

Data Privacy

Sharing sensitive data while keeping it private
Privacy Preserving Machine Learning

Didier Guillevic



Context

0. Context - Lack of Privacy - Data is Dangerous

Techniques aiming to keep data private

1. Data anonymization - De-identification
2. Differential Privacy
3. Federated Learning
4. Encrypted: Multi Party Computation
5. Encrypted: Homomorphic Encryption

0. Context: Location Data / tracking

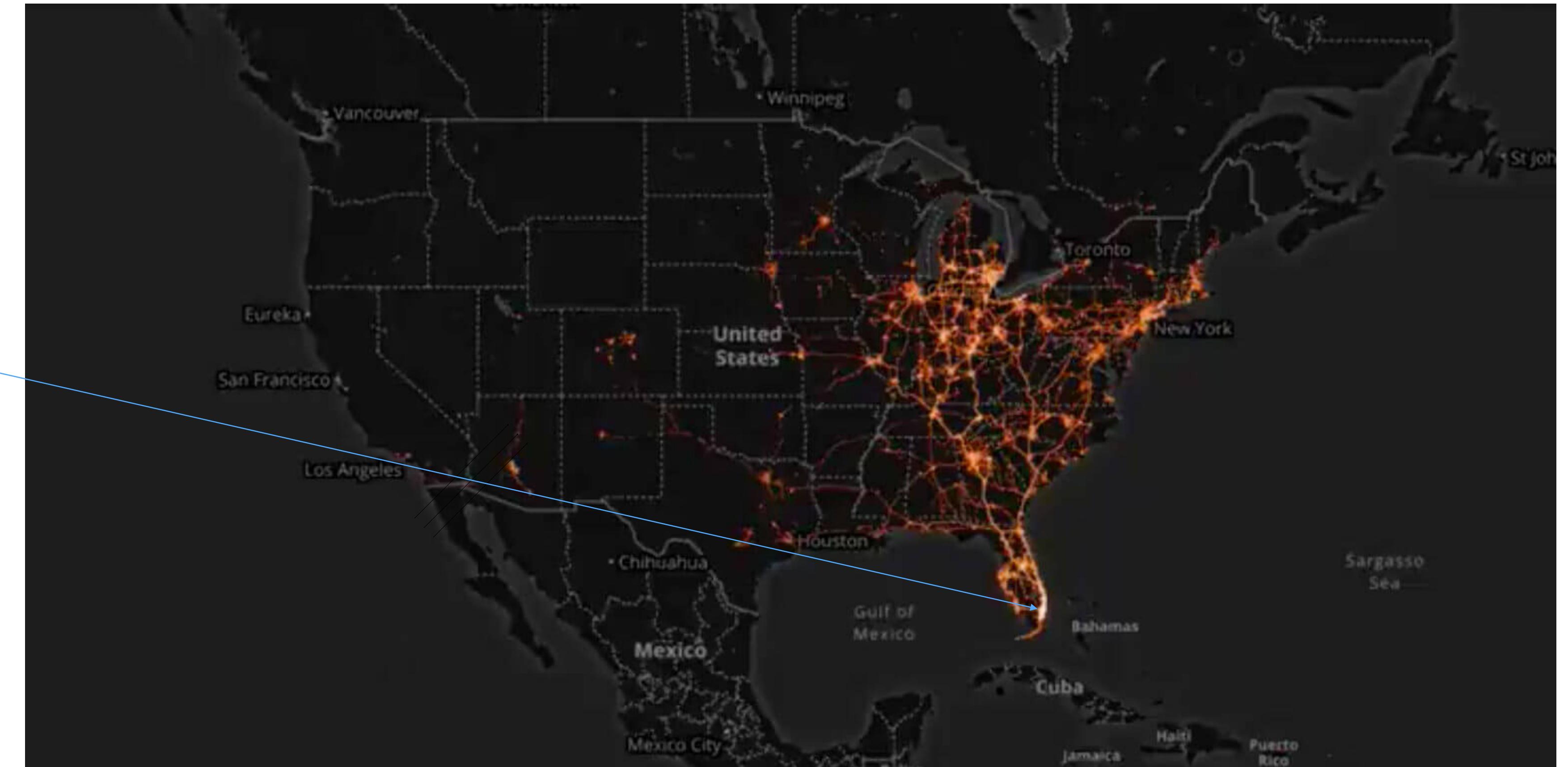
Big data: a chance to mine anonymized demographic, financial, medical, and other vast data sets

Fort Lauderdale - Spring Break 2020 - Cell phone heat map during coronavirus pandemic (Twitter)

Tectonix: Location data processing visualization



xmode.io: data

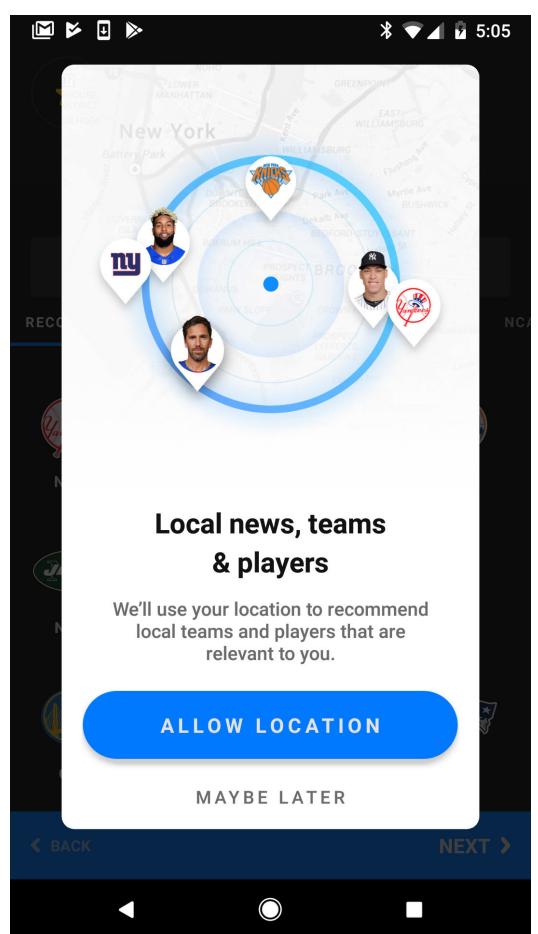
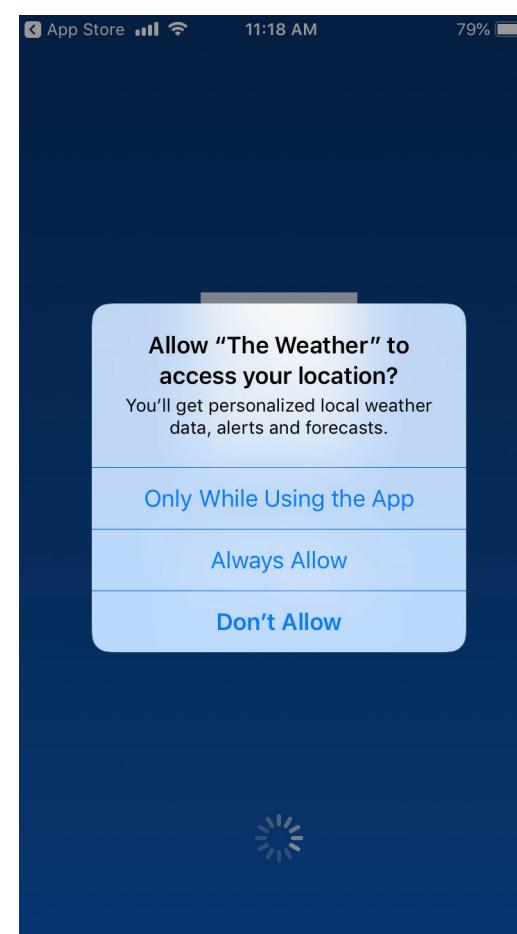


Day 0: 5,600 cell phones selected

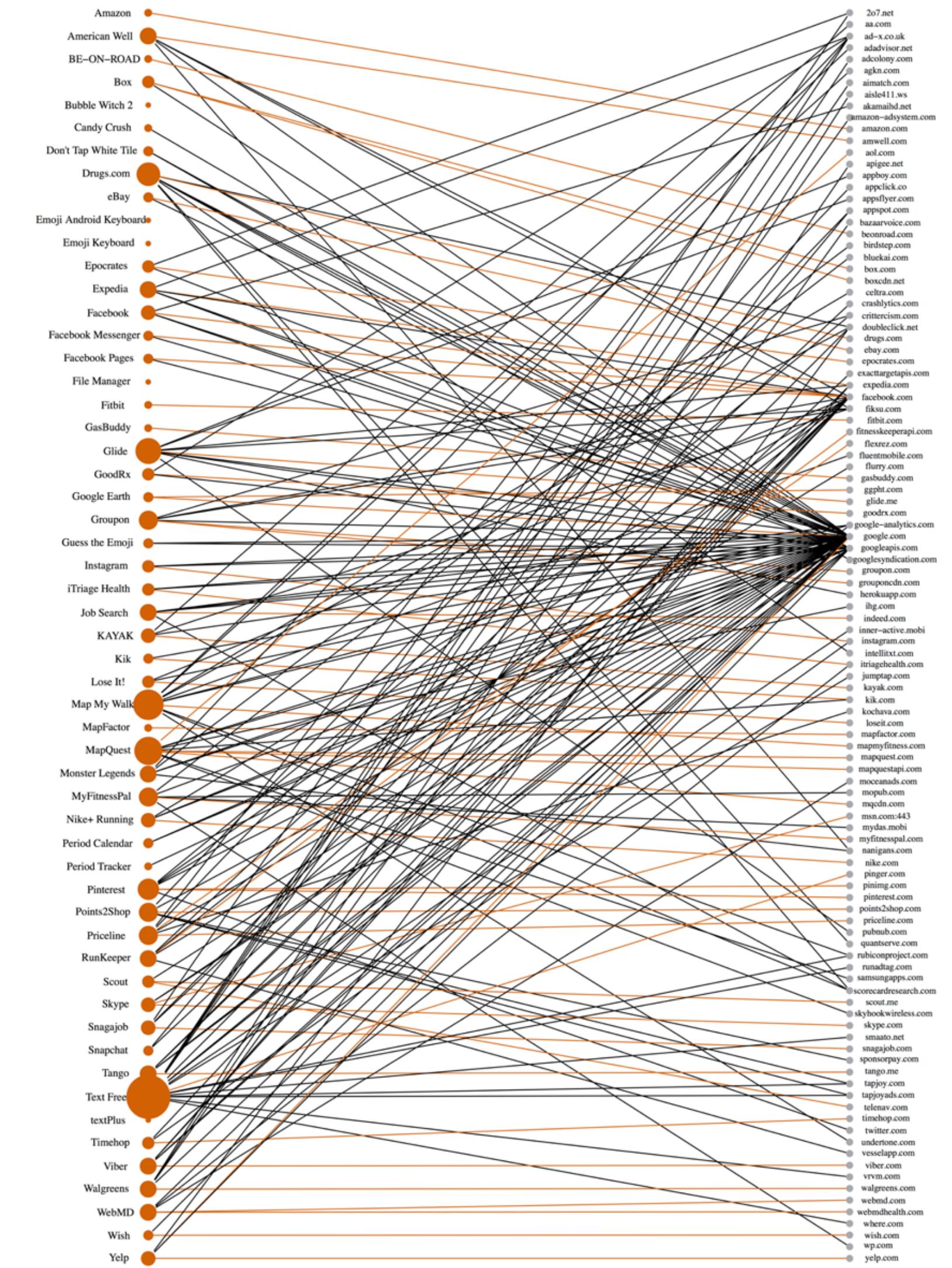
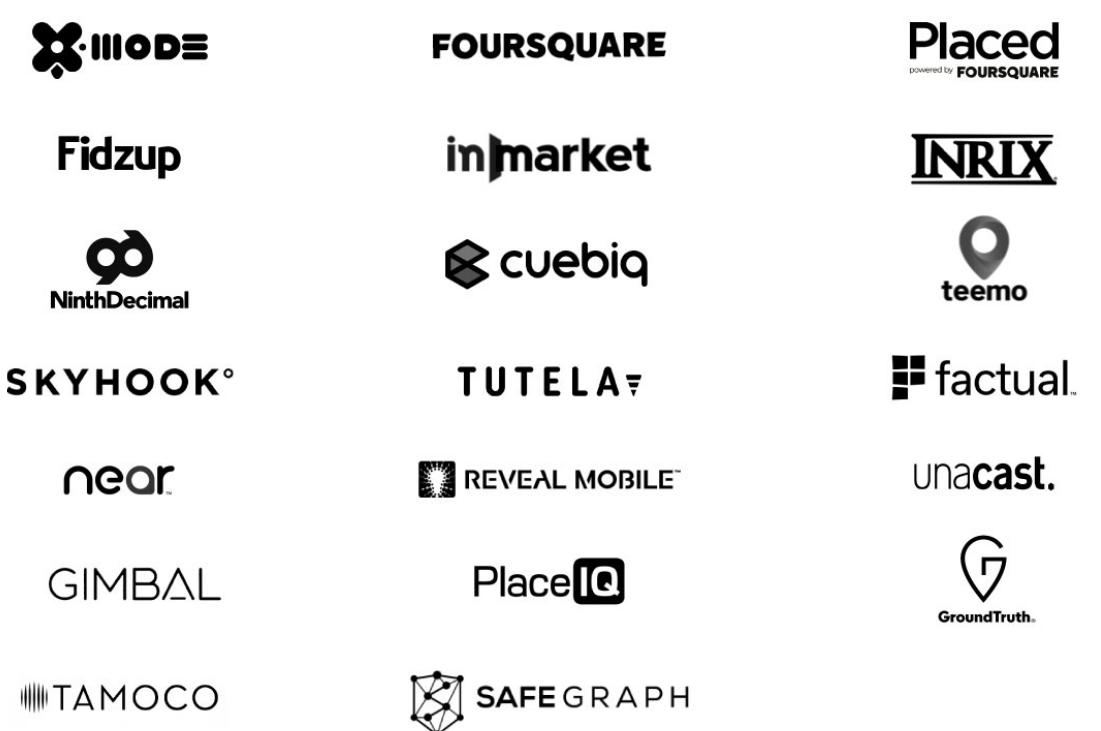
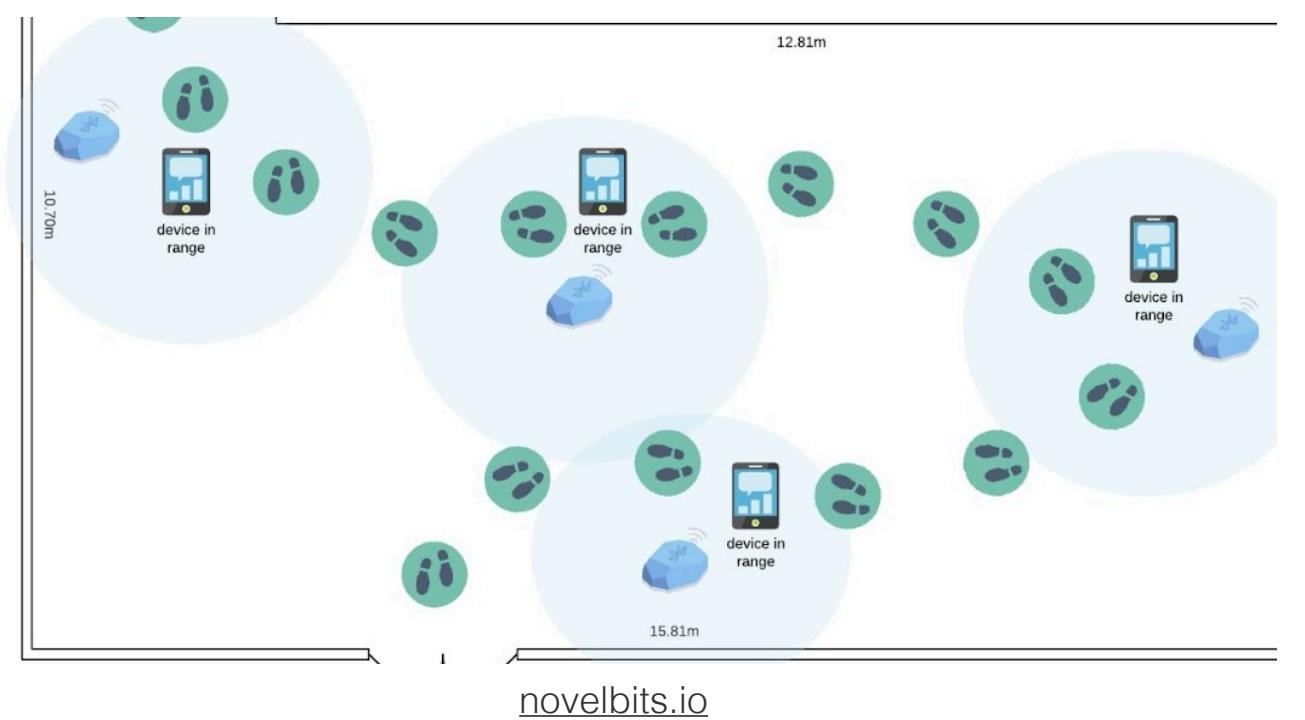
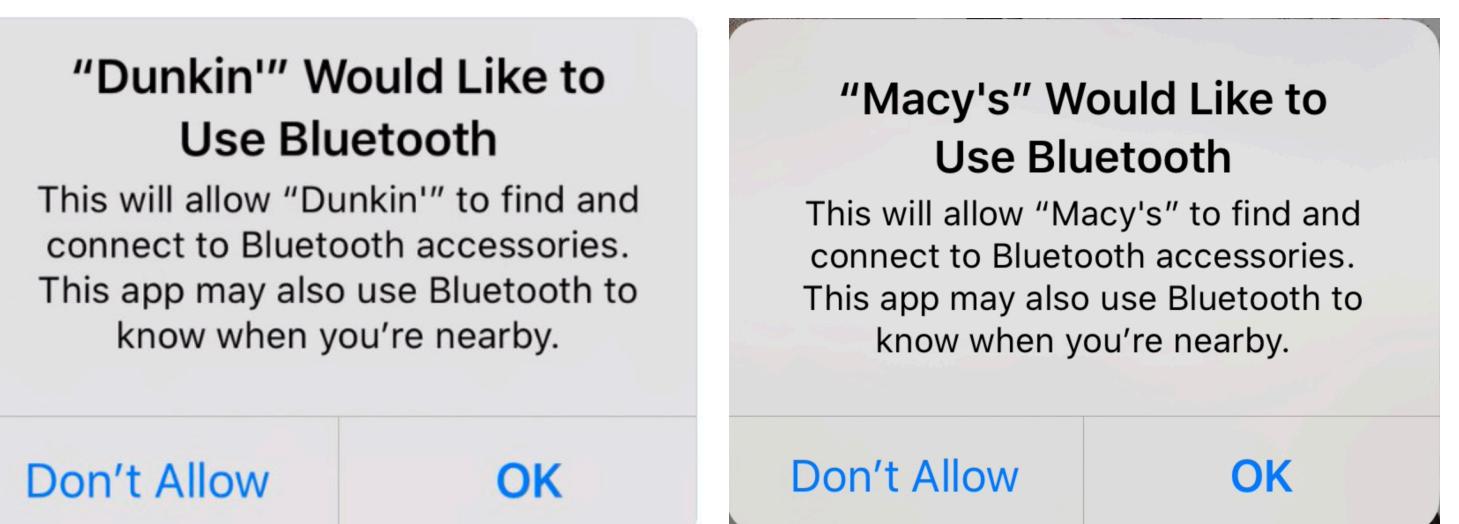
Day 14: Locations visited by those 5,600 call phones

Collecting Location Data: Smartphones Are Spies

1. Phone companies have the data
2. **Data brokers**: collected from mobile apps with location sharing enabled
3. Wifi handshaking: used by mobile apps
4. Bluetooth beacons: used by mobile apps when denied location sharing



Apps asking to use Bluetooth in iOS13 (2019-09)



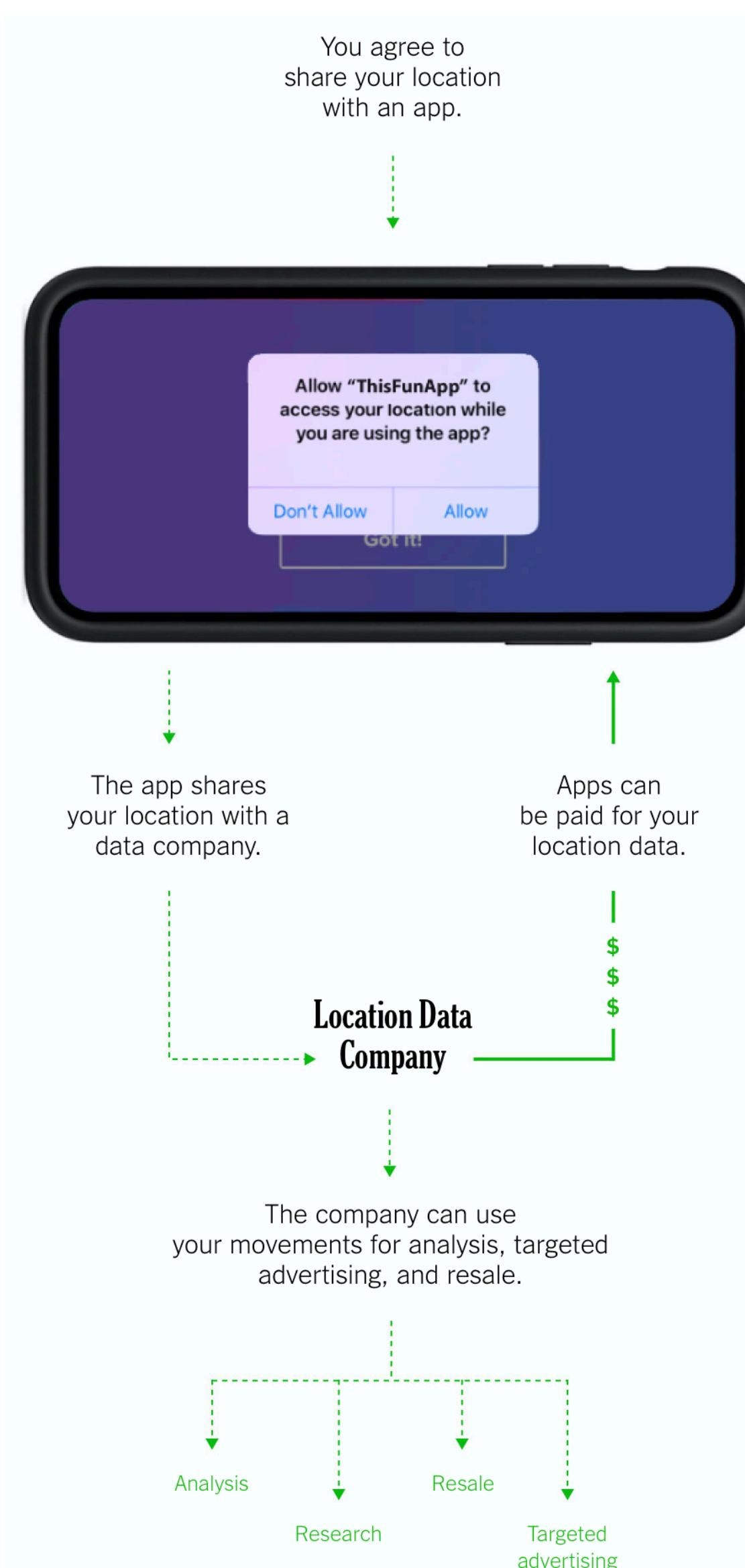
Who Knows What About Me?: Data sharing to third parties by mobile apps (2015)

(Harvard U.: tested 110 popular, free Android and iOS apps)

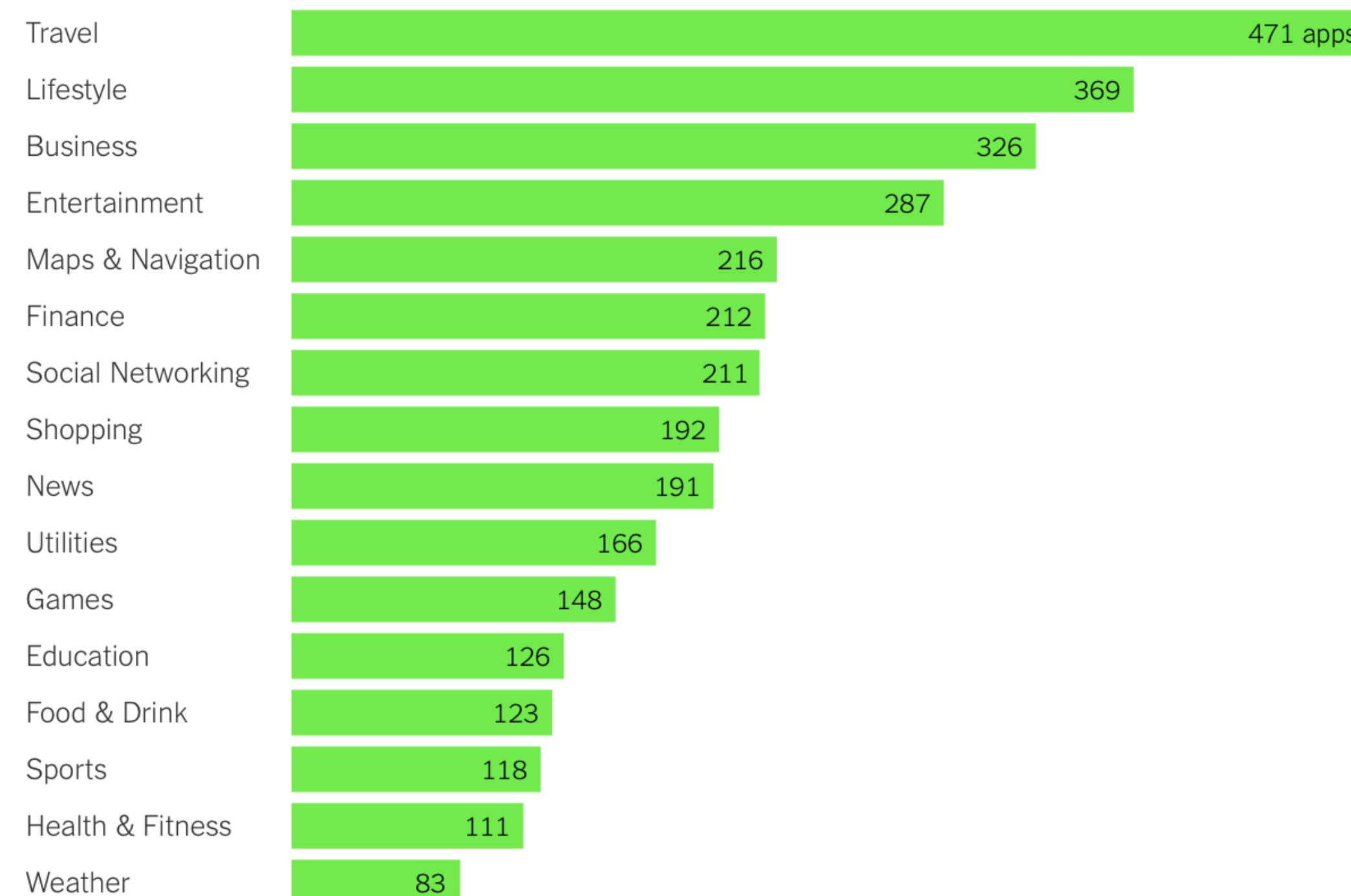
Your Apps Know Where You Were Last Night, and They're Not Keeping It Secret (NY Times 2018-12)

The Mobile Apps

Smartphones Are Spies. Here's Whom They Report To. ([NY Times - 2019-12-20](#))



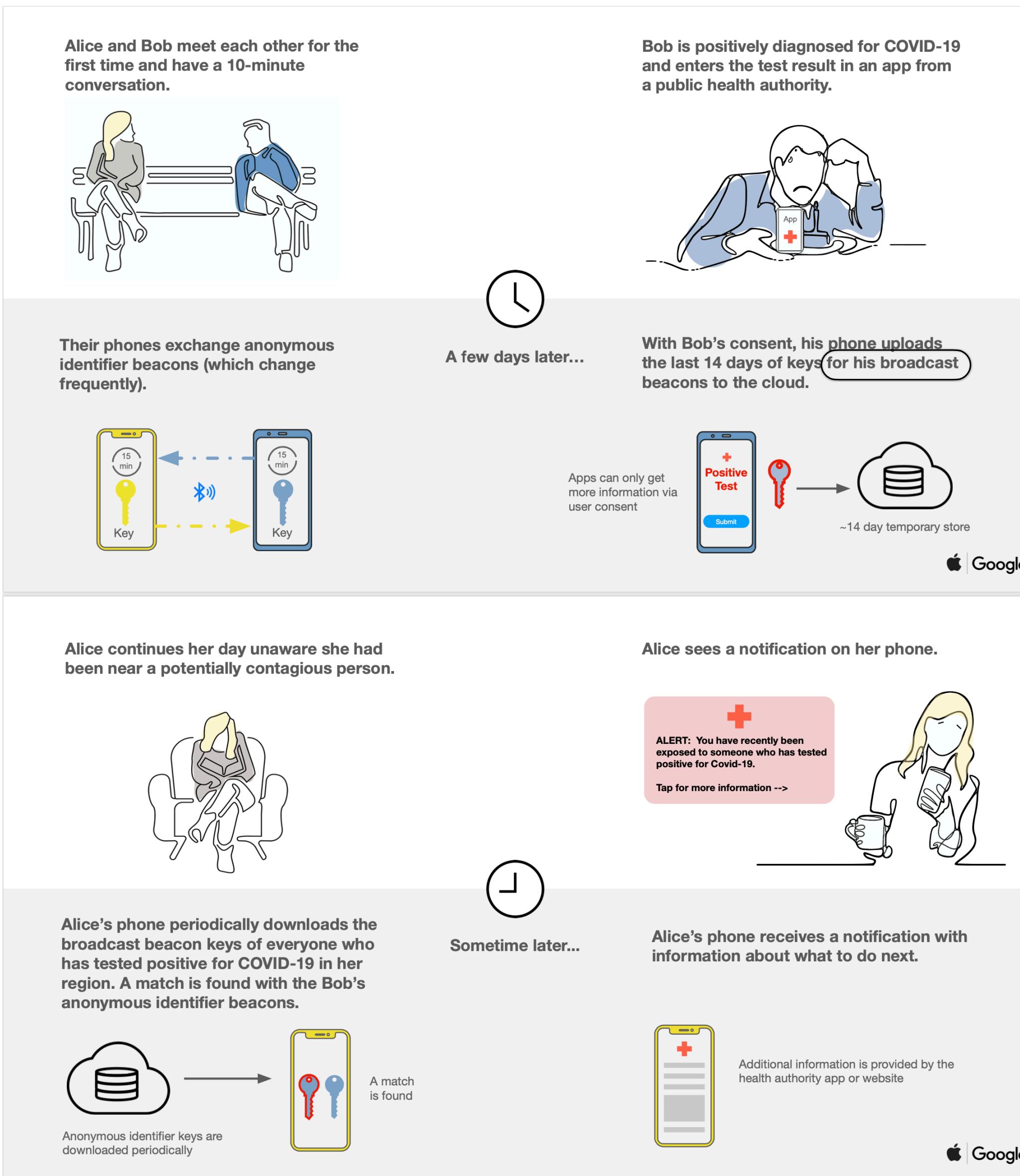
Apps With Location S.D.K.s Installed, by Category



Note: Some apps appear in more than one category. Excludes apps classified as "other." Source: Mighty Signal

Data For Good (privacy issues?)

