Conceptualizing beneficial interests in the political economy of data: A theoretical inquiry

Submission to GIGA Net 2020

Author's Note

The conference paper being submitted is not a full-fledged academic paper yet—very much a work-in-progress. This paper is a blueprint for a proposal to rethink the beneficial interests of different stakeholders in data, and how their corresponding conflicts may be resolved. We refer to several legal and regulatory theories, some of them well established, while others still in a state of nascent evolution. When we refer to a the principle of legitimate interests, or competition law, or anti-discrimination law, we are cognizant of specific limitations different versions of these legal and regulatory theories may face in a given context. However, we draw from first principles in these evolving legal theories to propose a coherent basis for recognising beneficial interests in data. Based on the feedback, there are several ways this proposal could evolve before publication, which we hope to do after obtaining feedback from the reviewers and from discussants and co-panelists at the conference.

Introduction

Perhaps there is no clearer indication of the primacy of data in this age than the overworked metaphors that are often used to describe it. In the last few years, data has been likened, aside from the hackneyed comparison to 'oil', to any manners of tangible entities such as mineral deposits,[1] dividend deposits,[2] currency,[3] and even the Alaskan Permanent Fund.[4] On the other end of the spectrum, commentators has also likened data to radioactive materials such as uranium[5] and pollutants such as carbon dioxide.[6] As tired or inventive these metaphors may be, they signify a desperate need for a clear conceptual model through which we can think through the legal, social and economic ramifications of data. This conceptual clarity is not just a theoretical pursuit but is necessary to identify various rights and interests that multiple stakeholders have in data across various contexts.

While most metaphors, including those mentioned in the paragraph above are inaccurate, there are some analogies that merit theoretical discussion. Jennifer Shkabatur has come up with the most well-articulated theoretical framing of data, in a paper that was originally presented at the Giga Net conference in 2018.[7] Through the conceptualization of a 'global data commons,' she attempts to frame data in a manner that a wide range of stakeholders can access and therefore derive benefits from user-generated data-benefits that are now limited to the tech behemoths who collect, aggregate and process it. Shkabatur also offers a workable framing to ensure that tech companies comply with this conception-by invoking the 'public utilities' doctrine, and by offering fiscal incentives, and adopting "naming and shaming" initiatives. Shkabatur's paper also successfully engages with possible critique of this model-along the lines of privacy, competition, and user consent.

However, Shkabatur's framing of a 'global data commons' does not account for the variety of stakeholders that retain a beneficial interest in the data or cull out, even at the theoretical level, mechanisms for resolving tensions across these beneficial interests. Looking at data solely as a 'global commons' does not enable us to construct beneficial interests or mechanisms for resolving them, which means that in certain contexts, alternate conceptions of data may need to be considered to understand the beneficial interests involved. By relying on theoretical conceptions of data, this paper seeks to clarify the beneficial interests in data

---

[1] Hooper, J (2017)," Data mining: How digging through big data can turn up new mineral deposits," Cosmos, Aug 2,https://cosmosmagazine.com/geoscience/data-mining-how-digging-through-big-data-can-turn-up-new-mineral-deposits/

[2] Sumagaysay J(2019), " Could californians get paid for data they share," The Mercury News,February 15,https://www.mercurynews.com/2019/02/15/could-californians-get-paid-for-data-they-share-with-facebook-google-and-others/

[3] Barratt J(2019)" Data as currency; What value are you getting,"Aug 27,Wharton podcast,https://knowledge.wharton.upenn.edu/article/barrett-data-as-currency/

[4] Hughes C (2018)," The wealth of our collective data should belong to all of us," The Guardian,Apr 27,https://www.theguardian.com/commentisfree/2018/apr/27/chris-hughes-facebook-google-data-tax-regulation.

[5] https://twitter.com/FiloSottile/status/1162404848073170944.

[6] Tisne M (2019), " Data isnt the new oil, its the new CO2," Jul 24,Medium,https://luminategroup.com/posts/blog/data-isnt-the-new-oil-its-the-new-co2.

[7] Shkabatur J(2019), "The Global Commons of Data," 22 Stan Tech L. Rev 354

which accrue to the following stakeholders-states, private corporations (data processors,) individuals, and communities.

Shkabatur also stops short of acknowledging the political power asymmetries that exist across the globe and should play a role in the framing of any global commons. The big technology companies processing data are largely located in the US, while the large numbers of people coming online in emerging economies indicates that vast swathes of data that get pooled into the commons are coming from citizens of the global south. Any theoretical framing of beneficial interests must account for these power asymmetries-something our paper seeks to do.

Our paper stops short of articulating practical policies or legal frameworks that can implement our theoretical framing of beneficial interests. We define beneficial interests as " any benefit (economic or otherwise) that an individual or entity should receive through its relationship with another individual or entity." The use of this term should not be confused with the traditional use of the term in contract law or trusts law which usually refers to a material benefit.

The manner of adoption and the scheme of implementation will vary across jurisdictions and socio-economic contexts-leaving such elaboration as a topic for future endeavours. **This paper is a blueprint for a proposal to rethink the beneficial interests of different stakeholders in data, and how their corresponding conflicts may be resolved. We refer to several legal and regulatory theories, some of them well established, while others still in a state of nascent evolution. When we refer to a the principle of legitimate interests, or competition law, or anti-discrimination law, we are cognizant of specific limitations different versions of these legal and regulatory theories may face in a given context. However, we draw from first principles in these evolving legal theories to propose a coherent basis for recognising beneficial interests in data.**

The paper is divided into two sections. The first part surveys existing conceptions of data. This survey is not exhaustive but studies the conceptions most suited towards understanding the beneficial interests involved across contexts.The second section looks at the beneficial interests available to each stakeholder using the appropriate conception of data explored in Part 1.

Part 1: Conceptions of data

Data as property.

The concept of 'data ownership' seems to have quite a lot of intuitive power. The economic theory of endowment effect describes that the owner of an object, (in this case our personal data), assigns it greater value than the possessor. The basis for this theory is an evolutionary response to surviving in competitive environments "in order to provide a strategic advantage in confrontation with others seeking to appropriate [the object in our possession]".[8] In the case

[8] Eswaran, Mukesh and Neary, Hugh M., An Economic Theory of the Evolutionary Origin of Property Rights. Available at https://www.isid.ac.in/~pu/conference/dec_11_conf/Papers/MukeshEswaran.pdf.

of data, the value attached to one's data is the benefits of privacy and financial gains. Unlike other commodities where ownership is often synonymous with control over the commodity due to its ability to be physically possessed or legal documentation enumerating its rightful owner, data cannot be possessed by just one person. The shared nature of creation of data by the data subject's interaction with an interface created by a data holder makes the answer to the question 'who is rightfully entitled to control over personal data' nuanced. The exploitative nature of mining personal data creates an imbalance in the benefits accrued by those whose data is utilized for financial gain and those monetizing on having access to personal data. There has been an upswell of discontent-particularly in the Global South with several commentators claiming that excessive focus on consent has skewed the discussion in favour of US based technology corporations who reap monetary dividends from data gathered from Global South citizens, thereby leading to accusations of 'data colonialism.'

The idea that individuals should receive fair compensation for the use of their personal data has received significant support in the last decade from a range of commentators.[9] Given that data about individuals have become a commercial asset for data processing organisations, it has been argued that data subjects must be given an instrument that would enable them to negotiate and bargain over the use of their data. It is also worth noting that despite academic discourse suggesting that legal frameworks do not favor propertisation of data, the business practices, particularly dealing with digitally available personal data suggest otherwise.[10] Data is very much a commodity, to be traded and valuations of early stage companies are often linked to the scale and nature of data they control.[11] The existing laws also permit corporations to contractually claim ownership over data, by virtue to their participation in creating them.

