TÉCNICO LISBOA

Prof. Margarida Catalão
Prof. Inês Carrilho Nunes

THE IMPACT OF BIG DATA ON MARKET STRUCTURES

ABSTRACT

In this report we take a look at how Big Data impacts market structures by firstly understanding the technical properties of Data, matching them with its Economic properties, and then looking at different examples from markets with heavily data-driven activities, such as eCommerce, social media and Search Engines. We close by looking at opportunities and challenges in the foreseeable future.

Group 14
Economics 2020/2021

Davis Gouveia (58240) | Miguel Marques (70376) | Sigurd Rimereit (101012) | Steinar Fjelldal (100921)

TABLE OF CONTENTS

1 INTRODUCTION

Data Production, Consumption and Distribution have been and continue to increase at astonishing rates [1]. Nowadays, we rely evermore on data to guide our decisions, from multinational giants' strategic decisions to everyday-life individual consumer choices.

As perfectly described by Reinsel et al (IDC) [1], this trend only shows signs of amplifying:

> *"The data-driven world will be always on, always tracking, always monitoring, always listening and always watching – because it will be always learning."*

This scenario opens up the opportunity - if not the necessity - for the emergence of a Data Economy, i.e., *"an ecosystem of different types of market players – such as manufacturers, researchers and infrastructure providers – collaborating to ensure that data is accessible and usable.",* as put by the European Commission [2].

Naturally, the devil is in the (not so small) details, and more than adding immediate opportunities, concerns and challenges for both Producers and Consumers across industries, the emergence of trends like Big Data brings serious, central implications to broader economic features like, for example, the market structure of existing industries/sectors.

The combination of the wide range of areas affected by Big Data, the potential implications this trend can have on entire markets and even economies, and the opportunity to explore core economic concepts in the real world, further understanding how they are affected/modified in the Big Data context, are the motivational pillars that led us to choose this topic for our Project.

We hope you enjoy reading it as much as we enjoyed building it - from research to discussion and writing.

2 BIG DATA

In this chapter, we take a look at the evolution of and properties of Big Data, from a technological perspective, and then explore how those properties translate into economic characteristics that make Big Data and the Data Economy such relevant matters nowadays.

2.1 WHAT IS BIG DATA?

Big Data is defined by Gartner [3] as:

> *"Big data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation."*

2.2 THE V'S OF BIG DATA

In 2001, Doug Laney introduced what would become the first 3Vs of Big Data[4]:

2.2.1 VOLUME

Possibly the most commonly associated feature with Big Data. It represents the massive amounts - ranging from tens of terabytes to hundreds of petabytes per organization [5] - of Data collected.

2.2.2 VELOCITY

Associated with the rate at which data is collected, processed, and sometimes acted upon, Big Data is characterised by its high velocity, implying real time - or at least near to real time - processing of the incoming data streams. Additionally, it is important to account also to the need for fast reactions, i.e., it's not only about the real-time collection and processing capacity in a steady pace, but also the ability to react promptly to the implications that each new data entry can have, as the relationships between existing data are also being affected in real-time.

2.2.3 VARIETY

Variety concerns the diversity of types of existing data. Big Data deals with high variety which means that data can come in many formats - structured (e.g., Excel, SQL), semi structured (e.g. JSON, XML, csv) or unstructured (e.g. rich media, medical records, emails).

Later on, this set got expanded, with the addition of two more "Vs"[6]:

2.2.4 VERACITY

With ever larger, faster paced and more diverse data sets, it becomes paramount to ensure that incoming data is accurate and trustworthy, as it becomes impractical (if not impossible) to systematically sanitize suboptimal inputs. Therefore, Big Data heavily relies also on the veracity of its sources.

2.2.5 VALUE

The most crucial feature, especially in business context. Big Data's purposefulness is directly connected to the value it is able to extract from its inputs. In other words, being able to collect and rapidly process enormous amounts of data, enables Big Data to extract unique value from its inputs.

