Generative AI Systems Usage of Personal Data: Utility, Efficiency, and Competition Concerns

Emanuele Giovannetti[1],

Anglia Ruskin University and Hughes Hall, University of Cambridge

Summary

The objective of this contribution is to share information about results on the economic impact of digitalization and the portability of user data, through Generative AI systems in different platform markets. This is an area where Generative AI systems may reinforce the risks of entrenching network bottlenecks, preventing entry, or reinforcing tendencies towards tipping in platform markets. These risks are mainly due to the dynamic effects obtained through algorithm training and Generative AI system parameter refinements via the access and processing of personal data by dominant digital platforms. These effects might also exert different impacts on consumers depending on their level of "captivity," expressed by their asymmetric degrees of lock-in, that are artificially created through the Generative AI systems' ability to profile and market services based on this personal data. On the other hand, access to personal data, might also help build trust on decentralized platforms that address relevant issues of digital and financial exclusion.

Given the conflicting evidence and results on the impact of the usage of personal digital data by Generative AI systems on the competitiveness of digital markets, we recommend that relevant regulatory and competition authorities' assessments of the impact of the collection and analysis of personal digital data and their role in the training of the algorithms and refinement of parameters of Generative AI systems carefully consider its many often-conflicting dimensions.

These include: the beneficial impact on the enhanced visibility and shaping of digital identities of the data originators; the systemic efficiencies digital personal data, when fed into Generative AI systems, bring in the management, planning, and forecasting of digital infrastructures, and the impact on the competitive landscape, due to the economic rent that providers might derive from being able to access, process, transform, and use these digital data, whereby this might create new barriers to entry and competition into these platform markets, for example by allowing data-based price and quality discrimination of services.

Keywords: *Personal digital data, Generative AI, data portability, digital platform, Asymmetric switching costs, lock-in costs, entry, Entrenched Market power, Network Centrality as a measure of Bottlenecks in Digital Ecosystems*

[1]Email address emanuele.giovannetti@aru.ac.uk

Postal Address, ARU, East Road, CB1 1PT, Cambridge, UK

1. Introduction

Generative AI systems typically enhance the visibility and public availability of algorithmic elaborations based on large and real-time-changing personal data sets within a digital network. This can bring significant economic benefits to the original data owners, as availability and visibility can help owners' reachability by other network agents, facilitating data traffic exchanges across digital supply chains (D'Ignazio and Giovannetti, 2014 and 2015) or between different sides of a trading platform (Giovannetti and Siciliani, 2020 and 2023).

Richer and more complex data analytics, obtained from applying algorithms to the original personal data, may also help in forming and shaping digital identities, whose features can be essential in determining success or failure in digital businesses. An example, particularly relevant in addressing the impact of the barrier to access to credit typically faced by women led microbusinesses, is that of online crowdfunding. In crowdfunding, the project proposer looking for online funders, publicly displays a digital identity that is often combined with additional publicly visible data, for example, relating to the network of funders and supporters of the project proposer. All these original and derived project data, when fed into appropriate algorithms, for example, to calculate a project's network centrality, contribute to forming the "Latent Network Capital" of a project, itself a trust enhancing signal, that is essential in determining a project's success in raising the necessary funds through crowdfunding platforms (Davies and Giovannetti, 2018 and 2022).

Generative AI systems' elaboration of digital user data also has critical economic value for the efficiency, planning, and forecasting of large digital infrastructures. For example, digitalization is important for integrating the intermittent nature of renewable energies because it helps grids better match and incentivize energy demand patterns that are linked to maps and forecasts of how strong the sun and wind will be (Llorca et al., 2023; IEA,

2017). Hence, Generative AI systems' elaboration of digital user data available through the digitalization of the energy grid is seen as a critical step for the European Green Deal, focusing on delivering the EU's 2030 climate targets and the Green Energy Transition. Digitalization is required for the integration of renewable energy production and usage and to integrate users' granular-level data into regional energy infrastructures (European Distributed Data Infrastructure for Energy, 2023, Rossetto and Reif, 2021).

However, in addition to the increased value of the services provided to the personal data owners and their essential role in planning, forecasting, and managing complex data spaces and infrastructure with improved efficiencies, the Generative AI systems processing of digital personal data also exerts a critical impact on the possibility of entrenching market power for gatekeepers and controllers of network bottlenecks, in the digital platform markets. This is the case as the Generative AI system outcomes, based on the algorithmic use and interaction of personal and derived/aggregated data, with the internal system parameters, (which are themselves revised and updated through the personal data, sometimes through Bayesian mechanisms), can be used by incumbent Platform and Service providers to supply improved profiled personalized services and ads, that, while improving user experience, also introduce lock-in effects and switching costs (Klemperer, 1987), making it more difficult for the original data owner to switch to competing providers, or potential entrants hence creating new barriers to entry into these markets for possible competitors and innovators.