For something to be appropriately classified as property, they must satisfy the following conditions:

A. Possession and Enjoyment: enjoy your possessions in a way that you choose

B. Exclusive use: exclude others from their use if you wish

C. Transferability: dispose of them by gift or sale to someone else...who becomes their owner

Property is effectively an interest in an object, whether tangible or intangible, that is enforceable against the rest of the world.[12] Let us consider the nature of personal (and non-personal data which once was personal data) as property. As per the above definitions of property, property rights are different from rights arising from privity. This distinction is important. The treatment of data as commodity and right of actors to trade in it so far, seems to draw more from contract rights, where the data collectors, by virtue of, unnegotiated broadly drawn terms of use, appropriates rights over the data collected. This however does not

[9] Scott J (2018), "You should be paid for your Facebook data," QUartz,Apr 11,
[10] Prins, J.E.J. (Corien), Property and Privacy: European Perspectives and the Commodification of Our Identity. Information Law Series, Vol. 16, pp. 223-257, 2006 , Available at SSRN: https://ssrn.com/abstract=929668.
[11]PWC, "Putting a value on data," https://www.pwc.co.uk/data-analytics/documents/putting-value-on-data.pdf.
[12] Henry Hansmann & Reinier Kraakman, Property, Contract, and Verification: The Numerus Clausus Problem and the Divisibility of Rights, 31 J. LEGAL STUD. S373, S374 (2002).

suggest that the data collectors have an intrinsic right over the data they collect, as it is not drawn from an 'interest of beneficial ownership' in the data.

In reality, data, including personal data, has been treated like a commodity for some time now. Once personal data becomes a commodity, questions arise regarding the necessity, if any, of legal limits on data trade.[13] There is also some widespread recognition of the market failure inherent in the commodification of data. This market failure is marked by the systemic incentives towards trade in data at great negative externalities in the forms to privacy harms to the data principal.

Next, let us consider the nature of the right to privacy, which is the most obvious legal complication that we must necessarily contend with before embarking on any discussion about interests in data. Information privacy entails that the use, transfer, and processing of the personal data must only occur with the informed consent of the individual. Conceptually, one key thing to remember about any kind of property interest in data is that it necessarily means that privacy as a 'value' is owned, and like any other piece of property can be bartered. The clear implication of vesting property rights in personal data would be that privacy is an alienable right.

There have been several definitions of the right to privacy, but perhaps a useful one for the purposes of our discussion would be based on the idea of privacy as individual control. It is our right to control access to and uses of physical places or locations, as well as, personal data about us.[14] It is however, important to remember that when this right is exercised to relinquish control, for example, by way of sharing some information, that does not make this not lead to waiver, relinquishment or forfeiture of the right itself or future claims to control the same data.

To further clarify, alienation is different from both waiver and forfeiture. Waiving a right has immediate consequences for the specific instance of the exercise of the right, it does not, by itself, impact future exercise of the same right even over the same particulars. This is different from alienation. Forfeiture on the other hand often involves losing access to a right due to illegal acts on part of the right holder. This is also different from alienation. Alienation implies "transferring the moral authority to engage in the general practice the right protects."[15] Therefore, the implications of data as property and consequently, privacy as an alienable right would be dire. It has also been pointed out by Tisne and others that given the current state of the data economy, owning, renting and selling personal data would lead to extremely exploitative and iniquitous consequences.[16] This extends to paying individuals for their data.

---

[13] Margaret Radin on commodification: "capable of being reduced to money without changing in value, and completely interchangeable with every other commodity in terms of exchange value."

[14] Alan Westin, Privacy and Freedom (New York: Atheneum, 1968); Anita Allen, Unpopular Privacy: What Must We Hide?, (Oxford University Press, 2011); Beate Rossler, The Value of Privacy (Polity Press, 2005).

[15] Moore, Adam D., Privacy, Interests, and Inalienable Rights (January 22, 2018). Available at SSRN: https://ssrn.com/abstract=3107324 or http://dx.doi.org/10.2139/ssrn.3107324.

[16] Tisne M (2018)," It's time for a data bill of rights,," Dec 14, MIT Technology Review,https://www.technologyreview.com/2018/12/14/138615/its-time-for-a-bill-of-data-rights/.

In a scenario, where corporations could pay individuals for their data, it would mean that those in lower income groups would be more willing to trade away their rights than the well-off. [17] This is fundamentally incompatible with the notion of an inherently inalienable right. Further, when trading one's data, how is the value or the cost to the individual to be determined? A cost to the user and a definite benefit to the private platform is through the aggregation of data. Data is far more valuable when aggregated. It is thus impossible to accurately compute the precise value of an individual's data. To the contrary, it is possible that data provided by an individual can be aggregated and used to conduct predatory practices against the group the individual belongs to.

Opponents of dictum that privacy is an inherent inalienable right point to cultural relativism of privacy, and that its nature, facets and scope vary with cultural contexts.[18] Schwartz,[19] and separately Roberts and Gregor[20] effectively respond to and resolve this question. They acknowledge that privacy, by its very nature, allows for deviations in order to sustain social establishments and group values. While the exact manner in which privacy as a right may manifest itself may be culturally influenced, the very need for privacy is not. As mentioned above, the right to control access to and uses of physical places or locations, as well as, personal data about us is essential to human dignity.[21]

Global Data Commons

Shkabatur's Global Data Commons Model

While others have advocated for use of the phrase 'data commons,'[22] Shkabatur has come up with the most comprehensive theoretical justification for the same. [23]User generated data has tremendous value in modern society across sectors.[24] As of now, requirements on platform companies to share data with users are limited to personal data. For example the GDPR provides users with the rights to access data collected about them, and third party recipients

[17]Elvy S-A (2017)," Paying for privacy and the personal data economy,"117 Columbia Law Review 6,1370
[18] Nileena M (2020),"From Aadhaar to Aarogya Setu, Vidhi's questionable role in technology-related policy making", Aug 24,https://caravanmagazine.in/technology/vidhi-aadhaar-aarogya-setu-arghya-sengupta-privacy-think-tank.
[19] Barry Schwartz, "The Social Psychology of Privacy," American Journal of Sociology 73, no. 6 (May 1968): 741–52.
[20] John Roberts and Thomas Gregor, "Privacy: A Cultural View," in Privacy: Nomos XIII, ed. J. Roland Pennock and John W. Chapman (New York: Atherton, 1971), 225.
[21] Adam D. Moore, "Privacy: Its Meaning and Value," American Philosophical Quarterly 40 (2003): 215–27.
[22] Singh, P (2019), " Data and digital intelligence commons (Making a case for their community ownership," Data Governance Network, Working Paper 2,
[23] https://law.stanford.edu/wp-content/uploads/2019/09/Shkabatur_Global-Commons_20190830-1.pdf
[24] (" The real resource at the core of digital economy, and its new relationships, therefore is digital intelligence. This intelligence is built from data. Data is something inherent in the concerned social relationships, left as digital traces over platforms from where it is collected and processed by digital companies.")Ibid at 23

of the data, rectify inaccuracies and demand erasure. [25]The justification for re-sharing personal data stems from the ability of individuals, upon receiving their personal data to draw useful conclusions about themselves.[26]Shkabatur argues that the same rationale can be extended for creating a global data commons owing to its usefulness for "our collective pursuit of knowledge" and importance in public decision-making.[27]Movements around the world, such as the Open Government Partnership have already started championing the benefits of open access to government data.[28]India developed the Open Government Data Platform as an outcome of the Indo-US Open Government dialogue in 2010 and was customised by the National informatics Centre as per the National Data Sharing Accessibility Policy.[29] Several other emerging economies, Ghana [30]and Rwanda[31] are also adopting open data platforms.

