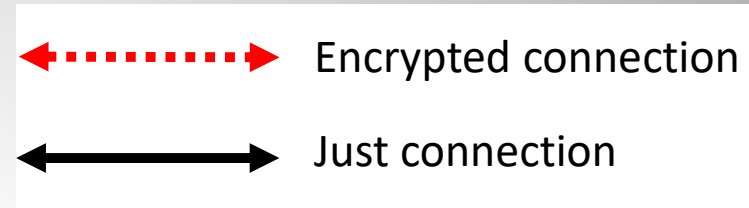


Endpoint #1

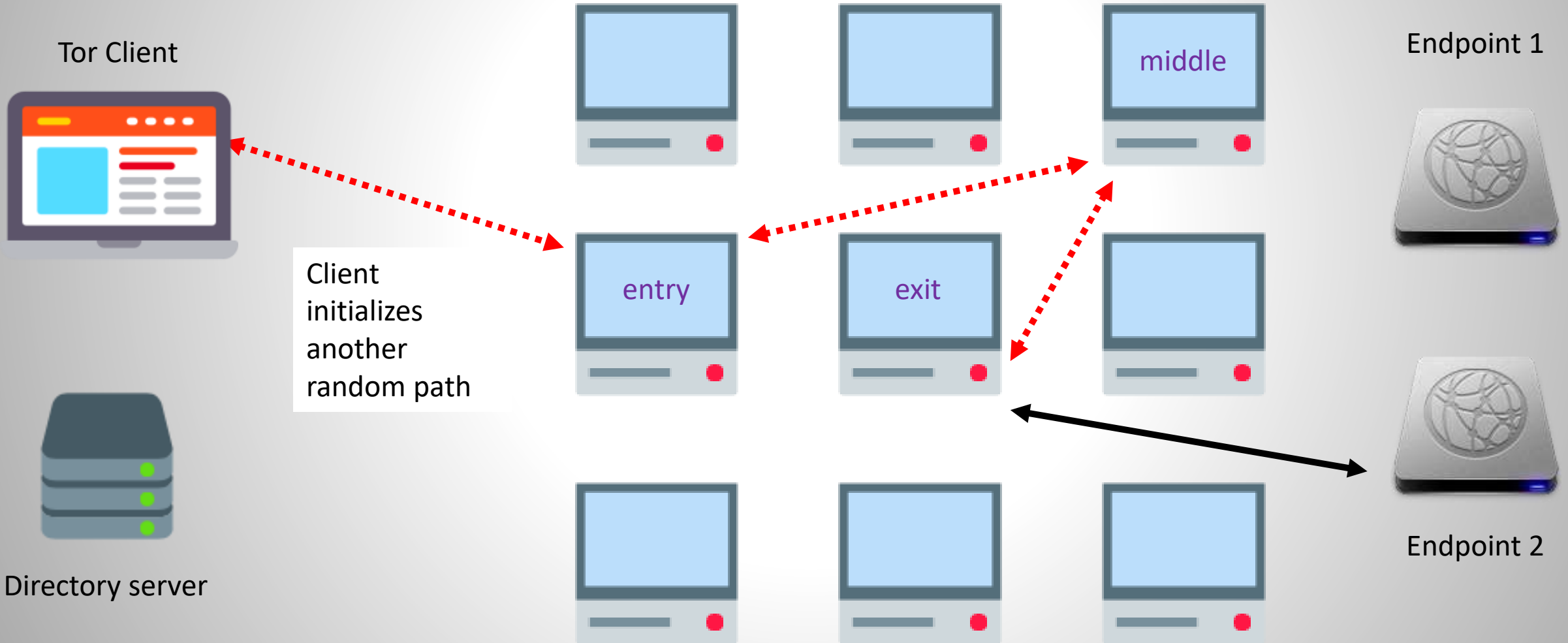
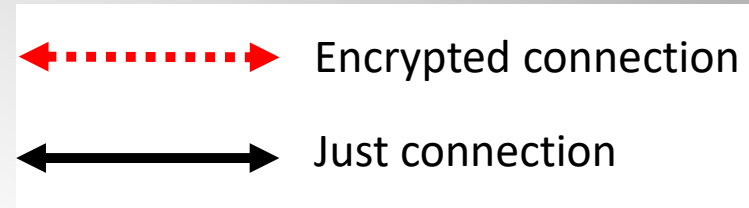


