Social Graph Portability:
AN ANALYSIS OF INDIA’S APPROACH

Aparajita Srivastava and Nehaa Chaudhari
October 2019
Aparajita Srivastava is Senior Associate, and Nehaa Chaudhari is Public Policy Lead at Ikigai Law. Ikigai Law is a technology focused law and policy firm in India. The views in this paper are those of the authors. The authors thank Arpit Gupta, Ratul Roshan, Vihang Jumle and Vijayant Singh, Associates at Ikigai Law, for their research and editorial assistance.
INTRODUCTION

The story of social graph portability in India is fundamentally one of data portability, and the story of data portability in turn is one where individual privacy meets the state’s desire to regulate large (and mostly foreign) technology companies, and promote India’s digital economy. Data portability is a recurring theme across multiple law and policy instruments in India over the past couple of years. Key among these instruments are the draft Personal Data Protection Bill, 2018 (“PDP Bill”) and the draft National e-Commerce Policy (“e-Commerce Policy”), for their impact as much as for their approach to data portability.

The draft PDP Bill grants individuals the right to data portability (“RDP”). This allows users to (a) receive their data from data controllers in a structured, commonly used and machine readable format; and (b) have this data transferred to any other data controller in such a format. An individual may exercise her RDP over three kinds of data: (a) which she has provided to the data controller; (b) which has been generated in the course of the data controller providing goods or services; and (c) which is a part of any profile of hers, or which the data controller has otherwise obtained.

Simultaneously, the draft PDP Bill limits an individual’s RDP on legal and technical considerations. It does so in three key ways. First, the exercise of this right is limited to processing undertaken by automated means. Second, an individual cannot exercise her RDP if it is technically unfeasible for the data controller to enable it. Third, an individual cannot exercise her RDP if doing so would reveal a trade secret of any data controller.

The Committee of Experts chaired by Justice B.N. Srikrishna that was tasked with framing India’s data protection law (“Srikrishna Committee”) identified four points of importance of the RDP. One, the RDP empowers individuals by “giving them greater control over their personal data.” It facilitates individuals’ “ability to move, copy or transmit personal data easily from one IT environment to another” and not be locked in to a particular service. Two, the RDP entrusts the individual with the power to hold a data controller accountable. Three, the RDP could improve consumer welfare as a result of increased competition between companies engaged in the same industry. Four, the RDP is “critical in making the digital economy seamless.”

While the Srikrishna Committee lists a variety of reasons for promoting the RDP, the draft PDP Bill situates such reasoning primarily in the concept of individual autonomy. In fact, even the Srikrishna Committee calls the RDP (together with the rights to object to processing, and not be solely subject to automated decisions) an “extension of the core principle of autonomy.” It identifies putting an individual in control over her data to be “at the core of [India’s] digital philosophy.” This view is grounded in the Indian Supreme Court’s understanding of privacy, especially in the Puttaswamy case. In Puttaswamy, privacy (including informational privacy) was held to be a fundamental right (and read to be an intrinsic part of other fundamental rights and freedoms, including those of liberty and equality) under the Indian Constitution.

In contrast to the draft PDP Bill, the draft e-Commerce Policy does not explicitly mention a RDP. It only articulates a wider vision for data governance, grounded in the economics and politics of data. It calls data the “most critical factor” for a business’ success and a “valuable resource” for companies, individuals, and governments. It opines that Indian businesses are disadvantaged by a lack of access to data, and are thus unable to innovate, despite a country’s data being its “national asset.” The draft e-Commerce Policy identifies the dominance of a few (early mover) companies given their scale and “network effects” to be the fundamental cause of this disadvantage. It aims to fix this by (a) empowering
Indians to control the “data they generate and own”;\(^{31}\) and (b) sharing data with Indian startups and entrepreneurs,\(^{32}\) which it believes will result in innovation.\(^{33}\)

Despite not being explicitly mentioned, a principled justification for data portability may be traced in the draft e-Commerce Policy’s engagement with data on economic and political levels, and its identification of an individual as the owner of her information.\(^{34}\) If an individual owns all the information that she generates, it stands to reason that she can port it across entities. Similarly, data portability is a mechanism for data sharing with Indian startups, which could counter the “network effects” of the early-mover (foreign) companies.\(^{35}\)

Significant data policy developments in India, and the Indian Constitution itself, thus offer a foundation – whether rights-based, economic, or political – for data portability, and by extension, social graph portability. However, this foundation is shaky. The RDP in the draft PDP Bill has largely been imported from the European Union’s General Data Protection Regulation, 2016 (“GDPR”),\(^{36}\) without a rigorous examination of its utility and applicability in India. Similarly, the e-Commerce Policy’s recommendations on data sharing also warrant close scrutiny, and not just because they utilize a data ownership paradigm, which has been rejected by the Indian Supreme Court\(^{37}\) and countries around the world.

