Santander X Innovation Xperts

Opportunities to Unlock the Value of urban data

DECEMBER 2022

By Santander

Table of contents

01 Introduction

This white paper aims to explore the opportunities deriving from urban data in the short and medium run (3 to 7 years' time horizon). By urban data we refer to a category of the broader data concept, defined as the data that provides information about a city that is proper to that city.[1]

The rapid adoption of new technologies – such as Internet of Things (IoT) devices, smartphones, high-resolution cameras, social media platforms or real-time geolocation – in cities is generating vast troves of data that can be used to improve city life through better mobility, energy use, air quality, etc. Unlocking that value will require the development of new products and services that are able to leverage that data through financially sustainable business models.

In section 2, the paper first explores the value of urban data and makes some conceptual clarifications around the notion of urban data and its lifecycle. The third section presents and analyzes a range of emerging products and services built around urban data. The fourth section lays out some of the main bottlenecks that currently exist to fully materialize the opportunities presented by urban data. This section also identifies some of the possible approaches to tackling these barriers, which can become business opportunities in and on themselves. The final section concludes with a succinct account of the main points made in the white paper.

Both the exponential growth of data generation and the increase in the market valuation of data analytics and management products point to an immense potential value deriving from urban data. Some of the emerging products and services presented in this paper describe the likely short to medium term scenario that will unfold in this sector.

1. Bruno Carballa Smichowski. The value of data: an analysis of closed-urban-data-based and open data- based business models. 2018. hal-01736484.

Despite these promises, unlocking the potential value of urban data will not be easy. The urban data landscape is fragmented, the regulation is nascent and therefore unclear, and there is a significant need to level up the investment on infrastructure, systems, skills, and capacities. These are not only challenges, but interesting opportunities that will require financing, sound policy decisions, strengthening of private and public sector capacities, and the establishment of effective collaborations across sectors.

02 Where is urban data today?

European data economy is estimated to reach €829 billion in 2025, representing 4 percent of the overall EU GDP.[2] In other regions such as Latin America, the big data and analytics market is forecast to reach $8.5 billion in 2023.[3] Open data, which is just a fraction of the total amount of data generated in an economy, has been estimated at a market value of €325 billion in the EU for the period between 2016 to 2020,[4] and between $3 to $5 trillion in economic value annually across seven sectors in the US.[5] Perhaps more important than these figures is the speed at which new data is created. The volume of data is growing exponentially (see figure 1), with some calculations estimating that 90 percent of existing data was generated in the previous two years.[6]

One of the main ways in which data can create value is through its use to improve and optimize current public and business processes, services, and practices. Several academic and practitioner analyses have found that the use of information technologies and data increases companies' productivity and competitive advantage.[7] At the same time, data itself has given way to the emergence of new business models, products, and services built on it.[8]

For example, many companies are already making money by selling the data that they collect regularly as part of their businesses to other companies, and almost every big company in the market is already buying data from other sources.[9]

---

2. See https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en#projected-figures-2025

3. See https://www.statista.com/statistics/993318/big-data-analytics-revenue-latin-america/

4. Berends, John, Wendy Carrara, and Cosmina Radu. Analytical Report 9: The Economic Benefits of Open Data. 2017. European Open Data Portal.

5. The World Bank. Starting an Open Data Initiative. January 24, 2019.

6. Bernard Marr. How much data do we create every day? Forbes, 21 May 2018, available at https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/

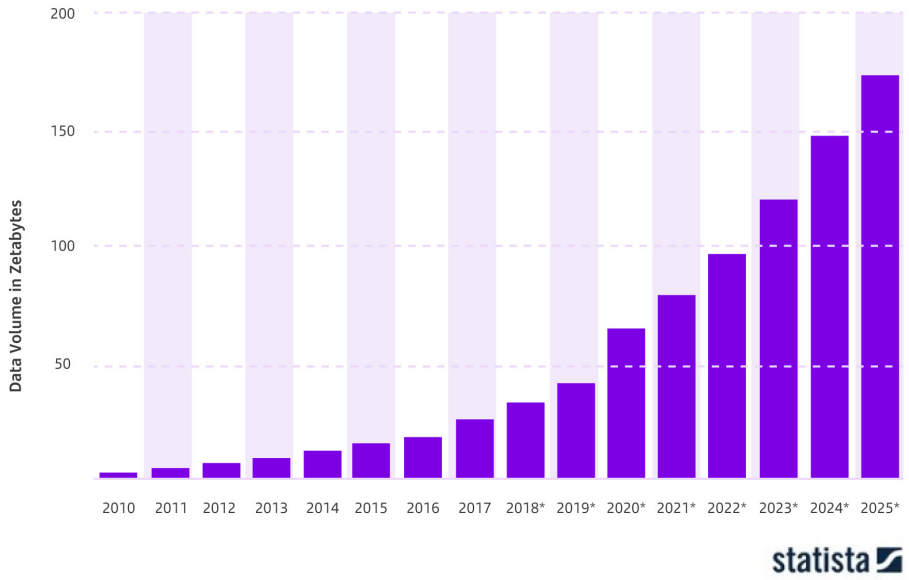
7. Erik Brynjolfsson and Kristina McElheran. The Rapid Adoption of Data-Driven Decision-Making. 2016. American Economic Review Vol. 106, No. 5 or James Manyika et al. Big Data: The Next Frontier of Innovation, Competition, and Productivity. 2011. McKinsey Global Institute, Amsterdam.

