Risk-Based Approach as a Solution to Secondary Use of Personal Data

Antti Antikainen
University of Helsinki
Master's Thesis
Law and Economics
Faculty of Law
08/2014

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta/Osasto Fakultet/Sektion – Faculty Faculty of Law		Laitos/Institution– Department
Tekijä/Författare – Author Antti Antikainen		
Työn nimi / Arbetets titel – Title Risk-based Approach as a Solution to Secondary Use of Personal Data		
Oppiaine /Läroämne – Subject Law and Economics		
Työn laji/Arbetets art – Level Master's thesis	Aika/Datum – Month and year 08/2014	Sivumäärä/ Sidoantal – Number of pages 92

Tiivistelmä/Referat – Abstract

The research question of this thesis is about the secondary use of data and a risk-based approach to the regulation of data protection. The intention of this thesis is to explore the current regulation of secondary use of data, which means uses of data that are outside the primary purpose for the collection of data. Law and economics is applied to frame and offer regulatory solutions to the research question. In the current changing environment, the right to privacy is at danger. A fundamental rights conflict has emerged between the right to privacy and the fundamental rights of the data controllers. The online economy is built around the use of personal data, also secondary use is widespread. This conflict needs to be solved, since if data is not used there will be substantial welfare losses to the whole society.

The thesis explores the current legislation, mainly the general data protection directive and the EU commission proposal for General data protection regulation. The emphasis is on the concept of purpose restriction and the legality of processing data. The current basis for secondary use of data is the 'legitimate use' article 7(f) of the directive, which is implemented in different Member States statutes. The secondary use of personal data must meet the conditions and fulfil the purpose restriction, this is however problematic since the purpose limitation principle limits the future uses of data, and also repurposing the data is problematic. Anonymization is explored as a current solution for the problem. In the field of anonymization there are two major problems, which are the devaluation of the value of the data if the anonymization is conducted robustly. Also the problem is with re-identification, which means that the anonymization is broken and an individual is found from the data set.

Enforcement is analyzed, since without functioning sanctions there are incentives for using data without complying with the data protection laws. Current level of sanctions is not sufficient. The increasing value of data calls for proper enforcement; however without new legal inventions a too strict regimen of sanctions will cause new problems. Monetary value of sanctions is seen as a highly important part of a functioning system of data protection. Future research would benefit the setting of effective level of sanctions.

The solution the thesis offers for the problematic purpose restriction and legitimacy of data use is based on risk-based regulation. A risk-based solution would allow more data uses while simultaneously protecting the fundamental right to privacy. The proposed model would classify data on the basis of risk the use causes for processing and regulate the different categories. Anonymization is used in certain categories to reduce the risk of processing. With legal inventions the increasing value of data can be harnessed while simultaneously protecting fundamental rights of the data subjects.

Avainsanat – Nyckelord – Keywords
privacy, data protection, law and economics, risk-based, EU, regulatory theory, fundamental rights, sanctions, big data

Säilytyspaikka – Förvaringställe – Where deposited
Helsinki University library

Muita tietoja – Övriga uppgifter – Additional information

Contents

Bibliography

Textbooks and Monographs:

Aarnio 1987

Aarnio, Aulis: The rational as reasonable : a treatise on legal justification. Reidel, Dordrecht, 1987.

Aarnio 1986

Aarnio, Aulis: Lain ja kohtuuden tähden, Werner Söderström, Porvoo 1986.

Baldwin, Cave and Lodge 2012

Baldwin, Robert, Cave, Martin and Lodge, Martin: Understanding regulation: theory, strategy, and practice (2nd ed. edn Oxford University Press, Oxford 2012).

Barnard 2007

Barnard, Catherine, The Substantive Law of the EU : the four freedoms (2nd ed. edn Oxford University Press, Oxford 2007).

Beck 1992

Beck, Ulrich: Risk society: towards a new modernity, Theory, culture & society, Sage, London 1992.

Bennett and Raab 2006

Bennett, Colin J. and Raab, Charles D: The Governance of Privacy (2.th edn The MIT Press, 2006).

Black 2010a

Black, Julia, 'Role of Risk in Regulatory Process' in Baldwin, Robert, Martin Lodge and Martin Cave (editors), Oxford Handbook of Regulation (2010).

Black 2010b

Black, Julia: 'Risk-based Regulation: Choices, Practices and Lessons Being Learnt' in Risk Regulation and Governance Institutions (OECD, 2010).

Bounds 2010

Bounds, Gregory: 'Challenges to Design Regulatory Policy Frameworks to Manage Risk' in Risk and Regulatory Policy Improving the Governance of Risk (OECD, 2010).

Brouwer 2011

Brouwer, Evelien: Legality and Data Protection Law: The Forgotten Purpose of Purpose Limitation, in Besselink, Leonard F. M., Pennings Frans and Prechal Sacha (editors): The Eclipse of the Legality Principle in the European Union, Wolter Kluwer, 2011.

Bygrave 2002

Bygrave, Lee A.: Data protection law: approaching its rationale, logic and limits, Information law series; 10, Kluwer Law International, The Hague 2002.

Cooter and Ulen 2014

Cooter, Robert and Ulen, Thomas: Law and economics, Pearson custom library. 6th ed. edn Pearson, Harlow, Essex 2014.

Craig P. and De Bruca 2011

Craig, Paul. P. and De Burca, Thomas: EU law: text, cases, and materials, 4th ed. edn, Oxford University Press, Oxford 2011.

Craig and Lundloff 2011

Craig,Terence and Lundloff, Mary E.: Big data and Privacy, O'Reilly, Sebastopol 2011.

Duncan, Elliot and Salazar 2011

Duncan, George T., Elliot, Mark and Salazar, Gonzalez Juan-Jose: Statistical confidentiality: principles and practice, Statistics for social and behavioral sciences, Springer, New York 2011.

Fisher 2010

Fisher, Elizabeth Dr: Risk Regulatory Concepts and the Law in Risk and Regulatory Policy - Improving the Governance of Risk (OECD Reviews of Regulatory Reform, OECD, 2010.

Friedman 2000

Friedman, David D.: Law's order: what economics has to do with law and why it matters, Princeton University Press, Princeton, NJ 2000.

Innanen and Saarimäki 2009

Innanen, Antti and Jarkko Saarimäki, Internet-oikeus, Edita Publishing Oy, Helsinki 2009.

Kuner 2013

Kuner, Christopher: Transborder data flows and data privacy law, Oxford University Press, Oxford 2013.

Kuner 2007

Kuner, Christopher: European data protection law: corporate compliance and regulation, end ed. edn, Oxford University Press, Oxford 2007.

Lessig 2006

Lessig, Lawrence: Code: version 2.0, 2nd. ed. edn Basic Books, New York 2006

Lessig 2002

Lessig, Lawrence: The future of ideas : the fate of the commons in a connected world, Vintage Books, New York 2002.

Länsineva 2011

Länsineva, Pekka: Omaisuudensuoja, Perusoikeudet in Hallberg Pekka, Länsineva Pekka, Karapuu Heikki, Ojanen Tuomas, and others, Perusoikeudet, Oikeuden perusteokset, 2., uud. p., Werner Söderström lakitieto WSLT, Helsinki 2011.

Mackaay 1982

Mackaay, Ejan: Economics of information and law, Kluwer Nijhoff, Boston, MA 1982.

Mayer-Scönberger and Cuckier 2013

Mayer-Schönberger, Viktor and Cukier Kenneth: Big data: a revolution that will transform how we live, work, and think, Houghton Mifflin Harcourt, Boston, Ma 2013.

Mercuro and Medema, 1997

Mercuro, Nicholas and Medema Steven G: Economics and the law : from Posner to post-modernism (Princeton University Press, Princeton, N.J 1997).

Millard 2013

Millard, Christopher, Cloud computing law, Oxford University Press, New York, NY, 2013.

Ojanen 2010

Ojanen, Tuomas: EU-oikeuden perusteita, Edita, Helsinki 2010.

Posner 1998

Posner, Richard A.: Economic analysis of law, 5. ed. edn, Aspen Law & Business, New York cop. 1998.

Power 2004

Power, Michael: The risk management of everything: rethinking the politics of uncertainty , Demos, London 2004.

Pöyhönen 2003

Pöyhönen, Juha: Uusi varallisuusoikeus, 2. p. edn Talentum, Helsinki 2003.

Pöysti, 1999

Pöysti, Tuomas: Tehokkuus, informaatio ja eurooppalainen oikeusalue, Helsingin yliopiston oikeustieteellisen tiedekunnan julkaisut, Helsingin yliopisto, oikeustieteellinen tiedekunta, Helsinki 1999.

Rule, 2007

Rule, James B: Privacy in peril, Oxford University Press, Oxford 2007.

Saarenpää, 2009a

Saarenpää, Ahti, 'Henkilö- ja persoonallisuusoikeus' in Martti, Minna-Greta (ed), Oikeusjärjestys. Osa 1, Lapin yliopiston oikeustieteellisiä julkaisuja. Sarja C; 52. 6. täyd. p. edn Lapin yliopisto, Rovaniemi 2009.

Saarenpää, 2009b

Saarenpää, Ahti, 'Oikeusinformatiikka' in Martti, Minna-Greta (ed), Oikeusjärjestys. Osa 1, Lapin yliopiston oikeustieteellisiä julkaisuja. Sarja C; 52., 6. täyd. p. edn Lapin yliopisto, Rovaniemi 2009.

Seipel, 2004

Seipel, Peter, Juridik och IT: introduktion till rättsinformatiken, 8., omarb. uppl. edn, Norstedts juridik, Stockholm 2004.

Seipel, 1977

Seipel, Peter, Computing law: perspectives on a new legal discipline, Liber, Stockholm 1977.

Shapiro and Varian, 1999

Shapiro, Carl and Varian Hal R: Information rules: a strategic guide to the network economy, Harvard Business School Press, Boston, MA 1999.

Siegel 2013

Siegel, Eric, Predictive analytics : the power to predict who will click, buy, lie, or die, Wiley, Hoboken, N.J, 2013.

Siltala 2013

Siltala, Raimo: Tutkijan tanssiaskeleet eli Oikeustieteellisen tutkimuksen Koreografia, Husa, Jaakko, and others (eds): Oikeuden avantgarde: juhlajulkaisu Juha Karhu 1953-6/4-2013, Talentum, Helsinki 2013.

Siltala 2011

Siltala, Raimo, Law, Truth, and Reason: A Treatise on Legal Argumentation, Law and Philosophy Library; 97, Springer, Dordrecht 2011.

Siltala 2003

Siltala, Raimo, Oikeustieteen tieteenteoria, Suomalaisen Lakimiesyhdistyksen julkaisuja. n:o 234, Suomalainen Lakimiesyhdistys, Helsinki 2003.

Solove and Schwartz 200,

Solove, Daniel J. and Paul M. Schwartz, Privacy, information, and technology, 2nd ed. edn, Aspen Publishers, New York 2009.

Solove 2008

Solove, Daniel J., Understanding Privacy, Harvard University Press, 2008.

Taleb 2008

Taleb, Nassim Nicholas: The black swan: the impact of the highly improbable. Penguin Books, London 2008.

Tolonen 2003

Tolonen, Hannu: Oikeuslähdeoppi. WSOY lakitieto, Helsinki 2003.

Vanto 2011,

Vanto, Jarno, Henkilötietolaki käytännössä. WSOYpro, Helsinki 2011.

Viljanen 2011

Viljanen, Veli-Pekka_ Perusoikeuksien rajoittaminen, Hallberg Pekka, Viljanen, Veli-Pekka, Länsineva Pekka, Karapuu Heikki, Ojanen Tuomas, and others: Perusoikeudet, Oikeuden perusteokset, 2., uud. p., Werner Söderström lakitieto WSLT, Helsinki 2011.

Voutilainen 2012

Voutilainen, Tomi: Oikeus tietoon: informaatio-oikeuden perusteet, Edita, Helsinki 2012.

Scholarly Articles & Conference Proceedings

Acquisti 2010

Alessandro, Acquisti: The Economics of Personal Data and the Economics of Privacy, Joint WPISP-WPIE Roundtable Background Paper 3.

Acquisti, Leslie and Loewenstein 2009

Acquisti, Alessandro, John Leslie, Loewenstein George: What is Privacy Worth?, Tweny First Workshop on Information Systems and Economics WISE, Phoenix, AZ 14-15, 2009.

Barth-Jones 2012

Barth-Jones, Daniel C., 'The "Re-identification" of Governor William Weld's Medical Information: A Critical Re-examination of Health Data Identification Risks and Privacy Protections, Then and Now' Pre-Publication Draft, 2012.

Black and Baldwin 2010

Black, Julia and Baldwin, Robert: Really Responsive Risk-Based Regulation, Law & Policy, Volume 32, Issue 2, April 2010.

Brinhack and Elkin-Koren 2011

Brinhack, Michael and Niva Elkin-Koren, 'Does Law Matter Online? Empirical Evidence on Privacy Law Compliance', Michigan Telecommunications and Technology Review, Volume 17, Issue 2. 337, 2011.

Bräutigam 2010

Bräutigam, Tobias: Getting High on Information? The European Commission's Proposal for Renewal of the Data Protection Legislation. JFT 5/2012.

Calzolari and Pavan 2006

Calzolari, Giacomo and Pavan, Alessandro: On the Optimality of Privacy in Sequential Contracting, Journal of Economic Theory, Volume 130, No. 1, 168, 2006.

Coase 1960

Coase, R. H.: Problem of Social Cost, The Journal of Law and Economics, Volume 3, 1960

Narayana and Shamatikov 2008

Narayanan, Arvind and Vitaly Shmatikov: Robust De-anonymization of Large Sparse Datasets, 2008 IEEE Symposium on Security and Privacy, 2008.

Muth 2009

Muth Karl T.: Googlestroika, Duquesne Law Review, Volume 47, Number 2, 337, Spring 2009.

Mähönen 2004

Mähönen, Jukka, 'Taloustiede lain tulkinnassa, p. *49*, *Lakimies 2004*/*1*.

Ohm 2009 - 2010

Ohm, Paul, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' (2009-2010) UCLA Law Review, Volume 57, 2009 - 2010.

