Dataccount System Overview

Problem statement

Many companies maintain databases of personal information about individuals, which they use for their own purposes:

- Web sites like Facebook, Amazon, and Google monitor your online activities, then sell the information to advertisers for their own profit, with no compensation to you.
- Roomba's robot vacuums make a map of the interior of your home and send it back to the company. Roomba uses it to improve their robot algorithms, again with no compensation.
- Amazon keeps a permanent database of your personal home address and credit card information, which they can use as they see fit.
- Doctors keep your sensitive medical history in their own databases. You have no control over it.

The first problem is that information about yourself is stored by many companies, who generally promise to keep it safe, but you have no concrete way of verifying it. Thus the companies may accidentally expose information which is private or embarrassing, or which allows other people to impersonate you. They can make money from your information without compensating you. You just have to trust them.

The second problem is for the companies. By keeping individuals' information, they themselves are exposed to scandal, lawsuits, etc. if they lose the information, for example to hackers.

Proposed solution

This document proposes a combination of computer systems and laws that would:

A. Allow companies to access personal information under the control of the individual.

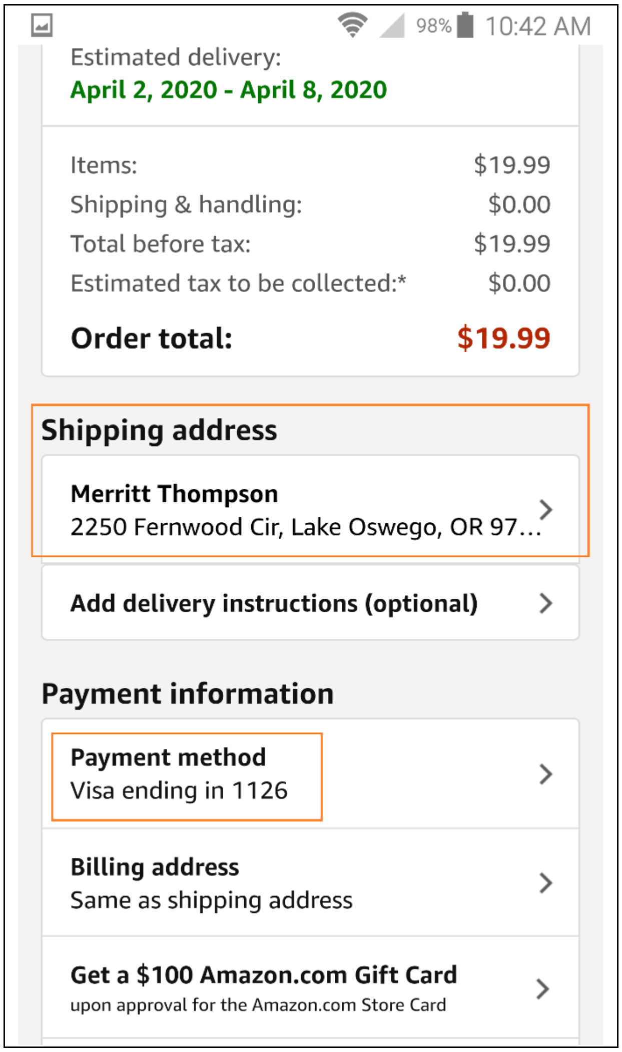
B. Allow individuals to track accesses to their own personal information.

C. Prevent companies from storing personal information.

A. Allowing access under individual control — Example

When I order something on Amazon today, they show me a screen that contains my address and credit card, as shown on the left. This shows how Amazon maintains my personal information in their private database.

The image on the right shows the same screen after Amazon has upgraded to use this system. All personal information is gone. Instead there is only a web address for a “dataccount” (“data” + “account”). This shows how Amazon doesn’t keep my personal information. They get it from my dataccount, which is at “https://merrittmarcthompson-dataccount”:



Estimated delivery:
April 2, 2020 - April 8, 2020

Items:	$19.99
Shipping & handling:	$0.00
Total before tax:	$19.99
Estimated tax to be collected:*	$0.00
Order total:	**$19.99**

