THE INDIA FORUM
A JOURNAL-MAGAZINE ON CONTEMPORARY ISSUES

October 20, 2020

The Digital Panopticon and How It Is Fuelled by Personal Data

By: Anurag Mehra

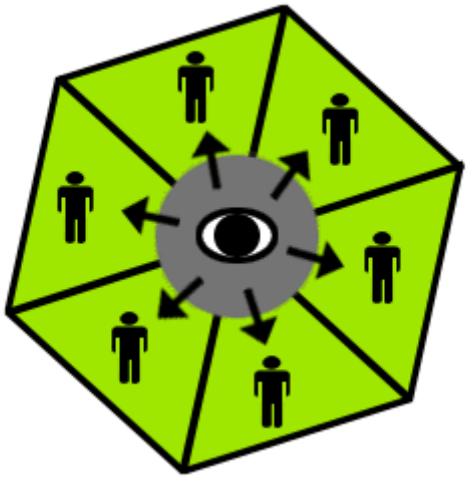
Our phones are continuously leaking data about who we are, where we are, what we read, what we buy, and a lot more. The data is being collected with and without our consent. It is sold for profit; more dangerously it can be used to modify our behaviour.

We live in a world that is overwhelmed by digital technologies that thrive only on personal data. This data is being extracted from us and processed, relentlessly by private companies, state agencies, and in public spaces, sometimes coercively and often without consent. Some specific data may be needed, for genuine reasons, by government agencies or private entities to provide a specific service. But the amount of data that is being taken is humungous and goes far, far beyond the routine.

The threats arising from this are not merely about the embarrassment that may be caused by some intimate details of our lives becoming public but that the extracted data can be used to manipulate and control us. In more harrowing situations it may lead to being discriminated against and be hounded by state agencies. The greatest threat is to our 'free will' and political freedoms, of expression and to dissent.

New tech, more data, more money

Even as you read this, some data *about you* is being recorded, stored, and processed, and even being sold, somewhere! You leak data just by being online. The great digital panopticon has already been built and we are all trapped inside it. Every day brings newer technologies that are more intrusive than the previous ones. They will capture even more data about us that will flow into data banks of the companies that sell us the technology and the app.



The most ostensible reason why apps take our data is because they customise their services and products to our 'needs'. If we read an article of a certain type, that choice data is used to show us more articles of that type. If we search for something to buy or actually make a purchase, then we will see ads of 'more like this' products or services, embedded in the websites we visit and the apps we use. Somewhere, an algorithm has recorded our choices and calculated what else we are likely to buy.

> Extracted data is stored, cleaned, classified, analysed, mined, shared, and sold. It is used for creating user profiles that can be eerily accurate…

Search engines like Google track and store data literally about everything we do as do social media platforms like Facebook, Twitter and sundry others. Google takes money from the advertiser and provides the ad placement service that shows us 'relevant' advertisements, based on this data. These ads are more effective than random ads because we are likely to have some interest in the

products on offer. Facebook allows companies to target their ads at a population segment that is most likely to be manipulated into making a purchase; they may even sense our mood and then make a sales pitch.

The huge revenues that accrue to these companies on account of such advertisement services have made them behemoths. This is one of the most significant uses of our personal data.

Amazon knows what we like to buy and how much we spend (too many books on politics). Zomato and Swiggy know what we like to eat and how much (too much fat). Uber and Ola know how much and where we travel to (where we work and live, and other places that we frequent).

Extracted data is stored, cleaned, classified, analysed, mined, shared, and sold. It is used for creating user profiles that can be eerily accurate in terms of estimating our psychological and behavioural traits; discovering the advertising 'hooks' that work on us; predicting what we may like to read or buy (intent); and even sense the dominant patterns of our moods (sentiment).

> Data, of all varieties, can be bought from data markets operated by thousands of data brokers, big and small, who source this data mostly from app makers.

Data is also used to train, validate and refine the algorithms (Machine Learning, ML; Artificial Intelligence, AI) that work 'behind the scenes' to 'improve the quality of experience', specifically, in carrying out the above tasks of pattern recognition and profiling.