8. Fang Liang et al. A Survey on Big Data Market: Pricing, Trading and Protection. 2018. IEEE Access

9. Markus Spiekermann. Data Marketplaces: Trends and Monetisation of Data Goods. 2019. Intereconomics, 4.

Figure. 1

Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025 (in zettabytes)

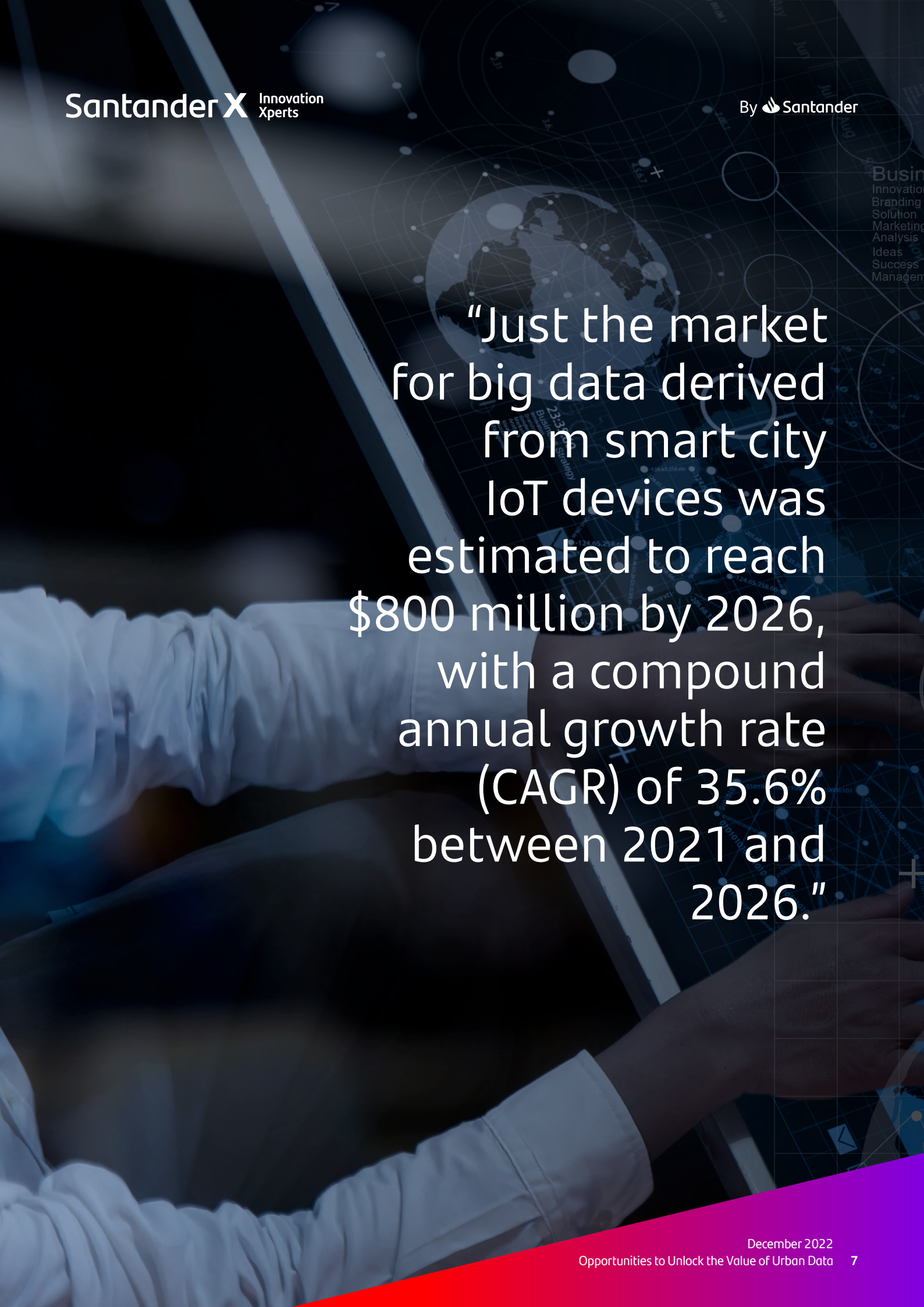


Interestingly, many of the opportunities created by data will materialize at the city level, where a large proportion of the data is being generated from IoT devices, public and private mobility operators, utilities and telcos, platforms operating in the accommodation, mobility, e-commerce or delivery sectors, the administrative data collected and increasingly opened up by governments, and all the data generated in smartphones and the web that can be tracked to physical places thanks to geolocation technologies. Just the market for big data derived from smart city IoT devices was estimated to reach $800 million by 2026, with a compound annual growth rate (CAGR) of 35.6% between 2021 and 2026.[10]

In addition, emerging technologies such as autonomous vehicles and drones largely operate on data-driven technologies, and at the same time, generate vast amounts of data that could be used to create additional value.

10. Industry Arc: https://www.industryarc.com/Report/17921/big-data-market-in-smarter-cities.html

Figure 1: Source. Statista (2002)



"Just the market for big data derived from smart city IoT devices was estimated to reach $800 million by 2026, with a compound annual growth rate (CAGR) of 35.6% between 2021 and 2026."

