From: Andy Williams
To: privacyrfc2018
Subject: 180821780-8780-01 proposals revision 3
Date: Wednesday, October 10, 2018 2:15:18 PM

Request for Comments on Developing the Administration's Approach to Consumer Privacy

"User" refers to U.S. citizens or customer of U.S. company regardless of citizenship.

User data may include personally identifiable information, behavioral analytics and telemetry, sentiment, financial transactions, personal communication transactions and device logs, sensor data, and identifiers.

The examples provided are not intended to single any company out. Using real-world examples simply gives context to universal concerns.

Policy is retroactive.
Security of Data

Specific requirements for user and administrator authentication. Data must be secured in transit and at rest.

Firewalls, intrusion detection, and prevention must be implemented.

Perform annual formal audit by one or more external third parties. Proof of remediation by X timeframe of all recommended critical, practical, or feasible vulnerabilities. todo: Establish a standard definition of types of vulnerabilities for all companies to address.

Ensure timely security updates to software products and hardware involved in collection, transmission, and storage of data.

Service providers must implement an authentication scheme to prevent spoofing. Neighbor spoofing, for example, is a common abuse in the telecom industry. If someone can pretend to be another subscriber, that is not sufficient authentication to establish a connection.

Systems must record user and administrator authentication and keep the audit immutable for X timeframe.

Social Security Numbers, in whole or in part, must not be used for user verification in commerce. A physical Time-based One-time Password authenticator would reduce the value of personally identifiable information when verifying users.
User Consent and Disclosure

Opt-in is preferred versus opt-out

Companies must prominently disclose any third parties and affiliates who have access to data and what specific types of data are collected and/or transferred to third parties prior to and remain accessible to the end user. The user must receive notification during collection and information on the frequency and amount of data prior to collection. All entities must be represented and individually consented at time of collection and previews of the data each entity collects must be accessible by the user.

Every company represented by a product or service must be disclosed and have individual consent, and ability for end user to preview the actual data that is sent from the user to the company.

If permissions are permanent they must give an additional option to consent for one-time use.

Example: LastPass Password Manager Android app contains code provided by and establishes network connections to at least 4 additional companies: Facebook Graph, Fiksu, Crashlytics, Segment This is a password manager that can have Accessibility permission, Draw Over Other Apps, and Usage Access on a mobile device- each of which has the ability to observe a lot of user interactions. Therefore, each company represented by the host app must present option for consent, provide a preview of the data they want to collect, and alert the user when that data is being collected.

Entities must report what data is collected of the user on an annual basis.

Permissions such as location must not collect more frequently nor more accurately than what is advertised as the primary purpose of the service. And thresholds must be adjustable by the user. Ex: If timezone is sufficient or accuracy within state, region, city, or a practical unit of measurement- sane defaults must be used. A weather app may only need accuracy within a few miles while a navigation service may need a more accurate measurement. If the user overrides the accuracy allowed the location may opt to fuzz a location within the user's desired threshold if the service requires a data point. App wants location in feet. User consents to no more than 25 miles accuracy. App should take a location somewhere at least 25 miles away.

If a service is offered that prompts for an identifier they cannot use it for purposes outside the scope of its intended function without the user giving explicit consent.

Examples: I go to Best Buy and they ask if I want my receipt emailed to me. That does not imply that they can use my email for marketing purposes as I did not consent to that. If a chat app requires my phone number for some sort of verification process, that does not allow them to also use the identifier for user tracking for marketing purposes.

This I think is a big problem: Policy that clearly prevents this from happening. I'm quoting this from a Mozilla bug since it describes the situation best and I believe this sort've behavior occurs often in analytic technologies:

> [triage] Potentially critical - I'm concerned about:
>
> > User disables telemetry
> > We still collect measurements
> > User re-enables telemetry
> > We send measurements collected while telemetry is disabled
>
> Which violates user consent. I'm less concerned about storing data that the user consented to be captured when they had telemetry enabled.

source cited: https://bugzilla.mozilla.org/show_bug.cgi?id=1371447#c1

Imagine if Google or Facebook have an opt-out that pauses collecting history data. In reality it's still logging your behaviors. If you toggle the setting back on to consent- even for an instant- they immediately pull all of that backlogged history.

Ability to Export Data

The user must have the ability to export the data in acceptable standard formats for each type of collected data if they so choose. Examples of standard formats include plain text, xml, json, csv, jpg, png, bmp, tiff, or any other open standard.

Entities Acting as Intermediary to Facilitate Transactions or Communication

If an entity acts as an intermediary to another entity to process the transaction, the intermediary must not be able to view or store the data and each transaction must use a unique identifier. Example: if a user makes a purchase on a credit card machine or online shopping cart, the merchant does not need to see the credit card number nor pass a credit card number. Instead, transmit a unique identifier for that transaction for that particular account so that an attacker cannot collect credit card numbers for later reuse.

Opting Out and Requests for Deletion

To opt-out or remove data, the user must not have to provide more information to verify their identity than what they have collected. That is, the user must not have to mail a copy of a driver's license, write a letter, or provide additional identifiers to verify identity than what the entity has.

To request deletion of user data the user must be able to request deletion electronically, it must be an easily discoverable feature. Opt-out or deletion procedures must be free of charge to the end-user.

Opt-out should not require persistence or maintenance. Technologies such as opt-out cookies or routine renewals put undue burden on a user.

Example: the Network Advertising Alliance http://optout.networkadvertising.org allows users to opt-out of data collection from several advertising and analytic companies via a browser cookie or mobile application. If the user uninstalls the mobile application or deletes the browser cookie, collection and personalization must not resume. Therefore, this is not a valid persistent opt-out. Google, The Digital Advertising Alliance, TRUSTe, Allspark are other entities that offer and require apps or browser extensions to maintain some form of opt-out control. This puts a lot of burden on users to have knowledge of each technology to maintain consent nor is it all-inclusive of the thousands of other companies that do not offer such mechanism.

If the user requests a company to delete data, regardless of anonymization, they must delete all rows and data in their systems and backups. Not a delete/hidden flag but a hard delete. In some industries, such as healthcare, databases are not designed to hard delete data. Patient data critical and requires auditing. There may be a "delete flag" that is flipped in a row in the database but the record still exists on the backend. If the database is compromised, the attacker can see this soft deleted data because it is still present. Behavioral analytics is not crucial to keep.

Anonymization does not supersede the user’s choice to opt-out or remove their data. If the user does not want their data collected, it must be removed even if it was anonymized.

Opt-out and deletion should take no longer than 30 days with an optional choice to rescind request within the 30 days at the company's discretion. An example might be the user requests deletion, has the choice to undo the request for 14 days, after which the company has 16 days to proceed with deleting.

If deletion is unfeasible, all data associated with the user may be scrambled or defaulted to an ambiguous value to maintain database integrity. That is, not simply attributed to an anonymized identifier but all data on that user is set to a default ambiguous value for that data type (ex: datetimes set to '1900-01-01 00:00:00:00', varchars set to 'ZZZZ'). A database query can simply be designed to exclude "where not equal to 'ZZZZ'" to exclude all data on users that have been made ambiguous.

https://chiefmartec.com/2018/04/marketing-technology-landscape-supergraphic-2018/

https://datatransparencylab.org/