A Purposeful Conspiracy?

PREFACE

"I can't take it anymore. He snores so loud. I just can't get any sleep. I don't know what to do. I know it's not his fault."

I laid under the covers in my bed trying to stay as still as possible.

This was the first time I was hearing about my apparent snoring problem.

My roommate was miserable. He couldn't get any sleep. Earplugs weren't working.

In a stroke of luck, I was hearing what my college roommate truly thought of me as he talked to his friends – oblivious to the fact that I was actually in the room at the time, barely waking up from a nap.

=====

When I was 12 years old, I came home from middle school and found that all my belongings had been searched through and reorganized by my mom.

She had a thing for cleanliness and thought she knew best of what was important enough to keep and what should be thrown away.

I felt violated.

There were things I wanted to keep – things that were important or had sentimental value to me. They were now gone.

This experience contributed to my habit of keeping digital records instead of physical notes and objects. My mom might be able to get into my dresser and desk – but she didn't know her way around a computer.

======
These experiences still replay vividly in my mind. They've contributed to my understanding of concepts such as secrets, privacy, and transparency.

Which brings us to now.

Today, we live in a world where everyone wants our data: Facebook. Google. The stores we shop at.

We're told this is a good thing. Data helps companies and the government improve our lives.

But is there a cost to free-flowing data?

Are there downsides? Maybe things like NSA surveillance or the Equifax hack.

Sometimes it seems like the world is against us. Like the powerful people are colluding to control the common man. We have no choice but to follow.

Secrets are kept from us. To protect us. But we aren't allowed to keep secrets. Privacy is dead we are told. But that is OK.

The only people who want privacy are the hackers trying to ransom us. They're the ones selling drugs or planning terrorist attacks.

And so, every time you access the internet, you are constantly being monitored and tracked. By governments, companies, and individuals. The use of the information may vary – but we have little to no say.

The question remains: Is there a conspiracy? Are people and businesses collaborating to build technology that violate privacy?

If so,

What will they gain?

What will we lose?

Who can we trust?

Chapter 1: Why People Want Your Data

This chapter discusses all of the reasons people want data and why it's so valuable. The reasons may vary

- Business Optimization: Businesses like Dell and Chik-Fil-A use data to predict when people will buy and optimize their supply chain so they buy materials at just the right time. Businesses also use data to help them understand which of their marketing efforts are working.

- Publicity: People and businesses use data to get in the news. Studies like chocolate is good for you, or that men regardless of age prefer to date women in their twenties (based on OKCupid's data) allow people to get in the news.
- Customer Service: Having a centralized database of customer information allows any customer service representative to help you because they can see all previous interactions you had with the company.
- Public Gain/Progress: Smart streets allow city planners to optimize the infrastructure to minimize crime and traffic. DNA analysis allows us to look for cures to horrible diseases. Law enforcement surveillance and data allows them to protect us from threats.
- Money: Data can tell us when the cheapest time to buy a flight is. Data can be stolen from others to enable identity theft. People and companies can compile data and sell it to businesses for optimization.

Data can have personal and communal benefits. Individuals and organizations take advantage of these benefits.

Chapter 2: Privacy vs Transparency

This chapter challenges the common perception that privacy is a tradeoff against *security*. A better way of looking at privacy is as a tradeoff between *transparency* and privacy.

Both transparency and privacy have their benefits. Privacy enables surprise, empathy, and change. Transparency enables accountability.

But both of these principles have major risks. Privacy enables *corruption* to go unchecked. Transparency exposes corruption but introduces the risk of *manipulation*.

Privacy & anonymity allows horrible, heinous things. Things like online harassment and bullying. The argument is that it people had to show their faces, they wouldn't say or do the same things. The KKK was cloaked in anonymity. Russian ads are cloaked in anonymity. Hackers are cloaked in anonymity. If we just identified everyone, if everything everyone ever did was attributed to them, the world would be a greater place. If we can connect every person's actions, words, and thoughts to a name, people can be held accountable and evil will be purged.

The big complicating for is that “corruption or "evil" is determined by individual cultures. Christianity may be seen as corruption in Muslim countries. Advocating for gender equality is corruption in Saudi Arabia.

Transparency in those cultures let people in power know who is for or against a certain position, they are able to quench rebellions, difference of thought, or diversity because openness also allows groups to be targeted.

Other forms of manipulation that come with transparency include: people being able to circumvent systems, taking things out of context, catering to an audience, and intimidation.

Throughout history, we see a cyclical pattern between gatekeepers and data democratization. Twenty years ago, used car dealers had more information than the average buyer. They were gatekeepers — people who learned about cars so regular people wouldn’t have to. But dealers could also manipulate buyers and take advantage of that knowledge. The internet solved this information asymmetry through data democratization. Making data freely available to all would solve all this corruption. But today we see a push back towards gatekeepers. People took advantage of the platforms meant to make data available to all. Russian ads. Online harassment. There is pressure for Facebook, Google, and Apple to become gatekeepers — to be responsible for every app in their app stores, to verify identities of ad buyers, to ban hate speech. If history is any teacher, there will be another wave of data democratization in the future.