Biglaiser et al. (2019) identified personal data as a potential source of incumbency advantage, as these data feed into platforms' algorithms, enhancing their ability to match users across various platforms. Web mapping services serve as typical examples of these advantages, training their algorithms with geolocation data to offer superior services to other users. Similarly, search engines develop network centrality metrics based on user

queries to build meaningful rankings for search results and targeted advertising. Hence: “If a user has been a client of a platform for some time, the platform knows his or her tastes and can give more prominence to goods or services that he or she prefers. Second, the platform can use the data stemming from other users to increase the quality of the service to each of its users” (Biglaiser et al., 2019). As a result, entrants into this market will have to overcome high lock-in effects, which might result in reduced innovation and competition in the long term.

2. Policy issues

The assessment of the economic relevance of personal data requires a better understanding of how Generative AI systems collect, analyze, and integrate this data into digital business models, as well as awareness of the intended and unintended consequences of these processes. For example, the consent given to service providers to use one’s personal data and to agree to the use of cookies is often granted without much attention, while it can be a key element in shaping competition and entrenching market power in the digital markets.

A customer's telephone number was the key antecedent to today's personal digital data. Traditionally, a network of other users could easily reach its owner with this key personal identifier, generating positive network externalities (Katz and Shapiro, 1994). The larger the network, the greater the benefits. Hence, the loss of a personal telephone number implies the loss of immediate reachability and costly information to readvertise a changed number, with the consequent loss of benefits from communication due to a reduced network of immediate peers. To address these problems and facilitate entry into previously monopolized markets, telecom regulators introduced number portability in many countries. The ITU Data Hub (2022) indicates that nearly 40% of the world countries surveyed have introduced Fixed Number Portability (FNP), “a process by which customers may keep their fixed telephone number when changing either service provider, service or location or both.”

54% of the world's countries require mobile number portability (MNP): “a service that allows a mobile service customer to change telecom carrier and keep the same phone number.” Since losing a personal number would require a consumer to spend time alerting all her contacts, as well as several essential service providers (e.g., banking, insurance, and utilities), about the changed contact details, number portability was an effective tool in increasing switching activity, especially in mobile markets (Buehler, Dewenter, and Haucap, 2006). However, as ICT markets have converged into multiple digital platform markets and Generative AI systems now use a variety and type of personal data to provide enhanced services, number portability has become only a small part of the personal data consumers may need to transfer to maintain their original benefits when changing providers. Therefore, in these converged markets, where Generative AI systems generate their offerings, the focus should be on reducing the overall "switching costs" that consumers incur when switching providers. This will include considering the portability of a full set of personal data, not just number portability.

Some experiences with the regulatory imposition of wider personal data portability did not encourage switching activity in some service markets, such as banking. For instance, since 2014, the industry-run Current Account Switching Service in the UK has automated the switching of personal current accounts, enabling consumers to transfer all their recurring transaction arrangements, including outgoing (such as utility bills and mortgage repayments) and incoming (such as monthly salary). However, over the last decade, switching has increased in other service markets, such as general insurance (e.g., car and home insurance) and retail energy (gas and electricity), even in the absence of regulated data portability. Arguably, a significant distinction lies in the timing of users' decision-making processes: personal current accounts do not automatically prompt for regular renewal, necessitating new searches, while consumers must periodically renew their

insurance policies and energy providers once the fixed-rate promotional period expires (Thaler and Sunstein, 2021). However, while regular renewals might facilitate switching, other services, whose tariffs and features are constantly reshaped and personalized in real time by Generative AI systems based on the latest data fed in real time, can be cognitively too costly to be assessed by users, and reintroduce higher levels of attention, hence switching costs, similarly as for the non-regularly updated renewal services. Generative AI services can manage the matching between contract timing flexibility and the flexibility of the quality and price of the services provided. This is another factor that can affect how well or how poorly mandating data portability works to fix market power imbalances and network congestion.

Recently, some countries have mandated data portability to facilitate the comparison of complex tariffs based on a specific usage profile, thus lowering search costs. This is an additional dimension of personal data that, once aggregated with additional, non-personal data, such as tariffs or average usage, allows providers to supply tailored tariffs that use both personal and non-personal data, aggregated through specific algorithms and forecasting techniques. The UK Competition and Markets Authority launched the Open Banking Regulation, leveraging regulatory advancements in personal data portability and sharing to enhance tariff comparability and boost switching activity in the market for personal current accounts. Under this data portability remedy, the largest incumbent banks are required to adopt standardized application program interfaces (APIs) to allow seamless access to user data (with consent) by third-party apps. In a recent article, Deloitte (2023) explores the state of Open Banking across the world, identifying two alternative approaches: market driven experiences, among which Singapore, Japan, and the USA, and regulatory-driven ones, EU, UK, Hong Kong, and Australia. Interestingly, the report identifies the wider scope of Australia's Consumer Data Rights Act (CDR), which will

allow consumers to share their data with any authorized third parties, without being restricted to financial services, becoming instead an overall data policy that will apply to the energy and telecommunication sectors as well.