Ever since, multiple approaches have built upon the 5Vs framework, attempting to exhaustively describe the features of Data, often increasing the “number of Vs''. Considering the scope of this report, we will be basing ourselves on the 5 V's framework, as we believe it is complete enough to support the economical properties of Big Data and the influence it has in today’s Data Economy.

2.3 ECONOMIC PROPERTIES OF DATA

Unsurprisingly, the technological properties of Big Data enable numerous business opportunities and, thus, it is worthwhile to explore how those technological properties translate into economic properties.

In a working paper from the UN [7], Cheng presented an interesting view of Data as a factor of Production - distinguishing it from labor, capital and land - due to its unique properties:

2.3.1 NON-RIVALRY

Data can be used by multiple entities and/or multiple times without exhausting it. In economic terms, this translates into the provision of any specific instance of data having a Marginal Cost tending to (if not reaching) zero.

A practical application of this property is the use of cookies by ecommerce websites, enabling them, for example, to provide a better customer experience to their visitors [8].

2.3.2 EXCLUDABILITY

An entity possessing a particular dataset has a degree of power to withhold it from third parties, i.e., preventing others from using it. The extent to which this is possible depends on several factors, like regulations (e.g., GDPR) or technology (e.g., encryption). It’s important to note that this property applies to data but not necessarily to the information it may contain, as is the classical example of consumer preferences (information) that can be inferred from a multitude of sources, such as web browsing patterns, purchase history or demographics (data).

Economically speaking, even partial excludability can have a significant impact on competition, in a similar fashion to patents, by granting the collector(s) of data some monopoly power.

2.3.3 INCREASING RETURNS TO SCALE

In data-driven activities, increasing the input of data (*ceteris paribus*) will yield a more than proportional increase in the output. This becomes particularly important in markets more reliant on data as a production factor where, together with the non-rivalry property (i.e., MC ≈ 0) and the significantly high fixed costs (due to high Volume, Velocity and Variety), the necessary conditions for a natural monopoly can be met.

This property is perhaps best illustrated with the cases of Facebook and Google and the way their data centered business models grant them enough competitive advantage to justify the argument that their respective industries might have become, indeed, natural monopolies [9].

2.3.4 LIMITED FUNGIBILITY

Each data set usually contains a unique information which makes it hard to substitute for another stream of data.

As an example, consumer preferences might be inferred from a multitude of data sources, as mentioned before (see Excludability), however, each individual piece of data contains different information and, therefore, is limitedly fungible.

2.3.5 TRADEABILITY

The technological nature of data makes it very easy (i.e., not costly) for it to be exchanged among different entities (from individuals to entire governments), at least in concept. On the other hand, tradeability also faces some barriers, mainly associated with other properties, such as Excludability (i.e., regulatory/technological factors might discourage or even forbid the exchange of data), Limited Fungibility or even its Uncertainty of Value.

2.3.6 EXPERIENCE GOOD

Data can be characterized as an experienced good given the difficulty of evaluating it before using it, i.e., like a meal at a restaurant, one can only truly know the value of a data set after experiencing it. At the same time, providing access to the data before the purchase would remove the incentive to pay for it afterwards.

2.3.7 DIFFICULT TO VALUATE

Given the potential to extract diverse insights from a single data set, depending on the context/purpose, makes it extremely difficult to precisely estimate Marginal Values from the production functions. On top of that, since data trading is not often associated with monetary transactions, it's nearly impossible to proxy the Value of Data through price discovery.

All in all, these properties of Data, amplified by the 5Vs can yield significant competitive advantage to the ones able to "harness" the potential of Big Data. Nevertheless, there are several trade-offs that cannot be overlooked, that may have significant impact on the markets, as will be discussed in the upcoming chapters of this report.

3 BIG DATA'S MARKET IMPACT

As the world is getting more and more digitized, another world of opportunities opens up. The retail market is shifting from face-to-face commerce to online shopping. When a customer enters a web page, he or she leaves huge amounts of data to the owner of the page. With the proper use of modern technology, such as artificial intelligence and machine learning, this data provides the suppliers valuable information about their customers.