Ohm 2013,

Ohm, Paul, The Underwhelming Benefits of Big Data, University of Pennsylvania Law Review Online, Volume 161, 2013.

Oker-Blom 2009

Oker-Blom, Max: Oikeustaloustieteen eli taloudellisten argumenttien merkityksestä Raimo Siltalan oikeuslähdeopissa, Teoksessa Oikeus ja kritiikki, Edilex, 2009.

Posner 1978

Posner, Richard A.: Economic Theory of Privacy, Regulation, May/June, 1978.

Posner 1977-1978

Posner, Richard A.: The Right of Privacy, 12 Georgia Law Review *393*, 1977-1978.

Schwartz 2003 - 2004

Schwartz, Paul M.: Property, Privacy, and Personal Data, 117 Harvard Law Review, 2056, 2003-2004

Schwartz 2012 - 2013

Schwartz, Paul M.: Information Privacy in the Cloud, University of Pennsylvania Law Review, Vol. 161, No. 1623, 2013.

Solove 2012 - 2013

Solove, Daniel J. :'Privacy Self-Management and the Consent Dilemma ' (2012-2013) 126 Harvard Law Review, 2012 - 2013.

Solove and Hartzog 2014,

Solove, Daniel J. and Woodrow Hartzog: The FTC and The New Common Law of Privacy, 114 Columbia Law Review 583, 2014.

Sweeney 2002

Sweeney, Latanya: k-anonymity: a model for protecting privacy, International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 10 (5), 2002; 557-570.

Warren and Brandeis 1890-1891

Warren, Samuel D. and Louis D. Brandeis: Right to Privacy, Harvard Law Review, Volume 4, 193, 1890-1891.

Wu 2013

Wu, Felix T: Defining Privacy and Utility in Data Sets, University of Colorado Law Review, Volume 84, 1117, 2013.

Yakowitz 2011-2012

Yakowitz Jane: Tragedy of the Data Commons, Harvard Journal of Law & Technology, Volume 25, Number 1, fall 2011.

Online resources & Newspapers (Checked on 20.8.2014)

Barbaro and Zeller 2006

Barbaro, Michael and Zeller Tom Jr: A Face is Exposed for AOL Searcher No. 4417749, New York Times,09.08.2006, http://www.nytimes.com/2006/08/09/technology/09aol.html?pagewanted=all&_r=0.

Gartner 2013

Gartner: 'Big data' http://www.gartner.com/it-glossary/big-data/, 2013.

Harford 2014

Harford, Tim: Big data: are we making a big mistake, Financial Times, 28.3.2014. http://www.ft.com/intl/cms/s/2/21a6e7d8-b479-11e3-a09a-00144feabdc0.html#axzz3ApnS6FyJ.

Innocenzio 2014

Innocenzio, Anne: Target CEO Resigns Amid Fallout From Massive Data Breach', Huffington Post, 05.05.2014. http://www.huffingtonpost.com/2014/05/05/target-ceo-resigns_n_5266229.html.

Kimball 2013

Kimball, Spencer, 'EU fines Microsoft for antitrust breach, DW, 06.03.2013, http://dw.de/p/17sKh.

Maass 2012

Maass, Peter: Your FTC Privacy Watchdogs: Low-Tech, Defensive, Toothless, Wired, 28.06.2012, http://www.wired.com/2012/06/ftc-fail/all/.

Morran 2013

Morran, Chris: 4 Ways Retail Stores Are Monitoring Your Every Move, Consumerist 27.03.2013, http://consumerist.com/2013/03/27/4-ways-retail-stores-are-monitoring-your-every-move/.

OECD Statistics Glossary

Data Utility, 2005, OECD http://stats.oecd.org/glossary/detail.asp?ID=6905.

Palmer 2006

Palmer, Michael. 'Data is the New Oil' Ana Marketing Maestros Blog, 3.11.2006 http://ana.blogs.com/maestros/2006/11/data_is_the_new.html.

Popescu 2013

Popescu, Adam, 'The Next Wave of Ads Knows Everything About You — Before You Do, Mashable, 26.7.2013 http://mashable.com/2013/07/26/inference-advertising/.

Schneier 2007

Schneier, Bruce: 'Why 'Anonymous' Data Sometimes Isn't', Wired, 13.12.2007, http://archive.wired.com/politics/security/commentary/securitymatters/2007/12/securitymatters_1213.

Singer 2012

Singer, Natasha: 'You for Sale; Mapping, and Sharing, the Consumer Genome, 16.6.2012, http://www.nytimes.com/2012/06/17/technology/acxiom-the-quiet-giant-of-consumer-database-marketing.html?pagewanted=all.

Statista 2014

http://www.statista.com/statistics/266206/googles-annual-global-revenue/

Laney 2012

Laney, Douglas: Infonomics: The Practice of Information Economics, Gartner, Forbes, 22.5.2012,

http://www.forbes.com/sites/gartnergroup/2012/05/22/infonomics-the-practice-of-information-economics/.

Lomas 2014

Lomas, Natasha: Facebook Data Privacy Class Action Joined By 11,000 And Counting, TechCrunch, 4.8.2014, http://techcrunch.com/2014/08/04/europe-vs-facebook-class-action/.

Varian 1996

Varian, Hal R. Economic Aspects of Personal Privacy, 1996, http://people.ischool.berkeley.edu/~hal/Papers/privacy/.

Official Material

EUROPEAN UNION

Article 29 Working Party

WP 29 Statement 2014

Article 29 Working Party, 'Statement on the role of a risk-based approach in data protection legal frameworks' in Volume 14/EN WP 218 (2014), Adopted on 30 May 2014.

WP 29 Opinion 05/2014

Article 29 Working Party: Opinion 05/2014 on Anonymisation Techniques, 0829/14/EN WP216, Adopted on 10 April 2014.

WP 29 Opinion 06/2014

Article 29 Working Party: Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, 844/14/EN WP 217, Adopted on 9 April 2014.

WP 29 Opinion 03/2013

Article 29 Working Party: 'Opinion 03/2013 on purpose limitation, 00569/13/EN WP 203, Adopted on 2 April 2013.

WP 29 Opinion 15/2011

Article 29 Working Party: Opinion 15/2011 on the definition of consent, 01197/11/EN WP187, Adopted on 13 July 2011.

WP 29 Opinion 1/2010

Article 29 Working Party: Opinion 1/2010 on the concepts of "controller" and "processor"' in Volume 00264/10/EN WP 169 (2010).

WP 29 Working Document 114, 1995

Article 29 Working Party, 'Working document on a common interpretation of Article 26(1) of Directive 95/46/EC of 24 October 1995, 2093/05/EN WP 114

WP 29 *Protection of Individuals 1999,*

Article 29 Working Party, 'Working Party on the Protection of Individuals with regard to the Processing of Personal data' in Volume 1/99, 1999.

European Union Agency for Fundamental Rights

FRA 2013 Handbook on European data protection law

European Union Agency for Fundamental Rights and Council of Europe: Handbook on European data protection law, Publications Office of the European Union, Belgium, 2013.

FRA 2010 Data Protection in the European Union,

European Union Agency for Fundamental Rights,
Data Protection in the European Union: the role of National Data Protection Authorities 2010

European Commission

Commission MEMO 14/186
Progress on EU data protection reform now irreversible following European Parliament vote, cl MEMO 14/186, 2014.

GDPR
European Commission, Brussels, 25.1.2012 COM(2012) 11 final 2012/0011 (COD)Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation).

Eurobarometer 359
European Commission: Special Eurobarometer 359, Attitudes on Data Protection and Electronic Identity in the European Union, Conducted by TNS Opinion & Social at the request of Directorate-General Justice, Information Society & Media and Joint Research Centre, Survey co-ordinated by Directorate-General Communication, 2011.

GDPR Impact Assessment 2010
European Commission, Brussels, 25.1.2012 SEC(2012) 72 final COMMISSION STAFF WORKING PAPER, Impact Assessment, Accompanying the GDPR.

Commission MEMO/09/235
Antitrust: Commission imposes fine of 1.06 billion euros on Intel for abuse of dominant position; orders Intel to cease illegal practices - questions and answers cl MEMO/09/235, 2009.

MEMBER STATES

Finland

He 96/1998
HE 96/1998 vp., Hallituksen esitys Eduskunnalle henkilötietolaiksi ja eräiksi siihen liittyviksi laeiksi.

Tietosuojavaltuutetun toimisto 2014
Tietosuojavaltuutetun toimisto: Tietosuojavaltuutetun lausunto liikenne- ja viestintäministeriön asettaman työryhmän valmistelemasta kansallisen big data -strategian luonnoksesta, Dnro 1518/05/14, 2014

Tietosuojavaltuutetun toimisto, 2/2001
Tietosuojavaltuutetuntoimisto, Hyvä tietää 2/2001, Henkilötietojen käsittely suostumuksen perusteella, Tietosuojavaltuutetuntoimisto, 2001.

Tietosuojvaltuutetun pohja tietosuojaselosteen laatimiseksi
Tietosuojavaltuutetun pohja tietosuojaselosteen laatimiseksi, http://www.tietosuoja.fi/31634.htm.

France

Commission Nationale de l'Informatique et des Libertés

CNIL, 2012a,

Commission Nationale de l'Informatique et des Libertés, 'Measures for the Privacy Risk Treatment, Translation of June 2012.

CNIL 2012b,

Commission Nationale de l'Informatique et des Libertés, Methodology for Privacy Risk Management, 2012.

United Kingdom

Information Commissioners Office

ICO Big data and data protection, 2014.

Information Commissioners Office, Big data and data protection, 2014

ICO Anonymisation: managing data protection risk 2012

Information Commissioners Officer, Anonymisation: managing data protection risk code of practice, 2012 .

HM Treasury

Hampton 2005

Hampton, Philip, 'Reducing administrative burdens: effective inspection and enforcement' in (HM Treasury, 2005).

Non-Member States

Cavoukian and El Emam 2011

Cavoukian, Ann and Khaled El Emam: Dispelling the Myths Surrounding De-identification: Anonymization Remains a Strong Tool for Protecting Privacy, Information and Privacy Commissioner, Ontario, Canada, 2011.

FTC 'Enforcing Privacy Promises'

Federal Trade Commission: 'Enforcing Privacy Promises' http://www.ftc.gov/news-events/media-resources/protecting-consumer-privacy/enforcing-privacy-promises, 2014.

Private Sector and NGOs

Institutes & Companies

CIPL 2014a

Center for Information Policy Leadership: A Risk-based Approach to Privacy? An Initial Issues Paper for Privacy Risk Framework and Risk-based Approach to Privacy Project, 20.3.2014, http://www.informationpolicycentre.com/privacy_risk_framework/.

CIPL 2014b

Center for Information Policy Leadership: A Risk-based Approach to Privacy: Improving Effectiveness in Practice, 19.7.2014, http://www.informationpolicycentre.com/privacy_risk_framework/

CIPL 2014c

Center for Information Policy Leadership: Response of the Centre for Information Policy Leadership to the National Telecommunications and Information Administration's Request for Public Comment on Big Data and

Consumer Privacy in the Internet Economy, DOCKET NO. 140514424–4424–01, 2014.

Deighton and Quelch 2009

Deighton, John and John Quelch: Economic Value of the Advertising-Supported Internet Ecosystem, The Interactive Advertizing Bureau, 2009, http://www.iab.net/media/file/Economic-Value-Report.pdf.

Epsilon 2014

About US' <http://www.epsilon.com/about> accessed on 16.8.2014.

Experian, 2014

Experian Information Solutions, Inc., http://www.experian.fi/

Manyika 2011

Manyika, James and others: Big data: The next frontier for innovation, competition, and productivity' in (McKinsey Global Institute, 2011).

Sedayo 2012

Sedayao, Jeff, Enhancing Cloud Security Using Data Anonymization, IT@Intel White Paper, IT Best Practices Cloud Computing and Information Security 2012.

SiSense 2014

SiSense. 'Data Justice for All', 2014, http://www.sisense.com/company/.

Telefonica 2014

Telefonica Insights, 2014 http://dynamicinsights.telefonica.com/479/about-us.

NGOs

IRGC Risk Governance Framework 2008

International Risk-Governance Council: Introduction of the IRGC Risk Governance Framework, 2008, http://www.irgc.org/risk-governance/irgc-risk-governance-framework/.

Persis, McLauglin and Levy, 2014

Persis Yu, Jillian McLaughlin and Marina Levy: 'Big Data; A Big Disapointment for Scoring Consumer Credit Risk, National Consumer Law Center, March 2014, http://www.nclc.org/images/pdf/pr-reports/report-big-data.pdf.

Rodriguez, 2010

Rodriguez, Katitza: European Privacy Officials: Google, Yahoo, and Microsoft Are Still Breaking European Privacy Law, Electronic Frontier Foundation, https://www.eff.org/deeplinks/2010/06/european-officials-google-yahoo-microsoft-breaking-law.

UKAN

UK Anonymization Network: About Anonymisation: for data about people, 2014, http://www.ukanon.net/key-information/.

Cases and Administrative Decisions

ECJ

ECJ Google

Court of Justice of the European Union, Judgment in Case C-131/12 Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja González, 13.5.2014.

ECJ Press Release No 70/14

Court of Justice of the European Union: ECJ C-131/12, Press Release No 70/14, 13.5.2014.

ECJ Seitliner and Other

Court of Justice of the European Union: Judgment in Joined Cases C-293/12 and C-594/12 Digital Rights Ireland and Seitlinger and Others, 8.4.2014.

ECJ Promusicae

Court of Justice of the European Union: C‑275/06, Productores de Música de España (Promusicae) v. Telefónica de España SAU, 29.1.2008.

Costa v E.N.E.L

Judgment of the Court of 15 July 1964. Flaminio Costa v. E.N.E.L Reference for a preliminary ruling: Giudice conciliatore di Milano - Italy, ECJ Case 6-64.

Van Gend

Judgment of the Court of 5 February 1963.NV Algemene Transport- en Expeditie Onderneming van Gend & Loos v Netherlands Inland Revenue Administration. Reference for a preliminary ruling: Tariefcommissie - Netherlands. Case 26-62.

European Court of Human Rights

Paeffgen v. Germany

Application nos. 25379/04, 21688/05, 21722/05 and 21770/05 by Paeffgen GMBH against Germany

DPAs

CNIL Délibération n. 2013-420

Commission Nationale de l'Informatique et des Libertés: Délibération n. 2013-420 de la formation restreinte prononçant à l'encontre de société Google Inc, 2013.