Despite these opportunities, the urban data market faces important bottlenecks to its long-term development: fragmentation, lack of common standards, data hoarding practices by different entities, and unclear regulations. According to an analysis by the World Economic Forum, of the 2.5 quintillion bytes of data that are generated daily, less than 1 percent is fully utilized due to these reasons.[11] Given the emerging nature of many of the economic opportunities presented by urban data, there is also a lack of clear and concrete understanding about what are some of the products, services, and revenue models emerging around it. Before analyzing some of these, however, it is important to make some clarifications around the concept of urban data and its lifecycle.

11. World Economic Forum. Unlocking the Shared Value of Smart City Data. June 2022. White Paper.

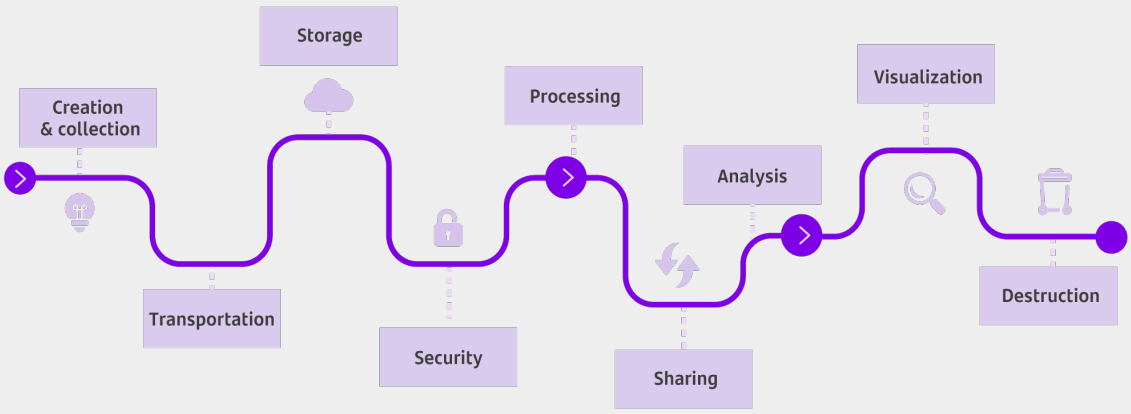
In focus

Urban data and its lifecycle

Data are usually defined as measures of the world that constitute the building blocks from which information, knowledge and wisdom are produced. Big data or urban data are specific types of the broader data concept (see Annex 1 for a more detailed discussion). The focus of this white paper is on urban data.

Figure. 2

Data Lifecycle



12. Rob Kitchin. The Data Revolution. 2021. Sage Publications.

Figure 2: Source. Adapted from Chignard & Benyayer, 2015.

We can distinguish several stages in the lifecycle of data: (1) creation and collection, (2) transportation, (3) storage, (4) security, (5) processing, (6) sharing, (7) analysis, (8) visualization, and (9) destruction.[13] There are opportunities to generate value – and therefore spaces for the creation of new businesses, products, and services – at each of these different stages. For example, some companies may create a new product by collecting data from novel sources which is then processed and sold or licensed to other organizations. Other companies may specialize in the stages of processing and analyzing data that is then used to develop a new product or service. The business model may also be centered around the exchange of data, for example by combining different data sources and selling new datasets or by creating and running a marketplace for data exchanges.

13. Adapted from Chignard, S., & Benyayer, L.-D. Datanomics. Les nouveaux business models des données. 2015. FYP editions.

03 Where is urban data headed?

This section explores some of the products and services focused on urban that are emerging and that have the potential to consolidate as the main business opportunities in a 3 to 7 years' time horizon. Some of these products and services are well-established but face important regulatory uncertainties (e.g., data brokers). Others have a less-proven business model but could become major sources of value generation (e.g., micro-mobility data pooling platforms). It is possible that some of these products will disappear or morph over the coming years, and that new ones not captured in this white paper may emerge. The picture provided in this section, however, contains some of the main traits of the urban data product and services' market in a short to medium term.

It is also important to note that a large part of the market potential arising from the use of technology in urban contexts is linked to investments in infrastructure, IoT devices, and other technologies, some of which are analyzed in detail in the other white papers of this series.[14] In this paper **we focus on those products or services that have urban data at the core of their value generation strategy.**

Building on the idea of the data lifecycle, we can distinguish several products and services built upon urban data:

Data provision

This refers to organizations who collect data to sell it, once processed, as a product or as a service. This product can be made of data originally collected by the data provider, or of a combination of such original data with other data collected by third parties (including open data).

14. For example, according to some estimates, the infrastructure segment accounts for 68.2% of the total market of the global smart cities market. See https://www.businesswire.com/news/home/20220622005660/en/Smart-Cities-Market-Research-Report-2022---2027-5G-AI-and-IoT-in-support-of-Smart-Cities-Market-Represents-Highest-ROI-Potential---ResearchAndMarkets.com

Sometimes this activity can be a byproduct of the company's main activity (e.g., telephone operating companies who sell anonymized geolocated data about their users). Data providers can make these data products or services available through data marketplaces or in bilateral exchanges with interested buyers or customers.