In this section we will discuss how big data affects the market and in what way this leads to a surplus for the customers, the small firms and the big firms. This we will present by explaining the outcome by theory and then present some real-world cases to enlighten the topics.

3.1 RISING DEMAND

Before the e-commerce era, there was a salesman who met the customers face to face. This gave the salesman the opportunity to get to know the customers before trying to convince them to buy a good. With online shopping this aspect of the trade has vanished, but even though the online trading is behind anonymous profiles, the customers leave personal data behind. By collecting this data from each customer, the suppliers can learn about preferences and behavior of the customers. This brings the online retailers back to 1-on-1 trading.

As you have probably experienced, advertisements often fit what you have been talking about with your friends or searched for on the internet. This may seem like a coincidence, but it often has a connection with big data handling, and AI algorithms. Take Facebook for example, if a Facebook friend has searched for a specific good, Facebook would assume you are into the same good and hence advertise the same good to you [10]. This is an example of targeted advertisement.

There are other techniques for stimulating the demand as well. By obtaining historical sales information, such as typical goods bought together, retailers may suggest suitable package deals and tying offers. These offers may not only appear in the website as any other offer, they often pop up at a certain time. The supplier has information of how much time each customer stays on their website and sends the offer right before the customer is leaving the website.

Studies have shown that targeted advertisements have a better effect on the customer than general ads. Gilt Groupe states that based on historical user data, over 3000 different customer messages are created at noon every day. Their goal is to make their customers feel like, "Wow, every time I come to Gilt they know me" [11]. Amazon is able to take this to new heights, because it is both a marketplace for smaller retailers and a retailer itself. This puts Amazon in a superposition where it has potentially full insight in the market and access to all the data. Amazon has been charged by the EU commission for abusing third party data to enhance its own sales. This is an example of concentration of the market share.

3.2 SELLING LOWER QUALITY AT A HIGHER PRICE: THE CASE OF AMAZON

With all the available information available online, users probably think that they will do the best possible deal if they price check Amazon. This is not always true. You could be buying lower quality products for higher prices. Big data makes it possible. Amazon has been accused of using third party data to boost its own sales [12]. The European commission is investigating Amazon for using sensitive data from retailers who are selling on their marketplace [13], [14]. Amazon's function as a marketplace gives them a lot of data they could use as a retailer. Big data makes it possible for Amazon to see what businesses and segments that are profitable. This will affect the market power if it's correct. Amazon is then able to reach a higher market share. This could lead to a transition from monopolistic competition towards an oligopoly.

It is accused of using its algorithms to boost its own sales compared to its competitors. In these circumstances, Amazon could potentially use this to increase demand for their own product, leading customers to buy amazon's product instead. This makes Amazon able to compete in a higher price range than they originally would. Amazon's marginal costs are lower thus leading to increased profits.

3.3 PRICING STRATEGIES

Every firm is trying to maximize its profit, and pricing strategies are a very common way to do so.

Price discrimination has been an efficient strategy to increase the quantity of goods sold, as well as selling goods at a higher price. Examples of this are student prices for movie tickets, and airline prices. These are examples where the customers are well aware of the price differences. With more and more e-commerce, new technology allows for taking the price discrimination to the next level. With information regarding the duration the customers visit the web page, historical purchases, and which products the customers are spectating, the seller gets a detailed profile of each customer. With this information the seller may perform dynamic pricing, maximizing producer surplus.

There is another form of price discrimination that many are not aware of. Companies will discriminate based on what they think the customer is willing to pay for the product. This is based on behavioral profiling. Area of living, age, income, job category, etc. Low income and vulnerable people are the most affected by this. Poorer areas will usually have fewer local retailers than richer areas. This means that they have less opportunities to buy other places and retailers could increase prices. The data broker industry has a "suffering seniors" list. This is used for targeting advertisements on people with Alzheimer's or similar maladies [15].

Big data is also used to explore the demand curve in an efficient way. Offline businesses have for a long time practiced this, but with modern technology this is done in a much more efficient way than before. Studies within the topic have unveiled a sales experiment on eBay where an identical item was listed up several times with different auction parameters presumably to learn how the variables affect the demand [16], [17].