FTC

Vision I Properties

UNITED STATES OF AMERICA, BEFORE FEDERAL TRADE COMMISSION COMMISSIONERS: Deborah Platt Majoras, Chairman Orson Swindle, Thomas B. Leary, Pamela Jones Harbou, Jon Leibowitz, In the Matter of VISION I PROPERTIES, LLC,d/b/a CARTMANAGER INTERNATIONAL, a corporation, DECISION AND ORDER, FTC Docket No. C-4135, April 19, 2005.

Google Inc. Safari FTC Statement

Statement of the Commission: United States of America v. Google Inc., United States District Court for the Northern District of California, In the Matter of Google Inc., FTC Docket No. C-4336 August 9, 2012

Other

Google Inc. Safari

District Order Approving Stipulated Order for Permanent Injunction -Google Safari, UNITED STATES DISTRICT COURT NORTHERN DISTRICT OF CALIFORNIA SAN JOSE DIVISION United States of America, Plaintiff, v. GOOGLE INC., Defendant./ No. CV 12-04177 SI Order Approving Stipulated Order For Permanent Injunction and Civil Penalty Judgment

Southern Illinoisan v. Illinois

Southern Illinoisan v. Illinois Department of Public Health (Justice MacMorrow). IN THE SUPREME COURT OF THE STATE OF ILLINOIS, Docket No. 98712, 2.2.2006.

Abbreviations

CIPL	Center for Information Policy Leadership
CNIL	Commission Nationale de l'Informatique et des Liberté
DPA	Data Protection Authority
DPD	Data Protection Directive
EU	European Union
ECJ	European Union Court of Justice
EEA	European Economic Area
ECtHR	Europea Court of Human Rights
FTC	Federal Trade Commission
FRA	European Union Agency for Fundamental Rights
GDPR	General Data Protection Regulation
ICO	Information Commissioners Office
IRGC	International Risk Governance Council
LOPD	Ley Orgánica de Protección de Datos de Carácter Personal
PbD	Privacy by Design
PET	Privacy Enhancing Technology
US	United States of America
UK	United Kingdom
UKAN	UK Anonymization Network

1 Introduction[1]

1.1 Big data and the conflict of fundamental rights

The problems privacy faces at the wake of 21th century are tremendous. Technological change is challenging the whole concept of privacy. The Internet allows cheap and easy monitoring of our behavior, which via behavioral advertising acts as the primary monetization motor for the free services offered across the cyberspace.[2] Internet users do not have control of their personal data or privacy, although this right is a fundamental one. At the same time, there exist economic incentives to both use more data and use the existing data to a greater extent; a too strict a privacy regulation could block this area of economic activity, which would be against the fundamental rights of the data controllers as legal persons also causing welfare losses to the whole community. The research question of this thesis is the secondary use of data, especially in the commercial context. Secondary use of data is a broad term. It can be understood as all of the further activities done with the data other than the initially stated purpose of the collection. The currently available solution for the conflict of principle rights is anonymization. [3] An alternative solution, from commercial perspective, in some cases, is non-compliance with the data protection regimen.[4] The current lack of enforcement is a big problem for a consumer's right to privacy. To be more specific, the problem is in the frequency sanctions are given and in the monetary value of the sanctions.[5] Without safeguarding, the right to privacy is a dead letter of the law, *de lege ferenda* strengthening enforcement could be the solution to non-compliant secondary use of data. Sanctions do not however facilitate the use of personal

[1]In addition to my supervisor Kristian Siikavirta, I would like to thank Tobias Bräutigam for excellent guidance and comments on this thesis and for good and educational discussions about data protection. Additional thanks go to Riikka Koulu for commenting the fundamental law perspective and structure of this thesis.

[2] Interestingly, majority of people using the internet in EU state that they would not like to disclose personal data in exchange for free services. This raises the question whether the users actually understands the amount of data collected. SPECIAL EUROBAROMETER 359 Attitudes on Data Protection and Electronic Identity in the European Union, 2011, p. 33.

[3] Anonymization is the technique rendering personal data in such a form that there are no personal identifiers left, and the data thus becomes anonymous. I will treat this issue more in the chapter 3.

[4]This solution can be used if the legislation set by data protection directive is followed. It is however important to note that the authorities have problems following and enforcing the data protection directive.

[5] For example in Finland there are no sanctions in place for violations of the Personal Data Act, the same scenario applies for Ireland and Austria as well. In certain European jurisdictions sanctions are given out more frequently – however in these areas the monetary value of the sanctions creates the problem.

data – there is a clear need for using data, and too strict rules in combination with strict enforcement could kill the free market in this area, and remove the 'free services' available for consumers. This also applies to the small and medium enterprises, which could be brought down by a single big sanction or additional compliance costs. Too strict data protection rules, in combination with strong enforcement, would prevent new players from entering the cyberspace market; big established companies have a much stronger position for adapting to regulatory changes. This brings forth the need for comprehensive solutions, which include both the text of the law, as well as enforcement actions (*law in action)*. A new risk-based approach to secondary use of data is needed, in combination with a strong and uniform enforcement regimen.

The emphasis of this thesis is on the regulatory side. The goal is to find guidance for a pan-European regulation on the secondary use of data, which would solve the conflict of privacy and utility. This conflict on the level of fundamental rights is essentially about weighting privacy against the importance of functioning online or information technology markets – that could bring wealth and efficiency across industries as well as to consumers. Big data highlights the importance of solving this question. Big data is collections of information of extreme size and complexity, usually so extensive that traditional computing methods would not be able to analyze the information. As defined in the IT glossary of the consulting firm Gartner:

> "Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making."[6]

I would supplement this definition by certain practical examples of big data. Technologies such as data visualization are needed to analyze the data. Analyzed data can be extremely useful in different processes from profiled marketing to predictive analytics.[7] The privacy aspects of big data arise when the data sets contain information that is identifiable or connectable to a specific individual. From the perspective of the individual, this data creates a risk of infringement of privacy. The data controller's economic interest is to monetize or otherwise use the data as a part of running their business. The data protection laws however put the controller in a position where their job is also to protect the fundamental rights of the data subjects. This dual role of the controller causes problems if

[6]Gartner, 2013 It-Glossary
[7]Siegel 2013

sufficient enforcement and administrative oversight is not in place, as the common saying goes:

"Like the fox guarding the chicken coop"

The data controller has interests from benefiting from the data, while simultaneously there are obligations to follow.

The term secondary use of data is derived from the purpose restriction defined in article 7 of the DPD or Section 7 of the Finnish Personal Data Act. Personal data can only be used for the predefined purposes, which the data subjects are aware of. Secondary uses of data are the possible uses for personal data, which have not been primary reasons for collection. As an example, an e-commerce company could collect personal data to deliver their customers products. As a secondary use this data could be used to create consumption models or profiles.

I want to draw attention to two aspects of big data. *First*, more and more information is collected; every web-based service contains increasing amounts of customer data; from purchase history to location data and records of Internet behavior. In addition, other service providers also collect data, such as credit card records, club membership data, location data from mobile devices, social media feeds and marked preferences, contacts or interests in services like Facebook[8]. *Second*, the multiple modern data processing technologies that have increased computing power in combination with advanced analytics, allow combining and analyzing sets of data and finding patterns and correlations from the data. New information is thus created from previously existing data.[9] This second aspect of big data is especially problematic from the perspective of personal data and privacy. This is due to the basis of our data protection legislation; data is collected with the *informed consent* for data collection and for a certain *purpose*. Especially important is the fact that data is collected only for a single purpose, as an example address data for deliveries of e-commerce goods.

The current legal framework does not sufficiently answer the problem of secondary use.[10] Law is lagging behind technological change. Computing capacity has significantly

[8]Feed is an entry to the social media, for example, Facebook allows sharing certain thoughts to social media contacts.

[9] Mayer-Schönberger and Cukier 2013, p. 107, for example Harford has claimed that Big Data does not remove the problems that exist within statistical research, the size of the data sets my in fact lead to more mistakes. Harford, 2014.

[10]Also the MacKinsey report on big data captures this aspect of regulation lagging behind: "Policy makers need to recognize the potential of harnessing big data to unleash the next wave of growth in their economies.

increased in the last 20 years.[11] The current European Union legislation was created before Internet usage was widespread. The current privacy legislation does not answer to the specific aspects of big data[12], profiling or cloud computing.[13] Even if the legislations were adequate, the problem would be its enforcement. If not enforced, regulation does not protect fundamental rights. Enforcement is seen as a key issue in the regulation of secondary use of data, since the current shortcomings of enforcement challenge the whole essence of data protection laws. Effective enforcement, in combination with risk-based approach, would allow the secondary use of data in a manner that both safeguards the data subjects' fundamental rights as well as allows economic growth in the next wave of data-based services and industrialization.[14]

The conflict between privacy laws and the information economy is worsened due to the fast growing value of data. This value is captured well in this popular quotation:

> "Data is the new oil" – Clive Humby[15]

Infonomics analyzes the value of data and sheds light on the claim that data is the new oil. This discipline or sub-science of economics assesses the value of information. If data is seen as the new oil of economy, it has to be valuated.[16]

The Internet is a powerful and unique environment for marketing, as it collects and contains an enormous wealth of data from an economics perspective; it thus facilitates segmented marketing, and as such captures a larger share of the surplus, as segmented marketing allows a much higher diffraction of prices for products.[17] The importance of

---

They need to provide the institutional framework to allow companies to easily create value out of data while protecting the privacy of citizens and providing data security. Manyika and others, 2011, p. 13 see also p. 63.

[11]For the increasing technologic capacity see Moore's law, which states that the number of transistors in circuits doubles approximately every two years. http://www.computerhistory.org/semiconductor/timeline/1965-Moore.html.

[12]There are multiple problems that big data create, equality of the consumer pricing for example as well as accuracy of the data used in cases when analyzing consumer credit risk might cause problems. For a review of the shortcomings of Big Data in the credit rating sector see for example Persis, McLaughlin and Levy 2014.

[13]Cloud computing may be defined as offering services thorough Internet. According to Milliard and Hong cloud computing may be one of the technologies allowing to benefit from big data, in more technical terms cloud computing is not dependent of the used location and the capacity can be used by demand. For more information on cloud computing see, Millard 2013, p. 1.

[14]Varian and Shapiro provide a good overview about Economics of the Information age and the old rules applying see Shapiro Varian, 1999, see of big data's role in increasing economic efficiency; Manyika and others, 2011.

[15]Palmer 2014.

[16] Laney's infonomics is an emerging theory that claims that information is an asset class. Information is seen as an actual asset; that can be quantified and internally accounted. Infonomics is not focused on personal data; the focus is on all information in general. The major thesis of the theory is that value of information should be valuated and analyzed, so the data can be used efficiently. See more, Laney, 2012.

[17]Craig and Lundloff 2011, p. 52 - 53.

infonomics arises from the fact that the importance of data collection and different data-mining techniques are on the rise. Companies have the possibility for improving performance and gaining savings with powerful utilization of big data. According to MacKinsey Global Institute, there are various areas in which personal data usage could efficiently be used for increasing productivity and economic gains.[18] These data groups include, for instance, location data in retail, health care data as a source for improving medical service, as well as financial data.[19]An example in the area of retail and marketing could be data about activities in both electronic and physical stores[20]

This thesis will analyze the legal construct of secondary use of data, and see whether anonymization can solve the conflict between freedom of trade and privacy.[21] In the secondary use of data, this thesis will focus especially on *purpose restriction* and the legitimate grounds for data processing. In addition, enforcement of privacy is seen as a vital part of the subject, since without proper enforcement the data controllers do not have the incentive to protect privacy in the form of effective anonymization. My thesis is that anonymization would give a clear answer to the conflict between commercial value of data and privacy. Big data increases the need for enforcement since the economic gains create incentive to break privacy regulation and without regulation the individual right of privacy cannot be fulfilled.

In the European Union data protection legislation is harmonized with a directive, which is implemented nationally in 28 different manners. In the following chapters, the European directive and the data protection rules and how they affect anonymization and enforcement are examined and explained.

From a constitutional perspective, this thesis interprets the conflict of economic value (rights for doing business and to functioning markets) against the fundamental right of privacy. Some limitations to the scope have to be done, since the thesis consists of various topics and they could easily be expanded.

---

[18]Manyika and others 2011, p. 100.
[19]Manyika and others 2011, pp. 39, 52, 63, 72 - 74 and 127.
[20] Modern technology allows monitoring movements both in the virtual e-commerce stores as well as in physical locations that contain for example Wi-Fi, also the store cameras can be used to analyzed eye movements. The stored information can for example includes the time spend looking certain items or the mouse movements or 'click-stream' in the online context. See more about monitoring in stores and targeted ads, Morran 2013 and Popescu, 2013.
[21]To some extent the question, whether anonymization alone is sufficient needs to be raised. Schneier has for example brought up the problems of anonymization Schneier 2007; see also Ohm 2010, p. 1704 - 1705.

1.2 Multilayered research question and previous research

My thesis falls under category of information law, with an economics and law focus on the analysis. In the background, there is a visible wider fundamental rights conflict between free market values and the right to privacy, and the changes in their balance due to technologic change. This greater "*superquestion*" locates this thesis to the field of information law, more specifically data protection law. Further elements are derived from EU law, since the harmonization efforts in Europe are driven by the EU. There are no previous studies either about the risk-based approach to privacy in Finland or about secondary use of data.[22] In fact, the majority of research about big data, privacy and anonymization is of Anglo-American origin. Pöysti has previously applied the economic method on data protection issues in his doctoral thesis "Tehokkuus, informaatio ja eurooppalainen oikeusalue".[23] In the area themes such as anonymization or purpose limitation, there are multiple international publications concerning these themes.[24] Data protection and privacy are popular subjects, so more research emerges every day both globally and locally.