According to an analysis from the Boston Consulting Group,[15] nearly half of all smart-city applications use data from a variety of sources, so single – but particularly multi-source – data will become an important input for this market, and therefore the demand for these product and services will likely grow in the coming years. Another interesting aspect of this business derives from data's economies of scope. Because a single investment in data collection and processing can allow a variety of analytical and visualization applications, companies can design data collection in a way that enables multiple uses, generating potentially enormous productivity gains.[16]

Data brokerage

There are organizations that do not collect new data, but instead focus on gathering data from multiple sources to create novel datasets through their combination. Data brokers collect personal, locational, and transactional data about consumers that then sell – both as finished products or as services – for a variety of purposes such as identity verification, marketing, or fraud detection.

15. Massimo Russo and Tian Feng. The risks and rewards of data sharing for smart cities. 2020. Boston Consulting Group.

16. See Carballa Smichowski, 2018, footnote 1 supra.

"Because a single investment in data collection and processing can allow a variety of analytical and visualization applications, companies can design data collection in a way that enables multiple uses, generating potentially enormous productivity gains."

In focus

Data brokerage in focus

According to a report by the US Federal Trade Commission published in 2014, these products generated a combined total of approximately $426 million in annual revenue in 2012 for the nine data brokers that it investigated. Estimates about the data brokerage market value forecast it to reach $42 billion at a CAGR of 6.8% during the period from 2022 to 2031.[17] Although data brokers do not exclusively focus on urban data, local governments are an important source of data, as well as recurrent clients of data brokers.

Companies like Oracle, Nielsen and Salesforce lead an industry whose growth has only boomed since the Federal Trade Commission's investigation, and yet, many of the concerns raised by the US government and others about data brokers have still not been addressed.[18] This is an area where further and clearer regulation is needed to tackle the risks of a largely opaque industry while enabling the benefits that advanced data processing and exchange can provide in a data-rich economy.

17. See https://www.globenewswire.com/news-release/2022/08/01/2489563/0/en/Data-Brokers-Market-Estimated-to-Reach-US-462-4-billion-by-2031-TMR-Report.html

18. See Data brokers: regulators try to rein in the ´privacy deathstars´ from the Financial Times (7 January 2019) available at https://www.ft.com/content/f1590694-fe68-11e8-aebf-99e208d3e521

Data marketplace

A data marketplace is a digital platform on which data products are exchanged or traded. These platforms are neutral intermediaries where users can upload, sell, and buy data products. Data marketplaces are often run by organizations that establish the standards, licensing models, and terms of access and use. While some data marketplaces focus on facilitating transactions between two parties (e.g., IOTA, Streamr, and DAWEX) others also provide data analysis, visualization, and processing functionalities within the platform infrastructure (e.g., Telekom Data Intelligence Hub, Advaneo, and Caruso).[19]

19. See Spiekermann, 2019, footnote 9 supra.

Case in point

The Copenhagen City Data Exchange (CDE)

An interesting example of an urban data marketplace was the Copenhagen City Data Exchange (CDE) developed in 2015 by the Municipality of Copenhagen, the Capital Region of Denmark, and Hitachi. According to an evaluation of the initiative, the CDE "supported local companies in exploring the value of their data by identifying potential customers and acted as a data broker by looking into challenges and opportunities and by matching their needs with data sources from both the public and private sectors".[20]

In 2018, the CDE was discontinued because it was not able to reach a critical mass of datasets and data purchasers to make the model sustainable.[21] The evaluation carried out by the CDE partners after the project ended found some key obstacles for the adoption of the marketplace, such as the immature and fragmented nature of the data market, the need to develop clearer use cases, the reluctance to share data by some organizations and the lack of skills and competencies.

20. Municipality of Copenhagen and Capital Region of Denmark. City Data Exchange. Lessons from a Public/Private Data Collaboration. 2018.

21. See case study published by Greater Manchester available at https://greatermanchester-ca.gov.uk/media/3529/copenhagen-city-data-exchange.pdf

Looking to the future, a key success factor for data marketplaces will be the provision of value-adding data-related services by these platforms. The future development of data marketplaces will likely be shaped also by the adoption of smart contracting and distributed ledger technologies that can enable faster and more secure execution of transactions.[22]

Data platform

Urban data platforms are those that enable the integration, linking, and use of data from different sources to address urban issues such as traffic management, pollution control, or energy use. There are, of course, many software products that city governments can use to manage their data. Over the last years, however, there has been an increase in the emergence of platforms that integrate data from multiple sources and break organizational siloes to unlock the potential value of better data management in urban settings. Some of these platforms are sold as software products, but many others are licensed as Software as a Service (SaaS). Urban data platforms may or may not include analytic services as part of their features, but most often do, somewhat overlapping with the next category (analytics provider).

Some of the most innovative data platforms (e.g., DKSR-OUP[23]) are built using open-source code and specialize in operating IoT-based applications with real-time data. Social media is also enabling the gathering of real-time feedback that can be processed through artificial intelligence to better understand populations' concerns and opinions on issues, policies, and decisions. That is the product developed by companies such as CitiBeats, who processes large quantities of unstructured data through ethical artificial intelligence to gather actionable insights into the population's concerns.[24] Other companies, like Vianova, develop and operate platforms that enable the collaboration between city governments and mobility operators by integrating private and public sector data and providing data management and analytic services in the mobility sector.