Apple Google partner on COVID-19 contact tracing technology (2020-04-10)



Apple - Google and the EU (2020-05)

Ensure contact tracing apps are: anonymized, voluntary, transparent, temporary, secure

Some governments prefer **mandatory**, without any data privacy:



AarogyaSetu (India) **mandatory**: 80 million downloads (2020-05) - Going for 350 millions

Qatar: mandatory, access to all photos on the phone

China, ...

Over governments opt for voluntary basis:



COVIDSafe (Australia)



TraceTogether (Singapore)

2020-05-20 (CBC): 22 countries planning to use Apple/Google tech. respecting privacy

MIT Technology Review Covid Tracing Tracker

Location	Name	Notes	Voluntary	Limited	Data destruction	Minimized	Transparent	Tech
Algeria	Algeria's App	Algeria's app was investigated by Amnesty International.	★	★	★	★	★	★
Australia	COVIDSafe	Australian experts have criticized the government for a lack of transparency and non-responsiveness to privacy issues.	★	★	★	★	★	Bluetooth
Austria	Stopp Corona	Austria was one of the first major European nations to align with the Google/Apple API.	★	★	★	★	★	Bluetooth, Google/Apple
Bahrain	BeAware	Though 25% of the country has downloaded BeAware, there is little public information about the app.	★	★	★	★	★	Bluetooth, Location
Bulgaria	Virusafe	Bulgaria began lifting movement restrictions in early May	★	★	★	★	★	Location
Canada	COVID Alert*		★	★	★	★	★	Bluetooth, Google/Apple
China	Chinese health code system	There is very little information available to the public about how China's technology works.	★	★	★	★	★	Location, Data mining

2020-06-18

Contact Tracing Apps and Privacy (COVID-19)

What you need to know about contact tracing apps and privacy (Protonmail.com 2020-05)



Privacy by design: Decentralized (data stays on phone), open source, no geo-location, data deleted after 14 days

Open protocols:



Decentralized Privacy-Preserving Proximity Tracing ([DP3-T](#))



Temporary Contact Numbers ([TCN](#))



Privacy-Preserving Contact Tracing (inspired by DP3-T)

Countries: Switzerland, Austria, Estonia, Finland, Italy, Canada, ...



Flawed approaches: Centralized server: massive database that can be exploited or abused

Open protocols:



Pan-European Privacy-Preserving Proximity Tracing ([PEPP-PT](#))



[BlueTrace](#)

Countries: UK, France, Australia, New Zealand, Singapore, ...

nicky-case.me/contact-tracing



Undermining privacy: No effort to protect privacy: cell phone data, credit card purchases history, surveillance cameras

Geo-location data from phone companies, call records from mobile phones, compulsory, access to photo library, ...

Countries: [South Korea](#), [Israel \(NSO\)](#), [China](#), [Qatar](#), ...

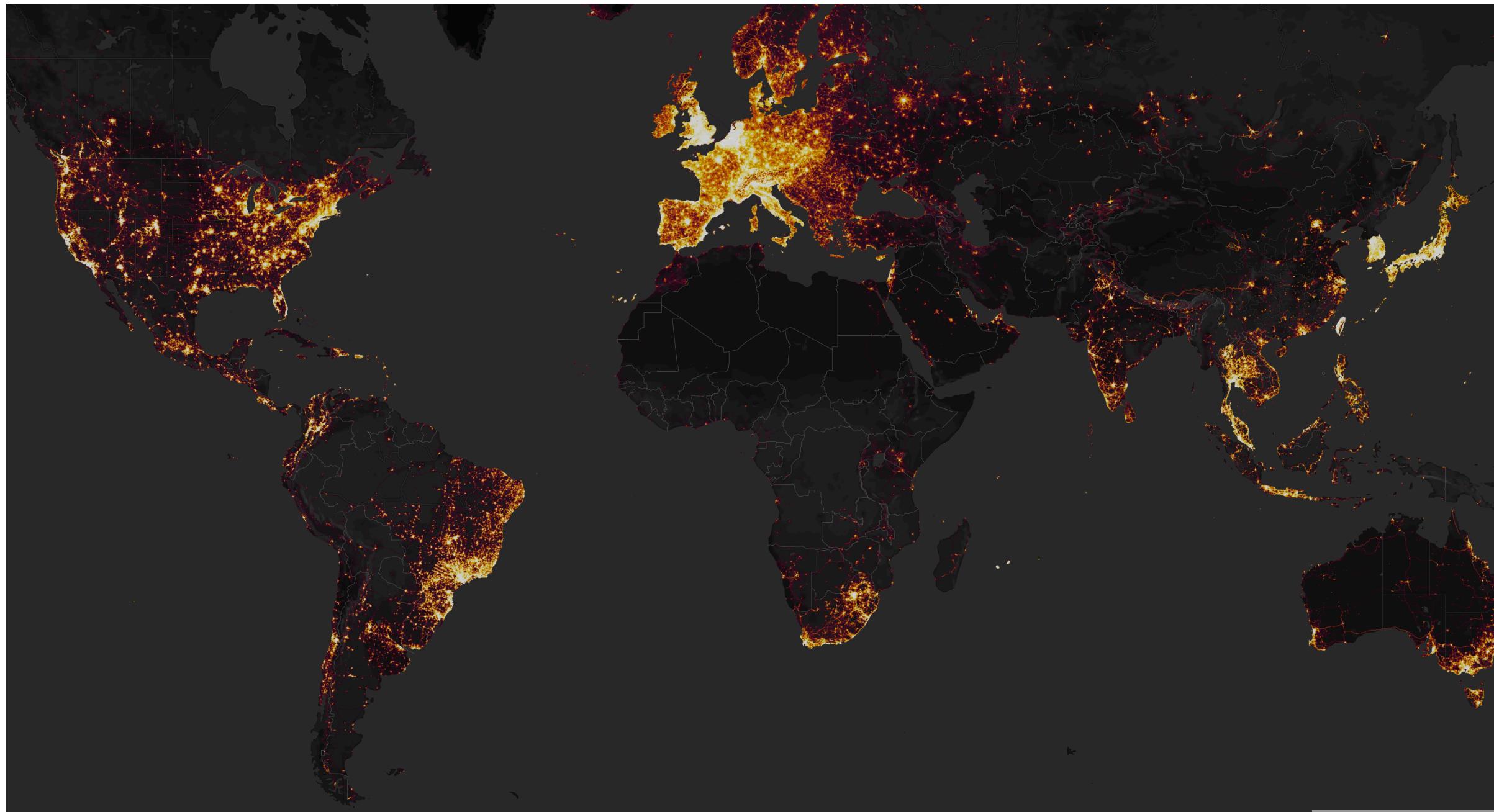
Sports Activity Mobile App: Location Tracking



The #1 app for runners and cyclists



Strava Heatmap (how to)

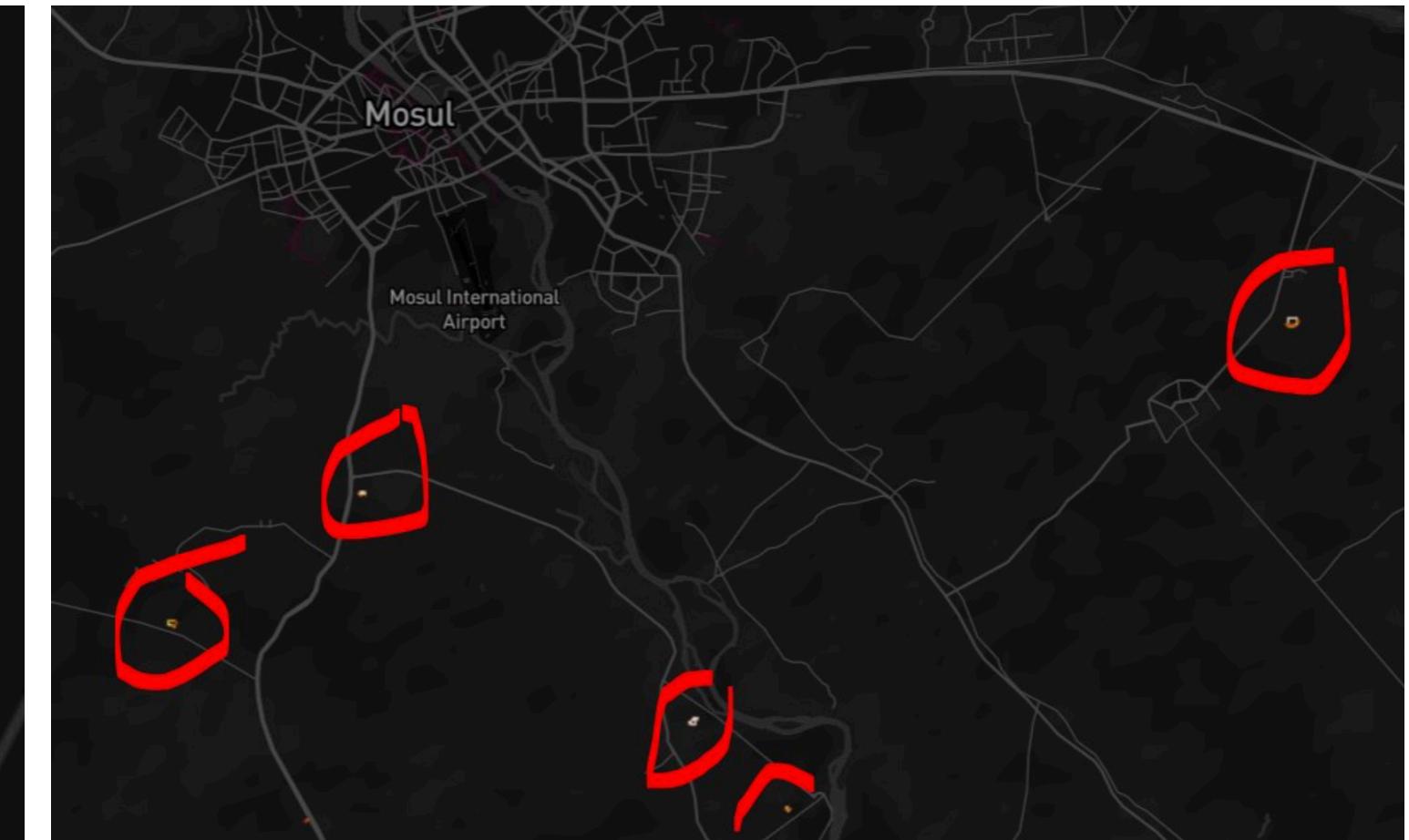


Locating a French military base in Africa



Tobias Schneider

Forward Operating Bases (FOB) in Afghanistan



Tobias Schneider

Alleged CIA black site in Djibouti



Strava/Fitbit-enabled
US Security patrols?
Equally-bright.

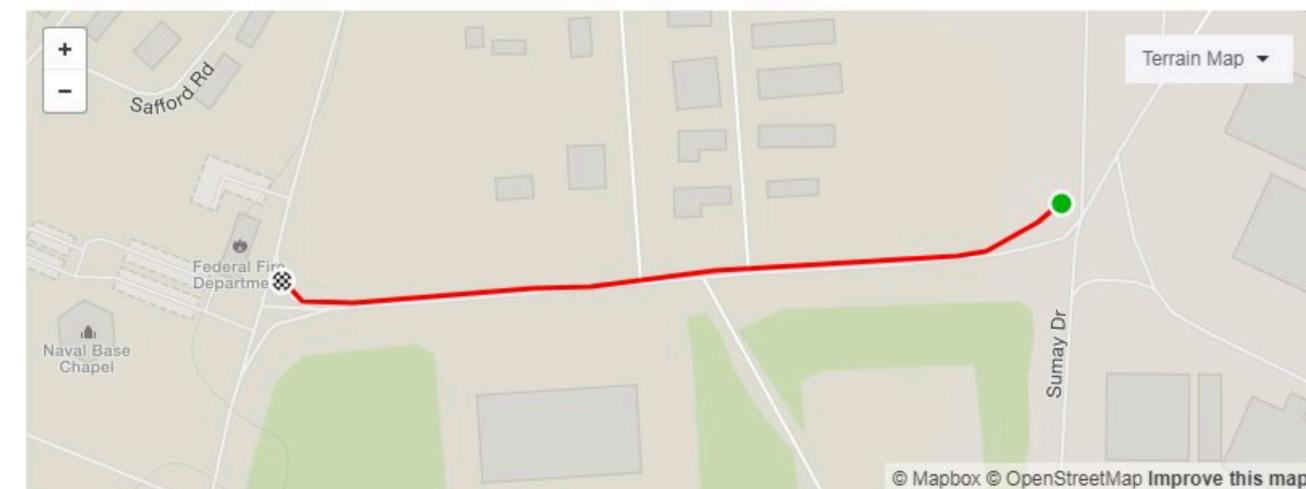
★ NBG HQ to Mini Mart

Run Segment Santa Rita, Guam

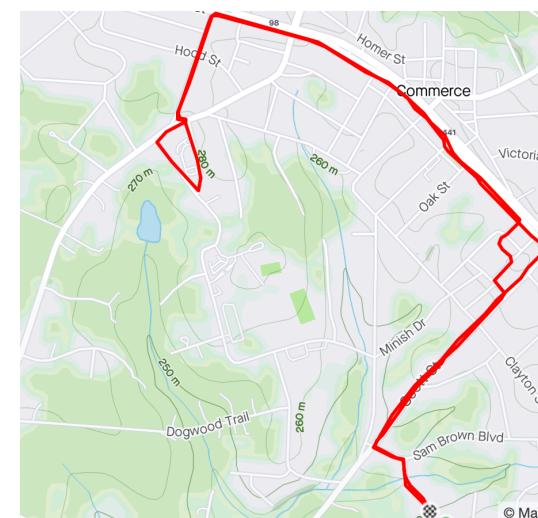
0.5km 0% 12m 14m 2m

Distance Avg Grade Lowest Elev Highest Elev Elev Difference 201 Attempts By 70 People

Inside military base with profiles of runners

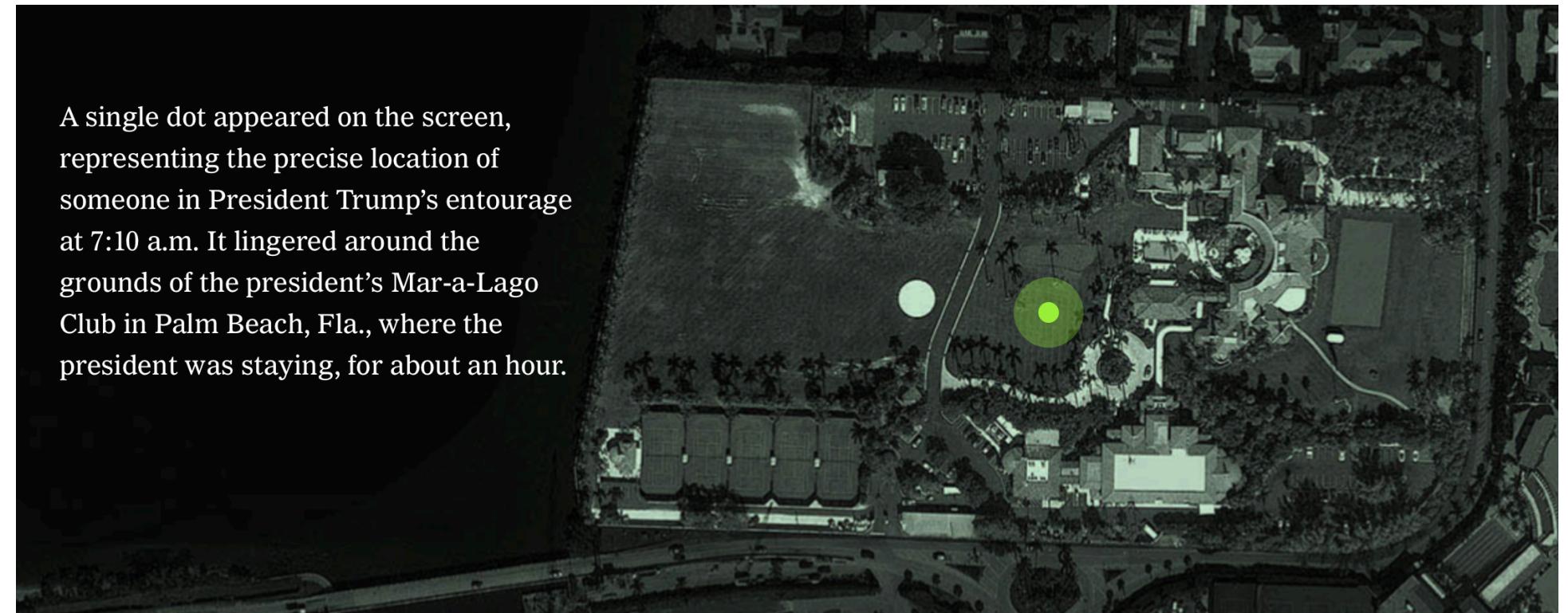


Sample user profile: runs from home

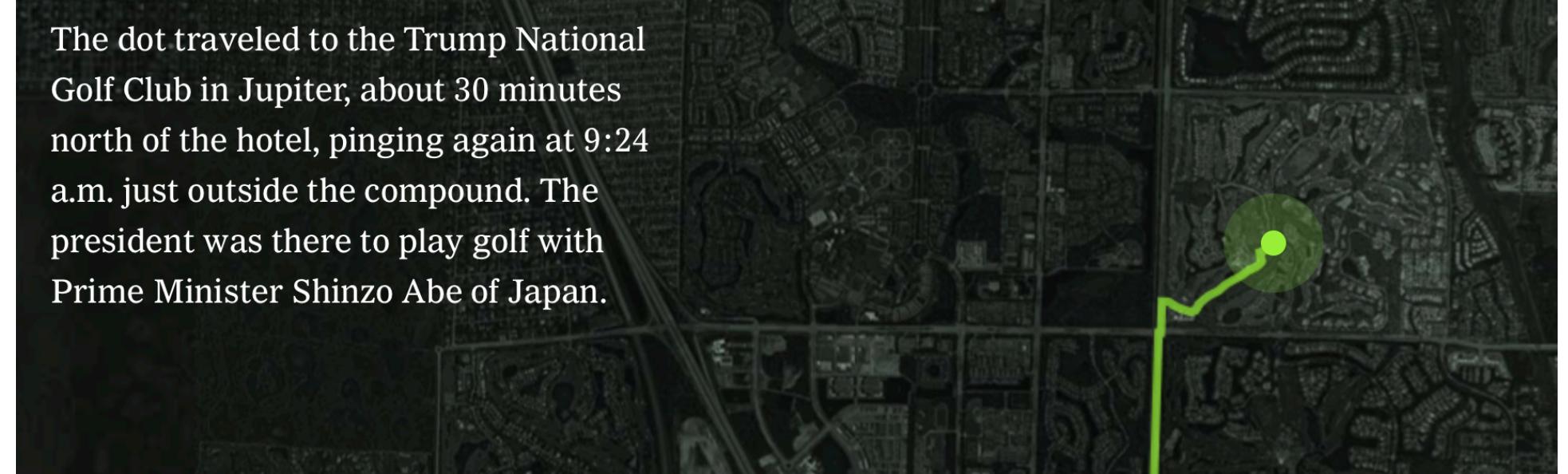


NY Times: How to Track President Trump

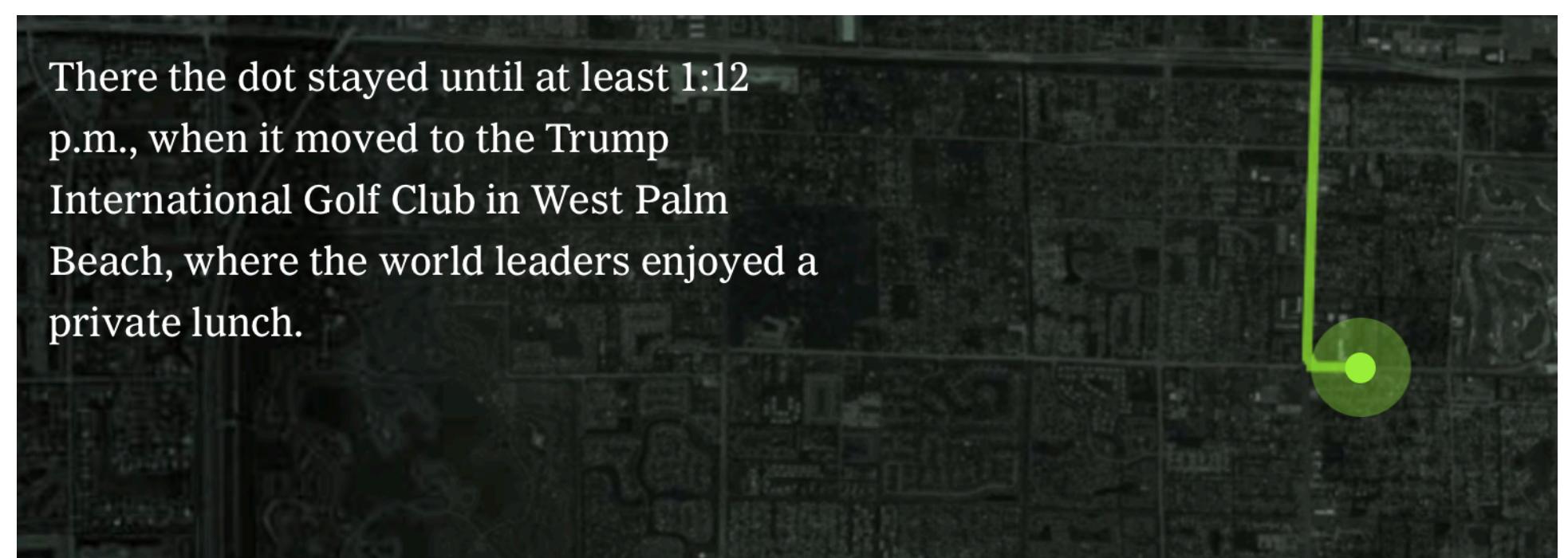
7:10 a.m.: Mar-a-Lago Club in Palm Beach



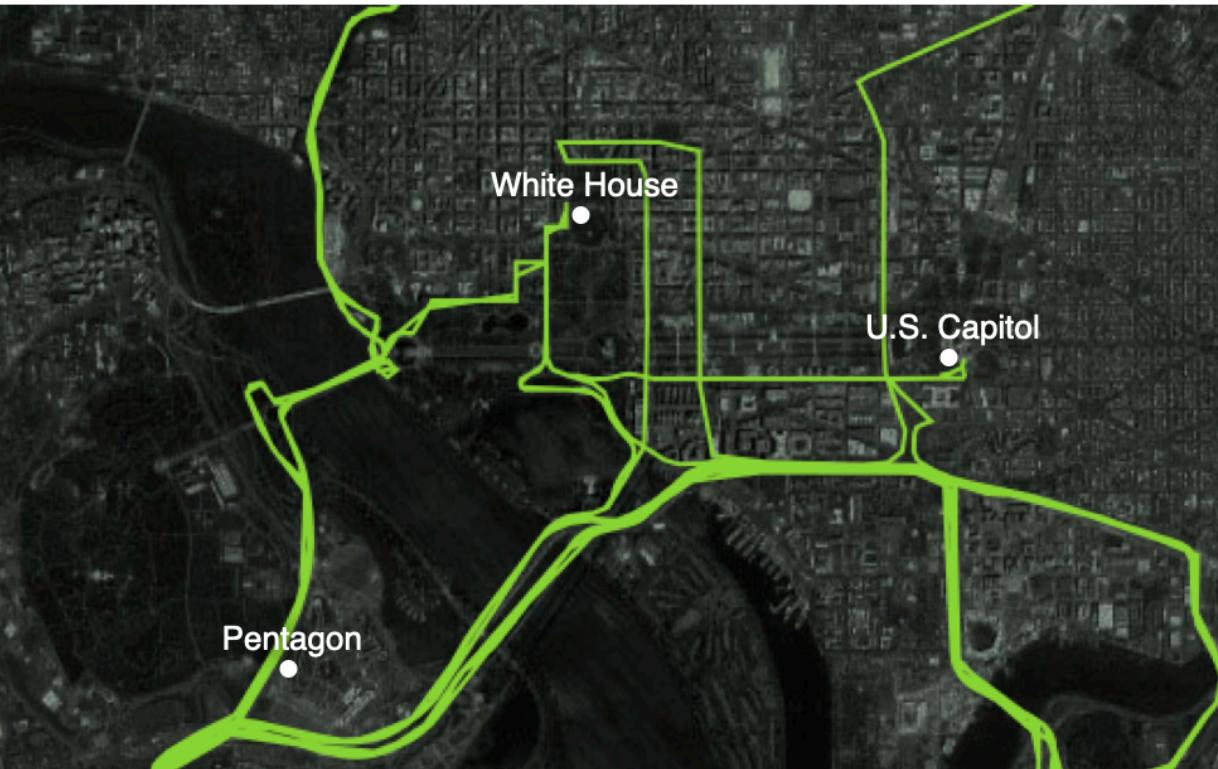
9:24 a.m.: Trump National Golf Club in Jupiter



