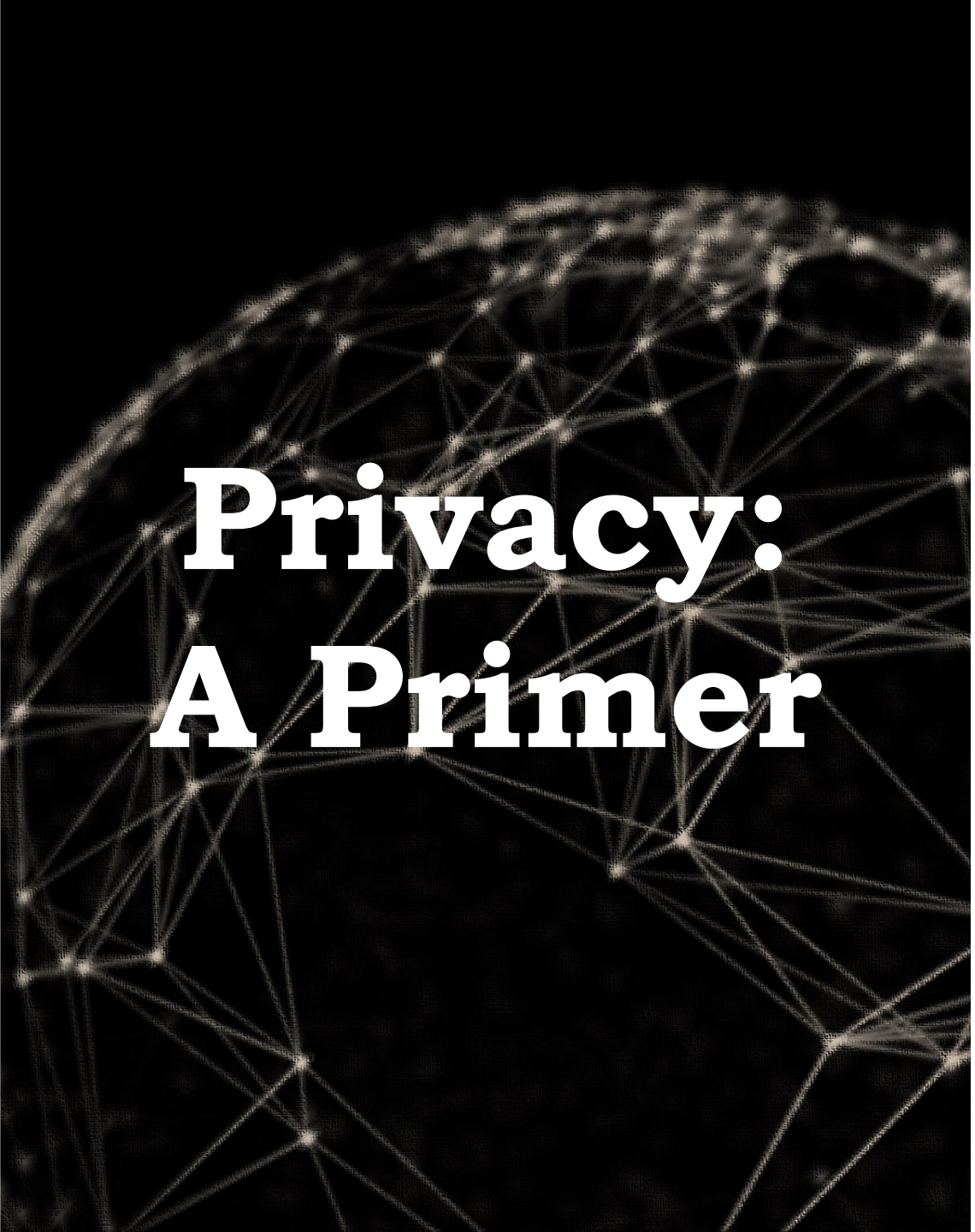
Defence Against the Dark Ads

Sandun Bambarandage

Who Am I?

- Princeton Class of 2020
- Former INTERFACE member (since SEoCS)
- I read privacy policies
- INTERFACE Meme Chair Emeritus

Tech Company: Does something unethical or illegal. CITP Professors:



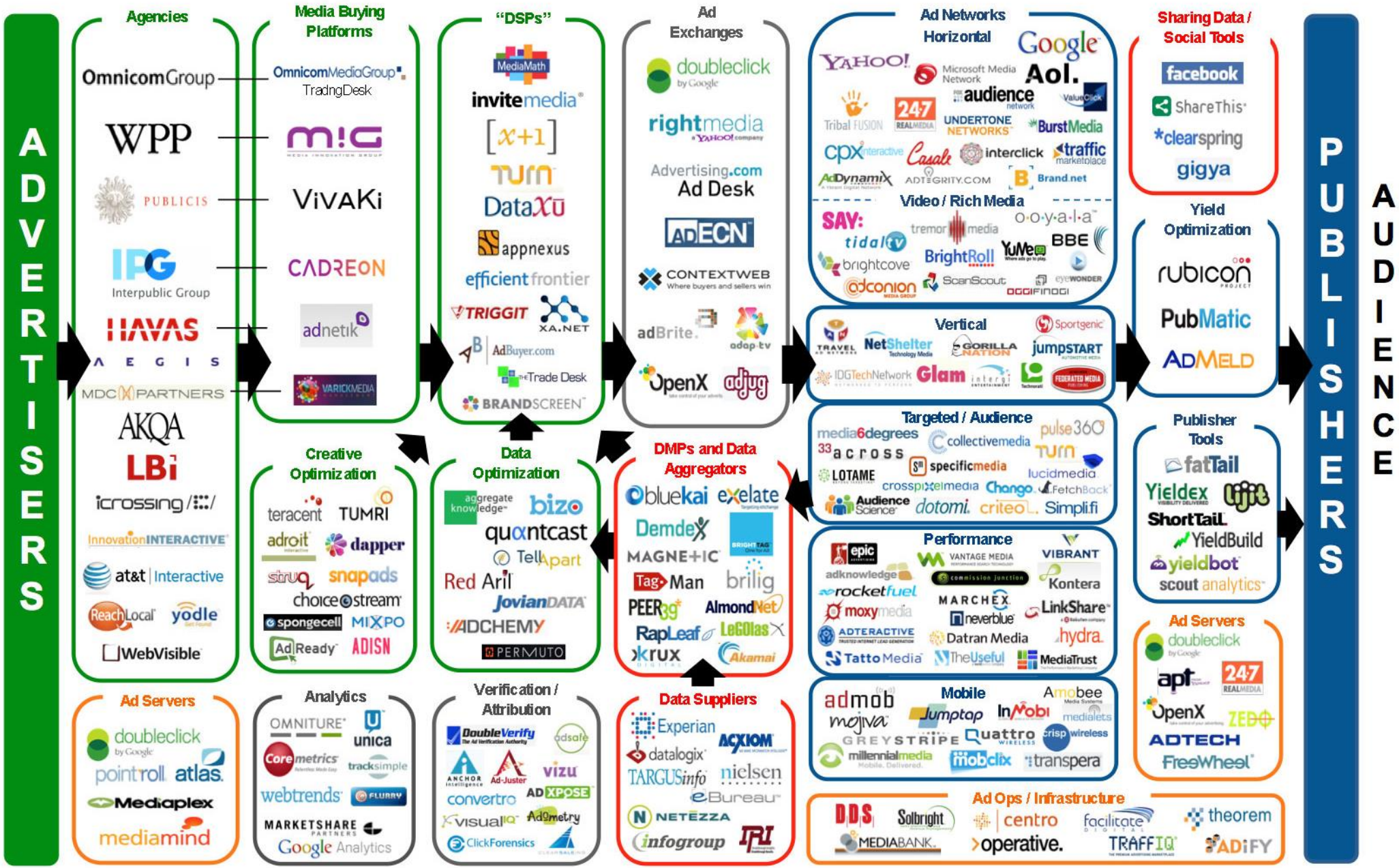
Privacy: A Primer

- Privacy is a human right
- Digital privacy is complicated
  - Business models of Big Tech
  - Moved: Fast. Things: Broken
  - Reactive approaches to violations and their consequences
  - Legislative lag
- Huge space for discussion
  - Advertising paradigms
  - Privacy and Security
  - Privacy and Ethics
  - Privacy and Philosophy
    - Targeted ads and free will