Chapter 3: Extracting Value Out of Data

This chapter exposes the main value of data as one of correlation. Even the most seemingly innocent/innocuous data points can be used to deduce surprising knowledge. Metadata about where you travel can reveal a lot about you. If you are in one location between 10pm and 7am, that’s likely your home. If you’re in another location from 9am to 5pm, that’s probably a job. If you go to a dentist, we can deduce you have teeth problems. The stores you shop at or the neighborhood you live in may reveal your income level.

Our data goes through three stages: collection, processing, and analysis.

Chapter 4: Technology and Privacy

This chapter challenges the notion that technology is developed explicitly to violate our privacy. Technology is usually built with good intentions. But there can be unintended consequences.

Technology is neither good or evil. Like guns and fire, our fight isn't against technology, it's against people. Technology is morally neutral. It's an amplifier and amplifies whatever is already there.

People find creative ways to misuse technology and appropriate it for their own purposes. Remember the tech axiom: *For every good use of technology, there is an equal or greater misuse/negative misuse*

Consider browser fingerprinting, which relies on finding about what browser you use, what operating system you use, what plugins you have installed, and what your screen resolution is. Giving this information to any website was not part of some grand plan to surveil you, but to give developers the ability to optimize a website for specific audiences.

Chapter 5: We Are Never Free

This chapter discusses the ways we can't control data about us. Data is generated by computers automatically without our permission. Every time you log into your email or edit something in the cloud, there are time stamps and IP logs.

Furthermore, data is usually exchanged – which means two people have access to the data. You might not be on Facebook, but a friend can still upload a photo of you. You might delete emails, text messages, and more but the other recipient still has them – not to mention service providers. You might not share your contact information with companies, but if a friend shares their phone's address book with an app, it gets out there.

Even if we don't generate data ourselves, other people will. When you reserve a hotel, fly, or go to school, data is generated. Equifax compiles data on you from banks and lenders, including sensitive permanent identifiers such as name, DOB, SSN, and drivers license numbers.

Using third-party providers mean you are under their rules. You can't say no even if you want to.

One day my dad came home from work. His Social Security Number was hacked. Not by his own doing but because someone in the corporate office - in the accounting department received a phishing email. It used the name of her boss and mimicked his email address. It asked for a list of employee W-2s. She complied and it was stolen. How can you prevent against that? Your employer needs your SSN to pay you and the government. It was in the hands of a third party. Better training and systems to develop trust are critical. On an individual level, you can choose not to work or find something under the table.

Chapter 6: Security & Privacy

Without security, you can't have privacy.

When I'm talking about security I am not talking about the concept/sense in which security means being protected from dangers by authorities who have access to private data — as is often postulated. Instead I am talking about security in the technological information systems sense.

Security is a kind of lock that ensures data is confidential and accessible by only authorized individuals, and that the data has not been tainted or corrupted by anyone else. Encryption is an example of this. Data cannot be seen by anyone except those with access to the data.

In today's world that means only you (or anyone with your password), as well as the tech companies whose services you are using will have access to your data.

Privacy is just one subset of security. Security is more about authenticity and reliability (more like confidentiality, integrity and availability of data). Only the person who should access it should access it. The ability to authenticate paradoxically means making one uniquely and personally identifiable. If the system doesn't know it's you, it can't give you access. If the makers of security systems analyze your data, they know you.

Is it security versus privacy? No. Security enables privacy (think encryption). Privacy disables security (think ID matching and how disabling identifying features means a security system can't authorize you with knowing it's you). In the old days you had an email and password you set up - nothing could be connected to you personally or uniquely identify you. Nowadays, 2FA means they have a phone number to connect everything with an individual.

Chapter 7: Power Structures and Privacy

This chapter exposes how power increases privacy available. Celebrities find ways of achieving privacy with tactics like buying homes under an LLC. Furthermore, organizations and governments like to force transparency of those under them but resist transparency for themselves.

Chapter 8: Spam: An Unlikely Hero

This chapter exposes that data isn't all its cracked up to be. Computers aren't omniscient. There are flaws. The best way to maximize those flaws are by learning from spam.

Modern data relies on statistics. In statistics, correlation doesn't equal causation. We need to reduce correlations by increasing correlations. Make more people have the same correlations. Crime shows rely on correlations. At the the right place, at the right time, with motive or prior knowledge of victim. Specific correlations like a particular plant or fertilizer molecules. But if multiple people have those same identifying marks, nothing can be proved.

Bad data is what companies are trying to fight. Bad links. Fake reviews. Ad Fraud. A lot of times we focus on when data gets it right – we don't realize that data gets it wrong too. Target knew one person was pregnant - but how many times were they wrong. Someone can report one thing, but do another.

Chapter 9: What's The Solution?

This chapter discusses the different approaches that have been considered for protecting privacy and the dangers of manipulation. Some say it's up to consumers, others favor government regulation, still others say developers are the key to solving privacy violations.

Chapter 10: Taking Action

This chapter introduces the principles of security on the consumer level. The three main principles are blocking, distortion, and keeping it local.

Chapter 11: An Ode To Trust

In the end, perfect privacy isn't possible. That's not what we should be chasing. We should be focusing on identifyingwho we can trust.

Appendix: Practical Privacy

This is a large section describing specific attacks, protections, and steps to take to maximize privacy.