3.4 LOWER OPERATIONAL RISK

Big data has become a well known buzzword due to its powerful market influence. Suppliers are able to collect information about their customers based on their website activity. This may seem frightening and just a way for e-commerce giants to manipulate the customers. But big data has other applications as well.

By spectating the whole shopping process, from the entrance of the web page to the delivery of the package, suppliers are able to detect retail frauds by the usage of AI and machine learning algorithms. Amazon for example collects more than 2000 historical and real-time data points on every order, trying to detect frauds [16].

3.5 HOW BIG DATA MAKES A BETTER PRODUCT

Behavioral profiling could be used for several reasons. Categorizing who are most likely to get subprime mortgages is one. But most companies use it to make a better product for the customer. And it could be used to make more customized products. Streaming services have been improved by big data and the personal experience keeps getting better. Spotify makes a new list every week with songs they think you will like. This is fine when listening to music, but a little alarming how much they can know about each consumer. You think it's random that Spotify recommends a band that you just have been to a concert with? It is not, they likely collected data from your Facebook or Instagram [18].

Tying is an important part of big data in sales. Big data have made additional online sales easier for both small and big companies. Information about what products customers buy together are important for additional sales. When the company knows what products a typical customer will buy in addition to a product, it could recommend it to every customer who buys it. The company will sell more, and customers will recommend products that fit well with what they have bought.

Amazon buy box is a perfect example of this. More than 80 percent of Amazon’s sales revenue comes from the buy box. It’s crucial for sellers to get in the buy box. A high buy box percentage will lead to high sales numbers [19].

3.6 DOES BIG DATA LEAD TO NATURAL MONOPOLIES?

Google has had a stable market share between 85% to 90% of the total search engine market the last ten years. This is due to many different reasons. The large amount of data and how it is implemented is a key factor.

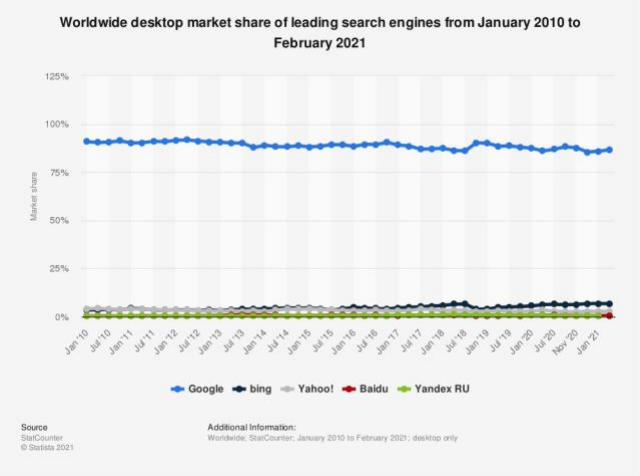

Figure 1 Market share of leading search engines between Jan 2010 and Feb 2021 (source: Statista [20]**)**

Google collects data from many different sources, maps, Gmail, android, search engine, etc. This makes it possible to create better personalized data than a lot of its competitors. Good data makes the user experience better. In 2008, Google processed 10 times more data every day than the total amount of information from all the US academic research libraries [21]. It processes more than 3,5 billion searches a day. The large amount of data needed creates high barriers to entry. Google used 13 billion dollars on new data centers and offices in 2019 alone. Additionally, the data centers need cooling and could use more than a million gallons of water a day [22].

In a similar situation, we have the widely known case of Facebook's acquisition of Instagram and WhatsApp, in 2012 and 2014 respectively. While the move itself was not appreciated by many regulatory authorities [23]- as it was considered anti-competitive behavior -, the increase in overall quality of the service offered hints at the possibility that, in these Data intensive markets, everyone might be better off by being served through a single service provider, making it more similar to a Utility instead [9].

It is hard to create a better product and experience without data. It's very unlikely that a new company will get the amount of data these Tech Giants have. The network effects of the big companies will make them better and better, and increasingly harder for others to compete with them, posing nearly insurmountable barriers of entry. The user experience gets better when more people use it. And when the user experience gets better, more people will use it.