Endpoint #2

Step #2



Step #3

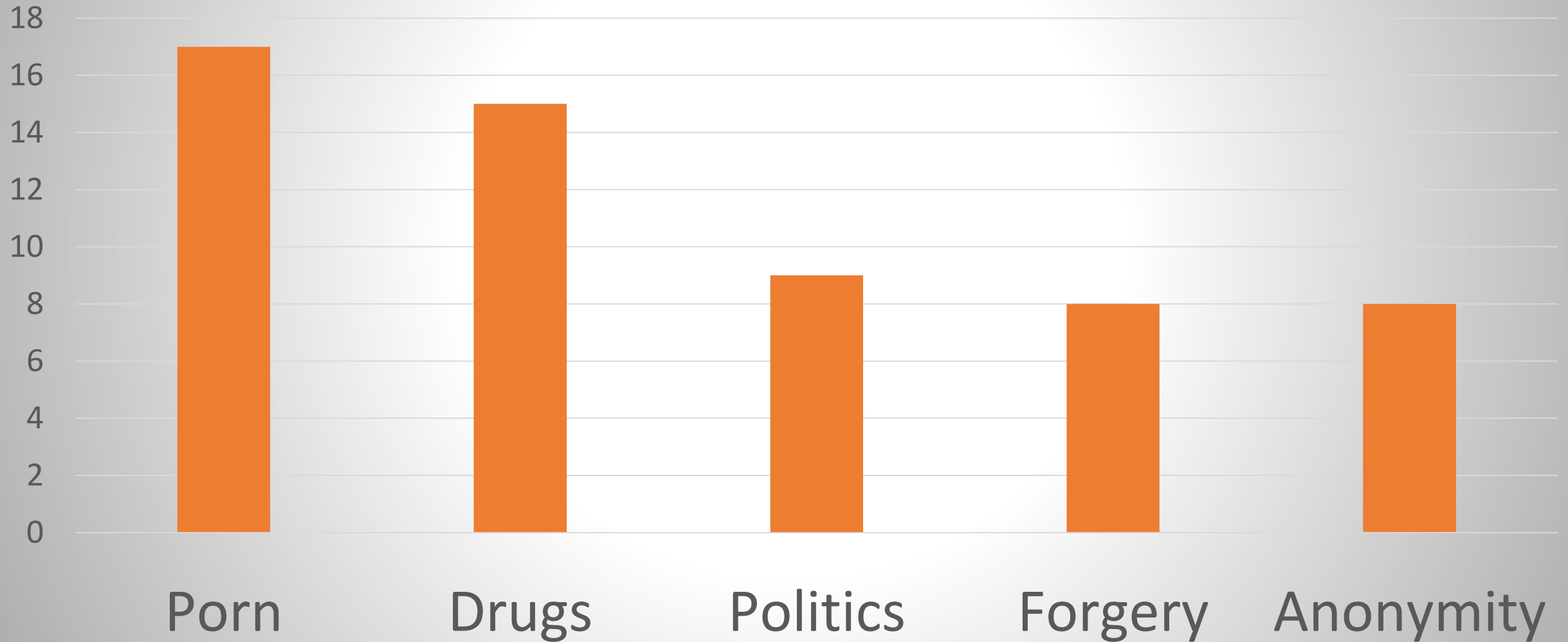


MYTH #1

ONLY
CRIMINALS
USE
TOR



The most popular content



MYTH #2

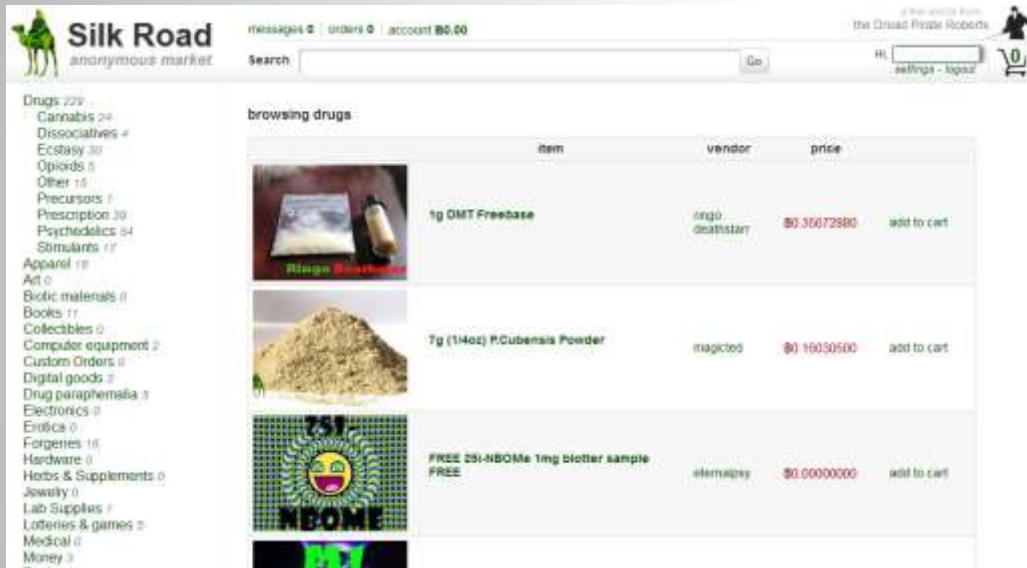
TOR IS