Data is used by app developers to design and fine tune their apps. Many apps take data of all kinds, often with permission, which is not relevant to the functioning of the app. For instance, a weather app may record your age, gender and network details, none of these have any relevance to reporting the weather. While phone makers—Google and Apple—have tried to restrict this, such data extraction is still widely prevalent.

Data, of all varieties, can be bought from data markets operated by thousands of data brokers, big and small, who source this data mostly from app makers. For app makers, who have harvested this data from their customers, this can be an important source of revenue. Data markets are a huge business often specialised across different data types. For instance, the Covid-19 pandemic has created even more opportunities in this market with brokers acquiring sensitive personal data, notably health data. A report, on India, cites a price of Rs 5,000 for data for one Lakh individuals.

Risks of unauthorised data access

We leave data trails all over the digital universe mostly unaware of the extent to which private and public agencies have an interest in them.

The threats emanating from state agencies and corporates having access to so much personal data are manifold. Imagine the consequences of profiles containing health or financial data being sold to hospitals, insurance companies, realtors, banks, and so on. In each of these cases, goods and services may be offered at prices determined by the data.

> The consequences of unauthorised access, by a variety of private and public entities, to … data can be devastating.

Health insurance may be denied, or the premiums made prohibitively expensive if the data suggests ill health. Food shopping profiles may suggest whether a person is likely to die of cardiac ailments, while a sex-toy app may be able to estimate the probability of a person's acquiring sexually transmitted diseases. Reading habits will readily reveal political orientation. The consequences of unauthorised access, by a variety of private and public entities, to such data can be devastating.

Political manipulation is fatal for democracy

It is one thing to be manipulated into buying commodities and services, but quite another into buying a president or a political party. Targeted ads that sell you a political idea or person use your data to estimate what you like and then show customised ads that only you and some others like you will see; public ads that lie or mislead can be exposed but these 'private' ones are meant only for you. Political ads exploit intensely emotional biases, of race, religion, caste, nation, and so on, and distort electoral outcomes by selling you leaders who should have been elected (or not) based only on rational considerations. We give our data and this gives the recipients the power to modify our behaviour! For this reason, targeted political ads on social media should be banned completely.

Your phone as a spy

Here is what your mobile phone—and the app environment it comes with— does (and this is not an exhaustive list) in terms of personal data extraction. It should tell you what an efficient little spy you have in your pocket.

Your mobile phone leaks out your location data; keeps track of your app/activity history; uses search services that record your browsing habits; provides access to social media apps that capture a variety of personal data; offers us health and fitness tracking apps that record and store a lot about your health indices.

Click here to see what your browser allows others to know about you, without needing any permissions from you. Read this to know more about the inadvertent digital trails you leave.

Miniaturised electronic chips and cheap sensors that have made wearable tech possible—like smartwatches and fitness bands—have enabled more detailed data to be extracted. Because these tiny things can sense or track stuff that the phone cannot, like the type of workout you are doing, electrocardiograms, sleep patterns etc. The recorded data is organised and then transmitted to app-affiliated data centres for analysis to find out what health problems you might have.

The lure of tech

Here is some upcoming 'connected' tech from the recent Consumer Electronics Show (CES-2020): Samsung's artificial 'digital' humans (NEONs) to assist you in tasks that need a 'human touch', meditation technology (Umay Rest—to soothe your eyes thermally to compensate for the ill-effects of excessive screen time), blood pressure monitoring with a little pump inside a fitness tracker (H2-BP band), sperm test (YO Sperm), body weight and posture monitoring (Mateo). All of these will record and transmit personal data in a big way. Not content with just these, tech companies are now coming for data related to your sex life as well, of course, all in exchange for technologically aided orgasms. As this Lovely app—which works in conjunction with sex toys—suggests, "Wear Lovely during sex, get Lovely suggestions after sex, and have even better sex next time". Already the debate about privacy and security focused on these connected gadgets has begun.

Why we surrender privacy: Convenience & pleasure

The simplest, near-term argument is that we 'willingly' share personal information with apps and websites in exchange for convenience. This convenience is not just about asking digital assistants to play music or sharing a file; it is much more about broader objectives of increasing productivity, enhancing the 'efficiency' and speed of tasks and 'being on top'. We allow Google to 'machine read' our documents and scan our photographs in exchange for free services like cloud storage and easy sharing. Convenience drives the entry of digital assistants who can hear all our conversations—like Alexa, Google Assistant, and Siri—into our homes because we can ask them to play music or operate gadgets through voice commands.