Both Singh and Shkabatur argue that till now, the private sector has chosen to share data as part of "data philanthropy initiatives"—an approach which needs to change. As argued by Shkabatur, as part of a global data commons, " user data would be responsibly managed in a manner that contributes both to business models of platform companies and to larger societal objectives."[32] This commons based approach has thus far only been discussed for scientific collaborations across the globe. High level principles have been developed by expert bodies for the purpose of data sharing. While she lauds this development, Shkabatur considers the limiting of a global data commons only for scientific purposes to be "narrow and restricting."[33]

She goes on to argue that the commons approach does not entail 'open access.'Under the data commons regime private associations, firms, researchers, and individuals will all hold distinct access and usage rights over separate tracts over user generated data. She suggests five modalities of data sharing-each with an incremental extent of sharing by the private player.[34]

The first is sharing internal data analysis-data platforms analyze data they process "their own data" and share insights derived but do not share the data itself. Shkabatur does not clarify what a platform's "own data'' might be as the primary data processed by platforms and used to derive insights stem from user-generated data. In the paper, Shkabatur uses the example of Mastercard's derivation of insights from aggregated transaction data but aggregated transaction data is essentially data generated by users of Mastercard in anonymized format.

---

[25] Shkabatur 383
[26] Omer Tene & Jules Polonetsky, Big Data for All: Privacy and User Control in the Age of Analytics, 11 Nw J. Tech&Intellectuall Property Law 239, 240, 247-50 (2013)
[27] Jane Yakowitz, Tragedy of the Data Commons, 25 Harvard Journal of Law&Tech. 1, 8-10 (2011 cited in Shkabatur 382
[28] https://www.opengovpartnership.org/open-data/
[29] https://www.nic.in/projects/open-government-data-ogd-platform-india/
[30] https://data.gov.gh/
[31] https://rwanda.opendataforafrica.org/
[32] Shkabatur 383
[33] Ibid 384
[34] Ibid, 385-395

Despite using the phrase "their own data," Shakabtur stops short of clarifying the beneficial interest Mastercard and its users have in data that might qualify as Mastercard's 'own data.'

The second model is releasing targeted data to address a concrete social problem or mitigate an emergency. This can be done by partnering with trusted organisations or inviting qualified individuals and organisations to develop apps and innovative uses of this data. Facebooks's Disaster Maps, for example, provides partner organisations with aggregated and de-identified data during a natural crisis. [35]

The third model mentioned Shakabatur are data pools- "a horizontal partnership between two or more companies or organizations that agree to share and analyze each other's data, and help fill knowledge gaps while minimizing duplicative efforts."[36] Relying on the analogy of patent pools, Mayer-Shonberger and Cukier argue that new firms may pool data from a number of consumers and provide an easy way to license it.[37]The example cited here is a collaboration between Esri, a mapping company and Waze, a community-based traffic and transport app with municipal governments which can access real time traffic data. [38]While Shkabatur rightly argues that the three entities involved in the pooling-governments and two private actors benefit,the beneficial interests of the users who generated the traffic data are not accounted for.

The fourth modality-granting access to public actors envisages private companies sharing data with specific "trusted partners." She goes on to suggest that independent government agencies or national regulators are well positioned to fulfil this position, and the power to make decisions about opening up the dataset vests solely in this body.This framing is inappropriate as often users may not want the state to access their data, even in anonymized form. This is particularly important in countries where states demonstrate authoritarian tendencies and have used digital surveillance as a method of clamping down on dissent and public participation. Any data sharing needs to be operationalised with clear and unambiguous consent from the user, rather than assuming bona fide intent on part of the state.Shkabatur's final modality that envisages the greatest degree of data sharing is termed open access-where the company provides "free, public and uncertified access" to certain portions of user-generated data.

Shkabatur invokes the 'public utilities doctrine' to compel private platforms to share user-generated data.[39]Due to their de facto de functioning as a common law understanding of 'public utilities,' private platforms owe public obligations such as non-discrimination and

[35] https://dataforgood.fb.com/tools/disaster-maps/
[36] Ibid 391
[37]Schonberger,V and Cukier K *Big Data: the essential guide to work,life and learning in the age of insigh*t,(John Murray,2013)
[38] See https://www.esri.com/about/newsroom/announcements/esri-and-waze-deliver-near-real-time-data-for-smarter-cities/
[39]Rahman K (2018) "Infrastructural regulation and new utilities," Yale Journal on Regulation 35

equal access that were imposed on traditional public utilities. The courts developed the 'public utilities' so that industries that provide essential goods and services to the public offer this service "under rates and practices that [are] just,reasonable, and non-discriminatory."[40] The two requirements for an industry to qualify as a public utility are that they must be a "natural monopoly" and "affected with public interest."[41] While traditional examples include electricity,water or telecommunications industries, Shkabatur rightly argues that modern day large private platforms also meet both criteria. This would include the five Silicon Valley companies collectively known as Big Tech-Google, Apple, Facebook,Amazon and Microsoft.[42] First, due to the high sunk costs, barriers to entry and high levels of market concentration by big tech companies,[43] they functionally operate as natural monopolies.[44]The leading firms exhibit network effects-which increase in value as more users utilise them. This makes it incredibly difficult for new entrants to be competitive.[45] Private platforms also meet the social necessity criteria due to the socio-economic significance of the power they possess, the potential for data processed by them to be leveraged for reducing socio-economic disparities, and their function as gatekeepers in the marketplace of ideas.[46]

Shkabatur also attempts to pre-emptively engage with three concerns with the commons approach. The first and challenge addressed is that of potential privacy concerns with user-generated data by recommending de-identification through pseudonymization and user consent.However, she fails to engage with the vast literature that points out that pseudonymization or even full anonymization is insufficient to prevent privacy violations.[47] Using the Netflix prize dataset, Narayana and Shmnatikov demonstrate how even with imprecise background information about a particular subscriber , an adversary could identify individual records through cross-correlation with other databases. [48]This challenge has also

---

[40] Joseph D. Kearney & Thomas W. Merrill, The Great Transformation of Regulated Industries Law, 98COLUM.L.REV.1323, 1331 (1998)

[41] Munn v.Illinois, 94 U.S.113, 130 (1877

[42] Apart from Saudi Aramco, they are the most valuable publicly traded companies in the world.Burszytynsky J (2020), " Apple surpasses Saudi Aramco to becme world's most valuable company," CNBC.com, Jul 31https://www.cnbc.com/2020/07/31/apple-surpasses-saudi-aramco-to-become-worlds-most-valuable-company.html

[43] As of April 2020, the largest five companies-Apple,Amazon,Facebook,Alphabet,(Google),Microsoft (the dreaded 'GAFAM") constitute twenty per cent of the stock market

[44] Taplin J(2017), " Natural monopolies: Time to break up Google?" Economic Times,Apr 20,https://economictimes.indiatimes.com/blogs/et-commentary/natural-monopolies-time-to-break-up-google/

[45] Ghosh D (2019), " Don't break up Facebook-treat it like a utility," HBR,May 30,https://hbr.org/2019/05/dont-break-up-facebook-treat-it-like-a-utility.