What To Expect

- An overview of:
  - Who collects your data
  - Why they collect your data
  - How they collect your data
- How to minimize data collection
  - Best practices when online
  - Which tools and technologies to use
  - How privacy-preserving tech works
- Trade-offs
  - Financial
  - Convenience
  - Security

Display Advertising Technology Landscape



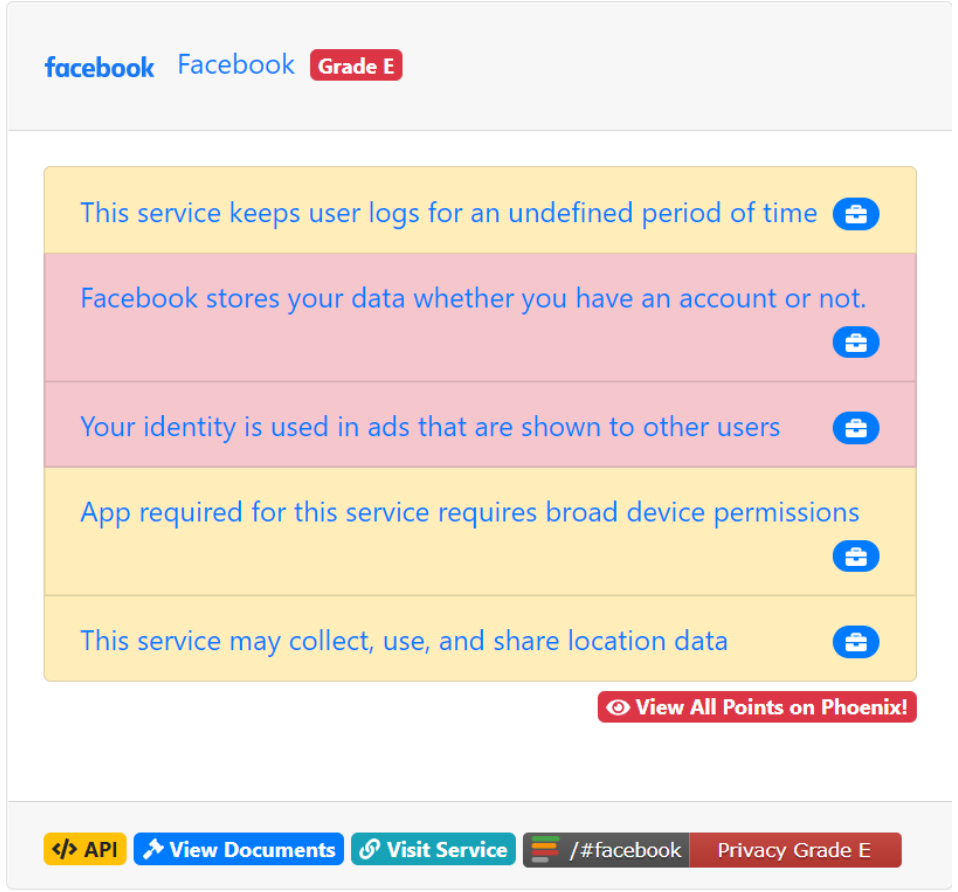
The Data Ecosystem

- Who collects your data?
  - Everyone
- Who gives your data to others?
  - Almost everyone
- Who can you avoid?
  - Most third-parties
- Who can’t you avoid?
  - (Unless you go to extremes)
  - One of Google/Apple/Microsoft
  - Zucc

- The entity whose service/device you are using
- Many legitimate uses!
  - Diagnostics
  - User Verification and Security
  - Product and UX improvement
  - Give data to second parties
    - Consultants, service providers etc
- Several not-quite legitimate uses
  - Preemptive/speculative data collection
  - Amazon and Kindle: https://www.theguardian.com/technology/2020/feb/03/amazon-kindle-data-reading-tracking-privacy

To Use Or Not To Use?

- Read the terms!
  - Privacy Policy != ToS
  - ToS;DR (https://tosdr.org/)



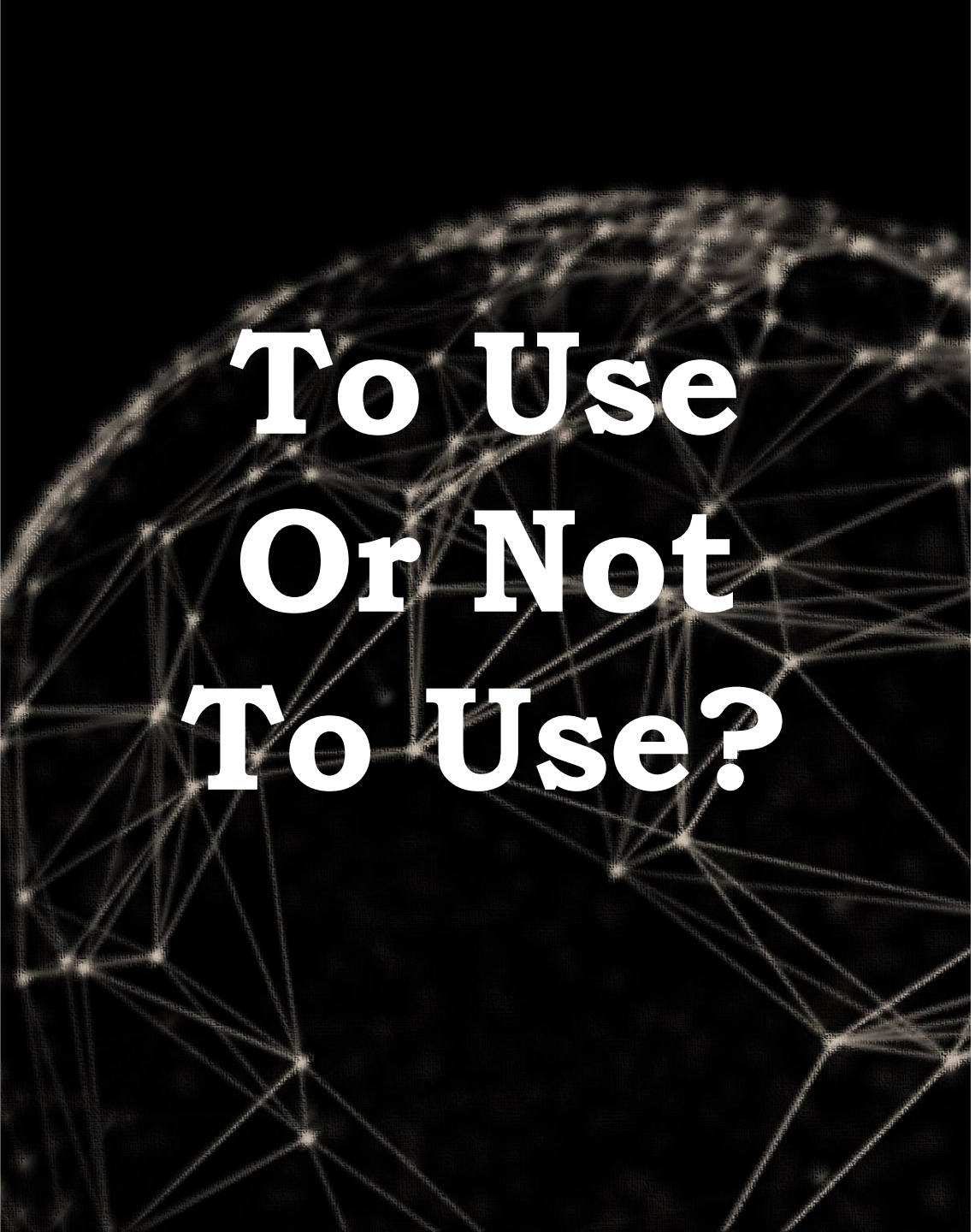
- https://www.paypal.com/ie/webapps/mpp/ua/third-parties-list

