Lecture 34 – The Dark Web 2

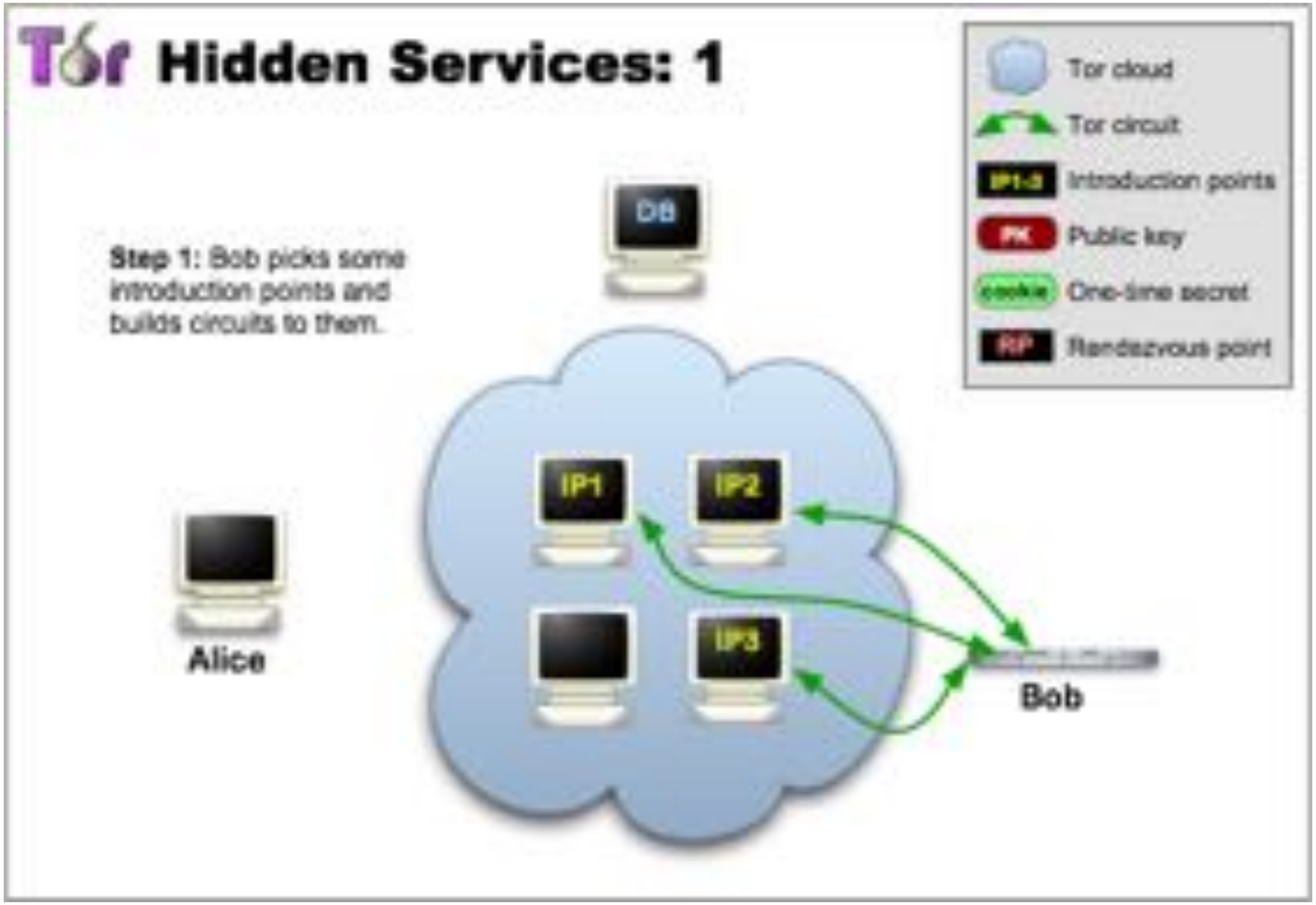
Ryan Cunningham
University of Illinois
ECE 422/CS 461 – Fall 2017

Security News

- Win 8 & 10 ASLR vulnerability
- 57M Uber users’ data breached, covered up breach and paid attackers 100k
- DOD left AWS S3 cloud storage buckets wide open
- Security researchers find flaws in IME

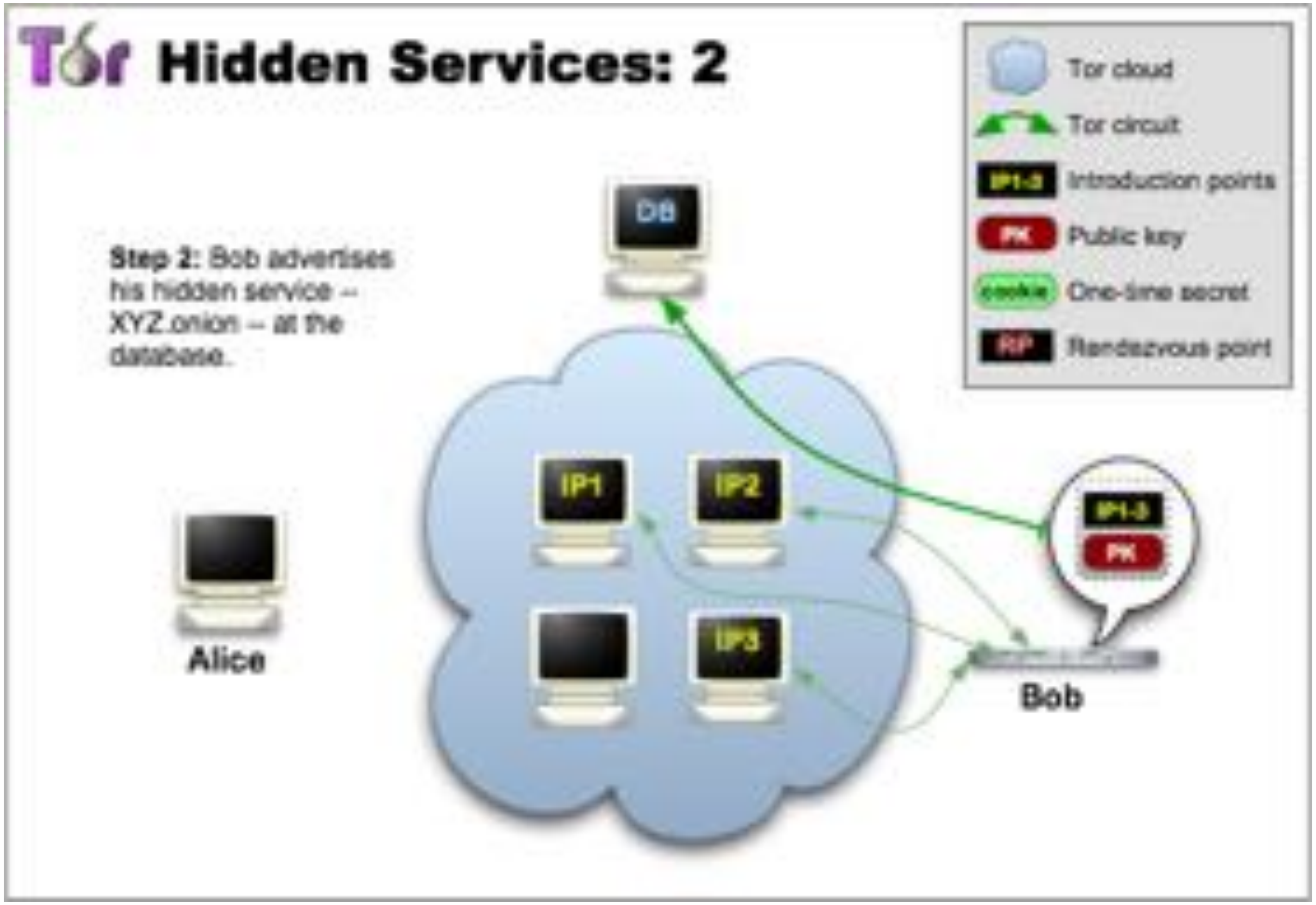
TOR HIDDEN SERVICES

Tor Hidden Services: Overview



Step 1: Bob picks some introduction points and builds circuits to them.

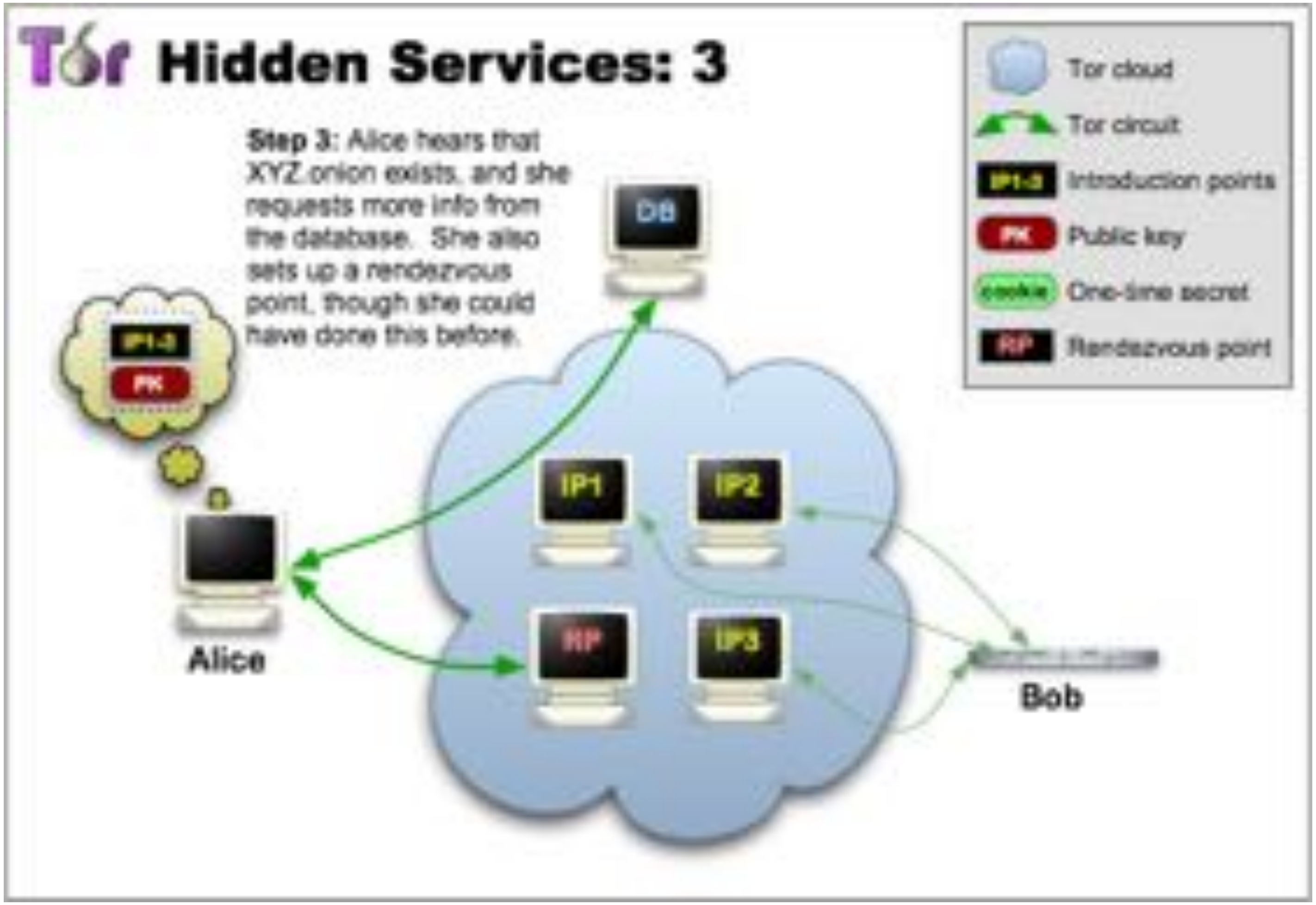
Tor Hidden Services



Tor Hidden Services: 2

Step 2: Bob advertises his hidden service -- XYZ.onion -- at the database.

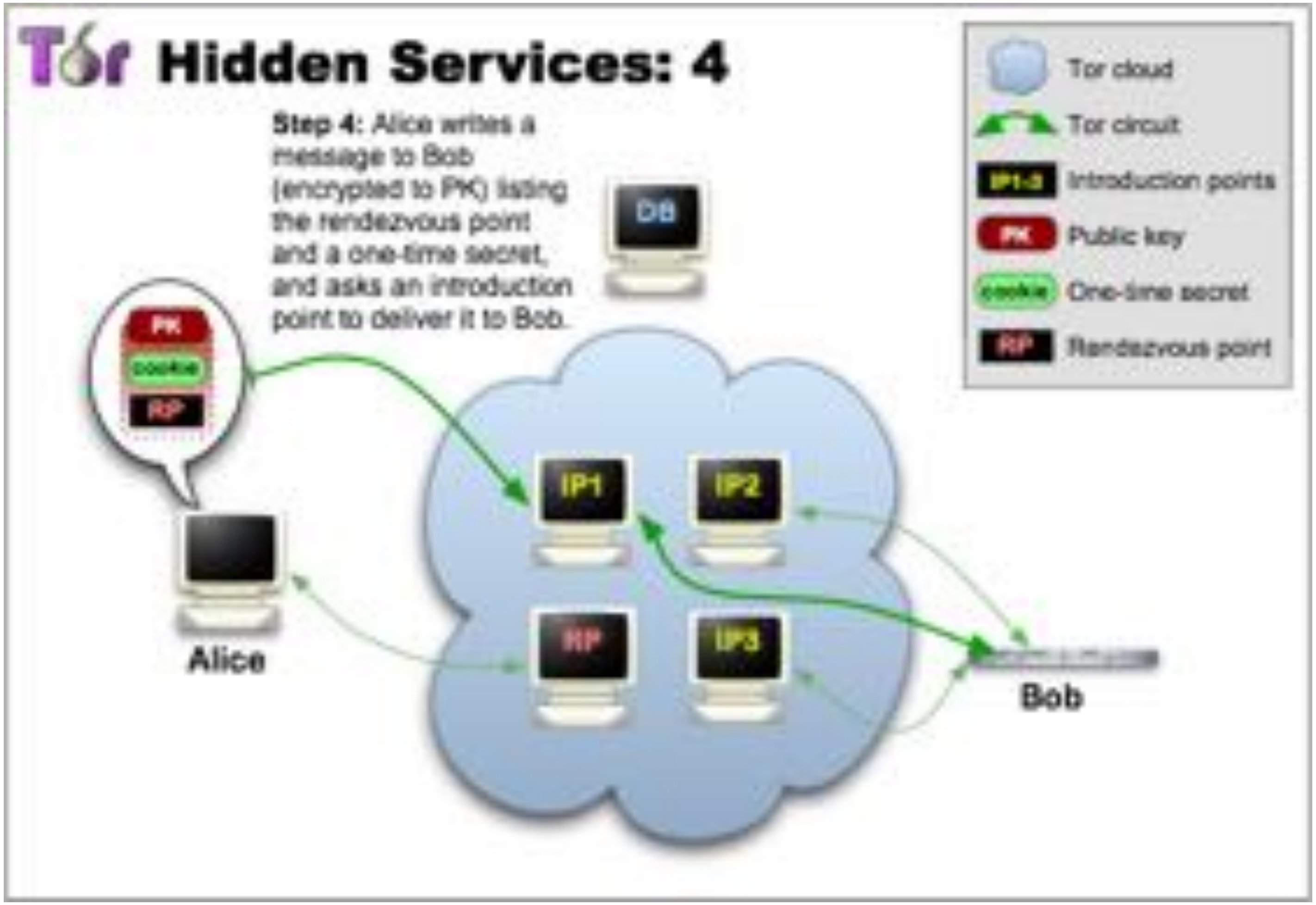
Tor Hidden Services



Tor Hidden Services: 3

Step 3: Alice hears that XYZ.onion exists, and she requests more info from the database. She also sets up a rendezvous point, though she could have done this before.

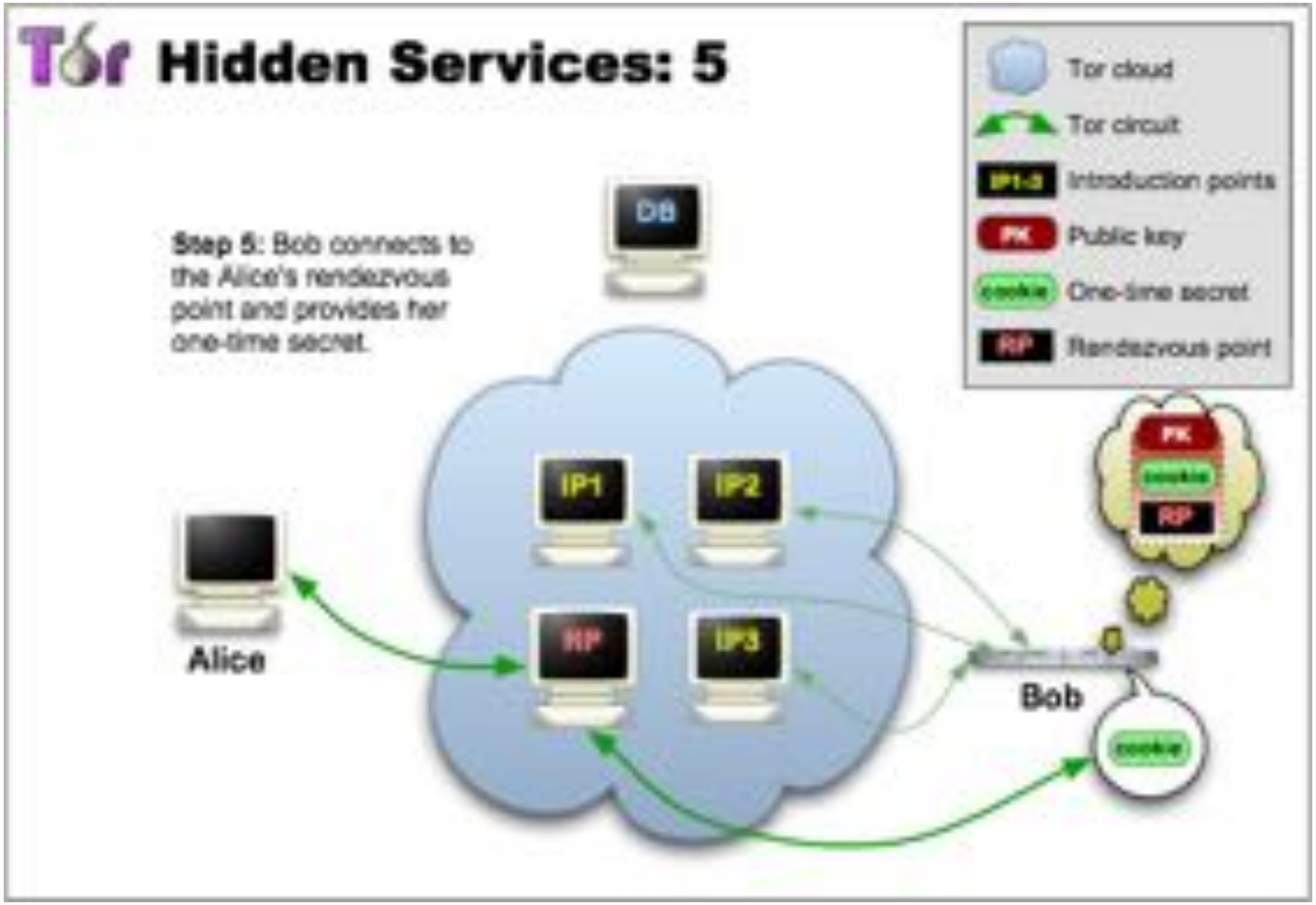
Tor Hidden Services



Tor Hidden Services: 4

Step 4: Alice writes a message to Bob (encrypted to PK) listing the rendezvous point and a one-time secret, and asks an introduction point to deliver it to Bob.