[46] (" Data and digital intelligence resources are social resources, implicit in social relationships forming a community, and an abstraction of them. In wrong hands, they can also cause great harm to the concerned community. In any case, their management in an appropriate manner is necessary for efficient and sustainable running of the digital economy and its various sub-systems, and fair allocation of benefits to different actors." See generally Singh

[47] Narayanan, Arvind and Shmatikov,Vitaly. 2008"Robust De-anonymization of Large Sparse Datasets" https://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf. Accessed August 1 2020.

[48] Ibid

been recognised in the GDPR and applied when determining the classification of personal data.[49]The GDPR, under Recital 26, has undertaken a risk-based approach to determine whether data is personal -an approach that has also been adopted by the British Information Commissioner's Office (ICO.) If risk assessment indicates that identification is 'reasonably likely' to occur, then anonymised data must receive GDPR data protection fully.The Article 29 Working Party however advocates for a far higher threshold, arguing that anonymised personal data can only qualify as non-personal data when "irreversible identification" is present.[50]

Some of these concerns are arguably solved by ensuring that the user provide explicit and unambiguous consent before data is shared with third party recipients, and the use it is put to, even when it is anonymised.Therefore, the solution to the privacy challenge lies in obtaining user consent whenever it is transferred to a third party or to the commons.

Data as a common property resource

Data is often characterised as a resource-a natural extension of the idea of a global commons.The term 'common pool resource,' as studied and articulated by Ostrom refers to a natural or man-made resource system that is sufficiently large to make it costly (but not impossible) to exclude potential beneficiaries from obtaining benefits from its use."[51]Indeed, data does bear some characteristics that are similar to those of natural resources-such as forests, water,or fisheries, as it can be extracted and monetized.The commons,therefore could theoretically function as a social system through which people can control, manage and distribute resources.Kapoor and Ramesh highlight two parallels between data and public goods resources. First, they argue that data is non-depletable and non competitive.[52] We disagree with this framing. Unlike water or forests, user generated data emanates from the body or persona of an individual, which means that the fundamental rights guaranteed to the individual also vests in their data. From this right stems a primary beneficial interest, which may be transferred or modified, even though the right itself is inalienable.Second, they state that " data is a resource that is more valuable when packaged together rather than siloed or broken down into individually owned chunks."[53] While the statement itself is valid and speaks to the network effects that have been responsible for propelling big tech to its position in the economic hierarchy, it suffers from the same drawbacks of framing ownership models or property rights over data-which we discussed in the previous section.

---

[49] Finck, Michele and Pallas, Frank.2020 " They who must not be identified-distinguishing personal from non-personal data under the GDPR," International Data Privacy Law.10

[50] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP216) 0829/14/EN, 11-12, 23-25.

[51] Ostrom E (1990) Governing the commons: The evolution of institutions for collective action (Cambridge University Press,1990,1st ed) 30

[52] Ramesh A and Kapoor A (2020), " Principles for revenue models of data stewardship," The Data Economy Lab,,https://thedataeconomylab.com/2020/07/31/principles-for-revenue-models-of-data-stewardship/

[53] Ibid

Data as exercise of decisional autonomy

Autonomous decision-making is a pre-requisite for respecting individuals as persons-as agents free to make their own choices.[54] Therefore, to guarantee that level of unrestrained decision-making, decisional privacy becomes an important interest, and has thus been recognised across jurisdictions as a core facet of a right to privacy. As a result, privacy discourse in the pre-Big Data era studied aspects of autonomy and decisional privacy in sync with each other.

Several scholars have identified several dimensions of informational privacy, including privacy of body,thought, and decision-making. [55]Decisional privacy has broadly been defined as the right against unwanted access or interference in an individual's decisions and actions. [56] Decisional privacy includes both narrow choices such a same-sex marriage, reproductive liberties,and child rearing.[57] A broader conception of decisional privacy encompasses not only these intimate choices but also actions, behaviour and lifestyle choices.[58] In essence decisional privacy grants the autonomy to carry out an individual's chosen life across social contexts and "a distinct type of privacy, which protects the autonomy of persons to make decisions about their body or other"[59] which echoes the constitutional ideal of autonomous decision-making.Roessler argues that while privacy cannot be reduced to another value like autonomy, individuals value privacy because of the autonomy that it provides. [60]

Indeed, there is a difference between autonomy and decisional privacy. The loss of decisional privacy does not necessarily entail an immediate consequential loss of autonomy.However, in several cases, this may be true. For example, if an individual is subjected to state surveillance,then decisional privacy is automatically violated, although that may not influence the autonomous decisions that individual makes.The extent to which a loss of decisional privacy impacts autonomy is determined by several other social,political and economic factors.In the present instance, if the individual were a social activist or journalist working against draconian laws implemented by the same government conducting the surveillance, the chilling effect resulting from a loss of decisional privacy would likely impact the autonomy of the decisions made by said individual.As another example,Chandrashekhar argues, a digital trail of a woman's health records creates " a digital trail of choices exercised about one's body

[54] Benn, S. I. (1971). Privacy, freedom and respect for persons. In F. Schoeman (Ed.), Philosophical dimensionsof privacy: An anthology. Cambridge: Cambridge University Press.
[55]Lanzig M (2019), " " Strongly recommended": Revisiting decisional privacy to judge hypernudging in self-tracking technologies," 32 Philos. Technol 549-568,556
[56] Allen, A. L. (1988). Uneasy access: privacy for women in a free society. Totowa: Rowman and Littlefield.
[57] Lanzig M (2019), "Strongly recommended": Revisiting decisional privacy to judge hypernudging in self-tracking technologies," 32 Philos. Technol 549-568,556
[58] Roessler, B. (2005). T*he value of privacy.* Cambridge: Polity Press.
[59] Koops, B.-J., et al. (2017). A typology of privacy. University of Pennsylvania Journal of International Law,38(2), 483–575.
[60] Roessler, B. (2005). *The value of privacy.* Cambridge: Polity Press.

under [the[state vision [and thus] hampers woman's autonomy to make decisions related to their bodies and life."[61]

Decisional privacy in essence speaks to the moment at which an individual chooses to do or not do something.Decision-making about one's body,lifestyle choices or intimate relationships are all connected to decisional privacy.Courts across jurisdictions have accepted the protection of privacy as decisional autonomy.In *Roe v Wade* the US Supreme Court held that "right of privacy, whether it be founded in the Fourteenth Amendment's concept of personal liberty and restrictions upon state action, as we feel it is, or, as the District Court determined, in the Ninth Amendment's reservation of rights to the people, is broad enough to encompass a woman's decision whether or not to terminate her pregnancy."[62]The same court held that the right to privacy also includes the right to engage in consensual sexual activity in one's home,regardless of sexual orientation.[63] In a different continent and century, the Indian Supreme Court also emphatically recognised decisional autonomy in its 2017 judgment *KS Puttaswamy v Union of India.*[64]The concept of decisional autonomy spilled across the concurrent opinions of three judges on the bench.[65]Justice Chelameswar spoke of privacy as "repose,sanctuary and intimate decision." (para 36) Justice Bobde and Justice Nariman (para 81) referred to the significance of choice in associated freedoms.

We posit that an individual's interest in their data cannot be analogised through financial or material conceptions such as oil, currency or property.instead, this interest protected by the right to decisional privacy can only be captured through the conception of data as an exercise of decisional autonomy. Any data generated and parted with by an user is fundamentally an exercise of that individual's decisional autonomy and should be treated as such by the state, private processors and other individuals.