An all-encompassing form of convenience is provided by a microchip implant under the skin, in Sweden. These chips 'speed up' daily lives because users can access their homes and offices merely by swiping their hands against digital readers.

Corporate myth: 'Inevitable evolution' of technology

The bigger, longer term narrative is that of the 'inevitability of evolution' of technology, stated often, as a (false) analogy with biological evolution. The emergence of new technology happens because of the corporate drive for profits, without regard for consequence and rarely because of random autonomous experiments. Facial recognition or smart diapers are not inevitable but products of conscious corporate decisions, though, of course, all this new tech is presented in the garb of more convenience and inevitable 'human progress'. This inevitability argument allows corporations to evade responsibility for the consequences of deploying their products that are made available to consumers, criminals or to security agencies. This myth makes us comfortable in giving away so much information about ourselves because we see it as a 'natural' process.

We are being tricked

End-user license agreements (EULAs), along with privacy and security policy statements, trick us into agreeing to stuff that we do not care to read, much less understand. Long winded, full of legalese and jargon, most EULAs try to extract consent to the maximum extent even as they cover their own liabilities. They have evolved into longer and longer documents ensuring that fewer and fewer users will read.

For instance, did you know that Uber used to track a rider's location even after the ride ended in a bid to figure out what people did after finishing their rides? Until Apple outed them for doing even more sinister stuff. People have literally sold even their souls probably because they did not read the EULAs!

Even children are not spared

Google's 'G Suite for Education' (Docs, Gmail, Meet) and Chromebooks dominate the school sector in the United States. The state of New Mexico recently sued Google and said that it had "collected a trove of students' personal information, including data on their physical locations, websites they visited, YouTube videos they watched and their voice recordings." Needless to say, that this data extraction was done without the consent of parents.

Earlier, in September 2019, Google was fined in a settlement with the Federal Trade and Commission and the New York attorney general, because Google-owned YouTube had extracted personal data relating to children, and had even targeted them with ads! This was a violation of the Children's Online Privacy Protection Act (CoPPA).

The problem is now suddenly aggravated, because of the ongoing pandemic, and children have to attend online classes. Parents are grudgingly accepting whatever terms and conditions software companies are imposing for the use of their education software, aware that their children's data is being harvested.

Location data to track & identify: Breaking anonymity

Location data is the one trail that is simply generated by a mobile device. Such data is typically available with mobile telephony service providers, with companies that manage devices (e.g. Google, Apple) and with apps that use location data (e.g. weather, health apps). Then there are rogue apps that will acquire this information on the sly and a device user will not even know. The web browser is an important window to sites that often demand location because they may want to display 'location relevant' information or an ad placement service wants to serve ads of local shops and services. Twitter has a geo-tagging feature available for tweets but this has recently been made an opt-in feature. Facebook, of course, does persistent and aggressive location tracking.

The location data market has big players who are networked with numerous smaller players, and even a secondary market.

> The greatest threat here [collection of location data] may be to physical safety if such data can be used to identify who we are.

Location data can reveal the frequency of our visits to, say, hospitals, clinics, and restaurants or wherever; hunting for simple patterns in such data will also reveal where we reside and where we work. In 2018, the US Supreme Court noted that such data provides "an intimate window into a person's life, revealing not only his particular movements, but through them his 'familial, political, professional, religious, and sexual associations'." The greatest threat here may be to physical safety if such data can be used to identify who we are. The effect on the freedom of expression will be chilling. In countries with authoritarian regimes, which are ubiquitous around the world these days, it could get you killed or 'disappeared'.

And lest you think it cannot happen, here is a genuinely frightening investigation carried out by the *New York Times* (NYT) on the tracking of individuals by just using the location trails of their mobile phones. The data that NYT accessed,

> ... originated from a location data company, one of dozens quietly collecting precise movements using software slipped onto mobile phone apps. ... to anyone who has access to this data, your life is an open book. They can see the places you go every moment of the day, whom you meet with or spend the night with, where you pray, whether you visit a methadone clinic, a psychiatrist's office or a massage parlour.

