PRIVACY E INNOVAZIONE

STUDI E RICERCHE SULLA PROTEZIONE
DEI DATI PERSONALI NELL'ERA DIGITALE

3

PRIVACY E INNOVAZIONE

STUDI E RICERCHE SULLA PROTEZIONE DEI DATI PERSONALI NELL'ERA DIGITALE

La collana ospita i risultati delle attività di studio e ricerca avviate o promosse dal Centro studi privacy e nuove tecnologie: monografie tematiche, paper e position paper che seguono l'evoluzione del tema della protezione dei dati personali in una prospettiva multidisciplinare (giuridica, politico–sociale, tecnologica) proponendo chiavi di lettura innovative. Sono inseriti in collana anche gli atti dei convegni organizzati dal Centro studi, gli interventi e i documenti presentati dai membri in occasione della partecipazione, in qualità di relatori, a conferenze e convegni. In "Privacy e Innovazione" sono pubblicate opere di alto livello scientifico, anche in lingua straniera, per facilitarne la diffusione internazionale.

Il direttore approva le opere e le sottopone alla revisione paritaria con il sistema del "doppio cieco" (*double blind peer review*) nel rispetto dell'anonimato sia dell'autore, sia dei due revisori che sceglie: l'uno da un elenco deliberato dal comitato scientifico, l'altro dallo stesso comitato in funzione di revisore interno. I revisori rivestono o devono aver rivestito la qualifica di professore universitario di prima fascia nelle università italiane o una qualifica equivalente nelle università straniere.

Ciascun revisore formulerà una delle seguenti valutazioni:

— pubblicabile senza modifiche;
— pubblicabile previo apporto di modifiche;
— da rivedere in maniera sostanziale;
— da rigettare;

tenendo conto della: a) significatività del tema nell'ambito disciplinare prescelto e originalità dell'opera; b) rilevanza scientifica nel panorama nazionale e internazionale; c) attenzione adeguata alla dottrina e all'apparato critico; d) adeguato aggiornamento normativo e giurisprudenziale; e) rigore metodologico; f) proprietà di linguaggio e fluidità del testo; g) uniformità dei criteri redazionali. Nel caso di giudizio discordante fra i due revisori, la decisione finale è assunta dal direttore, salvo casi particolari in cui questi provveda a nominare tempestivamente un terzo revisore a cui rimettere la valutazione dell'elaborato. Il termine per la valutazione non deve superare i venti giorni, decorsi i quali il direttore della collana, in assenza di osservazioni negative, ritiene approvata la proposta. Sono escluse dalla valutazione gli atti di convegno, le opere dei membri del comitato scientifico e le opere collettive di provenienza accademica. Il direttore, su sua responsabilità, può decidere di non assoggettare a revisione scritti pubblicati su invito o comunque di autori di particolare prestigio.

Elena Cirotti

Big Data: Excludability v. Access

A Comparative Overview of UE and US Competition, IP and Data Protection Laws

Preface by
Giorgio Resta

Aracne editrice

www.aracneeditrice.it

www.gioacchinoonoratieditore.it
info@gioacchinoonoratieditore.it

via Vittorio Veneto, 20
00020 Canterano (RM)
(06) 45551463

ISBN 978-88-255-3852-6

I[st] edition: December 2020

For those who saw in me something
that I'm not yet, but hope to become

To my parents

Table of contents

11 *Preface*
by Prof. Giorgio Resta

15 *Introduction*

21 Chapter I
The Pitfalls of Defining Data Markets I
1.1. Definition of Big Data, 23 – 1.2. Big Data in the Economy, 28 – 1.3. Defining the Relevant Market and Identifying Market Power, 31 – 1.4. Multi-sided Markets, 43 – 1.5. The Free Side of the Market, 50 – 1.6. Quality and Innovation, 54

69 Chapter II
Dominance in Data Markets II
2.1. Entry Barriers and Network Effects, 71 – 2.2. Data-Driven Mergers, 84 – 2.3. Monopolization and Dominance, 89 – 2.4. Dominance and Aggressive Market Strategies, 95

105 Chapter III
The Upsides and Downsides of Firms' Control over Data III
3.1. Database Right, 111 – 3.2. Copyright, 117 – 3.3. Trade Secrets, 121 – 3.4. Unfair Competition, 126 – 3.5. Private Tools: Contracts and Technological Measures, 129 – 3.6. Exclusivity v. Access, 131 – 3.7. Regulating Access to Data, 135

151 Chapter IV
Personal Data: Dignity or Commodification? IV
4.1. An Economic Interpretation of the Privacy Discourse, 154 – 4.2. A Fundamental Rights Discourse, 159 – 4.3. Data Protection across the Atlantic, 164 – 4.4. The Reality Gap, 181

185 *Conclusion*

191 *Bibliography*

Preface

Prof. Giorgio Resta[1]