22. See Spiekermann, 2019, footnote 9 supra.

23. The DKSR-OUP is an open-source urban data platform, with strengths operating IoT-based applications with real-time functionalities (such as environmentally sensitive traffic control). The DKSR-OUP focuses on minimal latency, meaning that the time between data acquisition and data provided is kept as low as possible, within a milliseconds range. In this way, sensor data can be processed, analyzed, and distributed in near real-time, and used to immediately respond to what is happening in the city. See https://www.dksr.city/en/the-dataplatform/

24. See https://www.citibeats.com

The most advanced features even enable the direct communication of geo-fenced policies from the city government to shared mobility operators, to ensure compliance with regulation in real time.[25]

Analytics provision

Many organizations operating in the urban data market collect data from multiple sources but do not to sell it as a product or provide a platform to another organization. Instead, they offer a service to an end user that is based on the analysis of such data. The service can be a recommendation of a travel route (e.g., Waze or CityMapper), the rating of a restaurant (e.g., Yelp or Trip Advisor), or information on real-time parking costs using garage occupancy data, current weather data or information on upcoming public events.[26] Most companies combine user data (e.g., geo-localization or route endpoints) with third-party data (e.g., open weather data). Mobility as a Service (MaaS)[27] companies fall under this category.

A second type of analytics' providers is formed by companies that provide tailored data analytics and visualization products or services, usually to governments or other companies. For example, a company named Aretian Urban Analytics and Design combines "data analytics, complexity science, and network theory-driven machine learning to produce high-resolution digital models of cities."[28] These models can be used to generate insights that can inform urban planning decisions such as infrastructure and building design.

A third category can be found in "digital twin technology", computer programs that use real-time 3D data to model urban environments or buildings and create simulations that can predict how a product or process will perform.[29]

25. See City of Stockholm. Data driven regulation of micromobility. A demonstration project with e-scooter providers in the City of Stockholm. February 2022.

26. See Russo and Feng, 2020, footnote 16 supra.

27. MaaS refers to a type of service that enables users to plan, book, and pay for multiple types of mobility services through a single integrated digital platform.

28. See https://www.aretian.com/

29. See https://www.govtech.com/sponsored/planning-urban-cities-smartly-with-digital-twins

Through augmented reality technologies, these platforms enable the monitoring and assessment of conditions and changes in the built environment.

This can be used to enable participatory planning but also to assess the potential impact of certain decisions. When combined with other technologies such as unmanned aerial vehicles, digital twins can be used to inspect areas or buildings that are difficult to access, for example in the aftermath of an earthquake.[30]

Data-driven coordination

One of the models of data-driven coordination is known as peer-to-peer urban platforms. It is used by companies such as AirBnB, Uber or Deliveroo, which operate in urban settings to leverage the large numbers of consumers and suppliers existing in cities through network effects,[31] and in turn, further deepen the density of these markets with a faster connection between them.[32] While some of these platforms are profitable, the financial sustainability of other companies has not been proven yet and is particularly vulnerable after many courts have ruled that they are not merely matchmaking-platforms but rather employers of drivers or riders.

30. Hoskere, V., Narazaki, Y., Spencer, B.F. Digital Twins as Testbeds for Vision-Based Post-earthquake Inspections of Buildings. In: Rizzo, P., Milazzo, A. (eds) European Workshop on Structural Health Monitoring. 2023. Lecture Notes in Civil Engineering, vol 254. Springer, Cham.

31. Nestor M. Davidson & John J. Infranca. The Sharing Economy as an Urban Phenomenon. 2016. Yale Law and Policy Review. 34:125.

32. Daniel E. Rauch & David Schleicher. Like Uber, but for Local Government Law: The Future of Local Regulation of the Sharing Economy. 2015. Ohio State Law Journal 76:901

Case in point

Trees as Infrastructure

A more recent, and therefore relatively less tested but promising, type of data-driven coordination platform can be found in those that are leveraging data and digital technologies to match investors (both institutional and private) with projects that can address urban challenges. An interesting example is Trees as Infrastructure[33] (TreesAI), a cloud-based platform that collects data from multiple sources to quantify and value the benefits—relating to carbon, water, health, energy, biodiversity, and the economy – of urban trees and forests. The platform then connects investors and urban residents who are willing to fund trees with the city governments or entities seeking to plant and maintain them.

Annex 2 contains a description of the main the revenue models deployed by the organizations who are innovating with the products and services described in this section.

33. See https://treesai.org/

04 Main bottlenecks and opportunities to unlock the value of urban data

Despite the potential value of urban data and the continuous emergence of new products and services, the full realization of their promises is hampered by important bottlenecks. In this section we highlight some of those barriers and suggest potential approaches to tackle them.

- **Fragmented urban data landscape**
  - **Opportunity 1:** Improve open data models
  - **Opportunity 2:** Increase private sector data sharing
  - **Opportunity 3:** More clear standards and technical interoperability
- **Need for better infrastructure and systems**
  - **Opportunity 4:** Public and private investment in infrastructure
  - **Opportunity 5:** Decentralized infrastructures
- **Unclear regulatory landscapes**
  - **Opportunity 6:** Clarify and strengthen the regulatory framework
- **Lack of clear skills and capacities**
  - **Opportunity 7:** Invest in data-related skills
  - **Opportunity 8:** Invest in capacities in organizations

Fragmented urban data landscape. Most data sit behind the walls of different organizations, who do not publish it openly nor share it with other entities. This limits the flow of data and the potential to unlock its value. Many organizations are wary about sharing their data because they think this can pose a risk to their competitive advantage or out of fear that such releases will result in a breach of their privacy protection obligations. With regards to public open data, often just publishing datasets is not enough to generate value. This failure to meet the expectations of the open data movement has sometimes resulted in the loss of interest by the actors involved in open data ecosystems. Finally, data exchange is often hampered by the lack of common standards or technical interoperability.