Category	Party Name and Jurisdiction (in brackets)	Purpose	Data Disclosed
1. Payment Processors			
	Barclays Bank Plc (UK), HSBC Bank Plc (UK, Ireland), HSBC Merchant Services LLP (UK), Bank of America N.A. (EMEA, USA BA Continuum India Private Limited (India), Discover Financial Services (USA), JPMorgan Chase Bank (UK, USA), BNP Paribas (France), Netgiro (Sweden), Wells Fargo (Ireland, USA), American Express (USA), National Westminster Bank PLC (UK), OmniPay Limited (Ireland), Australia and New Zealand Banking Group Limited (Australia), ANZ National Bank Limited (New Zealand), Transaction Network Services UK Limited (UK).	To allow payment processing settlement services, and fraud checking.	Name, address, email address, phone number, details of user funding instruments, and details of payment transactions, document identification number.
	BNP Paribas (France)	To allow payment processing, fraud checking and dispute handling for transactions of PayPal users when those users transact with a merchant who uses the PayPal service via the BNPP service.	Name, email address, phone number, physical address, IP address, date of birth, PayPal account number, details of user funding instruments and details of payment transactions.
	VTB24 (Russia)	To allow payment processing, fraud checking and dispute handling for transactions of PayPal users when those users transact with a merchant who uses the PayPal service via the VTB 24 service.	Name, address, email address, phone number, details of user funding instruments and details of payment transactions.
	Deutsche Bank AG (Germany, Netherlands, France, Spain)	To allow the processing of direct debits in the SEPA region.	Name, date of transaction, amount, currency and user's bank account information.
	Royal Bank of Scotland plc (UK) ("RBS")	To allow payment processing and dispute handling for transactions of PayPal users when those users transact with a merchant who uses the PayPal service via the RBS service.	All account information except details of user financial instruments.

- https://www.paypal.com/ie/webapps/mpp/ua/third-parties-list

	Facebook Inc (US)	To facilitate payment processing on Facebook for the joint customers of Facebook and PayPal.	Customer name. email address, account and transaction history

	Facebook, Inc (US), Facebook Ireland Limited (Ireland), Twitter, Inc. (US),(Ireland) AdRoll, Yahoo, Facebook Business Manager (SE).	To execute retargeting campaigns in order to deliver personalised advertising.	Advertising ID and device ID to segment user groups based on app behaviour, encrypted e-mail address associated with PayPal users (without indicating account relationship), IP Address, Anonymous ID generated by cookies, pixel tags or similar technologies embedded in webpages, ads and emails delivered to users. Mobile advertiser ID, IP Address and other metadata via Facebook SDK in mobile apps.



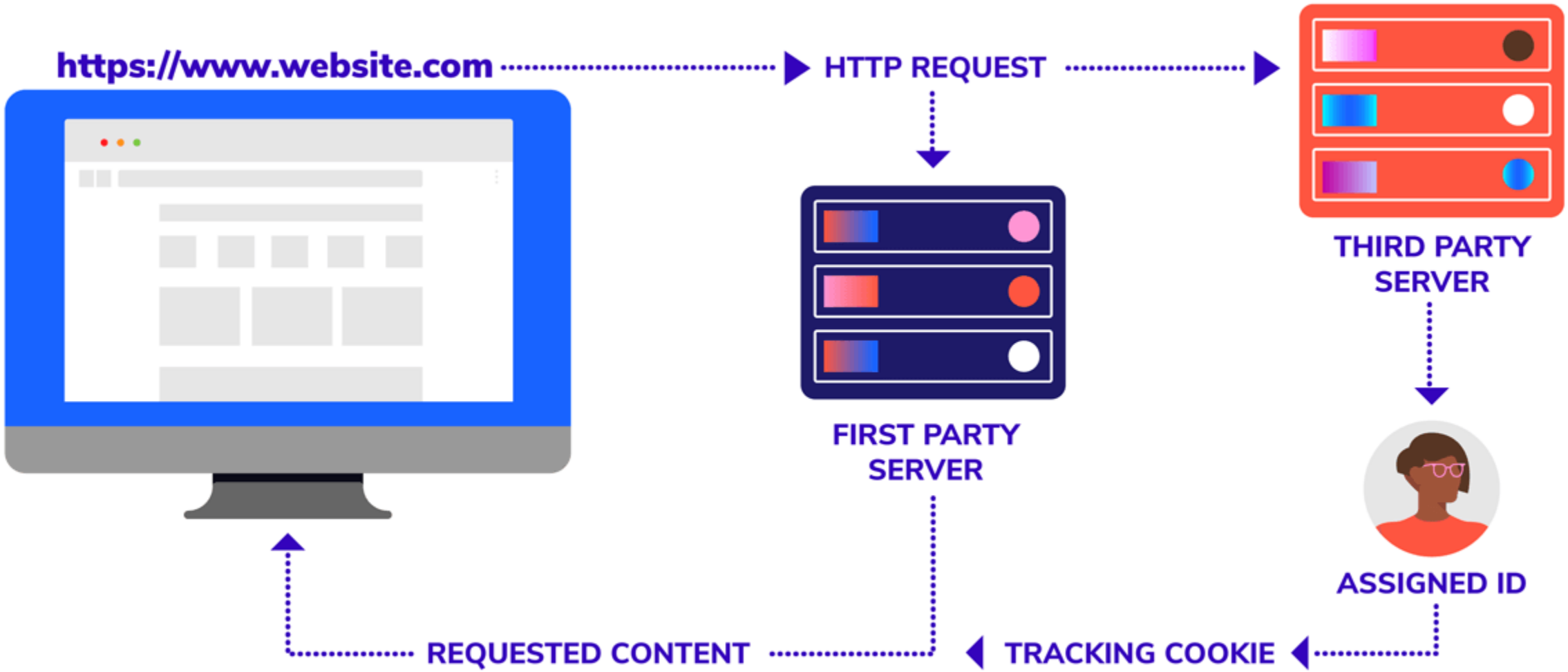
- Read the terms!
  - Privacy Policy != ToS
  - ToS;DR (https://tosdr.org/)
  - The Verge: Accept to Continue
- Limiting first-party collection: general tips
  - Privacy settings
  - Disable crash reporting (devices and apps)
  - Disable analytics
  - Improving voice/text services