Community Data

'Community data' is an Indian policy innovation that has emerged over the past two year from a desire to ensure that the data of Indian citizens and communities are used for their benefit and not monetized solely by private players. [66]Built on the edifice of 'group privacy,' this conception of data frames an important question-" "individuals are supposed to [control] their

[61] Chandrasekhar R (2018), "Here are the consequences of linking women's medical records to their Aadhaar," Indian Express,Apr 24https://indianexpress.com/article/gender/here-are-the-consequences-of-linking-womens-medical-records-to-their-aadhaar-5139360/

[62] Roe v Wade [410 U.S. 113 1973]

[63] Lawrence v Texas 539 U.S. 558 (2003)

[64] (2017) 10 SCC 1.

[65] Bhatia G (2017)," The Supreme Court's Right to Privacy Judgment – V: Privacy and Decisional Autonomy," Indconlawphil, Aug 31,https://indconlawphil.wordpress.com/2017/08/31/the-supreme-courts-right-to-privacy-judgment-v-privacy-and-decisional-autonomy/

[66] See Sinha A and Basu A (2020)," Community data and decisional autonomy: Dissecting an Indian legal innovation for emerging economies," **[File on source with author]**

data, why should data about groups/communities not, similarly, be [controlled] by the corresponding group/community?"[67]

We posit that community data as a conception that can preserve the beneficial interests of a community. Literature on a group right to privacy argues that it arises from the failings of traditional personal data protection frameworks in protecting group interests.[68]Big data and algorithms enable analysis that focus on attributes of personal data including membership of individuals in certain groups. Even in cases where individuals have provided informed consent on the processing of their personal data, this may be used to draw inferences about the group the individual is a part of, and by extension has implications for the individual themselves without the individuals knowing. A community interest in data therefore arises from power asymmetries in the data economy-where the private companies acting as data processors exercise beneficial interests-control,monetization, and alienation over individual data-insights from which can be used to subsequently target and discriminate against groups. After a series of unclear policy moves attempting to conceptualize community data, the Ministry of Electronics and Information Technology released a non-personal data framework that is the first such attempt at defining, constructing and charting the contours of community data-although several gaps in its framing remain.[69]The report defines a community as " any group of people that are bound by common interests and purposes and involved in social and/or economic interactions.It could be a geographic community, a community by life, livelihood, economic interactions or other social interests and objectives and/or an entirely virtual community."[70] This definition casts an ambiguous net on the notion of a community, and also fails to underscore the relationship between an individual and a community. It then goes on to note that "community non personal data" includes personal data that has been anonymised and non personal data about animate and inanimate phenomena. The example provided to justify this claim is unclear. It states that data collected by municipal corporations or private players such as ride-hailing companies falls within this framing of community data, which appears to suggest that all users of ride-hailing companies fall within an identifiable community.

The key challenge in conceptualizing community data therefore lies in identifying a community itself.The traditional understanding of a group stems from a degree of shared perception either by members of the group or by outsiders or by both.[71] In some cases, the members of a group are 'self-aware' and thus identify as one, thereby also claiming beneficial interests as a

[67] Singh, Parminder.2019. "Community data in the draft e-commerce policy." Medianama.https://www.medianama.com/2019/03/223-community-data-in-the-draft-e-commerce-policy/. Accessed August 1 2020.
[68] Floridi, L (2018)"Group Privacy: A Defence and an Interpretation."in Taylor, Linnet and Floridi,Luciano and Van Der Sloot, Bart. Group Privacy: New Challenges of Data Technologies.Springer International Publishing
[69] Available at https://ourgovdotin.files.wordpress.com/2020/07/kris-gopalakrishnan-committee-report-on-non-personal-data-governance-framework.pdf
[70] Ibid 14
[71] Taylor et al 38

collective.Groups might also not be 'self-aware'-when external perceptions result in the perception of a group.[72] For example, society or the government might brand a set of activists as a group of 'dissidents' or 'rioters,' and the individuals grouped might not associate either with the tag or with the others in the group.The advent of big data, and its ability to derive insights has brought out a third scenario-where neither the members of the group nor those external to it are aware of the group's existence. The group is only created through algorithmic processing of data that provides insights on trends in individual behaviour that may result in them being classified as groups.[73] For example, algorithmic processing might reveal similarities in the tastes and preferences of individuals within a certain income bracket who reside in a certain area. The existence of this group is not known either to the members themselves or any other individuals who are not part of the group. The group comes into existence only due to data processing.As we discuss in Part II, the interests that each kind of group and the individuals that make up each kind of group differ.

Data as sovereignty

When exploring the power asymmetries inherent in the data economy, we must highlight the fact that platform companies processing much of the world's data are all from the US. Through data generated by users in the global south, platform companies reap economic dividends through data stored and processed in their home jurisdiction, thus avoiding regulatory scrutiny from the state where they operate. This exploitative situation has been emphatically branded as 'data colonialism,' by several global south actors,likening the behaviour of Big Tech to private players that had served as the catalysts of colonialism several years ago.[74] The present power structures of the global digital economy undermines the sovereign rights of global south countries to guarantee both economic welfare and civil and political rights to its citizens. [75]This is not the first time that global governance has been compelled to confront this challenge to preserve the tenet of sovereign equality.

The principle of permanent sovereignty over natural resources marks one of the most hallowed developments in international law in the latter half of the twentieth century and is now firmly embedded within the notion of state sovereignty itself.[76] The principle was articulated by the recently decolonised developing countries in the 1950s to claim ownership over natural resources in their territories.[77] The articulation was fuelled by concerns that orthodox international law disciplines such as foreign investment law and the law governing the high

[72] Ibid
[73] Ibid 39
[74] https://swarajyamag.com/magazine/colonialism-20-truly
[75] Nick Couldry and Ulises Mejia Couldry, "Data colonialism: rethinking big data's relation to the contemporary subject" Television and New Media<https://eprints.lse.ac.uk/89511/1/Couldry_Data-colonialism_Accepted.pdf>
[76] . U.N. Int'l Law Comm'n, Draft Conclusions on Identification of Customary International Law, with Commentaries, U.N. Doc. A/73/10 (2018)
[77] L.S. Clark, 'International law and natural resources', 4 *Syracuse Journal of International Law and Commerce* (1976), 377–390 at 378.

seas at the time undermined the exercise of the state's sovereign rights, favouring capital exporting states and corporations.[78] PSNR was understood as being critical to enabling countries in the Global South to realize their development potential.

PSNR stems from the demand for economic sovereignty and the right to self-determination in developing countries. After decolonisation, states realised that self-determination was worth little if foreign exploitation of resources within their jurisdiction continued with aplomb.[79] Through this principle, developing countries asserted an 'inalineable,' 'absolute', and 'permanent right' over their natural resources[80].The present conception of PSNR evolved over several decades, starting from the 1950s.[81] Notably, the Declaration on Permanent Sovereignty over natural resources was adopted in 1962 and sought to balance the rights of capital exporting and importing countries by limiting expropriation only to instances where it was based on public interest and appropriate compensation was paid. [82]The declaration explicitly recognized the link between sef-determination and PSNR. In the 1970s, after incorporation into several other international instruments, it morphed into a wider political campaign aimed at the creation of a New International Economic Order (NIEO) that was more just and equitable. [83] Following this global campaign, there was a widespread realization among the global community through the turn of the century that the initial economic agreements inked during the colonial era were unjust and inequitable.[84]

As the PSNR doctrine became entrenched in international law,it had to strike a balance between the rights and corresponding duties of states that arose as a result of the construct.[85] First, in asserting claims to PSNR, a host state must respect the rights of other states,as per international law. Second, the state must ensure that the entire population benefits from the assertion of permanent sovereignty. International law has extended the Westphalian construct of state sovereignty to include non state actors and individuals as bearers of rights.[86] Therefore, any right asserted by the state, including a right of sovereignty over resources must be exercised for the benefit of its people-a principle captured in innovations such as the

---

[78] David P. Fidler, Revolt Against or From Within the West?: TWAIL, the Developing World, and the Future Direction of International Law, 2 CHINESE J. INT'L L. 29, 32–3 (2003).