This paper deliberates on the net benefit of social graph portability to users and platforms.
KEY ISSUES

1. Privity of contract

The draft PDP Bill, like most data protection laws around the world, envisages consent to be the primary basis for all data processing activities. This means that an individual (the data subject) must have consented to the manner and means by which a data collector may use her data. Typically, such consent is accorded when an individual consents to a company’s terms of use. This creates a direct contractual relationship between the data collector (the company) and the data subject (the individual). The company’s use of the individual’s data is governed by the terms of this contract, which restrain the data collector’s use of the individual’s data.

As a general rule, the doctrine of ‘privity of contract’ governs all contractual relationships. This means that a contract cannot confer or enforce rights against any person who is not a party to the contract. Therefore, a contract between a data collector and a data subject can govern the former’s use of only this specific individual’s data, and no other’s. Notably, social graph portability challenges this core principle of contract law.

When a user (“Primary User”) of a social networking platform (“Original Platform”) ports her entire historical social graph i.e. profile information, photographs, status updates and other activity from this platform to another platform (“New Platform”), she may also be porting data pertaining to other users (“Secondary Users”) of the Original Platform. For example, porting a Facebook Primary User’s entire social graph onto a new social media platform would mean porting data that may belong to her friends (who are Secondary Users), including their comments on status updates, posts, photographs that they are also tagged in, and locations visited together, among other things.

When the Primary User ports her data to the New Platform, she agrees to its terms of use, and thus has ‘privity of contract’ with the New Platform. On the other hand, Secondary Users have not consented to their data being moved to another platform, and may not even be aware that such porting has occurred. In effect, the Primary User has bound Secondary Users to the New Platform’s terms of use, which fundamentally cuts against the principle of ‘privity of contract’. India’s telecom regulator has also recognized this to be a serious challenge to data portability, and has observed that such Secondary Users have no control over how and where their data is being shared and used by the New Platform.

2. Porting user loaded data v. User Generated Data

User data on social networking platforms may be broadly categorized into (a) user loaded data; and (b) User Generated Data (defined below). User loaded data is that which a user brings to the platform from other sources, such as photographs uploaded to Instagram. A user’s access to this data is usually independent of its relationship with the platform – for example, the Instagram user may have a local copy of the photographs uploaded on the platform, or may have saved them on the cloud. On the other hand, User Generated Data, as the term implies, is created on the platform, and its access is contingent on the user’s active relationship with the platform.

User Generated Data may be further categorized into (a) data that a user generates on the platform (for example, the filters and captions that an Instagram user may add to her photographs); and (b) data that is generated by other users
of the platform while interacting with the Primary User or the content that she generates (for example, comments and reactions left on the Primary User’s Instagram photographs by her followers). Here, the data of Secondary Users which may be intrinsically linked to, or form a part of, the data generated by the Primary User on a platform (“Secondary User Data”) is bundled with the data generated by the Primary Users.

As seen earlier, data portability by the Primary User where Secondary User Data is involved exacerbates the concern with ‘privity of contract’, where only the Primary User has agreed to the terms of use of the New Platform, but Secondary User Data is also being used and processed by the New Platform. A possible solution to this could be that the Original Platform enables only the data generated by the Primary User to be ported without Secondary User Data. However, this may pose practical challenges, as the Primary User’s data may not make much sense without Secondary User Data.

Alternatively, only user loaded data may be made interoperable. Making such data interoperable may only mean making it easier for the user to extract it from one platform and populate it on another. Take the example of a user who wants to move her photographs from Facebook to Instagram. Allowing this user to port her photographs from Facebook to Instagram merely saves her the effort of first downloading them from Facebook and then uploading them to Instagram. She may not even need to download them from Facebook if she has a local copy.45

Undoubtedly, the manner in which different platforms may present this data will vary considerably, but this is not critical to the user’s ability to access the loaded data itself. The raw imported data is what the Primary User brought onto the platform and she represents to the data collector that she has the right to move it to a different platform. Even if this imported data set contains Secondary User Data, before importing this data onto the New Platform, the Primary User can declare or represent that she has the right to do so. Any dispute on this account will be dealt with between the Primary User and the disputing Secondary User, without any recourse to the platform – specifically since the platform does not have any role to play in the Primary User uploading the data on the platform.