The research question of this thesis is what is *commercial secondary use of personal data* and how could it benefit from *a risk-based approach*? The thesis aims to solve the conflict between privacy of individuals and commercial interests. The research question contains multiple aspects. The first is about the problems and challenges the current regulation is facing due increasing economic pressures, and whether it can efficiently solve those problems. The second question is whether it is possible to solve the described problems by a risk-based approach to regulation. The three areas analyzed in connection with the Data Protection Directive (DPD) are the provisions and principles that limit or allow the secondary use of data. Most notably DPD articles 6(b) *purpose limitation*, 7(a) *consent based processing* 7(f) *legitimate interest*. Anonymization is also analyzed, since it acts currently as the best possible solution for the secondary uses of data and it can provide much support for the risk-based approach. Enforcement has a dual role in the research question, first the lack of enforcement causes problems, second sufficient and efficient levels of sanctions could act as an important solution for the problems privacy is facing

[22] A thesis has been published about the recombination of data. See, Pakkanen, Tomi: Combination of Personal Data and the Data Protection Reform in the European Law, Edilex, 2014

[23] Pöysti, 1999.

[24] Just naming a few Brouwer, 2011, Ohm, 2009-2010, Ohm, 2012, Schwartz, 2012-2013 and Yakowitz, 2011-2012.

currently. I will also analyze if data protection regulation could benefit from a risk-based regulation. The chosen approach to the research question could be described as a multilayered approach.[25] A wide array of topics from anonymization to purpose restriction and enforcement are analyzed, this is however necessary for capturing the essence of secondary use of data.

The research question is of multidisciplinary nature, and it thus falls under the subjects of data protection and law and economics. Protection of personal data is commonly located under the concept of information law.[26] According to both Pöysti and Voutilainen information law, includes the questions of privacy, protection of personal data, publicity of data, tele-activities, electronic commerce and data security law.[27] Data protection law is the primary context of this thesis. Data protection law or privacy law[28] is part of administrative law or information law, but it is also strongly entwined with law of obligations, especially when analyzing consent.[29] As Bräutigam has noted data protection law is of technical nature and addresses information technology.[30] In the Nordic context, data protection law has been seen as part of computing law or information law.[31] There are, however, certain problems in this systemization and a computing law or information law school has not clearly emerged.[32] Saarenpää views protection of privacy as part of personality rights, but also partly as a problem analyzed in information law.[33] Saarenpää has also noted that the fact that the notion of privacy becomes judicial leads automatically to the fragmentation of the research of privacy.[34] From the perspective of this thesis it is important to note that privacy can be studied from multiple perspectives, the school of information law provides some support in the form of principles. It is, however, so that currently neither data

[25] Siltala has stated that academic text contains two levels: the first is 'metatext', the frame, borders, premises and goals of the research, the second level being the 'object', containing the true/untrue valuations. Siltala, 2013, p. 247.
[26] Although the school of information law, as well as data protection law are both young members of the legal doctrine and legal education
[27]Some differences exist between the classifications, for example, Voutilainen also adds freedom of speech under information law and does not include the commercial aspects as strongly as Pöysti. Pöysti, p. 368 – 370 and Voutilainen, 2012, pp. 34 - 35.
[28] Data protection is the European term used, in the Anglo-American context privacy law is used as a synonymous; I will through my thesis employ both terms privacy and data protection.
[29]Data protection law can be assigned to to various schools of law, for example, Bygrave has classified data protection law to be closely related with administrative law and human rights law, Bygrave, 2002,p. 166.
[30]Bräutigam, 2012, p. 415.
[31]Seipel, 1977, p. 124.
[32]Seipels views have, however, changed as can be seen from Seipel 2004, pp. 269 - 270, 272. Pöysti has also noted this Pöysti, 1999, p. 368 - 369, see also Saarenpää on the nature of information law Saarenpää, 2009b, p. 44.
[33]Saarenpää, 2009b, p. 7.
[34]Saarenpää, 2009a, p. 366.

protection nor information law have clearly taken their place as independent doctrines, in the future this may be the case, as information law becomes more vital for the tools of each lawyer.

The focus of the research question leaves certain areas outside of the research. I will conduct the analysis of law especially keeping in mind the commercial or so-called private sector data processing.[35] The end-goal is that a balance is struck between individuals' right to privacy and a functioning market. The rights that will be balanced are the property rights and economic rights of those processing the data; against the right to privacy the individuals (*data subjects*) carry. The fact that most of data collectors are private companies, and that the purposes for data analysis and collection are commercial, should be taken to account.

Although the recent news of NSA and other foreign and counter-intelligence operations receiving information from the commercial sector would justify a different focus, within the scope of the work in hand it is not possible to inspect this to a sufficient degree.[36] It is, however, important to mention since, in many cases, the question has been the access to private databases, the importance of secondary data collection is high. Companies collect data that the government then might access. The historical description of privacy and data protection laws is also left to a minimum, since the contribution offered would not justify the excursion to the history of data protection. In addition, such historical analysis would not help in solving the research question. Questions and areas balancing free speech and privacy are left untouched, since they do not have an important role when analyzing the secondary use of personal data.

The next subchapter will analyze the chosen methodology, which is used to frame and solve the research question. The chosen methodology could be defined briefly as risk-based law and economics analysis of regulatory activities and current statutes, with legal dogmatic analysis as a support function.

[35]This leaves for example the purely research use of for example medical data outside of the scope. Some examples of medical data being used after anonymization will be, however, used.

[36]The public sector processing of data is left outside the scope, even though especially the PRISM project and the so-called 2013 Snowden revelations have caused much discussion of privacy. I see these issues more political than in the field of law. More information about the Snowden revelations and the full history of Snowden can be read from the book of Snowden Files by Luke Harding.

1.3 Twofold methodology

I employ a twofold methodology to answer the research question. The core of methodology in my thesis is law and economics, which I apply for *de lege ferenda* analysis and for explaining the current environment in which the law operates.[37] Law and economics helps conceptually binding together the value (*and monetization*) of data as well as setting effective level of sanctions, that would be needed for safeguarding the e-consumers right to privacy. More specifically the used law and economics method leans towards a risk-based understanding of regulatory activities. I will additionally employ legal dogmatic analysis to analyze the current legal norms on the secondary use of data, enforcement of unauthorized data use, and anonymization.

This thesis analyzes fundamental right to privacy and weighs it against the economic reality, discussing the nature of privacy as a right. The role of law and economics is important since it helps to provide answers to the problems in the current legal reality, especially answering whether a statute is efficient and how it allocates resources and liabilities.[38] In addition, as stated by Pöysti, law and economics further clarifies the effects of a norm and acts as a comparative tool for the analysis of law.[39] Legal dogmatic analysis and law and economics have very different view on what is law; legal dogmatic analysis analyzes the law from within, law and economics from outside.[40] I use law and economics to analyze the question how the current law should be changed or if it should be changed. Law and economics can be defined as the application of economic theory to examine the law and the impacts of law to the surrounding economic reality.[41] Law and economics can be especially useful when analyzing efficiency of laws or sanctions. In addition, law and economics help to identify the economic effects of laws for different parties.[42]

The fundamental rights aspect of this thesis needs to be analyzed by the traditional legal dogmatic approach, due the institutional roots of human and fundamental rights. Legal dogmatic method acts as the secondary method of this thesis. I use it mainly as a tool to analyze the current directive, statutes and proposed EU regulation. It also helps balancing

[37] The chosen approach is following Mackaays description, institutional, since law and economics is used for the analysis of legal institutions and the way they affect individuals, Mackaay, 1982, p. 4.
[38] Mähönen, 2004, p. 49 and Mercuro and Medema, 1997, p. 23.
[39] For an example on use of law and economics on the field of information law see Pöysti, 1999, p. 49.
[40] Law and economics sees law as vaguely as possible to explain the effects of law. The whole sum of combined political-legal structure is seen as the object of study. Mercuro and Medema, pp. 21. - 22.
[41] Mercuro and Medema, 1997, p. 3.
[42] Friedman, 2000, pp. 18 - 19.

the conflicting fundamental rights.[43] Aarnio has described the legal dogmatic method to mean the description of legal norms; more precisely, it includes the definition and systemization of the legal norms.[44] In some parts of the thesis, I take comparative sources from the various European jurisdictions, bringing a comparative aspect to the used legal dogmatic method.[45] The purpose of these European excursions is to show examples of the implementation of EU law. However, the purpose is not to venture deep into legal comparison, since this would be impossible and it would not help further in answering the research question. Additionally, the method is in some scenarios used to bring analogy from environmental law and competition law to the discussion about sanctions, or for excursions to the United States context.

When analyzing the *superquestion of fundamental rights colliding*, the legal dogmatic method is influence by the idea of a systemization that is *responsive* to the fundamental rights. In his research Pöyhönen has analyzed this 'responsive approach' to fundamental rights in the context of property law:

> "The result of the responsive approach to the fundamental rights is that the norms of property law become more responsive to fundamental rights… The application of the norms of property law has to be analyzed following the principle of proportionality, by asking how much maintaining the fundamental rights of an interest party endangers or disrupts the fundamental rights of the other interest party." (Translated by the author)[46]

The essence of this idea is that when applying fundamental rights in the context of property law, we must ask the question of how the protection of fundamental rights applies to the fundamental rights of the other party. This approach works particularly well in the area of data protection and secondary use of data, where increasing the fundamental rights of the controller affects the fundamental rights of the data subject instantly and vice-versa. This is due the fact that information law is strongly in connection with fundamental rights.[47] In practice, this balancing of fundamental rights of different interest groups must be done in a way that rights are not limited in a draconian way.

---

[43]Siltala states that one of the basic functions of legal dogmatic analysis is the weighing and balancing of legal principles and fundamental rights that have enough institutional standing. See more Siltala, 2003, p. 138.

[44]Aarnio, 1986, pp. 110 - 111.

[45]This is due the fact that when doing research in the field of data protection (or information law), the legal dogmatic approach needs to take into account the international nature of information and the fact that much of the information laws have an international birth history. Saarenpää, 2009b, p. 15.

[46]Original Finnish text: "Perusoikeusmyönteisen systematisoinnin seuraus on varallisuusoikeuden normisto perusoikeusherkkyyden vahvistuminen… Varallisuusoikeudellisen normiston soveltamisharkinta on suhteellisuusperiaatteen mukaisesti, jäsennettävä kysymällä, missä määrin yhden intressitahon perusoikeussuojatun aseman ylläpitäminen haittaa tai vaarantaa toisen tason asemaa.", Pöyhönen, 2003, p. 78.

[47]Bygrave 2002, pp. 166 - 167.

The risk-based approach helps balancing the conflicting fundamental rights, in a way that would protect data subjects' rights, while at the same time allowing commercial secondary use of data. The adaptation of a risk-based approach can be justified by Beck's classification of social risks:

> "Risk may be defined as a systematic way of dealing with hazards and insecurities induced and introduced by modernization itself."[48]

In the context of my thesis, the risks are either man-made or technologic risks. The concept of risk is a powerful tool for understanding regulation. In the context of secondary use of data the risk-based methodology analyzes the potential uses of personal data versus the potential losses of individual privacy of the data subjects. This helps to further anchor risk to the fundamental rights, the loss of privacy is seen as the ultimate risk, and the amount of data and the sensitivity of data acts as the measuring instrument. Julia Black has classified risk in regulation to four categories:

> "Risk is an *object of regulation* in that much regulatory activity is defined in terms of risk. Risk plays a *justificatory role* in that it defines the object and purpose of, and provides a justification for, regulation, and thus frames regulatory policy making. Risk plays an *organisational and procedural role* in that risk provides the basis for the regulator to operationalise its objectives and for the introduction of particular sets of internal organisational policies and processes."[49]

This definition of the role of risk by Black is optimal since in privacy-related risk all of the four areas are fulfilled. First, privacy regulation should *de lege ferenda* concentrate on the management of risks; the justification of data protection regimens is to protect individuals against the loss of privacy. In addition, the organizational and procedural role of risk is extremely well in place for internal privacy processes. As well, this plays as a source to analyze the actions that should be done in relation to enforcement and anonymization. Hence risk-based regulation can be adjusted in several situations to meet changes of the environment internally (*institutional changes*) as well as when technology changes the surrounding reality (*technological changes)*. I will further analyze the risk-based regulation and the application of such approach in the chapter 5.

This thesis is built with a variety of sources, from different levels; the most important sources are the legislation and EU directives and other regulation on the issues at hand. The following subchapter will present the most important sources and official material

[48]Beck 1992, p. 21.
[49]Black, 2010a, p. 1, chapter 14.1

used to analyze the set research question. I will also analyze the hierarchy of norms, and choose an approach that serves the selected material and methodology best.

1.4 Sources of the research and hierarchy of norms

This thesis treats especially European data protection law from a law and economics perspective. The traditional Nordic or Finnish school of legal dogmatic of a Peznick-Aarnio-origin, it does not give international material, consequentialist arguments or for example EU law much value.[50] Siltala has given EU law, principles and international law more weight in his hierarchy of norms, "*dynamic hierarchy of norms* or *responsive hierarchy of norms*" noting the value of EU law as an institutional source of law. Siltala also includes international comparative material as a possible non-institutional source of law. His model also includes economic arguments and *consequentialist arguments as possible non-institutional sources*.[51] My thesis employs this *dynamic-model*, since the thesis derives from a multitude of international, EU and economic arguments.

The most important source for data protection laws is the EU. The EU has harmonized national data protection statutes by a directive, and the Commission aims to regulate data protection by a regulation. For this reason, the role of EU law has to be taken into account in the hierarchy of norms, even though the national implementation of data protection directive is not affected directly by EU law. EU law is an autonomic international legal system that can create obligations and rights to individuals within the Member States.[52] The doctrine that EU law can limit Member States sovereign rights was first discussed in *Van Gend en Loos*.[53] The ECJ first established the supremacy of EU law in *Costa v. Enel*.[54] In the legal hierarchy, the supremacy of the EU law sets it to the highest position, above national law.[55] In addition to EU law the role of law and economics in the hierarchy of norms has to be accounted. The role of EU law is important, but in the current state of affairs, the supremacy does not have similar role that it could have in future after the

[50]Aarnio 1987, pp. 89 - 90.
[51] Siltala 2003, pp. 896 - 897.
[52]Tuomas Ojanen 2010, p. 35.
[53]Judgment of the Court of 5 February 1963.NV Algemene Transport- en Expeditie Onderneming van Gend & Loos v Netherlands Inland Revenue Administration. Reference for a preliminary ruling: Tariefcommissie - Netherlands. Case 26-62.
[54]Judgment of the Court of 15 July 1964. Flaminio Costa v E.N.E.L Reference for a preliminary ruling: Giudice conciliatore di Milano - Italy. ECJ Case 6-64.
[55] P. Craig and De Burca 2011, pp. 256 - 257.

regulation is valid. Directives do not give directly rights to the citizens, unless the member state has not failed to implement them.[56]

According to Tolonen the role of law and economics in legal dogmatic argumentation falls under the category *consequentialist argument.* Siltala in the other hand has defined the role of economic arguments as independent non-institutional sources in his dynamic hierarchy of legal sources.[57] Siltala has analyzed the role of law and economics as method and its role in the hierarchy of law; he has noted that the usability of law and economic varies by the field of law, and that in many cases the institutional roots such as the constitution still needs to be considered. According to Siltala, the arguments of economics have more effect in areas that are connected with trade and commerce.[58] In my thesis, this is precisely the case, law and economics is important since data protection is strongly in connection with the online economy.