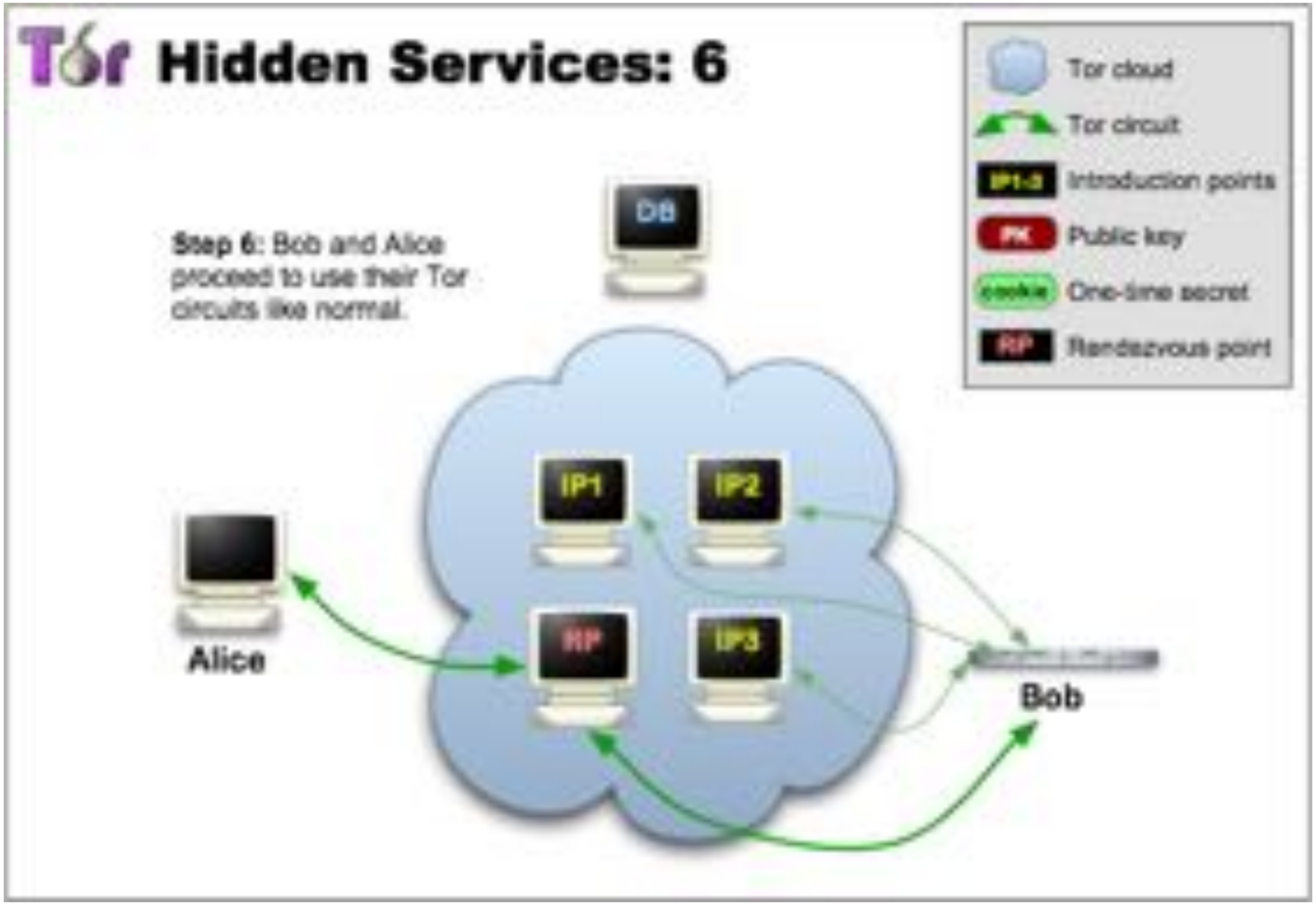
Tor Hidden Services



Tor Hidden Services: 5

Step 5: Bob connects to the Alice's rendezvous point and provides her one-time secret.

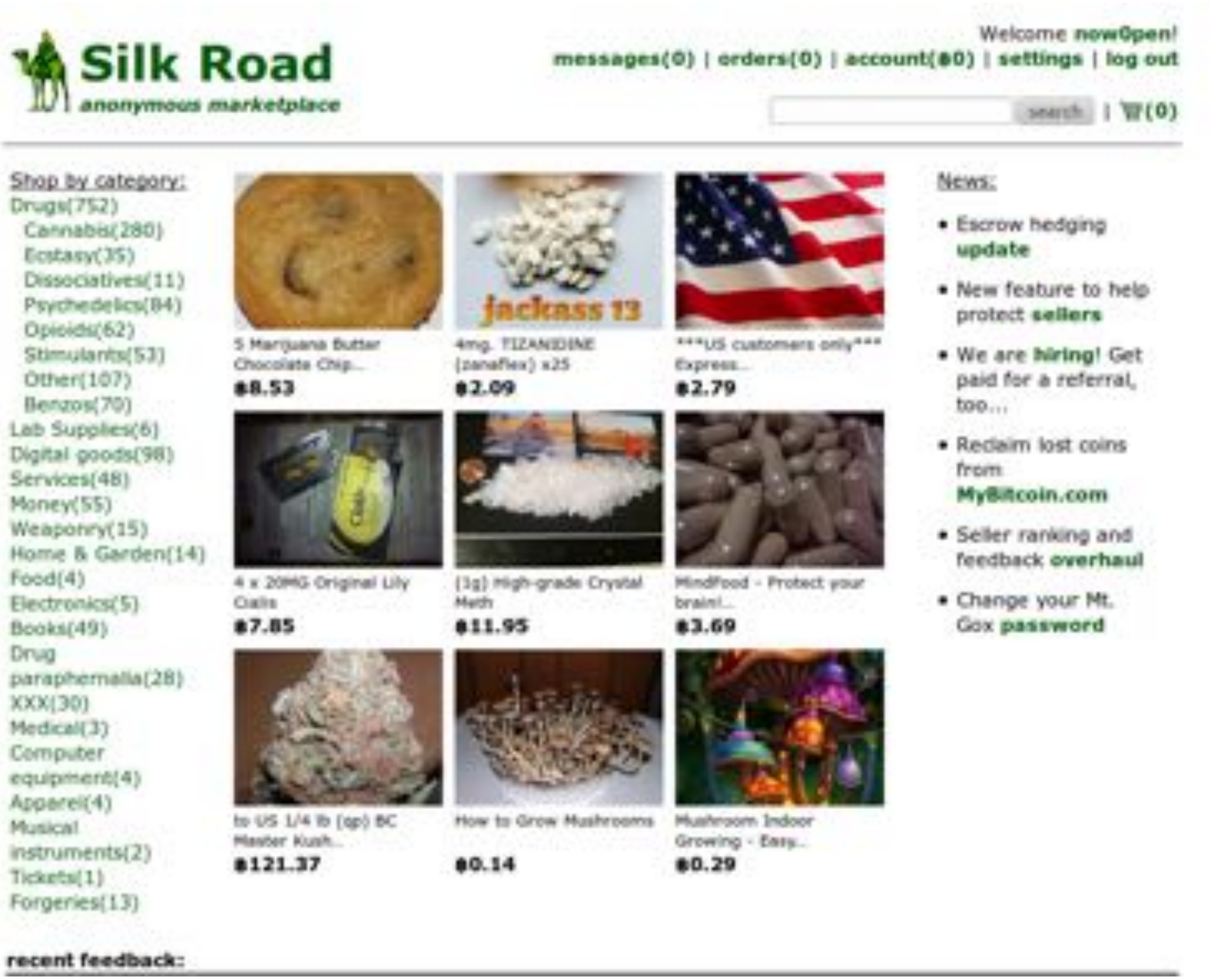
Tor Hidden Services



Tor Hidden Services: 6

Step 6: Bob and Alice proceed to use their Tor circuits like normal.

SilkRoad Marketplace



Silk Road
anonymous marketplace

Welcome **now0pen!**
messages(0) | orders(0) | account(฿0) | settings | log out

search | (0)

Shop by category:
Drugs(752)
Cannabis(280)
Ecstasy(35)
Dissociatives(11)
Psychedelics(84)
Opioids(62)
Stimulants(53)
Other(107)
Benzos(70)
Lab Supplies(6)
Digital goods(98)
Services(48)
Money(55)
Weaponry(15)
Home & Garden(14)
Food(4)
Electronics(5)
Books(49)
Drug paraphernalia(28)
XXX(30)
Medical(3)
Computer equipment(4)
Apparel(4)
Musical instruments(2)
Tickets(1)
Forgeries(13)

5 Marijuana Butter Chocolate Chip...
฿8.53

4mg. TIZANIDINE (zanaflex) x25
฿2.09

US customers only Express...
฿2.79

4 x 20MG Original Lilly Cialis
฿7.85

[1g] High-grade Crystal Meth
฿11.95

MindFood - Protect your brain!...
฿3.69

to US 1/4 lb (qp) BC Master Kush...
฿121.37

How to Grow Mushrooms
฿0.14

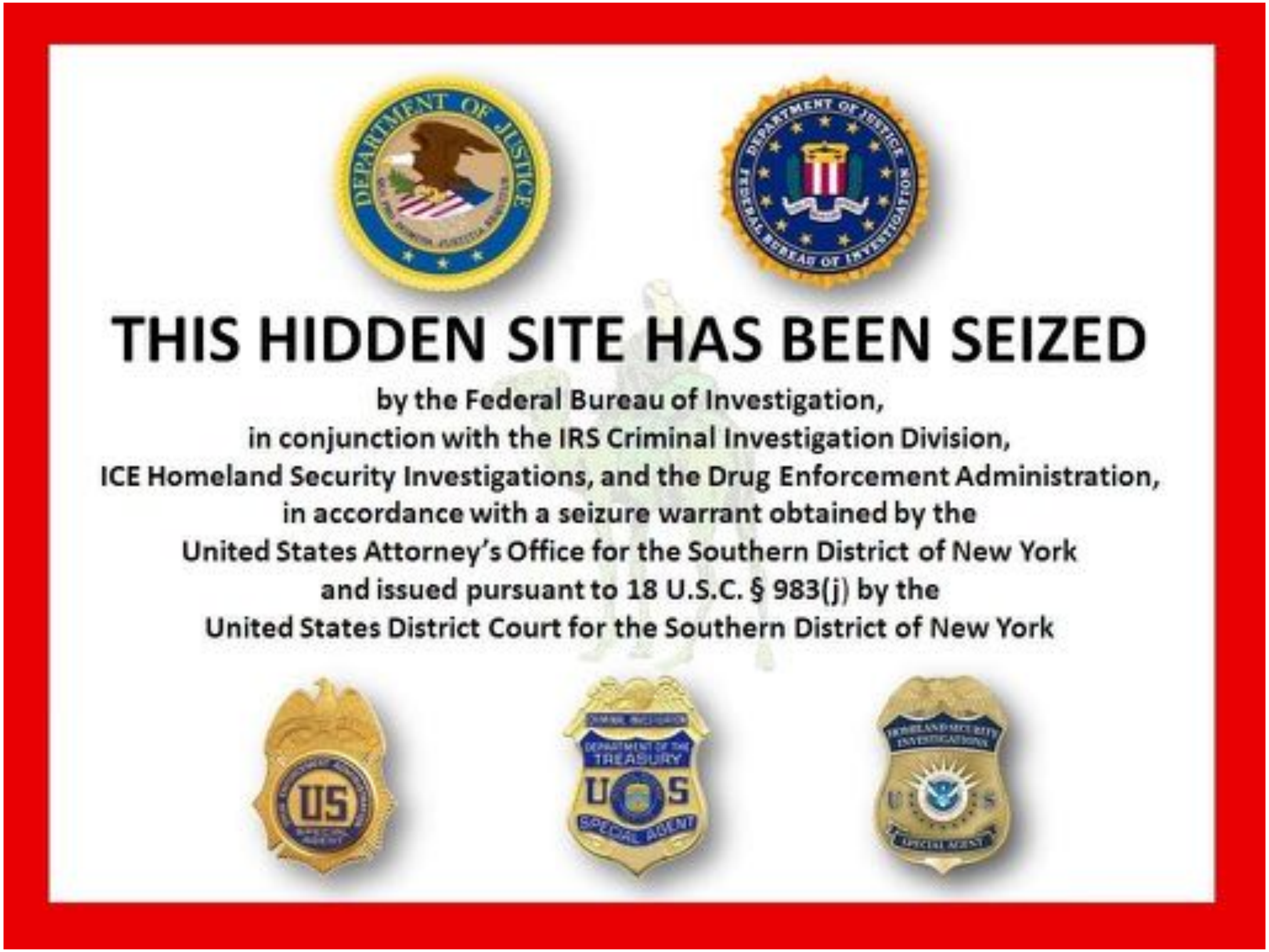
Mushroom Indoor Growing - Easy...
฿0.29

News:

- Escrow hedging **update**
- New feature to help protect **sellers**
- We are **hiring!** Get paid for a referral, too...
- Reclaim lost coins from **MyBitcoin.com**
- Seller ranking and feedback **overhaul**
- Change your Mt. Gox **password**

recent feedback:

SilkRoad Marketplace



DEPARTMENT OF JUSTICE
DEPARTMENT OF JUSTICE
FEDERAL BUREAU OF INVESTIGATION
US
DEPARTMENT OF THE TREASURY
SPECIAL AGENT
HOMELAND SECURITY INVESTIGATIONS
SPECIAL AGENT

THIS HIDDEN SITE HAS BEEN SEIZED

by the Federal Bureau of Investigation,
in conjunction with the IRS Criminal Investigation Division,
ICE Homeland Security Investigations, and the Drug Enforcement Administration,
in accordance with a seizure warrant obtained by the
United States Attorney's Office for the Southern District of New York
and issued pursuant to 18 U.S.C. § 983(j) by the
United States District Court for the Southern District of New York

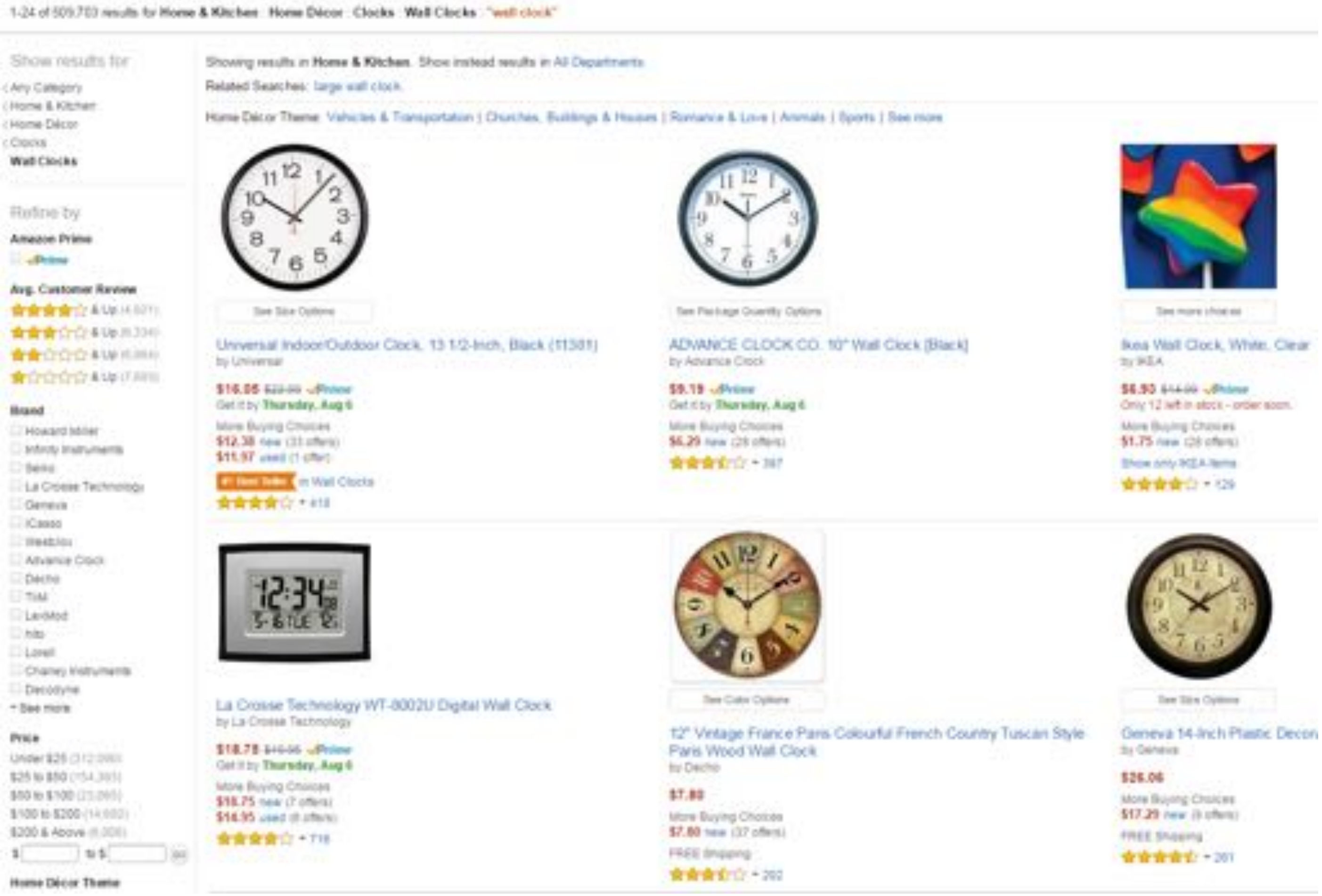
Measuring the Longitudinal Evolution of the Online Anonymous Marketplace Ecosystem

Kyle Soska
Carnegie Mellon University
ECE / Cylab
ksoska@cmu.edu

Nicolas Christin
Carnegie Mellon University
ECE / Cylab
nicolasc@cmu.edu

Conventional Commerce

Internet Commerce



1-24 of 509,703 results for **Home & Kitchen** : **Home Décor** : **Clocks** : **Wall Clocks** : "wall clock"

Show results for

‹ Any Category
‹ Home & Kitchen
‹ Home Décor
‹ Clocks
Wall Clocks

Refine by

Amazon Prime

□ ✓Prime

Avg. Customer Review

★★★★☆ & Up (4,971)
★★★☆☆ & Up (6,334)
★★☆☆☆ & Up (6,864)
★☆☆☆☆ & Up (7,695)

Brand

□ Howard Miller
□ Infinity Instruments
□ Seiko
□ La Crosse Technology
□ Geneva
□ iCasso
□ Westclox
□ Advance Clock
□ Decho
□ TiiM
□ LeviMod
□ hito
□ Lorell
□ Chaney Instruments
□ Decodyne
+ See more

Price

Under $25 (312,090)
$25 to $50 (154,393)
$50 to $100 (23,065)
$100 to $200 (14,693)
$200 & Above (6,306)
$ ______ to $ ______ Go

Home Décor Theme

Showing results in **Home & Kitchen**. Show instead results in All Departments

Related Searches: large wall clock.

Home Décor Theme: Vehicles & Transportation | Churches, Buildings & Houses | Romance & Love | Animals | Sports | See more

See Size Options

Universal Indoor/Outdoor Clock, 13 1/2-Inch, Black (11381)
by Universal

$16.05 ~~$22.99~~ ✓Prime
Get it by Thursday, Aug 6

More Buying Choices
$12.38 new (33 offers)
$11.97 used (1 offer)

#1 Best Seller in Wall Clocks

★★★★☆ ▾ 418

See Package Quantity Options

ADVANCE CLOCK CO. 10" Wall Clock [Black]
by Advance Clock

$9.19 ✓Prime
Get it by Thursday, Aug 6

More Buying Choices
$6.29 new (28 offers)

★★★½☆ ▾ 387

See more choices

Ikea Wall Clock, White, Clear
by IKEA

$6.93 ~~$14.99~~ ✓Prime
Only 12 left in stock - order soon.