4 FUTURE OPPORTUNITIES AND CHALLENGES

The amount of data created daily is increasing and there is not sign that growth will stop in the future. Originated from fonts from the internet, Internet of Things devices and sensors, they will contribute to a global big data market growth, which has already exceeded 50 billion dollars in size, according to Statista, and will exceed 100 billion dollars in size by 2027.

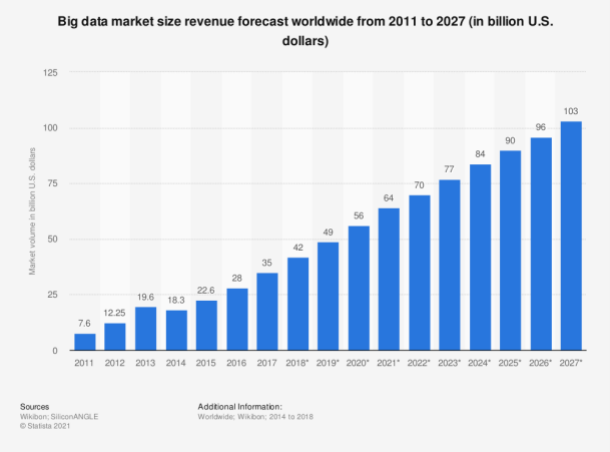

Figure 2 Big Data market size revenue forecast (source: Statista[24])

These predictions take into account several trends that will affect the future of big data. We decided to look at two of them: the rising of data volumes and data security and privacy.

4.1 INCREASE OF DATA VOLUMES

As has been happening until now, it's expected that the amount of data created will be growing exponentially in the future. IDC predicts, in their Data Age 2025 report for Seagate, that the global datasphere will grow from 33 Zettabytes (ZB) in 2018 to 175 by 2025 [1].

This rapid growth is based on one side on the increasing number of internet users, as well as an increasing number of unique mobile users and active social media users. In a society that already did more and more things online, from business communications, to shopping, to social networking, others areas evolved rapidly to this paradigm as online education.

Figure 3 Digital 2021 April Global Statshot Report (Source: DataReportal [25])

Adding to this, this growth comes also from the connection of billions of devices and embedded systems, that will generate, compile and distribute an abundance of IoT data constantly, everywhere.

IDC report predicts that enterprises will create and manage 60% of big data in the near future, which gives them the opportunity not only to provide the storing, but also to analyze massive volumes of data. But individuals will not have a passive role in this growth according to the report. It is estimated that 75% of the world's population, approximately 6 billion users, will be in some way interacting with online data every day by 2025. In another way to put things, each connected user will have at least one data interaction every 18 seconds [1].

This large amount of data presents several technical challenges when it comes to storage and processing. As many companies do not have the technical capabilities to manage this data internally, it led them to start transferring this data to the cloud. This led to a creation of a new market, currently dominated by Amazon Web Services, Microsoft Azure, and Google Cloud Platform, with their pay-as-you-go services, where companies can have cloud infrastructure with agility, scalability, and ease of use.

As this tendency will continue, some hybrid environments will be also used, where enterprises store most of their data in the cloud, but reserve sensitive information to their own infrastructure. Which led us to our second trend.

4.2 DATA SECURITY, PRIVACY AND OWNERSHIP

Privacy is a hot issue nowadays. Entities that create and manage big amounts of data are and will continue to be pressed to ensure data security and privacy. The protection of cyberattacks and intrusions will create more and more challenges for companies in the future.

As cyberattacks evolve, with threats that become more intricate by time, companies that ignore data security standards, even though governments are taking steps to regulate this area, and companies that do not have people trained in the area, there are reasons to worry.

According to Statista, for mid-sized companies, the average cyber loss in the last fiscal year amounted to $1.56 million, and $4.7 million across all company sizes, as of May 2019 [26].

