Web-Scraping as a Competition Law Offence

Thesis submitted in partial fulfilment of the requirements for the degree of LLM Law and Technology

Master Thesis
LLM Law and Technology
Tilburg University
The Netherlands
December 25, 2019

Supervisors: Inge Graef
Francisco Costa-Cabral

Student: Henrik Kiviniemi
SNR 2033648
ANR 766829

List of Abbreviations

EU	European Union
ToS/ToU	Terms of Service/Terms of Use
CFAA	Computer Fraud and Abuse Act
US	United States
TFEU	Treaty on the Functioning of the European Union
EC	European Commission
GDPR	General Data Protection Regulation
HR	Human Resource
CJEU	Court of Justice of the European Union

Table of Contents

Chapter One - Introduction

1.1 Background

Web scraping is a method that allows for the gathering of data from the Internet. Unlike human interpretation of browsers, scraping relies on machine to machine communication to retrieve data from a page. Through examined scripting and software production, web-scrapers can leverage loops, variables and conditions to obtain information from webpages effectively.[1] However, despite the evident power of web scraping, concerns about its legitimacy have somewhat shrouded its benefits in illegality.[2] The digital scene has tested legislators to appropriately balance technological resources with appropriate safeguards. Scraping, in particular, has produced concern in the fields of copyright, attracting the attention of courts worldwide.[3] It has even been stated that the method should be dealt with as a hacking offence[4], evoking criminal sanctions.

Companies may use automated web browsing products to collect web data for a broad diversity of uses. Examples from the industry include manufacturers tracking the performance ranking of goods in the search results of retail websites, companies observing data posted openly on social media to keep tabs on situations that require customer support, and companies staying up to date on news stories related to their industry across various origins.[5] E-commerce companies use automated web browsing to observe competitors pricing and inventory, and to aggregate data to help control supply chains.[6] Companies also use automated web browsers to obtain market information to help future plans.

Web-scraping is often seen negatively, and its reputation has gotten a lot worse in the past couple of years. Businesses increasingly use it to gain competitive advantage, and this means

---

[1] Mitchell, R. (2015). *Web Scraping with Python: Collecting Data From The Modern Web.* 1st ed. 1005 Gravenstein Highway North, Sebastopol, CA95472.: O'Reilly Media, Inc.,.

[2] F.Din, M. (2015). *Breaching and Entering: When Data ScrapingShould Be a Federal Computer Hacking Crime.* 81st ed. Brooklyn Law Review, pp.405-440.

[3] see, *hiQ Labs v. LinkedIn* [2017] (U.S. District)., *Field v. Google, Inc.* [2006] (United States District Court, D. Nevada.)., *Case C-30/14 Ryanair Ltd v PR Aviation BV* [2015] ECJ, *Dianping v Aibang* [2011] the Beijing Haidian District Court, *Ticketmaster Corp. v. Tickets.com, Inc.*, [2003] C.D. Cal.

[4] *Facebook, Inc. v. Power Ventures, Inc.* [2009] Northern Dist. of Cal.

[5] Data Use Cases, Gain a Competitive Edge With Web Scraped Product Data. (2019). [ebook] Scrapinghub Inc. Available at: https://cdn2.hubspot.net/hubfs/4367560/Case%20Studies/Data%20Use%20Case%20-%20Product%20Data.pdf. [Accessed 28 Aug. 2019].

[6] Zhao, Bo. (2017) *Web-Scraping*. Springer International Publishing AG (outside the USA)
L.A. Schintler, C.L. McNeely (eds.), Encyclopedia of Big Data.

that there is also a financial motive behind it. Not to mention that it is often in complete disregard of Terms of Service (ToS/ToU) or copyright laws. A large number of individuals and companies are running their own web-scrapers at the moment. This has been a pain in the flesh for those companies whose websites are being scraped, online stores like Amazon or social networks like Facebook.

Web-scraping of information inherently involves accessing a site that is hosted by another company. In some cases, the information that is web-scraped is private information of that site's user. Therefore web-scraping might not only give information about the private user to third parties, but might also give companies that web-scrape others all kind of information which can then be used for many different purposes — for example gaining knowledge of the rival companies website users or statistics that might give competitive benefit to the company that does the web-scraping. Even though some of this information was already given to the public and therefore one might argue that there is no infringement in copyright perspective, that does not mean that the actions in web-scraping process or the information that was gathered and then used were allowed by competition law.

So, do companies have any legal redress against web scrapers who extract information and data from the company's website? Valid examples can be found from Chinese courts which have supported this view and provided two routes for protection – namely, copyright infringement and unfair competition. In *Dianping v. Aibang*[7] the District Court approved that for restaurant reviews on Dianping.com, which are considered original in the sense that they show experience, emotion or personality, may be protected by copyright. Also in the case *Tencent v. Toutiao*[8], it was stated that news reviews which are considered to include authors personal comments are therefore to be regarded original and subject to copyright protection. However, there are still plenty of information and data which does not meet this criterion being subject to the protection of copyright. Therefore court continued its ruling in *Dianping v. Aibang* that customer reviews which are gathered and sorted on Dianping.com are the plaintiff's profits of labor which needs to be protected by the Anti-Unfair Competition Law since they have a high commercial value. It was also approved in the case of *Sina Weibo v. Mo Mo*[9] that a web platform has the right to collect users individual information, with its consent, for the purpose of its own business

[7] Vivien Chan & Co. Newsletter issue 03. (2018). [ebook] Available at: http://www.vcclawservices.com/sources/publications/2018issue03.pdf [Accessed 1 May 2019].
[8] Ibid.
[9] Ibid.

activities only if it is managed for commercial use or has commercial value. The web platform would have a valid claim if such information is 'scraped' without permission.

So, even though the scraped data itself does not always gain the protection of intellectual property, much of it is still valuable and innovative and can be used for a competitive purpose. For a situation where the usage of this data creates a scenario of anticompetitive means, there is a legitimate reason for this innovation, falling outside of the scope of intellectual property, to be protected by competition law. Businesses treat the Internet as one of their largest data source and the ability to compete in the digital economy will mainly depend on the ability to leverage that data. As Doug Laney, Vice President at Gartner Inc., said *"Your companies biggest database isn't your internal database, it's the Web itself."*[10]

1.2 Objectives and research questions

Legal issues to web-scraping have associated privacy claims and claims in the United States (US) associating the federal Computer Fraud and Abuse Act (CFAA)[11], in addition to competition claims concerning the need to collect public data to be able to compete freely. One of the big and recent legal decisions which involved web-scraping is *hiQ Labs v. LinkedIn*[12], from the US. The case indicates a shift in the means courts may be viewing efforts to limit data scraping, giving web-scrapers some leverage in their legal quiver to fight back against recent opinions about web-scraping. It also illustrates the significance of recognising general competition principles when acting in the range of big data, highlighting data as a competitive necessity in today's world.

Big data is a complex issue, different companies use it for different means from different sources. This means that the legality of some methods remains unclear. As big data's significance to the market continues to grow, it will become increasingly necessary to consider its effect on competition. Edith Ramirez from FTC[13] explained that *"there is no question that*

[10] Laney, D. (2015). *Gartner Predicts Three Big Data Trends for Business Intelligence*. [online] Forbes.com. Available at: https://www.forbes.com/sites/gartnergroup/2015/02/12/gartner-predicts-three-big-data-trends-for-business-intelligence/#6b3197fa6de4 [Accessed 11 Apr. 2019].
[11] Computer Fraud and Abuse Act 1986
[12] *hiQ Labs v. LinkedIn* [2017] (U.S. District).
[13] Federal Trade Commission

the aggregation of data may have important implications for competition."[14] Later on, Margrethe Vestager said that "*keep a close eye on how companies use data"* on here conference speech[15] on big data.

These claims were backed up by antitrust enforcement in 2017 when the European Commission laid to Google a 2.4 billion € fine for abusing its market dominance in Internet searches for unlawfully accelerate its own shopping service.[16]

The web-scraping issues proposed in the connection of the Google investigation afford a great illustration.[18] As a part of the investigation, the Commission hinted that managing third-party data without permission could amount to an abuse of a dominant position.[19] The Commission does not seem to debate that web-scraping is questionable because it could lead to foreclosure.[20] Alternatively, the Commission has explained that web-scraping could harm third parties motivation to innovate.[22]

This research analyses and tries to find a solution, if and to what extent, under what circumstances is web-scraping anti-competitive and therefore against EU competition law? Taking all the above into consideration, the main objectives of this thesis are:

---

[14] Ramirez, E. (2016). *Keynote Remarks of 43rd Annual Conference on International Antitrust Law and Policy, Fordham Competition Law Institute: Deconstructing the Antitrust Implications of Big Data.* Available at: https://www.ftc.gov/system/files/documents/public_statements/1000913/ramirez_fordham_speech_2016.pdf [Accessed 15 Apr. 2019].

[15] Vestager, M. (2016). In: *Conference on Big Data.* [online] Available at: https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/big-data-and-competition_en [Accessed 15 Apr. 2019].

[16] European Commission- press release (2017). *Antitrust: Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service.* [online] Available at: http://europa.eu/rapid/press-release_IP-17-1784_en.htm [Accessed 15 Apr. 2019].

[18] European Commission (2016). *Antitrust: Commission send Statement of Objections to Google on Android operating system and applications. European Commission Press Release.* [online] Available at: https://ec.europa.eu/commission/presscorner/detail/en/MEMO_16_1484 [Accessed 15 Apr. 2019].

[19] Ibáñez Colomo, Pablo (2016). *Restrictions on innovation in EU competition law*. European Law Review, 41 (2). pp. 201-219. ISSN 0307-5400 p.16

[20] European Commission (2012). *Statement of VP Almunia on the Google antitrust investigation*. [online] Available at: http://europa.eu/rapid/press-release_SPEECH-12-372_en.htm [Accessed 20 Dec. 2019].

[22] European Commission (2015). *Memo: Antitrust: Commission sends Statement of Objections to Google on comparison shopping service*. [online] Available at: https://ec.europa.eu/commission/presscorner/detail/en/MEMO_15_4781m [Accessed 16 Apr. 2019]. also see, where commission stated direct reduction of consumer choice as well as lower incentives to innovate, COMMISSION DECISION of relating to a proceeding under Article 82 of the EC Treaty and Article 54 of the EEA Agreement (COMP/C-3 /37.990 - Intel) para. 1616

1. *To expose and explore the possible available methods of web-scraping in a manner which they could fall within the scope of competition law and therefore generate a competition law offence.*
2. *To determine to what extent, if at all, the competition law should reach in situations where the scraped information at stake is not copyright protected intellectual property.*

The legal instruments essential for this analysis are Treaty on the Functioning of the European Union (TFEU)[23], the Enforcement Directive[24] among relevant regulations and guidelines for those instruments. Consequently, this research will answer the following research question:
To what extent should competition law address web-scraping considering the innovation falling outside of the scope of the protection of intellectual property?
Consequently, this means answering the following sub-questions:

1. *In the time of digital revolution, how to consider innovation in the application of competition law?*
2. *How is web-scraping regulated and why it is a competition law issue?*
3. *Should competition law create a second level of protection stricter than copyright law?*

1.3 Methodology

To answer these questions, doctrinal legal research has been decided to be as the primary method because a truly legitimate conclusion to this research question can only be found through a systematic analysis of legislation, case law, and literature. Doctrinal research is an especially reasonable approach to this.[25] To examine the relationship between web-scraping and competition law, the analysis of the legal framework is needed. Therefore TFEU, and Enforcement Directive is required. The legal framework is combined with a review of relevant information from websites of the European institutions, applicable regulations and guidelines for the above-said treaty and directive, as well as explaining relevant commentaries and literature. The other method used is an analysis of all the accessible literature dealing with web-

[23] Consolidated versions of the Treaty on European Union and the Treaty on the Functioning of the European Union (2012) C 326/01
[24] Enforcement Directive (2004) 48/EC
[25] Hutchinson, T. (2015). The Doctrinal Method: Incorporating Interdisciplinary Methods in Reforming the Law. 3 Erasmus Law Review 130.

scraping and its purpose, possibilities and legal issues. The approach of this thesis is the legal framework from the European perspective. There will be however, analysis of the situation of the legal framework in the US due to its much more "richer" case law. This will help to conclude the result of different situations as well as present a conclusion which answers to the objectives of this research.

1.4 Limitations and preliminary remarks

It should be noted that web-scraping does indeed address many areas of law, one of them being data privacy. However, the central question of this research is to consider that to what extent should competition law address web-scraping considering the innovation falling outside of the scope of the protection of intellectual property. Competition law is limited to a moderately narrow collection of economic matters. It is about finding balancing restraints of competition with countervailing economic efficiencies. Other concerns of public importance remain outside of its scope.[26] For this reason, and the limited length of this research, aspects relating to Data Privacy would only, therefore, be addressed when they are relevant in order to untangle the main questions and sub-questions or to give a better insight to the reader. For a more logical structure to the reader, international aspects such as case law and approaches to relevant issues will be addressed when they become topical through the research.

1.5 Structure

Based on the central research question and its sub-questions, the thesis will be structured in the following chapters. In chapter two author will answer to the first sub-question by explaining the importance of innovation in competition law and analysing the relationship between competition and innovation and why we commonly have so many problems to deal with innovation in competition law, despite the broad consensus that innovations are one of the significant benefits of market competition. Chapter three will start by distinguishing the importance of big data to companies since this is essential in order to understand the reasons for web-scraping. The author will then present and analyse what web-scraping is and how it is generally carried out and performed. The research will then carry on to analyse the benefits and threats of this practice to companies and how this could collide with EU Competition Law. The

[26] Lamarid, A. and Villiers, S. (2017). *Big Data Privacy and Competition Law: Do Competition Authorities know how to do it?*. [ebook] Available at: https://www.competitionpolicyinternational.com/wp-content/uploads/2017/01/CPI-Lamadrid-Villiers.pdf [Accessed 17 Apr. 2019].

author will also present a means of blocking hostile web-scraping or how companies are preventing it. After assessing the threat of web-scraping to the intellectual property, this chapter will try to answer to the second sub-question of how is web-scraping regulated and why it is a competition law issue? Chapter four examines and analyses firstly the data which is falling outside of the intellectual property and then distinguishes the difference between unfair competition and competition law in order to validate the arguments whether competition law could reach for help. In this chapter, the author also presents and discusses possible ways of forward and how the issues found could be tried to solve. Lastly, the conclusion of this research summarises the main findings of previous chapters. It will also answer the main research question as well as present some recommendations.

Chapter Two - How to consider innovation in the application of competition law?

Innovation plays an enormous role in creating value on the companies product or/and service and thrives for the competitive advantage. In this chapter, the author seeks to answer the question of how should we consider innovation in the application of competition law in the time of the digital era? In order to do so, we must first take ourselves into the benefits of innovation in the market competition before considering it more in-depth within competition law. To conclude this chapter, the author presents some of the problems arising with innovation in the hope that it will shed light on the other side of innovation, which is not so beneficial.

2.1 Competition law and innovation in market competition

Competition law enforcement deals with anti-competitive practices resulting from the acquisition or exercise of undue market power by companies that lead to consumer harm in the forms of higher prices, lower quality, limited options and lack of innovation.[27] In order to understand how competition law relates to innovation, it is useful to distinguish the three main branches of competition law and how does the innovation play a role in each of them. Firstly competition law prohibits agreements that restrict free trading and competition between companies.[28] This is associated often with issues related to cartels. Cartels reduce innovation,

[27] The benefit of competition policy for consumers. (2014). [online] United Nations Conference on Trade and Development. Available at: https://unctad.org/meetings/en/SessionalDocuments/ciclpd27_en.pdf [Accessed 16 Jun. 2019].