More Buying Choices
$1.75 new (28 offers)

Show only IKEA items

★★★★☆ ▾ 129

La Crosse Technology WT-8002U Digital Wall Clock
by La Crosse Technology

$18.78 ~~$46.95~~ ✓Prime
Get it by Thursday, Aug 6

More Buying Choices
$18.75 new (7 offers)
$14.95 used (8 offers)

★★★★☆ ▾ 718

See Color Options

12" Vintage France Paris Colourful French Country Tuscan Style Paris Wood Wall Clock
by Decho

$7.80

More Buying Choices
$7.80 new (37 offers)

FREE Shipping

★★★½☆ ▾ 292

See Size Options

Geneva 14-Inch Plastic Decor
by Geneva

$26.06

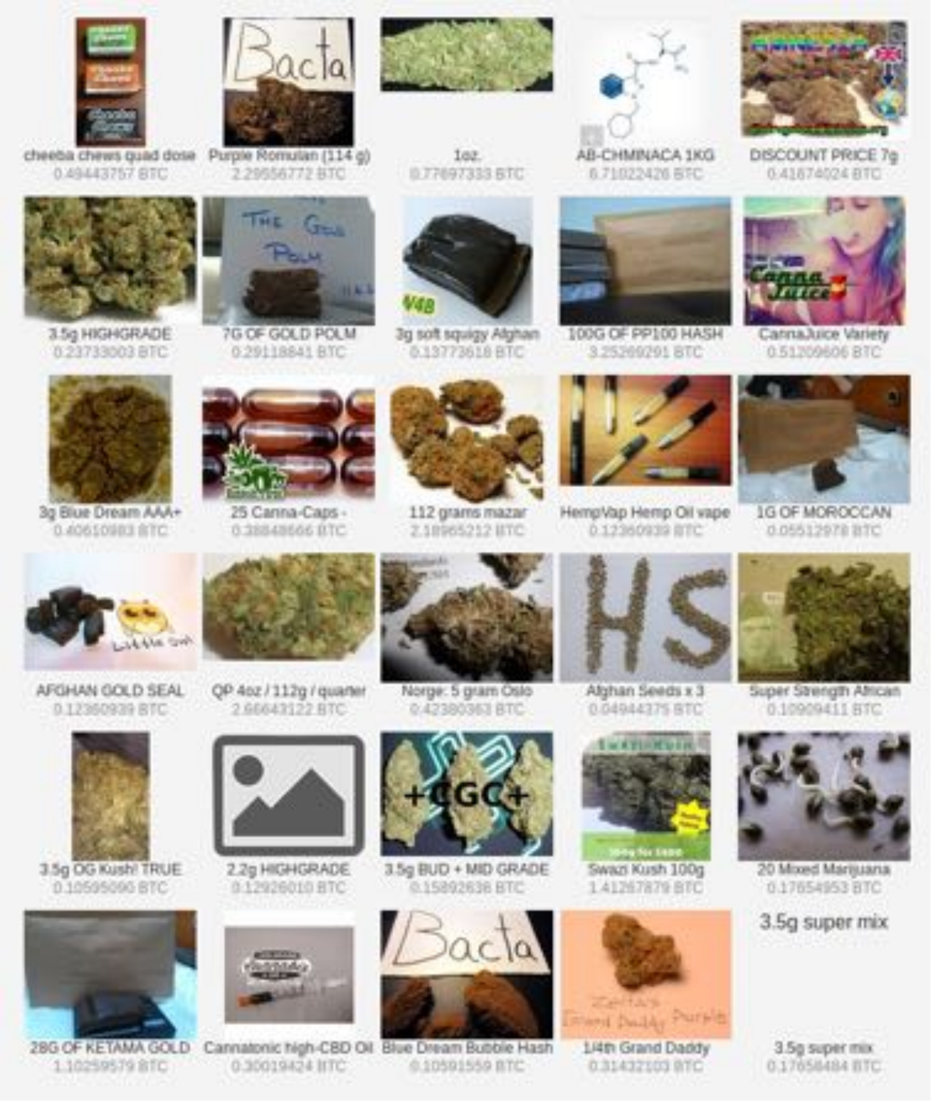
More Buying Choices
$17.29 new (9 offers)

FREE Shipping

★★★★½ ▾ 261

Conventional Illicit Commerce

Illicit Internet Commerce



cheeba chews quad dose 0.49443757 BTC	Purple Romulan (114 g) 2.29556772 BTC	1oz. 0.77697333 BTC	AB-CHMINACA 1KG 6.71022426 BTC	DISCOUNT PRICE 7g 0.41674024 BTC
3.5g HIGHGRADE 0.23733003 BTC	7G OF GOLD POLM 0.29118841 BTC	3g soft squigy Afghan 0.13773618 BTC	100G OF PP100 HASH 3.25269291 BTC	CannaJuice Variety 0.51209606 BTC
3g Blue Dream AAA+ 0.40610983 BTC	25 Canna-Caps - 0.38848666 BTC	112 grams mazar 2.18965212 BTC	HempVap Hemp Oil vape 0.12360939 BTC	1G OF MOROCCAN 0.05512978 BTC
AFGHAN GOLD SEAL 0.12360939 BTC	QP 4oz / 112g / quarter 2.66643122 BTC	Norge: 5 gram Oslo 0.42380363 BTC	Afghan Seeds x 3 0.04944375 BTC	Super Strength African 0.10909411 BTC
3.5g OG Kush! TRUE 0.10595090 BTC	2.2g HIGHGRADE 0.12926010 BTC	3.5g BUD + MID GRADE 0.15892638 BTC	Swazi Kush 100g 1.41267879 BTC	20 Mixed Marijuana 0.17654953 BTC
28G OF KETAMA GOLD 1.10259579 BTC	Cannatonic high-CBD Oil 0.30019424 BTC	Blue Dream Bubble Hash 0.10591559 BTC	1/4th Grand Daddy 0.31432103 BTC	3.5g super mix 3.5g super mix 0.17658484 BTC

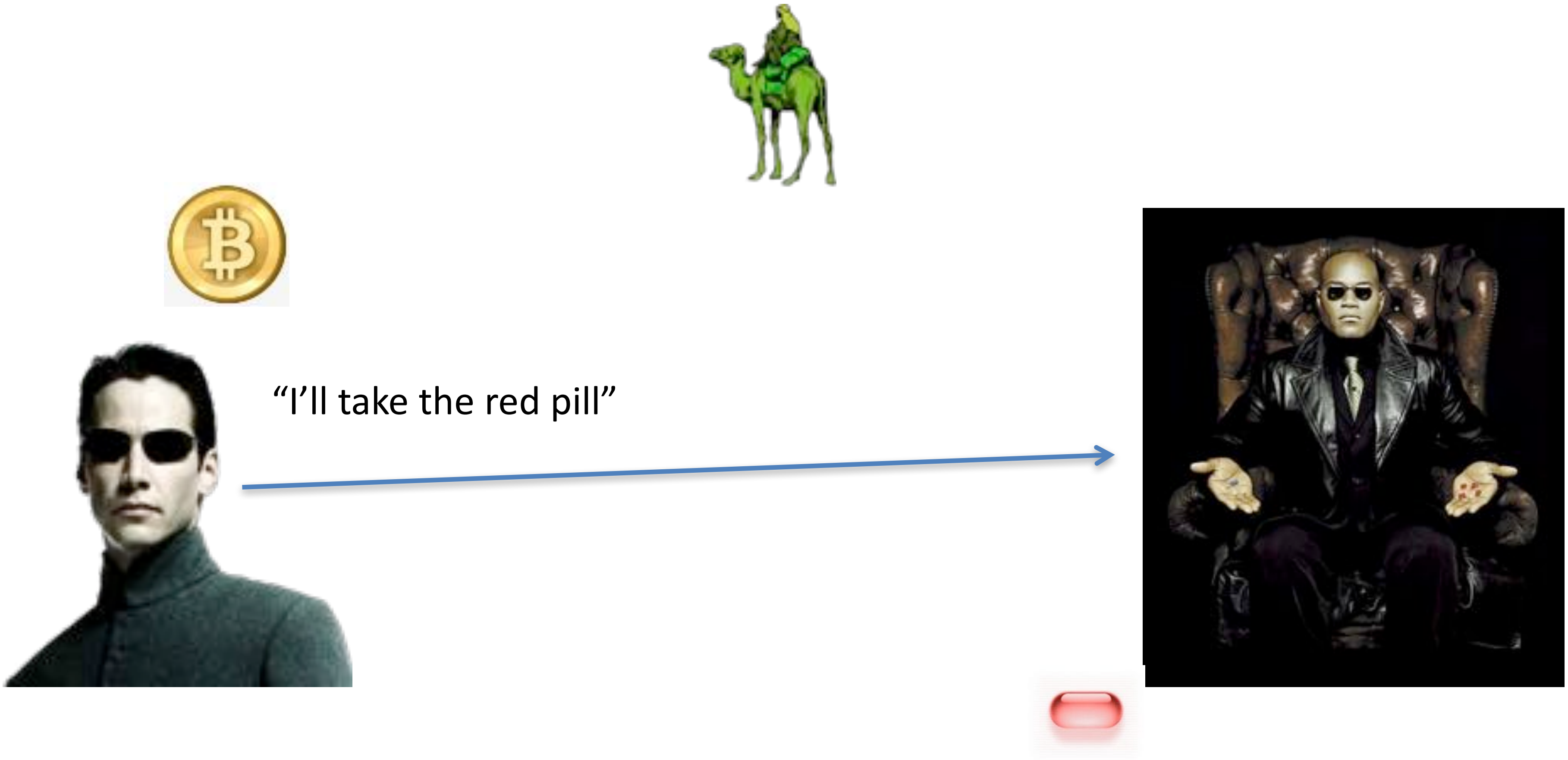
Anonymous Marketplaces

- Amazon.com of illegal goods
  - Drugs, CC’s & Fake IDs, Weapons, etc.
  - No Child Porn
- Safety
- Convenience
- Variety
- Accountability
- Competition

Anonymous Marketplace Technology

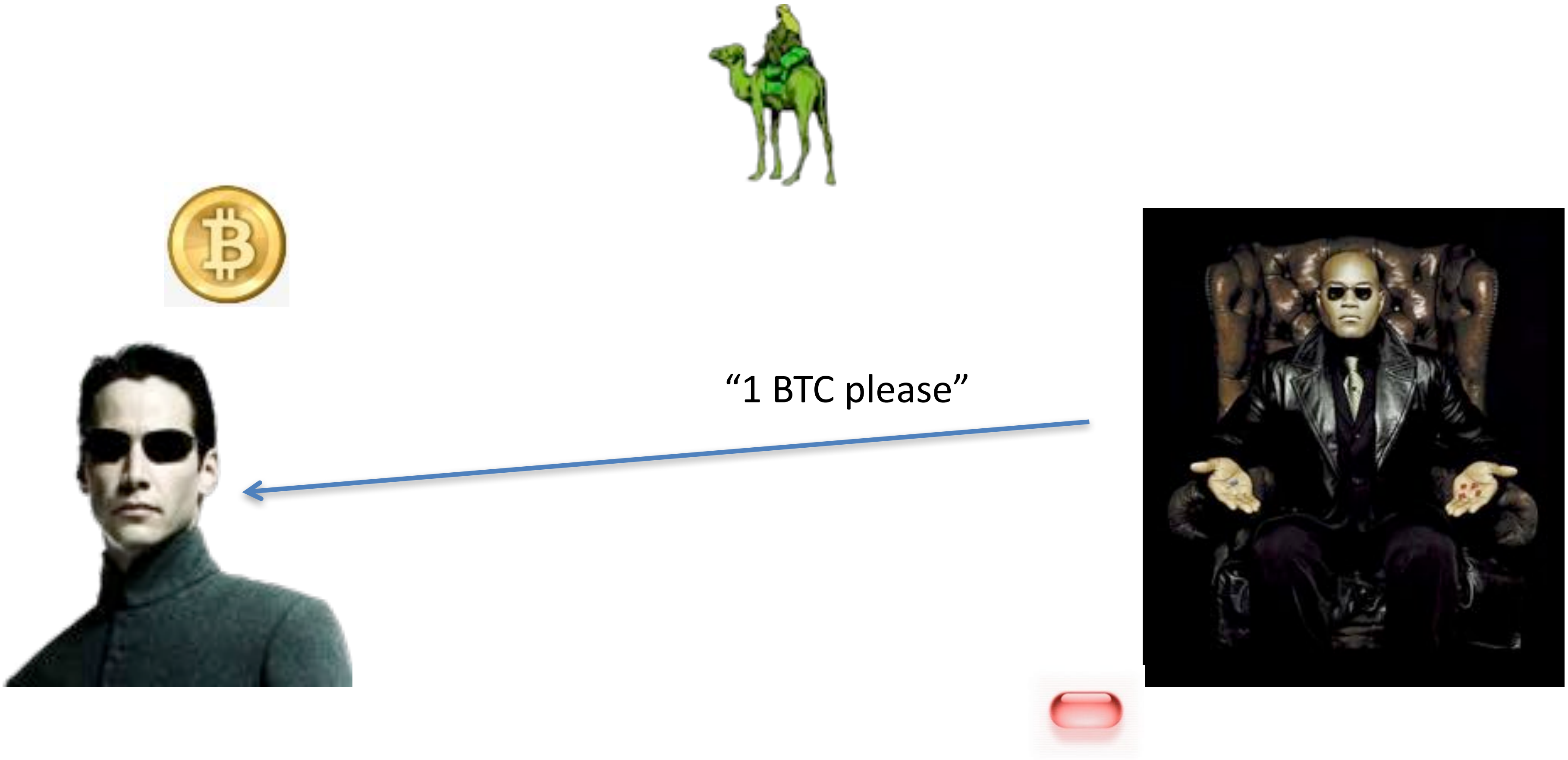
- Hidden Website (Tor Hidden Service, I2P)
  - Customers
    - No cost of creation
    - No information needed
  - Vendors
    - Vendor bonds required
    - Often invite only
    - Public feedback history
- Payments (Bitcoin)
  - Marketplaces often act as escrow agent
  - Escrow sometimes acts as a mixing service
- Hidden Messages(PGP)

Market Transactions



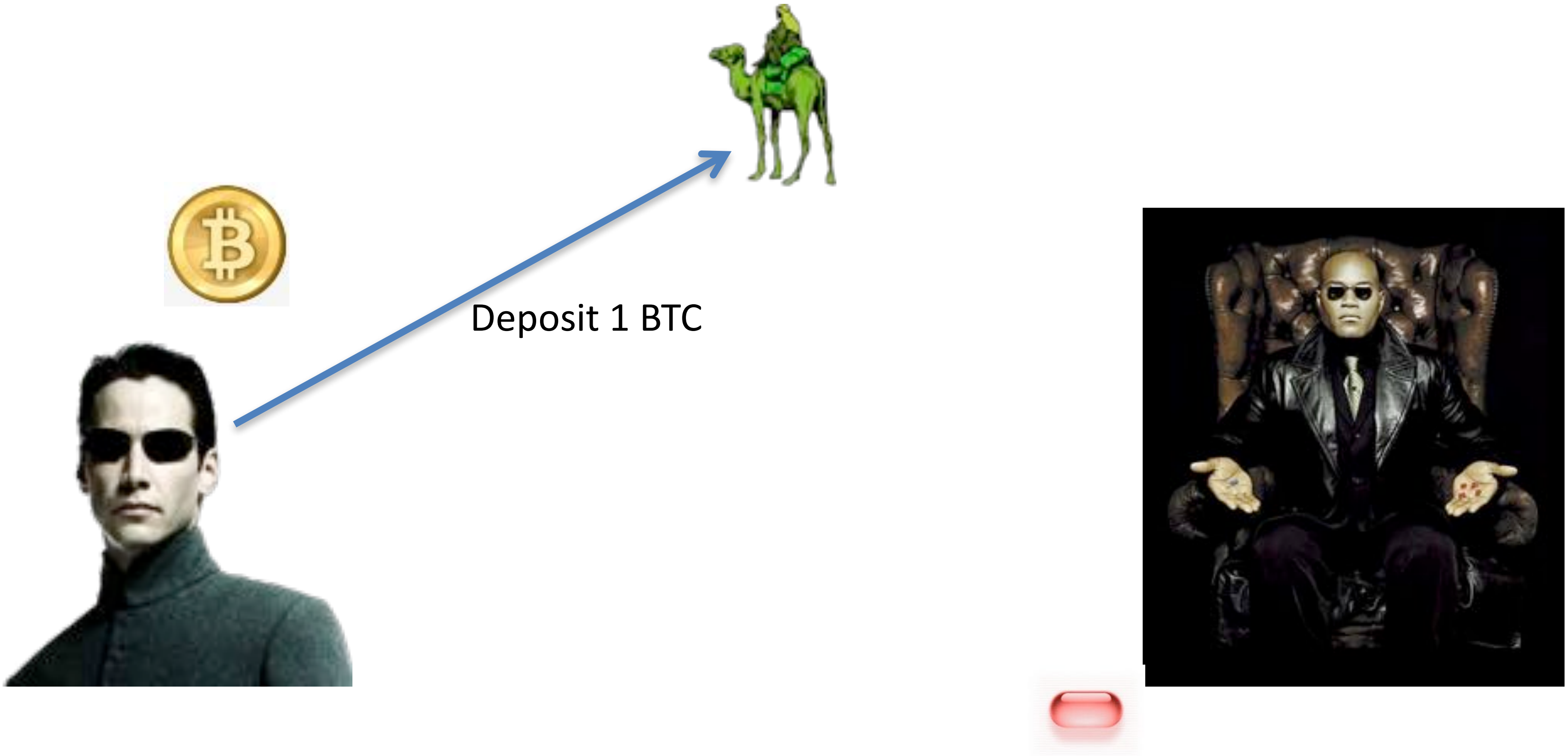
“I’ll take the red pill”

Market Transactions



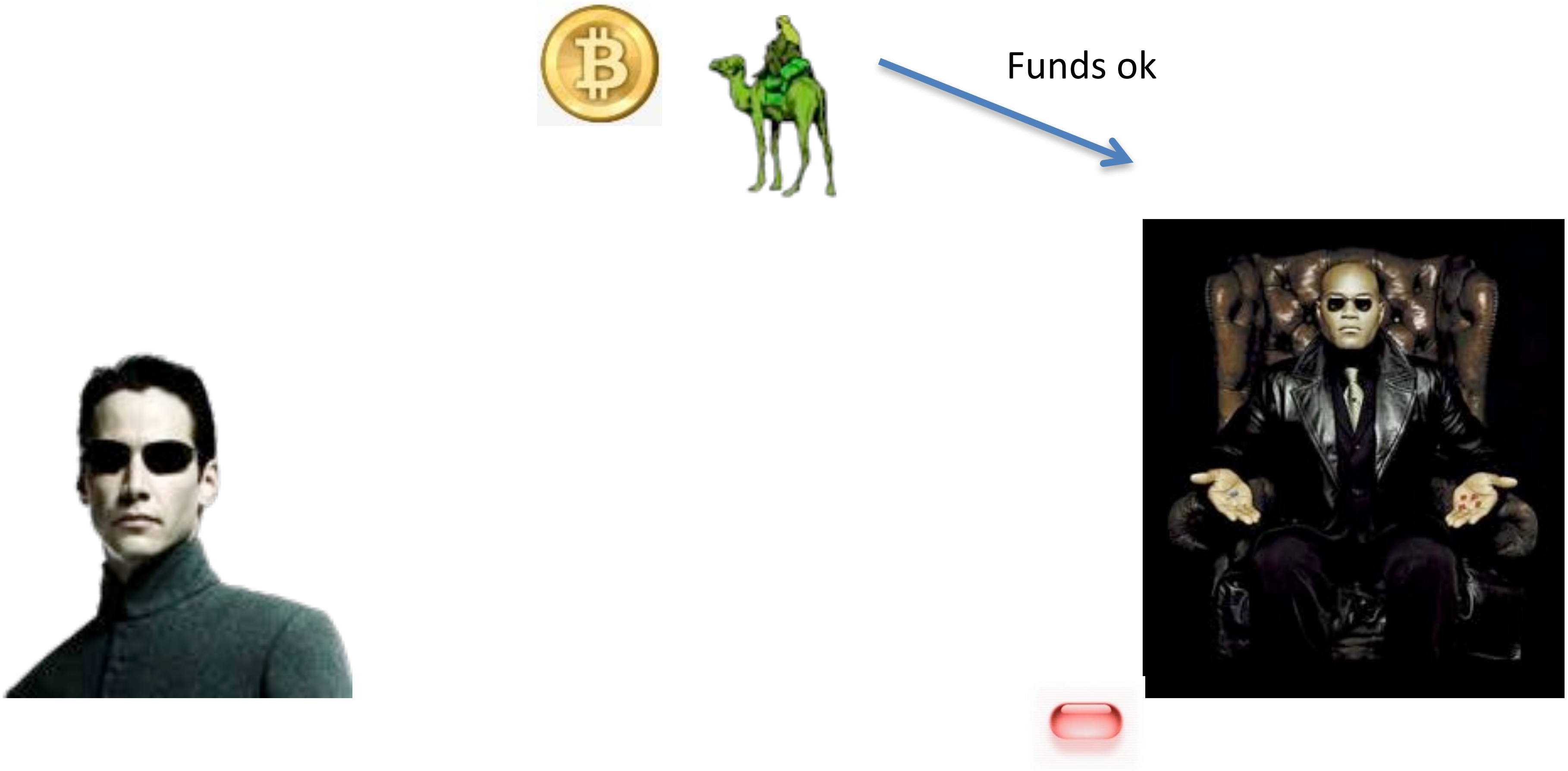
“1 BTC please”