The most important source of law for my thesis is the data protection directive, although there is variation in the implementation of the directive.[59] In addition to the directive, I will also especially interpret the Finnish Personal Data Act (Henkilötietolaki 1999/523) and in some cases for additional perspectives the Spanish Ley Orgánica de Protección de Datos de Carácter Personal (LOPD) and the United Kingdom Data Protection Act. These national laws are used as examples of implementation.[60] The future proposed EU regulation, the General Data Protection Regulation (GDPR) is also taken into account when considering the secondary use of data and sanctions. When analyzing the sanctions I will take into account the original Commission proposal and the Parliament amendments to the sanctions. GDPR has most recently been accepted in the European Parliament and the implementation is forecasted to start in 2015.[61] The GDPR contains many elements for the future of data processing and big data, most notably the articles 22 and 23 contain duties for increasing accountability and data protection by design in companies.[62]

Working Party 29 plays an important role in the interpretation of the DPD as well as advising on data protection issues. The Working Party is established by article 29 of the

[56] See more about the implementation of directives P. Craig and De Burca, 2011, p. 106 and Ojanen, 2010, pp. 43 - 44.
[57] Tolonen 2003, p. 152 and 164.
[58] Siltala, 2011, pp. 109 - 111.
[59] Even though there is variation in the implementation, Raab & Bennett have stated that the DPD is by far the most important instrument governing data protection, Bennett and Raab 2006, p. 93.
[60] The purpose is not analyzing the differences or to give detailed analysis about the national statutes.
[61] Commission, MEMO 14/186.
[62] I will not concentrate my analysis in these changes, since in this point is still two hard to say how they would affect the data processors see more, GDPR, p. 56.

DPD, its duties are listed in the article 30.WP 29 acts as an independent advisory body, consisting of representatives from the Member State DPAs.[63] As also Kuner notes, the working party has an influential role in data protection, since it often provides opinions on unclear issues or new problems caused by development of technology; often crystallizing the interpretation of the DPD.[64] However, it should be recognized that the WP 29 opinions are not binding by the nature of law. Even though, the opinions of WP 29 act as crucial source material for this thesis, shedding light on multiple issues arising in connection with secondary use of data and anonymization. Especially the recent opinions on anonymization, interpretation of article 7(f) and purpose restriction are used for analyzing the DPD.[65] It is however important to note that the opinions and recommendations the working party publishes, are not binding.[66] In a conflicting case, it could thus always be stated that the law should be interpret differently from the guidance of WP 29.

I will also use local DPAs official guidelines and decisions to look for guidance on the new issues. I will in this area derive especially from the guidance from the Information Commissioners Office (ICO) and Commission nationale de l'informatique et des libertés (CNIL).[67]There is few only very few material from the Finnish DPA regarding the questions around anonymization, risk management or big data.[68] As of official court decisions I will analyze several relevant cases from the ECJ.

When analyzing the effectiveness of sanctions in chapter 4, I will analyze certain US based Federal Trade Commission decisions alongside with sanctions issued by European DPAs. These will clarify the varying enforcement processes and levels of sanctions at the different sides of Atlantic. The goal is to shed light on the law in action of data protection. The US material is used as a complementary source that benefits *de lege ferenda* and illuminates to the potential ways of improving the current enforcement. In addition I will also use the preparatory materials of the GDPR and reports by the European Union Agency for Fundamental Rights (FRA) to analyze the current state of sanctions.

---

[63] WP 29 is independent of the local DPAs, Members States as well as the Commission see, e.g. Bygrave 2002, p. 73.
[64]Kuner 2007, p. 9.
[65]WP 29 Opinion 06/2014, Opinion 05/2014 and Opinion 03/2013.
[66]Bygrave 2002, p. 73.
[67]These guidelines are for the French and UK context, and thus not applicable law in Finland or other Member States, they however address issues that are similarly implemented in many jurisdictions. CNIL, 2012a and CNILb ICO 2012, Anonymisation: managing data protection risk code of practice 2012 and ICO 2014, Big data and data protection.
[68]There are the comments on the Finnish national big data strategy, Tietosuojavaltuutetun toimisto 2014, Tietosuojavaltuutetun lausunto liikenne- ja viestintäministeriön asettaman työryhmän valmistelemasta kansallisen big data -strategian luonnoksesta.

In the field of legal literature, the material is a mixture of Finnish and international publications and journals. Just to mention some of the most influential publications in relation with this thesis; in the field of anonymization a key article is the controversial paper written by Ohm[69], where he claims that anonymization is dead. Opposing Ohms view, I will consider the claims of Yakowitz.[70] From the Finnish scholarship Saarenpää and Pöysti the most stage time.[71] The material from economics of privacy varies highly, consisting critics of privacy, such as Posner or Acquisti who has used behavioral economics in his analysis.[72] When exploring regulation, the most important publications of this thesis are by Robert Baldwin and Julia Black, who have explored and analyzed risk-based regulation.[73] In the area of risk-based approach, I will also derive form the CIPL white papers on risk-based approach.[74] However, it should be noted that CIPL consists of private entities, which might have goals to influence the future of privacy regulation.

The following subchapter will briefly explain the structure of this thesis; it serves as a road map to this thesis.

1.5 Structure of the thesis

This thesis analyses first the current legal reality, venturing then via sanction to the analysis of future legal concepts that might facilitate the secondary use of data. In the second chapter of this thesis I will treat the current legislation around the secondary use of data. The fundamental rights conflict will be assessed and taken into account, also noting the goal for safeguarding the free flow of data as a goal for data protection laws. This acts as the *superquestion* or conflict of fundamental rights, which frames the whole issue. I will analyze the fundamental right of privacy and weight it against the fundamental rights of the data controller. The focus is on the DPD, and on some of its European implementations – most notably the Finnish implementation.

In the third chapter, anonymization is analyzed as the technical solution for allowing secondary use of data. In this area, the interest is on the current norms on anonymization. The risk of re-identification will also be treated. In the commercial context, the value of

[69]Ohm 2009-2010.
[70] Yakowitz 2011-2012.
[71]Pöysti 1999T and Saarenpää 2009b.
[72]Acquisti, Leslie and Loewenstein. 2009, Posner, 1977-1978 and Posner, 1978.
[73]Black and Baldwin, 2010, Black, 2010b and Black 2010a.
[74] CIPL 2014a and CIPL 2012b.

data is important and for that reason, the devaluation of data utility will be analyzed. Focus will be on the use of anonymization as a safeguard. The issue will be analyzed from the perspective of the risk of re-identification, which is an important part of the issue for the data subject and additionally from the perspective of the controllers seeking to authorize commercial activities by anonymizing data.

In the fourth chapter, I will analyze the enforcement of unauthorized data use, or more generally how are privacy violations sanctioned. The chapter also contains an excursion to the current data protection and enforcement in the United States. Especially the role of FTC and the sanctions given will be explored. The role of this US excursion is to shed comparative light on the monetary effects of sanctions. European sanctions will be compared to the sanctions given out by the FTC.

The fifth chapter will discuss the economic aspects of privacy and apply those as a background material to the risk-based approach. The risk-based regulation will provide input for *de lege ferenda solutions*. Finally, the goal is to apply the risk-based approach as a model for data use and provide solutions to the problems of the current DPD. The goal is to demonstrate how the risk-based approach could better solve the conflict of fundamental rights.

The next chapter will analyze the current norms on commercial secondary data use. The analysis starts by describing the fundamental rights conflict in question. Most notably, the right to privacy will be analyzed.

2 Commercial Secondary Use of Data

2.1 Colliding Fundamental Rights

This subchapter illuminates the conflicting fundamental rights that are behind the secondary commercial use of data. This is referred as the "*super-question or meta-question"* of this thesis, since the conflict of privacy and commercial interests also affects many other subjects than the use of secondary data.[75] This excursion to the conflicting fundamental rights of the data controllers and data subjects is needed for better understanding the collision of fundamental rights. This collision of fundamental rights is

[75]The conflict could also be formulated as the conflict between citizens' rights versus the rights of commerce or business. However, it should be noted that there are also societal benefits in the use of data and individuals do benefit from the free services offered to them.

caused by the dualistic nature of data protection law: on the other hand, data protection laws exist to protect the privacy of data subjects - the consumers using online services, at the same time, one of the goals of data protection laws is to allow economic activities and safeguard free movement of data.[76] This dualistic nature is shown further in for example when conducting the balancing test of DPD article 7(f) (*legitimate interest of controller)* and considering the correct levels of anonymization (*utility vs privacy)*.[77] Lee Bygrave has compared data protection laws to policies of sustainable development in the area of environmental law, these on one hand protect environment and on the other allow economic growth.[78]

Privacy is a fundamental right. Data protection can also be viewed as an individual fundamental right.[79] The freedom of trade and the free market can also be seen as a fundamental right, this is according to Länsineva is the institutional part of protection of property.[80] As noted by the Commission and in the article 52 of the Charter of Fundamental Rights in Europe no fundamental right is ultimate and the limitations in other rights need to be necessary.[81] In EU, privacy or right to privacy is a principle right protected by the article 7 of Charter of Fundamental Rights of the European Union, the article 8 acknowledges the right to data protection as a separate right. The charter also safeguards commercial activities in articles; 15 *right to choose occupation* article, *right to conduct business* article 16 and *right to property* article 17. European Convention of Human Rights protects privacy in the article 8. The section 10 of the Finnish constitution grants the right to privacy or integrity. The constitution also protects the right to conduct in commercial activity in article 18. In the case of secondary use of data, these two groups of fundamental rights are in a conflicting relationship. I would also like to bring a third level to this collusion – it is about the utility and societal benefits, use of data – even commercial use, has societal benefits it raises welfare and stimulates economic activities. This can be seen both in the generation of jobs in the internet sector and also in free services and cheaper prices for the consumers.[82]

[76]GDPR, p. 6 - 7.
[77]Subchapters 2.5 and 3.3.
[78]Bygrave 2002, p. 167.
[79]For the purpose of exploring the secondary use of data it is not important to distinguish privacy and data protection as fundamental rights.
[80]Länsineva 2011, p. 558.
[81]GDPR, p. 7.
[82]This aspect of the collision will be explored more in the subchapter 5.1, where I explore the economic aspects of privacy. See also Deighton and Quelch, 2009 and Varian. In the other side there are fears that big

There are two important aspects in the balancing of fundamental rights, first the legislator should be the party conducting the balancing test via legislation, and second, that the balancing should not affect the core of a fundamental right. The goal in such balancing activities is that both of the fundamental rights can operate as much as possible.[83] For operating and considering the balancing of privacy against other rights, it is essential to first explore the nature of privacy.

There are two different approaches, the first one is the US view of seeing privacy as liberty and a protection against the state; the second, European view, is seeing privacy more as a right to human dignity, which works both against the state and private parties.[84] This integrity based approach can be clearly seen in, for example, Sweden.[85] In comparison The US right, as Brandeis has formulated, is more about the idea of right to be left alone.[86] The individual is responsible for his or her conduct; privacy acts as a boundary between the individuals, and between public and private.[87]

In the current situation, right to privacy is not only battered by the nation state with the pretext of security, but also by the commerce collecting more-and-more data. Commerce is interested in collecting and analyzing data, since it can reveal spending patterns, and be used for profiling and predicting future behavior of consumer groups. The governmental interest is different: the interest is in behavior that is against the law or potentially a risk to security.[88] As mentioned, in the US context privacy is especially seen as a right not to be under the surveillance of the state, which has no right to access your data without having a legal basis for interfering. As Lessig puts this:

> "The traditional question of "privacy" was the limit the law placed upon the ability of others to penetrate your private space."[89]

This Lessigs view is at par with the traditional approach to privacy as the right to be left alone. Saarenpää has also advocated this view.[90] This is, however, a problematic definition

---

data might cause discrimination of consumers, see for example, Persis, McLauglin and Levy, 2014, p. 29 - 30.

[83]Viljanen, 2011, p. 139.

[84] Brinhack and Elkin-Koren 2011, p. 341 and Craig and Lundloff 2011, p. 18 - 19.

[85] See Swedish Data Protection Act, Section 1; states the purpose of the act is to protect the personal integrity – '*personlig integrihet'*.

[86]Warren and Brandeis, 1890-1891.

[87]Bennett and Raab, 2006, p. 5 and Warren and Brandeis, 1890-1891 pp. 193, 219 - 220, see also Bygrave, 2002, pp. 129 - 130.

[88] Commerce is generally interested about a pattern of consumption that can predict future behavior and help advertising. See i.e., Lessig, 2006, p. 216.