[28] Consolidated version of the Treaty on the Functioning of the European Union Article 101 (ex Article 81 TEC)

purely for the fact that by creating a so-called ”comfort zone” with their anti-competitive contracts restricting free trade and competition in the market, they have a much lower incentive to innovate or invest in it. Studies have shown that innovation and productivity drop notably when a member enters to the cartel.[29] Since cartels shelter its members from full exposure of the forces from the market, this reduces its pressure to innovate. Secondly, competition law restricts abusive behaviour by a company which already has a dominant position in the market or is conducting anti-competitive practices that will lead to a dominant position.[30] These abusive acts might be, for instance, tying[31] or predatory pricing[32] among many other practices. In a dominant position, a company might create barriers to entry[33] for the market, which tends to cause the existence of monopolies. A company holding a monopoly status in the market has a little to zero reasons to innovate since it already has a single control of a market. Lastly, competition law supervises the mergers and acquisitions of main companies.[34] Mergers that threaten the competition in the market can be prohibited completely or be subject to different obligations enforced by the competition authorities such as commitments to adjust the concentration in order to render it to be compatible with the market. Such commitments should be proportionate to the competition issue and eliminate it entirely.[35] Competition authorities in both Europe and the US, have raised their concerns on the aspect of innovation in mergers. Margrethe Vestager, European Unions (EU) competition commissioner, stated that we should not only look at the rise in prices with the high-tech mergers but also whether it has negative effects on innovation.[36] In the acquisition of GlaxoSmithKline’s business by Novartis, the European Commission (EC) recognised that there was a risk that Novartis might actually stop the development of two drugs that had the potential for an innovative treatment of skin and

[29] Report on the impact of hard core cartels and sanctions against cartels under national competition law. (2002). OECD. Available at: https://www.oecd.org/competition/cartels/1841891.pdf [Accessed 31 Oct. 2019].
[30] Consolidated version of the Treaty on the Functioning of the European Union Article 102 (ex Article 82 TEC)
[31] Practice of selling a product/service which creates a mandatory need of buying another product/service.
[32] Pricing strategy in which a product/service is sold very low with an intention of achieving new clients or achieving the foreclosure of other competitors and creating barriers to entry for potential new competitors.
[33] Barrier to entry is a fixed cost that has to be incurred by the new company, regardless of production or sales activities, into a market that current companies do not have nor had to incur.
[34] Council Regulation (EC) No 139/2004 of 20 January 2004 on the control of concentrations between undertakings (the EC Merger Regulation)
[35] Ibid. par.30
[36] Vestager, M. (2016). *Competition: The mother of invention*. speech delivered at the European Competition and Consumer Day. [online] Available at: https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/competition-mother-invention_en [Accessed 2 Nov. 2019].

ovarian cancer.[37] EC has also intervened and demanded remedies in order to preserve innovation in mergers as Bayer & Monsanto as well as Dow & DuPont.[38]

Perfectly competitive markets are said to meet economic efficiency, roughly meaning that nothing can be improved without something else being hurt. Encouraging economic efficiencies as part of the application of EU competition law echoes the neoclassical and neoliberal economic philosophy. As a primary importance in EU competition law, efficiency factors are connected together with the promotion of consumer welfare and adapted on consumers benefiting from them.

As Neelie Kroes stated in her speech at the European Consumer and Competition Day:

> "Competition policy serves one goal – to make sure that markets can operate as efficiently as possible to deliver these outcomes for our citizens. We want to help create a virtuous circle of economic growth and social welfare, in which the benefits are passed on within societies. [...] Consumer welfare is now well established as the standard the Commission applies when assessing mergers and infringements of the Treaty rules on cartels and monopolies. Our aim is simple: to protect competition in the market as a means of enhancing consumer welfare and ensuring an efficient allocation of resources. An effects-based approach, grounded in solid economics, ensures that citizens enjoy the benefits of a competitive, dynamic market economy."[39]

As such, they feed off the consumer welfare and wellbeing benchmarks. Such is the case in merger control where efficiencies may be examined, given they 'counteract the consequences on competition, and in particular the possible abuse to consumers.'[40]' `Efficiencies should be substantial and timely, and should, in particular, benefit consumers in those applicable markets

[37] EU Merger Control and Innovation. Competition Policy Brief. (2016). European Commission. Available at: https://ec.europa.eu/competition/publications/cpb/2016/2016_001_en.pdf [Accessed 2 Nov. 2019].
[38] Haucap, J., Stiebale, J. and Raschand, A. (2019). *How mergers affect innovation: theory and evidence from the pharmaceutical industry.* 63rd ed. p.284.
[39] Kroes, N. (2015). Delivering Better Markets and Better Choices'. Speech delivered at the European Consumer and Competition Day. European Competition Policy. Available at: https://goo.gl/3Fn75o [Accessed 24 Jun. 2019].
[40] Council Regulation (EC) No 139/2004 of 20 January 2004 on the control of concentrations between undertakings (the EC Merger Regulation) Recital 29

where it is otherwise possible that competition concerns would happen.´[41] Likewise, following Articles 101(3) and 102 TFEU[43] respectively, `consumers must receive a fair share of the resulting benefits,´ and as well, 'anti-competitive effects may be counterbalanced, or outweighed, by efficiencies which also benefit the customers.'[44]

Dynamic efficiencies have been the dynamic treatment in the context of the digital economy. A firm which is dynamically efficient will be reducing its cost curves by implementing new production processes. Such as innovation, which distinguishes several digital markets. Innovation processes spur dynamic markets, improve consumer welfare, and may help compensate otherwise diminishing marginal revenues. Not only as a crucial promoter of the competition but as well as for markets, innovation should be safeguarded and encouraged. Unquestionably, competition law has a major part to play in promoting competition in innovation [45] by encouraging the free market system, and by generating circumstances contributory to efficiency maximisation, market integrity, and competition on the merits.[47] A lot of process innovation is about reducing unit costs. This can be achieved by developing the ability of production to enable it to exploit economies of scale. Innovations also benefit companies in generating better quality products and a larger product range. Providing better products and a broader product range means consumers are more likely to find out that the company provides goods to their needs.

Competition law is not an end in itself, but an instrument intended to accomplish a certain public interest objective, consumer welfare. [48] Innovation was traditionally protected particularly under intellectual property laws, but this approach has increasingly reached its limit

[41] Guidelines on the assessment of horizontal mergers under the Council Regulation on the control of concentrations between undertakings (2004) C 31/03) para.79.

[43] Consolidated version of the Treaty on the Functioning of the European Union Article 101 (ex Article 81 TEC) and Consolidated version of the Treaty on the Functioning of the European Union Article 102 (ex Article 82 TEC)

[44] Ezrachi, A. (2019). *The Goals of EU Competition Law and the Digital Economy*. [ebook] BEUC Discussion Paper. Available at: https://www.beuc.eu/publications/beuc-x-2018-071_goals_of_eu_competition_law_and_digital_economy.pdf [Accessed 18 Dec. 2019]. See also: Case C-23/14 *Post Danmark* (2015) ECJ, electronic Reports of Cases, para 47; British Airways, para 69; Intel, para 140; Guidance Paper, para 30.

[45] European Commission, (2011). 'Guidelines on the applicability of Article 101 of the Treaty on the Functioning of the European Union to horizontal co-operation agreements'. OJ C 11/1, paras 119-122.

[47] Alford, R. (2018). *The Role of Antitrust in Promoting Innovation*. [online] Speech delivered at Kings College London. Available at: https://goo.gl/kcqtMQ [Accessed 19 Jun. 2019].

[48] Lowe, P. (2006). *Preserving and Promoting Competition : a European Response*. St Gallen Competition Law Forum. [online] European Commission. Available at: https://ec.europa.eu/competition/speeches/text/sp2006_018_en.pdf [Accessed 19 Jun. 2019].

when more innovation has to be spurred. Even though competition law, in its simplest aims to conduct that dishonestly obtains advantage of other competitors, and scholars have debated how innovation within the sphere of competition should be defined.[49] The OECD[50] has offered one of the most generally accepted explanation: ”innovation is changing the nature of the economic activity, creating new industries, expanding the scope for competition, and quickening product development”. [51] Generally accepted is also the notion that it truly remains one of the fundamental objectives of the primary market regulation to preserve innovation. By awarding exclusive rights to innovators, the primary market regulation preserves and spurs innovation in terms of its result. Such logic can be no better interpreted than by the famous saying of the ex-US president Abraham Lincoln that Intellectual Property law can: ”add the fuel of interest to the fire of genius, in the discovery and production of new and useful things”.[52]

2.2 Problems when dealing with innovation

Like in everything else, there is another side of the coin when dealing with innovation. A primary aspect of innovation is its unpredictability. Although it is logical to expect technology developments, it cannot be expected precisely if, when, and how innovation will take place. It can, therefore, be a burden for competition authorities to specify in a particular case whether changes in current market conditions are adequately likely to happen within the prediction phase. In a rapidly changing economic scene, the increase and development of the digital economy breed difficulties at competition enforcement. At the practical level, enforcers must confront the added complexity of managing their evaluations in a dynamic context. The shifting economic scene carries with it unavoidable uncertainty as to the nature of competitive tensions, the ability of markets to self-correct, possible harm, efficiencies, and disruptive innovation. Incentives to innovate ultimately depends on the capacity of the innovating company to earn additional earnings, which can be created from the innovation. This central concept is also a crucial explanation for most shielding measures such as patents, utility model protection, and copyrights.

---

[49] A. Baregheh, J. Rowley & S. Sambrook, (2009). *Towards a Multidisciplinary Definition of Innovation*. 47 Management Decision 1323.

[50] Organisation for Economic Co-operation and Development

[51] Report on Regulatory Reform Synthesis. (1997). OECD. Available at: https://www.oecd.org/gov/regulatory-policy/2391768.pdf [Accessed 18 Jun. 2019].

[52] A. Lincoln, (1953). *Lecture on Discoveries and Inventions*, in R. Baslter(ed) Collected Works of Abraham Lincoln, Vol 3, Rutgers. 363.

In the argument about the degree to which these kinds of shielding types should be available, this explanation is generally weighted toward the disadvantages of these possibilities. The implicit balancing among static and dynamic efficiency is as well an essential part of the debate on competition law perspectives.[53] In addition, for example, prices or the option and quality of goods and services, innovations act as a parameter of competition and market power, which is already acknowledged in competition law practice.[54] Therefore competition authorities must examine and weigh, for example, whether incentives and opportunities to innovate can be negatively affected by modifications to a market structure leading to more concentration or whether innovative competitors can challenge companies' powerful market positions.[55]

Some accuse competition law of privileging static over dynamic, , that is to say, merely keeping current markets competitive.[56] Grounds for this criticism can be explained by thinking that innovation may well happen outside, or create very diverse markets from, the market structure guarded by competition law.[57] Some scholars go on supporting that, to the degree of competition law's competence, innovation would be protected by preserving markets competitive structure.[58]

Despite the critique and different views, competition regulators have had a reason to interfere in several cases where innovation has played a role. As stated in the previous subchapter when distinguishing the three main branches of competition law. Competition law has facilitated

---

53 D. Encaoua & A. Hollander (2002). *Competition Policy and Innovation.* Oxford Review of Economic Policy. 18. page. 63-79.

54 Commission, Guidelines on the assessment of horizontal mergers, Official Journal C31/5, 05.02.2004, p. 5-14, paras 8. Guidelines on the assessment of non-horizontal mergers, Official Journal C265/7, 18.10.2008, p. 6-25, para. 10. OECD Competition Committee has also held many panel discussions on innovations and competition, also in the context of the Digital Economy, Oecd.org. (2019). *Digital Economy, Innovation and Competition - OECD*. [online] Available at: http://www.oecd.org/competition/digital-economy-innovation-and-competition.htm [Accessed 23 Jun. 2019].

55 Bundeskartellamt (2017). *Innovations-Challenges for competition law practice*. [ebook] Available at: https://www.bundeskartellamt.de/SharedDocs/Publikation/EN/Schriftenreihe_Digitales_II.pdf?__blob=publicationFile&v=3. [Accessed 14 Jun. 2019].

56 Costa-Cabral, F. (2017) *Innovation in EU competition law: The resource-based view and disruption.* Jean Monnet Working Paper 2/17 Available https://jeanmonnetprogram.org/wp-content/uploads/JMWP-02-Costa-Cabral.pdf Accessed 14 Oct. 2019. See also, Wu, T. (2012) *Taking Innovation Seriously: Antitrust Enforcement if Innovation Mattered Most*, 78 Antitrust Law Journal 328 page 313-328. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2166525. Accessed 25 June 2019.

57 Ibid.

58 Costa-Cabral, F. (2017) *Innovation in EU competition law: The resource-based view and disruption.* Jean Monnet Working Paper 2/17 Available https://jeanmonnetprogram.org/wp-content/uploads/JMWP-02-Costa-Cabral.pdf Accessed 14 Oct. 2019. See also Colomo, I. (2016). *Restrictions on Innovation in EU Competition Law*, 41 European Law Review. Available at https://www.lse.ac.uk/collections/law/wps/WPS2015-22_Colomo.pdf. [Accessed 26 June 2019]

remarkable discussions over innovation. National competition authorities and the Commission hold a fundamental part in determining the pace of innovation, This can be seen from the extremely hefty fine of € 2.42 billion against Google for favouring its own services in search engine results,[59] and later on in the merger between Dow & Dupont.[60] When co-operations and mergers are to be assessed, often there is a conflict of interest within the short term negative effects, like for example price increases and craved medium to long term effects such as the appearance of incentives to innovate. In every particular case, it has to be examined before the decision is taken whether it is needed to step in or, in the other hand, whether the likelihood of a dominant position can be rejected.

The difficulty for enforcement in the digital age concerns the challenges in understanding dynamic changes. In the digital world, methodological shortcomings may undermine one's ability to distinguish the consequences of specific actions on innovation. When considering the characteristic of dynamic efficiencies as well as scepticism surrounding disruptive innovation, whether competition law can contribute as a useful tool to secure competition for future innovation markets stays uncertain.[61] What makes things more complicated is the possibility to execute separation among pro-consumer and negative innovation. [62] In a digitalised environment, the separation among research and development that supports and benefits the consumer, it is true that from innovation that is used to generate exploitative technology or harmful exclusionary effects becomes clouded.[63]

The unpredictable characteristics of innovation call for circumspect intervention. The extent of markets and goods, existing and possible competition, the characteristics of competition, as well as the likely future actors, may change with new waves of innovation. This dynamism

[59] EC Press release, (2017). *Antitrust: Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service*. Available at http://europa.eu/rapid/press-release_IP-17-1784_en.htm. [Accessed 26 June 2019].
[60] EC Press release, (2017). *Mergers: Commission clears merger between Dow and DuPont, subject to conditions.* Available at http://europa.eu/rapid/press-release_IP-17-772_en.htm. [Accessed 26 June 2019].
[61] Generally speaking, modern economic literature suggests greater alignment with the Arrowian premise for most industries, which expects competition to stimulate innovative activity: Drexl, J. (2012). *Anti-competitive stumbling stones on the way to a cleaner world: protecting competition in innovation without a market.* Journal of Competition Law and Economics, Vol. 8, Issue 3. Page. 507-542.
[62] Schrepel, T. (2017). *Predatory Innovation: The Definite Need for Legal Recognition.* 21 SMU Sci. & Tech. L.Rev. 19.
[63] Ezrachi, A. (2018). *EU Competition Law Goals and The Digital Economy.* Available at https://www.academia.edu/36790134/Ezrachi_-_Goals_and_the_digital_economy_-_Working_paper.pdf. [Accessed 12 Jun. 2019].

promotes the protection of the competitive means as an independent assessment, for the sake of innovation and future efficiencies.[64] Also, what comes in the merger review, the chance that ex-ante intervention may dampen innovation has caused few to request for a more laissez-faire approach.[65] However, risks linked with major networks, platforms, data pools and their influence on competing innovators, neighbouring markets, market-entry, exclusion of possible competition as well as the tipping of the market in favour of the merged entity, have influenced others to call for higher scrutiny. Competition law can provide relevant markets a tool to allow for innovations to take place and as a means to guard against particular excesses of innovation. Still, while it is clear that the competition law can play a great role in the innovative process, there is an unanswered questions in economic studies regarding whether competition law could or should include directly or indirectly innovation-structured arguments into its analytical framework.[66]

2.3 Innovation in the digital era

As we have seen, the digital era has created its own challenges when competition authorities assess whether certain action could have an impact on innovation. Innovation has a crucial role in today's modern economic policymaking. When trying to mould innovative market conditions, competition law can intervene in multiple ways. Even though it is widely believed that market economies designated by undistorted competition contribute directly to innovation,[67] there are still unresolved questions between the connection of innovation and competition law. By answering the question of how we should consider innovation in the time of the digital revolution, we must first discuss some of the traditional theories by Arrow, Schumpeter and Aghion.

Generally, discussion on the connection with innovation and competition markets start from the approaches with Schumpeterian and Arrowian methods. According to the teachings of

[64] Larouche, P. and Schinkel, M. (2013). *Continental Drift in the Treatment of Dominant Firms: Article 102 TFEU in Contrast to § 2 Sherman Act.* TILEC Discussion Paper No. 2013-020 Available at https://ssrn.com/abstract=2293141. [Accessed 19 Jun. 2019].