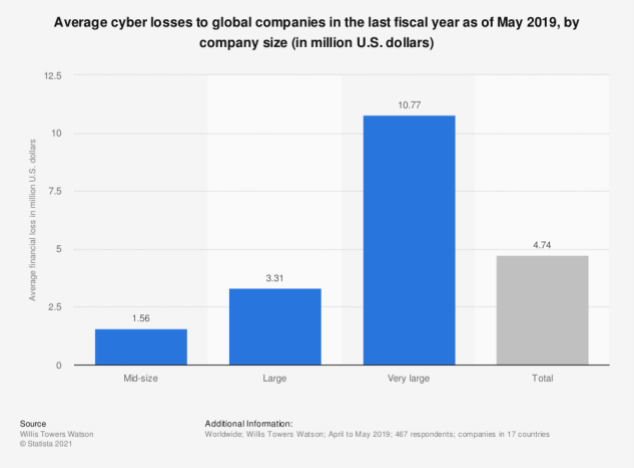

Figure 4 Online Incident financial losses in the fiscal year of 2018 per company size (Source: Statista [26]**)**

Beyond our own EU's General Data Protection Regulation, defined by the European Union as "the toughest privacy and security law in the world", many states in the US have created their own privacy protection laws, being the California Consumer Privacy Act one of the most known. These new laws made companies take data privacy more seriously, as severe fines must be paid for non-compliance.

These new laws also drew the attention of the topic to users, who began to demand accountability from the companies that manage their data. This concern for its reputation, and also to show transparency, has led more and more companies to permit a user-level control over data. In addition, investment in ensuring data privacy will be increasingly on the agenda of companies in the next few years.

Another topic that draws some controversy is data ownership. Notre Dame Cathedral recently suffered a tragedy that destroyed part of the famous monument. There is a public and political intention to rebuild the monument, which can be facilitated with the use of the technique of 3D scanning. Andrew Tallon, an art professor from Vassar College, was the first and last person to scan Notre Dame Cathedral digitally. The help of the collection of billions of data points can be crucial in this process, but who owns this data - Professor Andrew Tallon, Vassar College or Notre Dame? [27]

Also, who owns data we generated online after we die? There are no concrete rules about this issue today, in an area that technology has moved faster than the legal framework. The same dilemma exists on our data that is uploaded to the cloud. Is the data ours or the ownership passes to the virtual caretaker? In the same way as before, there is no international or national legal framework that regulates this issue, leaving the ownership definition to the individual terms and conditions that each individual company specifies.

The definition of the data ownership is an important issue for the security challenge discussed before, and designate who is responsible for guaranteeing it. Although in many cases the answer is not easy to solve, the growth of conflicts in the field of ownership will shape the future of the world of data, and the new regulations related to it.

4.3 BIG DATA AND THE RODRIK'S TRILEMMA

Big Data enables enhanced insights and better decision making. Conceptually, this added value increases with the amount of data, which, in turn, is often obtained from users. Therefore, it is not hard to make the argument that Big Data is more efficient (or even justifiable) in global markets, and these will, naturally, serve more customers, generate more data and yield higher revenue.

At the same time, we also saw that many of the industries that leverage on Big Data, specially the ones that rely almost exclusively on it (like Search Engines or Social Media), often operate best with a single producer, i.e. in a natural monopoly structure. If we combine this with the "globalised nature" of Big Data, it becomes extremely difficult to design effective policy and regulation to ensure efficient market functioning, or even define an entity accountable on such responsibility - the traditional approaches of governmental regulatory entities all limited by their own geographical borders, in a borderless market.

In contrast, we have increasing concerns with Data ownership and privacy on the consumer side. Data is the core of Big Data and in many (if not most) cases, that Data comes from the consumer. In such a scenario, consumers face a trade off between providing their data and the quality of the service.