Shipping address

Merritt Thompson
2250 Fernwood Cir, Lake Oswego, OR 97…

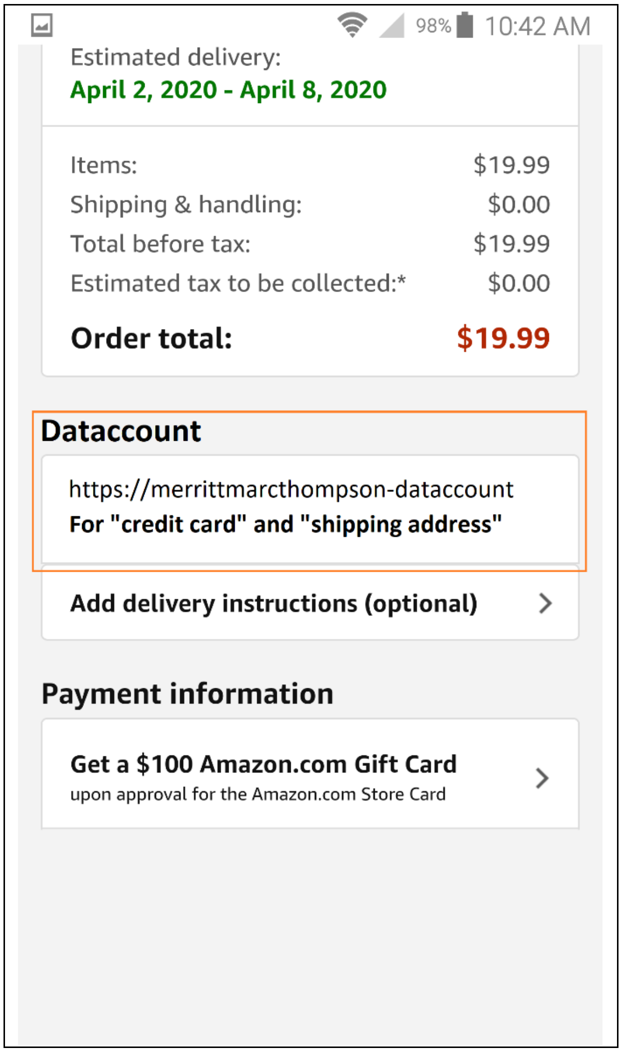
Add delivery instructions (optional)

Payment information

Payment method
Visa ending in 1126

Billing address
Same as shipping address

Get a $100 Amazon.com Gift Card
upon approval for the Amazon.com Store Card



Estimated delivery:
April 2, 2020 - April 8, 2020

Items:	$19.99
Shipping & handling:	$0.00
Total before tax:	$19.99
Estimated tax to be collected:*	$0.00
Order total:	**$19.99**

Dataccount

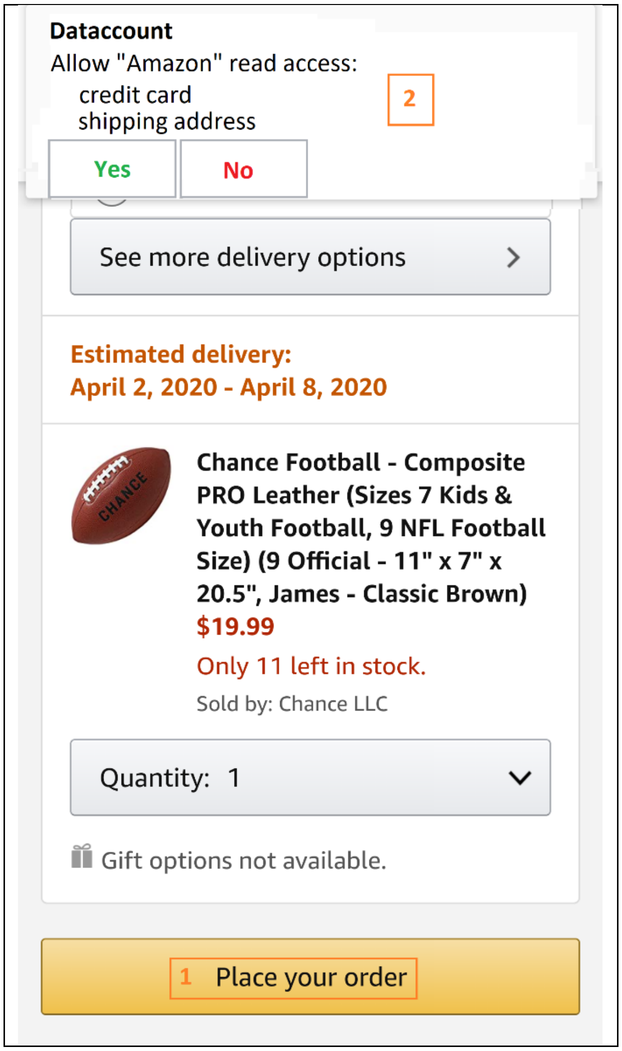
https://merrittmarcthompson-dataccount
For "credit card" and "shipping address"

Add delivery instructions (optional)

Payment information

Get a $100 Amazon.com Gift Card
upon approval for the Amazon.com Store Card

When I click the Amazon "Place your order" button (1 below), Amazon tries to access my dataccount. That causes a push notification from the Dataccount software on my phone (2), which lets me approve the access. This allows Amazon to bill my credit card and print a shipping label without permanently storing my personal data:



Dataccount
Allow "Amazon" read access:
credit card
shipping address

See more delivery options

Estimated delivery:
April 2, 2020 - April 8, 2020

Chance Football - Composite PRO Leather (Sizes 7 Kids & Youth Football, 9 NFL Football Size) (9 Official - 11" x 7" x 20.5", James - Classic Brown)
$19.99
Only 11 left in stock.
Sold by: Chance LLC

Quantity: 1

Gift options not available.