[65] Alford, R. (2018). *The Role of Antitrust in Promoting Innovation*. [online] Speech delivered at Kings College London. Available at: https://goo.gl/kcqtMQ [Accessed 19 Jun. 2019].

[66] Nihoul, P. & Van Cleynenbreugel, P. (2018) *The roles of Innovation in Competition Law Analysis*. ASCOLA Competition Law series. Edward Elgar.

[67] G. M. Swann, P. (2009). *The Economics of Innovation: An introduction*. Cheltenham, Edward Elgar Publishing, page. 8–20.

Joseph Schumpeter, product innovation is guaranteed best by less competition than more.[68] In line with his ideas on creative destruction and the replacement of old products with entirely novel ones at some point in time, Schumpeter debates that less competition affords more incentives to a company to engage in product innovation, as the incumbent company may be worried that its products will be substituted by a more novel and better product that would result in the creation of a new market and the destruction of the old product market.[69] As the European Commission summarised Schumpeter's views in 2016:

> *"Less competition increases the post-innovation rewards for the innovator, which in turn will increase the incentives to engage in research and development. Even if there is little price competition in the market, innovation competition from firms seeking to take over the leading supplier's role (competition for the market) will goad the current market leader to invest in innovation to stay ahead, or else lose its market position to rivals."*[70]

In that sense, product competition, or at least the risk of it, spurs innovation.

Kenneth Arrow, however, argues that intense price competition rather than product competition works as a means to spur innovation. According to his thoughts, the more pressure companies feel concerning their pricing decisions, because competitors are also trying to capture demand, the more incentives to innovate will remain with the companies concerned.[71] As was stated by the European Commission:

> *"In a competitive environment, a newly invented product will not cannibalise the firm's own profit as much as it would under a less competitive market structure. In a competitive market, an invention will allow the inventor to gain sales from competitors and will therefore be applied to a higher output. Innovation incentives depend not on post-innovation profits per se but on the difference between post-innovation and pre-innovation rents. For these reasons, less competition in the market would reduce the incentives to innovate."*[72]

---

[68] Schumpeter, J (2010). *Capitalism, Socialism and Democracy.* Routledge, Abingdon, page. 82–83

[69] Ibid.

[70] EU Merger Control and Innovation. Competition Policy Brief. (2016). European Commission. Available at: https://ec.europa.eu/competition/publications/cpb/2016/2016_001_en.pdf [Accessed 2 Nov. 2019].

[71] Arrow, K. Economic Welfare and the Allocation of Resources for Invention, RAND Paper 1959, available at https://www.rand.org/pubs/papers/ P1856.html, page. 15–20. [Accessed 14 June 2019].

[72] EU Merger Control and Innovation. Competition Policy Brief. (2016). European Commission. Available at: https://ec.europa.eu/competition/publications/cpb/2016/2016_001_en.pdf [Accessed 2 Nov. 2019].

Philippe Aghion and Rachel Griffith developed a theory which explained the Schumpeterian hypothesis with positive effect or market competition on innovative activity.[73] The model assumed that companies that are in intermediate goods market are able to manufacture different intermediate goods and that these companies can be categorised for two different types. The other one is competing directly with the company, and there is a small technological gap between these rivals. Companies have to be innovative in order to escape competition. In the other type, companies that have a considerable gap have little to nothing reasons to innovate since the returns that the company receives are minimal. Aghion and Griffith presented that this relationship with innovation and competition can be defined with an inverted U-curve meaning that increased competition spurs innovation. If however, the competition becomes too fierce, the increased competition reduces reasons to innovate since the rewards are small.[74]

Essentially the threat of competition in itself incentivises companies to act in more innovative working methods. According to economist Carl Shapiro, neither methods towards competitive markets and innovation are not reciprocally exclusive.[75] Both Schumpeter and Arrow mostly agree on the point that markets have to be contestable, indicating that they must enable new competitors to achieve market access and to use that particular market access to challenge incumbent companies. Also, to stimulate innovation, synergies within equivalent goods or services have to be distinguished and understood, since bringing together such synergetic goods may develop the likeliness of innovations to take form in the future.[76]

Competition law plays a crucial part here as well since it offers tools to take action upon market foreclosure and the competitive harm arising directly or indirectly from there and to justify synergies that equivalent goods or services can bring to spur innovation. Views from Arrow and Schumpeter agrees that companies need to be able to gain the advantage of their innovation and to benefit from the competitive advantage achieved from their innovative efforts. Companies have to be able to allocate the value of their innovation.[77] In order to do so,

---

[73] See Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P. (2002), *Competition and Innovation: An Inverted U Relationship*, NBER Working Papers 9269. And See also Aghion, P., Griffith, R. (2005), *Competition and Growth: Reconciling Theory and Evidence*, Cambridge, Mass.: MIT Press.

[74] Ibid.

[75] Shapiro, C. (2012). *Competition and Innovation. Did Arrow Hit the Bull's Eye?*, in Josh Lerner and Scott Stern, The Rate and Direction of Inventive Activity Revisited. NBER, 361–410.

[76] EU Merger Control and Innovation. Competition Policy Brief. (2016). European Commission. Available at: https://ec.europa.eu/competition/publications/cpb/2016/2016_001_en.pdf [Accessed 2 Nov. 2019].

[77] Auer, D. (2017). *Appropriability and the European Commission's Android Investigation*. Columbia Journal of European Law 23, page 647–680.

alongside competition law, the recognition of certain types of intellectual property rights protection and ways to keep relying on their innovations are to be spurred in that regard. Even though protecting innovations typically reduces competition to a certain degree since patents can also be abused and in particular hinder innovations which are based on patent-protected advancements. The link between intellectual property and competition law are still, two sides of the same coin. On the one hand, intellectual property law is intended to protect the profits of innovators innovative methods. On the other side, competition law intends to secure that markets remain to foster and spur innovation. So, therefore, it is clear that the encounters within both areas of the law are notably significant for the role of innovation in competition law analysis.

Not only do innovations stimulate the economy as a whole, but they also represent a significant part in the growth of individual companies. Innovations are significant factors for competition. According to the insights gained from the Arrow-Schumpeter debate, competitive tension is required in order to encourage companies to generate new or enhanced goods and technologies. Still, companies that want to be innovative must also have a sensible possibility of being compensated for such efforts and investment with an at least brief improvement in market power and, correspondingly, greater earnings as a consequence of the innovation. Even though some praise the view that the more prominent and most influential companies are the most innovative, the opening of former monopolies to competition, shows that there is a clear concrete relationship within competition and innovations. As an example can be shown the opening of the packaging disposal area to competition uniting it with competition instruments released a flow of innovation in this particular area. This then caused great cost reductions and developments in the quality of recycling.[78]

Whether a company has a dominant position in a market or what actually is the market in question, we must define the market. Otherwise, it is not possible to tell whether there is a reason to enforce the provision of competition law. Market definition is needed to identify what is the market that competition takes place. This tells us whether the innovating companies in question, whether it being a merger or any of the three branches mentioned previously, are directly competing companies. Discussions of the connection between innovation and

[78] Bundeskartellamt (2017). *Innovations-Challenges for competition law practice*. Page 1 Available at: https://www.bundeskartellamt.de/SharedDocs/Publikation/EN/Schriftenreihe_Digitales_II.pdf?__blob=publicationFile&v=3. [Accessed 14 Jun. 2019].

competition law had its peak in the aftermath debate of the *Dow/DuPont* merger. The *Dow/DuPont* merger was seen pointing a significant difference from past cases, according to economists as well as lawyers. According to the views presented in relation to *Dow/DuPont* merger[79], the competition law analysis concentrates on research and development abilities, as opposed to precisely defined product markets, it joins a speculative area. However, the fact that companies focus resources on the development of new and novel goods does not indicate that these attempts will be prosperous. In addition, there is not an evident connection in the level of concentration and the scale of innovation in a given industry. Arguments against this view from the opposite side have as well been eloquently advanced. The belief that competition law can assess the impact of transactions in light of companies research capabilities is undoubtedly already known. A framework based on this same approach has existed from the first version of the Guidelines on horizontal co-operation agreements. In addition, it has been accepted for a long period that competition law analysis concerns not only to price and output but also to other parameters, namely innovation.[80]

Strategic administration researches have contributed an essential insight by distinguishing in 'sustaining,' and 'disruptive' innovation: sustaining innovation takes place in the value network of companies, providing clients more or better in the attributes they already appreciate. Disruptive innovation takes position outside that value network, creating new favoured attributes.[81] When a company conducts these methods, they can either expand their markets by either innovating sustainably and generate better products for the market they are competing, or capture more customers by innovating in a disruptive manner and generating something different to a completely different market than they originally are known for. Competition law has a role in both scenarios preserving competition, and this comes in question especially when a company might gather great market share. Either having clients from its own market or as well from the new market by capturing market share, e.g. in a tying manner mentioned above. The company might innovate a novel product and then tie it to its other products in the same or different market.

[79] Colomo, P. *Competition Law and Innovation: Where do we stand?* Available at https://academic.oup.com/jeclap/article/9/9/561/5234221 [Accessed 12 Jun. 2019].
[80] Ibid.
[81] Streel & Larouche (2015). *Disruptive Innovation and Competition Policy Enforcement*, TILEC Discussion Paper 2015-21 5.

What needs to be said is what the courts have also stated, the dominant position is not illegal in itself.[82] By being innovative and creating a natural dominance is supposed to result from higher efficiency, an idealised running of the market that the Courts calls 'competition on the merits'.[83] However, a dominant status is defined by safeguards against competitor's productive efficiency, market power, and dynamic allocative efficiency, market barriers.[84] The reasonable result is that competition on the merits depends on innovation in order to challenge dominant positions. This is also the opinion of the US Supreme Court: "the opportunity to charge monopoly prices, at least for a short period, is what attracts "business acumen" in the first place; it induces risk-taking that produces innovation and economic growth".[85] This puts innovation in the sphere within the competition law's mandate.

In conclusion, competition does not just provide an incentive to companies on providing more and charge less. To beat the competition of its rivals also stimulates the company to innovate. Competition provides a powerful motive for companies to be more innovative. There are direct and indirect beneficiaries in this scenario. Companies and customers of the market benefit it directly since innovative companies generate more profit and buyers can purchase either cheaper or improved goods. Indirectly the rivals benefit this as well since they do gain insights into these innovations and how the market reacts to them.[86] They can then improve these innovations and be more innovative in their efforts on doing so, which then creates more direct and indirect benefits to the market. It is therefore clear that innovation as such can benefit from a competition law framework that guarantees the contestability of markets, recognises the potential to appropriate certain advantages and spurs synergies between equivalent goods. So competition law, therefore, serves above all as an instrument to enable and promote innovation in markets, both at the state of individual goods or at the state of competitive markets in general. Competition law should as well work as an instrument to fight against innovations that have as a result to distort competition as diminished contestability and foreclose markets. Competition

[82] Case 322/81, Michelin v. Commission, C:1983:313, ECJ

[83] Case C-209/10 Post Danmark I v. Konkurrencerådet , ECJ para.21.

[84] As such, an allocative efficiency explanation of why dominant positions are allowed does not hold: the rents from market power do provide an incentive for entry, but market barriers prevent it.

[85] Verizon Communications, Inc. v. Law Offices of Curtis V. Trinko, 540 U.S. 398, 407 (2004)

[86] Baker, J. *Beyond Schumpeter vs. Arrow: How Antitrust Fosters Innovation* (2007). Articles in Law Reviews & Other Academic Journals. Paper 276

law, in general, should therefore be acknowledged as an instrument that provides means and tools to make innovation competition possible.

For a successful future, the digital economy creates a central driver to innovations, efficiencies, and consumer welfare. It has revolutionised business structures, services, goods, communications, and social synergies. Digitalisation has also spurred a change in market dynamics, paving the way for the rise of essential platforms, networks, and the proliferation of multi-sided markets.[87] A growing data-driven economy is necessary for innovation, growth, and European competitiveness as well as for an effective Digital Single Market. The Communication "Towards a data-driven economy"[88] offers a concept for the data-driven economy as an ecosystem with different types of players such as data providers, data analytics companies, skilled data and software professionals, cloud service providers, driving to more business possibilities.[89] The availability of better quality, reliable, and interoperable datasets was particularly pointed as a critical enabler for novel data products.

The Commission is also enlarging merger control to innovation-related assets like data and intellectual property, as mergers of free service providers, risk escaping turnover thresholds.[90] The Commission does seem certain about the situation that it can position innovation under the current status quo of competition law.[91] Innovation in itself continues to be an open-ended notion. In principle, it comprises the commercialisation of recently invented or improved products, which can be classified as product innovations, or production and distribution processes, defined as process innovations. Innovations like these can take place in a most disruptive fashion, generating a completely new market for a new product or can be fairly incremental, which could change the whole dynamics on an existing market.[92] However, no

[87] Ezrachi, A. (2018). *EU Competition Law Goals and The Digital Economy.* Available at https://www.academia.edu/36790134/Ezrachi_-_Goals_and_the_digital_economy_-_Working_paper.pdf. [Accessed 12 Jun. 2019].
[88] European Commission, (2014). *Communication from the commission to the European Parliament, The Council, The European Economic and Social Committee and The Committee of The Regions.* Available at http://ec.europa.eu/transparency/regdoc/rep/1/2014/EN/1-2014-442-EN-F1-1.Pdf. [Accessed 19 June 2019].
[89] European Commission, (2017). *Commission Staff Working Document on the free flow of data and emerging issues of the European data economy.* Available at https://eur-lex.europa.eu/legal-content/LV/TXT/?uri=CELEX%3A52017SC0002. [Accessed 23 June 2019].
[90] EC Press release, (2016). *Mergers: Commission seeks feedback on certain aspects of EU merger control.* Available at http://europa.eu/rapid/press-release_IP-16-3337_en.htm. [Accessed 26 June 2019].
[91] European Commission, (2016). *Competition policy brief: Eu merger control and Innovation.* Available at http://ec.europa.eu/competition/publications/cpb/2016/2016_001_en.pdf. [Accessed 2 Nov. 2019].
[92] Ibid.

matter what classification of innovation one adheres to, it has the potential to shape markets or to at least switch the competitive dynamics within those markets. In the highly developing digital era, where data has been classified as a ”new oil”[93], web-scraping can be seen as an innovation which can affect the competitive dynamics in the market.

Chapter Three - How is web-scraping regulated and why it is a competition law issue?

Before distinguishing the key areas where web-scraping triggers competition law interest, we must define what actually is web-scraping. Also, a lot of discussions has been in the value of data and since web-scraping does include the gathering of a large amount of data, it is essential for the reader to comprehend how big data can be best understand to benefit companies and why there is an incentive for companies to collect large data sets of information from the web. After comprehensively explaining these two themes, the author has a ground to explain the benefits and threats of web-scraping and especially its threat to intellectual properties since this has triggered a large amount of discussion among scholars in the area of web-scraping. Lastly, this chapter will discuss web-scraping in the eyes of competition law and answer in a conclusive manner to the second sub-question: How is web-scraping regulated and why it is a competition law issue?

3.1 Methods and Means

For a much precise view for the reader, the author sees reasonable to shortly mention the aspect of big data in competition law sphere. Since web-scraping does include the gathering of large amounts of data from the web, it is essential to comprehend the basics of this asset.