[89]Lessig, 2006, p. 201

[90]Saarenpää 2009a, p. 371.

for privacy, when we take into account the development of the Internet and technology that allows easy surveillance as well as profiling and direct marketing. Rule has criticized the view and stated that technology itself dangers privacy, stating that the actual problem is not the fact that technology allows us to absorb, analyze, transmit and use personal data. The fact is that technology is always developed and used by humans. This view of Rule is well in line with the views of Lessig: the choice between privacy and no privacy can be made on the level of code and architecture. [91] This is the central idea of privacy by design (PbD). In practice, this means that privacy can be decided in the development of software and application. According to Lessig, the protection of privacy is hindered by the fact that there are no clear interest groups protecting privacy. There is a great amount of regulation to protect intellectual property, while in comparison privacy is lagging behind.[92]

The current legal literature seems to see the core of privacy in the right to be left alone, or personal autonomy. In the personal integrity model, much weight is placed on the protection of personal autonomy. Raab and Bennett have noted that current protection privacy protects privacy as an individual right, instead of seeing it as a larger societal safeguard for democracy.[93] The idea of privacy commons does not take privacy as a purely individual right: this challenges the concepts of *autonomy* and *informed consent*.[94] Regulation of a commons, with the idea of protecting democracy and privacy as a collective right of the society, would require different too. In my opinion, it would be good to view privacy also as a collective right, since it would help balancing the free market and right to privacy. A great deal of examples could found from environmental legislation.[95] This type of regulation would better take the need for balancing the need for using data.

The use of data as a currency is a commonplace practice. The majority of legal scholars are, however, skeptical about property rights to personal information.[96] Additionally, as Raab and Bennett state, although privacy or personal information may be given value, the construction of ownership can hardly be constructed in the context of personal data.[97] It would be problematic if personal data would be seen as property, disposable accordingly to

[91]Rule, 2007, p. 19.
[92]Lessig 2006, p. 200.
[93]Bennett and Raab, p. 23.
[94]For the separate discussion about cyberspace as commons, Lessig, 2002, pp. 19 - 23 and also Lessig, 2006, p.198.
[95]Muth has viewed the protection of privacy and compared it to natural resources regulation, Muth 2009, p. 350.
[96] Although there are also varying opinions on this issue, see for example Schwartz, 2003-2004, p. 2057.
[97]Bennett and Raab 2006, p. 10 and Schwartz 2003-2004, pp. 2076.

the freedoms of contract. Also Saarenpää argues that individuals have the ultimate property right to the information about them, when we disclose the information – we do not disclose our property right to the information. I would not go as far with the rebalancing of privacy and property rights The individual's right to privacy needs to be protected in future also, but this could also be done by using a different approach – that would allow more of the economic activities online. In my opinion, the core of privacy resides in the protection of democratic institutions and not in the acquisition of consent from individual for each activity. According to Bygrave, the complexity of the concept of privacy helps balancing the right to privacy, against other rights.[98] The balancing should be done according to the rules of balancing fundamental rights, in the level of legislation.[99]

The European Court of Justice has balanced in several decisions the individual's right to privacy against other fundamental rights. These cases provide additional input on the discussion of balancing privacy against property rights. Although by far ECJ has not resolved cases that would treat especially the economic activities online.[100] The trend in the judgments has been however giving privacy more weight against other fundamental rights.

In the two most recent resolutions on data protection, ECJ has strongly emphasized the principle right of privacy and data protection. In the so-called data retention decision[101], ECJ decided that the current data retention directive is unconstitutional. This was because the limitations on privacy were too high and did not meet the principle of proportionality. The cases places high emphasis on privacy and safeguards of privacy. The case gave importance to the proportionality principle and stated that privacy must be taken in account even when there are security reasons for legislation.

Additionally, in the ECJ Google case, privacy received a high standing.[102] ECJ decided that Google is required to delete search results from the search index. The case is in several ways a landmark judgment in the area of data protection, since it shows a more activist role of the ECJ in the field of data protection. The most important part of the ECJ Google case, in the perspective of secondary use of data, is the decision to interpret the establishment of a commercial entity and thus increase the scope of applying the DPD. In the ruling, ECJ

[98]Bygrave has stated, that it may actually help when balancing different rights for example the freedom of speech, see Bygrave 2002, p. 127.
[99] Viljanen 2011, p. 141 - 144.
[100] Cases have been followed until the spring 2014.
[101] *ECJ Judgment in Joined Cases C-293/12 and C-594/12 Digital Rights Ireland and Seitlinger and Others*
[102]*Judgment in Case C-131/12 Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja González*

established that the Spanish DPA had authority over Google, even though Google Inc. is established in the United States and only had advertising sales in the EU through its European subsidiary.[103]

It is important to note that neither of these cases address the issue of balancing the principle rights behind free markets against privacy, nor balance the free movement of data against the free movement principles. The property rights aspect has been balanced in the case In ECJ Case *Promusicae v. Telefónica de España SAU*[104]. The case dealt with the question of protection of copyrights, and to what extent can privacy be limited in the context of protecting property rights. Promusicae wanted Telefónica to disclose information about those individuals who were using KazaA to download music. The question was about traffic data retained by Telefonica and whether it should for the protection of copyrights disclose such information to Promusicae. The court decided against the claims of Promusicae, stating that it is not the purpose of the directives E-Commerce directive (Directives 2000/31), Copyright directive (2001/29), IPR Rights Enforcement Directive (2004/48) or Directive on privacy and electronic communication (2002/58) that operators would be required to disclose traffic data for the protection of copyrights.[105] Although the ECJ Promusicae case could be read in the defense of privacy versus property rights, I would see the scope of the case much more limited. First of all the question was about balancing privacy against copyrights protection instead of property rights in a wider sense. Second of all the disclosure of individuals data for potential criminal liability cases in a copyrights violation does more harm to the individual than the average commercial secondary use of data. In addition to the ECJ case praxis, it is important to note the European Court of Human Rights decision Germany v Pauefgen, where the commission has treated the question of what is property, stating clearly that the protection of property includes also non-tangible assets.[106]

In my opinion, none of these cases sufficiently balances the *free market fundamental rights* and data protection. New ECJ judgments balancing these rights would be useful for clarifying the balancing of fundamental rights in the area of secondary use of data. In my opinion, it is clear that some room needs to be left for the data controllers rights' as well.

---

[103] ECJ C-131/12 Press Release No 70/14, p. 2

[104] *C‑275/06, Productores de Música de España (Promusicae) v. Telefónica de España SAU*

[105] *C‑275/06, Productores de Música de España (Promusicae) v. Telefónica de España SAU, sections 45 and 70 of the judgment.*.

[106] The case was about the right to domain names, and as such it does not require additional analysis. See more, PAEFFGEN GMBH v. Germany.

Privacy should not become an obstacle for the development of internet commerce. As stated, the conflict of privacy is itself created by the dual nature of the Data Protection laws, most notably the DPD; the controllers of data are in a position to both safeguard the data subject's rights as well as conducting business. The risk-based approach to privacy would solve the conflict of fundamental rights much more efficiently.

The following subchapter further analyzes the free movement of data and the current legal norms on data protection in Europe, including the important definitions of controllers and processors of personal data. These definitions are explored, so that the parties using data and the different roles for data processing can be understood. Prior to that, a further argument for balancing commercial interests and privacy is derived from the European Union free movement of data principle. The emphasis is on the factors that conflict the value of data. The norms preventing secondary use of data are also analyzed.

2.2 Free Flow of Personal Data

The European general data protection directive (DPD) been enacted to harmonize data protection in the common market. The directive especially protects the free flow of data as well as safeguards data protection and privacy.[107] The free flow of data within the common market is seen as a part of creating an effective, functioning free-trade area.[108] These flows could be seen as a part of the four freedoms, since the data flows may be in some cases necessary for safeguarding freedom to provide and receive services, especially online services. Free flow of data is safeguarded within the common market, even if the US corporate perspective is that in many cases data protection acts as an entry barrier to the region.[109] In my opinion, the free flow of data has such important role in the current online economy, that it should be considered as the fifth freedom of trade in the EU. The Free movement of data acts as an important background for the secondary use of data: it shows that the economic incentive has been behind the data protection directive from the start.

[107]The free flow of data and privacy can be seen complementary values rather than conflicting values, see e.g. Bennett and Raab, 2006, p. 95.
[108]For the freedom to provide services, see Barnard 2007, p. 354.
[109] Lee Bygrave summarizes the US based critique of protectionism, and sums the claims to two aspects; data protection laws cause restrictions to trans-border data flows and increase the scope of such restrictions. See Bygrave, 2002, p. 114.

Free movement of data and fundamental rights are safeguarded in the article 1 of the DPD. Article 1 of the GDPR also protects the free movement of personal data.[110]

The DPD also contains the important definition of personal data, which is one of the important concepts of this thesis, since the purpose is to analyze the rules that restrict the commercial secondary use of personal data. The DPD article 3(a):

> (a) 'personal data' shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;

In the era of big data the concept of personal data is problematic, since it big portions of data fall under this category. This means that any identifiers or potential identifiers can be seen as personal data, which alone causes problems for the ever-expanding use of data.[111] An identifiable natural person is considered a data subject.[112] Any kind of data relating to an individual may be personal data, this includes, for example: address, social security number and IP-address. The form of the data is not important, it may be images, spoken communications, or in text format.[113] When the data controller processes personal data as part of analytics or profiling the data protection laws apply to the processing.

In addition to personal data, also processing of such data has been defined in the DPD. Most of the commercial secondary uses can be classified as processing of data, as the definition in DPD article 2(b) shows;

> Processing of data means any operations performed to personal data, whether the operations are concluded in automatic means or not, including such activities like collection, recording, *organization*, *storage*, *use*, *alignment* or *combination*, *adaptation* or *alteration* and also such activities like consultation and destruction of data." –cursivated by the author

The definition includes all activities. It is important in the specific context of this thesis since the secondary activates fall under the definition. The secondary use of data starts with the storage of data. Further activities include, for example, combination of data sets for gaining valuable insights on consumer behavior or preferences, this could be done by combining data from various sources. According to Kuner the processing should be understood as it is defined in the directive.[114] The definition does not distinguish between

[110]GDPR, p. 40.
[111]Ohm claims that the concept of PII or personal data is problematic, since it expands the DPD to many areas, Ohm 2009-2010, p. 1741.
[112]Kuner 2007, p. 76 - 77.
[113]FRA 2013, Handbook on data protection law, p. 42.
[114]Kuner 2007, p. 74 - 75.

data processed by automatic measures and not. This means that manual data analytics also falls under the category of processing.

The secondary user of data is often a data controller; in the abovementioned classification of this thesis, in two of the situations explored the party conducting the activity is a data controller. Data controller is the party that collects personal data. Controller is the party that controls and collects the personal data. The term deriving from the DPD is fairly similarly implemented in the UK Data Protection Act, article 1, Spanish LOPD, article 3 (d) and Finnish Personal Data Act section 3(4).

WP 29 states, that being identified as a controller, is an important first step defining who has responsibilities for the data.[115] The controller carries the responsibilities and has obligations, especially since articles 10 – 12 of the DPD have been written in a form that gives obligations to the controller. These articles contain rights for the data subject, and more precisely: rights to information, access, rectification, erasure and blocking, and the right to object to the processing of personal data. According to WP 29, this means that the concept defines who is primarily responsible, and thus the concept of controller allocates responsibility.[116] This definition of being the controller or processor is especially important in the scenario where a third-party analytics company provides services, which can be classified as secondary use of data. The party could be defined as the controller of the data and thus vesting the responsibility of such activities to the external third party.

According to the DPD article 2(e) processor is the party processing the personal data, on behalf of the controller. The most important consequence of being a controller or processor is the legal responsibility to comply with the obligations set by data protection law.[117] It is important to note that legal responsibilities might differ in some cases; however, in nearly every scenario the controller is responsible for the acts of the processor.[118] However, usually it is not easy to define whether a party actually is a controller, especially when the entities are multinational corporations with complex legal structures and multiple databases.[119]

---

[115]WP 29, Opinion 1/2010 on the concepts of "controller" and "processor", p. 4.
[116]In addition, the concept of controller helps to identify which is the applying national legislation, since the seat of establishment is of crucial importance when selecting the applicable national law. See also WP 29, Opinion 1/2010, pp. 4 - 5.
[117]Handbook on European data protection law, p. 49.
[118]It must be noted that the processing is usually is done in basis of an agreement that might shift the liability in form of potential compensations.
[119] Kuner 2007, p.70.

The distinction between controller and processor has also become increasingly blurry; because of technology allows joint-activities on data. The relationships are also changing fast. This is, according to Kuner, problematic, since in many scenarios the controller is responsible for many actions, such as giving notices and complying with the legislation.[120] My view is that this will not become easier in future, as varying data activities, such as profiling, analytics, outsourcing and sale of data and transfers will further blur these roles. The proposed European Regulation will not solve the conflict; in fact, the definitions will remain nearly unchanged.[121]

The following subchapter will open the concept of secondary use of data, which is derived from the purpose limitation. The subchapter will also provide three example scenarios of secondary use of data.

2.3 Secondary Use of Data

In the European context, secondary use of data is the use of data for other purposes than the ones specified when the data was collected. The *secondary use of data* is not specified in the DPD or in national law.[122] Under the DPD other uses than the initial use for a certain purpose are restricted by article 6 or *purpose specification* and *consent* article 7(a). The new Proposal for General Regulation does not provide answers to such secondary use, which would be needed for the big data context.[123] In the big data context, it is important that data can be used for other purposes than the one specified, and according to Mayer-Schönberger often the value of big data derives from the combination of different data sets.[124] *Prima facie*, the DPD acts in a preventive way against secondary use of data.[125]

The *commercial secondary use of data* is another concept used in this thesis. It encloses all the commercial activities that might be done with personal data collected. With the term

[120] Kuner 2007, pp. 72 - 73.
[121] GDPR, p. 41, article 2.
[122] In the US context, secondary use of data can be viewed as the uses of data in which the data subject has not consented to. See for example Daniel J. Solove, 2008, 131. Also see, for example, WP 29 for the repurposing of data use, Opinion 03/2013, pp. 20 - 21. WP 29 has used the term 'further' processing instead of secondary use of data.
[123] This is especially problematic since there are also beneficial effects for the consumers in the secondary use of data, these can for example fall of prices also the secondary market for data can create positive externalities to consumers. These externalities include for example convenient use of multiple services or relevant information about products, Acquisti 2009, p. 10 - 11.
[124] Mayer-Schönberger and Cukier 2013, pp. 106 - 107.
[125] Solove has stated the US view on limitations of secondary use of data is much 'less comprehensive' than the European equivalent see Solove, 2008, p. 133.

commercial, for example, research and medical research are excluded from the definition. Activities included are, for example, profiling, behavioral analytics and marketing, and customer feedback and personalization of services. Also combination of data archives would fall under this category. Predictive analytics would most likely also fall under this category. In the modern day internet economy, multiple data-based economic activities can be seen as secondary uses of data. The ICO has listed several activities that are big data analytics that would involve the use of personal data; these include location data, purchases data and loyalty card data. In addition the ICO also notes that due big data there is the possibility for the creation of new data.[126]

The possible secondary uses of data in the scope of my research question can be classified into three categories. Classification is based on the different data controllers or processors involved in the secondary use of the data. The first example is *internal secondary use of data*, in which data collected is used internally for purposes other than originally specified. This could be i.e. analytics or internal marketing operations with the data. For example, the data controller could collect data for providing products to customers or to survey their satisfaction, secondary use of the data could be analytics of customer behavior, which could then be used for targeted marketing.