COMPLETELY
ANONYMOUS



Gov. VS Tor



Silk Road



Used to be the **biggest Drug Store**



Revenue:
9.5 mln BTC

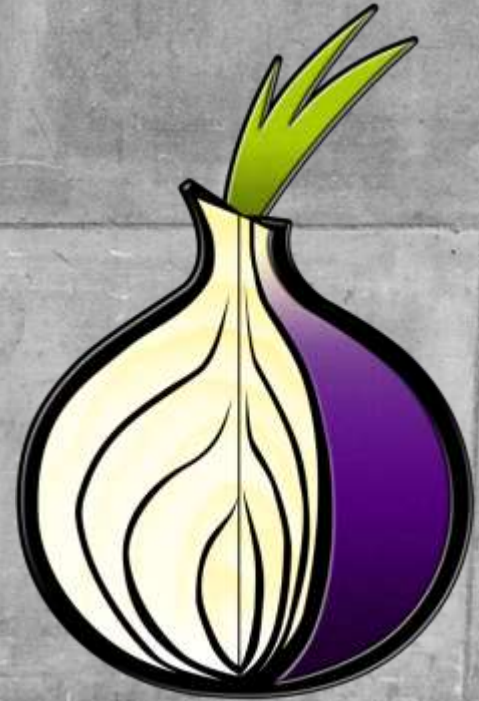
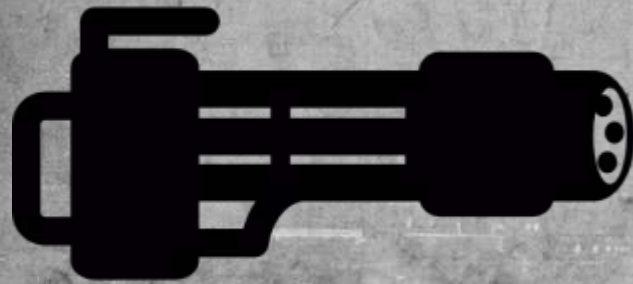