”We don't need a whole new competition rulebook for the big data world. Just as we didn't need one for a world of fax machines, or credit cards, or personal computers.”[94] Value and importance of data have grown massively in the past years. Data can now be seen as a new currency in the digital era and justifiable so. Massive amount of data is generated throughout the world by companies and consumers. Click on the websites, cookies, internet searches,

[93] Neelie Kroes Vice-President of the European Commission responsible for the Digital Agenda Digital Agenda and Open Data From Crisis of Trust to Open Governing Bratislava. (2012). [online] Available at: https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_12_149. [Accessed 20 Dec. 2019].
[94] Vestager, M. (2016). *Competition In A Big Data World*. European Commission. Available at: https://dld-media.com/events/opening-keynote-e5a8fd5c-c210-44f4-b93c-882669f06688 [Accessed 11 Nov. 2019].

profile information through websites, use of a mobile phone, e-commerce transactions. As FTC Chairwoman Edith Ramirez described in her Keynote Remarks at the Fordham Competition Law Institute's Annual Conference on International Antitrust Law and Policy in 2016, "there is no question that the aggregation of data may have important implications for competition."[95] Also, EU Competition Commissioner Margrethe Vestager has promised to "keep a close eye on how companies use data."[96]

Big data can help to comprehend better what consumers truly need, which also serves consumers and produces better services. Examples of this might be more precise recommendations when shopping online or personalised medical treatments, which are genuinely better for us.[97] When dealing with big data, these large collections of data might help companies to find patterns that they would not see in smaller individual data sets. "After all, the whole point of big data is that it has to be big. Because, with the right tools, you can find patterns in a large set of data that you just wouldn't see in a smaller one. And we don't want to discourage companies from putting in the effort to collect that data."[98] To put it in a simple frame, data is really a product, and therefore, the same competition law analysis can be implemented to it as is implemented to any other product. Data can also play a role as an incentive to another company to buy a rival due to its possession of data.[102] From a competition law perspective, a relevant question that issues is also whether the access and management of big data by companies can award them with market power and a competitive benefit over their rivals. As an example, if the data examining tools and the construction of complicated artificial intelligence algorithms demand significant investments, can access to big data result in highly concentrated markets with high entry barriers?[104]

---

[95] Ramirez, E. (2016). *Keynote Remarks of 43rd Annual Conference on International Antitrust Law and Policy, Fordham Competition Law Institute: Deconstructing the Antitrust Implications of Big Data.* Available at: https://www.ftc.gov/system/files/documents/public_statements/1000913/ramirez_fordham_speech_2016.pdf [Accessed 15 Apr. 2019].

[96] Vestager, M. (2016). In: *Conference on Big Data.* [online] Available at: https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/big-data-and-competition_en [Accessed 15 Apr. 2019].

[97] Vestager, M. (2016). In: *Conference on Big Data.* [online] Available at: https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/big-data-and-competition_en [Accessed 15 Apr. 2019].

[98] Ibid.

[102] EC Report by Crémer, J. de Montjoye, Y. Schweitzer, H. (2019). *Competition Policy for the Digital Era.* Available at http://ec.europa.eu/competition/publications/reports/kd0419345enn.pdf. [Accessed 12 July 2019].

[104] Autorité de la concurrence and Bunderkartellamt, (2016). *Competition Law and Data*. Available at http://www.autoritedelaconcurrence.fr/doc/reportcompetitionlawanddatafinal.pdf. pages 12-13. [Accessed 1 July 2019].

Competition issues commonly do not appear when companies gather more data, and competition does not usually impose on companies an commitment to share data that they have gathered and developed. In such situations, competition law could dampen innovation, which is the very behaviour that competition law is meant to protect.[105]

There is also the factor of companies willing to secure their datasets they have bought or collected by themselves. As a general example would be companies that whole business model is based on the monetisation of the collection of large data sets: "Some providers of online platforms try to shield data away from competitors: in the case of Facebook[106], for example, by prohibiting third parties in its general conditions from scraping content off its platform or, in the case of Google[107], by restricting the portability of advertising campaigns and by requiring websites to enter into exclusivity agreements for search advertisements".[108]

When it comes to innovation, the pressure of competition is one of the main incentives for competitors to launch innovative and more advanced products. Where again, the excessive market power and the over-concentration of the market creates a disincentive.[109] Previously stated data sets and a large amount of data holds a great possibility for companies to create new innovations. Therefore one of the innovations that companies have been keen on using is web-scraping. In order to educate ourselves with how actually companies could exploit these new types of innovations relating to data and web-scraping, we must shortly define web-scraping and its process so that we can pinpoint some of the legal issues behind it.

---

105 Competition Bureau, (2017). *Big data and Innovation*. Available at https://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/04304.html. [Accessed 3 July 2019].

106 Under Facebook's Statement of Rights and Responsibilities on Safety, Facebook prohibits automatic collection of user content: "You may not access or collect data from our Products using automated means (without our prior permission) or attempt to access data you do not have permission to access". available at https://www.facebook.com/terms.php. Accessed 11 June 2019.

107 Speech of former Competition Commissioner Almunia. (2012). *Statement of Commissioner Almunia on the Google antitrust investigation.* Press room Brussels, 21 May 2012, SPEECH/12/372. Available at http://europa.eu/rapid/press-release_SPEECH-12-372_en.htm. [Accessed 11 Jun. 2019].

108 Graef, I. (2015). *Market Definition and Market Power in Data: The Case of Online Platforms.* Page 479-480 Available https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2657732. [Accessed 23 Jun. 2019].

109 See as an example on investment disincentives due to a merger between the single companies that perform research and development in a sector, Commission Decision of 28 January 2015 under Article 6 (1) (b) in conjunction with Article 6 (2) of Council Regulation No 139/2004 (Case No COMP / M.7275 - Novartis / GlaxoSmithKline Oncology Business).

Commonly data sets are transferred over the Internet by some abstract structured data set which can then be read and processed without undue burden by a computer program. Currently, the Internet is full of data which is unstructured with pieces here and there without any abstract and well-structured data set. This data then, some useful some not, might be left out without a proper tool for researchers to process this significant bulk of data. Web-scraping is fulfilling that gap by providing researchers structured data sets which are human-readable, by automating the structuring of that extracted data from the Internet.[110] The process imitates how a web-browser functions by reaching through different webpages and then storing them to computers hard drive cache. Web-scraper enters through webpages, discovers specified data components on the page, extracts them, converts them if that is required, and eventually stores these data as a structured data set.[111] Web's fast growth has significantly shaped the methods we share, obtain, and distribute information and data. An enormous quantity of information is being collected online, both in structured and unstructured forms. To overcome this massive amount of untangled data, automated web scrapers are the tools in order to utilise this information. Without these techniques, it would be simply not possible to collect a significant quantity of data repeatedly and in a reasonable time. Without web-scraping there would not be a search engine. To identify and scrape the dedicated information from web-pages, the scraper uses its contents and structure to collect this data.

The techniques of web-scraping have evolved together with the Internet. Therefore different web-scrapers have also been developed, ranging from harvesting extremely diverse information from several websites to highly dedicated web-scrapers adjusted on selective items on a single site. Needs have also developed completely automatic web-scrapers which operate to programs that demand user cooperation or a one-time scraping for a research project to use in production for example. In today's hyper-competitive world, it is extremely important for companies to stay up-to-date about market trends and its changes, prices, and not only their own but rivals customers as well. Websites hold detailed and regularly renewed information that may be valuable to certain statistics which companies could then utilise. Companies can improve the outcomes of the traditional manually performed techniques with automatically gathered information. Web data may enable new types of indicators that are not possible to achieve with

[110] Mitchell, R. (2015). *Web Scraping with Python: Collecting Data From The Modern Web*. 1st ed. 1005 Gravenstein Highway North, Sebastopol, CA95472.: O'Reilly Media, Inc.,.

[111] Boeing, G. & Waddell, P. (2016). *New Insights into Rental Housing Markets across the United States:Web Scraping and Analyzing Craigslist Renal Listings*. p.3. Available at:https://arxiv.org/ftp/arxiv/papers/1605/1605.05397.pdf. [Accessed 28 Aug. 2019]

manually performed techniques. The gathering and analysis time is significantly faster than executing these assignments manually. This means for example fewer surveys done by customers which are used to give companies a much more precise picture of their markets. Companies see this as an asset since it is faster better and much more efficient.[112]

While innovation has developed scrapers for companies to use, countermeasures have also been created to prevent crawling the site as much as possible. Robots.txt is a tool which can be used to guide search engines on the company's website. It is stored in the root register of a domain. Therefore the first thing crawler sees when it enters the website is a robot.txt file. Not all search engines robots, unfortunately, follow these rules. If you use it for image files, it prevents these files from appearing in the Google search results. Therefore robots.txt is generally used to determine which parts of the website is shown on search engine result, limit these too much, and this has a negative effect on the ranking of your website. The instructions in a robots.txt file have a strong influence on Search Engine Optimisation as the file allows you to control search robots.

So, extracting data automatically from the web is not only much more cost-efficient than doing it manually but also more precise due to the amount of it. But why extract data and information from the web, what are the actual and concrete benefits of it to companies and are there any downside to it?

3.2 Benefits and threats to companies

As we defined web includes an enormous amount of information, and as data grows in amount, businesses have to concentrate on the diversity and importance of this data and find the pieces of it which matter to them the most. Companies that can recognise the full potential of this data and information and exploit the critical subset of that data will drive meaningful positive impact for user experience, solving complex issues, and creating new economics of scale.[113] Since the web also contains data that is not relevant, it is critical that companies concentrate and focus their attention to identifying information that benefits their agenda and provides them that

[112] Bosch, O. (2017). *An Introduction to Web Scraping, IT and Legal Aspects*. European Commission.
[113] Reinsel, D. & Gants, J. & Rydning, J. (2017). *Data Age 2025: The Evolution of Data to Life-Critical.* An IDC White Paper. Available at https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/workforce/Seagate-WP-DataAge2025-March-2017.pdf [Accessed 29 Jun. 2019].

crucial piece of information in order to thrive on the market. Not only that, but data has grown to be a vital resource for economic growth, societal development, as well as creating employment. Data analysis facilitates the optimisation of processes and choices, innovation, and the forecast of events happening in the future.[114] Therefore companies that are embracing on utilising new technological innovations such as, web-scraping, big data, and machine learning, are obtaining a market share, and are disturbing some of the previous generation companies that have been more hesitant to adapting and modifying their business methods and principals to digital changes that innovation has provided.[115] By the forecast of IDC, the global data sphere will continue to grow to 163 trillion gigabytes by the year 2025. This is ten times more than the previously generated data in 2016, which was 16 trillion gigabytes.[116] This data has a huge potential for business to reveal new business opportunities and ideas if properly utilised. Here is when web-scraping comes into the picture. With this unorganised growth, it is no longer possible to manually track and record all available sources. Automated scraping techniques allow the collection of a massive amount of data from the web compared to manual data extraction. Scraping data can provide companies various advantages, but the essential benefit of it is that it can place the business in a competitive position.[117] Web-scraping provides a possibility for a company to harvest data for market research which will then give the company an insight whether their business strategy is up to date for the current information on their industry or any other purpose. It is vital for businesses to have up to date information since, without it, they cannot get insights into the market changes in real time and therefore lashes out of the competition. When this information is provided to companies, they know what is happening on the market, and they can react to those changes accordingly and consequently minimise their losses or maximise their profits. Therefore they need to extract the latest information from the web so that strategies and business methods can then be adjusted accordingly to that specific market. Up to date data from the market is crucial for businesses

---

[114] European Commission, (2017). *Communication From The Commission To The European Parliament, The Council, The European Economic and Social Committee and The Committee of The Regions.* Available at https://ec.europa.eu/transparency/regdoc/rep/1/2017/EN/COM-2017-9-F1-EN-MAIN-PART-1.PDF. [Accessed 31 Aug. 2019].

[115] Picardo, E (2019). *10 of the Worlds Top Companies Are American*. Available at https://www.investopedia.com/articles/active-trading/111115/why-all-worlds-top-10-companies-are-american.asp. [Accessed 30 Aug. 2019].

[116] Reinsel, D. & Gants, J. & Rydning, J. (2017). *Data Age 2025: The Evolution of Data to Life-Critical.* An IDC White Paper. Available at https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/workforce/Seagate-WP-DataAge2025-March-2017.pdf [Accessed 29 Jun. 2019].

[117] Several services are provided by different web-scraping providers. see, Webscrapingexpert.com. (2019). *Web Data Scraping, Website Data Scraping, Web Page Scraping Services | Web Scraping Expert*. [online] Available at: http://www.webscrapingexpert.com/web-data-scraping/ [Accessed 27 Oct. 2019].

that have their business in price quality and competition. As most commonly used example is travel or hotel businesses that claim that they have the best price on the market. They have to get this information through data extraction so that they truly have the most appealing price on the market to the customer. As for retailers online, they must have the latest information about discounts and price comparison for their competitors and what specific product has been sold out so that they can adjust their own stock and prices for the current market trends.[118]

Companies want to learn about their current or possible markets as well as its consumers, and therefore, every consumer's opinion on what they like and what is not appealing to them is useful for the company. Consumers share their opinions and leave feedback for other consumers, but this is not only useful for the company and its other consumers, but also its rivals. Movie producers want to know opinions about their movies, data like this is widely shared on review portals and feedbacks on different portals.[119] When a company wants to update its prices, it needs to adjust its current plan for the markets, and for this, it needs targeted scraping of data from different online shops in order to update its price range. With web-scraping, the company can update its indexes much more frequently. Collecting data from online sources is one of the most efficient ways of conducting market research. Compared to traditional surveying it provides much immediate response, compared to classical surveying. Social media data can be obtained through publicly available sources by scraping. These social media platforms afford indicators about current and ongoing trends for various markets. Therefore these platforms form the major data source not only by communicating with others but as for companies to find insights about trends and opinions. Social media data are just like conventional data, which is a potential treasure chest but requires scraping and mining to uncover hidden treasures. [120] Web-scraping is seen as cost-effective support for such instruments.[121] Web scraping makes it possible for a company to study both their own as well as the appearance of competitors in headlines of news servers. It can also obtain details about

[118] See, company providing web scraping services presents common reasons to web-scrape Scraperapi.com. (2019). *What kind of companies use web scrapers? | Scraper API*. [online] Available at: https://www.scraperapi.com/blog/what-kind-of-companies-use-web-scrapers [Accessed 27 Oct. 2019]. Also, Scrapy providing information on its clients and how they have provided services. Scrapy.org. (2019). *Scrapy | Companies using Scrapy*. [online] Available at: https://scrapy.org/companies/ [Accessed 27 Oct. 2019].
[119] Yang, J. & Yecies, B. (2016). *Mining Chinese social media UGC: a big-data framework for analyzing Douban movie reviews.* Journal of Big Data 3.
[120] Huan, L. (2016). *The good the bad, and the ugly: uncovering novel research opportunities in social media mining.* Springer International Publishing Switzerland
[121] Raulamo-Jurvanen. & Kakkonen, K. & Mäntylä, M. (2016). *Using Surveys and Web-Scraping to Select Tools for Software Testing Consultancy*. M35, ITEE, University of Oulu.

rivals and even about its own employees. Human resource (HR) divisions in large companies process multiple job proposals for their companies and try to match the position with prospective employees. HR Offices also cooperate with third-party companies, which can offer them own lists of professionals. Contact scraping is an essential activity for such businesses. As an example, *hiQ* was a tech startup which collected and analysed public profile data on *LinkedIn* in order to provide its clients, insights about their employees.[122] These insights were information, such as which employees are likely to be lured by a rival company or which skills its employees have.

Web-scraping is even used in enterprise technology level. A water utility company PDAM Surabaya had its water supply system reservoirs monitored by a sensor device which can provide water quantity parameters value information, but since it is a passive sensor, it stores its data internally in the sensor itself. To solve the issue, an application of data logger was then introduced to control data collections online water quality monitoring system by utilising web scraping.[123] Data extraction procedures also appear as a fundamental tool to execute data analysis in competitive intelligence structures as well as for company process re-engineering.[124]

But those advantages may come at a cost to the owners of the websites that are being scraped. The company which is scraping data may obtain sensitive information which is owned by another company, distract that company's customer away from a scraped site, or cause a scraped site to crash completely or slow down so much that it isn't appealing to visit that website. That means that just like everything else in the world, the act of conducting web-scraping can be used for malicious and unethical intentions as well.It can also be used to access and acquire private, unauthorised data that is not public. Obtaining such information would be a violation of the General Data Protection Regulation (GDPR).

In general, scraping of data can be done without the permission of data owner and in complete violation of scraped websites terms of service. This would then concern contractual liability and possible claims according to contract law. There are several companies that expressly

[122] *hiQ Labs v. LinkedIn* [2017] (U.S. District).
[123] Nourma, P. & Budiarti, R. & Widyatmoko, N. & Hariadi, M. & Hery Purnomo, M. (2016). *Web scraping for automated water quality monitoring system: A case study of PDAM Surabaya.* pages. 641-648. Conference paper.
[124] Ferrara, E. & De Meo, P. & Fiumara, G. & Baumgartner, R. (2014). *Web Data Extraction, Applications and Techniques: A Survey.*

prohibit scraping within their website terms and conditions. Data scrapers can also put extremely heavy loads on a website's servers by requesting for data more times than a human does manually. This might then crash or slow down the scraped site and create damage claims depending on the situation and damages cost to the site owner. It is an offence to cause criminal damage to a computer, including damage to data, or to use a computer to access data without authorisation. Accordingly, data scraping could be a criminal offence as the website owner has not authorised access to the data.