1:12 p.m.: Trump International Golf Club in West Palm Beach



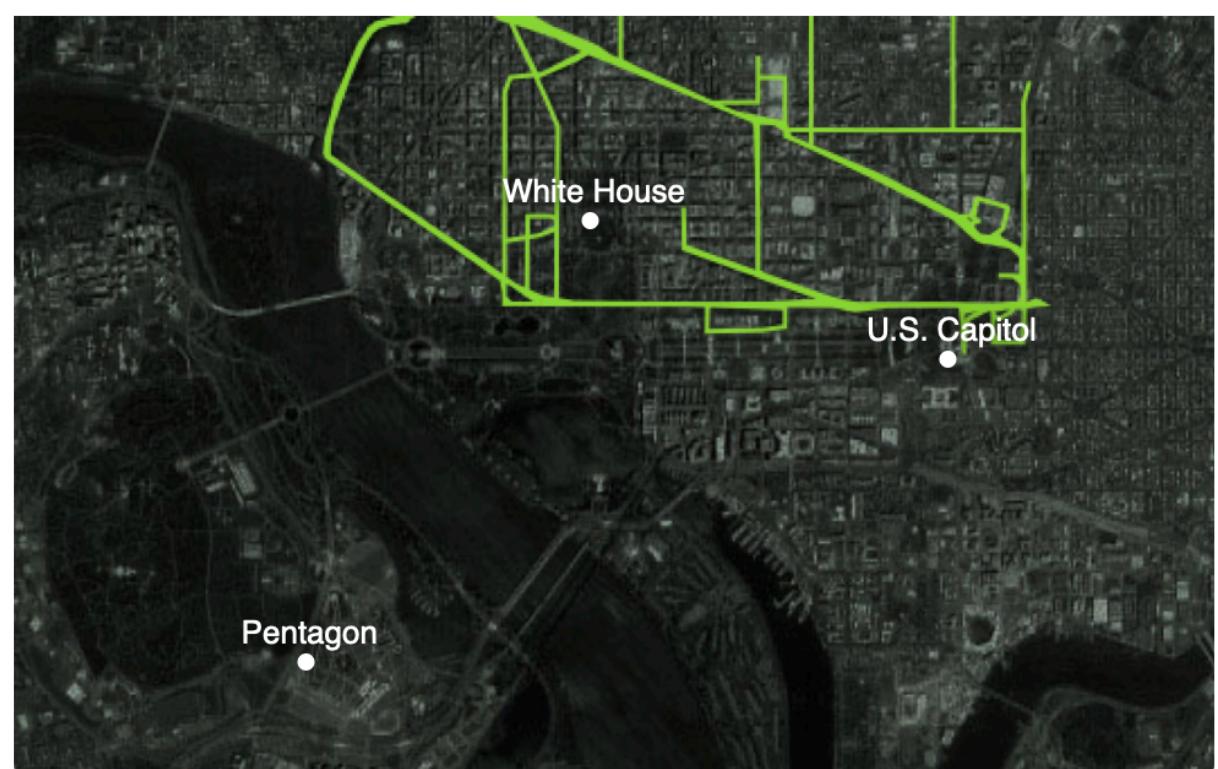
Most probably, smartphone a Secret Service agent, whose home was also clearly identifiable in the data



Technology at the Supreme Court

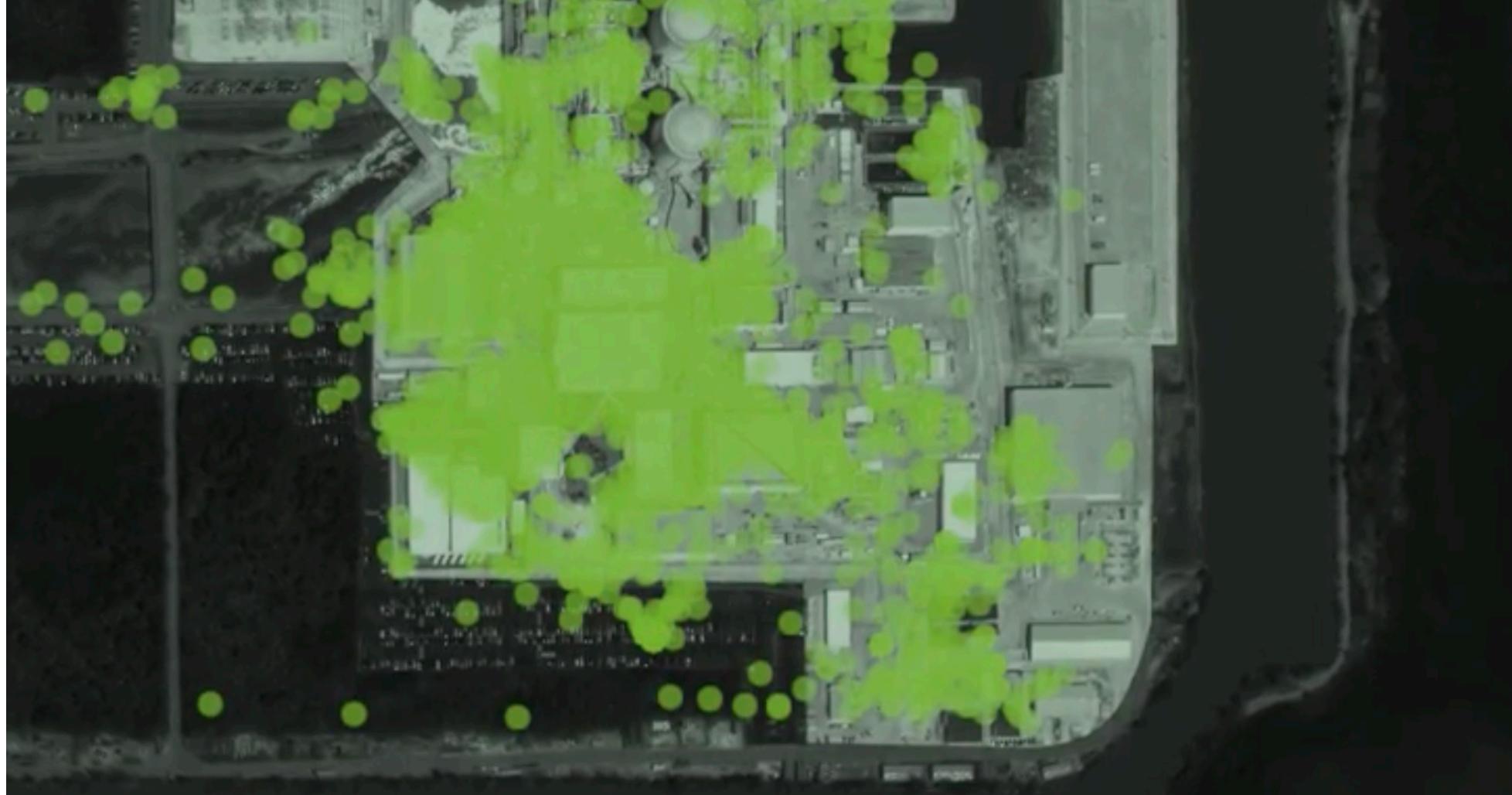
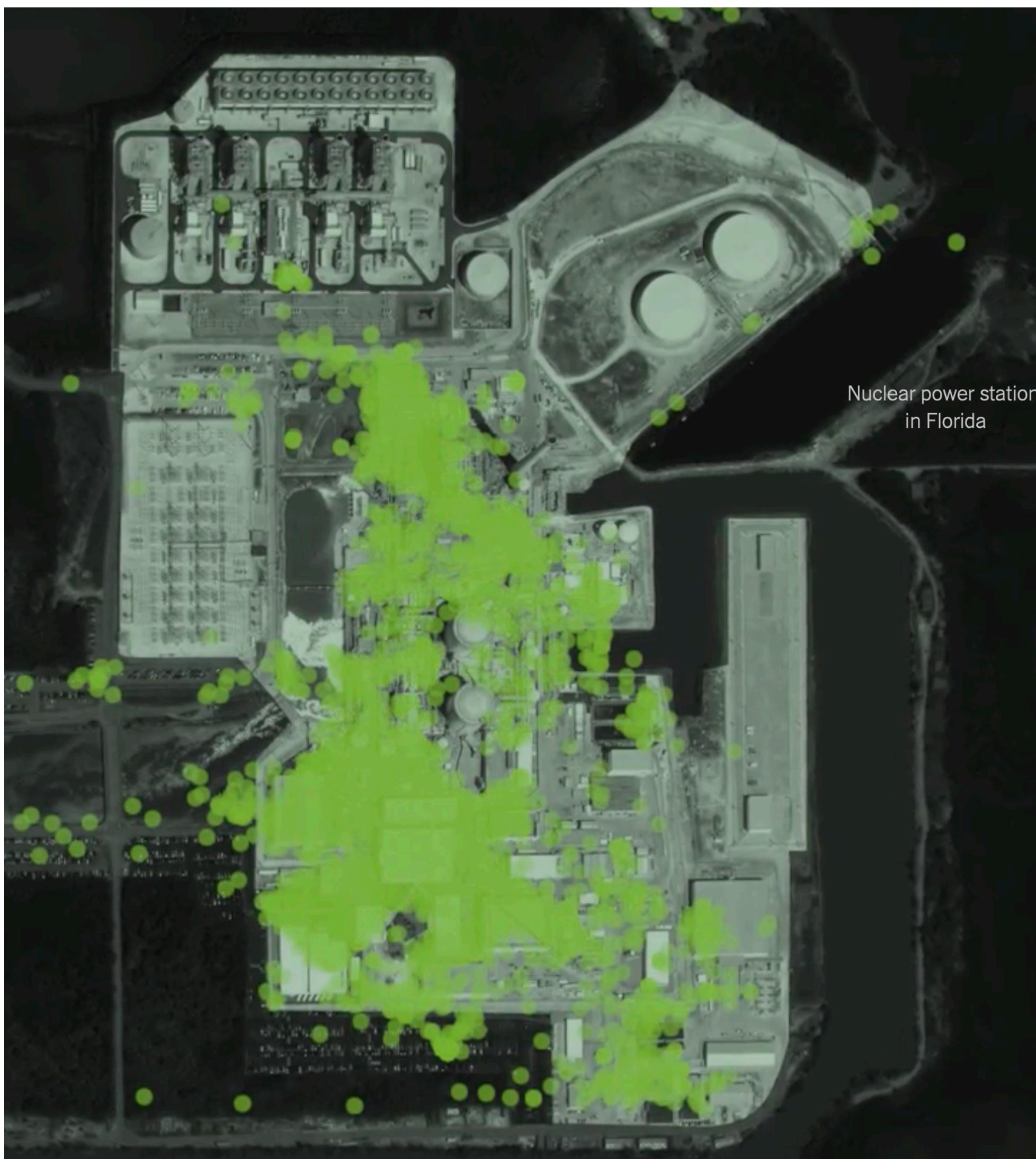


Advisor for a Senator



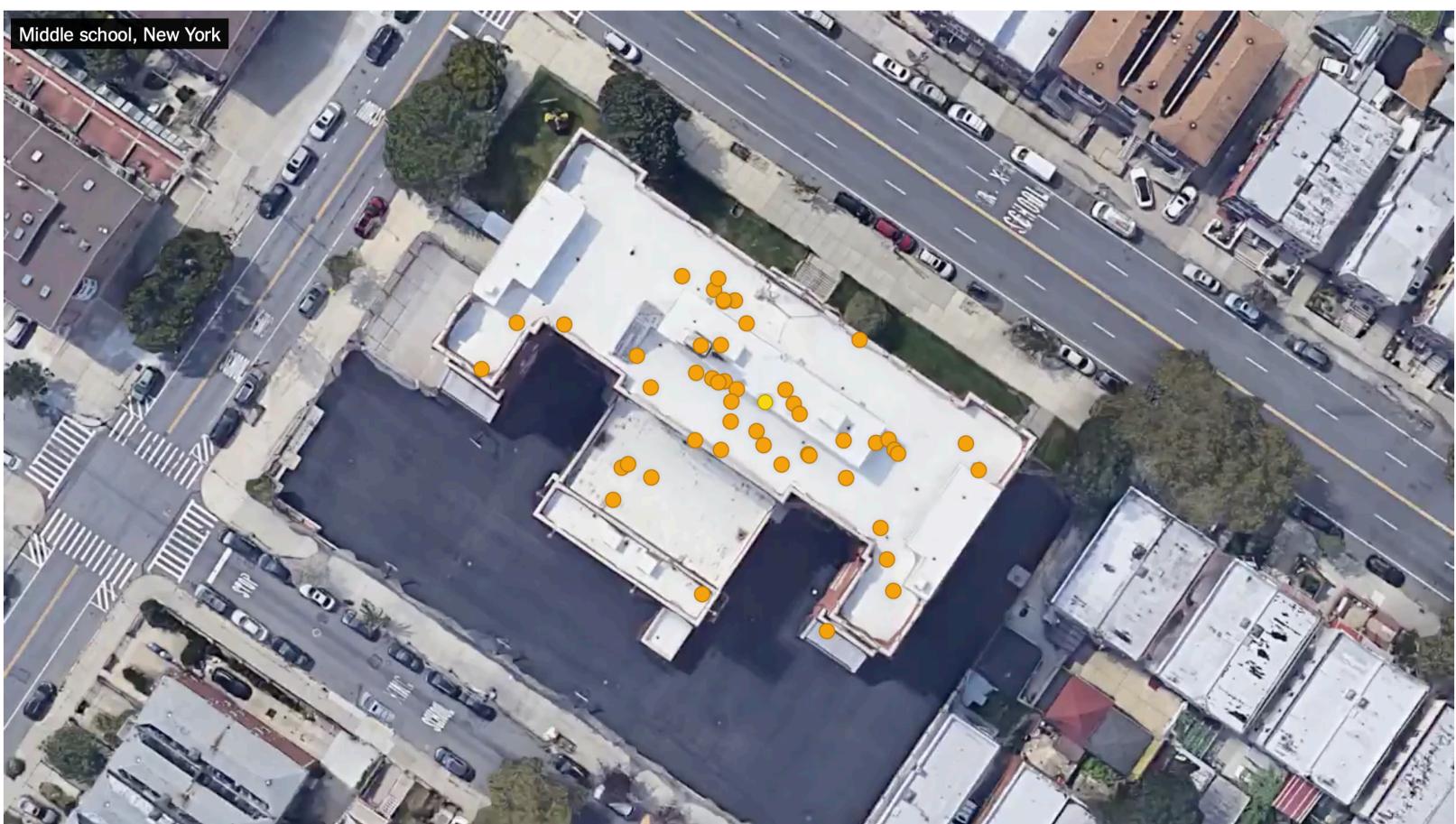
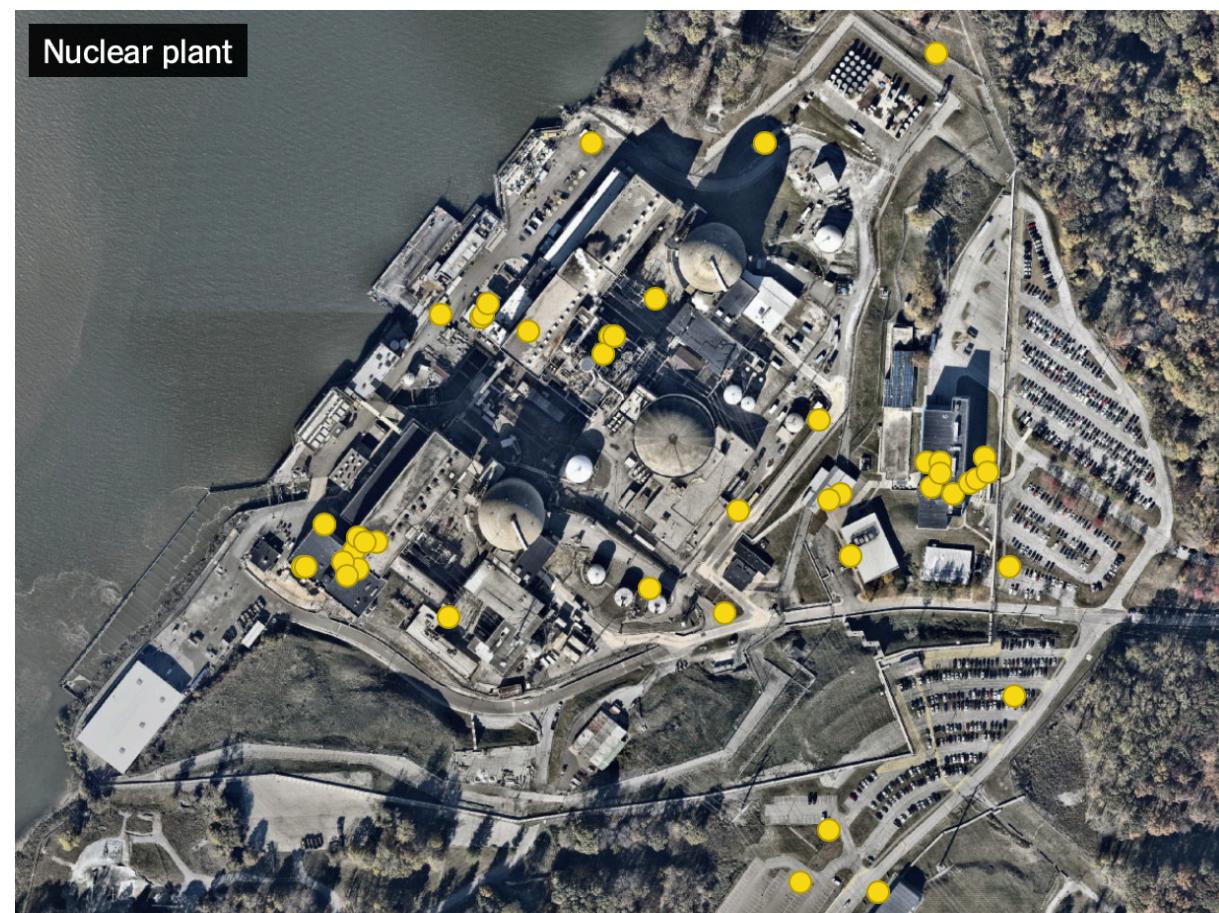
NY Times - Opinion - The Privacy Project (2019-12-20)

Some uses of location data: reconnaissance, recruitment (by foreign powers), social engineering, extortion, kidnapping, ...



NY Times - Opinion - The Privacy Project: How To Track President Trump (2019-12-20)

Your Apps Know Where You Were Last Night, and They're Not Keeping It Secret (NY Times 2018-12)



Lisa Magrin



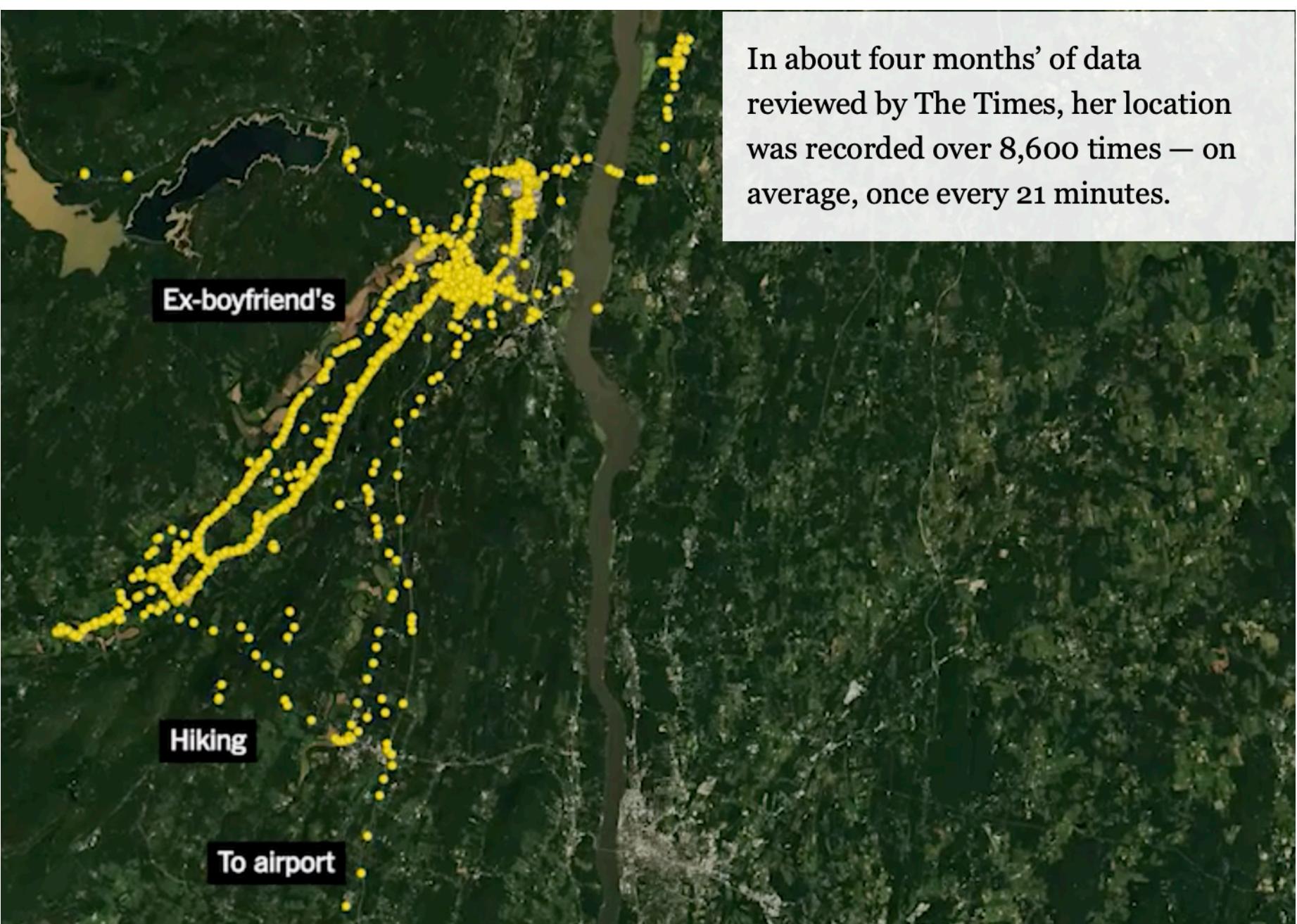
Where she lives



Where she works



Where she has been



Smartphone location tracking industry

Your smartphone can broadcast your exact location thousands of times per day, through hundreds of apps, instantaneously to dozens of different companies ([NY Times](#))
Each of those companies has the power to follow individual mobile phones wherever they go, in near-real time.

Thanks to the apps on your smartphone, **everyone** can know:

Where you live, where you work, where and for how long you take your lunches

All of the places you have ever been in the past few years (the data is not deleted)

For years past: where you have been, who you have met at those places

Each time you seek medical services (e.g. mental services, hospitalization, length of stay)

Whether you ever attended or currently attending Alcoholics Anonymous meetings

When and where you spend the night when not at home

Whether you check in and out of hotels (and which ones) at irregular times

Your precise movements 24 / 7 today and for several years in the past

Whether you attended some demonstration at a specific location and time, ...

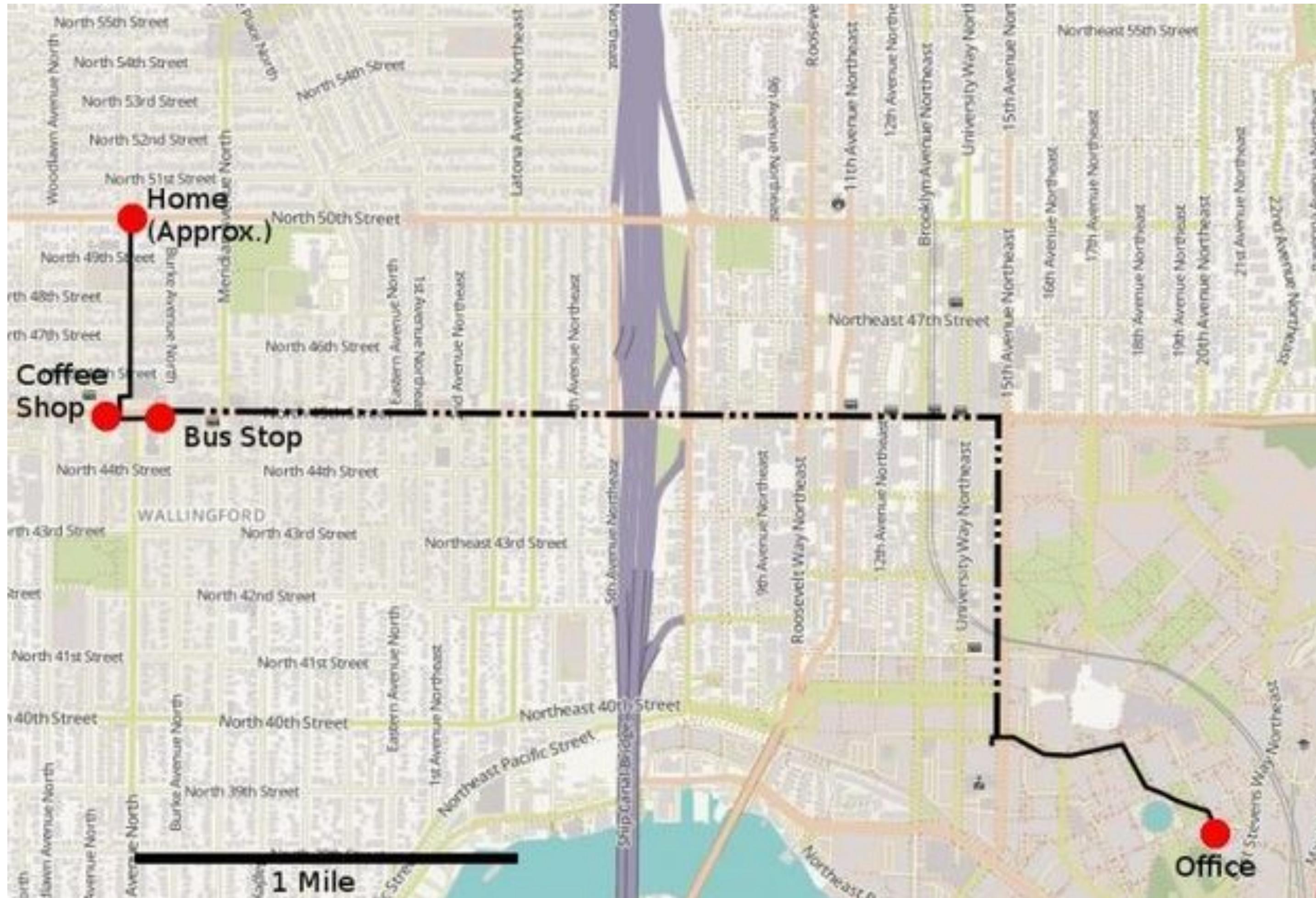


Hong Kong
How Your Phone Betrays Democracy ([NY Times](#))

This data is held by private companies and is for sale...

Proactively Track Anyone Anywhere using Adverts

Track Anyone, Anywhere Just By The Adverts They Receive (2017)



IDFA (ID For Advertisers): one ID per mobile phone

IDFV (ID For Vendors): one ID per mobile application vendor

iOS: Settings>Privacy>{Advertising, Location Services}

Android: Settings>Privacy>Advanced/Ads

Scheme: hyperlocal advertisements sent to one given phone

What: Track a person's location, apps being used,
know when a person leaves his/her home, surveillance, ...