Control over data is one of the main battlegrounds of this century. The ongoing datafication of everything, made it possible by digital technologies and the interconnection of the communication networks, is not only challenging traditional legal categories, such as the notion of "thing" (one might think at the impact of 3D printing technologies) or the notion of "territory" (the un-territoriality of data is the source of difficult jurisdictional issues). It is, more generally, raising critical issues in terms of power. As it happened in the last century with regard to oil, few multinational actors having data as their main assets are able to exert nowadays a tremendous influence over economics and politics. The Cambridge Analytics scandal is a clear demonstration of the capacity of data holders to strongly influence consumer behaviors and indirectly the workings of democratic processes. *Nihil sub sole novi*, one might object, recalling the objectionable uses of personal files made by dictatorships and sometimes even by democracies throughout the last century. Control over data, by the way, is an inherent component of the very substance of state power in modern times, as the history of statistics reminds us. Still, there is something unprecedented today, and this is the enormous ability of private powers to collect, process and use enormous amounts of data for commercial and non-commercial purposes. This is the source of economic and social power, that cannot be left unchecked in a democratic system.

[1] Full Professor of Comparative Law at University of Roma Tre Law Department.

Our legal systems have developed since the 1970's specific institutions aimed at controlling the collection and use of personal data mainly by public authorities. By contrast, the processing by private controllers has been regulated only recently and just by some legal systems, mainly located in Europe (other countries, such as the USA, have limited the scope of general data protection laws to the processing within the public sector). Also, very few provisions deal with the processing of non-personal data, such as data produced by sensors, IoT, etc. These have become increasingly valuable, especially for firms competing in the sector of automotive and smart cities, and as a result, controversies started to emerge with regard to the appropriation and exploitation of such assets.

All of this has transformed the law of data processing into one of the most important and controversial parts of contemporary law. This is one of the reasons why the analysis proposed by Elena Cirotti in this book — which is based on her Bachelor's degree thesis presented at the University of Roma Tre — deserves to be carefully considered. Differently from other works in the field, she takes a holistic perspective, with the aim of studying the impact of big data on the market, by considering the interplay between antitrust, intellectual property and data protection laws. The analysis is not limited to European law, but it takes also into account US law and other selected legal systems. The first part deals with antitrust and competition law, in order to illustrate whether and to what extent the techniques traditionally devised for curbing the abuses of market power can be fruitfully extended to data markets. Furthermore, the remedies offered by competition law have been recently resorted to by courts and administrative agencies (for instance, by the German Bundeskartellamt) as a mean to protect non-economic values, such as privacy and freedom of self-determination, and this has given rise to an heightened debate about the logic and limits of antitrust intervention. The second part of the book deals with data protection and sketches the most important features of the two Western cultures of privacy: the European dignitarian approach and the

American freedom–oriented perspective. The book is well written, carefully researched, and easily accessible also for non-specialists. Therefore, it can be praised as a significant contribution to the law of the digital society.

Introduction

Big Data refers to the collection and analysis of large amount of data, both structured and unstructured, in order to identify patterns and correlations. By combining such a variety and volume of data at a high speed, Big Data technologies have enabled to identify new insights, which were impossible to detect before. This revolution is due to two interrelated factors. Firstly, a relevant decrease in the costs of related technologies, such as cloud computing, has made it possible to obtain information «at nearly zero cost, virtually instantaneously and at a fine–grained level of disaggregation»[1]. Secondly, the widespread use of Internet and smart devices has boosted the near–time collection of billions of data. As a result, Big Data technologies have enabled something that even few years ago seemed unrealistic, given the poor availability of data and the expensive costs of processing.

Mayer–Schönberger and Cukier explain:

> Not only is the world awash with more information than ever before, but that information is growing faster. The change of scale has led to a change of state. The quantitative challenge has led to a qualitative one.[2]

Volume is the first characteristic to be emphasized about Big Data technologies. The initial idea was that the volume of data was so large that could not be contained in a memory of a computer. Thus, new processing technologies were designed, able to

[1] L. WU, E. BRYNJOLFSSON, *The Future of Prediction: How Google Searches Foreshadow Housing Prices and Sales*, A. Goldfarb, S. M. Greenstein and C. E. Tucker (eds.), *Economic Analysis of the Digital Economy, The University of Chicago Press,* 2015, 89.

[2] V. MAYER–SCHÖNBERGER, K. CUKIER, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, Mariner Books, 2014, 6.

process vast amounts of data. Additionally, data needed not anymore to be structured to be analyzed: unstructured data became valuable. Indeed, unstructured data — e.g., Facebook posts or tweets — allow multiple data combinations, supporting innovation and efficiency. Consequently, Internet companies started to employ these technologies to extract value from the millions of data at their disposal and became the leading users within the market.

At the same time, the large scale enabled things that were not possible at a smaller scale: to extract new insights or create new forms of value, in ways that revolutionized markets, our daily life, the relationship between citizens and governments, and more. Big Data Analytics, indeed, «is all about how these data can be stored, processed, and comprehended such than it can be used for predicting the future course of action with a great precision and acceptable time delay»[3].