Closed by FBI

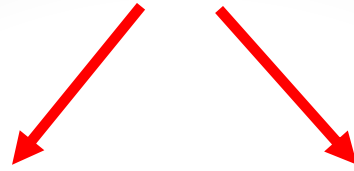


Founder is
life sentenced

Attacking Tor



Attacks

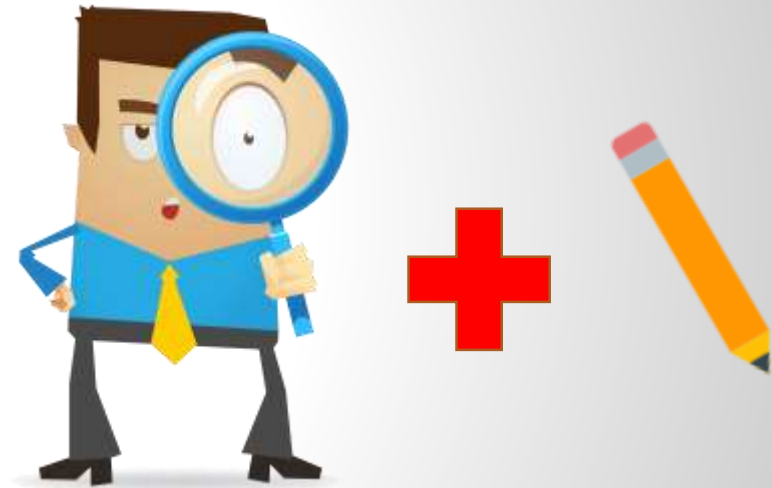


Passive



Attacker only **observes**
traffic, without
modifying it

Active



Attacker observes and
modifies traffic

Classification

#	Resources	Attacks
1	Corrupted entry guard	<ul style="list-style-type: none">• Website fingerprinting attack
2	Corrupted entry and exit nodes	<ul style="list-style-type: none">• Traffic analysis• Timing attack• Circuit fingerprinting attack• Tagging attack
3	Corrupted exit node	<ul style="list-style-type: none">• Sniffing of intercepted traffic
4	Corrupted entry and exit nodes, external server	<ul style="list-style-type: none">• Browser based timing attack with JavaScript injection• Browser based traffic analysis attack with JavaScript injection
5	Autonomous system	<ul style="list-style-type: none">• BGP hijacking• BGP interception• RAPTOR attack
6	Big number of various corrupted nodes	<ul style="list-style-type: none">• Packet spinning attack• CellFlood DoS attack• Other DoS and DDoS attacks

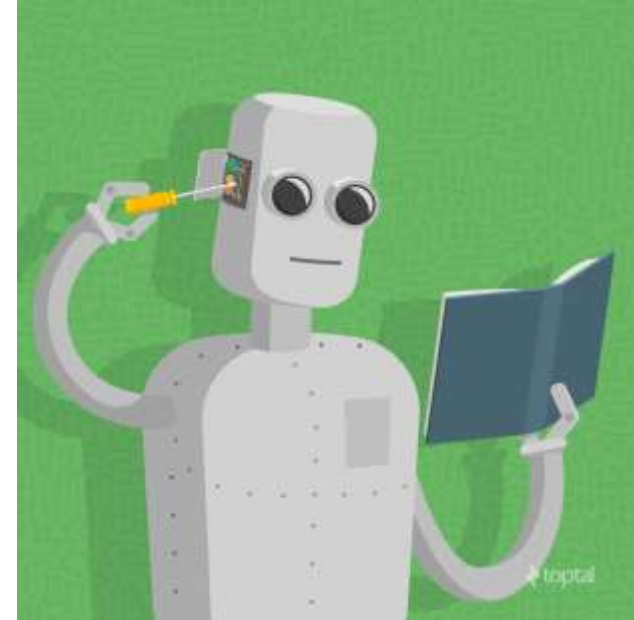
Website fingerprinting attack



The Idea:

