# Data matters: once more, unto the breach

Pranesh Prakash

2023-06-16

Hindustan Times

**Abstract**

The CoWIN data leak was not a freak, one-time breach. At the very least, we need to have the option to opt out of large databases that have collected our data.

---

Earlier this week, information about potentially 1.1 billion Indians — which would make it one of the largest data leaks ever — was made available via a bot on the chat app Telegram. The bot could be queried using a phone or an identity document number (like Aadhaar, PAN, etc.), and it would return the name, sex, date of birth, phone number, ID number and Covid-19 vaccination centre information registered on CoWIN.

Government sought to allay fears by saying the CoWIN app or database were secure, and not directly breached.

Clearly, the data was related to the CoWIN platform, which was used by the government to enable the massive Covid-19 vaccine roll-out during the pandemic. The leaked data contained information that may not have directly been stored on the platform, such as the date of birth.

The minister of state for Electronics and Information Technology Rajeev Chandrasekhar responded by saying that the data accessed by the bot was from a database with "previously stolen data".

He sought to allay fears by stating, "It does not appear that Cowin app or database has been directly breached," and later insisted that the "data to a large extent is fake" — contrary to reports of accurate data (and only accurate data, as far as one can ascertain) retrieved by multiple news outlets — without saying how that determination was made. The Union health ministry denied the allegations of a data leak, stating they were "without any basis and mischievous in nature."

Surely it is clear that the CoWIN website could be "secure" as could the computers hosting the CoWIN database, without the CoWIN data itself being secure. While ministries have said they are investigating the incident and an official explanation may or may not be forthcoming, news reports have suggested the compromise happened via a health ministry system that allowed data access to auxiliary nurse midwives or vaccinators. If the information that originated

from CoWIN is available via another system — whether run by the health ministry or not — that amounts to a leak of CoWIN data. In this case, the linkage of people's phone numbers to all the vaccine beneficiaries registered under that number — a single phone number could be used to register vaccination for six people — and vaccination site is information that originated with CoWIN (though this information was often sent over unsecured channels like SMS). Regardless of how the data leak happened, the impact on privacy and the data risks created due to the leak remains the same and risk mitigation will need to be undertaken by the government.

Some might legitimately ask: "So what if my birthdate, PAN, or Aadhaar number, etc. become public knowledge? After all, many celebrities' birth dates are publicly known, and thanks to social networks like Facebook, hundreds of people already know my birthday. I give my Aadhaar number everywhere."

The reality is that nothing is inherently private, it all depends on social context and norms. For instance, many financial institutions assume that information like 'date of birth' is private and thus, use it as passwords for sensitive encrypted PDFs that they send you over an e-mail; they use birth dates along with phone numbers or PAN or other ID numbers to authenticate their customers over a phone call with a customer service agent to minimise identity fraud. They could well have chosen any other information known only to you, but the fact remains that by choosing birth dates or Aadhaar numbers as information that helps them identify you before an interaction, it takes on the character of 'private' information.

Regulatory bodies like the Reserve Bank of India, Securities and Exchange Board of India (SEBI), and CERT-In (Indian Computer Emergency Response Team) should put out advisories cautioning against such assumptions. Or indeed, you might mistakenly assume that your (or your family member's) date of birth isn't public knowledge and use it as a PIN for UPI or your ATM or credit card or your phone. You should not. Or, you might forget that your information is publicly available, giving an opportunity to fraudsters to pretend to be from your bank or your insurance company. The more that potential attackers have access to your information, the more the risk of harm, including financial harm. Such misuse of personal data is a practical reality and forms part of many cases of financial fraud these days.

Though we do not have a strong general-basis data protection law in place, the Aadhaar (Sharing of Information) Regulations, 2016 require that no entity "share the Aadhaar number with any person without the consent of the Aadhaar number holder", which seems to have been violated in this instance, though it is unclear by whom exactly. If one believes an interview with the alleged hacker, a health worker had the ability to query a database for data related to 1.1 billion people, without their consent. Does any health ministry-run database allow this?

The Aadhaar regulations also require that no entity (including the government) store Aadhaar numbers for longer than is necessary for the purpose specified to the Aadhaar number holder at the time of obtaining consent. It is incumbent upon the government to clarify if it followed the law. If the Aadhaar numbers were collected for vaccination, once a person is fully vaccinated,

shouldn't the Aadhaar numbers be removed from the CoWIN database?

The issue of health data security is becoming ever more critical as health records are being digitised and interlinked via the Ayushman Bharat Digital Mission. This is not the first time such a leak has happened – there have been several instances of leaks from diagnostics labs and hospitals, such as in December 2022 from the database of a private hospital in Tamil Nadu, or in December 2016 from a Mumbai-based laboratory's website.

## How can we avert this?

Some would put forward decentralisation of data collection and storage as the answer to prevent further data leakage. Simply put, rather than a database holding every person's data, an encrypted smart card with the data should be in possession of that person. Such a large-scale data breach would then not be possible. While that addresses security concerns, it has its trade-offs. What happens if the person loses their card or it gets burnt in a fire? Would they then simply lose access to their health records? Could their paper records (if they've kept them) be re-digitized? Given these difficulties, we shouldn't force individuals to be responsible for the security of their data, but they should have the choice.

While we cannot prevent all data leaks — whether because of the impossibility of perfectly securing data or because of the impossibility of effective enforcement of data protection laws — we should aim to build social and financial systems that are more resilient to such leaks.

One small step towards this could be for the government to mandate the usage of the 16-digit revocable Aadhaar "virtual ID" — which is currently opt-in — instead of the 12-digit Aadhaar number, large amassments of which one should assume are lying unsecured in lakhs of databases. (One incidental benefit of this would be that people would be prevented from re-using infamous fake Aadhaar IDs as many were reported to have done on CoWIN.)

Regardless of one's approach, at a minimum, we urgently need a data protection law that requires notification to citizens upon a data breach, and allows people to hold entities holding and processing data to account if they fail to secure it. It should also instate common sense rules: minimising how much personal data is collected, restricting access to it on an as-per-need basis, etc. Lastly, we need openness and transparency from the government. They should make public the CERT-In report into this episode, and take action against those who violated the Aadhaar Act or the IT Act. We cannot speak of effective governance when the government reflexively obfuscates and denies as lies what we all can see to be the truth.

*Pranesh Prakash is the head of Anekaanta Advisory, a technology, law and policy consultancy based in Chennai, and was a co-founder of the Centre for Internet and Society and a fellow at the Yale Law School's Information Society Project.*