In short, these trade-offs tamper directly with the famous Rodrik's Trilemma [28], where we have, on one side, Big Data's natural tendency for Globalization of markets and the increased value provided to consumers through it; on another side, we have the need to keep data monopolies in check [7] (while not necessarily breaking them up) and the inherent difficulty to do so with the existing regulatory authorities, usually bounded by their respective economic areas. On the third side, we have consumer protection and the associated Data security, privacy and ownership issues that affect, not only the (lack of) control consumers have over the use of the data they provide, but also the possibility that they'll be charged higher prices in return due to the excessive market control from the producers or even behaviour manipulation, as a result of the use of that same data [29].

5 CONCLUSION

5.1 LIMITATIONS OF THE PROJECT

At the outset of our work, we quickly realized that it would be extremely difficult for us to have access to quantitative sources on different markets. This happened mainly due to the fact that such sources are often not free (in fact, some reports can reach thousands of dollars in price) and, as a consequence, this report would have to focus mainly on qualitative perspectives and second-hand analyses found on publicly available publications.

Ironically enough, even limiting ourselves to "qualitative" sources, Big Data is such a broad area, with so many intricacies - from technological details, all the way to macroeconomic effects of its adoption - that we struggled to find a balance between extensive coverage and in-depth exploration of this topic. As a result, despite including the most crucial matters, in our opinion, relating to the scope of the Economics course, some aspects might've been left out and, therefore, this report should not be regarded as an exhaustive analysis of the impact of Big Data on market Structures, but rather as an extrapolation of topics lectured in the course towards real life applications/observations, in the context of Big Data.

Finally, deriving from the previous point, we found it more interesting to focus on Big Data in a cross-sectional fashion (i.e., looking at its effect across multiple markets); nevertheless, we focused on examples the more simply would describe the phenomena being addressed, often resulting in a concentration of examples coming from a narrow set of players and markets (Google, Amazon, Facebook, etc.). Even though this was done "by design", we reckon that more diverse examples might exist, and, thus, a portion of the "global picture" might've been left out in this work.

5.2 OUTLOOK

All in all, this project was a very rewarding learning experience, granting us the opportunity to explore up to date economic challenges and their technological counter parts (which is always interesting, specially from an engineering background perspective), consolidate the concepts lectured in class, during the semester and even deepen our understanding of their real-life applications. Furthermore, this exposure motivates us to keep ourselves updated on these matters (i.e., Economics in general, Big Data trends, Data policy, etc.), as well as further explore them, even after this semester.

6 REFERENCES

[1] D. Reinsel, J. Gantz, and J. Rydning, "The Digitization of the World From Edge to Core," 2018.

[2] European Commission, "'BUILDING A EUROPEAN DATA ECONOMY,'" Brussels, Jan. 2017. Accessed: May 14, 2021. [Online]. Available: https://digital-strategy.ec.europa.eu/en/library/communication-building-european-data-economy.

[3] Gartner, "Definition of Big Data - IT Glossary | Gartner." https://www.gartner.com/en/information-technology/glossary/big-data (accessed May 14, 2021).

[4] D. Laney, "3D Data Management: Controlling Data Volume, Velocity and Variety," *META Group Research Note*, 2001.

[5] Oracle, "What Is Big Data?" https://www.oracle.com/big-data/what-is-big-data/#three (accessed May 14, 2021).

[6] Y. Demchenko, P. Grosso, C. de Laat, and P. Membrey, "Addressing big data issues in Scientific Data Infrastructure," in *Proceedings of the 2013 International Conference on Collaboration Technologies and Systems, CTS 2013*, 2013, pp. 48–55, doi: 10.1109/CTS.2013.6567203.

[7] H. W. J. Cheng, "Economic Properties of Data and the Monopolistic Tendencies of Data Economy: Policies to Limit an Orwellian Possibility," United Nations, Jun. 2020. doi: 10.18356/9E71DB8C-EN.

[8] "What is a cookie and why is it important? | BigCommerce." https://www.bigcommerce.com/ecommerce-answers/what-cookie-and-why-it-important/ (accessed May 14, 2021).

[9] D. Ghosh, "Don't Break Up Facebook — Treat It Like a Utility," *Harvard Business Review*, May 30, 2019.