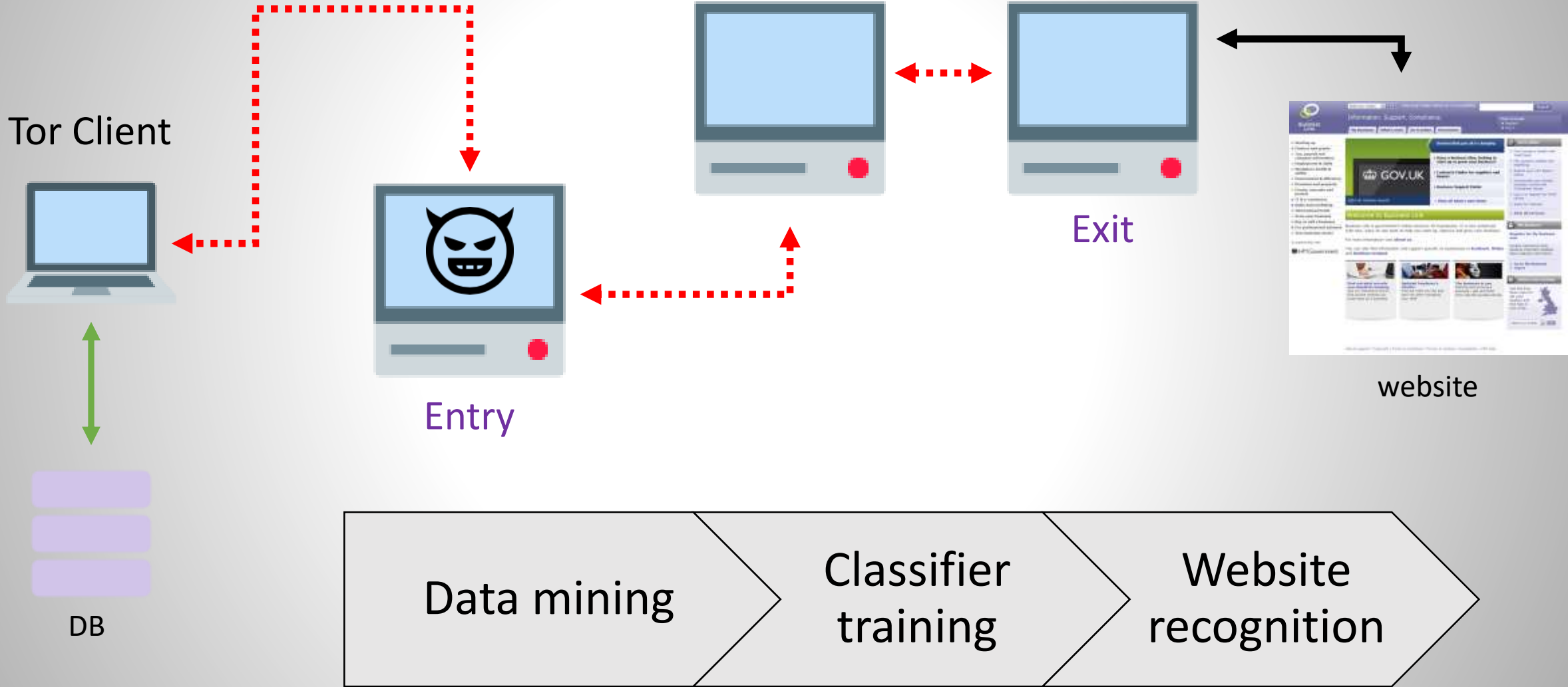


Data mining



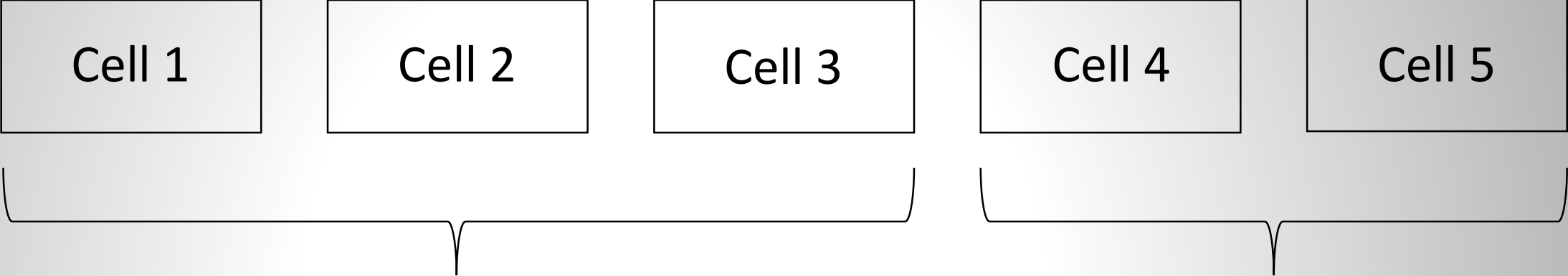
Machine learning

Attackers strategy

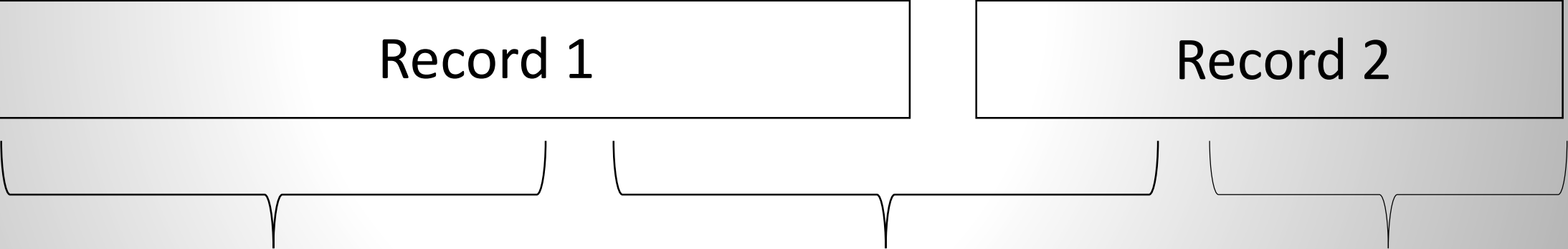