This supplied data was in anonymised form i.e. no personal identifier information was included. Yet, with little effort the NYT team was able to track individuals and identify many from a women's march protest event attended by nearly half a million. The location trails could be combined with publicly available information, like a home address, to identify people.

> Many notorious companies, known for their intrusive technologies, have jumped on to the Covid tracking bandwagon to provide surveillance tools.

This example also demonstrates that the idea of anonymity is mostly a myth: it may not take much effort to de-anonymise anonymised data. We consent to give away quite a lot when we sign a very common consent clause that says, "We may share aggregated, non-personally identifiable information publicly and with our partners – like publishers, advertisers or connected sites." While this clause may sound comforting, the example above shows that there is always a good likelihood of discovering personal identifiers in anonymised data. In a more general sense, this tells us that a significant consequence of combining separate pieces of data, each of which may have been obtained by due consent, may yield information the disclosure of which was not consented to.

But a pandemic makes tracking 'essential'

The pandemic sweeping the world has brought concerns related to location tracking technologies into sharp focus. The highly contagious nature of the disease necessitates that people who have caught the disease, or even those who are currently asymptomatic (but may become ill later), need to be in quarantine. One way of keeping track of the relevant people is to track their location by monitoring the movement trails of their mobile phones. Many nations, city and state administrations as well as central agencies have released apps that do this tracking with user 'awareness' (a user has to install an app on her smartphone and give permissions). The Aarogya Setu app, sponsored by the Indian government, focuses on acquiring proximity data, i.e., who were the people a person was in proximity to across days. This is to aid 'contact tracing' to establish a path of transmission of the novel coronavirus. The app has been criticised for extracting location data and for having ambiguous clauses in its privacy policy.

> There is also the ever-present danger that anonymised data can be easily de-anonymised, by linking it to databases that hold the mobile number…

An MIT-based project is critically documenting Covid tracking apps across the world. Sure enough, a variety of problems plague many of these well-intentioned apps: extraction of irrelevant personal details as seen by the plethora of permissions they demand to run; lack of clarity about the use and retention of the data as well as privacy rules that decide what can be made public and what cannot be. Open-ended privacy policy clauses such as "will be used for appropriate purposes" or "will be shared by relevant agencies" keeps the door open for the data to be shared with law enforcement or tax authorities, and thereby be used for coercive purposes it was never collected for.

These apps therefore create significant vulnerabilities for those who are being tracked, in terms of loss of privacy. For instance, an app used by the Karnataka government released names and addresses of patients along with location trails, resulting in harassment by neighbours. There have been many more instances. In an ethos where even the suspicion of being infected immediately invites social stigma and discrimination—rather than empathy—and sometimes physical violence at the hands of self-appointed vigilantes, release of personal details into the public sphere endangers lives.

There is also the ever-present danger that anonymised data can be easily de-anonymised, by linking it to databases that hold the mobile number (e.g. Aadhaar, Vahan, in India), or even as described above in the NYT investigation above. In some cases, the apps are hackable or amenable to be copycatted by fraudsters. Like any other kind of data, that collected by some Covid apps will invariably reach the data markets.

Many notorious companies, known for their intrusive technologies, have jumped on to the Covid tracking bandwagon to provide surveillance tools. The Israeli company NSO - of the Pegasus spyware fame - has offered its technology, named Fleming, for testing to many countries. Another is the Rome-based company Cy4Gate offering its product Human Interaction Tracking System (HITS) free to Italian authorities.

Data breaches are not uncommon

We often make the facile assumption that corporate or state entities always indulge in ethical practices. A scary report tells us about tracking of phones by bounty hunters using data apparently sold by mobile telephony companies themselves! A company as 'public' as Facebook has been involved in so many scandals featuring unethical and even illegal exchange of data with other companies, for instance, data sharing deals, and the Cambridge Analytica and the Brexit sagas. It would take a humongous leap of faith to believe such things happen only rarely.

A short list of the biggest data breaches ever shows that even security obsessed financial companies have been victims as also notables like Facebook and Yahoo. And it keeps on happening. As we keep on giving away more and more of our personal data into an ever-

increasing number of obscure app-related databases, our chances of becoming victims of privacy invasions keeps increasing disproportionately.