[79] L.S. Clark, 'International law and natural resources', 4 *Syracuse Journal of International Law and Commerce* (1976), 377–390 at 378.

[80] M Sornarajah, The Pursuit of Nationalized Property (Martinus Niijhoff,1986) 120

[81]Alam S and Faruque A (2019), " From Sovereignty to Self-Determination: Emergence of Collective Rights of Indigenous Peoples in Natural Resources Management," 32 The Georgetown Envt Law Review, at 64

[82] G.A. Res. 1803 (XVII), at 15 (Dec. 14, 1962), https://undocs.org/en/A/RES/1803(

[83] https://unesdoc.unesco.org/ark:/48223/pf0000035806_eng

[84]Garcia-Amador F (1980)," The Proposed New International Economic Order: A New Approach to the Law Governing Nationalization and Compensation", 12 U. Miami Inter-Am. L. Rev.1

[85] Alam S and Faruque A (2019), " From Sovereignty to Self-Determination: Emergence of Collective Rights of Indigenous Peoples in Natural Resources Management," 32 The Georgetown Envt Law Review

[86] Andrew Clapham, The Role of the Individual in International Law, *European Journal of International Law*, Volume 21, Issue 1, February 2010, Pages 25–30, https://doi.org/10.1093/ejil/chq001

public trust doctrine.Third, as the bearers of the right of PSNR, a duty has evolved to manage resources in an environmentally sustainable manner. Fourth, economic globalisation has altered the traditional conceptions of state sovereignty and laid an emphasis on international co-operation and interdependence.

While data cannot be analogised as a resource in any form, the evolution of PSNR has several bearings for the correction of power asymmetries between states. In essence, sovereignty over data indicates that a state has the right to govern data generated by its citizens or within its jurisdictions as per its domestic law and policy, in line with the principles of international law. This is the conception of data as sovereignty.

An extension of the PSNR principle in International law to the digital realm further dictates that states govern data in a manner that preserves civil and political rights and furthers economic welfare of its citizens. It also means that states must respect the rights of other states when applying data sovereignty with a view to furthering international co-operation. The final decision-maker applying these principles, however must be the state itself, without influence from other states or Big Tech. The conception of data as sovereignty is important for global south states to reclaim the power to tax, regulate, and govern Big Tech platforms in a manner consistent with the interests of its citizens.

Part 2: Unpacking beneficial interests

The table below provides an early blueprint of the matrix of relationships between each key stakeholder who may have a beneficial interest in data, and how this interest may be exercised as against other stakeholders. The need for an identification of such inherent beneficial interests in data outside the purview of contractual rights arises from the power structures that have emerged in the current data economy. The exploitative nature of mining personal data creates an imbalance in the benefits accrued by those whose data is utilized for financial gain and those monetizing on having access to personal data. There has been an upswell of discontent, particularly in the Global South with several commentators claiming that excessive focus on consent has skewed the discussion in favour of US based technology corporations who reap monetary dividends from data gathered from Global South citizens, thereby leading to accusations of 'data colonialism.'[87] Our proposal below stems from the recognition of a need to redistribute power over data in a manner that fair and equitous to all actors involved in its generation. In order to do so, we draw from established and evolving regulatory and legal theories. For each stakeholder, we articulate an appropriate conception of data that is best suited to preserve the beneficial interests of the stakeholder in its relationship with other stakeholders.

[87]Couldry N and Mejia U (2019), "Data colonialism: rethinking big data's relation to the contemporary subject" Television and New Media<https://eprints.lse.ac.uk/89511/1/Couldry_Data-colonialism_Accepted.pdf>

Individuals

To start with, the data of an individual in relationship to other stakeholders is best understood as an exercise of decisional autonomy-a far more appropriate analogy than any of the other hackneyed analogies that adopt a propertarian view of data. Data as an exercise of decisional autonomy enables the safeguarding of an individual's beneficial interests in all relationships.

First, with other individuals,autonomy should provide an individual the breathing space to exercise free choice in one's lifestyle, preferences,and other forms of decision-making. In today's datafied world, these choices often come through in the form of user-generated data. Gavison argues that decisional privacy, which as we discussed previously is an interest allied with decisional autonomy, protects an individual from the 'chilling effect' of being compelled to comply with social norms due to fear of social sanction.By treating data as an exercise of decisional autonomy, we capture the essential link between the individual and their data and thus preserve the right of decisional privacy.

The same framing applies when unpacking the individual's relationships both with the state and private processors.

Informed Consent

Most modern laws and data privacy principles seek to focus on individual control. In this context, the definition by the late Alan Westin, which characterises privacy as "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to other,"[88] is most apt. The idea of privacy as control is what finds articulation in data protection policies across jurisdictions beginning from the Fair Information Practice Principles (FIPP) from the United States.[89] Schwarz, called the FIPP the building blocks of modern information privacy law.[90] These principles trace their history to a report called 'Records, Computers and Rights of Citizens'[91] prepared by an Advisory Committee appointed by the US Department of Health, Education and Welfare in 1973 in response to the increasing automation in data systems containing information about individuals. The Committee's mandate was to "explore the impact of computers on record keeping about individuals and, in addition, to inquire into, and make recommendations regarding, the use of the Social Security number."[92] The most important legacy of this report was the articulation of five principles which would not only play a significant role in the privacy laws in US but also inform data protection law in most privacy regimes internationally like the OECD Privacy Guidelines, the EU Data Protection Principles and later, the GDPR, the FTC Privacy

[88] Westin A (1998)Privacy and Freedom, Atheneum, New York,
[89] FTC Fair Information Practice Principles (FIPP) available athttps://www.it.cornell.edu/policies/infoprivacy/principles.cfm.
[90] Schwartz P,""Privacy and Democracy in Cyberspace," 52 Vanderbilt Law Review 1607, 1614
[91] US Secretary's Advisory Committee on Automated Personal Data Systems, Records, Computers and the Rights of Citizens, available at http://www.justice.gov/opcl/docs/rec-com-rights.pdf.
[92] *Id*.

Principles, APEC Framework. Fred Cate effectively summarises the import of all of these privacy regimes as follows:

> *"All of these data protection instruments reflect the same approach: tell individuals what data you wish to collect or use, give them a choice, grant them access, secure those data with appropriate technologies and procedures, and be subject to third-party enforcement if you fail to comply with these requirements or individuals' expressed preferences"*[93]

This is intended to make the individual empowered and allows them to weigh their own interests in exercising their consent. The allure of this paradigm is that in one elegant stroke, it seeks to "ensure that consent is informed and free and thereby also to implement an acceptable tradeoff between privacy and competing concerns."[94] For the purposes of our discussion, informed consent is critical in both articulating and operationalising the beneficial interests of individuals in their data. A functioning model of informed consent involving both legal and technological tools, allows the individual to determine the ways in which her data will be used by both state and private parties.