Thus, Big Data is essentially about predictions. For instance, Wal–mart, the largest retailer in the world, analyzed the data collected about purchases and discovered that before a hurricane, people used to buy Pop–Tarts, a sugary snack, in addition to hurricane supplies. Consequently, before the hurricane, it started to sell the snacks close to hurricane supplies, thereby boosting its sales. Another often–cited example regards the American discount retailer Target. Target combined data available from its baby–shower registry to target pregnant women, since pregnancy is generally a moment of changes, including for shopping habits. By crawling the data, Target was able to detect the products that generally pregnant women purchase and to build a "pregnancy prediction" score. It was also able to estimate the due date so to send specific coupons for each stage of the pregnancy. These two examples show how the shift in the approach is consistent: it is not about causality anymore, rather about probability. Correlations, indeed, picture how often two phenomena are related, but

[3] A.K. BHADANI, D. JOTHIMANI, *Big Data: Challenges, Opportunities, and Realities*, Manoj Kumar Singh and Dileep Kumar G. (eds.), *Effective Data Managament And Opportunities For Implementation, Information Science Reference*, 2016, 2.

they do not explain why. A starting hypothesis, as in the classical scientific approach, is not needed. At the same time, Big Data discoveries should be carefully examined, given that the reliance of Big Data on probability arises risks of "false discoveries".

It is widely acknowledged that Big Data could be extremely beneficial by enabling more informed decisions by firms, governments and individuals. By knowing consumers' preferences, firms can create better products or services, or better target them with behavioral advertisements. Individuals, in turn, benefit from the use of innovative products and services. Even governments can use Big Data Analytics to allocate scarce resources and to predict disease outbreaks, climate impacts, job losses, etc. Yet, there are some Big Data downsides, as the possible harms to privacy or the potential use of data to discriminate certain categories of consumers.

Therefore, the scope of the present study is to analyze the Big Data impact on the market by considering the interplay between competition, intellectual property and data protection laws. It will examine the differences and similarities in approaching such phenomenon between the US and the EU jurisdictions, being the most flourish data markets.

First, Chapter I will first define what is Big Data and set the scenario within the analysis will be carried out. Indeed, the present study will focus on the use of Big Data technologies by private firms operating in the market. Then, the Chapter will specifically consider the key of the economic success of data firms, the main characteristics of such markets and their impact on competition. Specifically, after a general overview of the main tools of a competition analysis, the Chapter will consider their viability for Big Data markets. Indeed, Big Data markets have peculiar features that challenge a traditional approach. They generally show the structure of multi–sided markets, often with a free side so to attract more users and, hence, gain more value. Given their free side, they tend to compete on quality and innovation, rather than on price. This result into a more problematic antitrust analysis, since the latter has to focus on elements difficult to measure.

More in the antitrust natural territory, Chapter II will examine data markets entry barriers to understand whether they tip towards dominance or not. Having ascertained the concentrated nature of these markets, the Chapter will conclude by examining the possible anticompetitive conducts undertaken by dominant firms to shield their datasets, browsing through the most prominent EU and US cases in this regard.

Chapter III will then study the current tools used by firms to shield their datasets and whether this control has to be enhanced by creating a new property right or not. In particular, it will consider the current intellectual property protection — composed of database right, in the EU context, copyright, and trade secrets — that can be enjoyed by data firms. Yet, this protection does not cover any data nor any use of these data. Hence, the Chapter will examine how data firms enhance this level of protection by relying on unfair competition and on private tools, such as technological measures and contracts. Finally, the Chapter will focus on whether such wide control shall be promoted or reduced — the exclusivity v. access issue. Indeed, according to some commentators, this control shall be promoted by recognizing to data firms a property right on data, which will enhance businesses' incentive to invest and innovate. According to others, instead, data markets are already highly concentrated and a property right would only exacerbate the said situation. Embracing this second line of argument, the Chapter will conclude by examining the access theory, which sustains the importance for market participants to access unique datasets under certain circumstances so to compete viably.

Yet, mandating access to personal data could infringe upon data protection provisions. Hence, Chapter IV will consider the current EU and US data protection systems so to understand the relation between individuals and their data and its impact on competition. It will start by considering the current legislative framework and the relative market failures. Then, it will analyze the proposals to heal these market failures. In particular, it will examine the property right discourse as opposed to the funda-

mental right discourse. The first one relies on the idea that conferring a property right to individuals would enhance their control over their data and, thereby, their position in the market. On the other hand, the second one criticizes the resulting commodification of personal data, since data protection is a fundamental right, which cannot be confined into the market logic. To conclude, the Chapter will consider how data protection concerns cannot be remedied with the traditional solutions, which, given Big Data characteristic, suffer from a "reality gap" and urge the need to define new requirements to be complied with by data firms.