Steps to perform the attack:

1. Get the mobile Advertising ID of targeted phone
(e.g. when joining an insecure network like a coffee shop wifi)
2. Sign up with a large internet advertising company
3. Set adverts to specific locations, then watch in real-time
to see if the person has received the ad or not
4. Create a network of those ads to map a person's movements

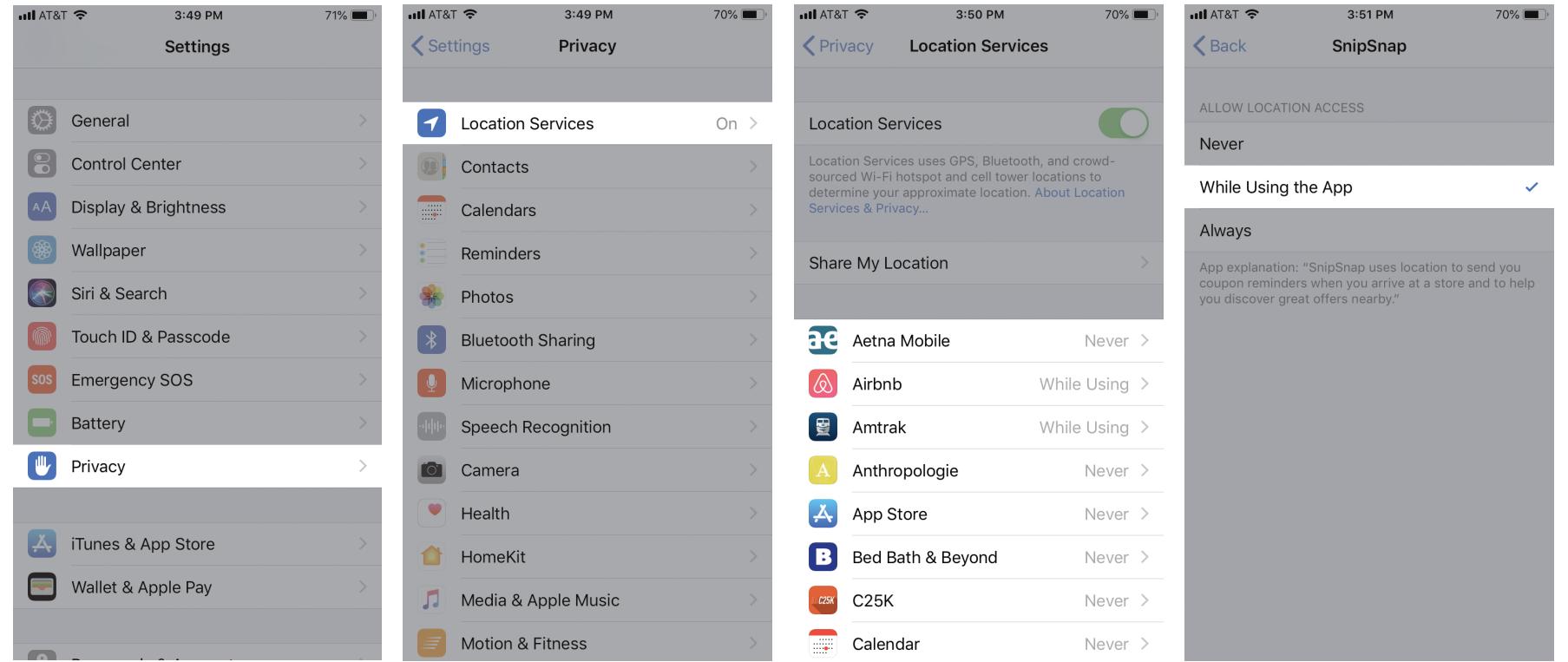
Specifics:

The target (person) does not need to interact with the ad

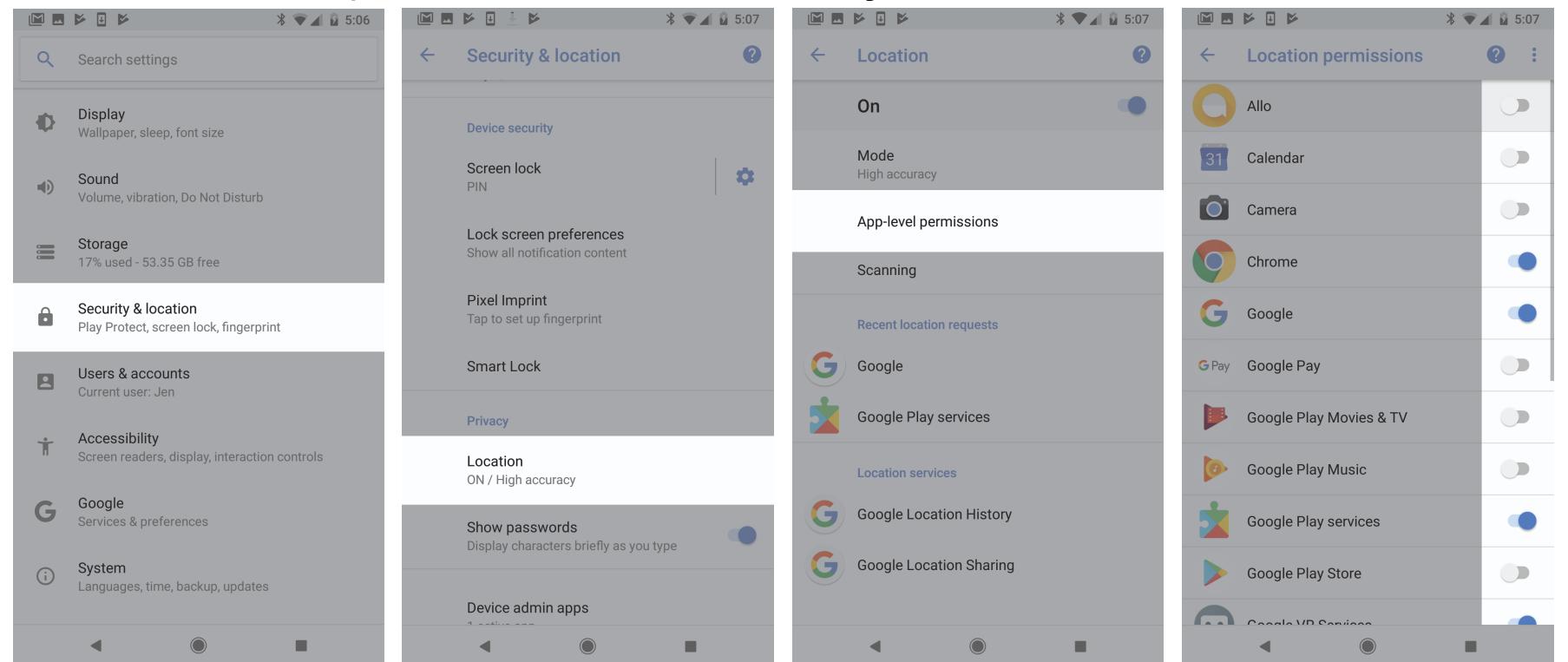
Location Data: Steps towards STOPPING apps from tracking your location

How to Stop Apps From Tracking Your Location ([NY Times 2018-12](#))

iPhone: 3 options: Never, While using the app, Always



Android: 2 options: Never, Always



Resetting the ID for Advertisers

[IDFA \(ID For Advertisers\)](#): one ID per mobile phone

[IDFV \(ID For Vendors\)](#): one ID per mobile application vendor

iOS: Settings>Privacy>{Advertising, Location Services}

Android: Settings>Privacy>Advanced/Ads

Freaked Out? 3 Steps to Protect Your Phone ([NY Times 2019-12](#))

1. Stop sharing your location with apps

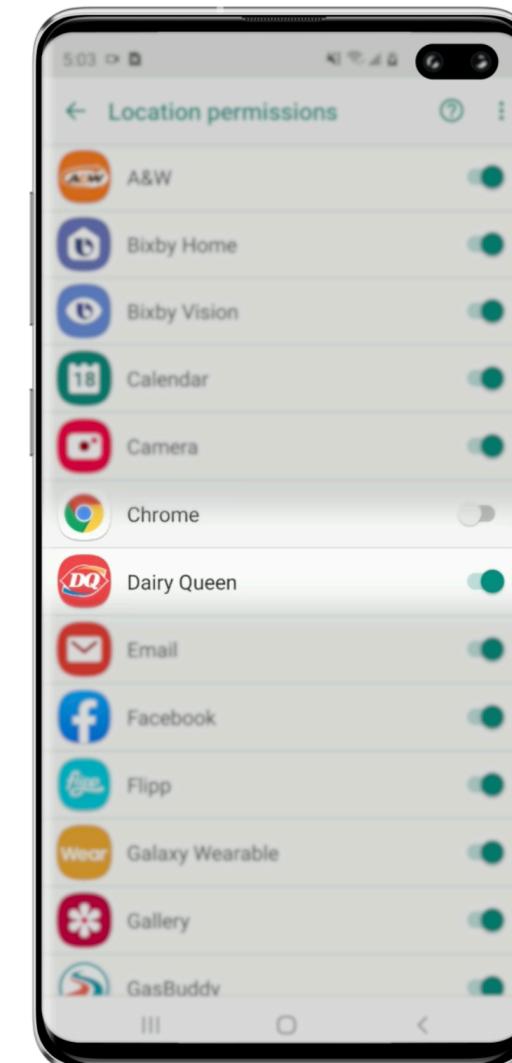
iPhone

To turn off location sharing, go to [Settings > Privacy > Location Services](#). You can choose when to share your location for each app.

Android

To turn off location sharing, go to [Settings > Biometrics and security > App permissions > Location](#). You can choose whether to share your location for each app.

2. Disable your mobile ad ID



Go to [Settings > Privacy > Advertising](#) and turn on [Limit Ad Tracking](#).

Go to [Settings > Google > Ads](#) and then turn on [Opt out of Ads Personalization](#).

3. Prevent Google from storing your location

<https://myaccount.google.com/activitycontrols/location>

4. Understand location tracking is hard to avoid

Cell phone companies ([T-Mobile](#), [Sprint](#), [AT&T](#)) selling location data to bounty hunters

Metadata may reveal a lot about an individual

Risk of re-identification

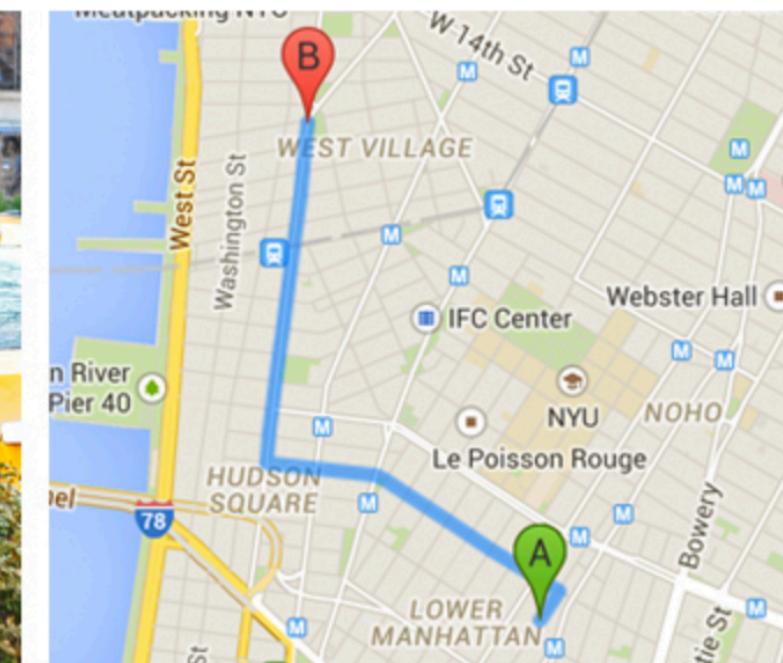
Personally Identifiable Information (PII): names, home addresses, phone numbers, ...

Lack of PII does not make it anonymous (NY Times) nor safe to release to the public or third parties

Public NYC Taxicab Database (2014)

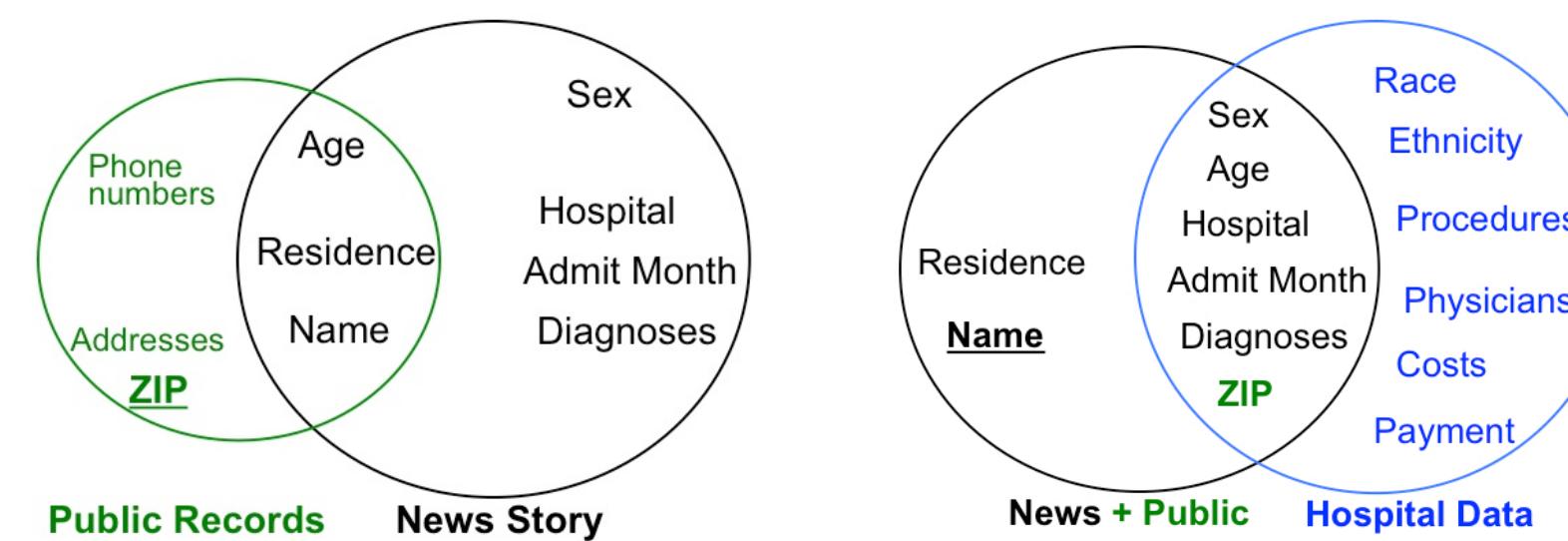


ASHLEE SIMPSON



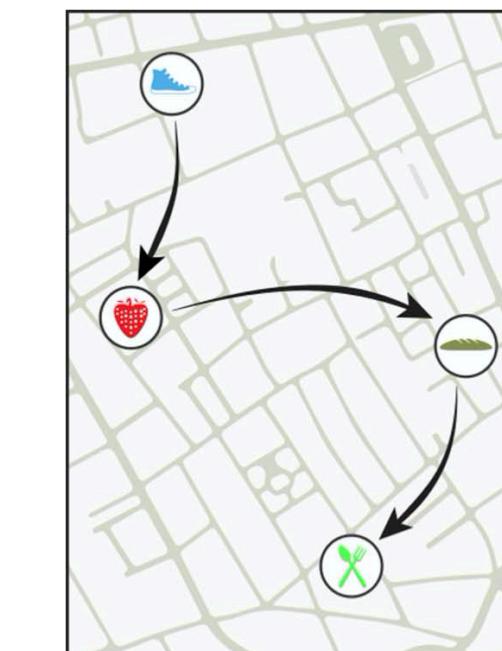
Matching Known Patients to Health Records

(Washington State - 2013)



The State of Washington sells patient-level health data for \$50: all hospitalizations occurring in the State in a given year, including patient demographics, diagnoses, procedures, attending physician, hospital, summary of charges, how the bill was paid.

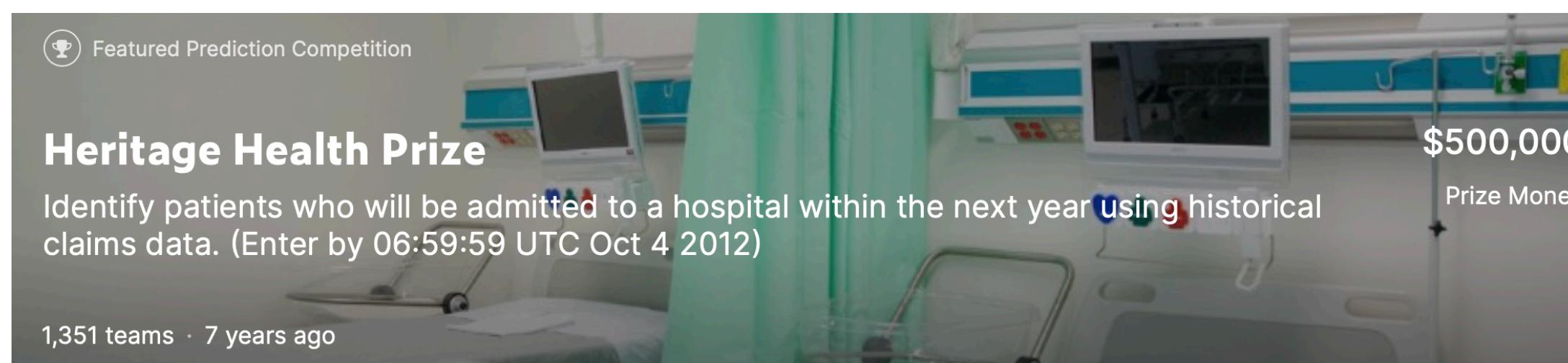
Credit card study blows holes in anonymity (2015)



shop	user_id	time	price	price_bin
blue square icon	7abc1a23	09/23	\$97.30	\$49 – \$146
red strawberry icon	7abc1a23	09/23	\$15.13	\$5 – \$16
green circle icon	3092fc10	09/23	\$43.78	\$16 – \$49
green circle icon	7abc1a23	09/23	\$4.33	\$2 – \$5
swim icon	4c7af72a	09/23	\$12.29	\$5 – \$16
mountain icon	89c0829c	09/24	\$3.66	\$2 – \$5
green X icon	7abc1a23	09/24	\$35.81	\$16 – \$49

Heritage Health Prize (Kaggle):

An adversarial Analysis of the Reidentifiability of the Heritage Health Prize Dataset (2011)



No silver bullet: De-identification still doesn't work (Princeton 2014)

No known effective method to anonymize location data

De-identification inadequate for high-dimensional data

Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule

Guidelines for De-identification of Health Information (HHS.gov 2014)

Genetic Privacy



(2006) Personal Genome Project

aims to sequence the genotypic and phenotypic information of 100,000 informed volunteers and display it publicly online in an extensive public database

(2013) Identifying Participants in the Personal Genome Project by Name (Sweeney et al.)

(2011) About half of the publicly available profiles had date of birth, gender and 5-digit ZIP

Recovering around 90% of the targeted profiles (names, medical and genomic info, medications, diseases)

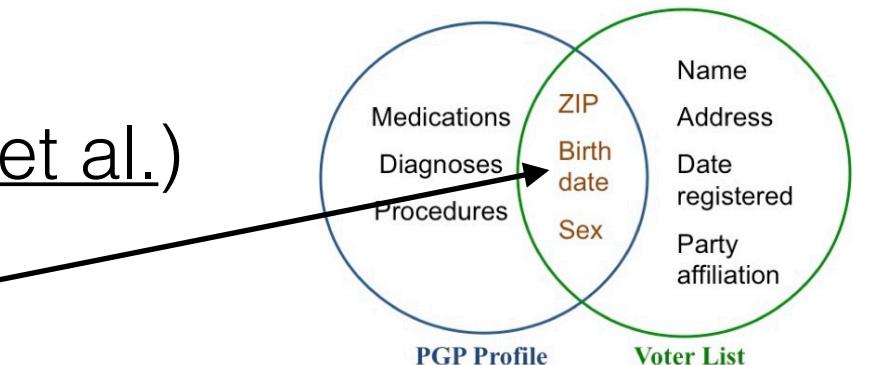
Linkage attack: voter registration data, mining for names hidden in attached documents, **demographics**

(2013) Identifying Personal Genomes by Surname Inference (Gymrek et al.)

Attack: profiling short tandem repeats on the Y chromosome (Y-STRs)

Linkage: querying recreational genetic genealogy databases

DNA genetic sequencing: **Health related risk assessments, testing for ancestry**



Risks if the data gets in the wrong hands:

Job: being denied a work position if genetic sequencing indicates risk for a given disease

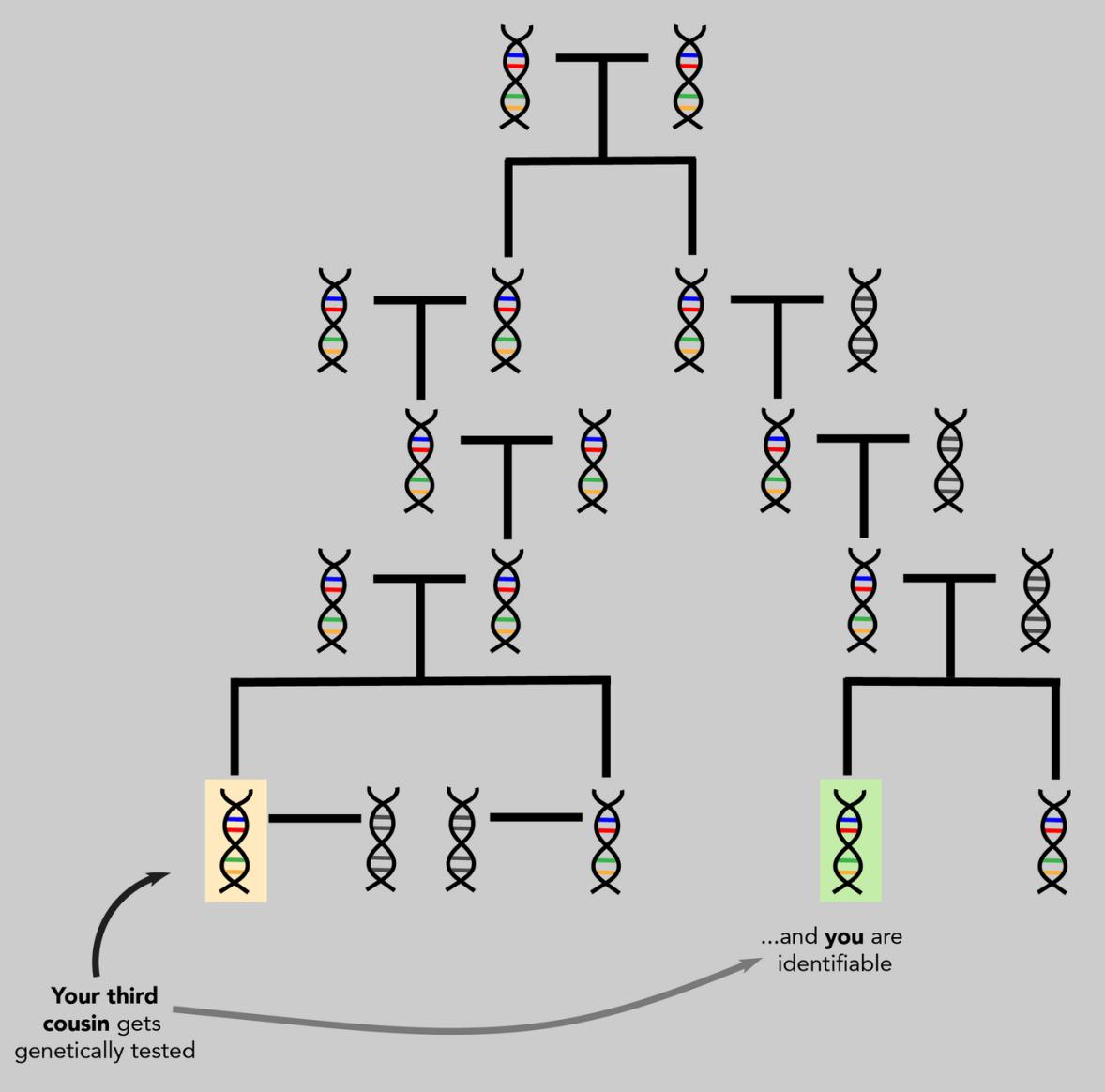
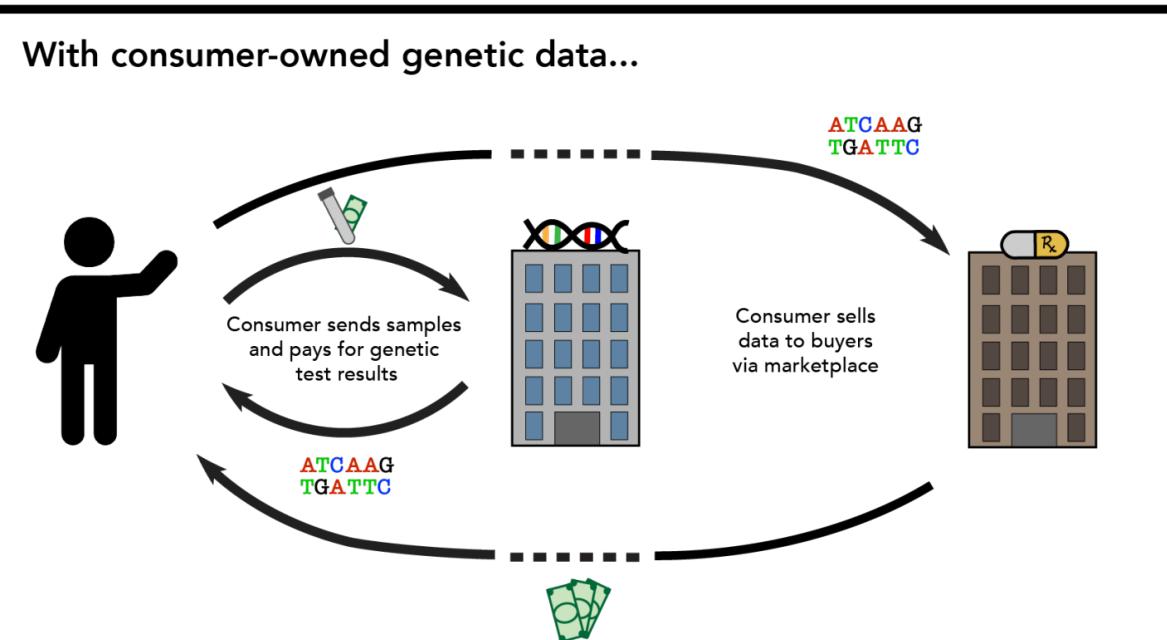
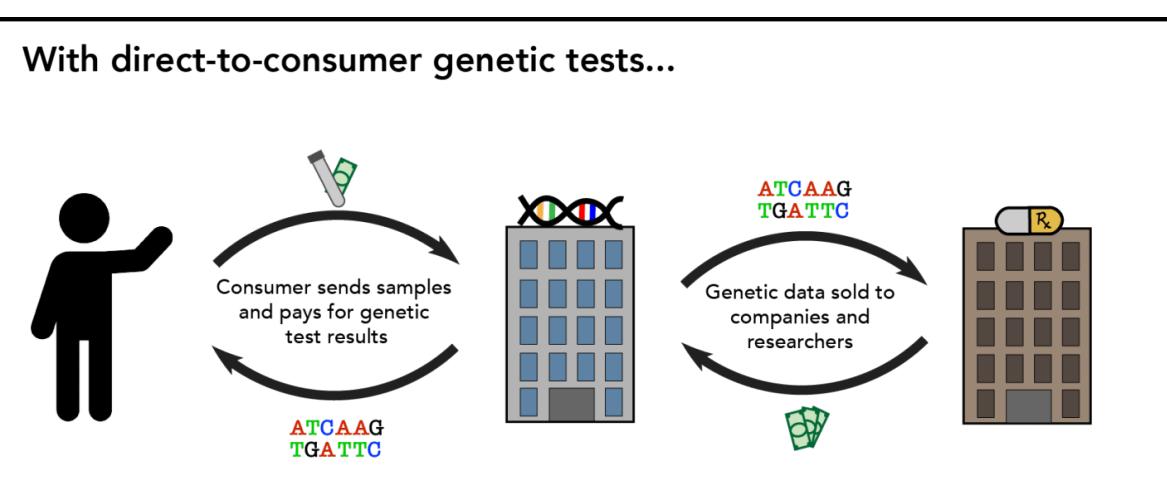
Health insurance: being denied insurance, higher premiums, ...

Wired 2017: To Protect Genetic Privacy, Encrypt Your DNA

iDASH 2017: (integrated Data Analysis, Anonymisation and Sharing conference): Secure Genome Analysis competition

Wired 2019: Urgent need for new genetic privacy laws

Understanding Ownership and Privacy of Genetic ([Harvard 2018](#))



Apps for Mobile Health (mHealth)

 **HealthEngine** Australia's most popular medical appointment booking app sharing clients' personal information with lawyers (2018)

University of Toronto study: 19 out of 24 of most popular mHealth applications found to be sharing user data to various companies (2019)

Table 2 | Types and frequency of user data shared with third parties in traffic analysis

User data type	Explanation	No (%) of apps sharing*
Device name	Name of device (eg, Google Pixel)	15 (63)
OS version	Version of device's Android operating system	10 (42)
Browsing	App related activity performed by user (eg, view pharmacies, search for medicines)	9 (38)
Email†‡	User's email address	9 (38)
Android ID†‡	Unique ID to each Android device (ie, used to identify devices for market downloads)	8 (33)
Drugs list‡	List of drugs taken by user	6 (25)
Name/Last name†‡	User's name and/or last name	5 (21)
Time zone	Time zone in which device is located (eg, GMT+11)	5 (21)
Connection type	Cellular data or wi-fi	4 (17)
Medical conditions‡	Users' medical conditions (eg, diabetes, depression)	4 (17)
Birthday‡	User's date of birth	3 (13)
Device ID†‡	Unique 15 digit International Mobile Equipment Identity code of device	3 (13)
Sex	User's sex	3 (13)
Carrier	Mobile network operator, provider of network communications services (eg, AT&T)	2 (8)
Country	Country in which device is located (eg, Australia)	2 (8)
Coarse grain location‡	Non-precise location. Usually city in which device is located (eg, Sydney)	2 (8)
Drug instructions	Instructions related to user's drugs (eg, orally, with food)	2 (8)
Drug schedule	Times for drug administration (eg, 8 pm, in the morning)	2 (8)
Personal conditions‡	Users' personal conditions (eg, smoker, pregnant)	2 (8)
Personal factors‡	Includes user's anthropometric measurements or vital signs (eg, height, weight, blood pressure)	2 (8)
Symptoms‡	User's symptoms (eg, headache, nausea)	2 (8)
Doctor's name‡	Name of the user's doctor	1 (4)
Doses‡	Dose of user's drug (eg, 100 mg aspirin per day).	1 (4)
Feelings	User's current feelings (eg, happy, sad, anxious)	1 (4)
Pharmacy name‡	Information about user's favourite pharmacies (eg, name, location)	1 (4)

*Total number is 24; percentages do not add to 100% as apps could share multiple types of user data.