Opportunity 1: Improve open data models

The EU has recognized the importance of high-value datasets and has taken legislative measures to increase their publication. While this is important, the focus has been on releasing these public assets free of charge. In some cases, however, it makes sense to establish fees or charges for the use of such data. This can increase the incentive of public bodies to release their data in valuable formats, as well as finance the costs derived from the process of publishing data in a way that can generate value.

In focus

From Open Data to Data as a Service

Cities like London are moving from open data portals that operate as catalogues of datasets to creating a model of "data as a service". In this approach the focus is on sourcing data based on user needs and linking users to where data is held, including the sharing of non-open data.[34] Governments steward and provide sustainable and equitable access to data for data seekers inside and outside the public sector.[35]

When data is available on demand it enables the development of different products, such as using that data to provide up-to-the-minute flood warning information by an external developer.[36] Companies can specialize in creating these new products and services based on the 'on demand data' but can also be pivotal in helping city governments transition from static open data inventories to Data as a Service models.

34. See https://cities-today.com/why-london-is-pivoting-its-data-approach/

35. See https://public.digital/2022/03/14/data-as-a-service-in-a-city-context

36. See https://www1.nyc.gov/assets/cto/downloads/annual-reports/2021_impact_report.pdf

Opportunity 2: Increase private sector data sharing

A large part of the valuable urban data is collected, generated, and controlled by private entities such as mobility operators, accommodation platforms, or energy companies. Increasing the sharing of such data can enable the creation of new products and services, enable more competitive markets, and improve the regulation of the negative externalities of certain activities.[37] In some cases data sharing can be made mandatory, such as those foreseen by the Data Act proposed by the European Commission[38] or those increasingly being established by cities for micro-mobility operators. In most instances, however, data sharing can be enhanced by the promotion of data collaboratives[39] where all parties can benefit from the exchange and pooling of data.

Opportunity 3: More clear standards and technical interoperability

The development of clear and common standards is a fundamental tool to overcome data fragmentation. Several initiatives are underway that can play a key role in tackling this challenge. For example, the work on common European data spaces will be key in creating a secure and privacy-preserving infrastructure to pool, access, share, process and use data, while establishing clear and trustworthy data governance mechanisms.[40] Other initiatives have also been launched at the local level to establish common standards, in particular for shared mobility operators (see for example the Mobility Data Specification[41] developed in the US and now applied in several European cities, or the CDS-M mobility standard currently being developed by five Dutch cities[42]). As some of these examples show, developing solutions for better data flow and sharing is also a business opportunity in and on itself.

37. See the Guidance on private sector data sharing released by the European Commission available at https://digital-strategy.ec.europa.eu/en/policies/private-sector-data-sharing

38. See https://ec.europa.eu/commission/presscorner/detail/en/ip_22_1113

39. See the work by the Gov Lab on data collaboratives available at https://datacollaboratives.org/

40. See the European Commission Staff working group document on European Common Data Spaces available at: https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces

41. See https://www.openmobilityfoundation.org/about-mds/#:~:text=MDS%E2%80%94%E2%80%9CMobility%2Data%2Specification%E2%80%9D,scooter%20and%20bike%20share%20companies.

42. See available documentation at https://openresearch.amsterdam/en/overview/70020

Need for better infrastructure and systems. Creating a more thriving economy around urban data will require significant investments in the infrastructure and technological tools required for upgrading the collection, processing, sharing, analysis, and use of data at scale.

Opportunity 4: Public and private investment in infrastructure

The EU has proposed to invest €2 billion (and to crowd-in an additional €4 billion by private investment) in a European High Impact Project to develop data processing infrastructures, data sharing tools, architectures, and governance mechanisms for thriving data sharing and to federate energy-efficient and trustworthy cloud infrastructures and related services. The potential for investment by the private sector is particularly promising since the business of managing data - just one of the components of the data lifecycle – has already become one of the fastest growing areas in enterprise infrastructure investments in 2021 and has an estimated worth of over $70 billion.[43]

Opportunity 5: Decentralized infrastructures

New developments in distributed ledger and privacy enhancing technologies are fostering interesting innovations in decentralized infrastructures for the creation of trustworthy data ecosystems that enable the creation of value while ensuring individuals' control over their data. One of the early examples in this space was the DECODE project funded by the EU and piloted in Amsterdam and Barcelona.[44] More recent projects, like Solid, launched by Tim Berners Lee, are among the most cutting-edge and promising approaches in this space.[45]

Unclear regulatory landscape. Given the nascent nature of the urban data market, there is still a high degree of uncertainty on applicable regulations for many of the activities performed around data. Different jurisdictions are establishing their own norms resulting in further fragmentation and confusion. In other cases, regulation lags new technologies, products, or business models, resulting in regulatory vacuums. For example, there is no property regime for data, with actors deploying two main legal strategies to exert certain rights over data: strategies based on intellectual property laws and strategies combining general terms of use with trade secrets.[46]

---

43. See detailed report by Andreessen Horowitz available at https://future.com/data50/
44. See https://decodeproject.eu/
45. See https://solidproject.org/
46. See Carballa Smichowski, 2018, foonote 1 supra, for a discussion

"The business of managing data - just one of the components of the data lifecycle – has already become one of the fastest growing areas in enterprise infrastructure investments in 2021 and has an estimated worth of over $70 billion."