So for companies and enterprises as well as organisations, web-scraping is able to afford vital commercial advantages. It can enable companies to produce indexes and list data, observe rivals 'services and prices, and acquire consumer and merchandise marketing insight from different sources such as forums, social media, blogs, and sites that hold reviews. Yet, while there are notable benefits to using website data through methods such as web scraping in innovation research, the literature on the use and legality of these methods is relatively undeveloped.[131] Especially, the aspect of intellectual property and the threat that web-scraping creates on it has been one of the discussions around this innovation.

3.3 Threat to the protection of intellectual property?

As hinted in previously, depending on the circumstances, screen scraping presents a range of possible legal risks. Before we discuss it in competition laws perspective, firstly it is sufficient that we address the issues web-scraping might raise up in intellectual properties perspective along with the cases that has lead the discussion to the side of competition law in order to gather a concrete image about the issues.

Claims in web-scraping situations can take a variety of forms, including copyright infringement, breach of contract, also in the US a violation of the CFAA[132] and also tort speculations such as trespass to chattels. But since copyright law gives its author the exclusive right to reproduce, display and distribute their copyrightable works,[133] the most obvious and common issue that web-scraping is usually linked with is copyright law, since it does at least

[131] Bosch, O. (2017). *An Introduction to Web Scraping, IT and Legal Aspects.* European Commission.
[132] Computer Fraud and Abuse Act 1986
[133] Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society.

to some extent include copying information on the scraped website. Situations where scraping party blends that information with its own content connection emerge. When data is copyrighted, it is then addressed accordingly with the intellectual property laws, since this would mean that the copyrighted data is actually ”copied”.[134] As scraping often deliberately means copying, it may cause copyright infringement.[135] It is then dependable about the facts and details of the issue at hand whether that particular data is copyright protected. A company conducting scraping may reproduce a website owner's trademarks without their permission, this would possibly lead to a trademark infringement case.[136] Authorial works which are protected by copyright include original literary and artistic works. Computer programs can be counted as such as well as website graphics and photographs. Since web-scraping gathers a large amount of data by copying it to another location that it can then be used by different purposes, the most logical way of approaching this issue is by taking data as a primary object of discussion. If the data set is a creative work, then this would mean that format which that data is presented is what is copyrighted, not the data itself. So by scraping mere facts from the work shouldn't get the scraper in any trouble. Or should it? Companies often have their own databases which have different value and contain different information and data for various purposes. By formulating data from different sources and then organising them into a database in a creative way, the database itself might receive protection secured by Database Directive.[137] A database right is infringed in a situation where a substantial part of a database is either extracted or/and reutilised without the consent of the owner.[138] Scraping data is usually repeated extraction or re-utilisation of insubstantial parts of a database which conflicts with the normal use of the database and may also infringe database rights.[139]

An extremely interesting case and judgement were given in *Ryanair Ltd v PR Aviation (Ryanair)*[140] where the Court of Justice of the European Union (CJEU) judged that there were no intellectual property rights subsisted in the data that were scraped. Ryanair's database

[134] Ibid.
[135] Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society.
[136] Directive 2008/95/EC of the European Parliament and of the Council of 22 October 2008 to approximate the laws of the Member States relating to trade marks.
[137] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of database
[138] Ibid.
[139] Ibid.
[140] *Case C-30/14 Ryanair Ltd v PR Aviation BV* [2015] ECJ

included information of flight times and prices, and therefore the company that was conducting the scraping of the data had not infringed Ryanair's intellectual properties.[141] This was because the database itself was not the result of the essential creative input, which is necessary in order to be awarded copyright protection.[142] Not only that, but the local court of Utrecht and the court of appeals of Amsterdam stated that since Ryanair claimed the protection under Database Directive, it had limitations which couldn't be derogated by contract since the scraping amounted to normal use of the said website.[143] This meant that the ToS/ToU where Ryanair included a clause forbidding the users of the website to scrape any information, was invalid. This gives companies, that own databases, an incentive not to actually claim the protection of Database Directive since this then would limit their contractual freedom.

However, with this judgement, the CJEU made clear, that it is possible for a website owner to limit the re-use of the scraped data with its ToS/ToU[144] Companies should keep this in mind, not only those who own the website but those as well scraping the data since the judgement indicates that there is a possibility to be liable for a breach of contract.[145] This paradox then indicates that databases without the protection of intellectual property might receive greater protection than the one which is granted by intellectual property law.

Arguments in favour of web-scraping are presented that websites, which are filled with data, include that data in their websites for the sole purpose of that other people would see that information. In the case of *hiQ v LinkedIn (LinkedIn)*,[146] their business strategy is to publish data from profiles and then derive an income by charging a piece of service that is sold to their users. For the user, the goal is to get their CV´s exposed to a possible employer. Therefore possessing any private data would not be in their business strategy, and by extension of any data available is therefore public.[147] Generally speaking, publicly published data does not raise

---

[141] *Case C-30/14 Ryanair Ltd v PR Aviation BV* [2015] ECJ

[142] Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society.

[143] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases Article 6 & 8 & 15

[144] *Case C-30/14 Ryanair Ltd v PR Aviation BV* [2015] ECJ

[145] Rezai, A. (2017). *Beware of the Spiders: Web Crawling and Screen Scraping the Legal Position*. [online] Paris Smith. Available at: https://parissmith.co.uk/blog/web-crawling-screen-scraping-legal-position/ [Accessed 16 Oct. 2019].

[146] *hiQ Labs v. LinkedIn* [2017] (U.S. District).

[147] *hiQ Labs v. LinkedIn* [2017] (U.S. District).

intellectual property related legal issues. If data is secured by the owner that data cannot then be collected. Owner of the data might secure its data behind a safety mechanism like a traditional username/password or access code related technique. By scraping data like such can risk the company performing scraping process in legal proceedings since that data is not then public.[148]

Due to the certain rifts that intellectual property law neglects with regard to pure non-copyrightable data, several website owners turn to contract law to protect what they see as their own content, especially because of the *LinkedIn* case. Several websites and other online services are controlled at least some sort of ToS/ToU. These contracts are designed between the website owner and anyone who confirms their theoretical agreement by viewing the site or using the service. More and more ToS/ToU contracts include clauses which are designed to prevent or control the use of scraping. ToS/ToU agreements are usually enforceable up to a certain point, but they are viewed with suspicion, since they might contain unusual terms or where the agreement is not sufficiently brought to users' attention.[149] What creates even more speculation is the event that even if a scraping robot is able of comprehending the ToS/ToU agreement, it could be argued whether it lacks the legal ability to enter into a binding contract. So ToS/ToU agreements might be enforceable to some extent, but first of all, it is hard to declare damages in those situations, and it still does leave room to speculations. This turns lawyers and legal theories to look at solutions from further.

In summary, there is no specific piece of legislation which forbids web-scraping to gather information. As long as companies do not reproduce the information, they should be in clear waters. The website owners, however, may have legal rights against the company under intellectual property law and contract law at some occasions when the scraped data is reproduced or used in bad faith against the intellectual property law or against contractual obligations. Each case will turn on its own facts though, and this is very much dependent upon what information is scraped from the websites.

3.4 Web-scraping in the eyes of competition law

[148] Ibid.

[149] see. In re Zappos.com, Inc., Customer Data Sec. Breach Litig., 893 F. Supp. 2d 1058 (D. Nev. 2012).

Web-scraping is increasingly being used for business purposes to gain a competitive advantage. So there's often a financial motive behind it. It would be extremely naive to only consider copyright law as an only possible instrument to deal with the act of web scraping. In order to prove this, we start by first discussing a case where a shift towards different areas of law can be seen to be implemented by courts.

For instance, in *Field v Google Inc (Google)*,[150] heavy consideration was given to the role of a website operator which was preventing web scraping. The case was about Google's method of web crawling which it relies upon to populate its search engine. It explicitly involved building a cache of the defendant's individual site so that it may then afterwards be searched. Not only that but the plaintiff claimed that the website contained published works and that those works copyright was allegedly infringed when the site was cached. The *Field v. Google* case is significant since it measures the state of web crawling and the tools that are used to block it. These tools are similar to those we discussed earlier in the form of robots.txt method. Google had a way of indexing sites that occurs automatically, by making use of several "spiders" that crawl through the web scanning for unique sites, changes to current sites or site deletions. To block a robot from indexing a specific site, a webmaster can use a code, to refuse access to the content. In this *Field v. Google* case, the plaintiff had knowledge of Google's tool for indexing sites and the capacity of robot.txt to block this. Still, the choice was made not to use robot.txt, which meant that the court viewed as the plaintiff giving the defendant an implied consent to both cache and index the site. By taking a step to a more contractual view of law as we discussed in the previous chapter, companies have found a way to use contract law as a safety measure or a tool on their web-scraping practises.

In contrast, however, if a company decides to use countermeasures too strictly, robots.txt for essence, the overall outcome could be negative for the website owner, the possibility of fewer visitors, fewer links from content aggregator websites and less income from advertising. Therefore, data hosts should only use legal actions against scrapers when the scraper presents a threat to the data host's core business, and the data host has a strong enough claim to prevail legally against the scraper. ToS/ToU can, and sometimes should, as the case proves, include restriction of web-scraping measures. A simple step can carry the protection of a company's data asset to a long way, and it doesn't require much resources, and it allows direct

[150] *Field v. Google, Inc.* [2006] (United States District Court, D. Nevada.).

argumentation at the court. From a competition law perspective, it is necessary to adjust the ToS/ToU on the websites. So that it doesn't reach the level of anticompetitive conduct. By taking restriction measures too far, there might be a risk of foreclosure of competitors

The concept of that case was recently affirmed in the *LinkedIn*[151] case. Which proves a shift towards a more competition law perspective and gives companies more legroom to stand, which face a threat from companies that use their dominant position in another market. Here, LinkedIn was obliged to withdraw anti-scraping tools which blocked the plaintiff from utilising data on that network. Arguments were that LinkedIns data was crucial to the operation of hiQ´s business plan, that developed methods which then determined which individual might leave their current employer. When LinkedIn conducted anti-scraping measures, they pushed hiQ out of the markets since they couldn't any longer service their customers. In the court document, it was reasoned that LinkedIn is unfairly using its power in its "own" market of professional networking, to improve its position in another market (data-analytics) and therefore gaining an anticompetitive advantage.[152] This can also be said to be an example of disruptive innovation explained in the second chapter of this research and how innovative measures can be used in the age of the digital era. The court stated that hiQ successfully argued that LinkedIn violated the essential facilities doctrine, which precludes a monopolist or attempted monopolist from denying access to a facility that it has control over and is essential to its rivals. This then results in a possibility that hiQ would face a threat of foreclosure. The court believed that the reason why LinkedIn terminated hiQ's access to its public member data mainly was because it wanted exclusive control over its data for its own purposes in the market.[153] The *LinkedIn* Case implies that if a company desires to make use of search engines and indexation, it must be done in a way which does not breach competition law by restricting other competitors from accessing their data in a similar way.

It is also remarkable to note that the Commission maintains its power to investigate Google[154], in case of possible abuse within the meaning of Article 102 TFEU[155], when concerning the

[151] *hiQ Labs v. LinkedIn* [2017] (U.S. District).
[152] Ibid. p.21-22
[153] Ibid,
[154] European Commission (2015). *Antitrust: Commission sends Statement of Objections to Google on comparison shopping service', MEMO/15/4781*. [online] Available at: https://ec.europa.eu/commission/presscorner/detail/en/MEMO_15_4781 [Accessed 19 Oct. 2019].
[155] Consolidated version of the Treaty on the Functioning of the European Union Article 102 (ex Article 82 TEC)

practice of scraping competing engines content in its own search engines.[156] That means that actions that we previously discussed would usually be addressed within the meaning of intellectual property rights or competing on the merits since company has gathered the asset which he uses now to gain advantage on a different market, possibly on an innovative and disruptive way, might currently be falling in the scope of competition law, namely abuse of dominant position. So we see that web-scraping has reached a level where competition law has a reason to intervene. Shall we call this a ”when all else fails” view when an intellectual property doesn't provide sufficient protection, companies can rely on contract law. But also when contract law is used in an abusive manner in view of another company so that it might face a fear of foreclosure, competition law is the one that these companies can rely on. At least up to a certain point when the practice of using innovative measures of companies own assets doesn't receive the status of competing on mere merits, but with an anticompetitive ways.

So in order to answer the second sub-question, we can see that there aren't really clear rules or a piece of legislation that sets the game rule for companies conducting web-scraping. Case law has however provided us with some guidelines that we can follow on and come to a conclusion that when the protection of intellectual property reaches its limits, contract law has a way of extending this protection. This issue steps in the front yard of competition law, when contract law is used on fraudulent purposes it has an obligation to intervene and alter that contract or forbid it completely.[157] Also, the finding on a paradox in *Ryanair*[158] case which indicated that databases without the protection of intellectual property might receive greater protection than the one which is granted by intellectual property law, namely with contract law. Companies should keep this in mind, not only those who own the website but those as well scraping the data. Are current laws too outdated so that they are not able to deal with this issue directly? Should there be new legislation drafted or designed to create more accurate and clear game rules for web-scrapers? Or could competition law satisfy the gaps that copyright law and contract law are unable to fulfil? Should companies themselves be able to deal with the problem of their data being scraped and create new innovations to prevent scraping more effectively?

Chapter Four - Should competition law create a second level of protection stricter than copyright law?

[156] IP/16/2532: "The Commission has at the time also opened proceedings and it will continue to investigate the favourable treatment by Google in its general search results of its other specialised search services, and concerns with regard to copying of rivals' web content (known as 'scraping'), and undue restrictions on advertisers. ", MEMO/15/4781:" the Commission continues its ongoing formal investigation under EU antitrust rules of other aspects of Google's behaviour in the EEA, including the favourable treatment by Google in its general search results of other specialised search services, and concerns with regard to copying of rivals' web content (known as 'scraping'), advertising exclusivity and undue restrictions on advertisers. ".

[157] *hiQ Labs v. LinkedIn* [2017] (U.S. District).

[158] *Case C-30/14 Ryanair Ltd v PR Aviation BV* [2015] ECJ

We are testifying an industrial reconstruction which is run by digital data economy, computation and automation. Human activities, industrial processes and research all are heading to gathering and processing of data on an unprecedented scale, this then spurs new products and services as well as new business processes.[159] For a company, it is essential to innovate new business plans and methods. Therefore they are scraping and analysing this collected data in order to understand different market trends and other useful information for their business plans. Analysing the data from possible market provides better results, processes and decisions for a company. This again then helps them to generate new ideas or solutions or to predict future events more accurately. As technology advances, entire business sectors are being reshaped by systematically building on analysing data scraped from the web.[160] Data-driven innovation is therefore essential not only to businesses but to consumers primarily, since it provides and helps to invent better services and products.[161] In order to answer the question whether competition law should create a second level of protection to data that companies hold we must distinguish some issues where intellectual property didn't provide sufficient protection to companies when data were scraped from their websites. Then a distinction must be made between nationally legislated unfair competition and EU competition law since most of the arguments are tackled merely by this fact. Then we continue to consider extending competition law for matters falling outside of the reach of intellectual property law. Lastly, the author considers how the legislators could possibly extend the protection of this data and how to answer the question of this chapter.