†Unique identifier.

‡May be considered personal data under the General Data Protection Rules—that is, “any information relating to an identified or identifiable natural person.”¹⁸

Table 3 | Data sharing practices of apps

No of installs* and apps	No of different types of user data shared†	No of unique transmissions (type/entity)‡	No of unique recipients§
500-1000:			
Dental Prescriber	0	0	0
Medsmart Meds & Pill Reminder App	14	25	4
myPharmacyLink	5	5	2
1000-5000:			
DrugDoses	0	0	0
MediTracker	4	6	3
MyMeds	5	8	3
5000-10 000:			
CredibleMeds	1	2	2
Med Helper Pro Pill Reminder	0	0	1
Nurse's Pocket Drug Guide 2015	0	0	3
Pedi Safe Medications	0	0	0
10 000-50 000:			
MIMS For Android	3	6	2
50 000-100 000:			
ListMeds-Free	0	0	0
MedicineWise	5	9	5
100 000-500 000:			
Dosecast-Medication Reminder	9	16	3
Lexicomp	3	6	3
MedAdvisor	8	20	3
My PillBox(Meds&Pill Reminder)	0	0	0
Nurse's Drug Handbook	4	9	5
Pill Identifier and Drug list	5	10	4
500 000-1 000 000:			
UpToDate for Android	5	11	3
1 000 000-5 000 000:			
Ada-Your Health Companion	15	27	13
Drugs.com	5	5	2
Epocrates Plus	8	14	3
5 000 000-10 000 000:			
Medscape	7	21	8

*As reported in Google Play store at time of sampling (November 2017).

†As detected in traffic analysis of 28 possible types.

‡As detected in traffic analysis and defined as sharing of unique type of data with an external entity—for example, app shares Device Name unique transmissions.

§Identified in traffic and privacy policy analysis.

Data Protection Regulation

European Union (2018): General Data Protection Regulation (GDPR)

Aims primarily to give control to individuals over their personal data

Canada (2019): Personal Information Protection and Electronic Documents Act (PIPEDA) (revised 2019)

Privacy Act is being reviewed (2019)

California (2020): California Consumer Privacy Act (CCPA) (The Economist - 2020-06-02)

Grants consumers in California the right to ask companies to disclose, transfer or delete personal data held about them

Lets consumers sue companies that misuse or mishandle such data

The law makes life harder for companies that make money from trading personal data

Many websites use cookies or trackers that send visitor data to other companies via data exchanges: considered as **sales!**

Anonymized data may still be considered as “personal” (re-identification risks)

(even with no cash exchange)

—> ADD “opt out” button to allow website visitors to block the sale of their data

Data Protection Regulation

California Consumer Privacy Act (CCPA) and GDPR Comparison Chart (Practical Law - 2019)

	CCPA	GDPR	Comparison	Practical Law Resources and Citations
Anonymous, Deidentified, Pseudonymous, or Aggregated Data	<p>The CCPA does not restrict a business's ability to collect, use, retain, sell, or disclose a consumer information that is deidentified or aggregated.</p> <p>However, the CCPA establishes a high bar for claiming data is deidentified or aggregated</p>	<p>Pseudonymous data is considered personal data.</p> <p>Anonymous data is not considered personal data.</p>	<p>The CCPA and GDPR pseudonymization definitions are very similar and both require technical controls to prevent reidentification to qualify.</p>	CCPA Cal. Civ. Code §§ 1798.140(a), (h), (o), (r), and 1798.145(a)(5). Practice Note, California Privacy and Data Security Law: Overview: Personal Information under CCPA (6-597-4106).
	<p>Pseudonymous data may qualify as personal information under the CCPA because it remains capable of being associated with a particular consumer or household. However, the statute does not clearly categorize or exclude pseudonymous data as personal information.</p>	<p>While the GDPR does not mention deidentified data, the CCPA definition is similar to GDPR's concept of anonymous data.</p>	<p>The CCPA primarily discusses pseudonymization in the context of using personal information collected from a consumer for other purposes, for research. It does not appear to help businesses generally avoid the CCPA's requirements.</p> <p>At this point, it is unclear how different the position under the GDPR is.</p>	GDPR Article 4(5). Practice Note, Anonymization and Pseudonymization under the GDPR (W-007-4624).

Pseudonymization: personally identifiable information (PII) fields are replaced by one artificial identifier, or pseudonyms

Pseudonymized data can be restored to its original state with the addition of information which then allows individuals to be re-identified

Anonymized data can never be restored to its original state

Techniques aiming to keep data private

1. Data anonymization - De-identification
2. Differential Privacy
3. Federated Learning
4. Encrypted: Multi Party Computation
5. Encrypted: Homomorphic Encryption

1. Data anonymisation (De-identification)

Original document

Five people have been charged in connection with an anti-fracking protest at an oil refinery in Grangemouth.

The protesters climbed on to the chimney at Ineos petrochemicals plant on 9 October.

Police Scotland said two men, aged 21 and 24, had been charged in connection with the protest and would appear at Falkirk Sheriff Court on Monday.

Three women, aged 40, 44 and 48, were charged with breach of the peace and will also appear at court.

The 40-year-old was also charged with breach of the peace.

The other three were arrested in connection with a protest outside a hotel in Glasgow on 10 October.

Anonymized document

Five people have been charged in connection with an anti-fracking protest at an oil refinery in LOCATION.

The protesters climbed on to the chimney at ORGANISATION plant on DATE.

ORGANISATION said two men, aged DD and DD, had been charged in connection with the protest and would appear at LOCATION on DATE.

NUMBER women, aged DD, DD and DD, were charged with breach of the peace and will also appear at court.

The DD-year-old was also charged with breach of the peace.

The other three were arrested in connection with a protest outside a hotel in CITY on DATE.

[OpenAI GPT-3 Language Model](#)

Electronic Health Record (EHR)

Original EHR

Name	Age	Sex	Zipcode	Disease
James Smith	27	M	10024	Pneumonia
Mary Johnson	31	F	11211	Diabetes
John Brown	52	M	11238	Liver disease
Robert Miller	67	M	11356	Scleroderma
Patricia Wilson	41	F	10029	Diabetes
Richard Anderson	23	M	11417	Asthma
Susan Taylor	36	F	10925	Sclerosis

“Anonymized” EHR

Name	Age	Sex	Zipcode	Disease
P0001	30	M	100**	Pneumonia
P0002	30	F	112**	Diabetes
P0003	50	M	112**	Liver disease
P0004	70	M	113**	Scleroderma
P0005	40	F	100**	Diabetes
P0006	20	M	114**	Asthma
P0007	40	F	109**	Sclerosis

Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule

Guidelines for De-identification of Health Information (HHS.gov 2014)

Re-identification potential

Anonymising and sharing individual patient data - El Emam, Rodgers, Malin - BMJ 2015

Table 2 | Changes in probability of re-identification of anonymised data in BORN (Ontario birth registry dataset) for different levels of generalisation of quasi-identifiers

Scenario	Mother's date of birth or age	Baby's date of birth	Mother's postal code	Baby's sex	Probability of re-identification*
S1	Year	day, month, year	3 character	Unchanged	0.973
S2	Year	month, year	3 character	Unchanged	0.677
S3	Age in 5-year groups	month, year	3 character	Unchanged	0.327
S4	Age ≤19, 20–30, 30–40, >40	month, year	3 character	Unchanged	0.23
S5	Age ≤19, 20–30, 30–40, >40	month, year	1 character	Unchanged	0.007
S6	Year	month, year	1 character	Unchanged	0.034
S7	Age in 5-year groups	quarter, year	3 character	Unchanged	0.152
S8	Age ≤19, 20–30, 30–40, >40	quarter, year	3 character	Unchanged	0.1

*Probability was measured using the average re-identification risk metric defined elsewhere.⁷

Re-identification - De-anonymization

Linkage attack: using auxiliary information from other databases to compromise privacy of a given database ([Cynthia Dwork 2007](#))

Netflix prize dataset

100,480,507 movie ratings, from 480,189 users, on 17,770 movies

CustomerID replaced with a randomly assigned user ID

movielID:
userID, rating, date

9211:
1277134,1,2003-12-02
2435457,2,2005-06-01
2338545,3,2001-02-17
2218269,1,2002-12-27

Auxiliary information: International Movie Database ([IMDb.com](#))

[Narayanan & Shmatokov - 2008](#): Robust De-anonymization of Large Sparse Datasets

very little auxiliary information needed to de-anonymize an average subscriber record from the Netflix Prize dataset

“With 8 movie ratings (of which 2 may be completely wrong) and dates that may have a 14-day error, 99% of records can be uniquely identified in the dataset”

With 2 movie ratings and dates that may have a 3-day error, 68% of records can be uniquely identified in the dataset”

Re-identification - De-anonymization

DataPrivacyLab.org (Latanya Sweeney - Harvard U.)

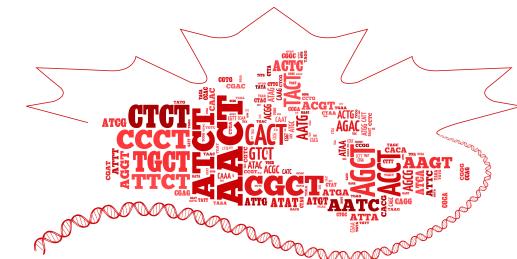
Re-identifying the personal medical records of the Governor of Massachusetts (including diagnoses and prescriptions) (1997)

Dataset: Group Insurance Commission" (GIC) data

Linkage attack: voter registration data



Personal Genome Project (2006): Recovering around 90% of the profiles (names, medical and genomic info, medications, diseases) (2013)



Linkage attack: voter registration data, mining for names hidden in attached documents, **demographics**

aboutmyinfo.org: How unique am I? (enter Date Of Birth, Gender and Zipcode)

Gender: Male

ZIP Code: 90011 (pop. 103892)

Birthdate	4/14/1952	Easily identifiable by birthdate (about 1)
Birth Year	1952	Lots with your birth year (about 267)
Range	1952 to 1953	Lots in the same age range as you (about 535)

==> Good old fashioned data anonymization is not sufficient...

2. Differential Privacy (DP)

Local Differential Privacy

Global Differential Privacy

Adding random noise to guarantee anonymity to a certain degree

Use Case 1: Survey population for sensitive issue (e.g. crime, disease)

Social Sciences Project: what is ratio of a given population that tested positive for XXX?



Fact: people might not want to answer truthfully

Issues:

Trust in the database curator that he/she is not going to divulge the information

Trust that the database is not going to get hacked and data released to the public

Local Differential Privacy

(Google Blog 2014): Learning statistics with privacy, aided by the flip of a coin

Local Differential Privacy: ADD random noise to each answer (data point). Data aggregator not trusted.

Instructions: 2 coin FLIPS

FLIP a coin:

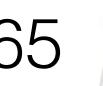
IF HEAD: answer truthfully

IF TAIL: answer randomly according to second coin flip:

HEAD: Yes

TAIL: No

Experiment: 100 participants

Responses: 35  65 

Observed_ratio: $35 / (35 + 65) = 35\% = \text{Observed_ratio}$

$\text{Observed_ratio} = 1/2 \text{ True_ratio} + 1/2 \text{ random_choice (50\%)}$

$\text{True_ratio} = 2 \text{ Observed_ratio} - \text{random_choice (50\%)}$

$\text{True_ratio} = 2 * 35\% - 50\% = 70\% - 50\% = 20\% = \text{True_ratio}$

Thanks to: The Law of Large Numbers

$\text{Observed_ratio} = \frac{1}{2} \text{True_ratio} + \frac{1}{2} \text{random_choice (50\%)}$

coin flips

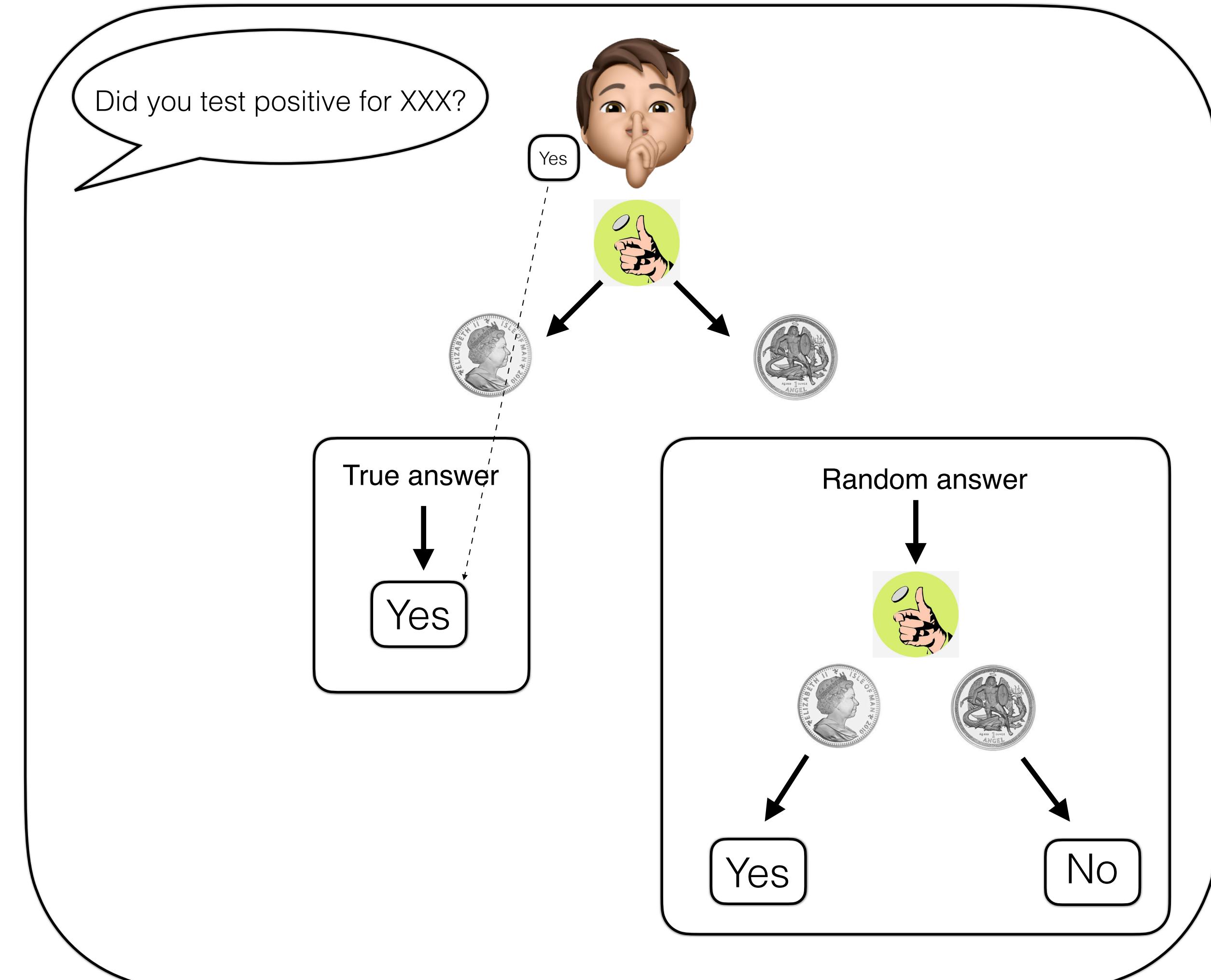


First coin: 10  40 

Second coin: 25  25 



Responses: 35  65 



Local Differential Privacy in the industry

Used by companies such as Apple and Google; e.g. to learn from mobile phone usage

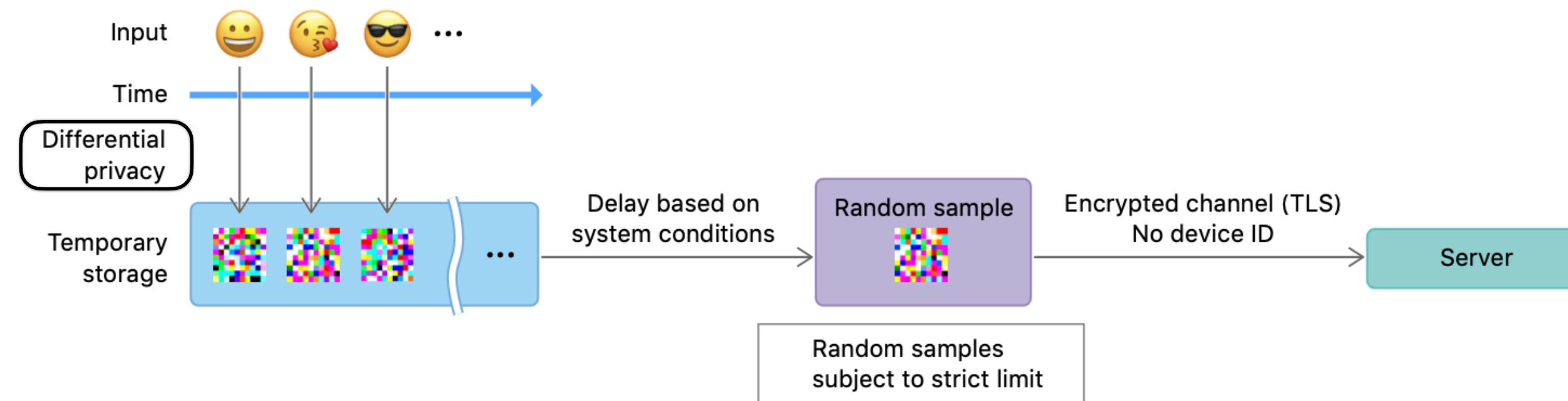
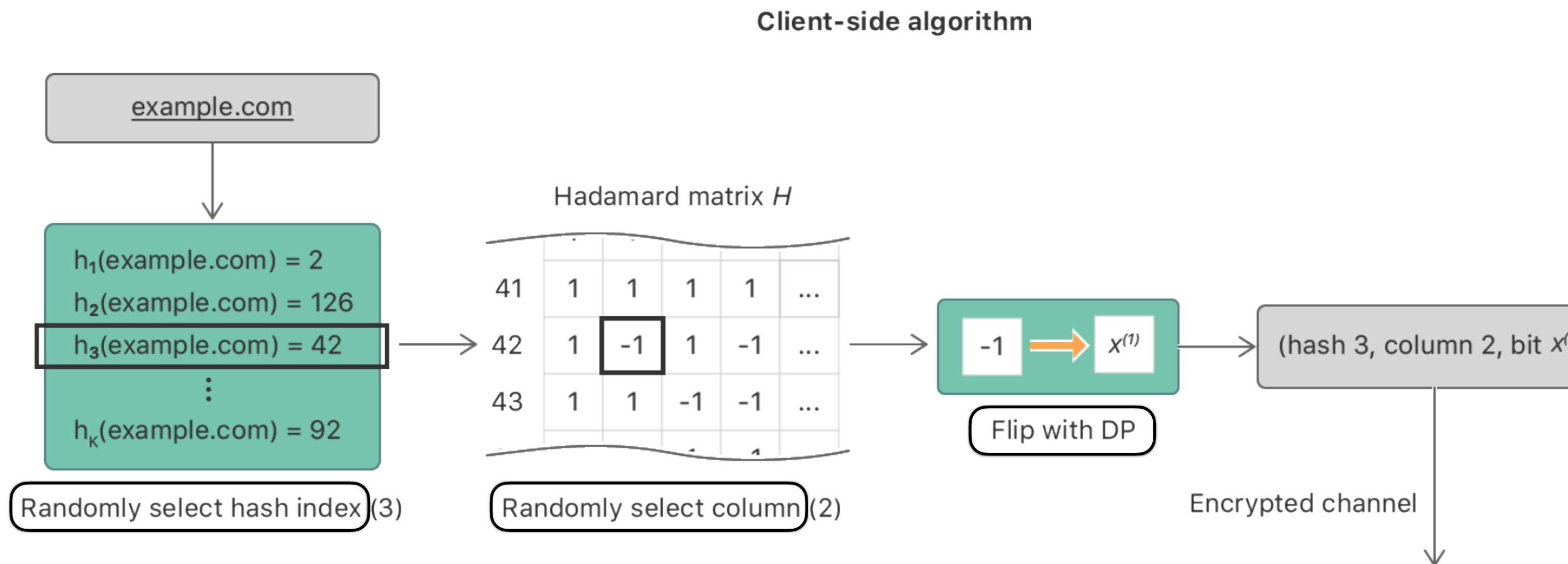
Applications: identifying popular emojis, discovering new words (improve auto-correction), identifying popular/troublesome web sites

K=65,536: K different hash functions

M=1024: Hash(string) \rightarrow [0, 1023]

$\epsilon=4$: privacy parameter

DP: Flip values with probability $\frac{1}{e^{\epsilon/2} + 1}$



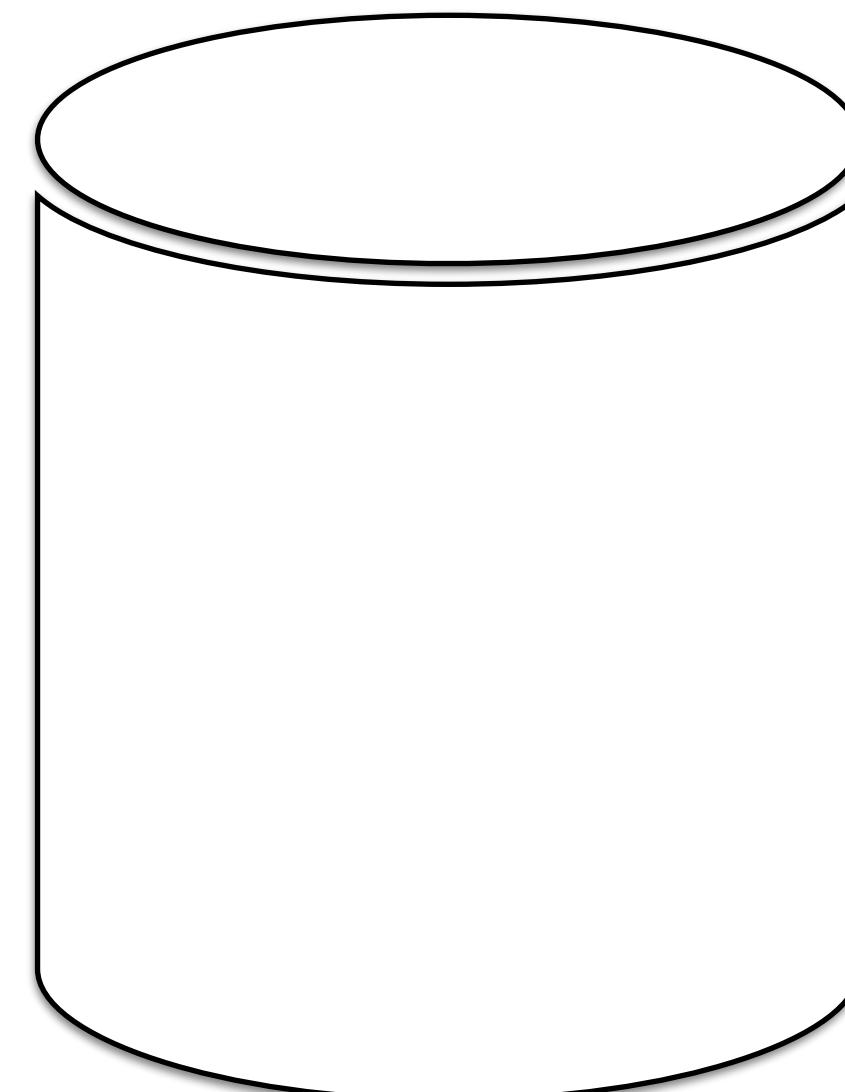
Use Case 2: Query Medical DB

Database: medical records with information about patient names, conditions, diseases, medications, ...

Access: Query database for **statistics** over patients. NOT individual patient data.

**** Search **:** Find if a given person e.g. “John Doe” has a given disease e.g. Asthma

Known facts (e.g. from social media): D.O.B. April 1, 1970, Male, living in Colchester VT (05401)



Medical Information DB

Query “Does John Doe has asthma?” —>  ONLY STATS QUERY ALLOWED

Query “number of persons with asthma” —> 173

Query “number of persons with asthma AND male” —> 91

Query “number of persons with asthma AND male AND zipcode == 054XX” —> 19

Query “number of persons with asthma AND male AND zipcode == 054XX AND DOB == 1970/04/XX” —> 1

==> John Doe has asthma!

Gender: Male	ZIP Code: 90011 (pop. 103892)
Birthdate	4/14/1952
Birth Year	1952
Range	1952 to 1953

Easily identifiable by birthdate (about 1)
Lots with your birth year (about 267)
Lots in the same age range as you (about 535)

aboutmyinfo.org

Definition of Global Differential Privacy

Keywords: statistical disclosure control, inference control, **privacy-preserving data mining**, private data analysis

Principal motivating scenario: statistical database

Original Definition

Privacy is preserved if: After the analysis, the analyzer doesn't know anything about the people in the dataset. They remain “unobserved”.

Cynthia Dwork & Aaron Roth's Definition (2014): The Algorithmic Foundations of Differential Privacy (book)

“Differential Privacy” describes a promise, made by a data holder, or curator, to a data subject, and the promise is like this:

“You will not be affected, adversely or otherwise, by allowing your data to be used in any study or analysis, no matter what other studies, data sets, or information sources, are available.”

← Linkage attack (Cynthia Dwork 2007)

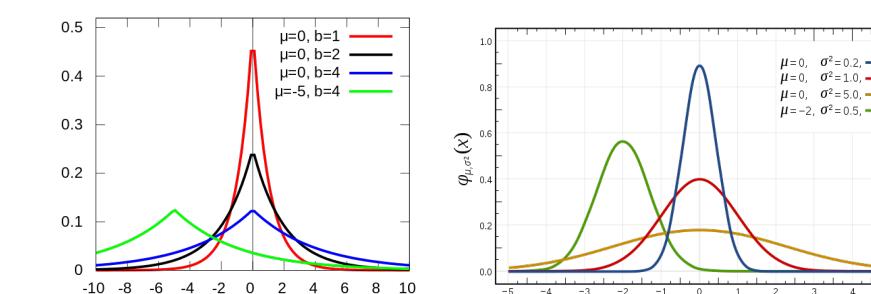
Global Differential Privacy (ϵ -differential privacy)

Trusted curator of the database

Add noise the **latest possible** in the processing chain (on the result of a query)

How much noise should we add? It depends on:

1. Type of noise added: Gaussian or Laplacian
2. Sensitivity of the query type (a “mean” query has lower sensitivity than a “count” query)
3. Desired epsilon and delta (a lower epsilon implies less data leakage and consequently greater variations (amplitude) in the noise added)



Differential Privacy for Deep Learning

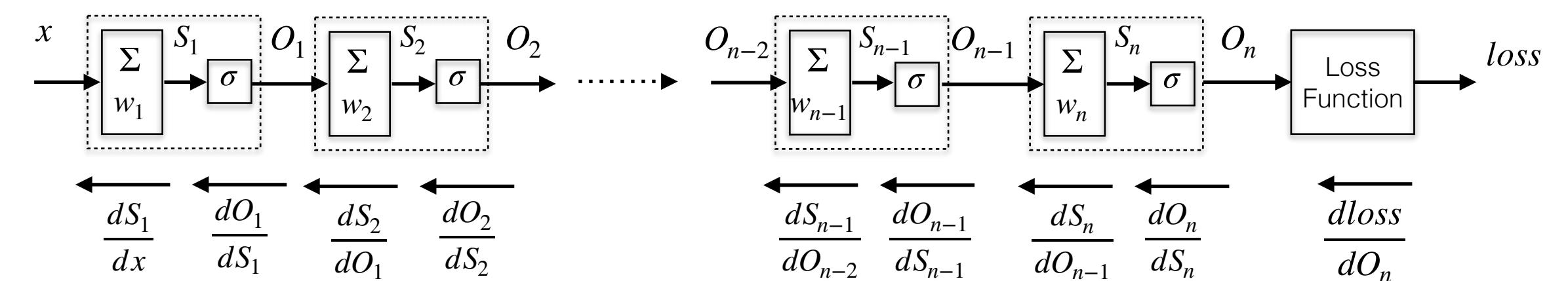
Perfect Privacy: Training a model should return the **same** model even if we remove any person from the training dataset.