Market Transactions



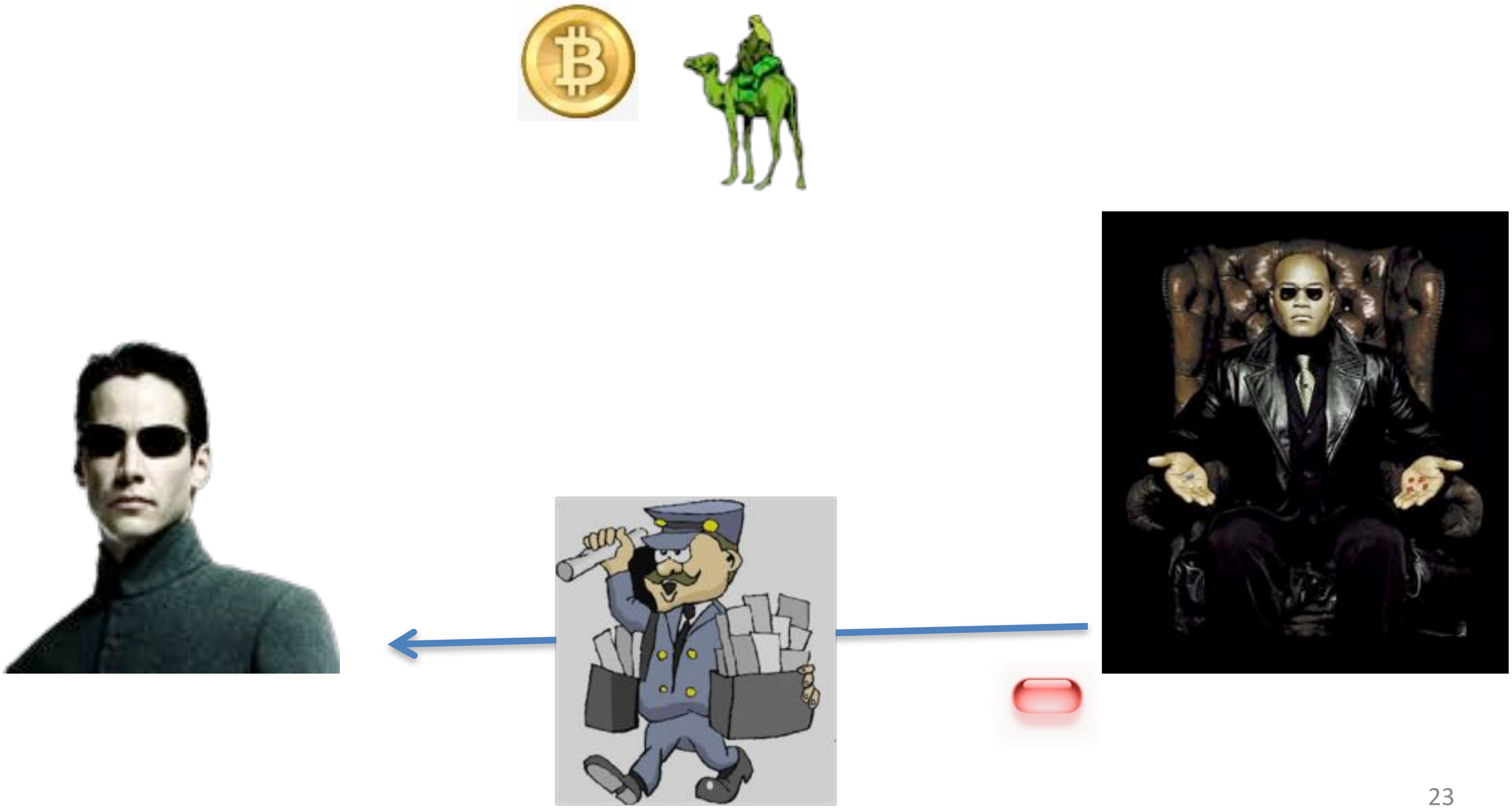
Deposit 1 BTC

Market Transactions

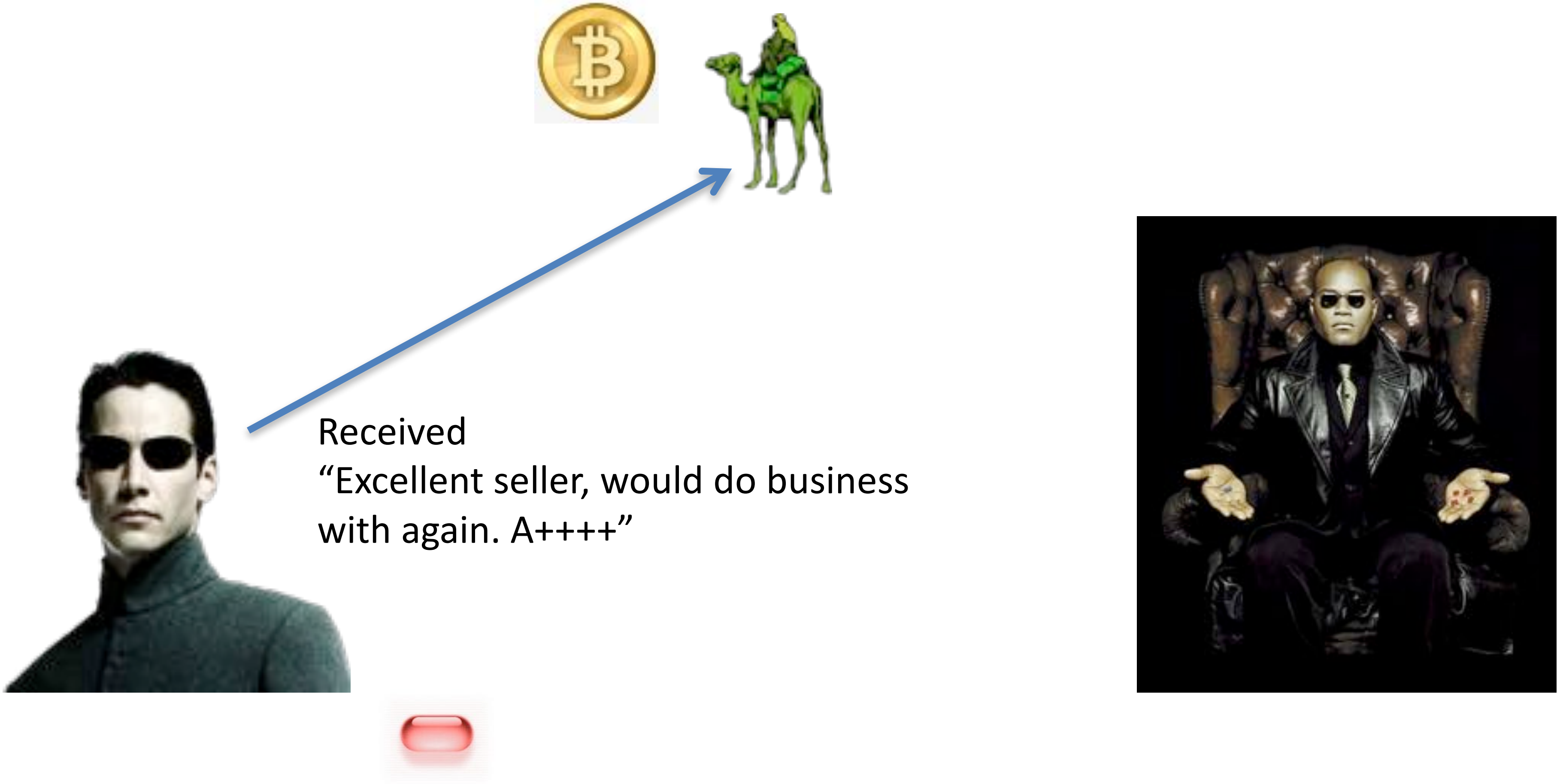


Funds ok

Market Transactions

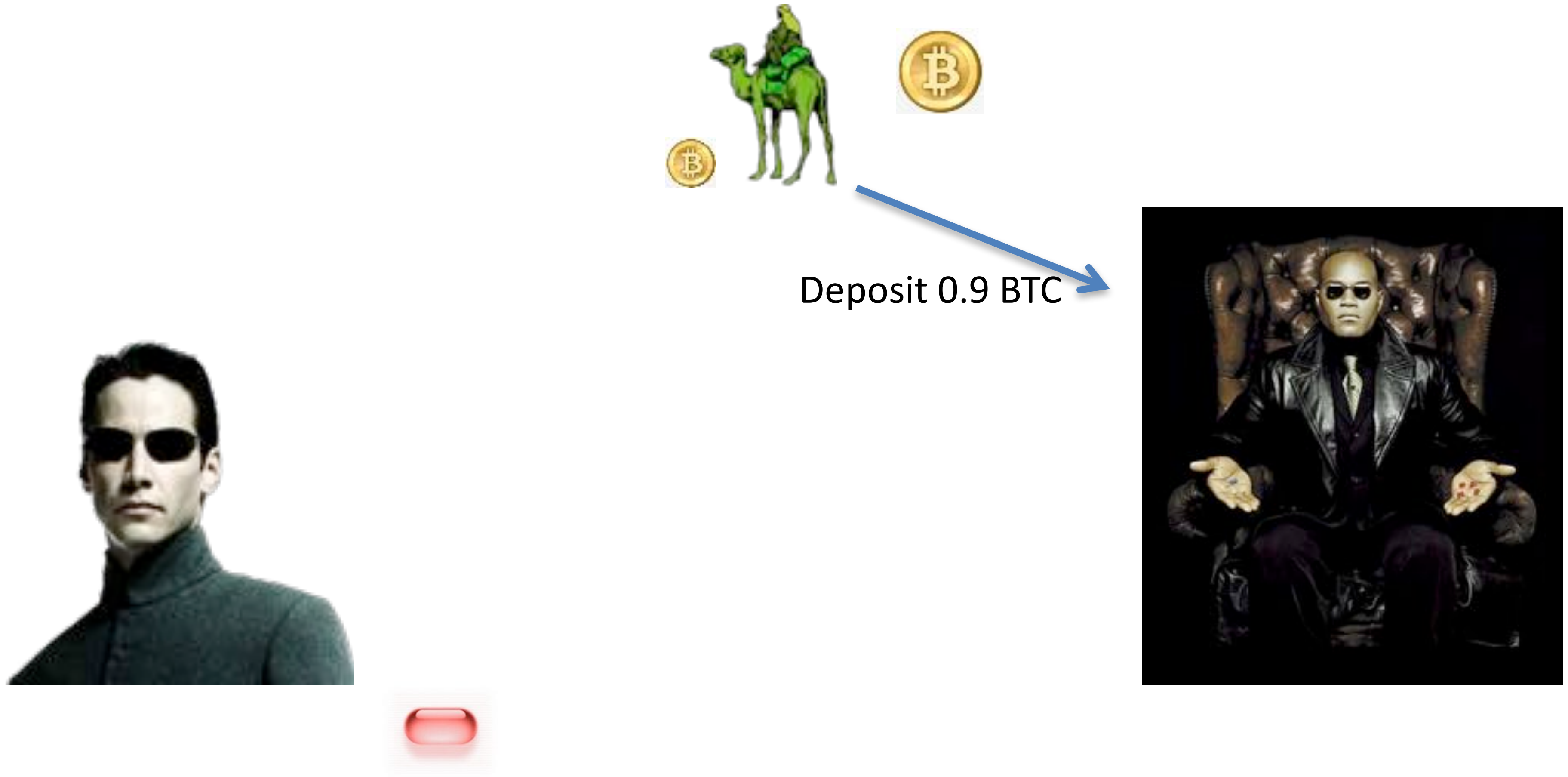


Market Transactions



Received
“Excellent seller, would do business with again. A++++”

Market Transactions

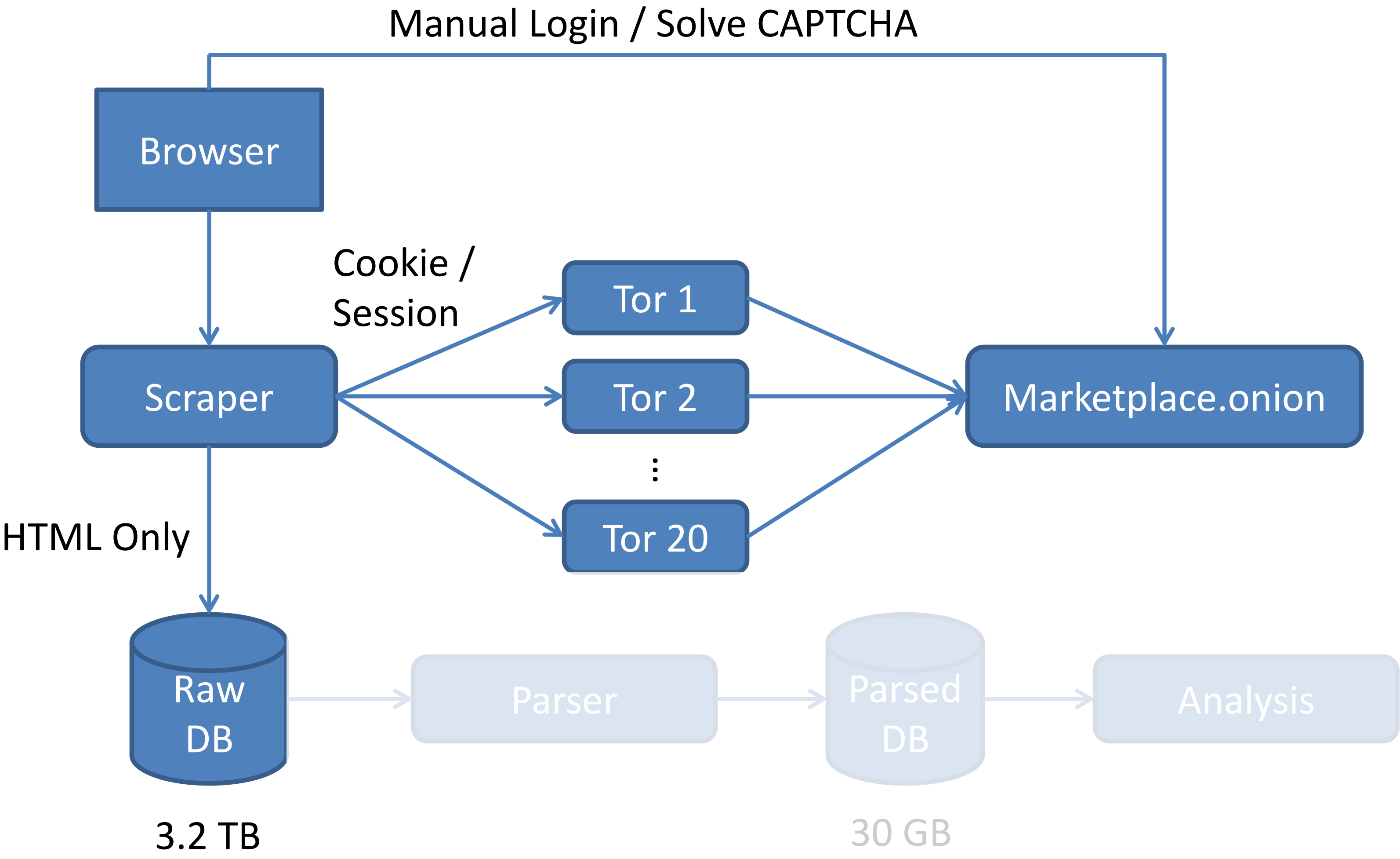


Deposit 0.9 BTC

Questions

- How much is being sold?
- What is being sold?
- How many vendors are relevant?
- What do vendors sell?

Measurement Platform Overview



Measurements

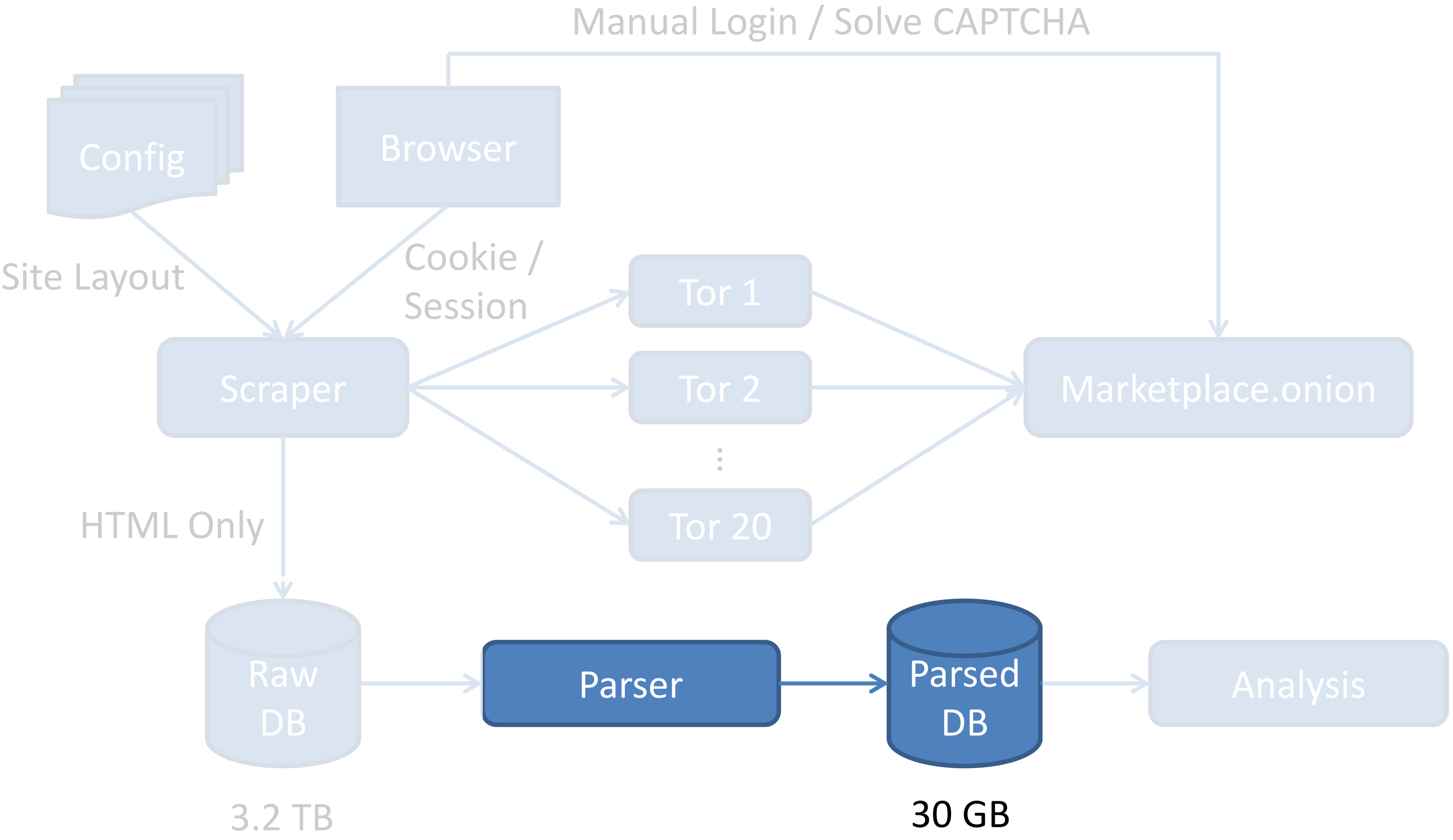
- **Stealth**
  - Indistinguishable from real user
  - Random delays, scrape slowly
  - Popular User Agent
  - Browse website “normally”
- **Complete and instantaneous**
  - Dynamic marketplace, moving target
  - Scrape quickly
  - Site availability as low as 70%