4.1 Data falling outside of the protection of Intellectual property

As stated previously, for a company to establish a copyright infringement claim against data scrapers, the party asserting rights over the data must prove that the scraped data is copyrighted. These are exclusive rights given to the creator of creative work to reproduce the work for a limited time. Copyright is designed to protect the original expression of an idea in the form of

[159] European Commission, (2014). *Communication from the commission to the European Parliament, The Council, The European Economic and Social Committee and The Committee of The Regions.* Available at http://ec.europa.eu/transparency/regdoc/rep/1/2014/EN/1-2014-442-EN-F1-1.Pdf. [Accessed 19 June 2019].
[160] Businesses that use 'data-driven decision-making' enjoy a 5-6% increase in productivity,Tene, O. and Polonetsky, J. (2012). Big Data for All: Privacy and User Control in the Age of Analytics. *11 Northwestern Journal of Technology and Intellectual Property 239 (2013)*.
[161] SIIA- Software and Information Industry Association (2013). *Data-Driven InnovationA Guide for Policymakers: Understanding and Enabling the Economic and Social Value of Data.* [online] Available at: https://www.siia.net/Portals/0/ppt/Policy/Data%20Driven%20Innovation/data-driven-innovation.pdf?ver=2013-07-08-135946-000 [Accessed 21 Oct. 2019].

creative work, but not the idea itself.[162] It is subject to some limitations based on public interest considerations.[163] Web-scraping is collecting, and copying of extracted data from websites, so naturally, scraping copyrighted works are subject to copyrights laws and may lead to infringement claims.[164] Case law has revealed that also restaurant reviews that show emotion, personality, and that are original, may also be subject to copyright laws.[165] Scraping data which include authors personal opinions and comments does also have the aspect of originality and have also been considered to be protected by copyright in some situations.[166] This does not, however, mean that regular factual affairs on current news or happenings would also be copyrightable. It is also required that the company claiming to be the owner of the data need to have exclusive ownership of the copyright or be authorised by the original data owner to claim that the copyright is in possession of the company on the original data owners behalf. Generally, when proceeding claims upon scraped copyrighted data, a data holder has to prove that the data was an actual work entitled to copyright by an individual or a company. This means that first of all, in order for it be possible to provide protection by intellectual property, There must be an aspect of creative work. If a company that web-scraped copyrightable material, then intellectual property law should have enough means to provide protection towards it. As stated, there can be no valid copyright in pure facts. However, the presentation and substantial investment on a gathering of those facts may be subject to copyright.[167] The sui-generis database right is an exclusive right that protects databases against unauthorised extraction and re-utilisation of their content. This Database Directive gives protection for the database and not for the data itself as such. It is distinct and independent from copyright, which protects original works. In addition, web platforms and companies may say that they have valid copyrights ownership over their users created works if the ToS/ToU include an automatic transfer provision. This was tried by *Dianping v Aibang*[168] which renewed its ToS/ToU, pursuant to which users agree to assign their copyright in restaurant reviews exclusively to Dianping. In *Dianping*[169] the court denied the general enforceability of the updated ToS/ToU, recognising

[162] Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society.
[163] Ibid.
[164] Dreyer, A. & Stockton, J. *Internet Data Scraping: A Primer for Counseling Clients*, N.Y.L.J.,LITIG. 2013
[165] *Dianping v Aibang* [2011] the Beijing Haidian District Court.
[166] *Tencent v Toutiao* Court of Beijing
[167] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of database
[168] *Dianping v Aibang* [2011] the Beijing Haidian District Court.
[169] Ibid.

Dianping's exclusive copyright only in a few reviews where the authors had expressly declared their distribution of copyright to Dianping in writing.[170] A defendant's principal protection in copyright law in the US is a fair use defence. This is more of an open-ended doctrine contrasted to the copyright law limitations laid down in the EU jurisdiction. In the US, Courts consider a number of factors in determining if a use is fair. Points taken into account are: the meaning and nature of the use; type of the protected work; the volume of the work used; and the market value of the use.[171] Many of the commercial web-scrapers have had victory citing a fair use argument to a web platform's copyright claims outside of EU.[172]

In the EU, the *Ryanair*[173] price comparison case, which we discussed in the previous chapter, gives an excellent illustration of the situation on how businesses defend themselves against allegations of abusive web-scraping in the EU. Dutch company PR Aviation, a website where customers could search through flight data of low-cost airlines as well as compare fares and reserve a flight, extracted relevant data of flights from third-parties websites by methods of web-scraping. Therefore Ryanair alleged that PR Aviation's actions resulted in infringement of copyright relating to the formation and construction of the database and of the sui-generis database right under the Database Directive. The Local Court, Utrecht, and the Court of Appeals of Amsterdam both rejected Ryanair's claims. The Court of Appeals, mentioned PR Aviation's scraping of Ryanair's website amounting to a "normal use" of that website in the meaning of the lawful user limitations under Sections 6 and 8 of the Database Directive[174]. These Sections cannot be derogated by contract[175], which Ryanair declared that existed with PR Aviation when they entered and used the said website. That specific clause in the contract between the user and the owner of the website was that the user had to tick a box, of the airline's general

[170] Chan, V. & Co, *Fight against Web-Scraping: Legal Redress*, (2018)

[171] Zamora, A. *Making Room For Big Data: Web Scraping and an Affirmative Right to Access Publicly Available Information Online*, 12 J. Bus. Entrepreneurship & L.203 (2019) Available https://digitalcommons.pepperdine.edu/jbel/vol12/iss1/8 [Accessed 24 Oct. 2019].

[172] See, e.g., *Ticketmaster Corp. v. Tickets.com, Inc*., [2003] C.D. Cal. (holding that a data aggregator that scraped plaintiff's ticket purchasing platform in order to acquire event information was protected from plaintiff's copyright claim by a fair use defence, even though use was for a commercial purpose and only slightly transformative when source code was downloaded, final display was only of plaintiff's aggregated non-copyrightable information, and defendant's final product did not damage the market value of plaintiff's product).

[173] *Case C-30/14 Ryanair Ltd v PR Aviation BV* [2015] ECJ

[174] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases Article 6 & 8

[175] Ibid. Article 15

ToS/ToU, which amongst others prohibited unauthorised screen scraping practices for commercial purposes, was required when accessing the website.[176]

After Ryanair's appeal, ECJ declared that the Database Directive is not applicable to this situation since Ryanair's website did not fulfil the steps required to be awarded the protection of a Database Directive. Therefore databases which are not protected either by copyright or by the sui-generis database right, clearly cannot enjoy the protection of said Directive.[177] So, since the Ryanair's database did not enjoy the protection of copyright, exceptions to restricted acts that are set by Sections 6 and 8 of the Directive[178] do not prevent Ryanair as a database owner from establishing contractual limitations on its use by third parties. Therefore, owners of databases that could not rely on intellectual property protection may have opportunity to contractually restrict extraction and scraping of content from their online databases. As we found out earlier, this gives databases greater protection than what intellectual property protection might give by relying on contract law.

However, the legality of contractual limitations obstructing access and reuse of data through web-scraping methods should be assessed from then the viewpoint of competition law. An clear indication and example can also be found on this regard from the judgement of the Court of Milan which declared that Ryanair's refusal to grant access to its own database for the online travel agency Viaggiare S.r.l., was an abuse of dominant position which effects could be seen in the downstream market of information and intermediation on flights.[179] It is clear however that discretion should be used when considering and weighing the need to compensate for the efforts and investments that were required from the creator of the database with the interest of third parties, possible scrapers, to be allowed with access to this database and its information. A balance between protecting the investment and work efforts but also to protect consumers and grant market access. This is a similar view that was also taken in *LinkedIn v. hiQ*[180]. We have seen that web-scraping triggers issues that copyright alone cannot cover, not only abuse of dominant position by excluding its rivals, but also unfairness and free-riding on other

[176] De Santis, F. (2015). *ECJ clarifies Database Directive scope in screen scraping case*. [online] Globallegalpost.com. Available at: http://www.globallegalpost.com/blogs/global-view/ecj-clarifies-database-directive-scope-in-screen-scraping-case-128701/ [Accessed 25 Oct. 2019].
[177] *Case C-30/14 Ryanair Ltd v PR Aviation BV* [2015] ECJ
[178] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases Article 6 & 8
[179] *Viaggiare S.r.l. vs Ryanair Ltd.* Decision of June 4, 2013 of the Court of Milan.
[180] *hiQ Labs v. LinkedIn* [2017] (U.S. District).

companies inputs which is one of the main arguments when scholars discuss on web-scraping. It is only logical to dive a bit deeper into the unfair competition aspect and make the distinction between harmonised competition law before reaching towards the end of this research with conclusive arguments to answer the research question.

4.2 Unfair competition

What needs to be distinguished is that unfair competition is traditionally legislated at national level instead of an harmonised EU wide treaty. It often applies to dishonest or fraudulent rivalry and does not typically refer to competition law legislation in particular prohibition of restrictive practises, abuse of dominance or merger review. Distinction must therefore be done in order to properly answer to the research question since arguments about free riding on another companies labour for example is something that is decided nationally within its legislation and not by EU competition law as such which is the view that we want to answer the research question. Therefore we want to exclude plain arguments as such from our conclusive answer. It is still true however, that in a case of web-scraping it is hard to see that plain copying of data and using publicly accessed data from rival might lead to a foreclosure, and therefore the issue at hand might more often be easier to explain by unfairness or free riding of another companies efforts. In this way, a behaviour that would traditionally qualify as an infringement of intellectual property rules could fall within the scope of unfair competition. In order to give the reader a better view we discuss some case law in unfair competitions side and then focus again on the main issue within EU competition law.

Unlike in a copyright infringement claim, in an unfair competition claim the data owner does not need to have the exclusive ownership of the data or even express permission from the copyright owner for the claim. It would simply have a legitimate claim within unfair competition in so far as its economic interest can be proved to have been violated by data scraping by another company. The Chinese Court declared in *Dianping v. Aibang,*[181] that customer reviews which were gathered and filed on dianping.com website are plaintiff's own investments of work with important economic value and therefore should be protected within unfair competition protection. So even though no intellectual property rights could probably be awarded to the reviews in the website, due for the lack of originality, the work of gathering and

[181] *Dianping v Aibang* [2011] the Beijing Haidian District Court.

labour in itself for filing and creating something of economic value was seen as something that should be protected in order to prevent unfair use and free-riding on another businesses inputs. Insights from this can be drawn that courts also in the EU might consider more and more in the future that investment on work and companies asset should be protected by competition law related aspects due to its economic value. Legislative framework in EU has already shown that investments that is made in for example compiling a database is covered by the Database Directive even though the investment does not reach to a copyright protection.

One of the examples of an possible basis for an unfair competition claim, could be that the scraped data or content from the website is a substantive substitute of that data owners website. Then the users of this website can get the relevant information they possibly are seeking for from the scraper's site, and they do not need to go to the data owners site that often, this again then reduces traffic. It would seem highly unethical and against fair competition practise by just copying the whole business practise of a rival even if the scraper would dodge the intellectual property claims by disguising its conduct and only scrape and use data which can be considered not to reach copyrightable aspect. Sure it would depend then on the merits of the case how national court would look and judge these issues but this is one of the main concerns of the companies whom data are being scraped. They have conducted the work and innovative ideas behind their business practise, why would it be justified for another company to copy it and step on a same market to gather customers. The *LinkedIn v. hiQ* case we discussed earlier gave an ruling of the situation where a hiQ were scraping data to be innovative on different market, namely data analytics, and the court justified this partly because they did so to manage their business conduct on a different market than LinkedIn originally. So using scraped data on a different market could therefore be seen justified. However, it was said in the case concerning Baidu Map,[182] that this kind of substitution forms unfair competition because it explicitly harms the data owner's own interest by violating common commercial ethics and disrupts the economic order and market competition in the network environment. When courts are determining whether there might be a scenario of unfair competition, they ordinarily analyse the balance of different interests. While acknowledging the positive effect of a new business scheme of Baidu Map,[183] the Shanghai Intellectual Property Court ruled that the schemes

[182] Chinese innovation comparable to Google maps. See, Chan, V. & Co, *Fight against Web-Scraping: Legal Redress*, (2018)
[183] Ibid.

positive effect was unbalanced to the harm it caused to Dianping's legal, commercial rights. The Court added that the same advantages could be obtained through different means without harming Dianping. The display of restaurant reviews on Baidu Map led to unbalanced damage and less user traffic to Dianping. In cases involving several aspects of data, courts consider the balance of interests between consumers, businesses and fair market competition. Even though these two cases are not comparable as such, the author prefers the courts view in *LinkedIn*, namely because the court allowed hiQ to be innovative and use already existing publicly available data by scraping it to thrive on a different market where that data was originally used and where the LinkedIn was conducting its own business. It was not shown that any harm to LinkedIns own market were done by using its scraped data in a different market where hiQ were doing business. We move on to discuss whether and on what means competition law could extend to cover the areas when other fields of laws, namely intellectual property law and contract law, fail to do so.

4.3 Competition law reaching for help

Like most court cases, especially on competition law, each case has its own facts and therefore is ruled individually based on these facts. But as said, that is the beauty of competition law, it balances the interest of the parties and rights in question and finds a solution that could not be found in other fields of law. As we have seen, web-scraping can be justified by the fact that it benefits the market and generates competition which then again benefits the consumers, and that is the aim of competition law. Then again, there is a conflict of interest since companies whose data is being scraped feel that the rival will get unjustified benefit by using their gathered data using scraping tools. Intellectual property law protects to some extent, the content owner, whereas competition law seems, in the cases we presented, protect the scraper who conducts his web-scraping in order to thrive in the market lawfully. Owners of copyright might ague that intellectual property law should not be engaged but that the balance set by it should be respected. To the extent where one would scrape copyrighted material lets say e.q. a database protected content, competition law should not intervene and justify this act. Since it would create overlapping legislation and diminish the value that intellectual property law grants to authors. It is part of the protection of intellectual property that it grants a right to authorial work which cannot be duplicated without the authorisation of the owner. However, when the data does not reach an intellectual protection as such, discretion should be used when considering

and weighing the need to compensate for the efforts and investments that were required from the creator of the database with the interest of third parties, possible scrapers, to be allowed with access to this database and its information. There should be a balance between protecting the investment and work efforts but also to protect consumers and grant market access.

Depending on the facts of the case and situation at hand, web-scraping can therefore be seen as an objectively justified business practice and considered as competition on the merits as shown in the previous chapters with the case of *LinkedIn v. hiQ* where essentially it was seen that hiQ were not in the same market and LinkedIn preventing access to scraped data would have resulted in an abuse of dominant position. In contrast, when scraping is used on abusive manner, and companies lose their commercial potential when their data is stolen in an intrusive manner, case law has shown that they might need to turn to contract law in order to protect their data sufficiently. Which was hinted in *Ryanair*[184] when it claimed the protection under Database Directive, it had limitations which couldn't be derogated by contract since the scraping amounted to normal use of the said website.[185] This meant that the ToS/ToU where Ryanair included a clause forbidding the users of the website to scrape any information, was invalid. This then meant that contract law could be the tool providing more sufficient protection than intellectual property protection in certain situations. Subject to requirements provided by competition law that contractual clauses do not interfere with any of the three branches mentioned in Chapter 2, namely article 101 TFEU considering cartels, control of collusion and other anticompetitive practises.

Based on findings previously mentioned on different interests of consumers and companies in the market competition authorities could take into account a number of factors such as whether the copied and scraped material is adequately creative, or if the material was essentially factual in nature (e.q. statistics or reviews); whether the principal webpage earned some significant commercial advantage from holding the information publicly accessible that might deny its claim against a scraper; whether the reproduced material was utilised by a bad actor to remove market share or otherwise hurt the business of the principal webpage or if it was utilised in the making of a novel product; and whether there is a suggestion that anticompetitive reasons are

[184] *Case C-30/14 Ryanair Ltd v PR Aviation BV* [2015] ECJ

[185] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases Article 6 & 8 & 15

the purpose of the complaint.[186] These factors seem particularly suitable since then new entrant to the market is not being prevented from existing, but also that companies do not get undue support so that competition law would preserve inefficient or un-innovative companies in the market. Both of these situations benefits the welfare of consumers, which is the ultimate objective of competition law, since it creates more market players and therefore increases competition. We should preserve the good actors in the market and get rid of the bad actors. The future of online competition requires restraints to litigation driven by anticompetitive intent.[187] As several web-scrapers are commonly smaller companies than the websites that are targeted for scraping, a rule allowing access to publicly accessible factual data would efficiently limit risk for these smaller companies so that they would not end up in lengthy anticompetitive prosecution.[188] A consideration allowing new businesses to expand themselves based on information extracted from other websites would honour the traditions of openness in which the Internet was originally built.[189] As the Organization for Economic Co-operation and Development (OECD) declared, "the Internet is fundamentally designed to be open and global, which has enabled it to be an engine of economic growth and innovation."[190]

With new technologies and innovations, there comes issues that are unregulated or partly unresolved. But do we always need entirely new rules? Frank H. Easterbrook said in Cyberlaw conference that we should avoid "the law of the horse".[191] What he meant by this is that we can try to stretch existing rules, so they incorporate the new technologies and are able to regulate the unresolved issues in the current state of legislation. Competition law is doing exactly this. It means that we can act immediately and that the judges and competition authorities have the tools and possibilities to address issues at hand without having to wait for lengthy proceedings on new legislation to be implemented. The author thinks that one of the assets of competition law is that it can result in some occasions "second-guess" other fields of law. Where e.q. Contract law would not possibly find an agreement, competition law might solve it, or in a

---

[186] Zamora, A. *Making Room For Big Data: Web Scraping and an Affirmative Right to Access Publicly Available Information Online*, 12 J. Bus. Entrepreneurship & L.203 (2019) Available https://digitalcommons.pepperdine.edu/jbel/vol12/iss1/8 [Accessed 24 Oct. 2019].