Training of models requires **large representative** datasets that may contain sensitive information

Differentially Private Stochastic Gradient Descent algorithm ([Google 2016](#))

Gaussian Differential Privacy ([U. of Penn 2019](#))

Research still ongoing in that field...

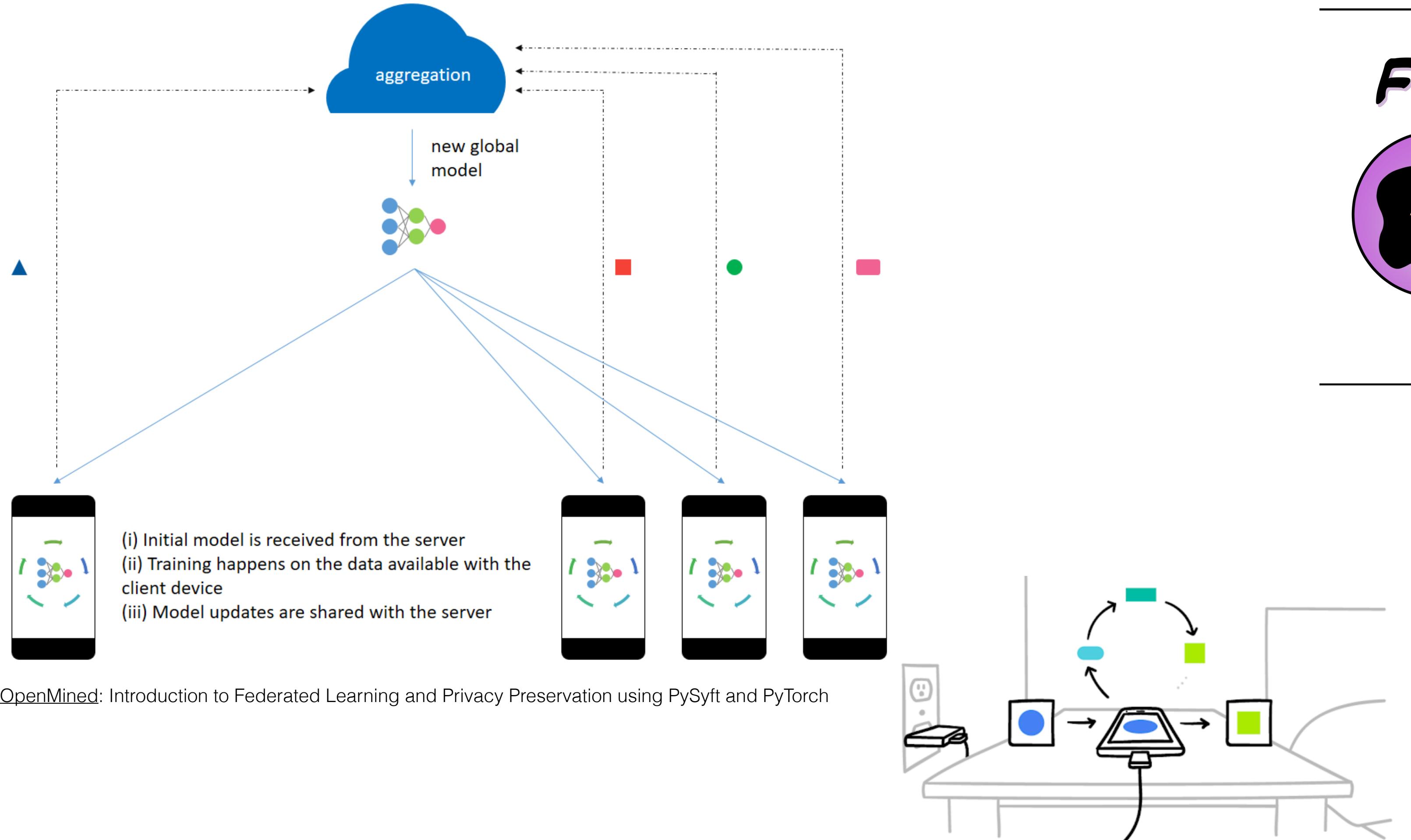


[Deep Learning \(2019\)](#)

3. Federated Learning

Federated Learning: Technique for training Machine Learning models on data which you do not have access to

Models trained on smartphones: next word prediction on keyboard, suggesting search query, face detection, voice recognition... Personal data does not leave the phone!



Aggregation of gradients (I/O intensive): Federated Stochastic Gradient Descent ([FedSGD](#))

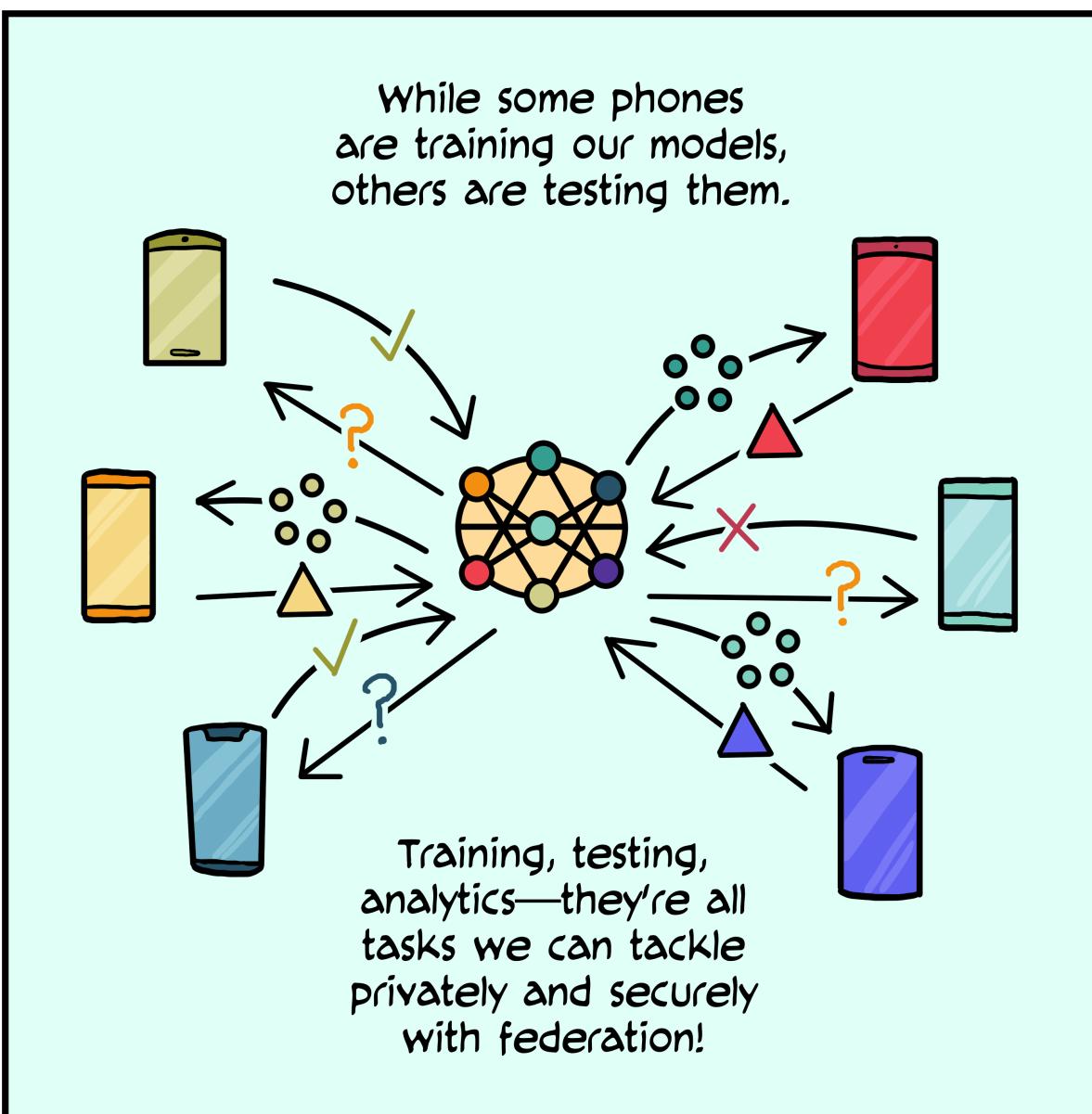
Aggregation of weights (CPU intensive): Federated Averaging ([FedAvg](#))

[Google Blog 2017](#)

Federated Learning



An online comic from Google AI



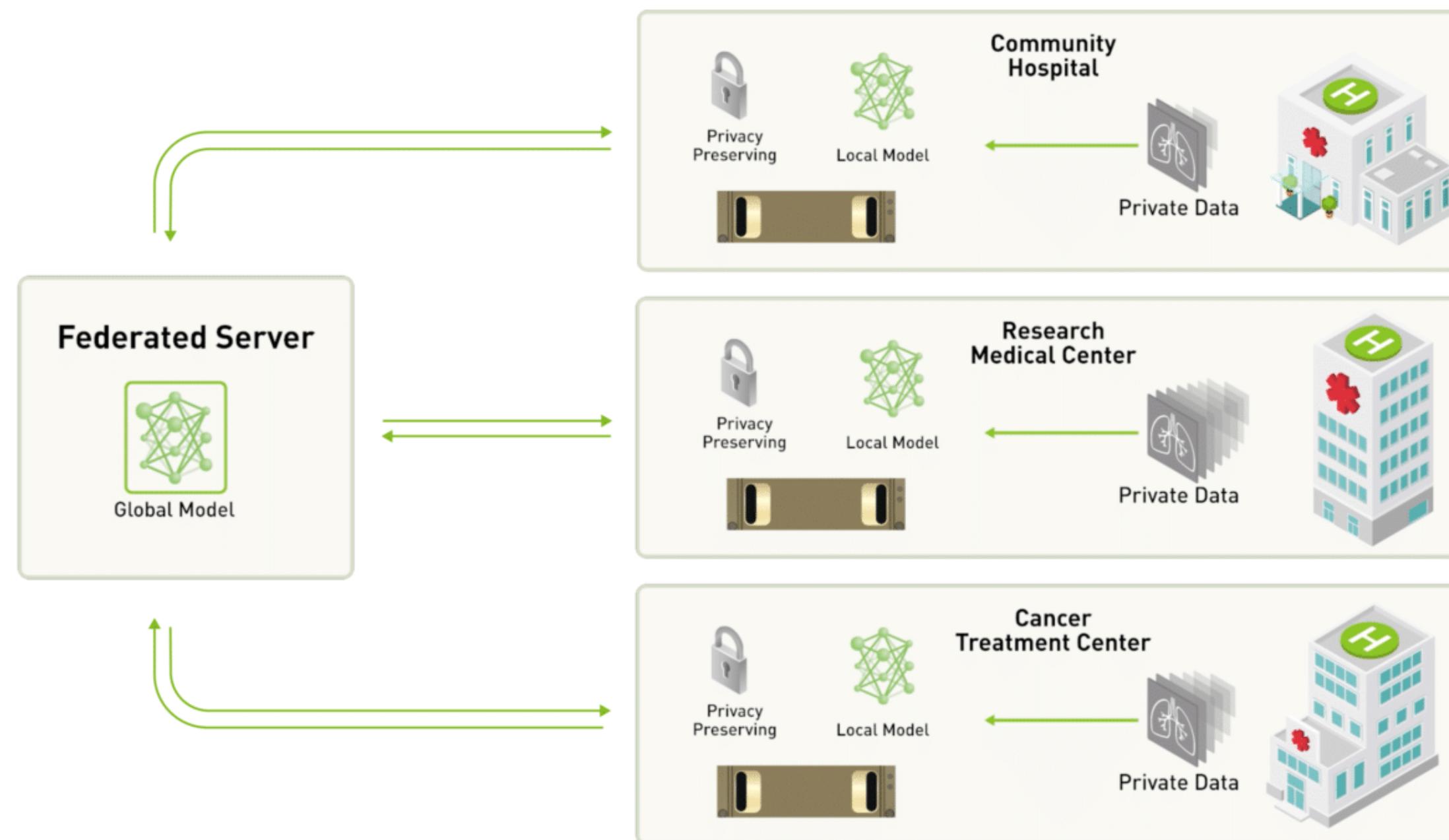
[federated.withgoogle.com](#)

Federated Learning: Use Cases: Transforming Healthcare

Federated learning makes it possible for AI algorithms to gain experience from a vast range of data located at different sites ([NVIDIA](#))

Federated learning decentralizes deep learning by removing the need to pool data into a single location

What is Federated Learning (NVIDIA 2019-10) ([video 2 mins](#))



Collaboration: NVIDIA and [King's College London AI Centre for Value Based Healthcare](#)

Dataset: [BraTS](#): Brain Tumor Segmentation database

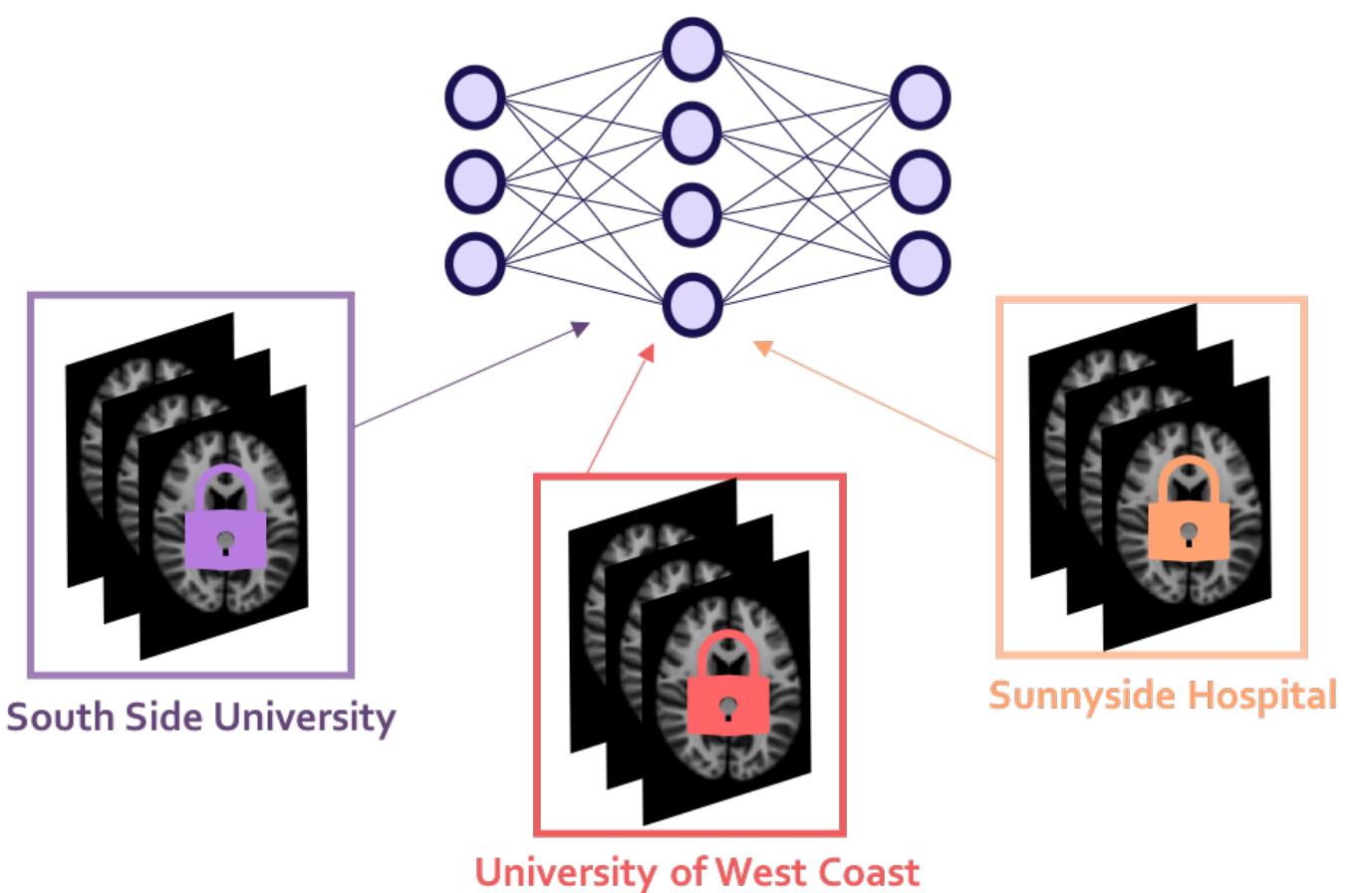
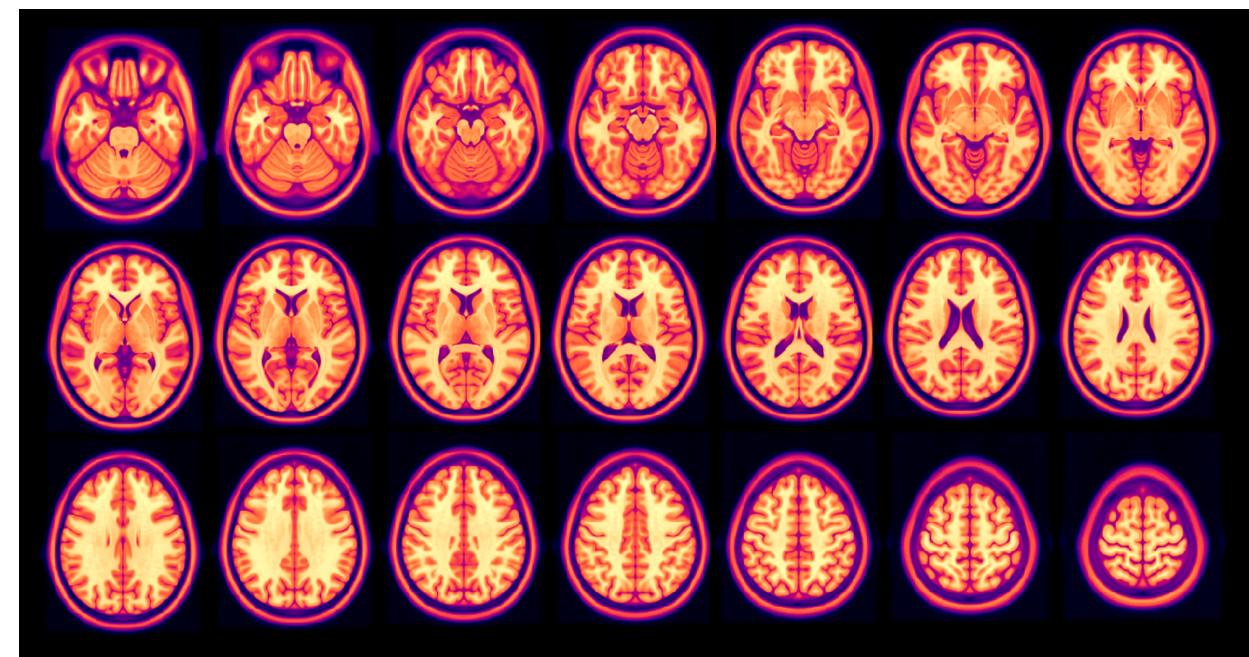
Simulating access to large, varied, high quality datasets spread over several hospitals

Privacy Preserving Federated Learning

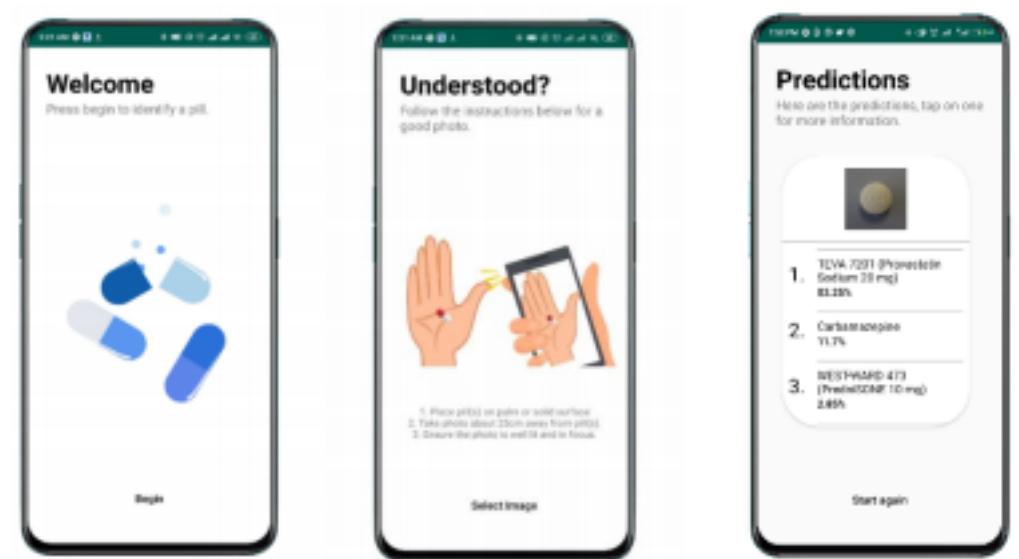
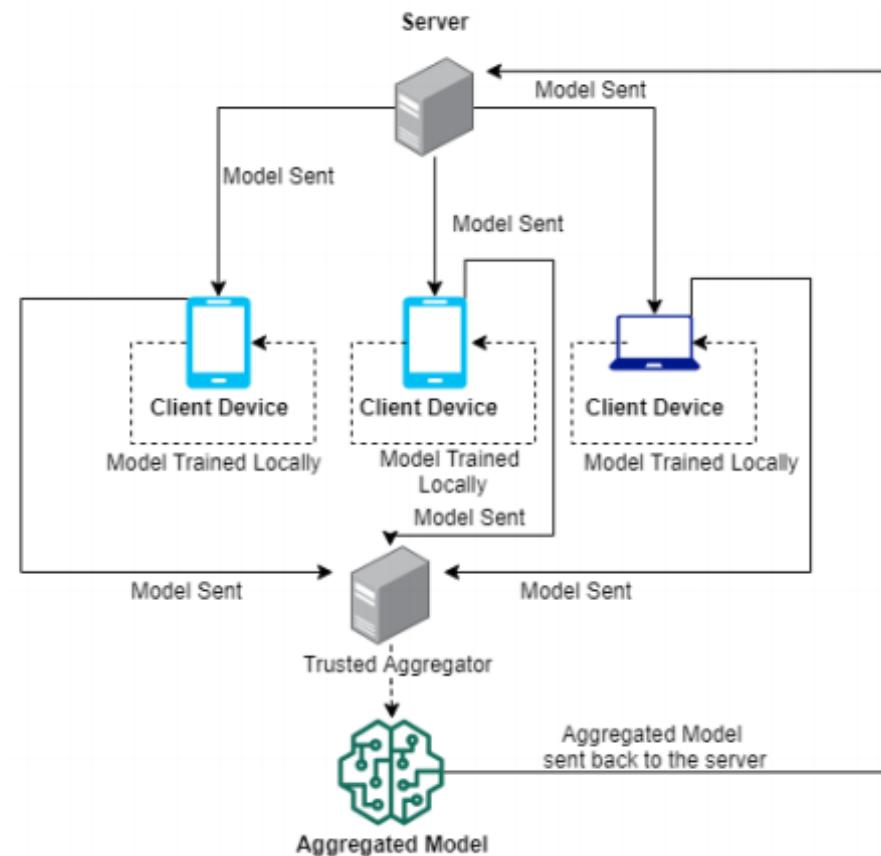
Paper: The Future of Digital Health with Federated Learning ([2020-03](#))

Federated Learning: Use Cases

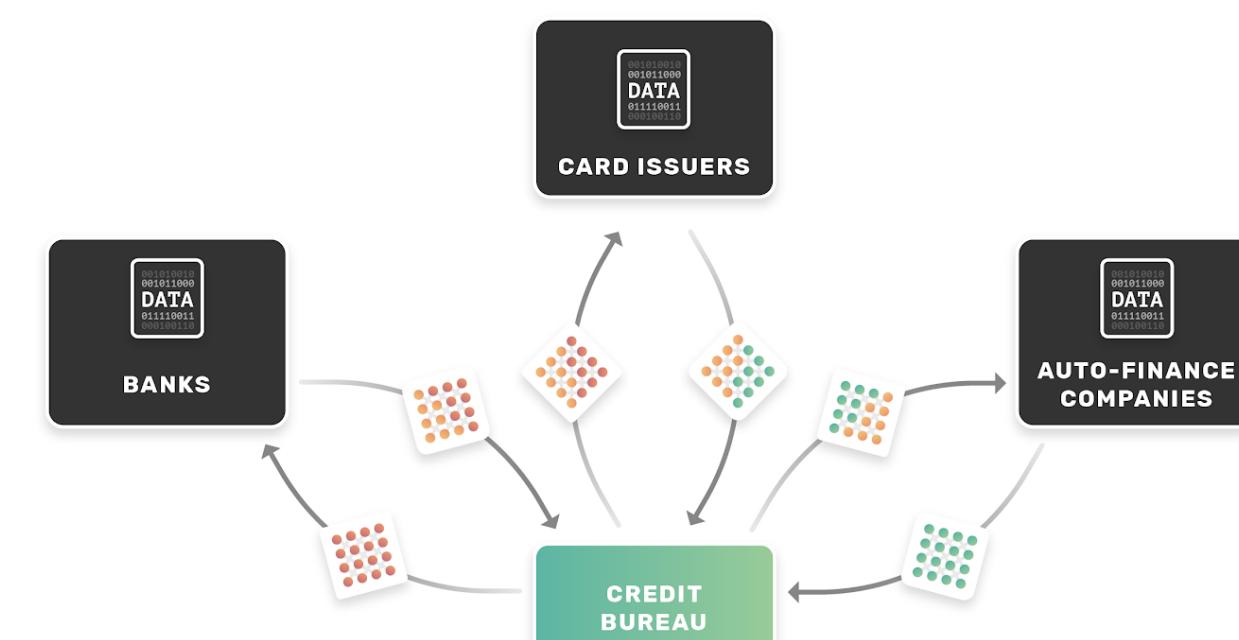
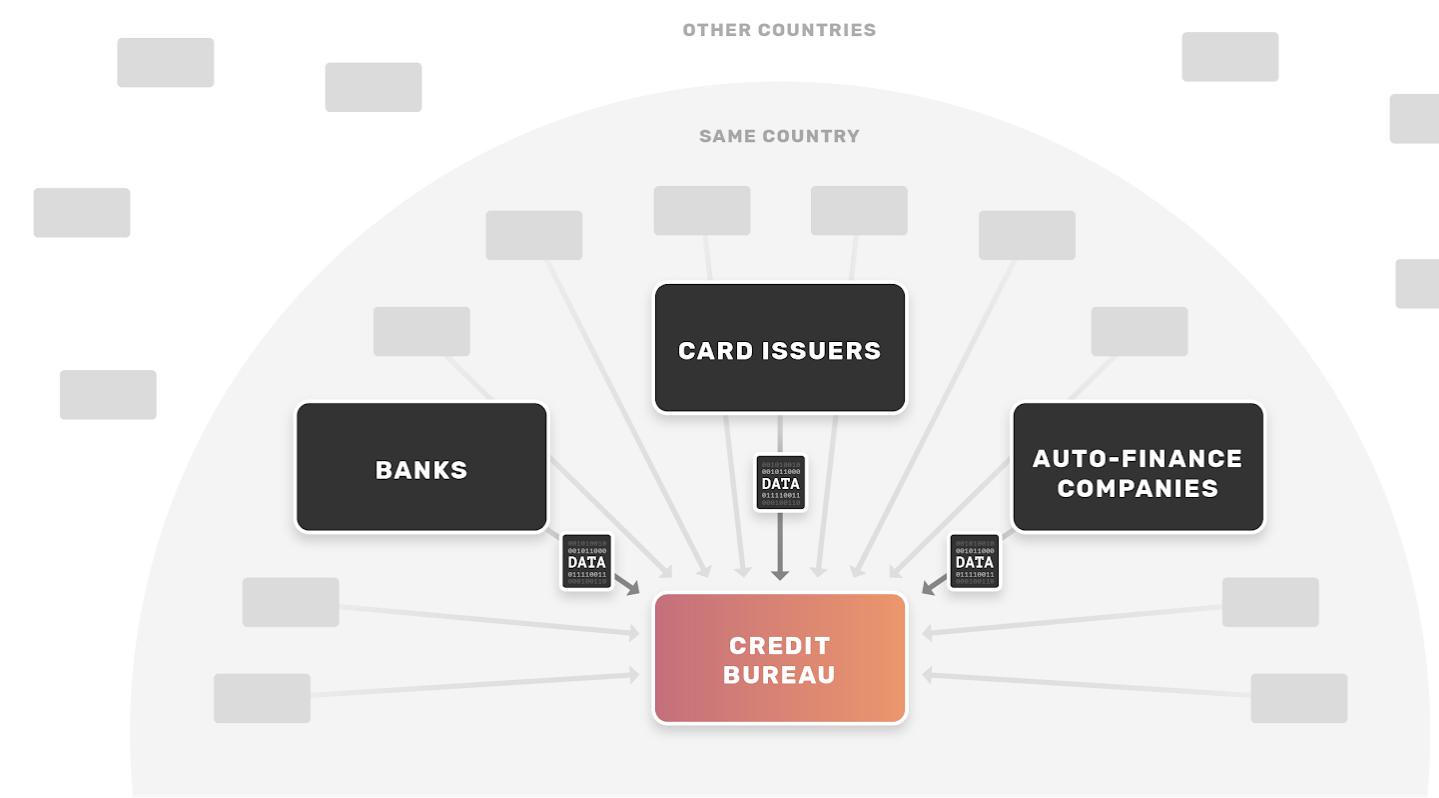
[OpenMined Blog 2019-08-09: Privacy Preserving AI in Medical Imaging](#)



[OpenMined Blog 2020-05-25: Medical Health App: Pill Identification](#)

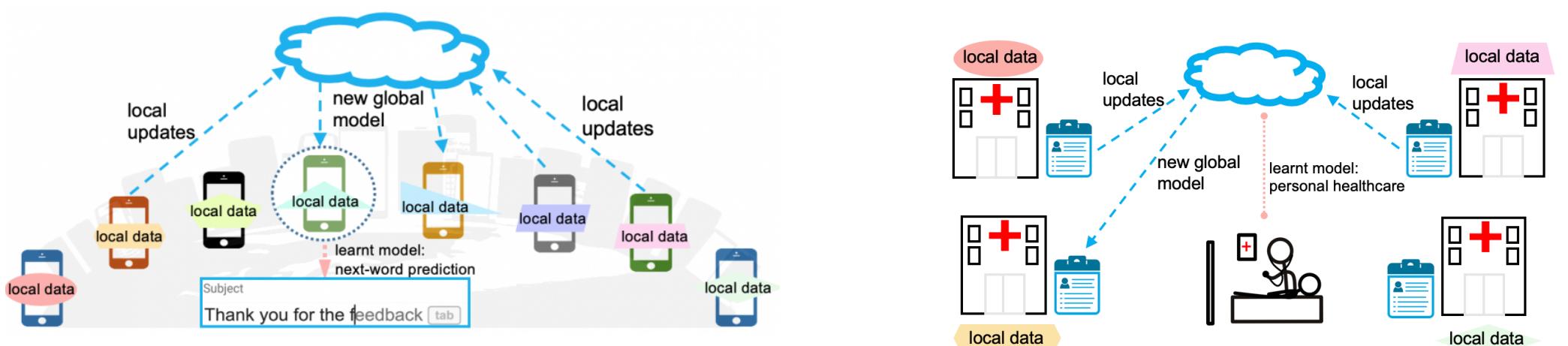


[OpenMined Blog 2020-05-26: Credit Scoring](#)



Federated Learning: Challenges, Issues, Open Problems

Federated Learning: Challenges and Future Directions (CMU 2019-11)



Four fundamental challenges in federated learning:



Necessary to develop communication efficient methods

Expensive Communication



Device may unreliable, unavailable, various speeds, timezones

Systems Heterogeneity



Non-identically distributed; e.g. varied use of language, types of pictures taken, ...