- Data collected by:
  - Manufacturer (OEM)
  - OS Vendor
  - Installed applications
- Preventive measures:
  - Clean install (Windows)
  - Disable optional data collection

Data Category	Description	Examples
Browsing history data	This type of **Optional diagnostic data** includes details about web browsing in the Microsoft browsers.	• Browser activity, including browsing history and search terms in Microsoft browsers. • Changes to browser configuration impacting search experiences.
Device connectivity and configuration data	This type of **Optional diagnostic data** includes details about the device, its configuration, and connectivity capabilities.	• More detailed information about device settings and configurations.
Inking, typing, and speech utterance data	This type of **Optional diagnostic data** includes details about the voice, inking, and typing input features on the device.	• Samples of the content you type, write, or dictate on the device. • Details about status of transcribing input into text.
Product and service performance data	This type of **Optional diagnostic data** includes details about device or service health and performance.	• More detailed information about device and service health.
Product and service usage data	This type of **Optional diagnostic data** includes details about the usage of the device, operating system, applications, and services.	• App activity, such as which apps are launched. • Operating system and services usage.
Software setup and inventory data	This type of **Optional diagnostic data** includes software installation and update information on the device.	• More detailed information about the operating system, apps and services installed.

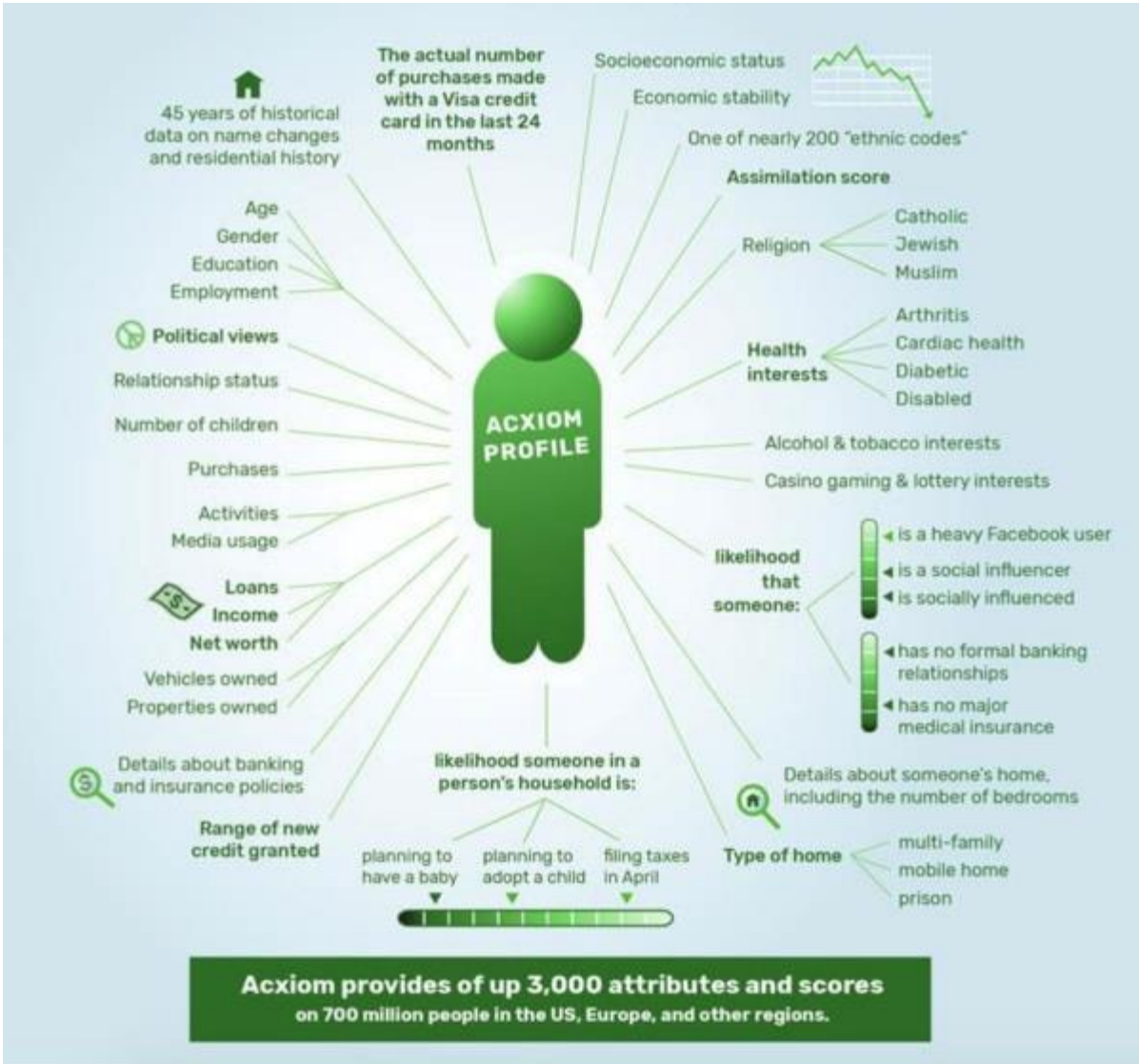
- Data collected by:
  - Manufacturer (OEM)
  - OS Vendor
  - Installed applications
- Preventive measures:
  - Clean install (Windows)
  - Disable optional data collection
- Alternatives:
  - Linux-based operating systems
    - Some of these are also very commercialized
  - Non-OEM Linux will generally prevent OEM tracking

- Cookies
  - First Party – set by the site that you visit
  - Third Party – set by a different site
  - Enables cross-site tracking


https://www.website.com
HTTP REQUEST
THIRD PARTY SERVER
FIRST PARTY SERVER
ASSIGNED ID
TRACKING COOKIE
REQUESTED CONTENT

- Cookies
  - First Party – set by the site that you visit
  - Third Party – set by a different site
  - Enables cross-site tracking
  - Can be blocked on every browser
    - Safari blocks all by default
    - Google plans to block all on Chrome
- Third party cookies are everywhere
  - Google – 74% of porn sites (https://www.nytimes.com/2019/07/17/opinion/google-facebook-sex-websites.html)
  - Blocking them is the bare minimum
- Newer, better technologies
  - “Facebook Pixel”
- https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html

CONSUMER DATA AND ANALYTICS INDUSTRY
Ad Networks
SSPs
Turn
Rocket Fuel
Nielsen
Neustar
ADVERTISING TECHNOLOGY
DSPs
DMPs
Personalization
MailChimp
Predictive Marketing
Call Center
Adobe
OpenX
Lotame
Cross-device tracking
Microsoft
Managed Services
Deloitte
Oracle
McKinsey
Accenture
CUSTOMER MANAGEMENT
CRM
MDM
BUSINESS IT
DATA INTEGRATION AND IDENTITY MATCHING
Auth
SAS
Health & Insurance Analytics
Loyality Programs
Salesforce
IBM
FICO
Palantir
Epsilon
Merkle
Acxiom
Identity Verification
Fraud Detection
Data Append
ID Analytics
Segmentation
Cyber Security
Database Services
Agencies
Experian
Telematics Data
TransUnion
Equifax
RISK DATA
Credit Scoring
WPP
MARKETING DATA
List Rental
Direct Marketing
Tenant screening
LexisNexis
Lead Generation
Background Checks