[187] O´Rourke, M. (2001). Property Rights and Competition on the Internet: In Search of an Appropriate Analogy. BERKELEY TECH.L.J.

[188] Ibid. p.612

[189] Ibid. p.616.

[190] Economic and Social Benefits of Internet Openness: 2016 Ministerial Meeting on the Digital Economy Background Report. (2016). [ebook] OECD DIGITAL ECONOMY PAPERS No.257: OECD. Available at: https://www.oecd-ilibrary.org/docserver/5jlwqf2r97g5-en.pdf [Accessed 26 Oct. 2019].

[191] Easterbrook, F. (1996). *Cyberspace and the Law of the Horse.* University of Chicago Legal Forum.

situation where an agreement is reached within the limits of contract law, competition law might find it anticompetitive and illegal within its aims. So something legal in other fields of the law might get turned around if it has anticompetitive measures for the market. This is, of course, up to the point where the conduct or case at hand might actually lead to anticompetitive scenarios and effects.

Competition law has its own specific collective aim, the protection of consumers and prevention of anticompetitive means towards other companies and sometimes it might require some other generally considered lawful practice to be stopped. When there is not a clear set of regulation concerning web-scraping, above-mentioned case law has provided some guidance on how companies as well as competition law might deal with web-scraping issues. This can be seen as a lifeline for small and big companies on planning how to use web-scraping, and what are the possible limits to do so. However, it is not conclusive on the matter. Therefore we should move on to propose thoughts on how competition law could assert even further on coming court cases and settlements on the issue of web-scraping and data owner related aspects.

4.4 Possible ways forward

The Communication "Building a European Data Economy"[192] announces possible future EU framework for accessing data, identifying the most effective ways to achieve the objectives such as: facilitate and incentivise the sharing of machine-generated data; and protect investments and assets. Absence of legal protection in relation to non-personal or anonymised machine-generated data is that it is not yet structured and therefore cannot enjoy the protection of a database. It is also not, therefore, protected through intellectual property rights. Some scholars have explained that a potential way to solve this could be the drafting of a new data producer right that has as its principal aim of enhancing the tradability of non-personal or anonymised machine-generated data as an economic good.[193] This could be seen as a rem right.[194] It could then be considered as an exclusive right for certain data on websites. Data as

[192] European Commission, (2017). *Commission Staff Working Document on the free flow of data and emerging issues of the European data economy.* Available at https://eur-lex.europa.eu/legal-content/LV/TXT/?uri=CELEX%3A52017SC0002. [Accessed 23 June 2019].
[193] De Franceschi, A. (2016). European contract law and the Digital Single Market. 1st ed. Cambridge: Intersentia, pp.51-79.
[194] A right in rem is the term used for property rights.

such could then be licensed, and unapproved access and use of such data by scraping means could then raise claims of infringement.

Data producers right has created a lot of discussions. It has been said that rather than creating a new data producers rights purely for the reasons of trusting in big data, we should trust in our existing toolkit which offers trade secret protection, contract and technological protection as securing the exclusivity.[195] Interesting view was also pointed out by arguing that in a global scene by introducing such rights it might actually lead to anticompetitive effects where situations might occur where a European data user is obligated to buy a licence in order to access the data while it would be freely available to US competitors in their territory.[196] It has also been said that it would have disruptive overlaps within intellectual property law.[197] As an example by aggregating data from the stock market in the database could be awarded the protection of Sui genesis database right and data producers right since data is recorded automatically by computer stock exchange.

As we have seen through case law, challenges that data-driven innovation and web-scraping have concerning access to data can concern in principle any non-personal or anonymised data. Therefore too general and broad approach could then make any data mentioned above subject to regulated access. The reality is that generating and creating data has value. It is therefore also an asset to companies. Regulating access should be examined only for particular categories of data. Even though data is reusable and it doesn't diminish its value or quality in general view. Granting exclusivity of access to this specific data may give a competitive advantage when players are in the same market. It is clear that by allowing an exclusivity on one actor on the market, can provide this actor with a competitive advantage, especially when companies are in the same market. CJEU jurisprudence[198] has also pointed into situations when re-use of data is done by actors that are not in the same market as the data-owner. However, access to this kind

[195] Drexl, J. (2017) *Designing Competitive Markets for Industrial Data – Between Propertization and Access.* 8 JIPITEC 257 para 1.
[196] P. Bernt Hugenholtz. P. Bernt. Data Property: Unwelcome Guest in the House of IP. Institute for Information Law (IViR). Available https://www.ivir.nl/publicaties/download/Data_property_Muenster.pdf. [Accessed 23 Oct. 2019].
[197] Drexl, J. (2017) *Designing Competitive Markets for Industrial Data – Between Propertization and Access.* 8 JIPITEC 257 para 1.
[198] See section 3 of European Commission, (2017). *Commission Staff Working Document on the free flow of data and emerging issues of the European data economy.* Available at https://eur-lex.europa.eu/legal-content/LV/TXT/?uri=CELEX%3A52017SC0002. [Accessed 23 June 2019].

of data which is held commercially and is essential to market entry or to spur competition could be implemented as an obligation on data holders to licence the use of the data.

It would not be the first time in EU legal sphere when an obligation to license would be issued. This has already been given in payment services in the internal market. Directive 2015/2366[199] provides an obligation to provide certain information, which is held by banks, on the basis of objective, non-discriminatory and proportionate, in order to provide market access for novel payment solutions to consumers and retailers, and this consequently lowers charges for payments.

In the absence of such methods, competition law can reach out to protect the side of the infringed one. Reaching out to situations when company conducting web-scraping is being denied to access data and therefore cannot compete with its rivals or enter the market. This balancing act that competition law affords to the market and consumers as well as companies as market players, creates a legal protection within competition law for web-scrapers that are collecting data from publicly accessible sources and acting in good faith. This legal protection also honours the function and original intent of the internet as an open place built on sharing. History has proven that through sharing information between others, several of the great insights of our age were discovered, new rights are formulated and enforced, and not only companies but people throughout the world can share their ideas and be connected.

Chapter Five - Conclusion

For a successful future, the digital economy creates a central driver to innovations, efficiencies, and consumer welfare. It has revolutionised business structures, services, goods, communications, and social synergies. Digitalisation has also spurred a change in market dynamics, paving the way for the rise of essential platforms, networks, and the proliferation of multi-sided markets.[200] A growing data-driven economy is necessary for innovation, growth,

[199] Articles 35, 36 of Directive (EU) 2015/2366 of 25 November 2015 on payment services in the internal market, amending Directives 2002/65/EC, 2009/110/EC and 2013/36/EU and Regulation (EU) No 1093/2010, and repealing Directive 2007/64/EC, OJ L 337, 23.12.2015, p. 35; access is to be given to "payment services and "credit institutions 'payment accounts services"

[200] Ezrachi, A. (2018). *EU Competition Law Goals and The Digital Economy.* Available at https://www.academia.edu/36790134/Ezrachi_-_Goals_and_the_digital_economy_-_Working_paper.pdf. [Accessed 12 Jun. 2019].

and European competitiveness as well as for an effective Digital Single Market. The Communication "Towards a data-driven economy"[201] offers a concept for the data-driven economy as an ecosystem with different types of players such as data providers, data analytics companies, skilled data and software professionals, cloud service providers, driving to more business possibilities.[202] The availability of better quality, reliable, and interoperable datasets was particularly pointed as a critical enabler for novel data products.

The Commission is also enlarging merger control to innovation-related assets like data and intellectual property, as mergers of free service providers, risk escaping turnover thresholds.[204] The Commission does seem certain about the situation that it can position innovation under the current status quo of competition law.[205] Innovation in itself continues to be an open-ended notion. In principle, it comprises the commercialisation of recently invented or improved products, which can be classified as product innovations, or production and distribution processes, defined as process innovations. Innovations like these can take place in a most disruptive fashion, generating a completely new market for a new product or can be fairly incremental, which could change the whole dynamics on an existing market.[209] Therefore, no matter what classification of innovation one adheres to, it has the potential to shape markets or to at least switch the competitive dynamics within those markets. In the highly developing digital era, where data has been classified as a ”new oil”[211], web-scraping can be seen as an innovation which can affect the competitive dynamics in the market.

Therefore web-scraping has reached a level where competition law has a reason to intervene. In a situation of ”when all else fails”, when an intellectual property right does not provide sufficient protection, companies can rely on contract law. But also when contract law is used

[201] European Commission, (2014). *Communication from the commission to the European Parliament, The Council, The European Economic and Social Committee and The Committee of The Regions.* Available at http://ec.europa.eu/transparency/regdoc/rep/1/2014/EN/1-2014-442-EN-F1-1.Pdf. [Accessed 19 June 2019].
[202] European Commission, (2017). *Commission Staff Working Document on the free flow of data and emerging issues of the European data economy.* Available at https://eur-lex.europa.eu/legal-content/LV/TXT/?uri=CELEX%3A52017SC0002. [Accessed 23 June 2019].
[204] EC Press release, (2016). *Mergers: Commission seeks feedback on certain aspects of EU merger control.* Available at http://europa.eu/rapid/press-release_IP-16-3337_en.htm. [Accessed 26 June 2019].
[205] European Commission, (2016). *Competition policy brief: Eu merger control and Innovation.* Available at http://ec.europa.eu/competition/publications/cpb/2016/2016_001_en.pdf. [Accessed 2 Nov. 2019].
[209] Ibid.
[211] Neelie Kroes Vice-President of the European Commission responsible for the Digital Agenda Digital Agenda and Open Data From Crisis of Trust to Open Governing Bratislava. (2012). [online] Available at: https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_12_149. [Accessed 20 Dec. 2019].

in an abusive manner in view of another company so that it might face a fear of foreclosure, competition law is the one that these companies can rely on. At least up to a certain point when the practice of using innovative measures of companies own assets doesn't receive the status of competing on mere merits, but with an anticompetitive ways. Web-scraping can be justified by the fact that it benefits the market and generates competition which then again benefits the consumers, and that is the aim of competition law. Owners of copyright might ague that intellectual property law should not be engaged but that the balance set by it should be respected. To the extent where one would scrape copyrighted material lets say e.q. a database protected content, competition law should not intervene and justify this act. Since it would create overlapping legislation and diminish the value that intellectual property law grants to authors. It is part of the protection of intellectual property that it grants a right to authorial work which cannot be duplicated without the authorisation of the owner. However, when the data does not reach an intellectual protection as such, discretion should be used when considering and weighing the need to compensate for the efforts and investments that were required from the creator of the database with the interest of third parties, possible scrapers, to be allowed with access to this database and its information. There should be a balance between protecting the investment and work efforts but also to protect consumers and grant market access

There aren't really clear rules or a piece of legislation that sets the game rule for companies conducting web-scraping. Case law has however provided us with some guidelines that we can follow on and come to a conclusion that when the protection of intellectual property reaches its limits, contract law has a way of extending this protection. This issue steps in the front yard of competition law, when contract law is used on fraudulent purposes it has an obligation to intervene and alter that contract or forbid it completely.[212] Also, the finding on a paradox in *Ryanair*[214] case which indicated that databases without the protection of intellectual property might receive greater protection than the one which is granted by intellectual property law, namely with contract law. Companies should keep this in mind, not only those who own the website but those as well scraping the data.

[212] *hiQ Labs v. LinkedIn* [2017] (U.S. District).
[214] *Case C-30/14 Ryanair Ltd v PR Aviation BV* [2015] ECJ

Bibliography

Legislation:

- Computer Fraud and Abuse Act 1986
- Consolidated versions of the Treaty on European Union and the Treaty on the Functioning of the European Union (2012) C 326/01
- Enforcement Directive (2004) 48/EC
- Consolidated version of the Treaty on the Functioning of the European Union Article 101 (ex Article 81 TEC)
- Consolidated version of the Treaty on the Functioning of the European Union Article 102 (ex Article 82 TEC)
- Council Regulation (EC) No 139/2004 of 20 January 2004 on the control of concentrations between undertakings (the EC Merger Regulation)
- Guidelines on the assessment of horizontal mergers under the Council Regulation on the control of concentrations between undertakings (2004) C 31/03) para.79.
- European Commission, (2011). 'Guidelines on the applicability of Article 101 of the Treaty on the Functioning of the European Union to horizontal co-operation agreements'. OJ C 11/1, paras 119-122.
- Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society.
- Directive 2008/95/EC of the European Parliament and of the Council of 22 October 2008 to approximate the laws of the Member States relating to trade marks.
- Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of database
- Directive (EU) 2015/2366 of 25 November 2015 on payment services in the internal market, amending Directives 2002/65/EC, 2009/110/EC and 2013/36/EU and Regulation (EU) No 1093/2010, and repealing Directive 2007/64/EC, OJ L 337, 23.12.2015,

Case Law:

- *hiQ Labs v. LinkedIn* [2017] (U.S. District).
- *Field v. Google, Inc.* [2006] (United States District Court, D. Nevada.).
- *Case C-30/14 Ryanair Ltd v PR Aviation BV* [2015] ECJ
- *Dianping v Aibang* [2011] the Beijing Haidian District Court.
- *Ticketmaster Corp. v. Tickets.com, Inc.*, [2003] C.D. Cal.
- Facebook, Inc. v. Power Ventures, Inc. [2009] Northern Dist. of Cal.
- Case 322/81, Michelin v. Commission, C:1983:313, ECJ
- Case C-209/10 Post Danmark I v. Konkurrencerådet , ECJ
- Verizon Communications, Inc. v. Law Offices of Curtis V. Trinko, 540 U.S. 398, 407 (2004)
- Commission Decision of 28 January 2015 under Article 6 (1) (b) in conjunction with Article 6 (2) of Council Regulation No 139/2004 (Case No COMP / M.7275 - Novartis / GlaxoSmithKline Oncology Business).
- In re Zappos.com, Inc., Customer Data Sec. Breach Litig., 893 F. Supp. 2d 1058 (D. Nev. 2012).
- *Tencent v Toutiao* Court of Beijing
- *Viaggiare S.r.l. vs Ryanair Ltd.* Decision of June 4, 2013 of the Court of Milan.

Official documents and opinions:

- Ramirez, E. (2016). *Keynote Remarks of 43rd Annual Conference on International Antitrust Law and Policy, Fordham Competition Law Institute: Deconstructing the Antitrust Implications of Big Data.* Available at: https://www.ftc.gov/system/files/documents/public_statements/1000913/ramirez_fordham_speech_2016.pdf [Accessed 15 Apr. 2019].

- Vestager, M. (2016). In: *Conference on Big Data*. [online] Available at: https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/big-data-and-competition_en [Accessed 15 Apr. 2019].
- European Commission- press release (2017). *Antitrust: Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service.* [online] Available at: http://europa.eu/rapid/press-release_IP-17-1784_en.htm [Accessed 15 Apr. 2019].
- European Commission (2016). *Antitrust: Commission send Statement of Objections to Google on Android operating system and applications. European Commission Press Release.* [online] Available at: https://ec.europa.eu/commission/presscorner/detail/en/MEMO_16_1484 [Accessed 15 Apr. 2019].
- European Commission (2012). *Statement of VP Almunia on the Google antitrust investigation*. [online] Available at: http://europa.eu/rapid/press-release_SPEECH-12-372_en.htm [Accessed 20 Dec. 2019].
- European Commission (2015). *Memo: Antitrust: Commission sends Statement of Objections to Google on comparison shopping service*. [online] Available at: https://ec.europa.eu/commission/presscorner/detail/en/MEMO_15_4781m [Accessed 16 Apr. 2019].
- COMMISSION DECISION of relating to a proceeding under Article 82 of the EC Treaty and Article 54 of the EEA Agreement (COMP/C-3 /37.990 - Intel) para. 1616
- The benefit of competition policy for consumers. (2014). [online] United Nations Conference on Trade and Development. Available at: https://unctad.org/meetings/en/SessionalDocuments/ciclpd27_en.pdf [Accessed 16 Jun. 2019].
- Report on the impact of hard core cartels and sanctions against cartels under national competition law. (2002). OECD. Available at: https://www.oecd.org/competition/cartels/1841891.pdf [Accessed 31 Oct. 2019].
- Vestager, M. (2016). *Competition: The mother of invention*. speech delivered at the European Competition and Consumer Day. [online] Available at:

https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/competition-mother-invention_en [Accessed 2 Nov. 2019].