Feature extraction levels

Cells



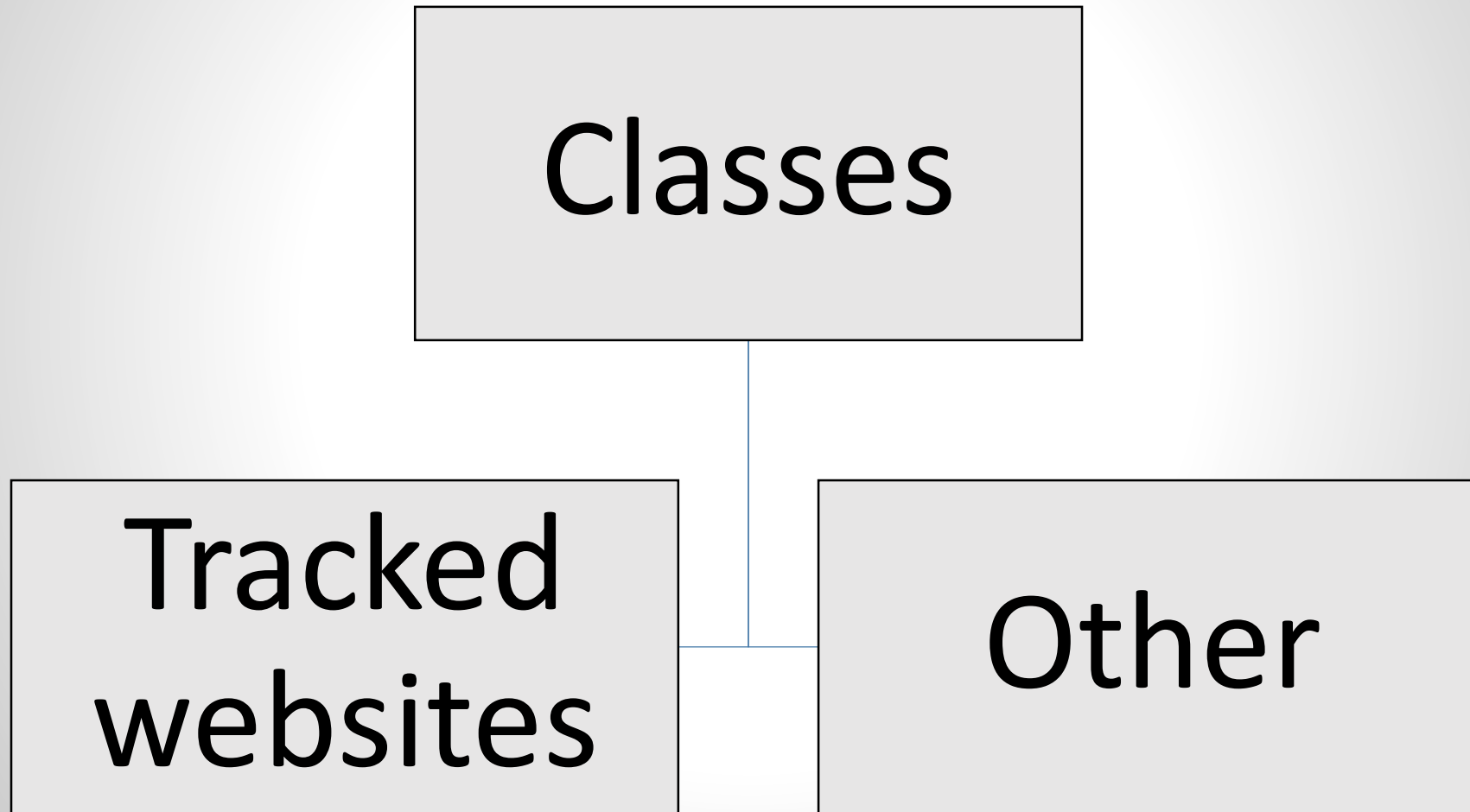
TLS



TCP



Attack as a classification problem



Problem?