Measurements

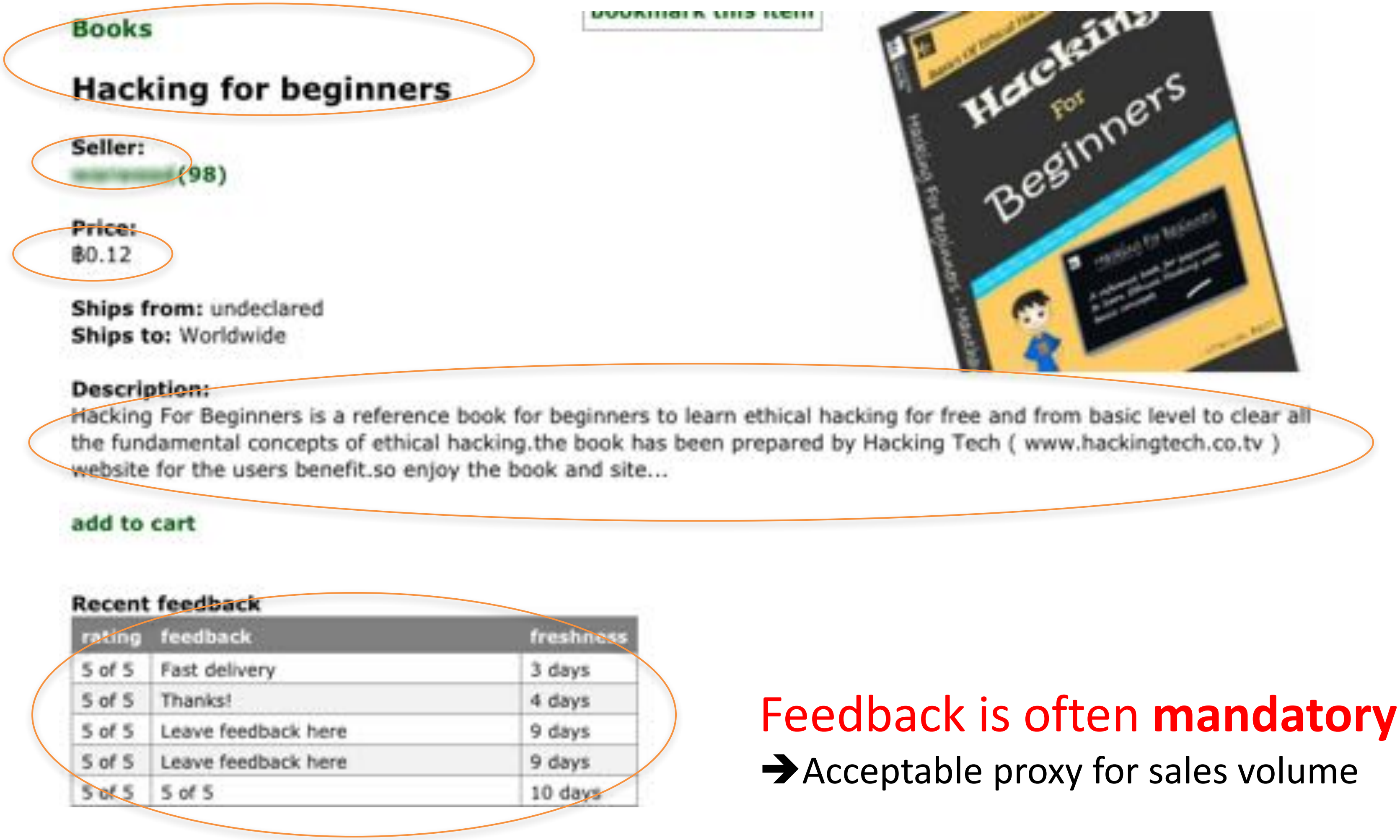
- **Anti-Scraping Encountered**
  - Rate Limits
  - Cookie Timeout
  - User Account Suspension

- **Totals**
  - 35 Marketplaces 1,908 scrapes total – 3.2 TB
  - 27 – 331,691 pages per scrape
  - 11/22/11 – present

Parsing


Manual Login / Solve CAPTCHA
Config
Browser
Site Layout
Cookie / Session
Tor 1
Tor 2
⋮
Tor 20
Scraper
Marketplace.onion
HTML Only
Raw DB
Parser
Parsed DB
Analysis
3.2 TB
30 GB

Silk Road Available Data



Books

Hacking for beginners

Seller:
[redacted](98)

Price:
฿0.12

Ships from: undeclared
Ships to: Worldwide

Description:
Hacking For Beginners is a reference book for beginners to learn ethical hacking for free and from basic level to clear all the fundamental concepts of ethical hacking.the book has been prepared by Hacking Tech (www.hackingtech.co.tv) website for the users benefit.so enjoy the book and site...

add to cart

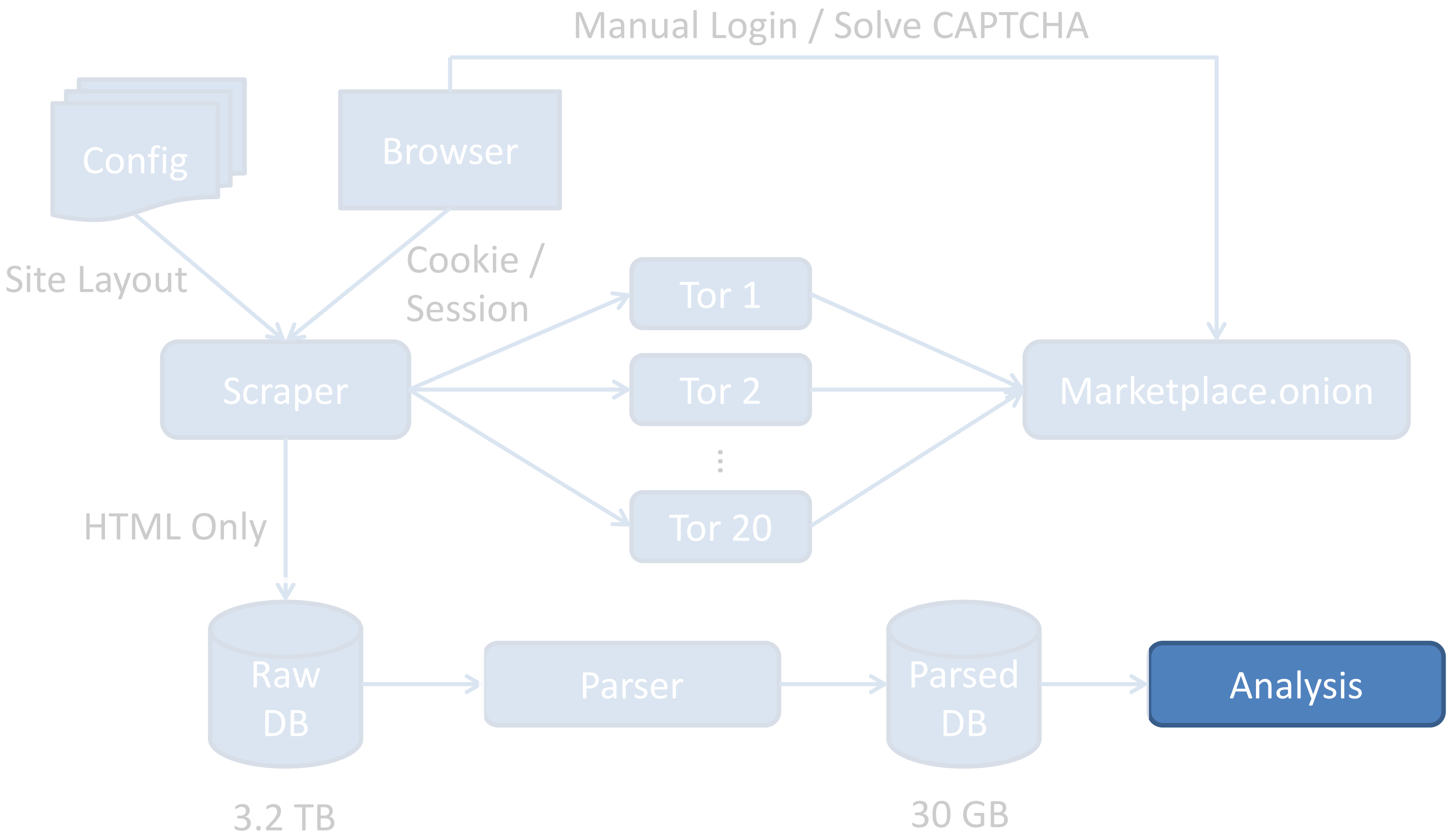
Recent feedback

rating	feedback	freshness
5 of 5	Fast delivery	3 days
5 of 5	Thanks!	4 days
5 of 5	Leave feedback here	9 days
5 of 5	Leave feedback here	9 days
5 of 5	5 of 5	10 days

Feedback is often **mandatory!**

➔Acceptable proxy for sales volume

Analysis


Manual Login / Solve CAPTCHA
Config
Browser
Site Layout
Cookie / Session
Tor 1
Tor 2
⋮
Tor 20
Scraper
Marketplace.onion
HTML Only
Raw DB
Parser
Parsed DB
Analysis
3.2 TB
30 GB

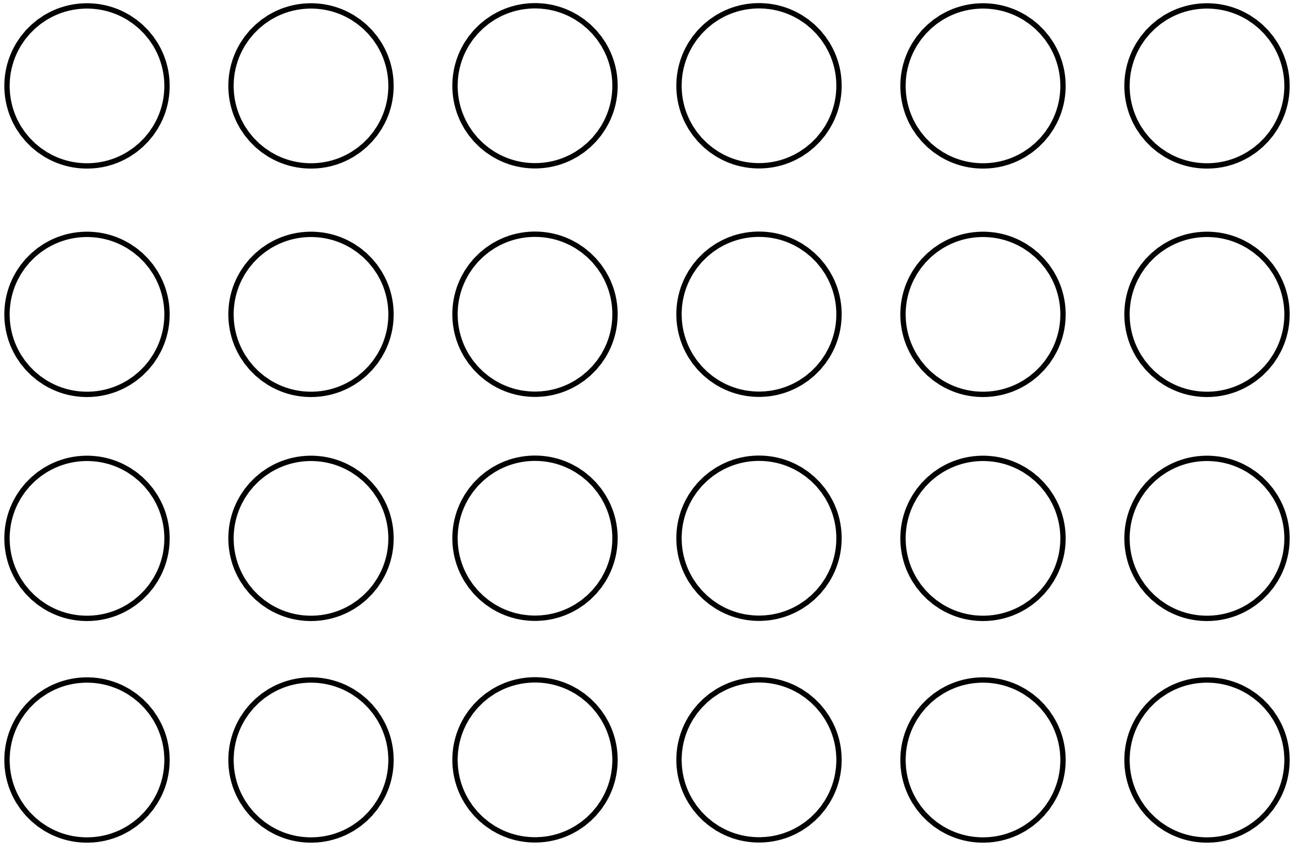
Data Completeness

- **How complete is the data?**
  - Unreliable dynamic marketplaces that take days to scrape
  - Empirical observations - lower bound

- **Idea:** Estimate population via mark and recapture
  - Schnabel Estimator allows multiple recapture