- European Commission, (2016). *Competition policy brief: Eu merger control and Innovation.* Available at http://ec.europa.eu/competition/publications/cpb/2016/2016_001_en.pdf. [Accessed 2 Nov. 2019].
- Kroes, N. (2015). Delivering Better Markets and Better Choices'. Speech delivered at the European Consumer and Competition Day. European Competition Policy. Available at: https://goo.gl/3Fn75o [Accessed 24 Jun. 2019].
- Ezrachi, A. (2019). *The Goals of EU Competition Law and the Digital Economy*. [ebook] BEUC Discussion Paper. Available at: https://www.beuc.eu/publications/beuc-x-2018-071_goals_of_eu_competition_law_and_digital_economy.pdf [Accessed 18 Dec. 2019].
- Lowe, P. (2006). *Preserving and Promoting Competition : a European Response*. St Gallen Competition Law Forum. [online] European Commission. Available at: https://ec.europa.eu/competition/speeches/text/sp2006_018_en.pdf [Accessed 19 Jun. 2019].
- Report on Regulatory Reform Synthesis. (1997). [online] OECD. Available at: https://www.oecd.org/gov/regulatory-policy/2391768.pdf [Accessed 18 Jun. 2019].
- EC Press release, (2017). *Mergers: Commission clears merger between Dow and DuPont, subject to conditions.* Available at http://europa.eu/rapid/press-release_IP-17-772_en.htm. [Accessed 26 June 2019].
- European Commission, (2014). *Communication from the commission to the European Parliament, The Council, The European Economic and Social Committee and The Committee of The Regions.* Available at http://ec.europa.eu/transparency/regdoc/rep/1/2014/EN/1-2014-442-EN-F1-1.Pdf. [Accessed 19 June 2019].
- European Commission, (2017). *Commission Staff Working Document on the free flow of data and emerging issues of the European data economy.* Available at https://eur-lex.europa.eu/legal-content/LV/TXT/?uri=CELEX%3A52017SC0002. [Accessed 23 June 2019].

- EC Press release, (2016). *Mergers: Commission seeks feedback on certain aspects of EU merger control.* Available at http://europa.eu/rapid/press-release_IP-16-3337_en.htm. [Accessed 26 June 2019].
- Neelie Kroes Vice-President of the European Commission responsible for the Digital Agenda Digital Agenda and Open Data From Crisis of Trust to Open Governing Bratislava. (2012). [online] Available at: https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_12_149. [Accessed 20 Dec. 2019].
- Vestager, M. (2016). *Competition In A Big Data World*. European Commission. Available at: https://dld-media.com/events/opening-keynote-e5a8fd5c-c210-44f4-b93c-882669f06688 [Accessed 11 Nov. 2019].
- EC Report by Crémer, J. de Montjoye, Y. Schweitzer, H. (2019). *Competition Policy for the Digital Era.* Available at http://ec.europa.eu/competition/publications/reports/kd0419345enn.pdf. [Accessed 12 July 2019]
- Competition Bureau, (2017). *Big data and Innovation*. Available at https://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/04304.html. [Accessed 3 July 2019]
- Speech of former Competition Commissioner Almunia. (2012). *Statement of Commissioner Almunia on the Google antitrust investigation.* Press room Brussels, 21 May 2012, SPEECH/12/372. Available at http://europa.eu/rapid/press-release_SPEECH-12-372_en.htm. [Accessed 11 Jun. 2019].
- Bosch, O. (2017). *An Introduction to Web Scraping, IT and Legal Aspects.* European Commission.
- European Commission, (2017). *Communication From The Commission To The European Parliament, The Council, The European Economic and Social Committee and The Committee of The Regions.* Available at https://ec.europa.eu/transparency/regdoc/rep/1/2017/EN/COM-2017-9-F1-EN-MAIN-PART-1.PDF. [Accessed 31 Aug. 2019].

- European Commission (2015). *Antitrust: Commission sends Statement of Objections to Google on comparison shopping service', MEMO/15/4781*. [online] Available at: https://ec.europa.eu/commission/presscorner/detail/en/MEMO_15_4781 [Accessed 19 Oct. 2019].

Books, articles and websites:

- Mitchell, R. (2015). Web Scraping with Python: Collecting Data From The Modern Web. [online] Available at: https://yanfei.site/docs/dpsa/references/PyWebScrapingBook.pdf [Accessed 28 Aug. 2019].
- F.Din, M. (2015). Breaching and Entering: When Data ScrapingShould Be a Federal Computer Hacking Crime. 81st ed. Brooklyn Law Review, pp.405-440.
- Data Use Cases, Gain a Competitive Edge With Web Scraped Product Data. (2019). [ebook] Scrapinghub Inc. Available at: https://cdn2.hubspot.net/hubfs/4367560/Case%20Studies/Data%20Use%20Case%20-%20Product%20Data.pdf. [Accessed 28 Aug. 2019].
- Zhao, Bo. (2017) *Web-Scraping*. Springer International Publishing AG (outside the USA)
- L.A. Schintler, C.L. McNeely (eds.), Encyclopedia of Big Data.

- Vivien Chan & Co. Newsletter issue 03. (2018). [ebook] Available at: http://www.vcclawservices.com/sources/publications/2018issue03.pdf [Accessed 1 May 2019].
- Laney, D. (2015). *Gartner Predicts Three Big Data Trends for Business Intelligence*. [online] Forbes.com. Available at: https://www.forbes.com/sites/gartnergroup/2015/02/12/gartner-predicts-three-big-data-trends-for-business-intelligence/#6b3197fa6de4 [Accessed 11 Apr. 2019].
- Ibáñez Colomo, Pablo (2016). *Restrictions on innovation in EU competition law*. European Law Review, 41 (2). pp. 201-219. ISSN 0307-5400 p.16
- Hutchinson, T. (2015). The Doctrinal Method: Incorporating Interdisciplinary Methods in Reforming the Law. 3 Erasmus Law Review 130.

- Lamarid, A. and Villiers, S. (2017). *Big Data Privacy and Competition Law: Do Competition Authorities know how to do it?*. [ebook] Available at: https://www.competitionpolicyinternational.com/wp-content/uploads/2017/01/CPI-Lamadrid-Villiers.pdf [Accessed 17 Apr. 2019].
- Haucap, J., Stiebale, J. and Raschand, A. (2019). *How mergers affect innovation: theory and evidence from the pharmaceutical industry.* 63rd ed. p.284.
- Alford, R. (2018). *The Role of Antitrust in Promoting Innovation*. [online] Speech delivered at Kings College London. Available at: https://goo.gl/kcqtMQ [Accessed 19 Jun. 2019].
- Baregheh, J. Rowley & S. Sambrook, (2009). *Towards a Multidisciplinary Definition of Innovation*. 47 Management Decision 1323.
- Lincoln, (1953). *Lecture on Discoveries and Inventions*, in R. Baslter(ed) Collected Works of Abraham Lincoln, Vol 3, Rutgers. 363.
- D. Encaoua & A. Hollander (2002). *Competition Policy and Innovation.* Oxford Review of Economic Policy. 18. page. 63-79.
- Oecd.org. (2019). *Digital Economy, Innovation and Competition - OECD*. [online] Available at: http://www.oecd.org/competition/digital-economy-innovation-and-competition.htm [Accessed 23 Jun. 2019].
- Bundeskartellamt (2017). *Innovations-Challenges for competition law practice*. [ebook] Available at: https://www.bundeskartellamt.de/SharedDocs/Publikation/EN/Schriftenreihe_Digitales_II.pdf?__blob=publicationFile&v=3. [Accessed 14 Jun. 2019].
- Costa-Cabral, F. (2017) *Innovation in EU competition law: The resource-based view and disruption.* Jean Monnet Working Paper 2/17 Available https://jeanmonnetprogram.org/wp-content/uploads/JMWP-02-Costa-Cabral.pdf Accessed 14 Oct. 2019.
- Wu, T. (2012) *Taking Innovation Seriously: Antitrust Enforcement if Innovation Mattered Most*, 78 Antitrust Law Journal 328 page 313-328. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2166525.[Accessed 25 June 2019]

- Colomo, I. (2016). *Restrictions on Innovation in EU Competition Law*, 41 European Law Review. Available at https://www.lse.ac.uk/collections/law/wps/WPS2015-22_Colomo.pdf. [Accessed 26 June 2019]
- Drexl, J. (2012). *Anti-competitive stumbling stones on the way to a cleaner world: protecting competition in innovation without a market.* Journal of Competition Law and Economics, Vol. 8, Issue 3. Page. 507-542.
- Schrepel, T. (2017). *Predatory Innovation: The Definite Need for Legal Recognition.* 21 SMU Sci. & Tech. L.Rev. 19.
- Ezrachi, A. (2018). *EU Competition Law Goals and The Digital Economy.* Available at https://www.academia.edu/36790134/Ezrachi_-_Goals_and_the_digital_economy_-_Working_paper.pdf. [Accessed 12 Jun. 2019].
- Larouche, P. and Schinkel, M. (2013). *Continental Drift in the Treatment of Dominant Firms: Article 102 TFEU in Contrast to § 2 Sherman Act.* TILEC Discussion Paper No. 2013-020 Available at https://ssrn.com/abstract=2293141. [Accessed 19 Jun. 2019].
- Nihoul, P. & Van Cleynenbreugel, P. (2018) *The roles of Innovation in Competition Law Analysis*. ASCOLA Competition Law series. Edward Elgar.
- G. M. Swann, P. (2009). *The Economics of Innovation: An introduction*. Cheltenham, Edward Elgar Publishing, page. 8–20.
- Schumpeter, J (2010). *Capitalism, Socialism and Democracy.* Routledge, Abingdon, page. 82–83
- Arrow, K. Economic Welfare and the Allocation of Resources for Invention, RAND Paper 1959, available at https://www.rand.org/pubs/papers/ P1856.html, page. 15–20. [Accessed 14 June 2019].
- Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P. (2002), *Competition and Innovation: An Inverted U Relationship*, NBER Working Papers 9269.
- Aghion, P., Griffith, R. (2005), *Competition and Growth: Reconciling Theory and Evidence*, Cambridge, Mass.: MIT Press.

- Shapiro, C. (2012). *Competition and Innovation. Did Arrow Hit the Bull's Eye?*, in Josh Lerner and Scott Stern, The Rate and Direction of Inventive Activity Revisited. NBER, 361–410.
- Auer, D. (2017). *Appropriability and the European Commission's Android Investigation*. Columbia Journal of European Law 23, page 647–680.
- Colomo, P. *Competition Law and Innovation: Where do we stand?* Available at https://academic.oup.com/jeclap/article/9/9/561/5234221 [Accessed 12 Jun. 2019].
- Streel & Larouche (2015). *Disruptive Innovation and Competition Policy Enforcement*, TILEC Discussion Paper 2015-21 5.
- Baker, J. *Beyond Schumpeter vs. Arrow: How Antitrust Fosters Innovation* (2007). Articles in Law Reviews & Other Academic Journals. Paper 276
- Autorité de la concurrence and Bunderkartellamt, (2016). *Competition Law and Data*. Available at http://www.autoritedelaconcurrence.fr/doc/reportcompetitionlawanddatafinal.pdf. pages 12-13. Accessed 1 July 2019.
- Facebook. (2019). *Terms of Service*. [online] Available at: https://www.facebook.com/terms.php. [Accessed 11 Jun. 2019].
- Graef, I. (2015). *Market Definition and Market Power in Data: The Case of Online Platforms.* Page 479-480 Available https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2657732. [Accessed 23 Jun. 2019].
- Boeing, G. & Waddell, P. (2016). *New Insights into Rental Housing Markets across the United States:Web Scraping and Analyzing Craigslist Renal Listings*. p.3. Available at:https://arxiv.org/ftp/arxiv/papers/1605/1605.05397.pdf. [Accessed 28 Aug. 2019]
- Reinsel, D. & Gants, J. & Rydning, J. (2017). *Data Age 2025: The Evolution of Data to Life-Critical.* An IDC White Paper. Available at https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/workforce/Seagate-WP-DataAge2025-March-2017.pdf [Accessed 29 Jun. 2019].

- Picardo, E (2019). *10 of the Worlds Top Companies Are American*. Available at https://www.investopedia.com/articles/active-trading/111115/why-all-worlds-top-10-companies-are-american.asp. [Accessed 30 Aug. 2019].

- Webscrapingexpert.com. (2019). *Web Data Scraping, Website Data Scraping, Web Page Scraping Services | Web Scraping Expert*. [online] Available at: http://www.webscrapingexpert.com/web-data-scraping/ [Accessed 27 Oct. 2019].
- Scraperapi.com. (2019). *What kind of companies use web scrapers? | Scraper API*. [online] Available at: https://www.scraperapi.com/blog/what-kind-of-companies-use-web-scrapers [Accessed 27 Oct. 2019]
- Scrapy providing information on its clients and how they have provided services. Scrapy.org. (2019). *Scrapy | Companies using Scrapy*. [online] Available at: https://scrapy.org/companies/ [Accessed 27 Oct. 2019].
- Yang, J. & Yecies, B. (2016). *Mining Chinese social media UGC: a big-data framework for analyzing Douban movie reviews.* Journal of Big Data 3.
- Huan, L. (2016). *The good the bad, and the ugly: uncovering novel research opportunities in social media mining.* Springer International Publishing Switzerland
- Raulamo-Jurvanen. & Kakkonen, K. & Mäntylä, M. (2016). *Using Surveys and Web-Scraping to Select Tools for Software Testing Consultancy*. M35, ITEE, University of Oulu.
- Nourma, P. & Budiarti, R. & Widyatmoko, N. & Hariadi, M. & Hery Purnomo, M. (2016). *Web scraping for automated water quality monitoring system: A case study of PDAM Surabaya.* pages. 641-648. Conference paper.
- Ferrara, E. & De Meo, P. & Fiumara, G. & Baumgartner, R. (2014). *Web Data Extraction, Applications and Techniques: A Survey.*
- Rezai, A. (2017). *Beware of the Spiders: Web Crawling and Screen Scraping the Legal Position*. [online] Paris Smith. Available at: https://parissmith.co.uk/blog/web-crawling-screen-scraping-legal-position/ [Accessed 16 Oct. 2019].

  Businesses that use 'data-driven decision-making' enjoy a 5-6% increase in productivity,Tene, O. and Polonetsky, J. (2012). Big Data for All: Privacy and User Control

in the Age of Analytics. *11 Northwestern Journal of Technology and Intellectual Property 239 (2013)*

- SIIA- Software and Information Industry Association (2013). *Data-Driven InnovationA Guide for Policymakers: Understanding and Enabling the Economic and Social Value of Data*. [online] Available at: https://www.siia.net/Portals/0/ppt/Policy/Data%20Driven%20Innovation/data-driven-innovation.pdf?ver=2013-07-08-135946-000 [Accessed 21 Oct. 2019].
- Chan, V. & Co, *Fight against Web-Scraping: Legal Redress*, (2018)
- Zamora, A. *Making Room For Big Data: Web Scraping and an Affirmative Right to Access Publicly Available Information Online*, 12 J. Bus. Entrepreneurship & L.203 (2019) Available https://digitalcommons.pepperdine.edu/jbel/vol12/iss1/8 [Accessed 24 Oct. 2019].
- De Santis, F. (2015). *ECJ clarifies Database Directive scope in screen scraping case*. [online] Globallegalpost.com. Available at: http://www.globallegalpost.com/blogs/global-view/ecj-clarifies-database-directive-scope-in-screen-scraping-case-128701/ [Accessed 25 Oct. 2019].
- O´Rourke, M. (2001). Property Rights and Competition on the Internet: In Search of an Appropriate Analogy. BERKELEY TECH.L.J.
- Economic and Social Benefits of Internet Openness: 2016 Ministerial Meeting on the Digital Economy Background Repor. (2016). [ebook] OECD DIGITAL ECONOMY PAPERS No.257: OECD. Available at: https://www.oecd-ilibrary.org/docserver/5jlwqf2r97g5-en.pdf [Accessed 26 Oct. 2019].
- Easterbrook, F. (1996). *Cyberspace and the Law of the Horse.* University of Chicago Legal Forum
- De Franceschi, A. (2016). European contract law and the Digital Single Market. 1st ed. Cambridge: Intersentia, pp.51-79.
- Drexl, J. (2017) *Designing Competitive Markets for Industrial Data – Between Propertization and Access.* 8 JIPITEC 257 para 1.

• P. Bernt Hugenholtz. P. Bernt. Data Property: Unwelcome Guest in the House of IP. Institute for Information Law (IViR). Available https://www.ivir.nl/publicaties/download/Data_property_Muenster.pdf. [Accessed 23 Oct. 2019].