Opportunity 6: Clarify and strengthen the regulatory framework

A well-functioning urban data market needs clear legislation on data governance, access, sharing and use. Some of these regulations need to be established at the supra-national level, and the EU is already taking the lead through a set of important norms.[47] In addition, cities like Barcelona have taken action to establish measures on the ethical management of data.[48] This initiative by city governments is important, particularly in the face of legal vacuums. Yet, to avoid fragmentation, the coordination and development of common regulatory frameworks that can be replicated by other cities will be very important. In this space, supporting initiatives such as the UN's Cities Coalition for Digital Rights[49] will be particularly important.

Lack of clear skills and capacities. Gaps in skills and capacities are another significant bottleneck to the full unlocking of the value of urban data. First, there is a widely acknowledged gap in data-related skills among the population. A report from the UN has called for increasing global data literacy as a necessary condition to catalyze the "data revolution",[50] with a report by McKinsey estimating the data skills gap in millions of unfulfilled posts.[51] In addition, many businesses and governments lack the capacity (understood as the personnel and tools) needed to perform their critical roles in the urban data economy. In this regard, a recent flagship report by the World Bank published in 2021 notes that a shortage of skilled data scientists, statisticians, and economists across public data systems is a critical constraint, which is compounded by constraints in technology and infrastructure.[52]

47. See, for example, the General Data Protection Regulation (Regulation (EU) 2016/679), Regulation on the Free Flow of Non-personal Data (Regulation (EU) 2018/1807), Cybersecurity Act (Regulation (EU) 2019/881), Open Data Directive (Directive (EU) 2019/1024), and the Data Governance Act (Regulation (EU) 2020/0340) to be complemented by the Data Act in drafting process at the time of writing and the Proposal for a Regulation on AI currently under development).

48. See https://www.barcelona.cat/digitalstandards/en/data-management/0.1/objectives

49. See https://citiesfordigitalrights.org/

50. Data Revolution Group. A World That Counts: Mobilising the Data Revolution for Sustainable Development. . 2014. New York: United Nations Secretary General's Independent Expert Advisory Group.

51. See Manyika et al., 2011, footnote 7 supra.

52. The World Development Report 2021. Data for Better Lives. World Bank

Opportunity 7: Invest in data-related skills

Organizations need to invest in the skills related to finding, manipulating, managing, and interpreting data which will be ''an integral aspect of every business function and activity''.[53] In the EU, this means moving from 5.7 million data professionals in 2018 to 10.9 million data professionals in 2025.[54] This of course represents a huge investment opportunity in training, reskilling, and upskilling activities.

Opportunity 8: Invest in capacities in organizations

Meeting this need will require large investments in (i) staff training, (ii) proper staff remuneration, (ii) career incentives, and (iv) technology tools inside public and private organizations.

53. Harvard Business Review. Data is useless without the skills to analyze it. 2012.

54. See available figures at https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en#projected-figures-2025

05 Conclusion

Urban data offers huge potential, and it is reasonable to expect the consolidation of existing innovative products and services and the emergence of new ones in the short to medium term.

To unlock this value, substantial barriers need to be overcome. Regulation needs to be issued and harmonized, and standards and technology need to be developed to increase interoperability and data sharing. All this will enable increased secure data flows, which will in turn favor the consolidation of existing products and services and innovation with new ones around data provision, exchange, management, and analysis. These businesses will demand capital, and the broader market will require substantial investments in infrastructures, technologies, and skills.

The rapid growth on sectors such as data management already reveal the future that may unfold. If the bottlenecks described in this paper are successfully tackled – through a mix of public sector decisions and private sector actions – such future will likely be defined by the exponential growth in the urban data products and services' market in the 3 to 7 next years.

Authors

Fernando Fernandez-Monge,

Senior Associate with the Bloomberg Harvard City Leadership Initiative and a Ph.D. candidate at IIPP at the University College London.

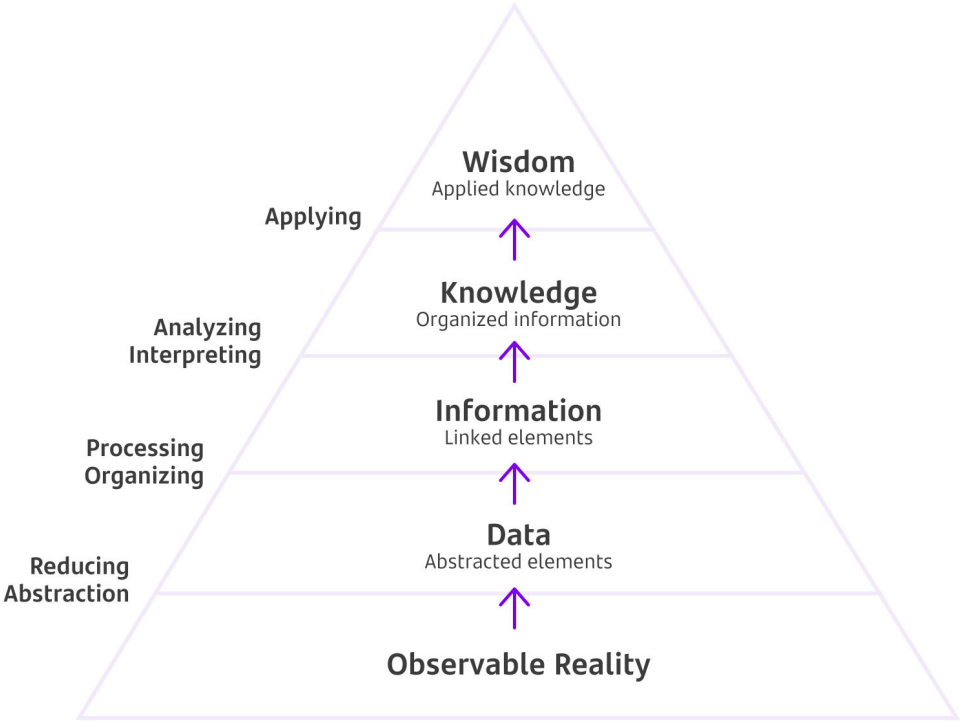
He teaches urban governance and innovation at the Harvard Extension School and the IE School of Global and Public Affairs in Madrid. He has written several teaching cases, book chapters and academic and policy papers on public innovation, data governance, urban governance, and city leadership and management.