3. RDP and ‘consent’ under the draft PDP Bill

Under the draft PDP Bill, ‘personal data’ may be processed on the basis of ‘consent of the data principal’.46 This consent is valid only if it is ‘capable of being withdrawn by the data principal’.47 As mentioned earlier, when the Primary User ports her personal data (which may be comingled with Secondary User Data) the Secondary User has not consented to use of her data by the New Platform. In order to overcome this, the Original Platform and the New Platform may require the Primary User to represent that she has taken the Secondary Users’ consent to port their data. The platforms may therefore shift the liability to solicit the Secondary Users consent onto the Primary User.

However, this manner of taking the Secondary Users’ consent may be flawed, given that the Secondary Users’ consent is taken through the Primary User and is not capable of being withdrawn by the Secondary User, since she has no visibility of where her data has been ported or how it is being processed by the New Platform. Therefore, the Secondary User is deprived of her right to ‘withdraw consent’ from the New Platform. Even if the Original or New Platform assert that consent to port data has been granted by the Primary User on behalf of the Secondary User, such consent is still incapable of being withdrawn by the Secondary User, and therefore falls foul of a fundamental element of ‘valid consent’ identified in the draft PDP Bill.
4. RDP under GDPR and the draft PDP Bill

Article 20(1) of the GDPR and section 26 of the draft PDP Bill provide users with the RDP. Under the draft PDP Bill the RDP extends to data which the user (i) provides a platform; (ii) forms part of her profile, or which the platform has otherwise obtained; and (iii) generates on the platform (“User Generated Data”).

The draft PDP Bill does not clarify whether the RDP extends to Secondary User Data or if it is restricted to data provided or generated by the Primary User on the platform. Porting such data would mean that data of multiple users (including those who have not made the porting request) is moved to the New Platform.

The RDP under GDPR, on the other hand, extends only to data provided by a user to a platform. Therefore, only data generated by the Primary User can be ported to another platform. The data generated by the Primary User must be distinguished and extracted from the Secondary User Data. However, extracting the Primary User’s data may pose technical challenges and would also represent an incomplete version of the user’s historical social graph – which may defeat the purpose of the RDP.

The RDP is also restrained under Article 20(4) of the GDPR, which states that the RDP “shall not adversely affect the rights and freedoms of others”. This provision can be interpreted to further strengthen the argument that Secondary User Data cannot be ported without their consent since it may adversely affect their data privacy rights.

5. Security threat

RDP may also pose a significant cyber security risk, since a single breach – for example, at the point of porting data from one platform to another -- may compromise the entire historical social graph of a user.

6. Platform Generated Data

Data on users’ behavioural patterns observed by the platform (e.g., the number of hours a user spends on the platform, the advertisements a user has clicked on, etc.) (“Platform Generated Data”) is not provided by users and in our view must not be subject to the RDP.

An interpretation of the RDP provisions under the GDPR supports this view, since the RDP only extends to data provided by the user. In principle, the restriction on providing only that data which is “provided” by the users may safeguard the intellectual property of data controllers, and avoid a situation where the intellectual property of a digital service provider (data inferred about consumers using complex algorithms) could be lawfully disclosed to competitive businesses for free. Therefore, under GDPR the RDP does not include any additional data extrapolated by the data controller based on the data a Primary User provides. For example, if a data controller uses data to create a user profile, such a profile is not in scope of the RDP.

However, the RDP given in the draft PDP Bill extends to data which the platform “has otherwise obtained”. This provision leaves room for ambiguity on whether Platform Generated Data is also subject to the RDP.
In our view, the RDP should not be used to compel platforms to share Platform Generated Data or the manner in which they structure or process data, since this data is not provided or generated by users and can be claimed as the intellectual property of the platform.

7. Technological challenges

The RDP under GDPR and the draft PDP Bill gives users the right to have their data transmitted directly from one platform to another. However, Article 20(2) of the GDPR and section 26(2)(c) of the draft PDP Bill accords this right to users only if it is “technically feasible”. In other words, data controllers can prevent the full exercise of users’ RDP if they prove that in a given situation the level of technological development of their organisation makes a direct transmission of data to another data controller technically unfeasible because, for instance, interoperable formats (encouraged, but not imposed) have not yet been developed.55
ANALYSIS

The concept of social graph portability stems from two fundamental principles: users must (a) have control over their personal data; and correspondingly (b) be permitted to port that data interoperably between platforms. Accordingly, users should not be locked into ‘walled gardens’ created by social networking platforms. Data portability intends to neutralize the harms of 'network effects' in which users stay within the walled gardens of a platform not because of its superior user experience, but because they are locked in. The RDP would reduce the friction of switching from one platform to another. For instance, “the increased threat of people ditching Facebook for competitors would create a much stronger incentive to protect users and society”.56

The longer a user stays on a social networking platform, the deeper her social interactions and historical social trail. As a result, over time, it becomes increasingly difficult for engaged users to switch platforms. For example, if a platform abuses a user’s privacy, she is faced with an all or nothing proposition: she either stays on the platform with the compromised privacy safeguards or terminates her account, thereby losing her entire historical social graph built over time. Social graph portability may be seen as the panacea to platform dominance.