45 years of historical data on name changes and residential history
The actual number of purchases made with a Visa credit card in the last 24 months
Socioeconomic status
Economic stability
One of nearly 200 "ethnic codes"
Assimilation score
Age
Gender
Education
Employment
Religion
Catholic
Jewish
Muslim
Political views
Relationship status
Number of children
Purchases
Activities
Media usage
Health interests
Arthritis
Cardiac health
Diabetic
Disabled
ACXIOM PROFILE
Alcohol & tobacco interests
Casino gaming & lottery interests
Loans
Income
Net worth
Vehicles owned
Properties owned
likelihood that someone:
is a heavy Facebook user
is a social influencer
is socially influenced
has no formal banking relationships
has no major medical insurance
Details about banking and insurance policies
likelihood someone in a person's household is:
Details about someone's home, including the number of bedrooms
Range of new credit granted
planning to have a baby
planning to adopt a child
filing taxes in April
Type of home
multi-family
mobile home
prison
Acxiom provides of up 3,000 attributes and scores
on 700 million people in the US, Europe, and other regions.

- Digital Fingerprinting
  - Gathering browser characteristics that are collectively unique
    - User agent string
    - Installed fonts
    - CPU and GPU details
    - “Canvas hash”
  - Test at https://coveryourtracks.eff.org/
- Anti-fingerprinting
  - Built-in measures in some browsers
  - Blocking script execution
  - Blocking tracking domains

- Privacy-preserving browsers
  - Firefox*
  - Tor (https://www.torproject.org/)
    - Privacy + anonymity
    - Slow, impractical for regular use
  - Brave (https://brave.com/)
    - Built-in adblocker
    - Similar features to Firefox, enabled by default
    - Pay users in crypto to watch ads
    - Ad overlaying controversy
- Anti-tracking extensions
  - uBlock Origin, Privacy Badger
    - List-based blocking
  - NoScript
  - DecentralEyes – CDN emulation

Not-So-Private Browsing

- Search Engines
  - Use search terms for targeted advertising
  - Build profile, deliver targeted ads
- Private Search Engines
  - DuckDuckGo (https://duckduckgo.com/)
    - Uses Bing search results
    - Non-targeted ads
    - ‘Bangs’ for anonymized queries
  - StartPage (https://www.startpage.com/)
    - Uses Google search results
    - Non-targeted ads

Startpage data flow

1 **A user performs a search**

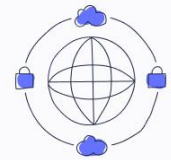
- Encrypted connection
- Perfect Forward Secrecy (PFS)
- A+ from SSL Labs for TLS cipher suite

2 **"Premise servers" remove IP address, and perform additional anonymization and fuzzing**

- Encrypted connection
- Perfect Forward Secrecy (PFS)
- A+ rated TLS cipher suite

3 **"Application Servers" anonymously retrieve search results from search providers, and compose a webpage with results**

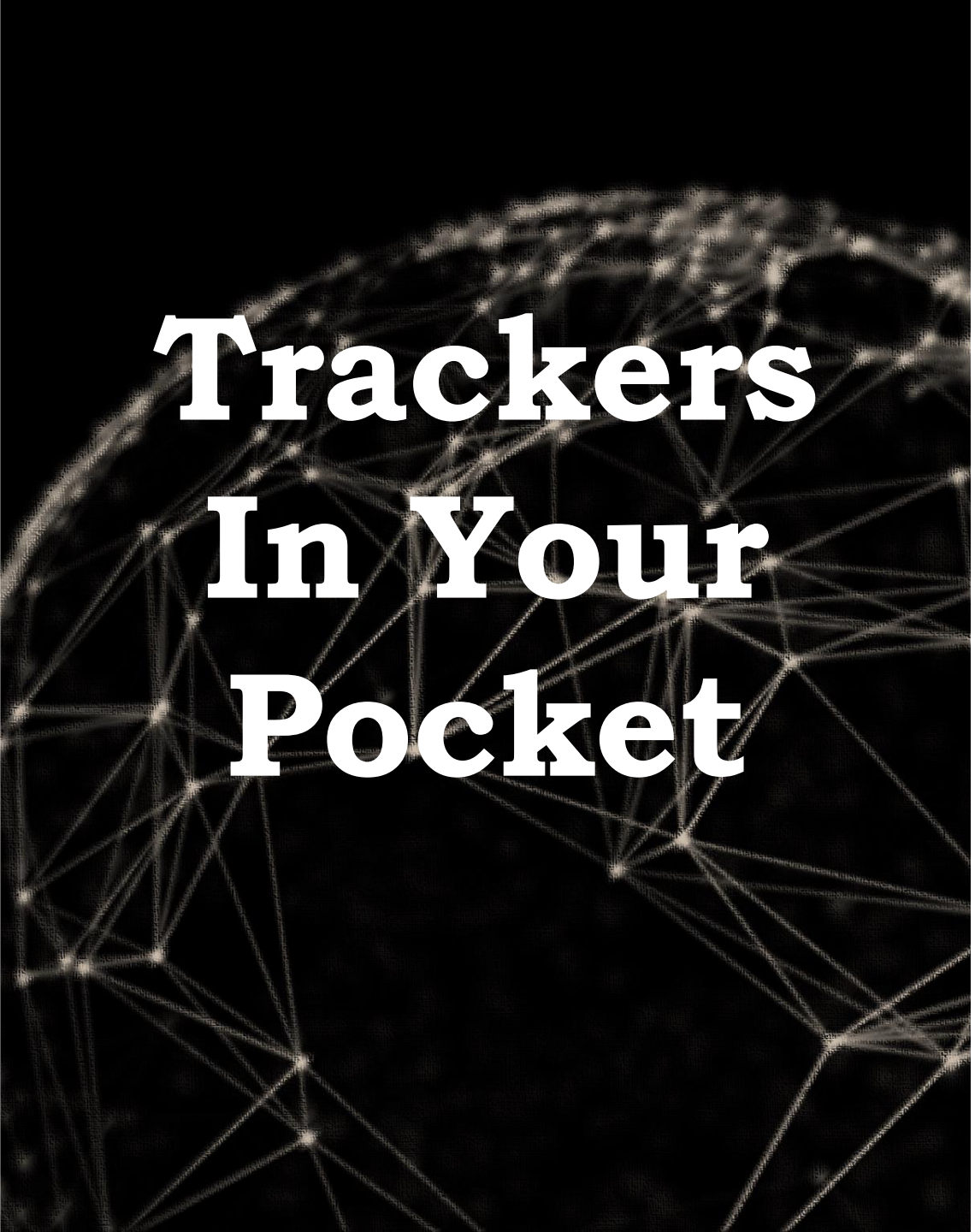
- Encrypted connection
- Perfect Forward Secrecy (PFS)
- A+ rated TLS cipher suite