Annex 1- Understanding urban data

Data are usually defined as measures of the world that constitute the building blocks from which information, knowledge and wisdom are produced.[55] This process of turning data into information, then knowledge and finally wisdom has been described using the image of a pyramid (see figure 3), where the transition from each stage to the next stage above is carried out by performing different actions (abstraction, processing, analysis, interpretation, and application).

Figure. 3

Data-Information-Knowledge-Wisdom pyramid



55. See Kitchin, 2021, footnote 12 supra.

Figure 3: Source: Kitchin (2021)

Data can be classified according to different criteria or characteristics. At a very basic level, we can group data according to the sector that they pertain to (e.g., water, mobility, energy, etc.). Data can also be grouped based on their form (structured, semi-structured or unstructured), type (indexical, attribute or metadata), and other criteria. For our purposes, four main characteristics are particularly important. First, whether data are historical or real-time (that is, data representing phenomena at the same instance as these phenomena occur). Second, whether data is controlled[56] by a public organization, a private organization, or a private individual. Third, whether data is considered personal or non-personal.[57] Fourth, whether data is open, shared (voluntarily or due to mandatory legal requirements) or closed.[58]

There are two additional definitions that we need to make. Although often confounded, not all data is big data. There is no single definition of big data, but it is often defined in relation to three characteristics known as the 3Vs: huge in Volume, high in Velocity, and diverse in Variety.[59] Other definitions include more attributes, and some definitions of big data incorporate the technological advances in computing, storage and analysis that have enabled the handling of massive amounts of data in real time.[60]Similarly, urban data is a subset of the broader data concept. We define urban data as that which provides information about a city that is proper to that city.[61]

56. There is no legal property regime over data, so is more accurate to say that data is controlled rather than owned. By control we refer to the ability to determine the conditions of access and use of the data.

57. The EU General Data Protection Regulation (GDPR) defines personal data as "any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person."

58. See the Data Spectrum developed by the Open Data Institute available at https://theodi.org/about-the-odi/the-data-spectrum/

59. Laney, D. 3D data management: Controlling data volume, velocity, and variety. 2001. Meta Group.

60. Crawford, K., Miltner, K., and Gray, M.L. Critiquing big data: Politics, ethics, epistemology. 2014. International Journal of Communication, 8: 1663-1672.

61. See Carballa Smichowski, 2018, footnote 1 supra.

Annex 2- Revenue models around urban data

Although there is no single definition of the concept of business model, most approaches are quite comprehensive in nature. For example, the framework proposed by Osterwalder and co-authors[62] contains nine building blocks of business models: key partners, key activities, key resources, value propositions, customer relationships, channels, customer segments, cost structure and revenue models. Analyzing each of those components for all emerging urban data products and services exceeds the scope of this article, so we will zoom in on the revenue models deployed by the organizations who are innovating with the products and services described in section 3:

1. Asset Sales:

This model is used, for example, when a data asset (e.g., a dataset or an analytical product) is sold to a customer. A few data platforms may also sell their platform as a software package and earn additional revenue by charging for maintenance and updates.

2. Usage Fees:

This model is often used by data marketplaces and data-driven coordination companies, who charge a fee for every transaction that they facilitate.

3. Licensing:

This is often used by data platforms, who license the use of their platform through SaaS arrangements.

4. Subscription Fees:

In some data marketplaces, subscription fees are also charged for membership, listing of data products, storage space or use of data services. A sub-category here are freemium models in which platform users can use basic functions at no cost but must pay a fee for an extended or full range of functions.

62. Osterwalder, A., Pigneur, Y., & Tucci, C. Clarifying Business Models: Origins, Present, and Future of the Concept. 2005. Communications of the Association for Information Systems, 16.

5. Advertising Fees:

Some of the analytic providers have a difficult time charging their users a fee (e.g., Yelp or Trip Advisor), so they often derive their revenue from selling visibility and targeted advertisement to other companies.[63]

6. Data-targeted complementary services:

Those data-driven coordination companies that collect continuous data from a user can gain precise insights about users' habits. These companies can then leverage these insights to sell complementary services to those users.

63. Sometimes they also sell the datasets that they generate from their users to third parties.

esade

Santander X Innovation Xperts

www.santander.com/santander-x-innovation-xperts-en

By Santander