[10] T. Klosowski, "How Facebook Uses Your Data to Target Ads, Even Offline," *lifehacker*, Nov. 04, 2013. https://lifehacker.com/how-facebook-uses-your-data-to-target-ads-even-offline-5994380 (accessed May 14, 2021).

[11] A. Maybank, "Gilt Groupe: Using Big Data, mobile, and social media to reinvent shopping," in *Big Data, Analytics, and the Future of Marketing & Sales*, McKinsey&Company, 2015, pp. 72–74.

[12] J. Espinoza, "EU struggles to build antitrust case against Amazon," *Financial Times*, Brussels, Mar. 11, 2021.

[13] European Commission, "Antitrust: Commission sends Statement of Objections to Amazon for the use of non-public independent seller data and opens second investigation into its e-commerce business practices," Nov. 10, 2020. https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2077 (accessed May 14, 2021).

[14] European Commission, “Antitrust: EC opens formal investigation against Amazon,” Jul. 17, 2019. https://ec.europa.eu/commission/presscorner/detail/en/ip_19_4291 (accessed May 14, 2021).

[15] N. Newman, “How Big Data Enables Economic Harm to Consumers, Especially to Low-Income and Other Vulnerable Sectors of the Population.” Accessed: May 14, 2021. [Online]. Available: https://www.ftc.gov/system/files/documents/public_comments/2014/08/00015-92370.pdf.

[16] “BIG DATA ANDDIFFERENTIAL PRICING,” Feb. 2015.

[17] F. Tubbs, “The Impact of Big Data on Marketing,” *The Data Administration Newsletter*, Nov. 21, 2018. https://tdan.com/the-impact-of-big-data-on-marketing/24044 (accessed May 14, 2021).

[18] A. Slynchuk, “Big brother brands report: which companies track our personal data the most?,” *Clario*, May 05, 2021. https://clario.co/blog/which-company-uses-most-data/ (accessed May 14, 2021).

[19] P. Pant, “Understanding Customers with Big Data - The Amazon Way,” *Great Learning Blog*, Nov. 21, 2020. https://www.mygreatlearning.com/blog/understanding-customers-with-big-data-the-amazon-way/ (accessed May 14, 2021).

[20] “• Search engine market share worldwide | Statista,” Mar. 12, 2021. https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/ (accessed May 14, 2021).

[21] B. Marr, “How Much Data Is There In the World?” https://www.bernardmarr.com/default.asp?contentID=1846 (accessed May 14, 2021).

[22] N. Sattiraju and Bloomberg, “Secret Cost of Google’s Data Centers: Billions of Gallons of Water,” *Time*, Apr. 02, 2020. https://time.com/5814276/google-data-centers-water/ (accessed May 14, 2021).

[23] G. Canon, “Facebook faces antitrust allegations over deals for Instagram and WhatsApp | Facebook | The Guardian,” *The Guardian*, San Francisco, Dec. 08, 2020.

[24] “Global Big Data market size 2011-2027.” https://www.statista.com/statistics/254266/global-big-data-market-forecast/ (accessed May 14, 2021).

[25] S. Kemp, “Digital 2021 April Statshot Report,” Apr. 2021. Accessed: May 14, 2021. [Online]. Available: https://datareportal.com/reports/digital-2021-april-global-statshot.

[26] J. Johnson, “Online incident financial loss company size 2019,” *Statista*, Jan. 25, 2021. https://www.statista.com/statistics/387871/cyber-incident-loss-company-size/ (accessed May 14, 2021).

[27] “Andrew Tallon.” https://pages.vassar.edu/antallon/ (accessed May 14, 2021).

[28] D. Rodrik, “Dani Rodrik’s weblog: The inescapable trilemma of the world economy,” Jun. 27, 2007. https://rodrik.typepad.com/dani_rodriks_weblog/2007/06/the-inescapable.html (accessed May 14, 2021).

[29] I. Lapowsky, “How Cambridge Analytica Sparked the Great Privacy Awakening,” *Wired*, Mar. 17, 2019. https://www.wired.com/story/cambridge-analytica-facebook-privacy-awakening/ (accessed May 14, 2021).