Open Banking's approach to data portability frequently serves as a model for regulatory intervention in broader digital platform markets, as noted by Gans (2018), Coyle (2019), and Scott Morton et al. (2019). In these markets, the presence of network effects often strengthens an incumbency advantage, both within the same user category (direct network effects, such as connecting with social peers) and across separate ones (indirect network effects under multi-sided platform competition, such as e-marketplace). In this sense, switching costs and network effects feed off each other to sustain the incumbency advantage. This is especially true when the same platform offers a bundle of personalized services that rely on the creation of a shared, detailed, and multifaceted digital profile of users' identities and individual preferences. This is often accomplished with the use of trackers that are largely run by a few Big Tech firms (OECD, 2020).

Similar to switching costs, network effects can also give rise to a first-mover advantage due to the belief that the challenger platform might fail to reach a viable scale. In these cases, data portability is not only aimed at lowering switching costs, but also at allowing entrants into the platform markets to match the quality of the incumbent’s Generative AI-based match-making service; that is, the ported data is used to improve the precision of its matching and predictive algorithms.

The key problem in defining the scope of relevant personal data portability is the fact that the attributes of a digital identity are not only a reflection of the original data inputs provided by the user but also the results of Generative AI inferences obtained from proprietary algorithms and statistical aggregation. For example, location services, browsing

histories, site reviews, dedicated advertising, and driving directions are all different tailored services based on Generative AI-based profiling relying on personal data gathered through tracking methods. Therefore, changing platform could entail a deterioration of the relevance of these personalized services. Arguably, this new type of 'lock-in effect' increases as the customer relationship with the platform in question continues. The EU Digital Markets Act (DMA) addresses the competitive value of personal data through algorithms and aggregation by prohibiting gatekeepers' platforms from amalgamating personal data from different services, preventing them from using data collected from third-party merchants for competitive practices, and requiring them to allow users to download apps from rival platforms.

Moreover, it is essential for policy considerations to note that digital platforms have different degrees of unavoidability in a networked market due to their betweenness centrality (D'Ignazio & Giovannetti, 2006 and 2014). These centralities are the key metrics to assess a provider's network relevance, as they capture more appropriately entrenched market power, than the traditionally used metrics of market shares, since these have a vague definition within complex digital ecosystem markets, whose boundaries are dynamically reshaped through the establishment of network links through Generative AI newly established relationships among the different ecosystem players.

Last but not least, not only gatekeepers network centralities vary in digital ecosystem markets, possibly leading to exosystemic tipping effects, but, on the demand side, the switching costs due to the profiling from Generative AI systems use of personal data can also be very different for different users. These different degrees of switching costs may reflect differences in knowledge, time availability, and cognitive and behavioral differences in users when dealing with complex choices across multi-dimensional, personalized contracts, services, and tariffs supplied through Generative AI systems.

Hence, Generative AI systems introduce an additional regulatory complication arising because they increase heterogeneity in switching costs both across users on one side of a digital platform, for example, sellers, across users on the other side, buyers, and in-between users within each one of the two sides. This level of complexity is shown by Giovannetti and Siciliani (2023), who created a singlehoming model to look at the incumbency advantage on two-sided platforms. In these models, agents have different switching costs and, more importantly, different ranges on each side of the platform. In this scenario, reducing switching costs for the most vulnerable users, who are more reliant on the current provider, could potentially hinder the entrant's ability to establish a sustainable presence in the market. This is because the incumbent naturally responds to the reduction in switching costs by setting lower prices, thus squeezing out the entrant. However, from a distributional perspective, users with high switching costs (i.e., those retained by the incumbent) always benefit from a reduction in switching costs.

3. Conclusions and recommendations

There is a lot of conflicting evidence and results about how using Generative AI systems for personal digital data affects the digital ecosystem markets. Because of this, the regulatory and competition authorities should look at all of these areas when they judge the effects of Generative AI systems collecting and analyzing personal digital data:

a. The beneficial impact on the visibility and digital identities of the data originators.
b. They bring systemic efficiencies to digital infrastructure management, planning, and forecasting.
c. And the economic impact on the competitive landscape, due to the economic rent that providers might derive from being able to access, process, transform, and use this digital data, whereby this might create new barriers to entry and competition into these

platform markets, for example, by allowing the use of Generative AI systems to implement data-based price and quality discrimination of services.