The RDP is often seen as a simplistic binary solution to reduce lock-in effect, enable consumer choice and promote competition among platforms. The idea is that the RDP allows market forces and competition to check pernicious business practices by dominant platforms; instead of relying solely on regulation to do so. This is particularly critical since technology significantly outpaces regulation. Therefore, data portability attempts to strike at the jugular of platform market dominance.

With low switching friction, users can move their historical social profile to another platform with better privacy controls. This seems like an obvious way to empower users and give them full control over their historical social data. However, the inherent challenge with the RDP lies in its implementation.

Porting Secondary User Data may not adequately safeguard the Secondary Users’ right to data privacy; in particular, their right to effectively grant and withdraw consent. It may also exacerbate data security by creating a single window for cyber-attacks. On the other hand, it may be technically unfeasible to restrict data portability to the data provided by Primary Users. Moreover, allowing only the data of Primary Users to be ported may be an incomplete solution, as it fragments historical social graphs and renders them only partly interoperable.

Additionally, the RDP may result in the creation of multiple new and smaller social networks. Such social networks may have lower value for users and businesses compared to a ubiquitous social network such as Facebook, because the latter allows users and businesses to benefit from network effects such as zero costs passed on to users, a large community of interconnected users on a single platform, and the ability to detect patterns in user behavior and data, among others. In case of large social networks the whole is greater than the sum of its parts.57 However, the RDP may result in breaking social networks into smaller, less valuable parts.
CONCLUSION

In our view, the RDP may be a credible solution to the lock-in effect of ubiquitous social networking platforms. However, our analysis shows that it may pose several practical and technical challenges which may outweigh its benefits.
ENDNOTES


7 At the time of writing this paper, both the PDP Bill and the E-Commerce Policy are yet to take final shape. According to news reports, the federal government has told the Parliament that it is formulating a data protection law. See ANI, Comprehensive legislation on data privacy under formulation: Ravi Shankar Prasad, The Economic Times, dated 03 July 2019, available at https://economictimes.indiatimes.com/news/economy/policy/comprehensive-legislation-on-data-privacy-under-formulation-ravi-shankar-prasad/articleshow/70057338.cms. Meanwhile, the commerce and industry minister said in June, 2019 that India’s ecommerce policy will be finalized within twelve months. See PTI, Govt to come out with national ecommerce policy within 12 months, Livemint, dated 25 June 2019, available at https://www.livemint.com/news/india/govt-to-come-out-with-national-ecommerce-policy-within-12-months-1561468426495.html.


38 Section 12, ministry of electronics and information technology, the Personal Data Protection Bill, 2018 available at https://meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill%2C2018%2C0.pdf.

39 Berkson v. Gogo LLC, 97 F. Supp. 3d 359 (E.D.N.Y. 2015), available at https://casetext.com/case/berkson-v-gogo-llc | for a detailed discussion on when electronic contracts of adhesion/standard form contracts, such as ‘terms of use’, are considered to be valid.

40 For instance, Clause 1 of Uber’s terms and conditions, titled ‘Contractual Relationship’, contains a clause incorporating its privacy policy, which governs its collection and use of personal information; see ‘Terms and Conditions, Uber B.V.’, available at https://www.uber.com/legal/terms/in/. Similarly, WhatsApp’s terms of service also incorporate a reference to its privacy policy, stating that users agree to its data practices, including the use of their information as described in WhatsApp’s privacy policy. See WhatsApp Terms of Service, available at https://www.whatsapp.com/legal/#terms-of-service.


42 Dunlop Pneumatic Tyre v Selfridge & Co. [1915] AC 1847.


46 Section 12, ministry of electronics and information technology, the Personal Data Protection Bill, 2018 available at https://meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill%2C2018%2C0.pdf.

47 Section 12(2)(e), ministry of electronics and information technology, the Personal Data Protection Bill, 2018 available at https://meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill%2C2018%2C0.pdf.

Analysing the Net Benefit of Social Graph Portability

Section 26, ministry of electronics and information technology, the Personal Data Protection Bill, 2018 available at https://meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill%2C2018%2C2018_0.pdf.


Section 26(1)(a)(iii), ministry of electronics and information technology, the Personal Data Protection Bill, 2018 available at https://meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill%2C2018%2C2018_0.pdf.


J Constine, Friend portability is the must-have Facebook regulation, TechCrunch, available at https://techcrunch.com/2019/05/12/friends-wherever/.