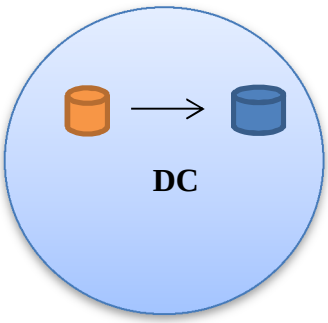
The second category is a so-called controller-to-controller, secondary use, in which the data is transferred from the initial controller to another, who then controls the data and conducts operations on it. An example would be the sale of data from a car retailer to the factory that builds the cars. In this question, the risks are in the transfer of the data and in the internal data processing of the manufacturer. The question is of joint controllership of the data, which has been analyzed in the previous chapter. In case of international data transfer the rules for international data transfer would apply, most notably the national implementations of the articles 25 and 26 on international data transfers. Kuner has analyzed data transfers.[127] In the scope of the European single market, the data transfers can be done freely, since the EU acts as a single market in relation to data as well.[128]

In the third category, data is processed on behalf of the original controller, by a subcontractor, who, for example, offers marketing or data-analytics services. In this

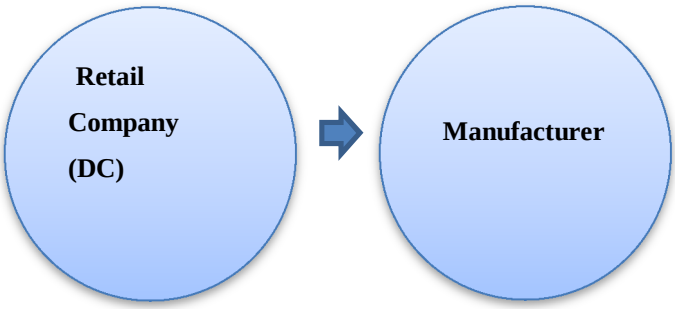
[126] ICO 2014, Big data and data protection, p. 11.
[127]For the scope of this thesis the data transfers cannot be treated, a short introduction would not cover the complexity of this topic and a loner one would not fit the scope of a master's thesis. See more in Kuner, 2013
[128]It should however be taken to account that the purpose restriction and legality of data processing has to be fulfilled in the context of data transfers as well.

question, the risks are same as in the previous ones, however, taking into account that the subcontractor is acting with a mandate given by the controller, and the need for relevant contractual clauses in place.

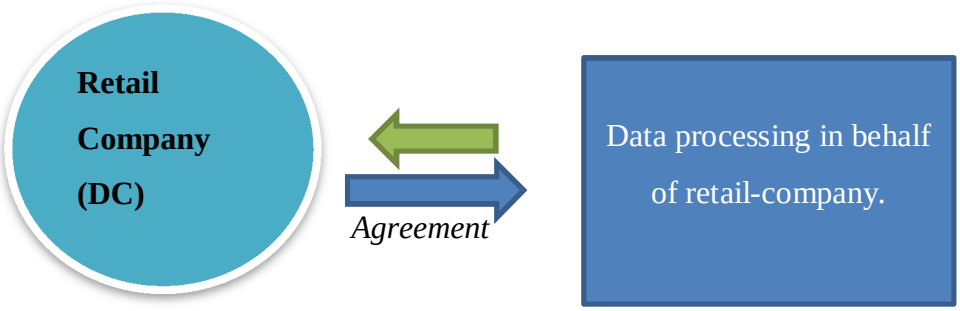
The following picture 2.1 shows three different scenarios of secondary data usage:



1. Internal secondary use of data



2. Data transfer and secondary use



3. Analytics of retail data by external analytics company

The purpose restriction and minimality apply to all of these scenarios. Also *Prima facie*, it seems that the GDPR does not either solve the problems in connection with the secondary use of data. The approach remains overall as it was previously; the definitions of personal data, data controller and processor all remain same.[129] A special emphasis is given to those principles entwined with the problems caused by the value of data, big data and anonymization, especially the purpose restriction.

2.4 Purpose Limitation and Minimality

The big data use of personal data needs to be compliant with the principles of data protection laws, which implement the DPD in the Member States. Most important

[129]GDPR, p.41 - 42.

principles in connection with my research question are: legitimacy of data collection, minimality and purpose limitation or specification, all of these principles are in a problematic and challenged by the technologic development.[130] Data minimality and purpose limitation are analyzed in this subchapter, fairness and lawfulness of data collection is analyzed in the next subchapter. Other principles include, for example, the accuracy of data, security of the data, informing the person concerned and the accountability of the data processor.[131] These principles are important as they affect the data controller and give rights to the data subject. Because the scope of this thesis, I will not open these principles more, since they not provide answers on the secondary use of data.[132]

The principles of the DPD are pan-European, but the national implementations of these principles might slightly differ. As also, Bygrave states, the member state implementations vary.[133] This further complicates the aforementioned conflict between value of data and the purpose specification; since data collection for marketing and profiling purposes can easily be cross-border and the purpose limitation may be differently interpreted in different European jurisdictions.[134] The following analysis is done based on the DPD principles, and national implementations will be viewed as examples of the implementation. These principles need to be accounted when planning activities that fall under the scope of the data protection directive "automatic processing of personal data" (or its national implementations), such as the secondary use of data. The goal of the core principles is to safeguard personal autonomy of the data subjects and privacy.[135]

The big data use of personal data could be conducted compliant to the DPD if the purpose of collection could be stated in a vague enough way, stating that data is collected for all purposes and for unlimited time and at the same time ensuring that the processing of data is legitimate. However, this would go strongly against the fundamental right to privacy and

[130]Mayer-Schönberger and Cukier 2013, p. 155.
[131]Bygrave, 2002, p. 2.
[132]In many cases, the information security is an important condition that the commerce needs to fulfil if they do not want to suffer high PR-damages, Innocenzio, 2014.
[133]Bygrave, 2002, p. 61.
[134]My view is that, since there is a proposal for harmonizing the data protection across EU, it is not important to venture more in the differences – the more important part is the analysis of the elements that will remain same even if the proposed GDPR will pass into force.
[135]It should be noted that in the area of data protection principles data protection law comes close to administrative law. The principles govern the data processors activities, much in the same ways as the principles for good governance. Bygrave 2002, p. 167.

against the whole notion of purpose limitation.[136] The purpose limitation is in its place to protect the consumers from unexpected data uses and also against aimless collection of big amounts of data, for the reason of potential future processing. The goal is also preventing the data creep or function creep, which means the expansion of data uses or blurring of the purpose restriction.[137] However, for the business it is seldom possible to define all of the possible purposes for the collection of data prior the collection of data. This arises the following questions; can there be multiple purposes, how vague can the purpose be and is the *repurposing* of data possible. *Repurposing* means that the purpose is changed from the original purpose.[138]

Purpose limitation answers the question of why data is collected and how is the data used, and for what purposes. It also sheds light to the types of data collected as well as the retention times of the data and processing operations.[139] The purpose limitation extends the whole lifecycle of data use, starting from pre-collection and extending to the disclosure and secondary use of data. The principle is strongly entwined with concepts of notice and consent.[140] Purpose specification means that personal data shall be collected for specified and lawful purposes, and it shall be processed within the purpose of collection.[141] If data is used for other purposes than the specified ones, it causes the data processing to be unlawful.[142] This means, for example, that if data were collected for delivering items from e-commerce, the data should not be for marketing, if the data subject has not been aware of that purpose. This is problematic since, there is a commercial incentive to use the primary data for secondary purposes and even sell the data.[143] The principle of purpose specification derives from the article 6(1)(b) of the DPD:

> Article 6.
>
> ---
>
> (b) collected for *specified*, *explicit* and *legitimate* purposes and not further processed in a way incompatible with those purposes. Further processing of data for historical, statistical or scientific purposes shall not be considered as incompatible if Member States provide appropriate safeguards.

[136]In addition, it has been stated that the purpose defined cannot be defined in such vague way.
[137]Brouwer 2011, p. 274 and 277.
[138]ICO 2014, Big data and data protection, p. 9.
[139]WP 29 Opinion 03/2013 on purpose limitation, pp. 11 - 12.
[140] Notice is given to show the purpose to the data user. In the Finnish context for the planning of data processing, see Innanen and Saarimäki 2009, pp. 91 - 92. See also Finnish DPA, Tietosuojavaltuutetun pohja tietosuojaselosteen laatimiseksi.
[141]Bygrave 2002 p. 61.
[142]WP 29 Opinion 03/2011, p. 36.
[143]Mayer-Schönberger and Cukier, 2013, pp. 104 - 106.

The article 6 means that data should be collected and used only for the purpose specified to the data subject.[144] The purpose specification can be separated into three different sub-criterions: firstly, the purpose needs to be specific, secondly explicit and thirdly legitimate.[145] The first criterion means that the data needs to be collected for a specific purpose, and unnecessary data should not be collected. The second criterion means that not only the data controller should know the purpose, but also it needs to be clearly expressed in a clearly understandable format to the data subject, authorities and third parties. The third criterion of legitimacy is connected with article 7 of the DPD, which contains the legitimate basis for data processing. WP 29 extends the requirement stating that the purpose needs to be in accordance with all data protection laws in every scenario and phase of the data collection and use.[146] In fact, according to WP 29 articles 6 and 7 are cumulative, both criterions have to be met simultaneously – it is not enough that secondary use is legitimate; it has to be also done accordingly with the predefined purpose.[147]

The Finnish Data Protection Act implements the purpose restriction in section 6 and the exclusivity of the purpose in section 7. Defining the purpose for data processing should be done prior to collecting the data:

> Section 6 Defined purpose of processing
>
> It must be appropriate and justified to process personal data in the operations of the controller. The purpose of the processing of personal data, the regular sources of personal data and the regular recipients of recorded personal data shall be defined before the collection of the personal data intended to be recorded in the file or their organisation into a personal data file. The purpose of the processing shall be defined so that those operations of the controller in which the personal data are being processed are made clear.

The exclusivity of purpose restriction is the most problematic part of the purpose restriction. This is defined in the section 7 of the Personal Data Act:

> Section 7 — Exclusivity of purpose
>
> Personal data must not be used or otherwise processed in a manner incompatible with the purposes referred to in section 6. Later processing for purposes of historical, scientific or statistical research is not deemed incompatible with the original purposes.

The Spanish article 4 of LOPD and Schedule 1 of the UK Data Protection Act contain similar purpose restrictions. The proposed GDPR article 5 also contains the principle

[144]This also fulfils the controller's duty to inform the data subjects.
[145]Brouwer interestingly distinguishes the purpose restriction differently: to five different aspects. See more, Brouwer 2011, p. 277 and Bygrave 2002, p. 338.
[146]Bygrave 2002, pp. 338 – 339 and WP Opinion 03/2013, pp. 15- 16 and 19 - 20.
[147]Working Party 29, 2013Opinion 03/2013 on purpose limitation, 36.

purpose restriction.[148] There are exemptions to the purpose restriction for statistic and research purposes. Quite clearly, these exemptions permit the use of data for secondary purposes in specific scenarios; this may be extremely useful in the field of medical research, for example.[149]

Secondary uses (such as analytics or combining data) can be included in the purpose. The interesting question is whether such methods can be used on legacy data, meaning that the data has been collected prior such techniques for analytics existed. The vagueness of purpose, is an element of the purpose restriction the WP 29 has in fact stated in its opinion that it is possible to collect and process data for multiple purposes, they should, however, be specified. In addition, it is avoidable to have a broad purpose just to justify possible data uses that are not related to the core purpose.[150] This option to collect data for multiple purposes does to some extent ease the pressure caused by purpose restriction. It is, however, impossible to state and include all the possible uses of data and especially in the case of legacy data, the analytics and other secondary uses are not usually included in the purpose defined. WP 29 has also proposed changes to the current language of the DPD; the changes would further clarify the secondary use of data.[151] These factors hinder the secondary use of data.