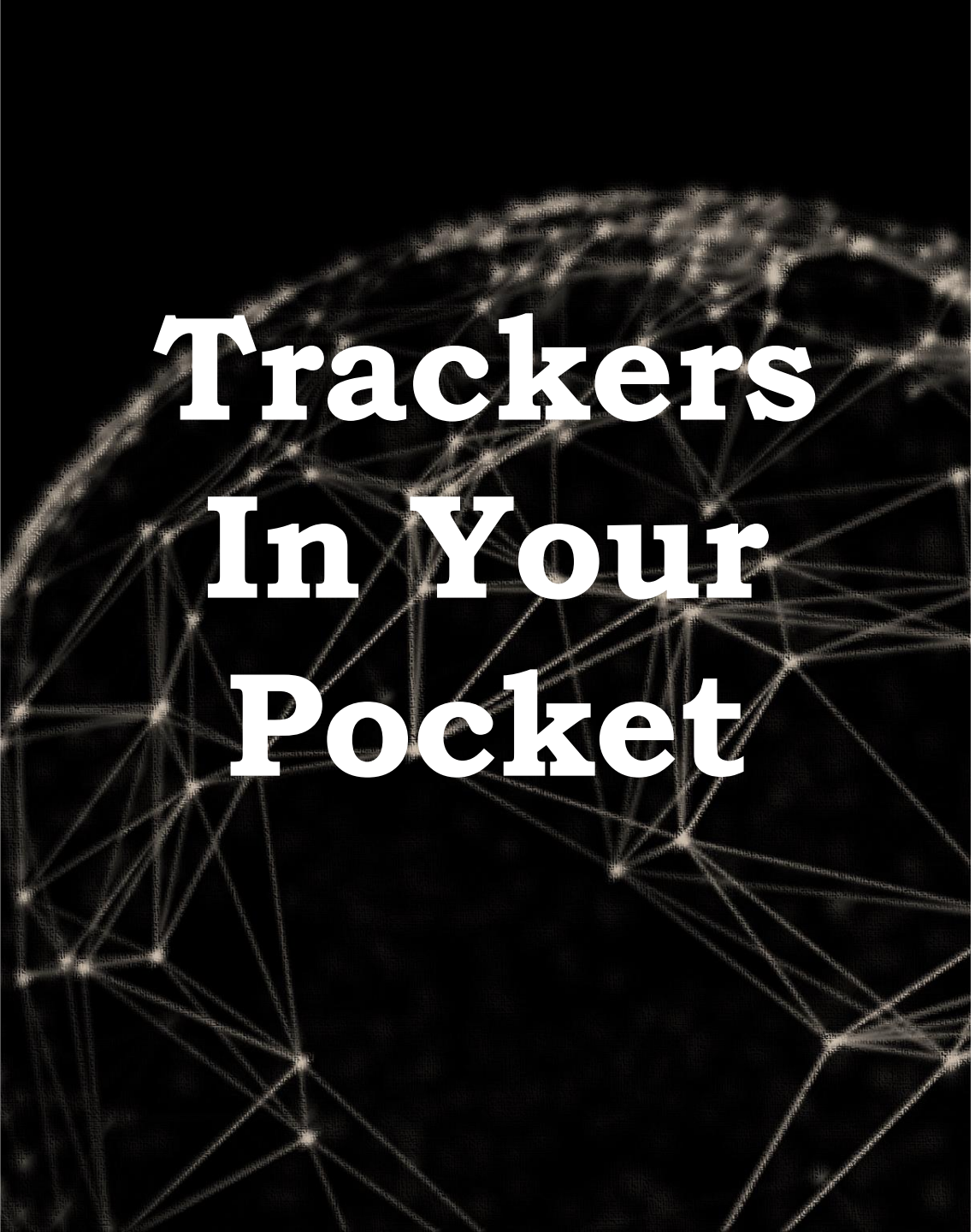
4 **"Premise servers" receive the results and return it to the user**

- Encrypted connection
- Perfect Forward Secrecy (PFS)
- A+ rated TLS cipher suite

5 **The user receives results**



- Data collected by:
  - Manufacturer (OEM)
  - OS Vendor
  - Installed applications
- Uninstallable applications
  - Chrome runs in background on Android phones
- Precise Location Tracking
  - Turn off location when not in use
  - Manage background permissions
- Advertising Identifier (IDFA/AAID)
  - Replaced Apple's UDID in 2012



- IDFA Opt-Out
  - Announced 2020
  - Delayed to Spring 2021
  - Opt-In to Share IDFA with advertisers
    - Cannot penalize customers for opting out
    - Cannot replace IDFA with other identifier
  - Enforcement, penalties unclear
  - 70% of iOS users share IDFA
  - Estimated drop to 10 -15%
  - Facebook launched ad campaign against it
    - Exploring legal options

Apple vs. the free internet

Apple plans to roll out a forced software update that will change the internet as we know it—for the worse.

Take your favorite cooking sites or sports blogs. Most are free because they show advertisements.

Apple's change will limit their ability to run personalized ads. To make ends meet, many will have to start charging you subscription fees or adding more in-app purchases, making the internet much more expensive and reducing high-quality free content.

Beyond hurting apps and websites, **many in the small business community say this change will be devastating for them too, at a time when they face enormous challenges.** They need to be able to effectively reach the people most interested in their products and services to grow.

Forty-four percent of small to medium businesses started or increased their usage of personalized ads on social media during the pandemic, according to a new Deloitte study. Without personalized ads, Facebook data shows that the **average small business advertiser stands to see a cut of over 60% in their sales for every dollar they spend.**

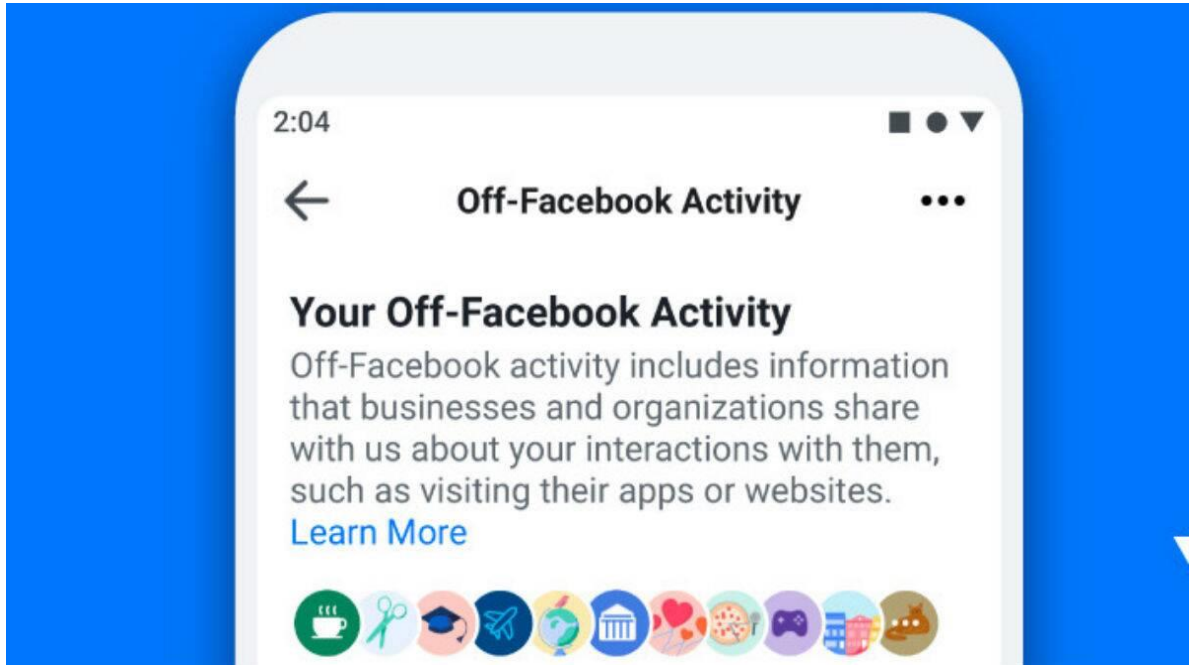
Small businesses deserve to be heard.
We're standing up to Apple for our small business customers and our communities.