Legitimate Interests

The notion of legitimate interest was first introduced in the EU Directive 95/46/EC under Article 7[95] and has subsequently been adopted in the GDPR under Article 6. In both contexts, the purpose of legitimate interests is to provide an additional ground for processing of personal data which also includes consent, contractual arrangement, legal obligation or other specifically identified rationale as other grounds for processing. In brief, legitimate interests involves taking into consideration the nature and source of the legitimate interest, the impact on the interest, fundamental rights and freedom of the data subject and the nature of safeguards in a balancing test before a conclusion is reached on whether the legitimate interest can be a lawful ground for processing in that specific instance. Due to confusion in its interpretation, the Article 29 Working Party, in 2014,[96] looked into the role of legitimate interest and arrived at the following factors to determine the presence of a legitimate interest— a) the status of the individual (employee, consumer, patient) and the controller (employer, company in a dominant position, healthcare service); b) the circumstances surrounding the data

[93] Cate FThe Failure of Information Practice Principles, available athttp://papers.ssrn.com/sol3/papers.cfm?abstract_id=1156972.
[94] Sloan R and Warner R, Beyong Notice and Choice: Privacy, Norms and Consent, 2014, available at https://www.suffolk.edu/documents/jhtl_publications/SloanWarner.pdf.
[95] EU Directive 95/46/EC – The Data Protection Directive, https://www.dataprotection.ie/docs/EU-Directive-95-46-EC-Chapter-2/93.htm.
[96] Article 29 Data Protection Working Party, "Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC," http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf.

processing (contract relationship of data subject and processor); c) the legitimate expectations of the individual.

Moerel and Prins recommend five factors[97] as relevant in determining the legitimate interest. Of the five, the following three are relevant to the present discussion:

- Collective Interest — A cost-benefit analysis should be conducted, which examines the implications for privacy for the data subjects as well as the society, as a whole.
- The nature of the data — Rather than having specific categories of data, the nature of data needs to be assessed contextually to determine legitimate interest.
- Contractual relationship and consent not independent grounds — This test has two parts. First, in case of contractual relationship between data subject and data controller: the more specific the contractual relationship, the more restrictions apply to the use of the data. Second, consent does not function as a separate principle which, once satisfied, need not be revisited. The nature of the consent (opportunities made available to data subject, opt in/opt out, and others) will continue to play a role in determining legitimate interest.

Legitimate interests as a ground of processing adds an additional layer of protection through which the limitation of informed consent can potentially be addressed. It provides an alternative model to determine use limitations in data processing where the reasonable expectations of the individual, and other principles of balancing tests can be used to ensure that the beneficial interests of the individual are not harmed.

The individual's relationship with the community and the protection of its interests can also be understood through the lens of decisional autonomy.If we consider the import of jurisprudence on decisional autonomy, it is clear that when individual rights and group rights are in conflict, individual rights prevail. A woman's right to choose prevails over group concerns on the institution of marriage.Therefore, any interest being claimed by the community on behalf of the individual must be with the individual's explicit consent. The Non Personal Data Committee Report in India appears to endorse this conception of decisional autonomy and extends the jurisprudence put forward by courts in *Lawrence v Texas* or *Puttaswamy,* without explicitly referring to it.It does so by adopting a midway between the contrasting approaches adopted in Recital 26 GDPR and the Article 29 Working Report on de-anonymisation which we discussed in the previous section. The NPD report recognizes the challenges with irreversible de-identification and seeks to solve this issue by stating that there should be explicit consent from an individual when their anonymised data is being processed.

[97] Moerel, E.M.L. and Prins, J.E.J. (Corien), Privacy for the Homo Digitalis: Proposal for a New Regulatory Framework for Data Protection in the Light of Big Data and the Internet of Things (May 25, 2016). Available at SSRN: https://ssrn.com/abstract=2784123 or http://dx.doi.org/10.2139/ssrn.2784123

State

The conception of data when governing a state's relationship with other stakeholders varies with regard to context. When understanding inter-state dynamics, the conception is data as sovereignty, along with the associated rights and obligations that we discussed in the previous section.Against private actors, the state has rights accruing out of the conceptions of a 'data commons' through which the private actor would need to open up datasets so that other individuals and organisations can benefit.

The state plays two key roles here. The first is in the enforcement of the public utilities doctrine to ensure that large technology players do share data with commons. The second role is as a facilitator of this commons towards ensuring that data is being shared and governed efficiently. Further,through clearly articulated policy can also enable state entities to use the data commons to take public policy decisions using insights derived from the data. The state's relationship with the individual and with the community should be conceptualised through the lens of decisional autonomy and community data respectively. In both cases-both entities have clear rights against the state. The community enjoys the right to self-determination.Any exceptions to informed consent, which is an aspect of the right of decisional privacy must be made only in case of legitimate interests.

Community

Community data brings with it several beneficial interests for the community-which fall within the broad umbrella of a right to self-determination. Self determination is a core principle of international law.Shaw has defined it as "a people's pursuit of its political, economic, social and cultural development within the framework of an existing state."[98]While it was initially limited to situations where "people" overthrow one form of government and opt for another-now termed as 'external self-determination.' [99] However, now some have posited a notion of internal self-determination that provides groups continuous political,social rights and allow minority groups to enjoy state protections and also enjoy a degree of autonomous as a group. This understanding is enmeshed in Article 1 of the ICCPR. While internal self-determination has been contested in international law,we believe that a right to self-determination for groups is the appropriate beneficial interest that a community should have in its data. The manifestation of this interest will differ depending on whether the group is self-aware, externally determined or algorithmically created.For self-aware groups this includes the following

[98] In 1962, the United Nations General Assembly recognised the "right of peoples and nations to permanent sovereignty over their natural wealth and resources." It is clear articulation of not only of group interests but also its right to have it say over resources deemed crucial to the collective interests of the group. Shaw, Malcolm.2003.International Law, Fifth Edition. Cambridge University Press.Cambridge
[99] Ibid

1.A right to group privacy which involves informed consent by the community as a whole. This includes processing with informed consent and legitimate interes,

2. It also includes a beneficial interest in insights derived from community data that is additional to the general interest individuals and communities would have in data that is shared into the commons.This is particularly significant for indigenous and vulnerable communities.[100] These beneficial interests need to be determined by the group itself and could include right to use and access insights from the data,an obligation to consult groups regularly when using data,acknowledgment of the use of the group's data in the framing of public policy and so on.

3.Further, self-aware groups also have an interest in ensuring that any data collected from the community or individuals within the community should not be used to illegally disadvantage that community. For example, insights generated about extremist groups from data generated by individuals belonging to said community might be used by the government to track,monitor and intervene to disadvantage the group. However, this disadvantage is not illegal. This is very different from a scenario where data collected about people of colour residing in a certain residential area is used to illegally conduct police surveillance on these communities.Additionally, most jurisdictions have anti-discrimination laws against protected categories so using data to discriminate against them would amount to an illegal disadvantage.[101]

In the case of groups that are algorithmically created or externally perceived, it is difficult to envisage a way in which the community itself can bargain for. Therefore, the beneficial interests associated with community data that are present in the case of self-aware groups cannot be conceptualised here. However, the obligation to protect privacy and not illegally disadvantage the community using data generated by them remains.

For instance, data about the income and preferences of white men living in Long Island may be used by the state to take or justify policy decisions, which may include affirmative action measures such as quotas for people of colour in educational institutions or government.While this is a disadvantage for the community, it is not an illegal disadvantage regardless of whether the men in Long Island are aware of their algorithmic existence as a group. Legality should be determined based on the legal and constitutional framework of a state in conjunction with international law.