References

Biglaiser, G., Calvano, E. and Crémer, J. (2019), "Incumbency advantage and its value", *Journal of Economics & Management Strategy*, vol. 28, pp41-48.

Buehler, S., R. Dewenter, J. Haucap (2006), "Mobile number portability in Europe", Telecommunications Policy, 30, 385-399.

Coyle, D. (2018), "Practical competition policy implications of digital platforms", Antitrust Law Journal, 82, 835-860.

Davies, W. E. and Giovannetti, E. (2018). Signalling experience & reciprocity to temper asymmetric information in crowdfunding evidence from 10,000 projects. *Technological Forecasting and Social Change* Volume 133, August 2018, Pages 118-131

Davies, W. E. and Giovannetti, E. (2022). "Latent Network Capital and Gender in Crowdfunding: evidence from the Kiva platform". *Technological Forecasting and Social Change;* Volume 182, September 2022.

Deloitte Open Banking around the world Towards a cross-industry data sharing ecosystem, (2023) https://www2.deloitte.com/tw/en/pages/financial-services/articles/open-banking-around-the-world.html

D'Ignazio, A. and Giovannetti, E. (2006) "Antitrust Analysis for the Internet Upstream Market: A Border Gateway Protocol Approach" *Journal of Competition Law & Economics*, Volume 2, Issue 1, pp. 43-69.

D'Ignazio, A. and Giovannetti E. (2015) "Predicting Internet Commercial Connectivity Wars: the Impact of Trust and Operators Asymmetry" *International Journal of Forecasting* Volume 31, Issue 4, October–December 2015, Pages 1127–1137.

D'Ignazio A. and Giovannetti E. (2014) "Continental Differences in the Clusters of Integration: Empirical Evidence from the Digital Commodities Global Supply Chain Networks" *International Journal of Production Economics*, Volume 147-B, pp 486–497.

European Distributed Data Infrastructure for Energy, (2023), EDDIE https://eddie.energy/about Accessed on 9 October 2023

Gans, J. (2018), "Enhancing Competition with Data and Identity Portability", The Hamilton Project, Brookings, available at https://www.brookings.edu/wp-content/uploads/2018/06/ES_THP_20180611_Gans.pdf

Giovannetti, E. and Siciliani, P. (2020) "The Impact of Data Portability on Platform Competition" Antitrust Chronicle, Fall 2020, Volume 2, Number 2

Giovannetti, E., & Siciliani, P. (2023). Platform Competition and Incumbency Advantage under Heterogeneous Lock-in effects. *Information Economics and Policy*, 101031.

IEA (2017), Digitalisation and Energy, IEA, Paris https://www.iea.org/reports/digitalisation-and-energy

International Telecommunication Union, ITU DataHub, (2022) https://datahub.itu.int/. Accessed, 8 October 2023.

Katz, M.L. and Shapiro, C. (1994). "Systems competition and network effects". *Journal of economic perspectives*, 8(2), 93-115.

Klemperer, P. (1987). "Markets with consumer switching costs". *The Quarterly Journal of Economics*, 102(2), 375-394.

Llorca M., Soroush, G., Giovannetti E., Jamasb T., Davi-Arderius D. (2023). "Digitalisation and Economic Regulation in the Energy Sector", forthcoming in the Forsyningstilsynet (Danish Utility Regulator) Anthology on better regulation in the energy sector

Llorca, M., Soroush, G., Giovannetti, E., Jamasb, T., & Davi-Arderius, D. (2024). Energy Sector Digitalisation, Green Transition and Regulatory Trade-offs (No. 5-2024). Copenhagen Business School, Department of Economics.

OECD (2020), "Consumer Data Rights and Competition - Background note", available at https://one.oecd.org/document/DAF/COMP(2020)1/en/pdf

OECD (2021), *Data Portability, Interoperability and Digital Platform Competition*, OECD Competition Committee Discussion Paper, http://oe.cd/dpic

Rochet, J.C. and Tirole, J. (2003). "Platform competition in two-sided markets". *Journal of the European Economic Association*, 1(4), 990-1029.

Rossetto, N., Reif, V., (2021) Digitalization of the electricity infrastructure : a key enabler for the decarbonization and decentralization of the power sector, EUI RSC, 2021/47, Florence School of Regulation, [Electricity] - https://hdl.handle.net/1814/70736

Scott Morton, F., P. Bouvier, A. Ezrachi, B. Jullien, R. Katz, G. Kimmelman, A.D. Melamed and J. Morgenstern (2019), Committee for the Study of Digital Platforms - Market Structure and Antitrust Subcommittee, Report, George J. Stigler Center for the Study of the Economy and the State, available at https://www.judiciary.senate.gov/imo/media/doc/market-structure-report%20-15-may-2019.pdf

Thaler, R. H., & Sunstein, C. R. (2021). *Nudge: The final edition*. Yale University Press.