B. Tracking access to personal information

When a company accesses my data, my dataccount software logs the access and its purpose:

Access date/time	Accessor	Data Item	Access	Allowed?	Purpose
4/7/2025 19:03:27	www.amazon.com	credit card	Read	Yes	Purchase $19.99 football
4/7/2025 19:03:27	www.amazon.com	shipping address	Read	Yes	Print delivery label

C. Restricting storage of personal information

Thus far, this system gives external entities a way to access necessary information without storing it, but it doesn’t stop them from doing so anyway. I.e. they can still access the information from the dataccount, then save it privately for their own use or profit.

This would be prevented by new laws that would criminalize storage of other people’s personal data for purposes other than the stated ones.

In the example above, if Amazon got your shipping address, then used it for any purpose other than printing a delivery label (for example, if they sold it to a mass-mailing company), they would have committed a crime.

This law would criminalize what is now an extremely common practice. You might think that companies would oppose this restriction, but it is more likely that they would approve of it, because as long as they follow the rules, they are completely absolved of liability for individuals’ data losses.

The goal of this system is not to magically track your data wherever it goes via a complex encryption scheme. The goal is to make not storing individuals’ data the norm in society. The law establishes the norm and the computer system provides a practical way to obey the law.

About compensation

In this scheme, your Roomba vacuum could make a floor plan of your house, then send it directly to your dataccount (sending it anywhere else would be illegal). Roomba could then buy your floor plan from you. I’m not sure if that would make you much money, but Roomba might give customers perks of some kind for it. You could also sell it to researchers other than Roomba, since it is your data, not Roomba’s.

Another example like this would be selling medical data to researchers. If you have a rare or interesting condition, you might get a reasonable amount of money for your medical records.

Implementation overview

Dataccount location

A dataccount can go on any server that is on the web, so long it supports the dataccount web API. Thus, you could have your own server locked in a closet in your house. Or you could contract with a “data bank” company that provides encrypted dataccount services for many individuals.

Data format

Personal data would be stored in the form of key/value pairs, for example:

"Home address"="Jane Smith, 123 Shady Lane, Tulsa, Oklahoma, 74101"

Nondisclosure of data in keys

Imagine that you want to send a diamond ring to your mistress Sheila's home address. Let’s say you already have a data key called "My mistress Sheila's home address". So, you might fill out a field like this on the Amazon web site:

Shipping Address Data Key: `My mistress Sheila's home address`

This would be less than ideal as it would expose personal details of your life. To prevent this, you could create an *alias* in your dataccount:

"Amazon shipping address"=@"My mistress Sheila's home address"

This means, “Whenever someone asks for ‘Amazon shipping address’, give them ‘My mistress Sheila’s home address’ instead”.

Amazon would default their form to the following, so you probably wouldn’t have to enter anything:

Shipping Address Data Key: `Amazon shipping address`

Web APIs

There would be two APIs:

- One for individuals (“owners”) to manage their own personal information. This includes functions that support the on-the-fly push notifications for Allow and Deny illustrated above, plus data item maintenance, viewing logs, etc.
- One for companies (“accessors”) to access personal information of others. This is just one function to access owner data, providing the names of the data items, what access is requested for each (read, write, or both), and the purpose of the access for each.

Owner authentication

Every owner API call is authenticated using either HTTP Basic authentication or OAuth. Encrypted HTTPS transport must also be used for HTTP Basic to work, as the password is merely obfuscated.

Accessor authentication

The most important thing for the accessor API is to authenticate the accessor. For example, if someone tried to access your dataccount, said they were “Amazon”, and the dataccount simply believed it without authentication, anyone could impersonate Amazon and get your information.

Transport Layer Security (TLS), the basis of the HTTPS communication protocol, normally requires the server to prove who they are to the client before any data can pass between them. The server does this by using a secret private key to encrypt some data. If the client can successfully decrypt the data using the public key, it proves that the server is who they claim to be.

But that process is the opposite of what this system requires. In this system the server (the dataccount) must obtain proof that the client (the accessor, ex. Amazon) is really who they claim to be, not the other way around.

Fortunately, there is a “two-way” variant of TLS which provides exactly this function. It both verifies that the server is who they claim to be to the client (as normal) and the client is who they claim to be to the server (needed for this system).

Push notifications

When a company contacts your dataccount web API to access data, your server must immediately notify your mobile device for approval. It’s possible to set up a “service worker” via JavaScript, which always runs and waits for such notifications.