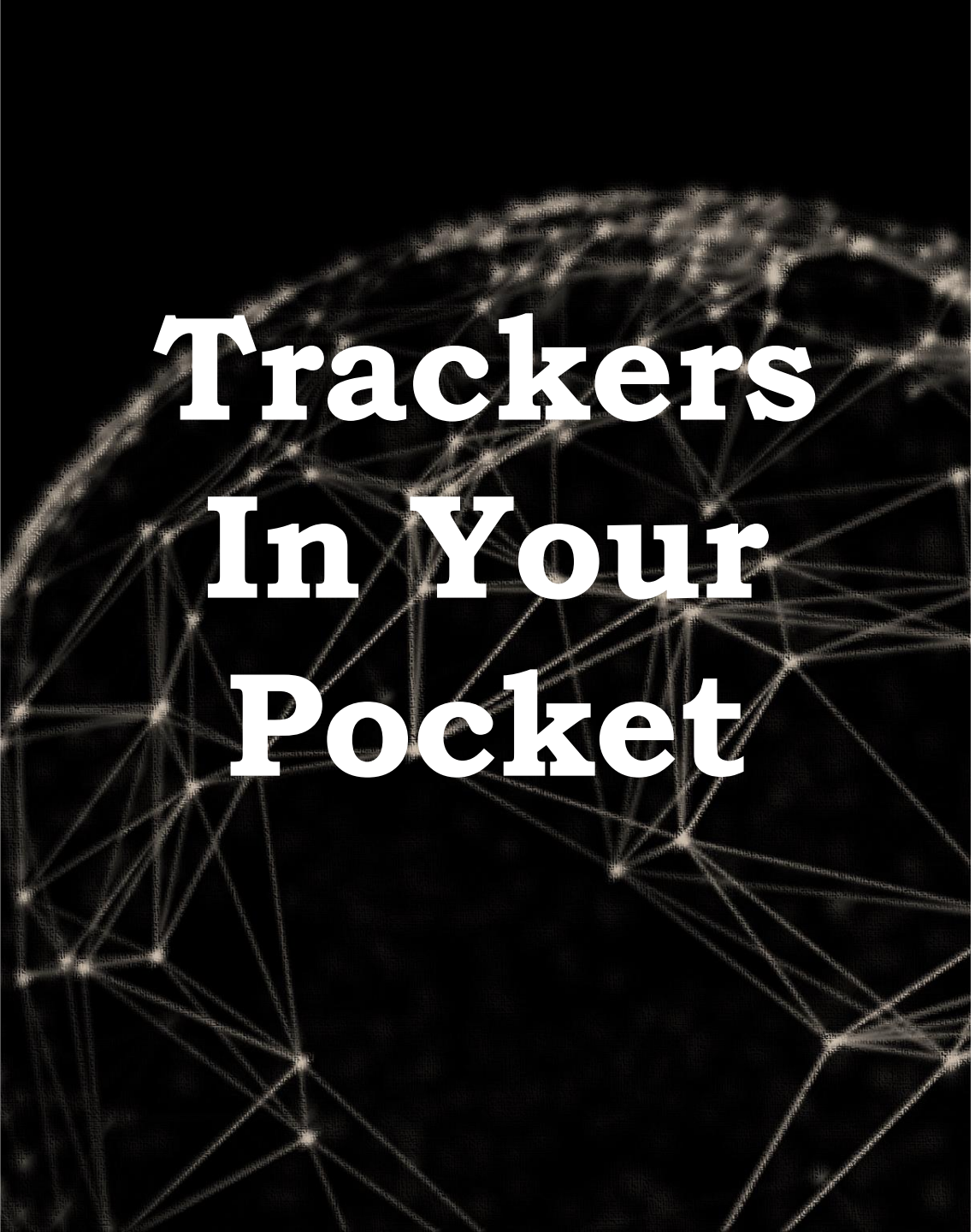
Get the full story at fb.com/ApplePolicyUpdate

FACEBOOK

Trackers In Your Pocket

- Software Development Kits
  - Google Analytics SDK
  - Facebook Analytics
  - User profile augmented with app usage data
  - “Off-Facebook Activity”

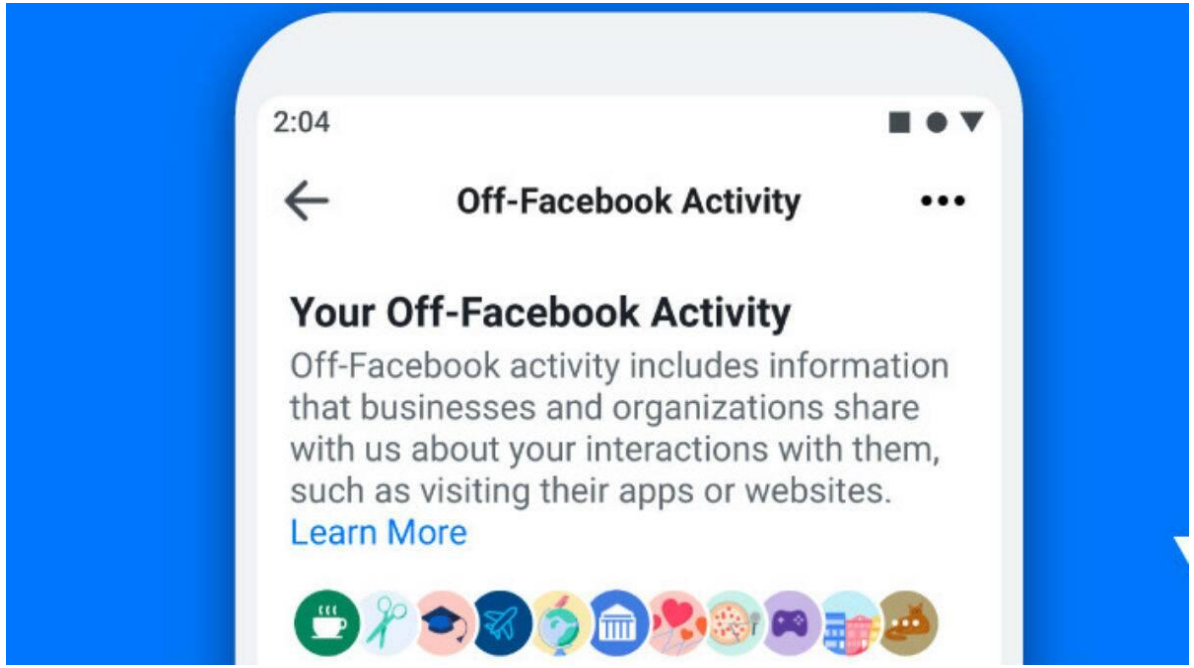




- Click the arrow in the top right of Facebook.
- Select Settings & Privacy, then click Settings.
- Click Your Facebook Information in the left column, then click Off-Facebook Activity.