Issue for modelling, analysis and evaluation



Privacy of model updates through differential privacy or multi-party computation.

However, cost for model performance, reduced efficiency.

Understanding, balancing trade-offs.

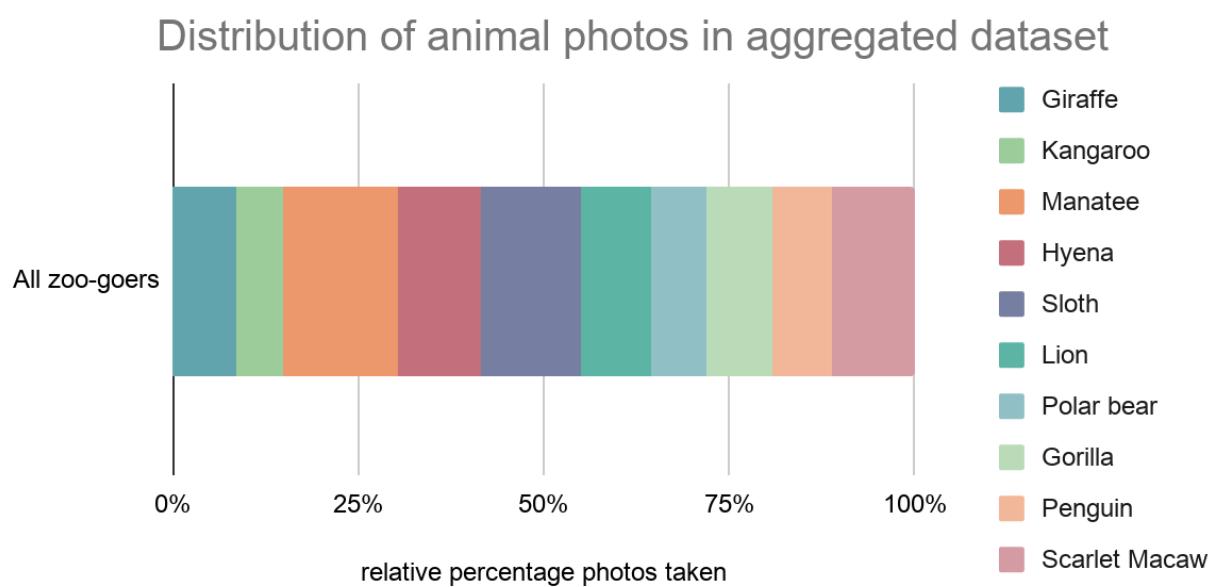
Future directions:

Communications schemes, novel models of asynchrony, heterogeneity diagnostics

Granular privacy constraints, productizing federal learning

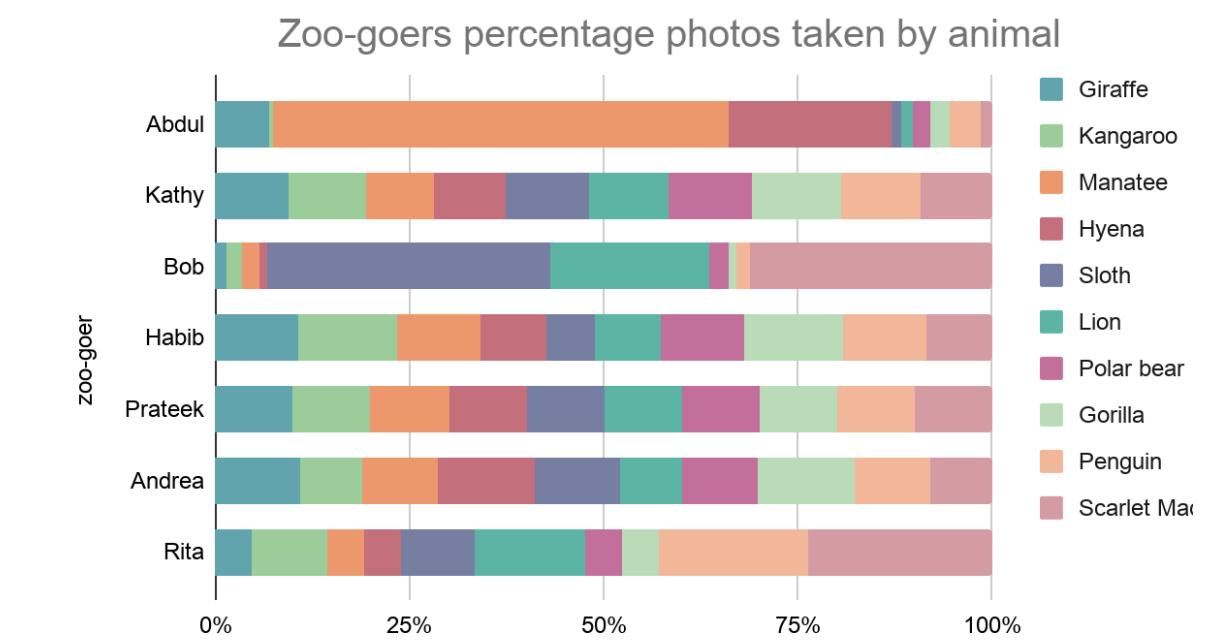
Non Independently and Identically Distributed (Non-IID) Data (OpenMined 2020-05-26)

Centralized data

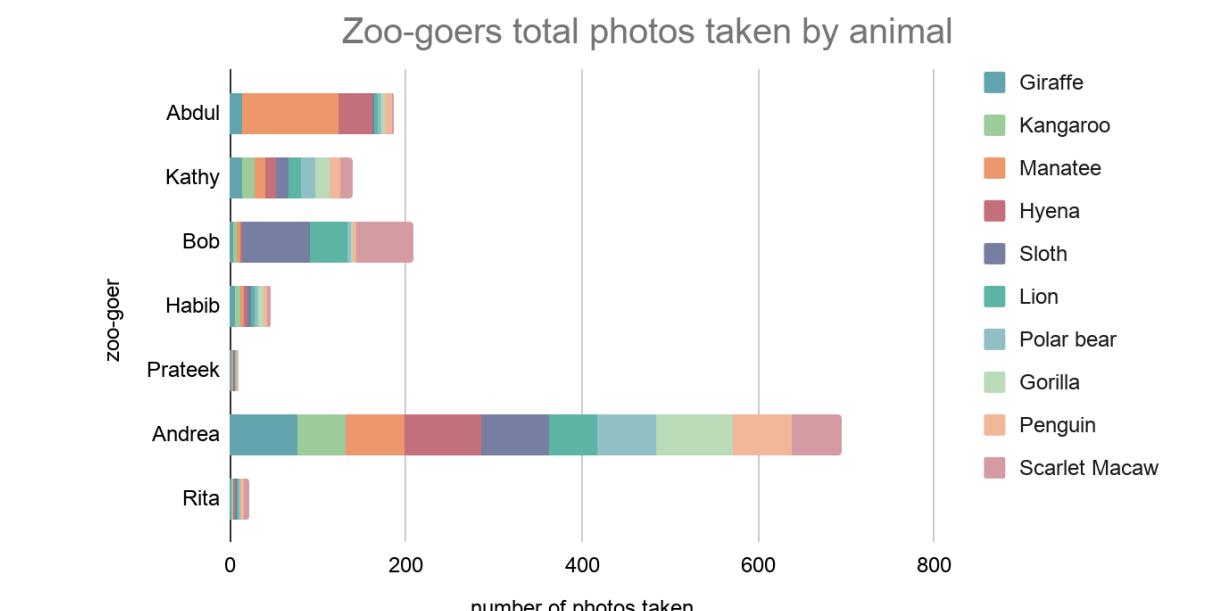


Distributed data

Skew in the distribution of classes of animals



Skew in the quantity of images available



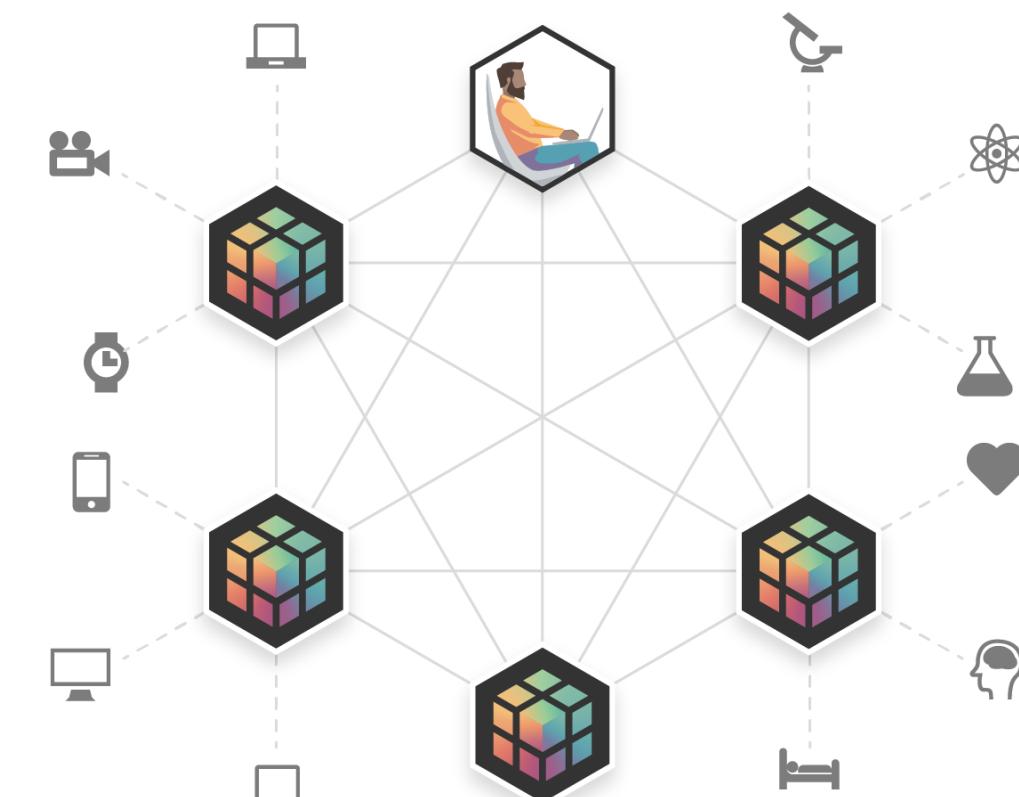
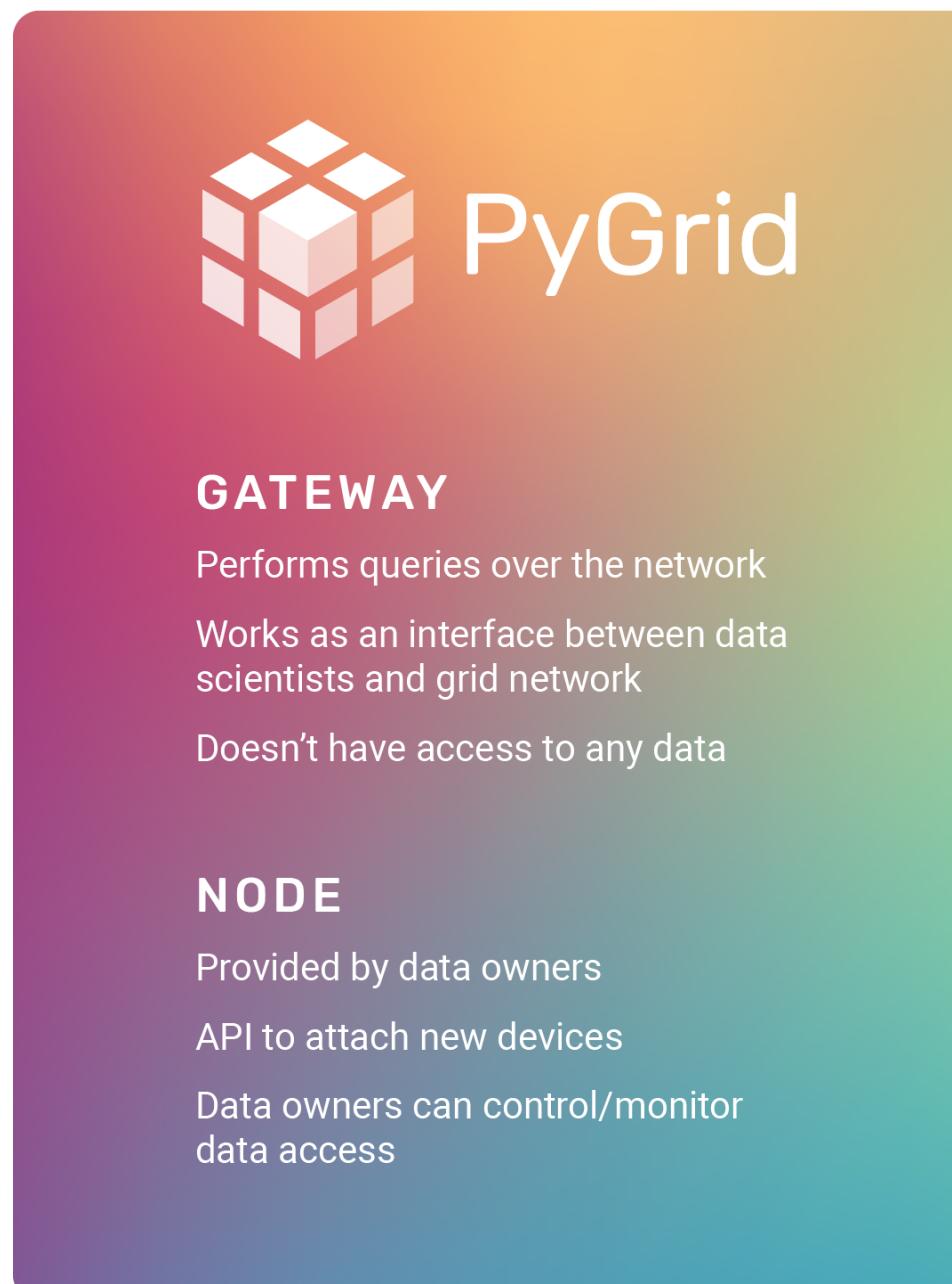
Papers: Advances and Open Problems in Federated Learning (2019-12), The Non-IID Data Quagmire of Decentralized Machine Learning (2019-10), Federated Learning: Challenges, Methods, and Future Directions (2019-08)

Federated Learning: Tools

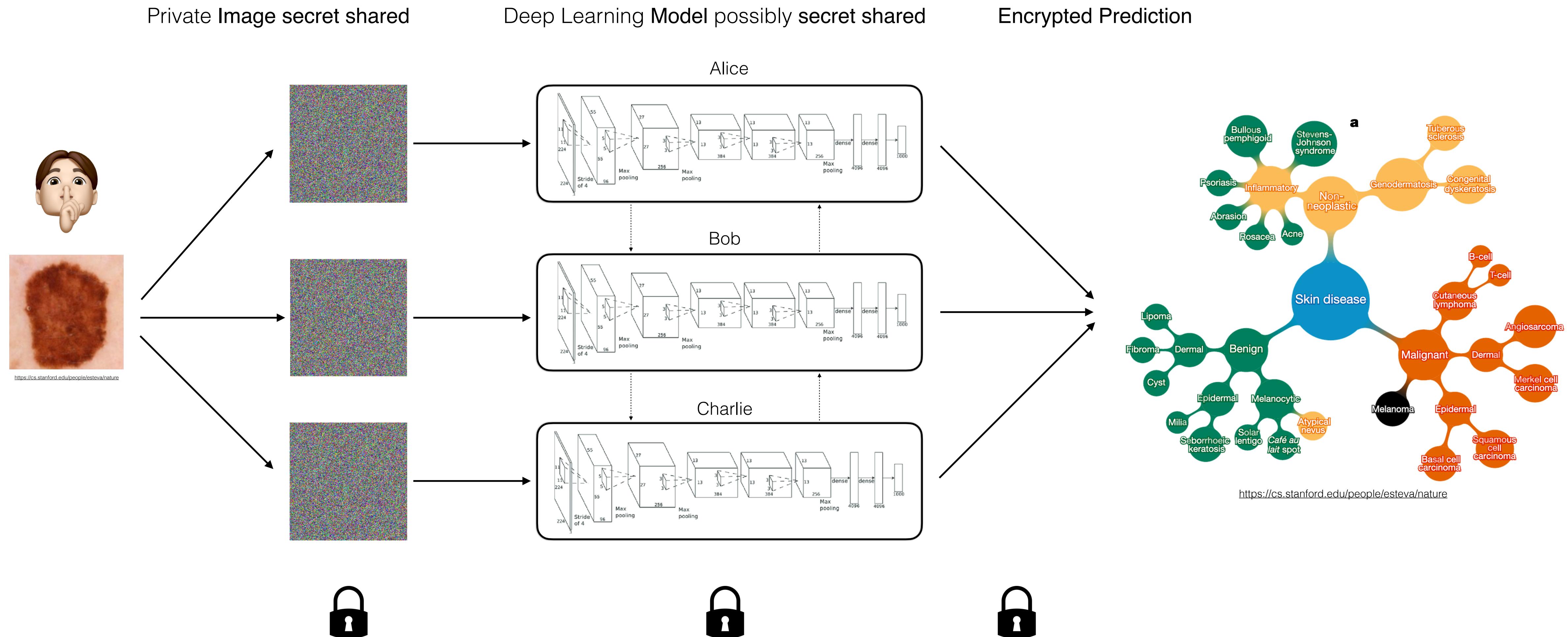


<https://blog.openmined.org/what-is-pygrid-demo>

Peer-to-peer platform for data science and federated learning



4. Encryption: Secure Multi Party Computation (MPC)



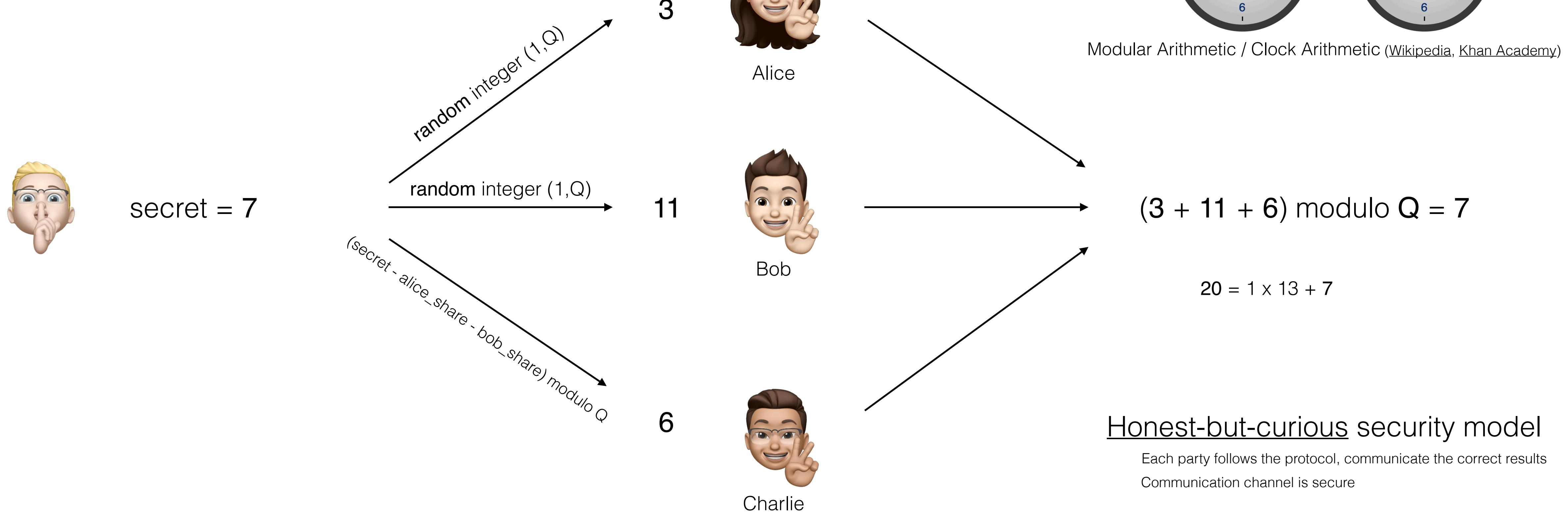
Privacy-preserving Machine Learning as a Service

4. Encryption: Secure Multi Party Computation (MPC) (Additive Secret Sharing)

Problem: How to store a secret... E.g. encryption keys, missile launch codes, numbered bank accounts, safe combination

$Q = 13$ (some very large prime number)

Field Arithmetic



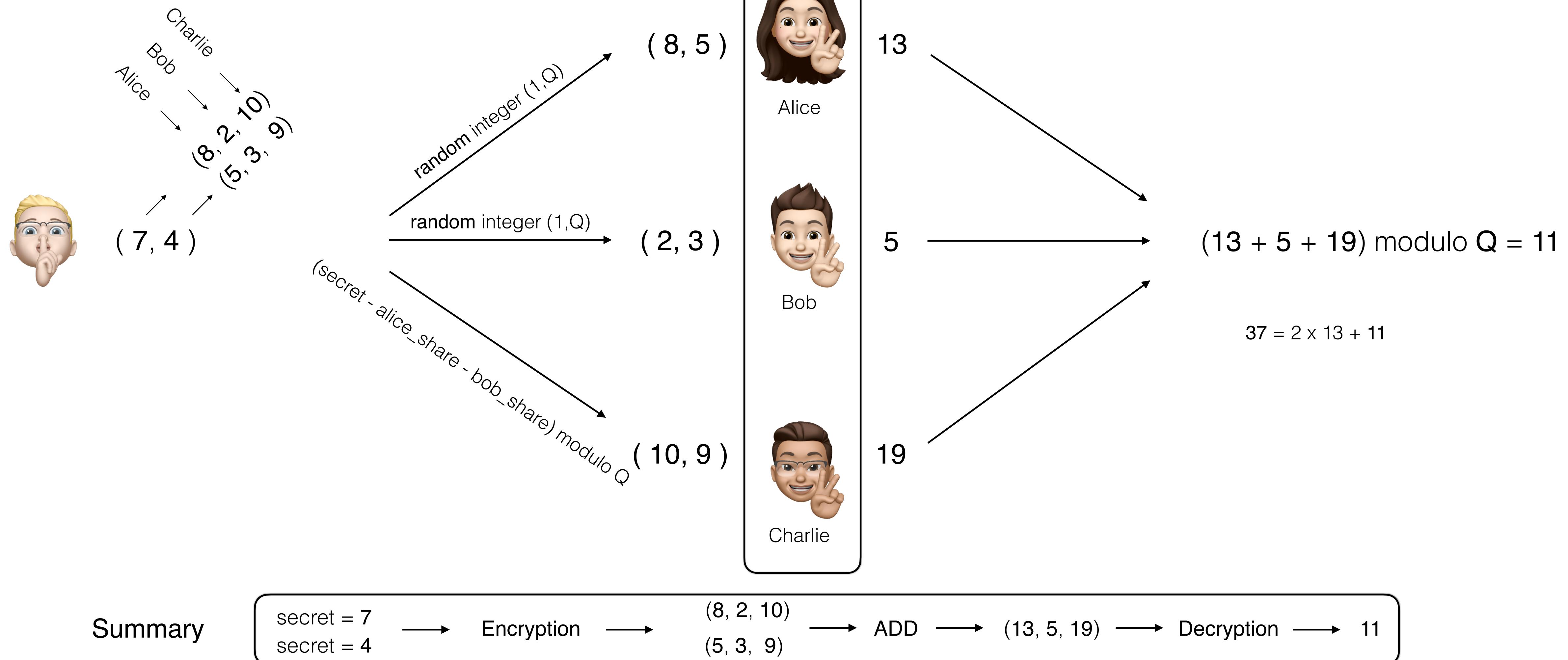
Summary

secret = 7 → Encryption → (3, 11, 6) → Decryption → 7

MPC: ADDITION of two numbers

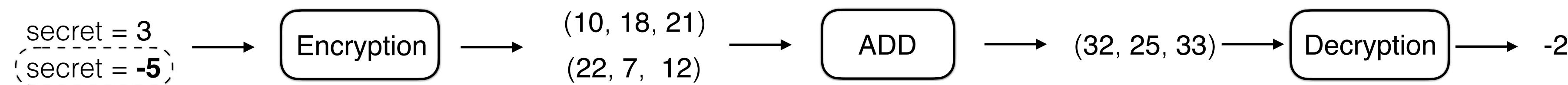
!!! We can ADD two numbers using three workers while keeping the values private (encrypted - additive secret sharing) !!!