1	0...	192.168.4.53	95.85.8.226	TLSv1.2	599	Application Data
2	0...	95.85.8.226	192.168.4.53	TCP	56	443 → 63506 [ACK] Seq=1 Ack=544 Win=164 Len=0 TSval...
3	0...	95.85.8.226	192.168.4.53	TLSv1.2	599	Application Data
4	0...	192.168.4.53	95.85.8.226	TCP	56	63506 → 443 [ACK] Seq=544 Ack=544 Win=4079 Len=0 TS...
5	0...	192.168.4.53	95.85.8.226	TLSv1.2	599	Application Data
6	0...	95.85.8.226	192.168.4.53	TCP	56	443 → 63506 [ACK] Seq=544 Ack=1087 Win=164 Len=0 TS...
7	0...	192.168.4.53	95.85.8.226	TLSv1.2	599	Application Data
8	0...	95.85.8.226	192.168.4.53	TCP	56	443 → 63506 [ACK] Seq=544 Ack=1630 Win=162 Len=0 TS...
9	0...	95.85.8.226	192.168.4.53	TLSv1.2	599	Application Data
...	0...	192.168.4.53	95.85.8.226	TCP	56	63506 → 443 [ACK] Seq=1630 Ack=1087 Win=4079 Len=0 ...
...	0...	192.168.4.53	95.85.8.226	TLSv1.2	599	Application Data
...	0...	95.85.8.226	192.168.4.53	TLSv1.2	599	Application Data
...	0...	192.168.4.53	95.85.8.226	TCP	56	63506 → 443 [ACK] Seq=2173 Ack=1630 Win=4079 Len=0 ...
...	0...	192.168.4.53	95.85.8.226	TLSv1.2	599	Application Data
...	1...	95.85.8.226	192.168.4.53	TLSv1.2	599	Application Data
...	1...	192.168.4.53	95.85.8.226	TCP	56	63506 → 443 [ACK] Seq=2716 Ack=2173 Win=4079 Len=0 ...
...	1...	95.85.8.226	192.168.4.53	TCP	56	443 → 63506 [ACK] Seq=2173 Ack=2716 Win=166 Len=0 T...
...	1...	192.168.4.53	95.85.8.226	TLSv1.2	599	Application Data
...	1...	95.85.8.226	192.168.4.53	TLSv1.2	599	Application Data
...	1...	192.168.4.53	95.85.8.226	TCP	56	63506 → 443 [ACK] Seq=3259 Ack=2716 Win=4079 Len=0 ...
...	1...	95.85.8.226	192.168.4.53	TCP	56	443 → 63506 [ACK] Seq=2716 Ack=3259 Win=164 Len=0 T...
...	1...	95.85.8.226	192.168.4.53	TLSv1.2	599	Application Data

The Oracle problem!

Data collection

Choose number of instances



Choose mining folder

Mining options are not configured yet. You have to choose folder.

Progress of uploading data will be shown here

There will be 1 traffic instance collected for each website

Start Mining

Model

Classifier is not configured

Choose already created classifier

Choose

Create new classifier

yourclassifiername.plk

Train file

Create

Update classifier

Choose

Train set

Update classifier

Test

Enter url for testing classifier

http://torWfTool.onion/

Predicted label will be displayed here

Test classifier on entered url

Test classifier on a dataset

Choose dataset

No dataset to use for testing

Predicted accuracy unavailable

Test classifier on a dataset

Monitored webpages

[19/05/2016 00:16:41]: Program is ready.

Edit

Choose

7

Websites

5

Men

5

Seconds
split

1

Relay

80

Traffic
Instances

5

Uploads
per website

0.71

Accuracy

THANK
YOU!