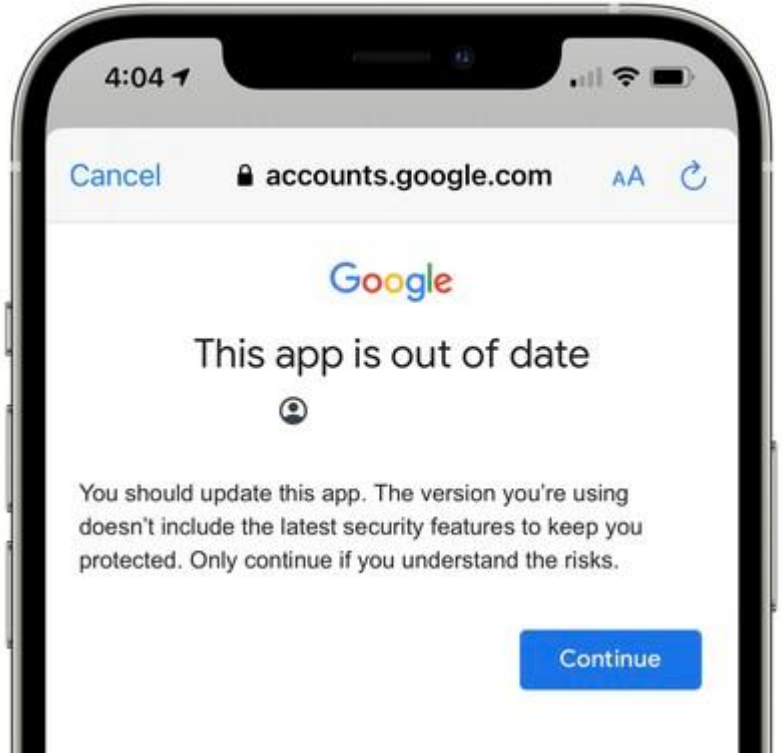
Trackers In Your Pocket

- Software Development Kits
  - Google Analytics SDK
  - Facebook Analytics
  - User profile augmented with app usage data
  - “Off-Facebook Activity”



Good Apps, Bad Apps

- App privacy labels (Apple)
  - All apps required to classify how they collect user data
  - New apps, old apps on next update
  - Google waited 2 months



Signal
'Data Linked To You'

iMessage
'Data Linked To You'

WhatsApp
'Data Linked To You'

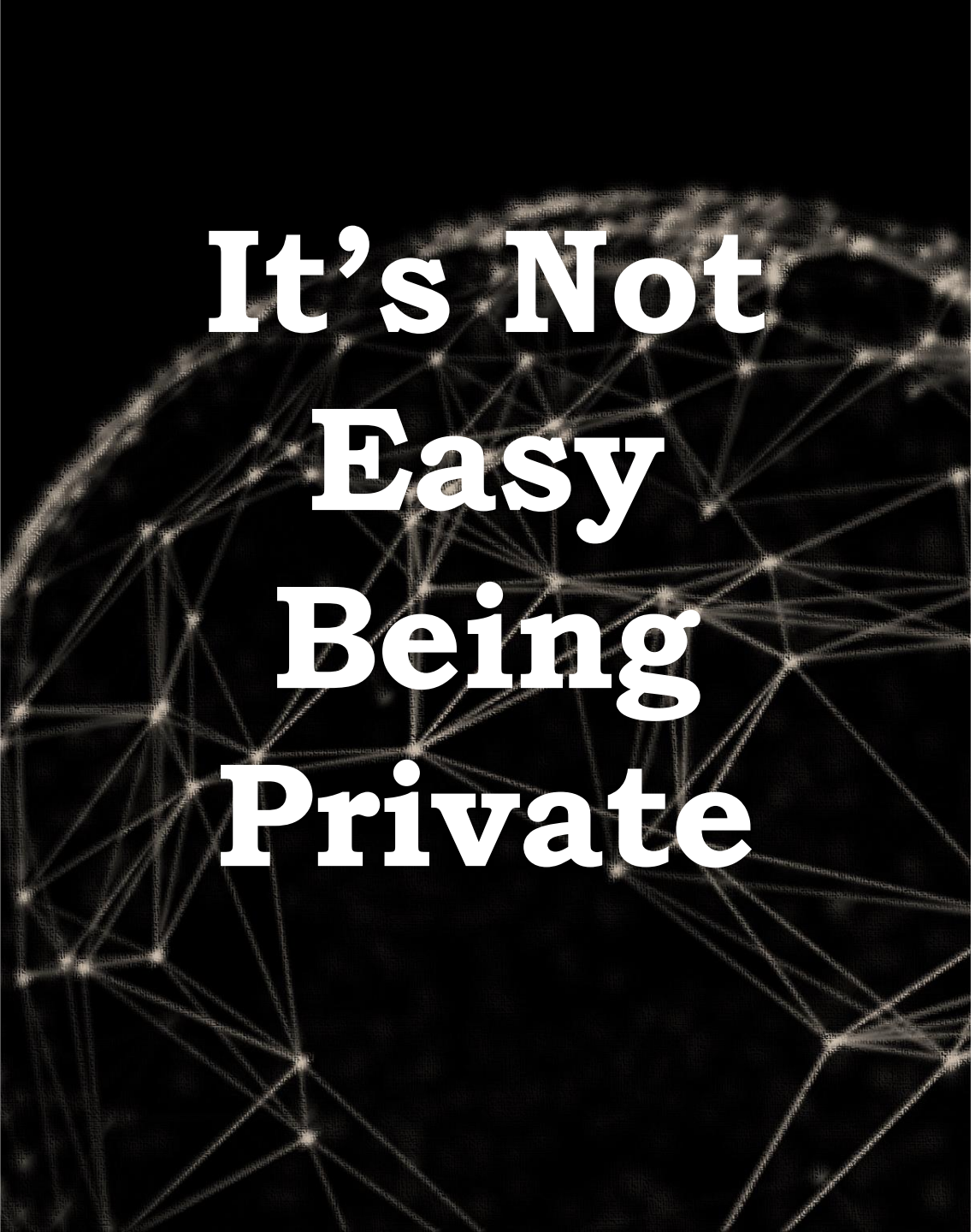
Facebook Messenger
'Data Linked To You'

Good Apps, Bad Apps

- Messaging Apps
  - iMessage
    - End-to-End encrypted
    - iCloud backups use Apple's keys
  - Signal

    - End-to-End encrypted
    - Signal protocol
    - Open Source
  - Telegram
    - End-to-End encrypted
    - More chat-like features
    - Not Open Source

- Email Apps
  - Most major email providers scan your email
  - https://www.zdnet.com/article/yandex-said-it-caught-an-employee-selling-access-to-users-inboxes/
  - Protonmail
    - Zero-knowledge mailbox (SRP protocol)
    - Open-source
    - End-to-End encryption between users
    - Encrypted message sending to outside users
    - Based in Switzerland
  - Tutanota
    - Similar to Protonmail
    - Encrypted headers, proprietary messaging protocol
    - Based in Germany



- Barriers to switching
  - Loss of convenience
    - Big Tech services have better funding, better design (in general)
    - Search results are less relevant for privacy-preserving services
  - Limited userbase
    - Low uptake of privacy-preserving apps
    - Nobody can spy on you…
    - … but there's nobody to talk to ☹
  - Private options cost more
    - Email: storage is more expensive
    - Cost of Apple devices
    - When you're not the product, you have to pay up!