$Q = 13$ (some very large prime number)



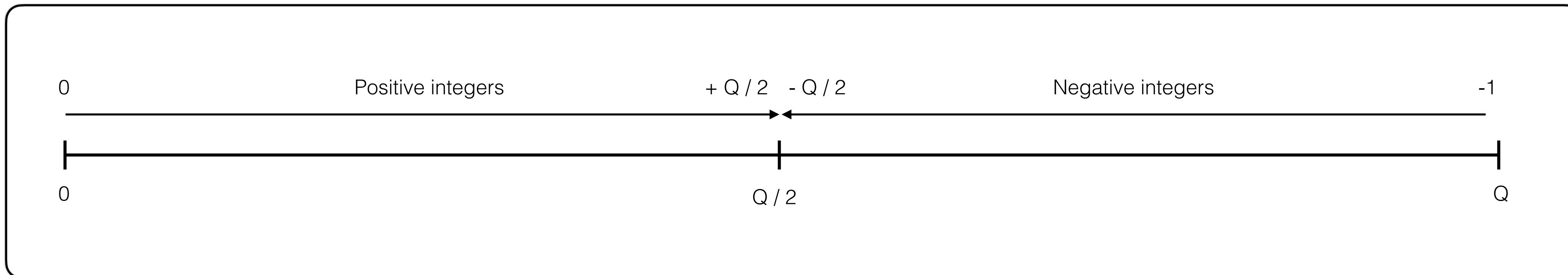
MPC: ADDITION of two integers (positive or negative)

$Q = 23$ (some very large prime number)



```
def encrypt(x, nb_shares=3, Q=23):
    shares = [random.randrange(Q) for _ in range(nb_shares - 1)]
    last_share = (x - sum(shares)) % Q
    shares.append(last_share)
    return shares
```

```
def decrypt(shares, Q=23):
    x = sum(shares) % Q
    return x if x <= Q/2 else x - Q
```



MPC: Fixed Precision Encoding

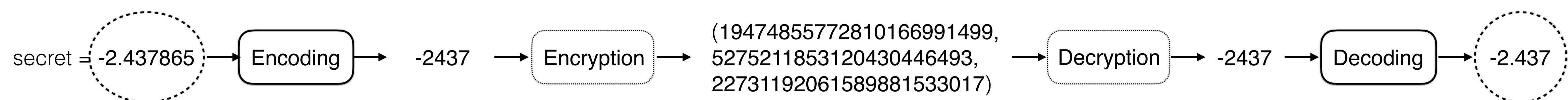
$Q = 23740629843760239486723$ (some very large prime number)

Fixed Precision Encoding

BASE = 10

PRECISION = 3

(3 digits after the decimal point)



MPC: OPERATIONS needed for Deep Learning

“Native” operations in secret sharing protocol:

Addition: we know how to securely add numbers (see previous slide)

Multiplication: By a public number (e.g. $x \times 3$): trivial: same as addition

Between two private numbers (e.g. $x \times y$): slightly more complicated (SPDZ protocol) [1,2,3,4,5,6,7]

secret sharing between workers, interactive computation, communication cost, synchronisation between servers

Other operations expressed using addition and multiplication:

Comparison (e.g. ReLU): more complicated and expensive

Could replace ReLU with sigmoid. Replace MaxPooling with AveragePooling [2]

Other functions (e.g. exponential): using Taylor series expansion; e.g.: $\exp(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \dots$

Still in the early stages of deploying MPC solutions to real problems...[6]

Some operations are surprisingly expensive in the encrypted setting... [2]

- Addition
- Multiplication
- Division
- Subtraction
- Sigmoid
- Tanh
- Exponential

[1] Private Deep Learning with MPC: A Simple Tutorial from Scratch ([Morten Dahl - 2017-04](#))

[2] Private Image Analysis with MPC: Training CNNs on Sensitive Data ([Morten Dahl - 2017-09](#))

[3] What is Secure Multi-Party Computation? A Tutorial for Encrypted Deep Learning ([OpenMined 2020-05](#))

[4] Building Safe A.I.: A Tutorial for Encrypted Deep Learning? ([Andrew Trask - 2017-03](#))

[5] Secret Sharing Explained ([Ben DeCoste - 2018-11](#))

[6] A Pragmatic Introduction to Secure Multi-Party Computation ([Evans & al. - 2020-04](#))

[7] Secure Multiparty Matrix Multiplication Based on Strassen-Winograd Algorithm ([Dumas & al. - 2019-04](#))

[8] Secure multi-party computation ([Wikipedia](#))

MPC for Deep Learning: Issues / avenues

Communication cost, synchronisation between servers for the multiplication operation

Computation cost: orders of magnitude slower, ensure that each private value is only sent masked once

Comparison operations

When training: issues with cost function, softmax operations

Division by private value: can opt for some privacy leakage (some computation performed by a trusted server)

Optimizer: choosing sub-optimal optimizers for easier operations in the encrypted domain (e.g. momentum SGD instead of Adam)

Transfer learning: pre-train on public dataset with optimal optimizer

Use of a crypto-provider to perform some computation potentially in advance

Synchronisation of servers for dropout operations

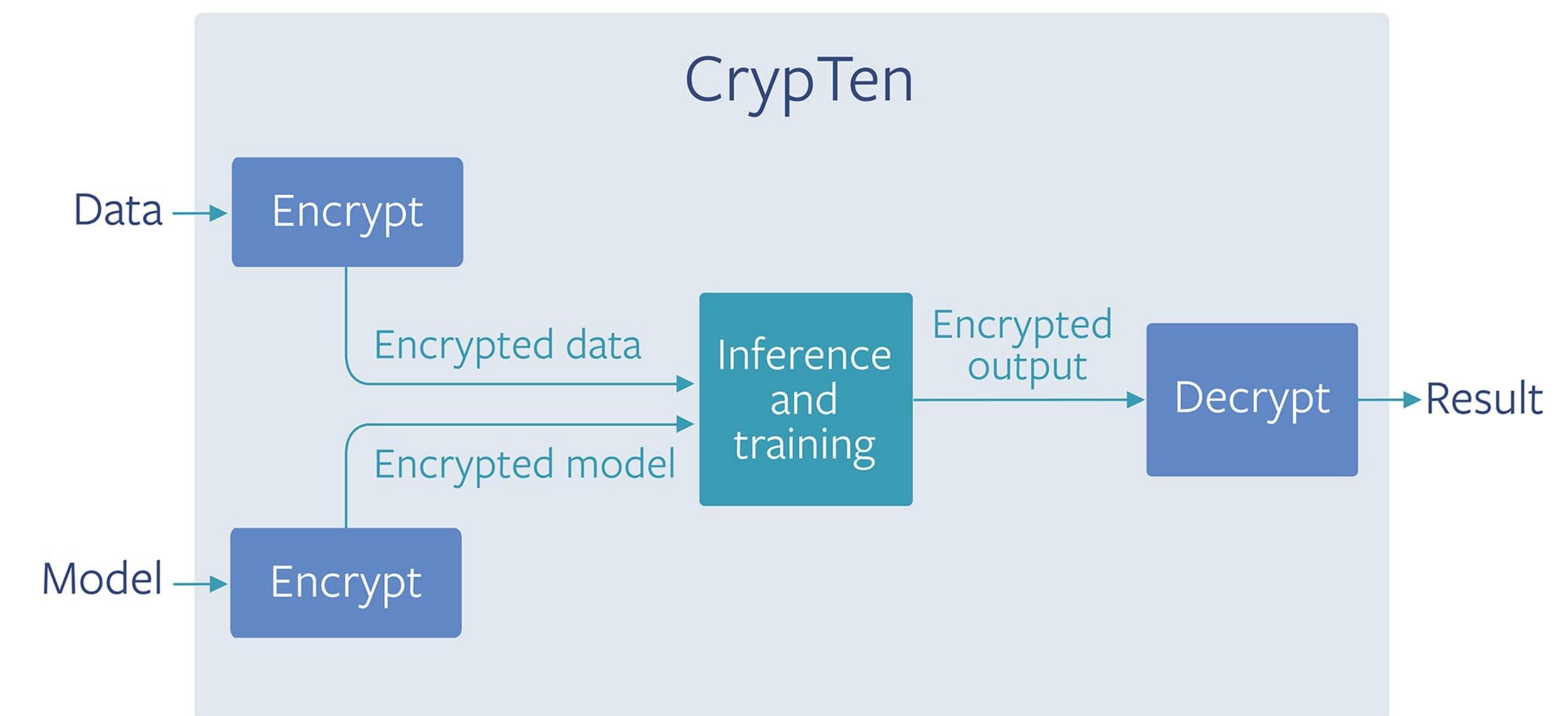
Optimisation of computation in dense layers (classification heads): trick in secret sharing the matrices

Activation functions: designing / altering functions for efficient computation in the encrypted domain

Secure Multi-Party Computation: Tools

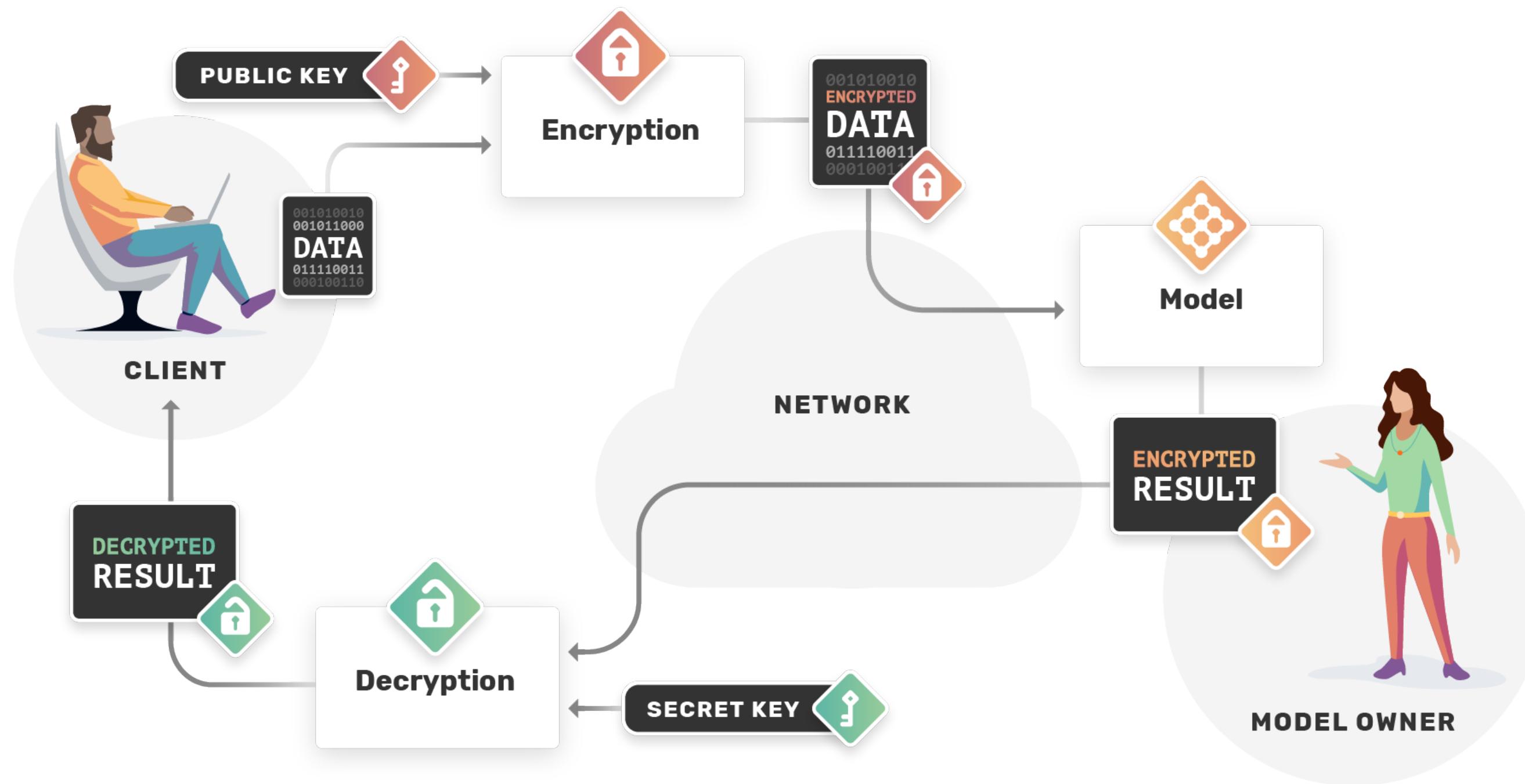


Facebook
Open Source

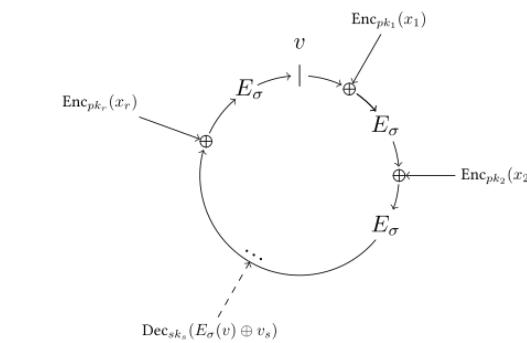


5. Homomorphic Encryption

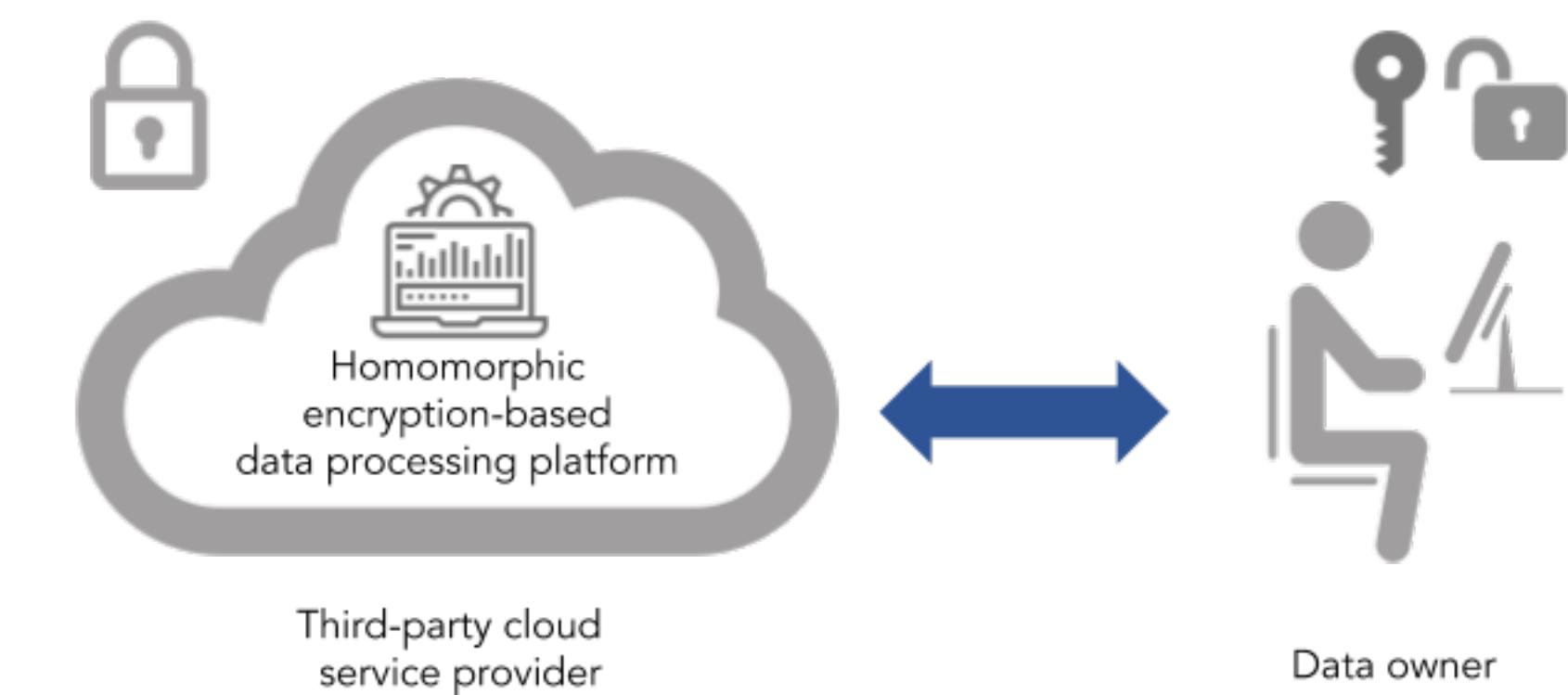
Privacy-preserving Machine Learning



[Homomorphic Encryption in PySyft with SEAL and PyTorch - OpenMined 2020-04](#)



- [Post-quantum cryptography](#)
- [Lattice-based cryptography](#)
- [Ring learning with errors](#)



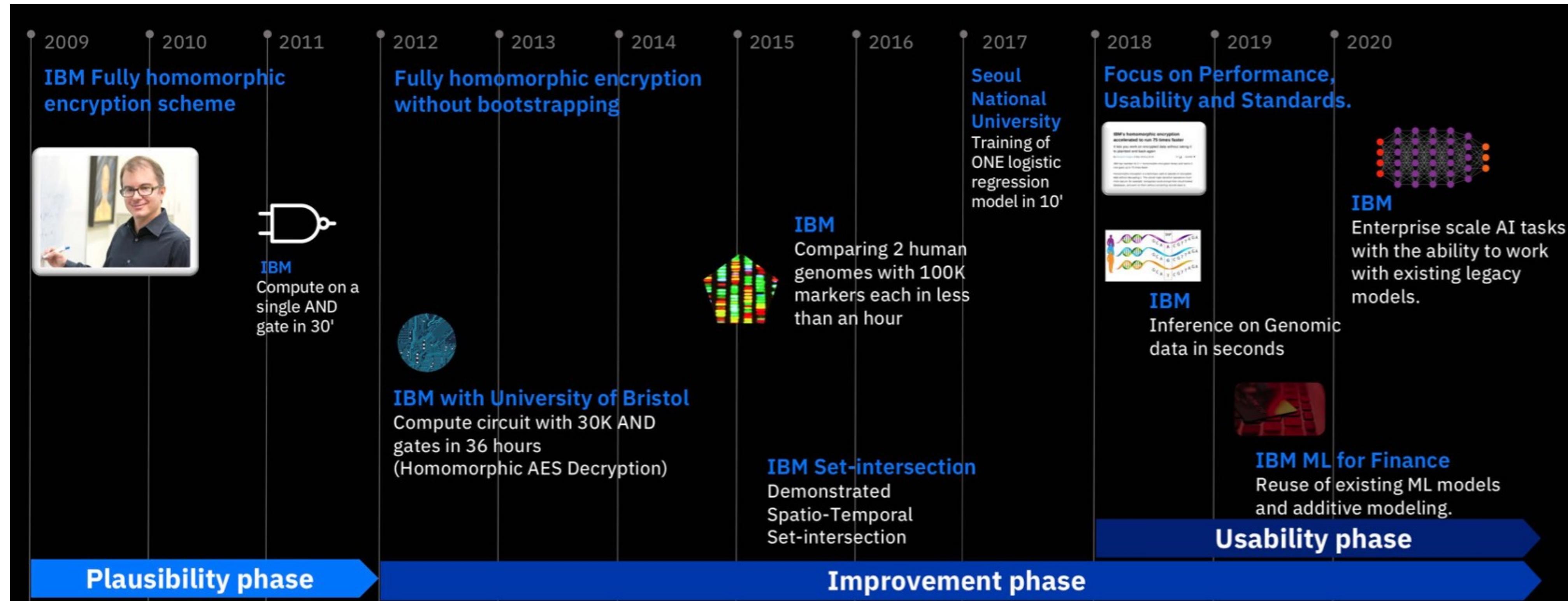
[Homomorphic Encryption for Beginners: A Practical Guide \(Part 1\) \(2018-12\)](#)

Fully Homomorphic Encryption: supports arbitrary computation on encrypted data

[CKKS method \(2017\)](#)

Homomorphic Encryption: History / Use case

Craig Gentry (2009) - Stanford - IBM



IBM Research (2020)

Use case: IBM study with top Brazilian bank (2020)

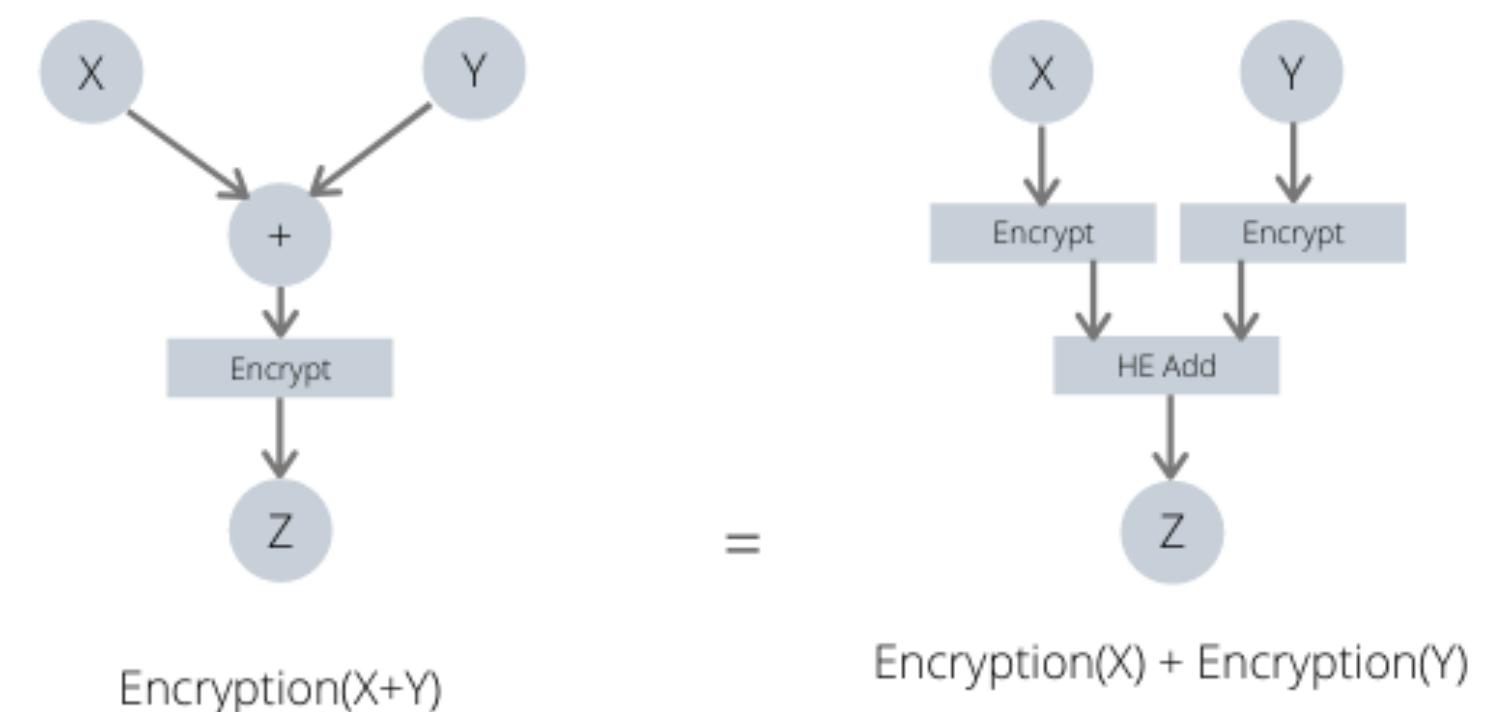
Homomorphic Encryption: Tools



TenSEAL

Homomorphic Encryption in PySyft with SEAL and PyTorch - OpenMined 2020-04

addition, subtraction and multiplication of encrypted vectors of either integers (using BFV) or real numbers (using CKKS)



Open Source implementations of Homomorphic Encryption (HE) schemes:

Palisade: widely used Lattice Crypto Software Library - DARPA funded - supports leading homomorphic encryption schemes (BGV, BFV, and CKKS, TFHE)

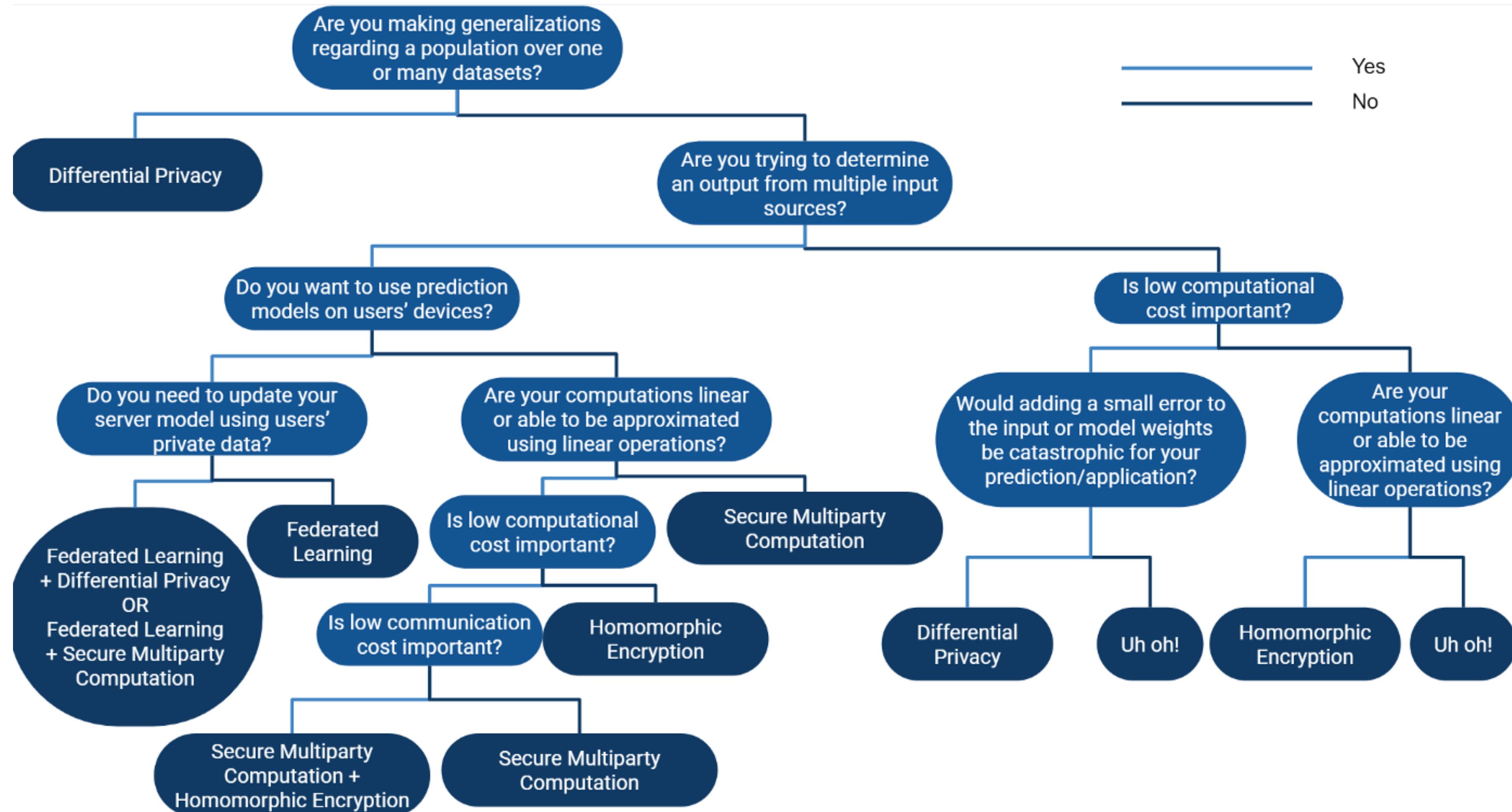
SEAL (Microsoft): widely used open source library from Microsoft that supports the BFV and the CKKS schemes

HElib: early and widely used library from IBM that supports the BGV scheme and bootstrapping

Full list on the Homomorphic Encryption Standardization's website: homomorphicencryption.org

In active development...

Which Privacy Preserving Method To Use?



Patricia Thaine (2019-01)

Context

0. You are being tracked

Privacy preserving techniques

- | | | |
|--|---|--------------------------|
| 1. Data anonymisation: | NO guarantee (linkage attack) | Numer.ai |
| 2. Differential Privacy: | Privacy guaranteed to a certain degree (ϵ) | |
| 3. Federated Learning: | Edge computing, the data stays private | |
| 4. Encrypted: Multi Party Computation: | <u>Honest-but-curious</u> security model | |
| 5. Encrypted: Homomorphic Encryption | Perfect Privacy Preserving Machine Learning | |

Resources

Course: [Secure and Private AI \(Udacity\)](#)

Data: [DataCollaboratives.org](#)

Publications: Center for International Governance Innovation ([CIGI](#) - Canada): [Data Is Dangerous](#) (2020-04)

Open source community: [OpenMined.org](#): Lowering the barrier to entry to privacy preserving machine learning technologies

[Awesome MPC](#): A curated list of multi party computation resources and links.

New York Times - Opinion - The Privacy Project - One Nation Tracked (2019-12-20):

[An Investigation into the smartphone tracking industry](#)

Governance:

Europe: General Data Protection Regulation ([GDPR](#) 2018),

Canada: Personal Information Protection and Electronic Documents Act ([PIPEDA](#) 2019), [Privacy Act](#) (2019),

California: California Consumer Privacy Act ([CCPA](#) 2020)

Keep learning!