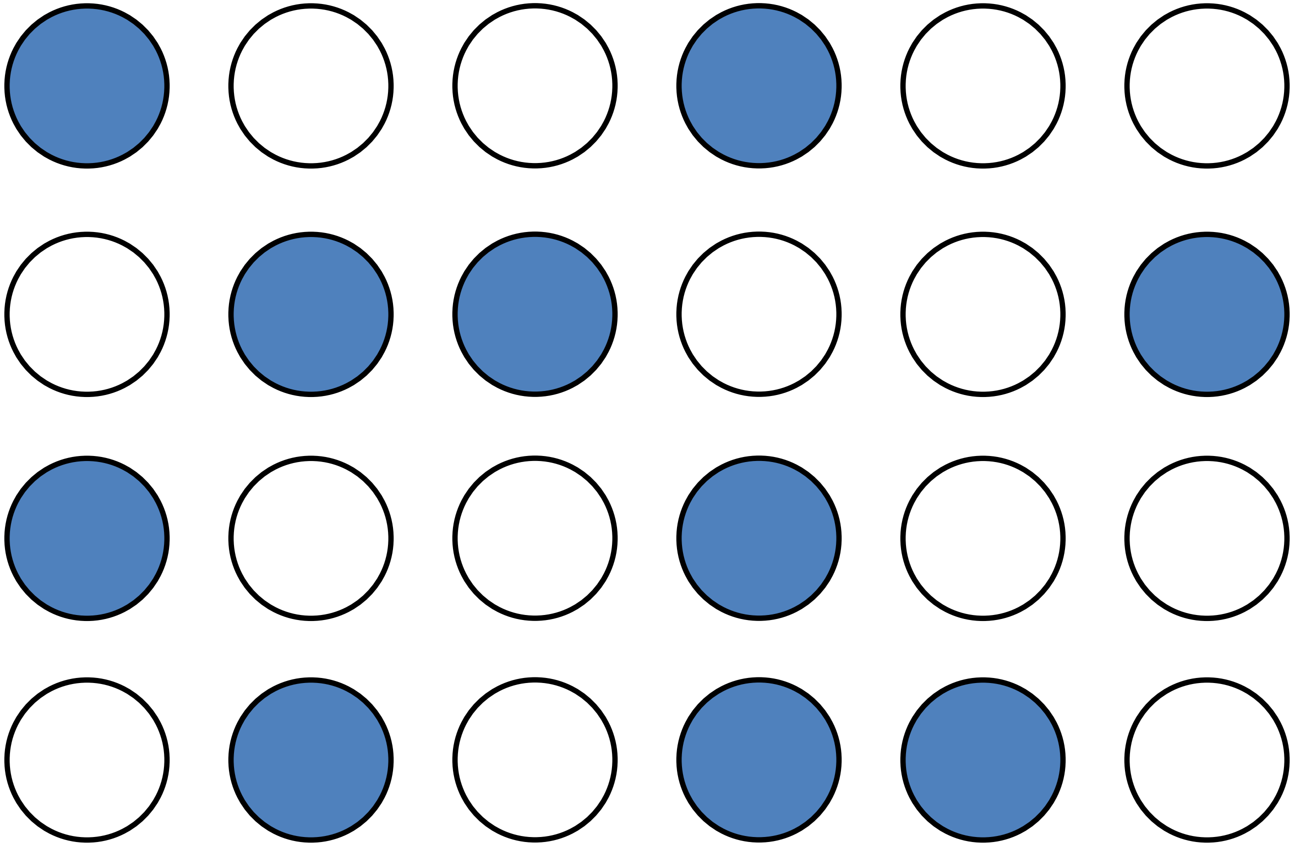
Mark and Recapture

Population Size = 24



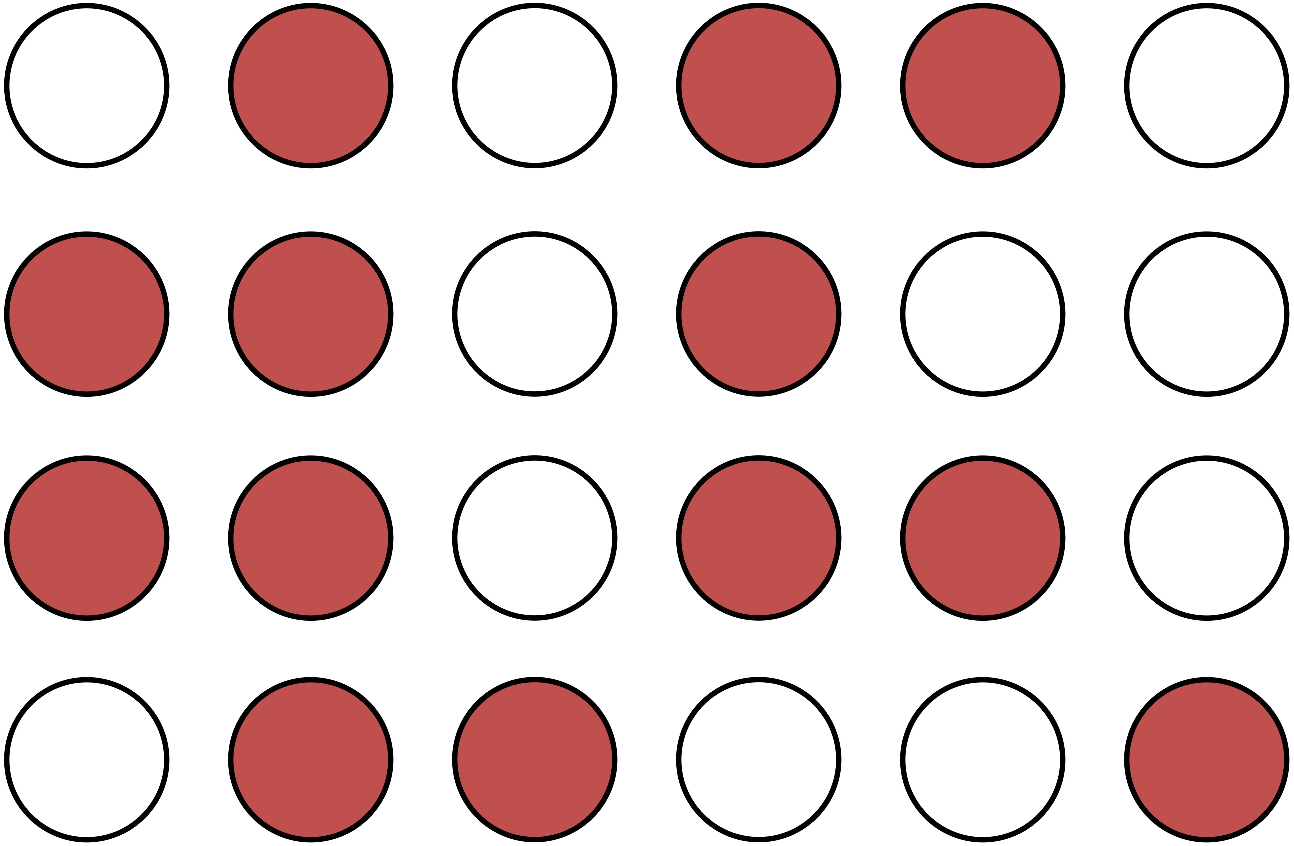
Mark and Recapture

Sample Size = 10



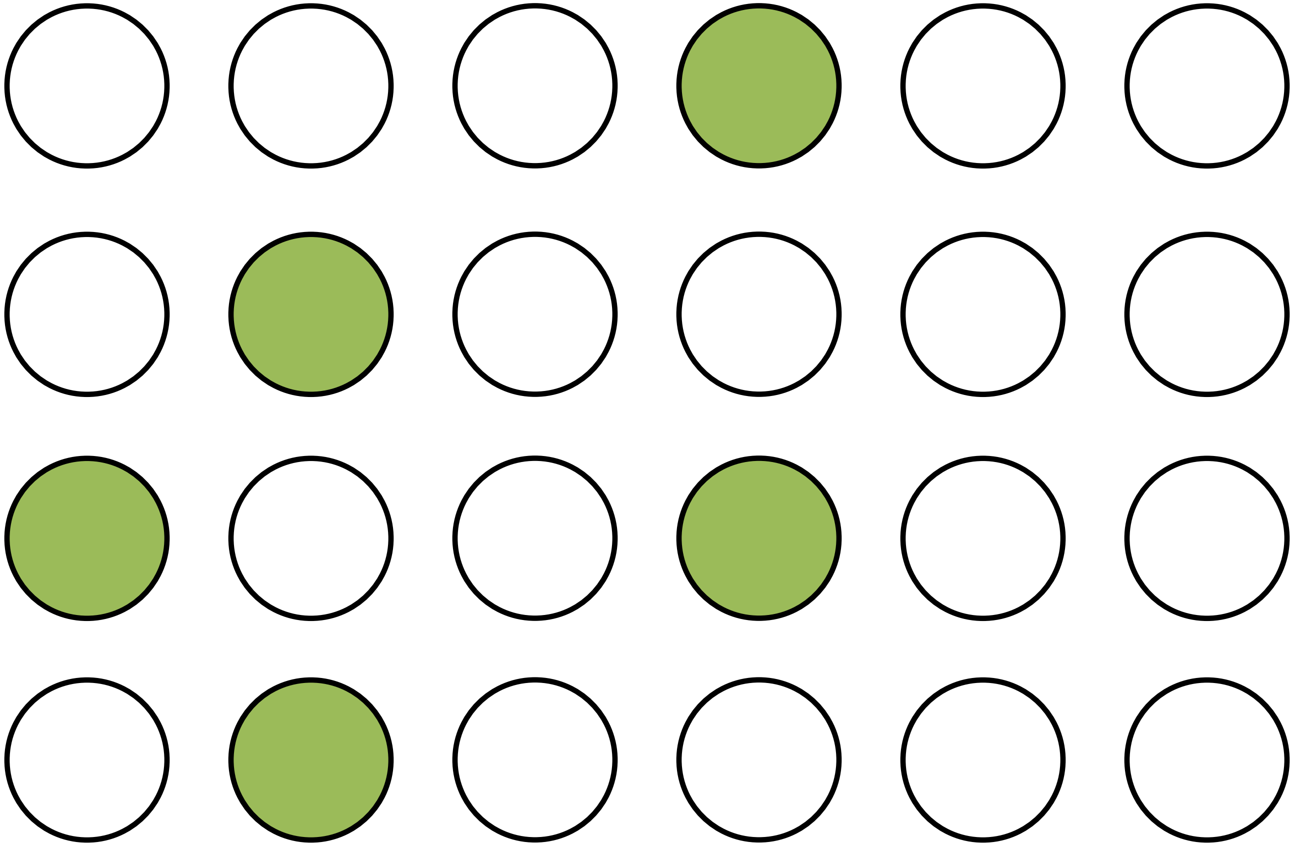
Mark and Recapture

Sample Size = 13

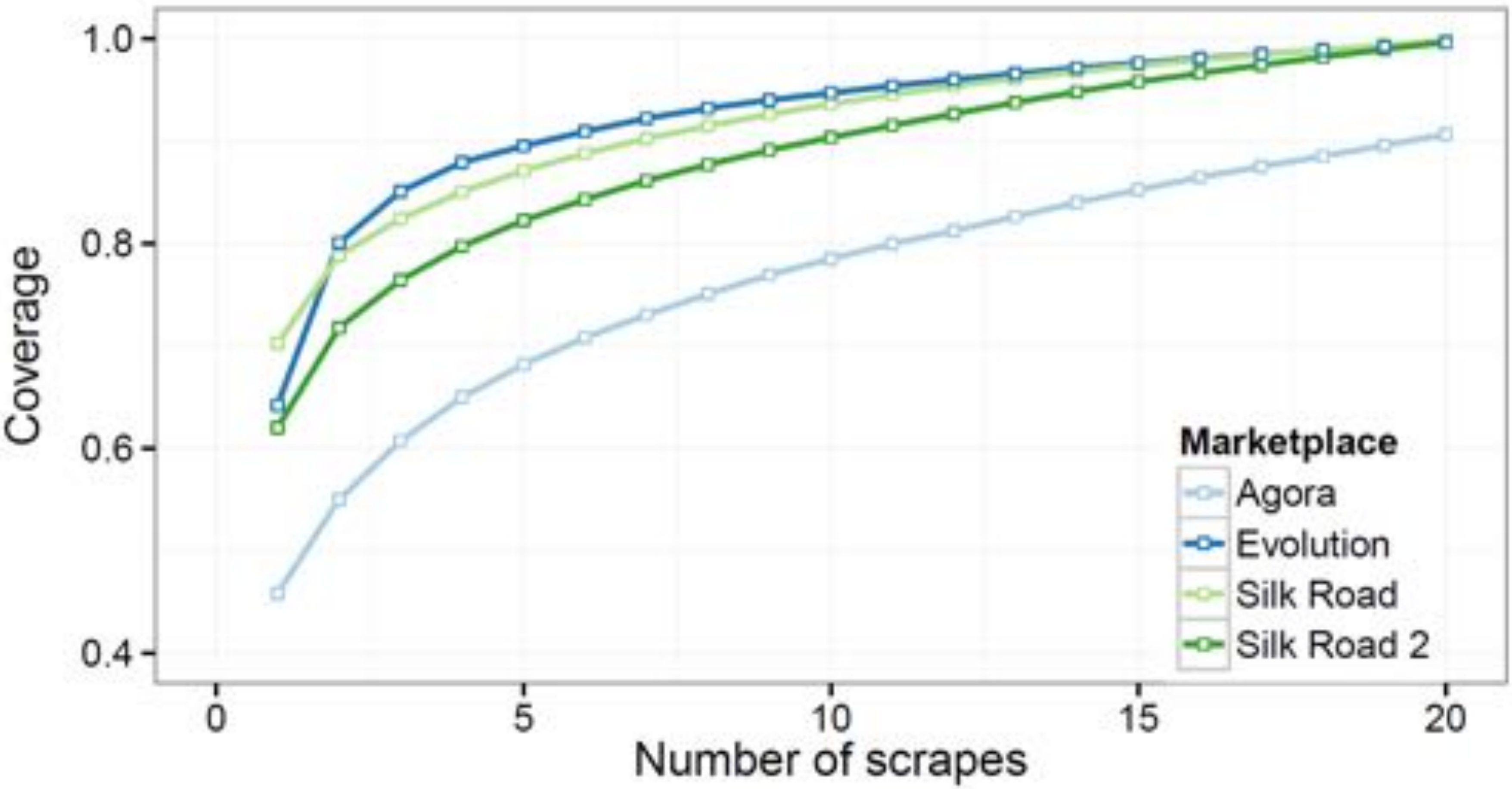


Mark and Recapture

Overlap = 5, Population Estimate = 26



Data Completeness



Analysis

- **Assumption:** Each feedback corresponds to precisely one transaction
  - Anonymity requires strictly enforced feedback system to establish reputation
  - Possible on many marketplaces to purchase several quantities of item and leave 1 feedback, conservative estimate

Alternative Transaction Proxies

- **Counting # Item Listings**
  - Very efficient and convenient
  - Assumes that there exists some stable ratio between transaction volume and # listings
  - Daily $\frac{volume}{\#\ Listings}$ for The Evolution Marketplace in July 2014 and September 2014 differ by factor of 4

Uniqueness

- **Problem:**
  - 100s of observations of same feedback
  - Double counting leads to over-estimations
  - Feedback may be updated, deleted

- **Solution:**
  - Automatically detect updated feedbacks
    - Only keep most recent version
  - Hash {timestamp, title, vendor, message, rating}

Holding Prices

- Feedbacks are useful to vendors but are destroyed when the listing is removed
- Vendors raise listing prices prohibitively high

$0.02 -> $1,000.00

$1,100.00 -> $1,000,000.00

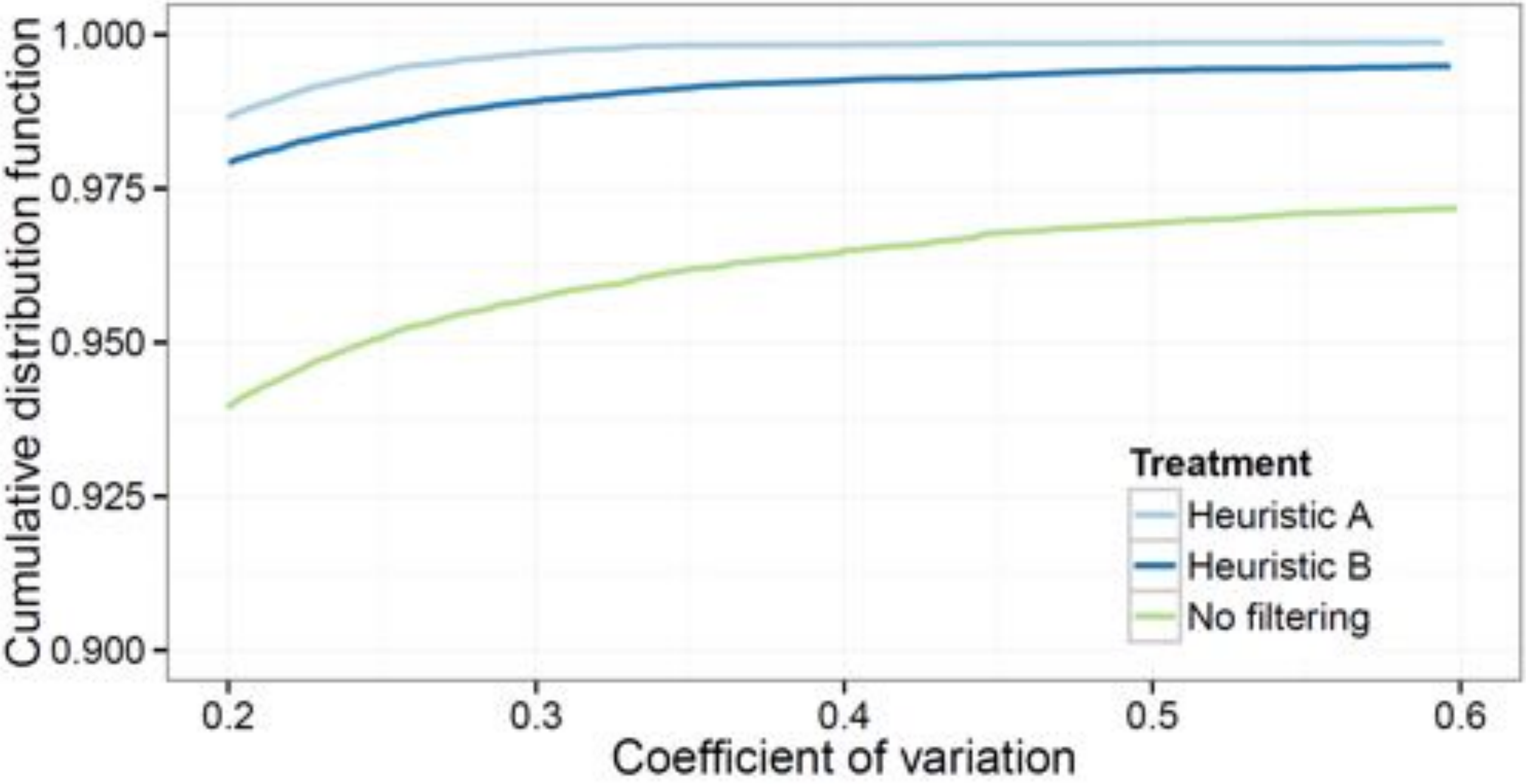
- Need to look at historical price for item

Holding Prices

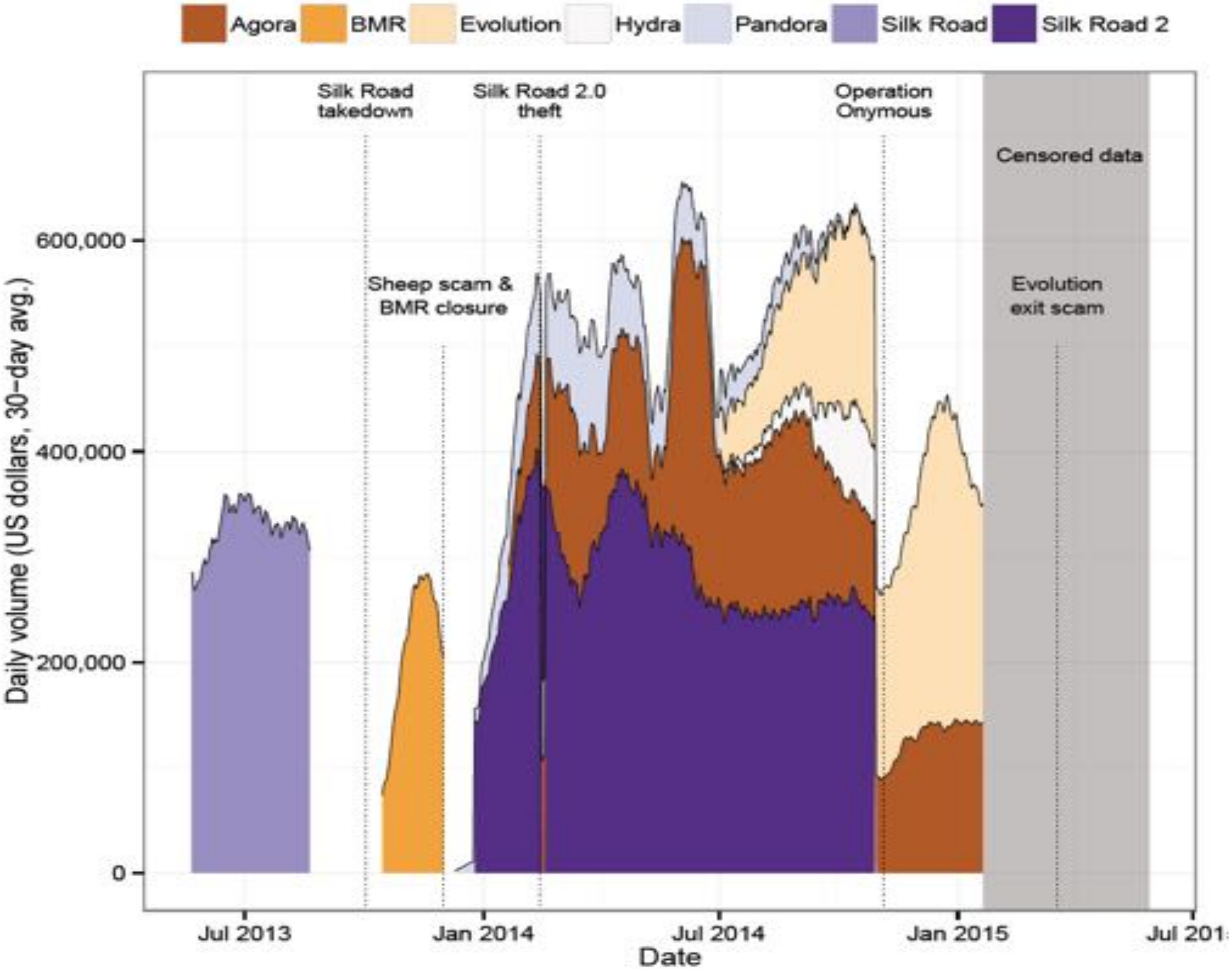
- Heuristic A:
  - Remove all free things
  - Remove all things > $100,000
  - Calculate median of remaining prices
  - Remove everything greater than 5x median
  - Remove things less than 25% of median
- Heuristic B:
  - Remove all things > $100,000
  - Remove upper quartile
  - Remove everything greater than 100x cheapest non-zero price