---

[100] SeeTsosie R,,"Tribal Data Governance and Informational Privacy: Constructing "Indigenous Data Sovereignty", 80 Mont. L. Rev. 229(2019)

[101] Khaitan, Tarunabh, A Theory of Discrimination Law (May 15, 2015). A Theory of Discrimination Law (Oxford University Press 2015), Oxford Legal Studies Research Paper No. 40/2015, Available at SSRN: https://ssrn.com/abstract=2628112

Private Firm

Currently, the nature of data ecosystem and data driven business models severely privileges the interests of the private firms which collect data, particularly the BigTech firms which hold significant competitive advantages over smaller firms through their use of data as market power. The size of network power that the big technology firms control is unprecedented. Google drives 90 per cent of the internet search;[102] 95 per cent of young adults on the internet use some product owned by Facebook, and Amazon.com now accounts for 75 per cent of electronic book sales. Apple and Microsoft Corp. supply 95 per cent of desktop operating systems, while 99 per cent of mobile phone operating systems are shared between Google and Apple.[103] Earlier anti-competitive policies only allowed companies to either create markets, or participate in them, but they could not play both roles. By having a creator of the market also participate, there would be an unfair and inherent advantage. The Chicago school's benevolent view of monopolies allows Facebook and Google to both create markets and be a participant. When Google acquired DoubleClick, it was essentially acquiring a means to influence a market in which it was already a participant. It allowed Google to favour its own businesses in the online advertising space through its use of DoubleClick. Similarly, Google's purchase of YouTube was the acquisition of a new market, which enabled it to tweak the YouTube algorithms to give preferential treatment to its own content on the platform. These are clear examples of "two-sided markets". Two distinct user groups that when brought together, provide each other with network benefits. Consumers are both the source of data, as well as the product. The advertisers are the customers, and they provide the market's revenue and depend on the scale of the market. The kind of scale that BigTech companies enjoy, provides advantages which are irreversible in a competitive system without interventions.

In most competition law regimes, dominant positions are determined on the basis of market share, size and commercial advantages.With digital business models, an emerging indicator of dominant position is also access to data. Given the network effects, access to data has the potential to be a significant barrier to entry. Germany's Federal Cartel Office held in 2019 that Facebook was exploiting consumers by forcing them to consent to the terms of data collection if they wanted to open an account. The German regulator also held that Facebook used its vast data collection to set up its position of dominance, effectively taking meaningful choice away from the consumers.[104]

This evolving understanding of competition law which addresses the value of data as market power, and accounts for network effects in defining the 'relevant market' can be instructive in

[102] Odrozek, Kasia. "More than 90% of the World Uses Google Search." Mozilla. May 04, 2018. Accessed July 26, 2019. https://internethealthreport.org/2018/90-of-the-world-uses-google-search/.
[103] Ip, Greg. "The Antitrust Case Against Facebook, Google and Amazon." The Wall Street Journal. January 16, 2018. Accessed July 26, 2019. https://www.wsj.com/articles/the-antitrust-case-against-facebook-google-amazon-and-apple-1516121561?mod=e2fb&fbclid=IwAR2-BlQ_kuQynGEmI49U450cXmWWvHhfjhw-iy5odR0bUSxoG1ZuNj266es.
[104]https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2019/07_02_2019_Facebook.html

how we must think of the limitations to a private firm's beneficial interest in the data it has amassed.

Definitions of personal data which are too prescriptive, such as catalogue approach used in the Massachusetts breach notifications statute[105] or the Information Technology Act in India,[106] are too restrictive or likely to be outdated every few years. A better model is to look at three kinds of data being captured — volunteered data (data actively provided by individuals such details in a form when they sign up for a service); observed data (behavioral data generated through individual's use of the service); and inferred data (data neither actively nor passively provided by the individual, but arrived at through analysis of collected data by data processors). This classification is also endorsed by the Article 29 Working party's report on data portability.[107] The report clearly excludes inferred or derived data from the scope of data portability rights of individuals, thus indicating that it falls outside of data on which there are clear rights of individuals. While inferred data is based on the personal data provided by the individual, it involves labour and intellectual effort of the data controllers, thus implying a clear beneficial interest in such data. However, any beneficial interest in the insights obtained data must be extracted with respect for the rights of individuals, communities and the state.

Conclusion

Power asymmetries define today's data economy. Traditional conceptions of relationships between various actors in political economy do not capture the myriad ways in which these power asymmetries are defined today. Existing conceptions of data as a resource through hackneyed analogies with natural resources miss the essence of an individual's relationship with their data—an exercise of their decisional autonomy. Through this framing, an individual is able to assert a right to decisional privacy as an inalienable right rather than trading away this right in a situation where one's bargaining power is determined by income status. Other stakeholders have beneficial interests as well, and coming up with appropriate conceptions of data is imperative for articulating the nature of these beneficial interests.

Data cannot be constrained by a single analogy or conception. The conception should be context specific, created to account for a specific power asymmetry between two stakeholders. The table below summarizes these relationships, as discussed in Part 2 of the paper. Like with any theoretical abstraction, there will be practical problems our framing will run into, and those problems have practical solutions, rooted in the jurisprudence of the countries and the social, political and economic framework to implement this jurisprudence. This paper does not attempt to solve all public policy challenges but aims to equip stakeholders, including

[105] https://www.mass.gov/service-details/requirements-for-data-breach-notifications.
[106] https://www.wipo.int/edocs/lexdocs/laws/en/in/in098en.pdf.
[107] Article 29 Data Protection Working Party, Guidelines on the right to data portability. Adopted on 13 December 2016. As last Revised and adopted on 5 April 2017. Available at https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611233.

researchers and civil society with a tool to understand the contours of the policy challenges themselves.

RECIPIENT OF BENEFICIAL INTEREST	**Against Individual**	**Against State**	**Against community**	**Against private platform**
Individual	**Conception:** Data as exercise of decisional autonomy, **Beneficial Interest:** Decisional Privacy, informed consent, legitimate interest in processing	**Conception:** Data as exercise of decisional autonomy, **Beneficial Interest:** Decisional Privacy, informed consent, legitimate interest in processing	**Conception:** Data as exercise of decisional autonomy, **Beneficial Interest:** Decisional Privacy, informed consent	**Conception:** Data as exercise of decisional autonomy, **Beneficial Interest:** Decisional Privacy, informed consent, access to data as part of commons
State	**Conception:** Data as decisional autonomy **Beneficial interest**:None without informed consent of the individual	**Conception:** Data as sovereignty **Beneficial interest:** Right to govern as per sovereign framework	**Conception:**Community data **Beneficial interest:** Public trust	**Conception:** Data commons (public utility doctrine) **Beneficial interest:** Data sharing under public trust doctrine
Community	**Conception:** Data as decisional autonomy **Beneficial interest:** None without informed consent of the individual	**Conception:** Community data, **Beneficial interest:** Group privacy,right against using data for illegally disadvantaging the community	**Conception:** Community Data **Beneficial interest:**Group privacy	**Conception:C**o mmunity Data, **Beneficial Interest:** Group Privacy, access to data as part of commons
Private Platform	**Conception:** Data as	**Conception:** Data commons	**Conception:C**om munity Data	**Conception: Data Commons**

	decisional autonomy, **Beneficial interests:** Financial or proprietary interest in insights derived from data	**Beneficial interest:** Financial or proprietary interest in insights derived from data	**Beneficial interest:**Financial or proprietary interest in insights derived from data	**Beneficial interest:** Financial or proprietary insight in insights derived from data, access to data as part of commons