Small breaches: The IoT quagmire

When Alexa malfunctions and starts recording voices without you ever having said the 'wake' word, it is capturing data without your consent. The Amazon Ring camera that is available in the US will capture footage of your front door as well if it happens to be in the field of view of the camera, even though you have never consented to such surveillance; there is nothing that you can do about it. It is not surprising that such devices have not only normalised the idea of everyday surveillance but many users confess that they have become voyeurs themselves.

Internet-of-Things (IoT) promises to bring more of this into your living spaces. Soon your refrigerator will send data to the supermarket about your food stocks and consumption patterns. Your electricity company will know how much time you spend in which room by analysing energy consumption patterns inside your home, based on the data sent out by domestic gadgets like air conditioners and smart lights. What looks like dream may turn into a privacy and security nightmare.

Morphed images & deepfake videos

Newer technologies have made some problems more sinister. Things that were difficult to do have become feasible in a rapidly growing ecosystem of cheap hardware (fast processors and more memory), cloud services (storage and processing), easily available AI algorithms, rapid search engines, and apps that provide access to this prowess.

> [P]olitical deepfakes have already arrived, and a Delhi election speech may have been a pioneering one. Should you believe what you hear or see?

In 2019, the DeepNude app used artificial intelligence to 'undress' clothed images of women. It was withdrawn amidst a huge backlash. The technology has now reappeared in even simpler form where a bot—automated, artificial intelligence, in this case—produces nudes, from clothed images, in a matter of minutes. Deepfake technology can replace the face of a person in a video by someone else's. The day is already here when a mobile phone app can do this, and easy tutorials are available. And so, the problem which arises is this: so many of us have so many of our photos and videos all over the internet, much of it voluntary. It did not matter much previously. But now these are input data for producing images and videos, featuring us, that can be terrifying. Your next great worry may just be finding yourself starring in a publicly available pornographic video. Facebook, Twitter and YouTube have just banned (some) deepfake videos from their platforms, and the detection is likely to be automated through AI. But what will happen when the fake is perfect and 'undetectable'? Meanwhile, political deepfakes have already arrived, and a Delhi election speech may have been a pioneering one.

Should you believe what you hear or see?

Breaking the digital cage

Convenience and pleasure are hard to fight against, as is the pursuit of vanity. It is a complex problem and not just a conflict between rational thoughts and irrational desires. Some have spoken about it as the 'Tyranny of Convenience'. Others have modelled social media obsessions as the 'Happiness Industry'.

As we immerse ourselves deeper into more and more data creating apps the dangers increase disproportionately. We spend more time on an app and it extracts even more data from us. The most proximate effects of using these apps is addiction: whether it is scrolling mindlessly through never ending feeds or an obsession with metrics that go much beyond the vain counting and giving of 'likes'. Wearables, for instance, create an obsessive pursuit of numerical goals.

A succinct analytical essay sums it up nicely, "wearables could fuel an unhealthy obsession with personal wellness." It adds, "As a result, we pour more attention into monitoring and controlling ourselves, giving us less time to do the things which actually make us happy." The real trade-off, which we usually fail to notice, is between convenience and happiness.

> Secure your privacy controls on all accounts and firewall your browser so that it does not track, leaves no fingerprints, and uses cookies only when essential.

Many of us may be aware of how to get out of this digital cage, but it seems next to impossible to actually do so. A drastic way is to get rid of the smartphone in your pocket and replace it with a feature phone. Another important recommendation is to avoid the Faustian bargain of using 'free services' in exchange for personal data: read your news from a regular newspaper that has an editor—and is likely to be behind a paywall—and not from a feed curated by an algorithm that will show only what you would like to see; pay to use software or services so that the service provider extracts no data; use paid messaging apps (remember SMS) where people are less likely to send toxic forwards and fake news. Try using only free software developed by programmer communities (e.g. the Linux operating system) that demand no data for their use.

Get rid of social media and make real friends. Control how much Google and Facebook can know about you. Secure your privacy controls on all accounts and firewall your browser so that it does not track, leaves no fingerprints, and uses cookies only when essential.

Perhaps we should also ask governments to guarantee basic email, data and cloud services to every citizen, as a legal obligation, with minimal data extraction (much like paid-for public utilities such as, electricity, water). Can we make this happen?

This is the first in a two-part series. Read the second part here.