- Evaluation
  - Coefficient of Variation

Holding Prices CDF

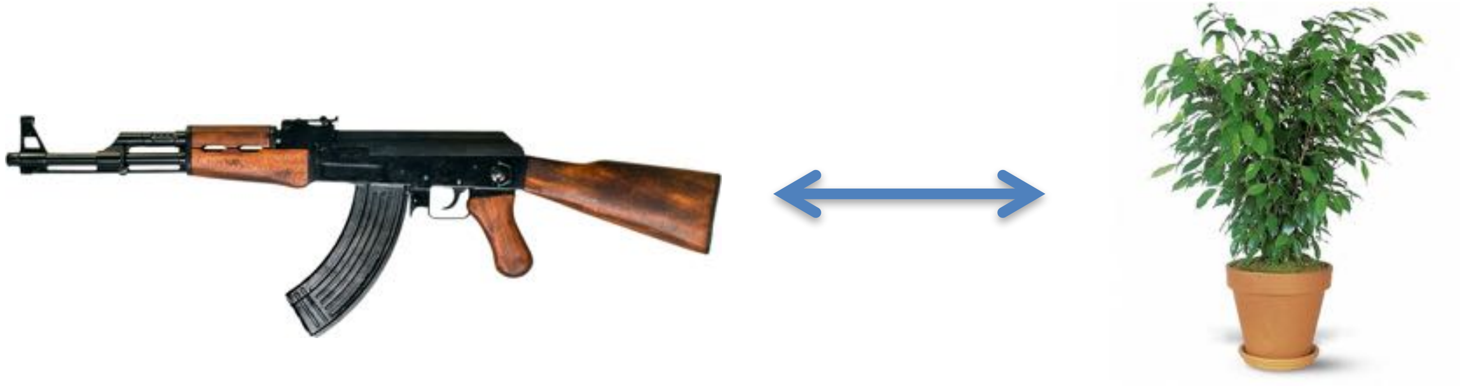


Sales Volume



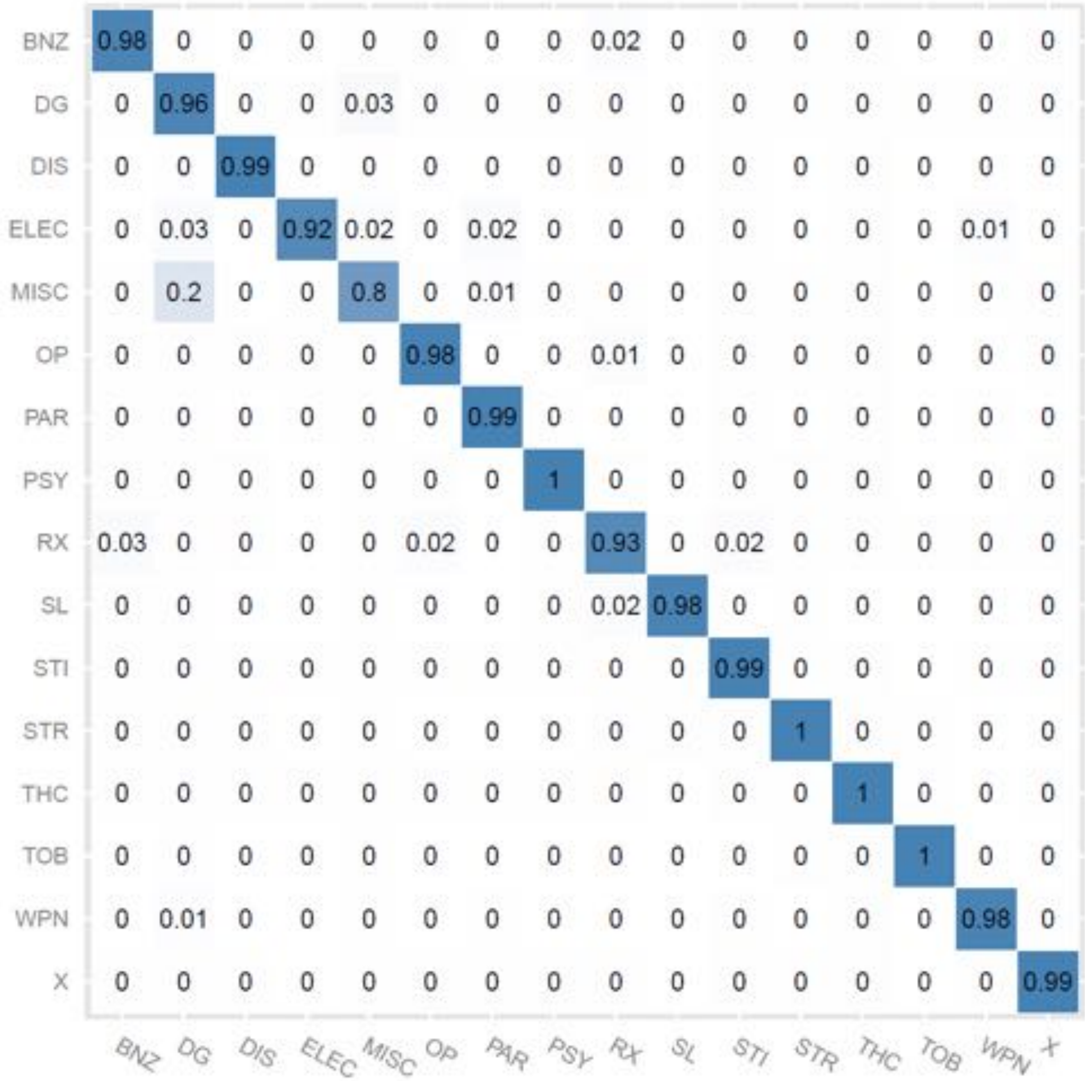
Product Categories

- **What is being sold?**
  - Product labels are often unavailable or inaccurate



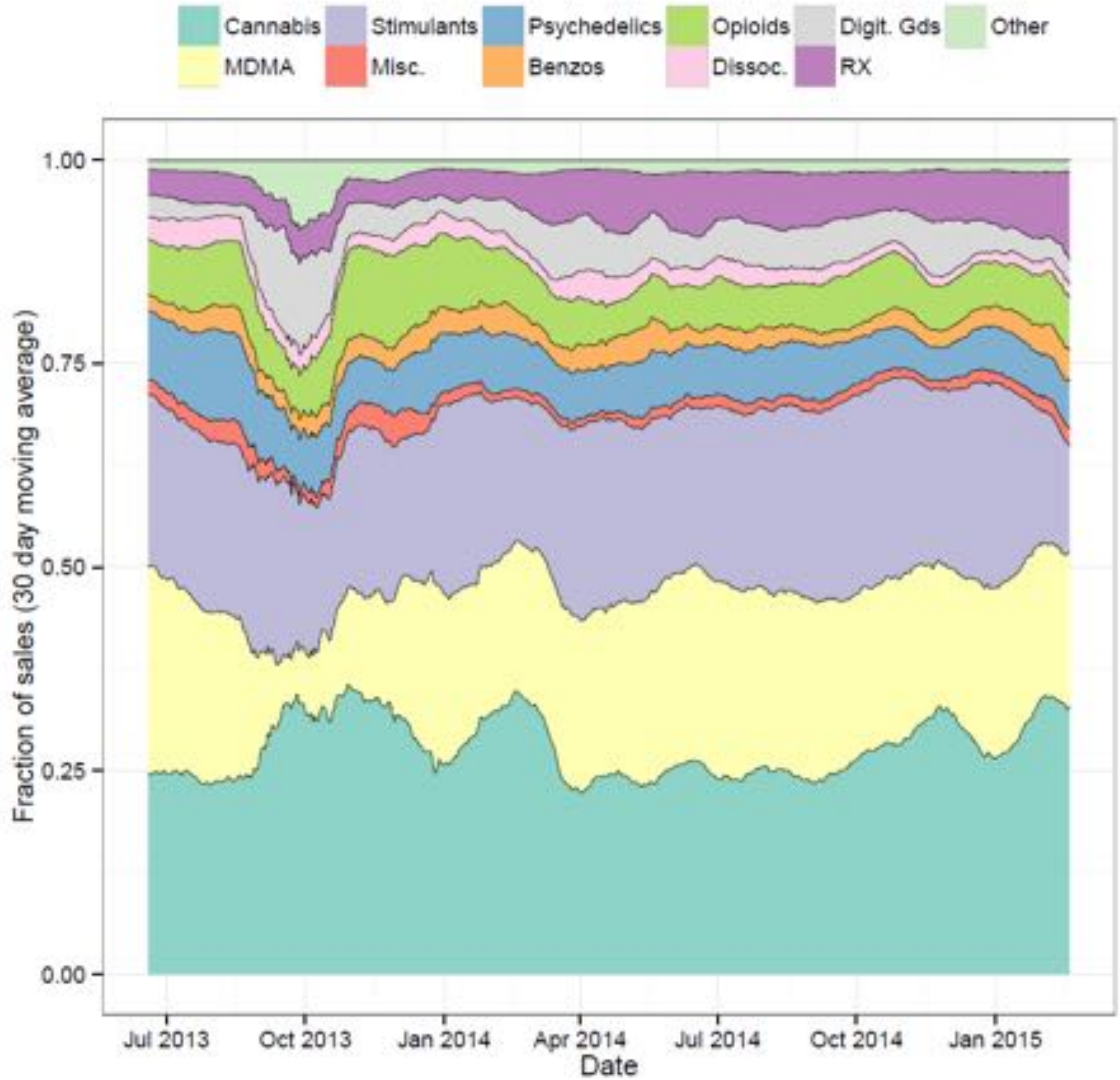
- Classifier trained from Agora and The Evolution Marketplace
  - Listing title and description concatenated and tfidf
  - 1,941,538 unique samples, 162,198 words tokenized
  - Predicts 16 class labels

Confusion Matrix

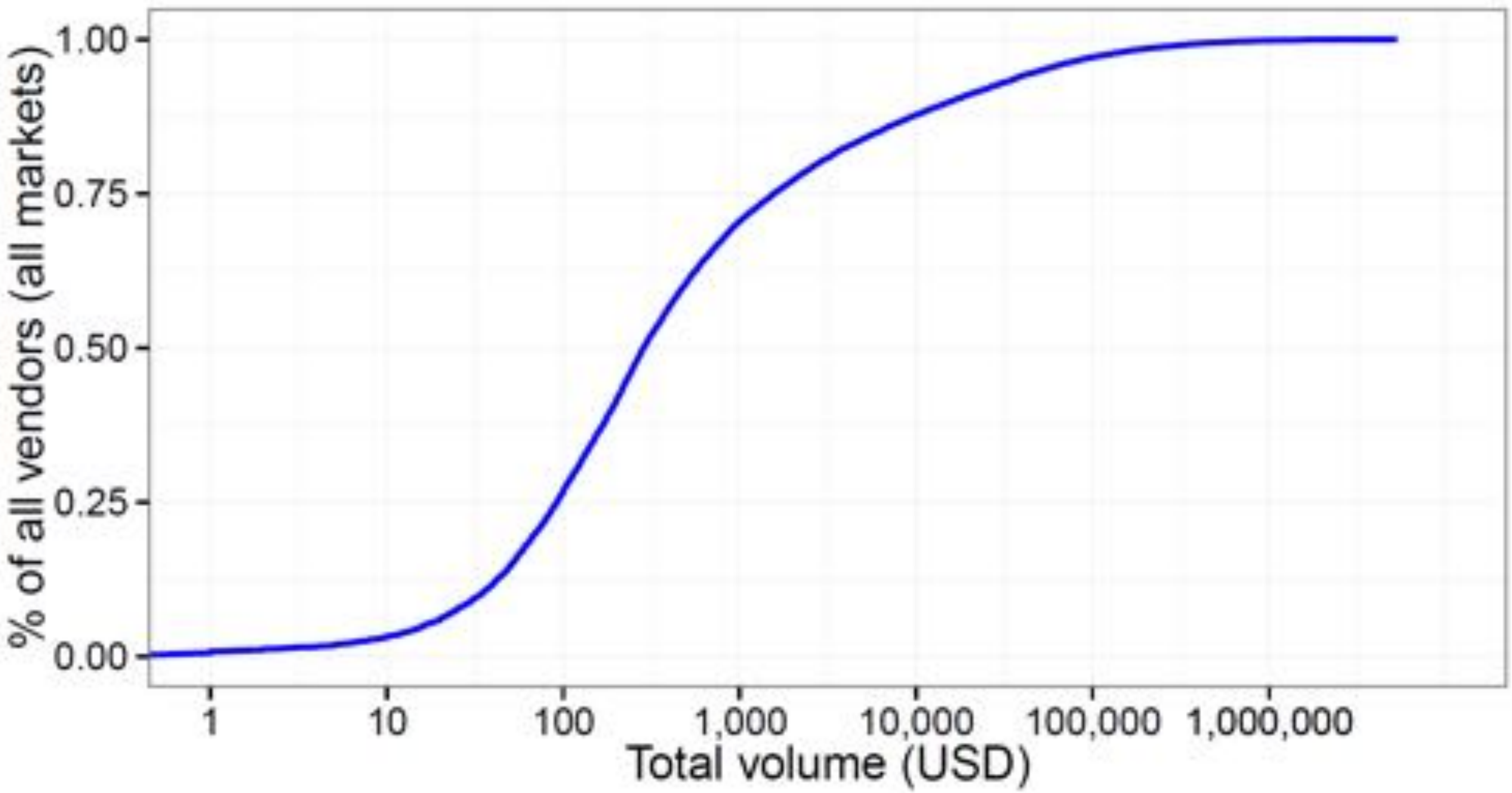


	BNZ	DG	DIS	ELEC	MISC	OP	PAR	PSY	RX	SL	STI	STR	THC	TOB	WPN	X
BNZ	0.98	0	0	0	0	0	0	0	0.02	0	0	0	0	0	0	0
DG	0	0.96	0	0	0.03	0	0	0	0	0	0	0	0	0	0	0
DIS	0	0	0.99	0	0	0	0	0	0	0	0	0	0	0	0	0
ELEC	0	0.03	0	0.92	0.02	0	0.02	0	0	0	0	0	0	0	0.01	0
MISC	0	0.2	0	0	0.8	0	0.01	0	0	0	0	0	0	0	0	0
OP	0	0	0	0	0	0.98	0	0	0.01	0	0	0	0	0	0	0
PAR	0	0	0	0	0	0	0.99	0	0	0	0	0	0	0	0	0
PSY	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
RX	0.03	0	0	0	0	0.02	0	0	0.93	0	0.02	0	0	0	0	0
SL	0	0	0	0	0	0	0	0	0.02	0.98	0	0	0	0	0	0
STI	0	0	0	0	0	0	0	0	0	0	0.99	0	0	0	0	0
STR	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
THC	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
TOB	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
WPN	0	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0.98	0
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.99

Item Sales Per Category



Vendor Volumes CDF

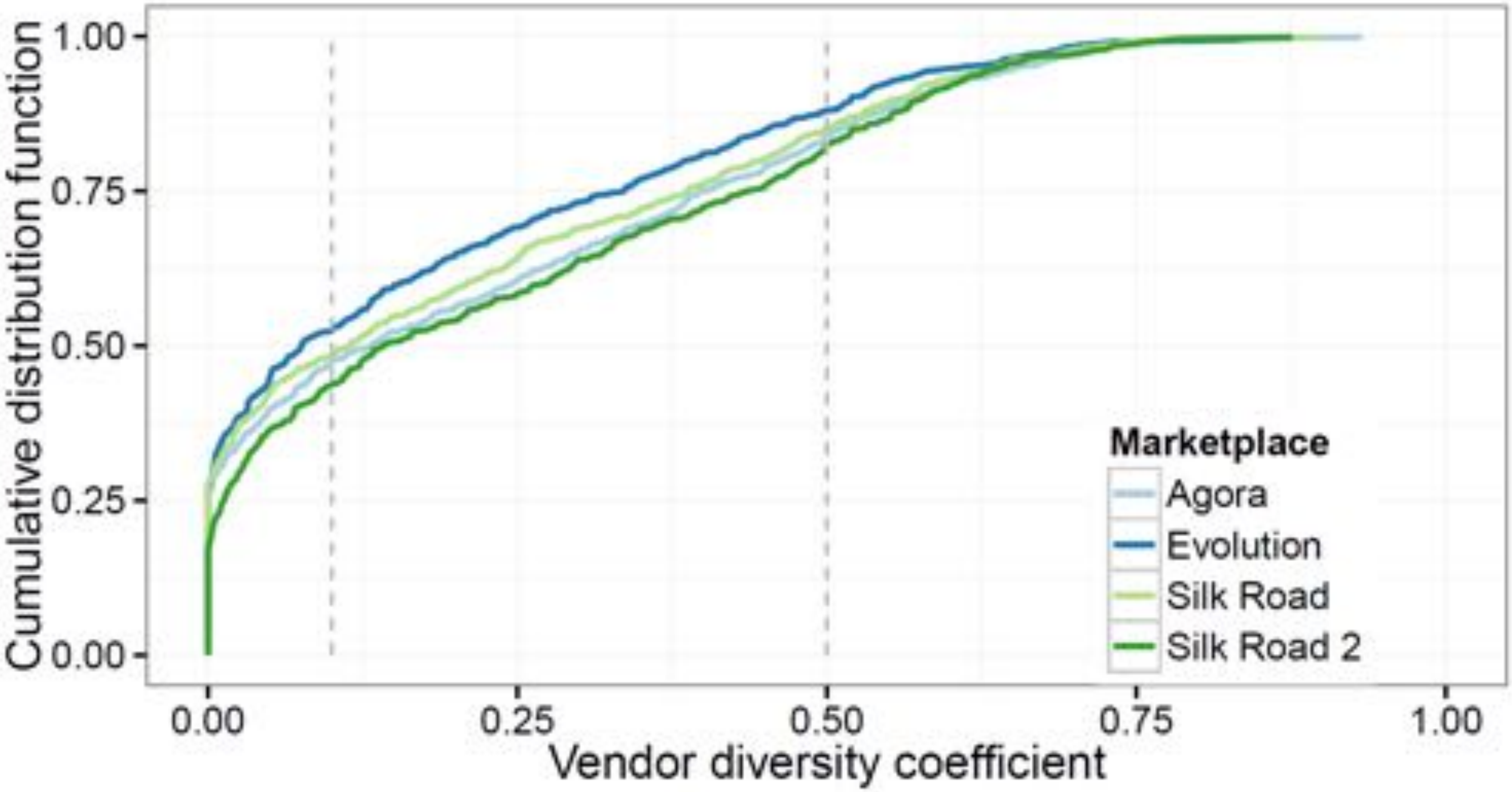


Vendor Diversity

- **Do vendors specialize in what they are selling?**
  - Do vendors sell what they make?
  - Does a single online presence sell goods for several diversified suppliers?

- **Coefficient of Diversity**
  - 0 – all sales from same category
  - 1 – equal sales from each category
  - Only vendors > $10,000 total sales considered

Vendor Diversity CDF



Validation

- Trial evidence GX226A, GX227C places Silk Road 1 weekly volumes at $475,000/week in late March 2012, consistent with our estimates

- Administrator reports Silk Road 2 daily volumes of around $250,000 in September 2014, similar to our estimated $270,000

- Leaked Agora vendor page shows sales total on June 5, 2014 to be $3,460, our observations yielded $3,408

Takeaways

- Anonymous Marketplaces are very easy to setup and use and have wide customer appeal

- Anonymous Marketplace ecosystem transacts in excess of $500,000 / day

- Anonymous Marketplaces are primarily used (~75%) for recreational drugs

- Anonymous Marketplace ecosystem has historically recovered from takedown efforts and scams

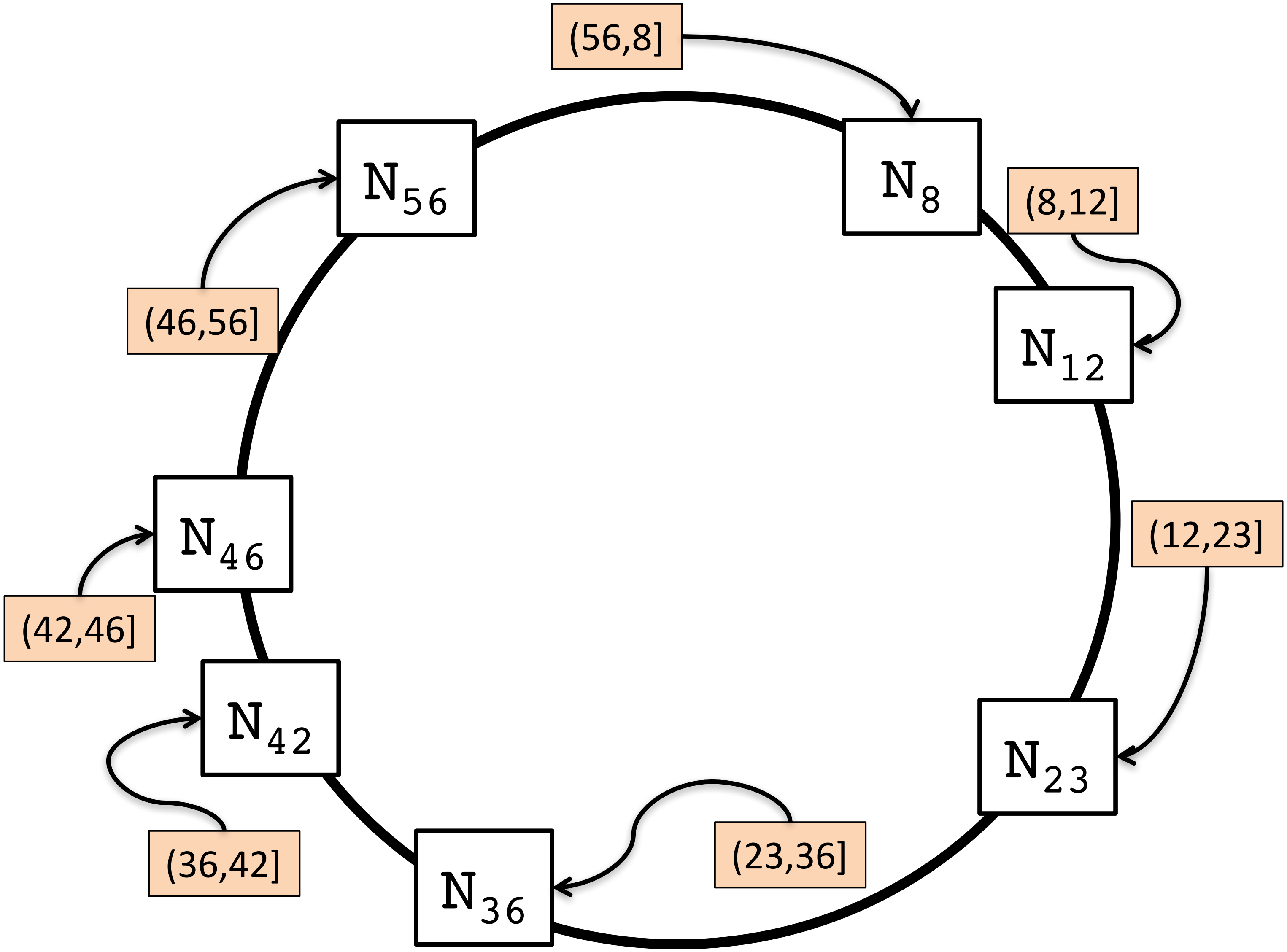
- Anonymous Marketplaces are controlled by small set of highly influential vendors

Kyle Soska – ksoska@cmu.edu

Non-Hidden Hidden Services Considered Harmful

Filippo Valsorda
George Tankersley

Example



What is Tor?

- **T**he **O**nion **R**outer
- Provides client anonymity
- Works by routing your connection though other machines

Hidden Services

- Provide *bidirectional* anonymity

- Supports generic TCP services

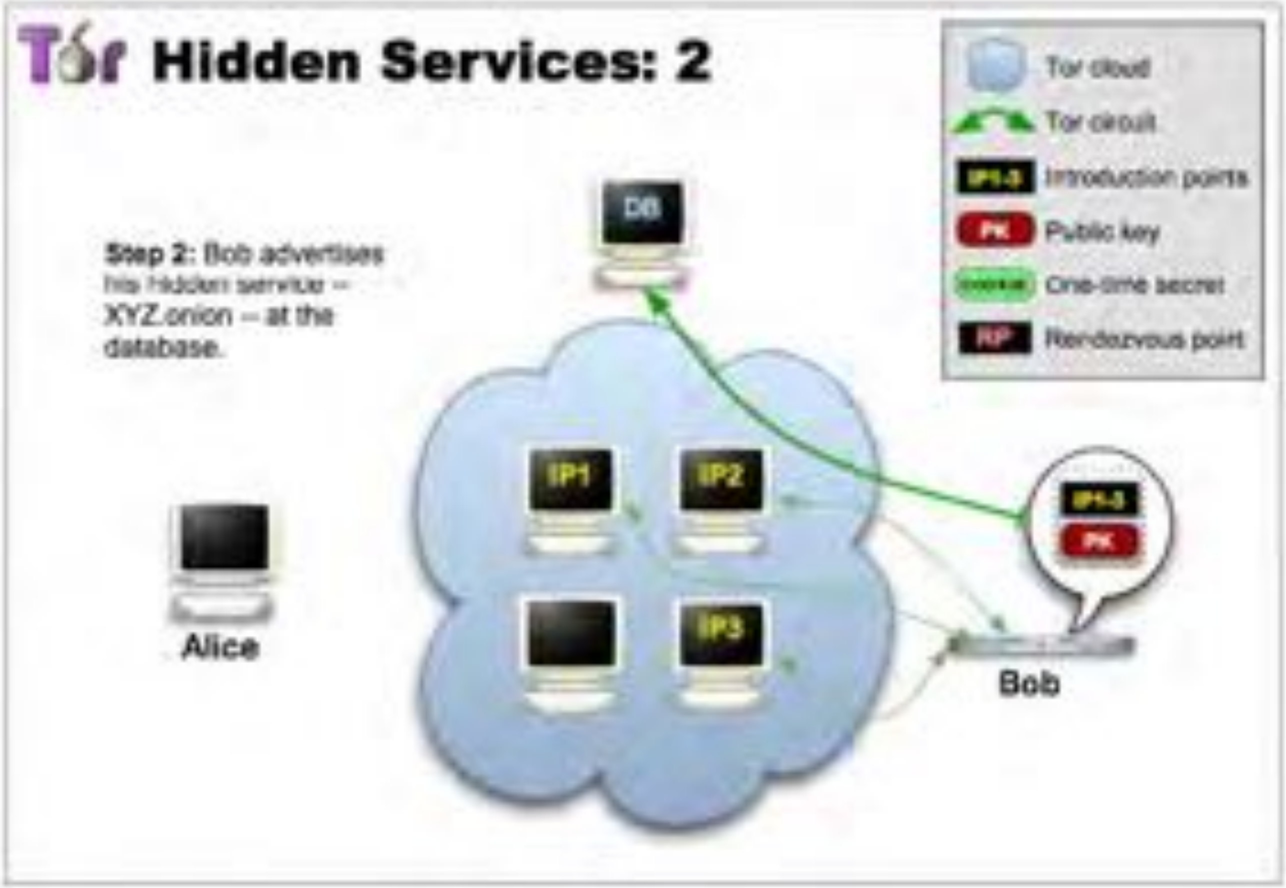
- Famous for drug markets
  - Silk Road
  - Silk Road 2

Hidden Services

But they're actually used for good

- Whistleblowing (SecureDrop)
- Private chat (Ricochet, XMPP-over-HS)
- Anonymous publishing (of course!)

Hidden Services



Tor Hidden Services: 2

Step 2: Bob advertises his hidden service -- XYZ.onion -- at the database.

Hidden Services

The “database” is a DHT made up of stable relays

- directory authorities grant *HSDir* flag
- not related to *Stable* flag

How do we choose where to publish?

HSDir selection

Choose two sets of 3 relays with *HSDir* flag

Think “consistent hashing”

- relays arranged in a ring sorted by identity

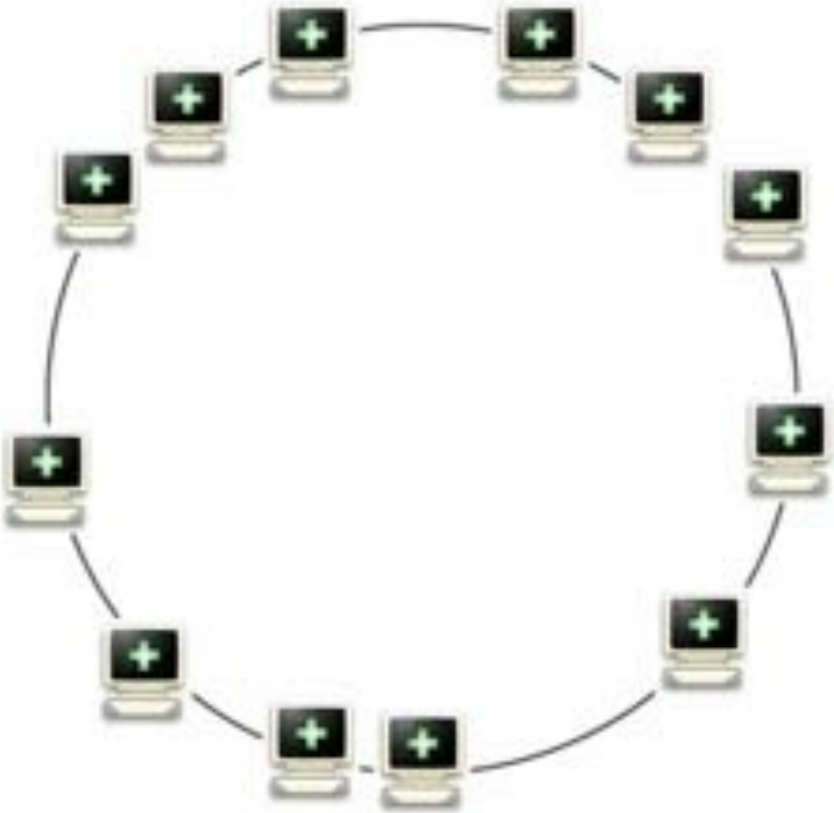
Based on a predictable formula (#8244)

HSDir selection

hs-descriptor-id =
    SHA1(id || SHA1(time-period || replica))

id: first 80 bits of SHA1(public key)
time-period: days since epoch (+offset)
replica: which set of HSDirs

HSDir selection



HSDir selection

facebookcorewwwi.onion

descriptor-id =

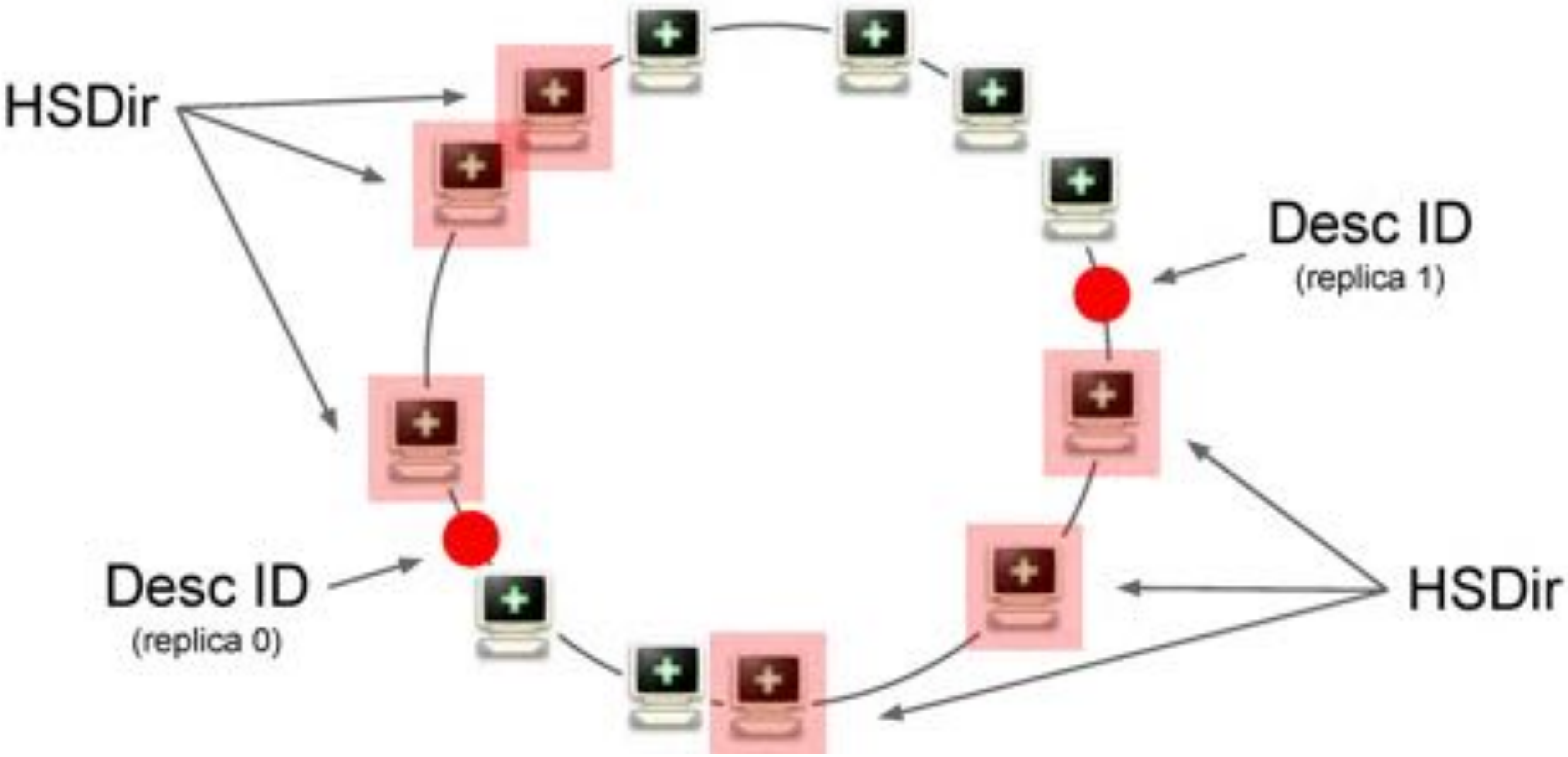
SHA1(facebookcorewwwi || SHA1(16583 || 0))

SHA1(facebookcorewwwi || SHA1(16583 || 1))

replica 0: ys5pml4c6txpw5hnq5v4zn2htytfejf2

replica 1: fq7r4ki5uwcxdxibdl7b7ndvf2mvw2k2

HSDir selection



Why did he just explain all this?

Point of the talk!

Hidden service users face a greater risk of targeted deanonymization than normal Tor users.

Vulnerability of Tor

Low-latency implies correlation attacks

Correlation attacks

in Tor, “both ends” means we’re usually just worried about entry nodes and exit nodes

- **entry nodes** see when a connection starts
- **exit nodes** see when it terminates

Correlation attacks

worried about entry nodes and exit nodes

- ***entry nodes** see when a connection starts*
- ***exit nodes** see when it terminates*

Tor has protections for entry/exit positions

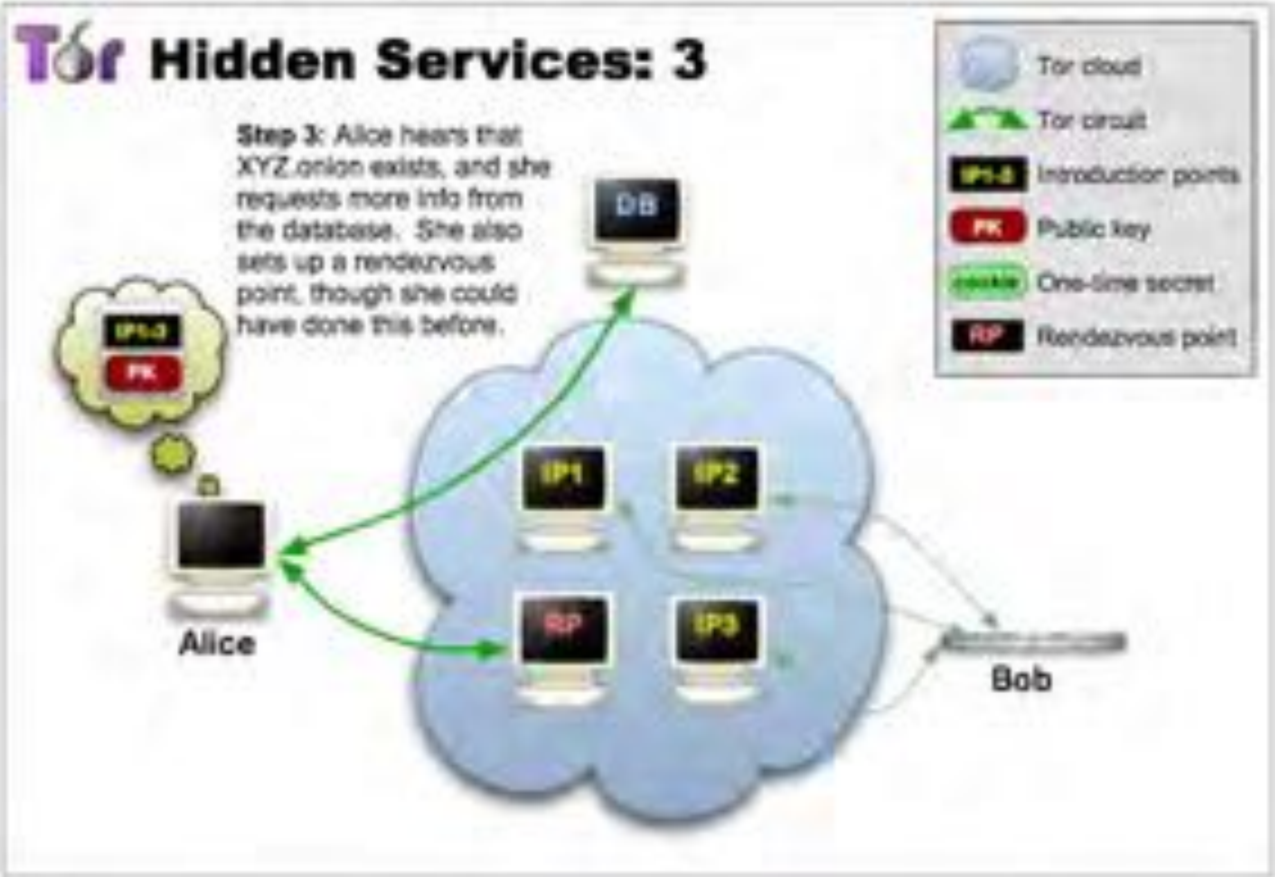
- entry guards, bad relay monitoring, size of network

Correlation attacks

It is hard to become both ends of a circuit.

What else can see when connections happen?

Hidden Services



Tor Hidden Services: 3

Step 3: Alice hears that XYZ.onion exists, and she requests more info from the database. She also sets up a rendezvous point, though she could have done this before.

Hidden Services

An HSDir for a hidden service gets a lookup on ⅙ of requests for information about the hidden service

A lookup indicates a user trying to connect to the hidden service

Correlation attacks

worried about entry nodes and exit nodes

- ***entry nodes** see when a connection starts*
- ***exit nodes** see when it terminates*

For a hidden service, the HSDir can see when a connection happens

Correlation attacks

*worried about entry nodes and **HSDir***

- ***entry nodes** see when a connection starts*
- ***HSDir** see when it terminates*

For a hidden service, the HSDir can see when a connection happens

Correlation attacks

If your target uses a hidden service, don't need exit relay to see when the connection happens.

Instead, be an HSDir.

Hidden Services

It is very easy to become HSDir

- You just need 4 days uptime
- It should be harder than it is (#8243)

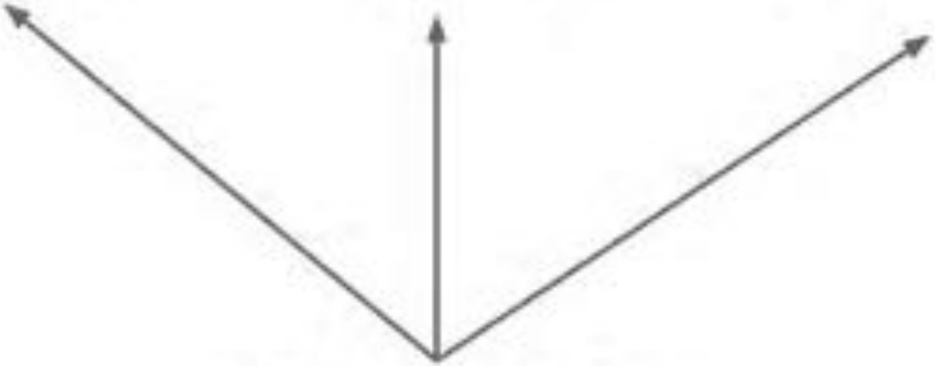
In fact, very easy to become *specific* HSDir

Positioning attack

SHA1(id || SHA1(time-period || replica))

Positioning attack

SHA1(**id** || SHA1(**time-period** || **replica**))



PREDICTABLE

Positioning attack

Predictable and fast? Bruteforce it!

1) Calculate descriptor IDs for the service
2) Generate random 1024-bit RSA key
3) Check if hash precedes the first real descriptor ID in the DHT
4) If not, goto 2

Correlation attacks

If your target uses a hidden service, don’t need exit relay to see when the connection happens.

Instead, be **their** HSDir.

Correlation attacks

If your target uses a hidden service, don’t need exit relay to see when the connection happens.

Instead, be **every** HSDir.

Vulnerability of Tor

worried about entry nodes and HSDir

- ***entry nodes*** *see when a connection starts*
- *HSDir see when it terminates*

Vulnerability of Tor

worried about entry nodes and HSDir

- ***many people*** *see when a connection starts*
- *HSDir see when it terminates*

Vulnerability of Tor

worried about entry nodes and HSDir

- ***many people*** *see when a connection starts*
- *HSDir see when it terminates*

“entry” does not just mean your entry node

- ISP, malicious access point, pen register…

Summarizing all of that

1) HSDirs can serve the same purpose against a hidden service as a malicious exit relay would in a basic correlation attack

2) The "entry side" of a Tor connection can be monitored by means other than compromising guards

Summarizing all of that

It's actually ***worse***, because it's way easier to be the user's HSDir.

Hidden service users face a greater risk of targeted deanonymization than normal Tor users.