However according to WP 29, purpose limitation is a crucial building block for privacy and the potential future uses can be included in data-collection phase.[152] The WP 29 argues that purpose specification is not in conflict with the recent development of the information economy. The working party states that purpose specification is a balanced approach and the problem is the disharmonized approach in the member states.[153] The opinion of WP 29 shows the problematic nature of the purpose specification: a vague purpose would best serve the commercial interest, but this cannot be accepted from the perspective of data protection. Especially the future processing and use of data is hard to include in the purpose and this conflicts with the economic value of information, as it is often impossible to know possible future uses of data. According to Mayer-Schönberger, purpose specification is a principle challenged especially because often the value of data is in the

[148]GDPR, p.43.
[149]For the scope of the research question, it is impossible to deeper venture to this topic of statistical, historical or scientific exemption to the purpose restriction. For more information on statistical research, see Duncan, Elliot and Salazar, 2011.
[150]WP 29 Opinion 03/2013, p. 16.
[151]WP 29, Opinion 03/2013, pp. 41 - 43.
[152]WP 29, Opinion 03/2013, p.3.
[153]WP 29, Opinion 03/2013, pp. 4 -5.

secondary, rather than in the primary use.[154] In the light of the WP 29 opinion on repurposing of data use Mayer-Schönbergers claim might in fact be a slight overstate. WP 29 and ICO have stated that that the future changes to the purposes to the processing do not need to be completely compatible with the initial purpose as long as they are not incompatible with the purpose, this gives some flexibility to the controller for changing the activities conducted with the data.[155] According to ICO the repurposing of data use can be done, if the new use is not unreasonable considering the expectations of the data subject.[156] WP 29 has also given out specific guidance for repurposing, which is especially crucial for the secondary use of data, in cases where the secondary use would be done on legacy data. I see that WP 29 has found a good balance to the principle of purpose restriction and their suggestion for the modification of the purpose restriction would serve as a good hotfix for data use in the era of big data. In my opinion, the purpose restriction has an important role in the future of data protection since; in all the vagueness, the principle also has flexibility.[157]

In addition to the purpose restriction, data minimality set in the article 6(1)c also greatly affects the commercial secondary use of data. The principle affects the quantity of data collected and the retention times.[158] The data minimization of the DPD can be interpreted so that the processing of personal data must be restricted to the necessary minimum to achieve the purpose.[159] Especially the temporary effect is important in the context of big data, where there would be an incentive for collecting data as long as possible. The other aspect of data minimization deals with the act of collecting and storing the data, which should be conducted in as minimalistic form as possible. This can also be interpreted to mean the infrastructure for processing as Kuner has pointed out.[160] In its guidance, CNIL has given an important role to the principle of data minimality, stating that it is one of the important prerequisites for data processing.[161] In the Finnish context, Data Protection Act Section 9(1) contains the limitation for collecting other data than needed. Vanto has analyzed this and stated that data other than the defined data needed does not need to be collected. The data minimality affects the data retention times, since according to this

[154]Mayer-Schönberger and Cukier, 2013, pp. 151 - 152.
[155]ICO 2014, Big data and data protection, p. 22 and WP 29 Opinion 03/2013, p. 21.
[156]ICO 2014, Big data and data protection , p. 22.
[157]Although this flexibility can be lost if the purpose restriction is applied too strictly, as for example Brouwer would suggests, Brouwer 2011, p. 274.
[158] Bygrave 2002, pp. 59 - 60 and ICO 2014, Big data and data protection, p.23.
[159]Kuner, 2007, p. 74.
[160]WP29 1999, Protection of Individuals and Kuner 2007, p. 74.
[161]CNIL 2012b. pp. 5 - 7.

principle data should not be retained after the possible use for data is fulfilled. For example, if phone numbers or addresses and other personal data were collected for an e-commerce for shipping, and disclosed to a courier service, the courier service should not retain the data after the service has been carried out.[162] Data minimality greatly safeguards the fundamental rights of the data subject; it lessens the risk that data would be used against the fundamental rights. The big data context however creates the need for storing data for a long time, since as stated; the value of data can be often discovered after a long time. For example, the courier company could use the data for analyzing routes and increasing efficiency. According to ICO anonymization may act as a solution for retaining personal data for longer periods and not violating the data minimality principle.[163]

There are however certain problems with the purpose restriction and data minimality principle which, restrict the uses of data. Some restrictions of data use are needed for safeguarding the fundamental right to privacy. The problem is however that more data is needed for analytics and almost anything can act as an identifier, this creates a scenario where a whole database needs to be processed accordingly to the limitations set by the data protection laws even though there not so many identifiers in the database.

The next element analyzed, is the legitimacy of processing data, and what is the required basis for such data processing. The legitimacy of data use is located in the center of the conflicting rights of privacy and freedom to conduct business. Consent representing the ultimate autonomy and legitimate interest being a balancing test for the data controllers' interests and individuals rights. The requirement that data processing should be legitimate is well in line with the purpose restriction as well as a data protection principle. The analyzed article is especially DPD 7, which is left untouched in the proposed GDPR. The recent WP 29 opinion further clarified the wording of purpose restriction also proposing the removal of unclear wording, which would more clearly lead to the interpretation that both article 7 and article 6 need to be fulfilled cumulatively by the data controller so that the data processing is legitimate.

[162] Vanto, 2011, p. 53.
[163] ICO 2014, Big data and data protection, p. 24.

2.5 Legitimate Commercial Secondary Use of Data

Since secondary use of data is data processing, it must meet the legal basis for data processing. The DPD states that personal data may only processed when there is a legal basis for processing the data. The different bases are listed in the article 7 of the DPD. The two bases for data processing analyzed in this thesis are consent and legitimate interest for processing; both of these bases have their benefits. From a fundamental rights perspective, consent best ensures the personal autonomy of the data subjects.[164] From the viewpoint of a data processor, consent is, however, a burdensome way to gain compliance with the relevant data protection laws. This chapter concentrates on the article 7(a) (*consent)* and on the article 7(f) (*legitimate interest*).[165] Other legitimate basis for data processing include, for example, that processing is based on contract, vital interests of data subject or in public interest and the processing is done by an official authority.[166]

Consent has been defined in the DPD. Consent is one of the bases that acts as a general ground for the legitimacy of data processing, as well as a special ground in several cases, most notably article 8 cases where processed data is considered as sensitive.[167] The DPD article 3 defines consent and sets the conditions for valid consent:

> "(h) 'the data subject's consent' shall mean any *freely given* specific and *informed indication* of his wishes by which the data subject signifies his agreement to personal data relating to him being processed." (Curs. by the author)

The most important aspect of consent in the context of commercial secondary data use is the scope of given consent. How explicit must the consent be so that it allows the use of data subject's data in different areas? Another important question is the temporal-effect of consent: when does the individual need to give his or her consent for the activities?

WP 29 has defined consent more in their clarifying opinion on consent about the four valid aspects of consent; it needs to be a clear and unambiguous indication of wishes, freely given, specific and informed.[168] WP 29 has further clarified the requirements of consent:

> "There is in principle no limits as to the form consent can take. However, for consent to be valid, in accordance with the Directive, it should be an

[164]Saarenpää 2009a, pp. 410 - 411.
[165] It must be noted that the previous discussion of the concept of controller applies also to this requirement 7(f) see e.g. WP 29, Opinion 1/2010, p. 5.
[166] Innanen and Saarimäki 2009, pp. 89 - 90 and Vanto 2011, pp. 46 - 47.
[167]Sensitive data will not be treated as part of this thesis since it falls out of the scope of this research.
[168]WP 29, Working Document 114, p. 67.

> indication. Even if it can be "any" form of indication, it should be clear what exactly can fall within the definition of an indication."[169]

However, it should be taken into account that according to WP 29, passivity cannot be seen as an indication. The European Commission has adopted a similar view, stating in the recital 25 of the proposal regulation that passivity cannot be accepted as an indication.[170] The requirement for a freely given and informed consent varies across jurisdictions. Especially in the case of online services, the requirements for consent may be different; in some jurisdictions, opt-in is required. In practice, in many Internet services the consent is given by ticking a box. In other jurisdiction passivity, such as using a service can be seen as consent.

Consent and purpose restriction have a problematic relationship in the environment of secondary commercial data use. The data subject consents only for the specified purpose.[171] This creates a strong interconnection between the purpose restriction and consent. The data subject needs to be informed well of the purpose of data processing. However, this might be problematic in the big data context, where there is an interest of changing the purpose later on, or defining the purpose in a vague manner. The data subject cannot consent to activities that are not clear, since this would not fulfill the requirement of *informed consent* and thus the personal autonomy and control aspect of consent. The individual has a clear interest of knowing, so that they may have control over the activities.

Another problematic area of consent is the temporality of consent. The temporal effect of consent requires that it must be acquired prior to processing.[172] In addition, if the individual informs the processor later on that they do not want their data processed, the processing must cease, since losing the consent would mean losing the legal basis for processing. The GDPR article 7 would further clarify this by stating that the individual may in any time withdraw their consent and that this does not affect the legality of the processing done prior the withdrawal of the consent.[173]

Consent is also a notion in contract law, in this area it is especially analyzed from the perspective of validity. WP 29 has clarified the relationship between civil law consent and DPD stating that the directive does not address the conditions of validity of consent from

---

[169]WP 29 Opinion 15/2011, p. 11.
[170]WP 29, Opinion 15/2011, p. 12 and GDPR, p. 21.
[171]Kuner, 2007, p. 68. For the Finnish context on consent see, Vanto 2011, p. 44 - 45 and HE 96/1998, p. 38 - 39.
[172]WP 29, Opinion 15/2011, p. 31.
[173]In addition, the article 19 of the GDPR contains the right to object to the processing of data. In my opinion, this will in the future create additional problems in the area of big data processing, GDPR, p. 45 and 56.

the perspective of civil law. There is an overlap, which means that also the conditions of civil law apply.[174] This interconnection with contract law should be taken into account. Although consent is valid from the perspective of the DPD it might still be invalid if it can be shown to be invalid from the perspective of contract law. This could be for example due to lacking competence, which can be caused by the fact that the data subject has been a minor or lacking the mental capacity for making a valid agreement. In this area, there is a strong interconnection between data protection and consumer protection.

This way of justifying data processing is extremely problematic in the era of big data for two reasons: it is not effective, and it does not give real autonomy to the subject, as it might be hard not to consent for the use of everyday services that need your data.[175] It could be even claimed that consent does not work for any of the parties included in the transaction.[176] For active purpose limitation the consumers would need to know of the purposes the consent is not valid if you do not know what you are consenting to. Even so the importance of consent is amplified by the fact that consent legitimizes nearly all uses of data.[177] The basis is the consent of the data subjects, legitimizing any kind of data use, this is however problematic. In many cases, consent is given without much consideration. For example behavioral economic studies indicate that people are not equipped to deal with the tradeoff of short-term gains and long-term loss of privacy.[178] As an economic argument to the issue of consent, Posner, has stated, that since the costs of disclosure of information to individuals are small, but since the costs of obtaining consent from the individuals for the secondary use of the information are high, such consent shouldn't be necessary to obtain.[179] Also already in the area of data transfers, the WP 29 has noted, that in many cases, it may be burdensome to rely on consent when carrying out regular data transfers and thus saying that consent may be a false good solution.[180]

We must take into account, the two major problems with consent; the temporal effect, and how to get explicit consent for activities that might not be clear. The most applicable solution would *prima facie* seem be the DPDs article 7(f), which is often referred as the legitimate interest ground for data processing, or as connection requirement in the Section

[174]WP 29, Opinion 15/2011, p. 6.
[175]Rule has stated, the we cannot really say no for most to the services that as a requirement for use need our consent and approval, for example opting out for the use of credit card or bank account. Rule 2007, p. 170.
[176] Mayer-Schönberger and Cukier, 2013, pp. 154 - 155, 173.
[177]Solove 2012-2013, p. 1880.
[178] Acquisti, 2010, p. 6.
[179]Posner 1977-1978 p. 398.
[180]WP 29, 2005, *Working Document 114,* p. 11.

8(5) of the Finnish Personal Data Act.[181] In cases where the data controller has a legitimate interest, they might process the data. By using the article 7(f) and conducting the balancing test correctly, with additional safeguards (such as anonymization for example), data processor may be able to conduct commercial secondary use of data, requiring that the purpose restriction of article 6 is fulfilled. The WP 29 has recently published a guideline on the article 7(f), firstly correcting the popular misconception that 7(f) would act as a last resort – instead it is stated that the article has its own specific uses and that it should be treated as an equally good ground for data processing.[182]As stated by WP 29, it is better to use article 7(f) as grounds for processing instead of misusing other legal grounds:

> "Appropriate use of Article 7(f), in the right circumstances and subject to adequate safeguards, may also help prevent misuse and over-reliance on other legal grounds"[183]

The article 7(f) acts as a balancing test that takes into account the legitimacy of the interest of the data controller, and balances it against the fundamental rights of the data subject.[184] The balancing is done on a case-by-case basis, since the article would otherwise have a very broad scope of application.[185] Prior to conducting the balancing test, it must be explored what could be considered as a legitimate interest of the processor in the commercial secondary use context. In addition, the rights of the data subject need to be balanced. First, the interests of the processor need to be clearly defined. For example, WP 29 has listed economic benefit of the processing one of such interests; a company has, for example, the economic interest of knowing the customers as well as possible.[186] The balancing test may offer some relief to the data controller wishing to conduct secondary activities.

WP 29 has stated that the interest of controller is legitimate as long as the controller may pursue their interest in accordance with the data protection and other applicable laws, such as consumer protection laws. According to WP 29, online and offline marketing activities can be seen as legitimate, as long as appropriate safeguards for the consumers are in place.

---

[181]Interestingly in the Finnish law this has been translated as *connection requirement* "asiallinen yhteys" which could be understood differently as the legitimate interests. It could in this way question whether the Finnish law fulfils the DPD in this area. This area is originally from the previous law in force. More on the scope of this requirement see, for example, Vanto 2011 pp. 47 - 48 and HE 96/1998, p. 40.
[182]WP 29 Opinion 06/2014, p. 10.
[183]WP 29Opinion 06/2014, p. 9.
[184]This balancing test resonates well with the fundamental rights super-question, which is the big picture conflict of the interests in this thesis. See more in the previous chapters 1.1 and 2.1.
[185]WP 29 Opinion 06/2014, p. 22.
[186]Other interests would include for example societal benefits of the processing, but I am doubtful of such benefits in the area of commercial secondary use of data, WP 29 Opinion 06/2014, p. 24.

WP 29 has stated that this does not mean that the controller could combine under this basis a vast amount of data collected in different contexts under different purposes, and create extensive profiles of consumers, and for example, sell these profiles. This would, according to WP 29, likely act as a significant intrusion into the privacy of the data subjects, and thus be unjustified and weight the balancing against the interests of the data controller.[187] *Prima facie,* the WP 29 opinion could be read to allow certain minor commercial secondary activities on data. However, activities that are not transparent and do not contain the possible safeguards are not allowed, and cannot be constituted to balance the scale in favor of the data controller. For example, profiling that leads to price discrimination, without the necessary transparency and right to object, is not allowed. The balancing will be done based on the interest of the data controller, against the fundamental rights or interests of the data subject. The balancing test is done by assessing four aspects, as stated by WP 29. The aspects are first of all controllers' legitimate interest, impact on data subjects, provisional balance and additional safeguards.[188] I will now analyze the criterion set by WP 29, in the light of commercial secondary use of data.

The data controller's right to conduct business and the right to property is the right in the other side of the balancing test. It safeguards the data controllers' right to conduct online business and use data processing as part of the economic activities conducted. According to WP 29, data processing needs to be necessary and proportionate.[189] At the side of the individual, the protected right is the right to privacy, and more specifically the nature of the personal data processed, as well as expectations of the data subject. For example, in the scope of secondary use of data for profiling, the negative emotional effects of such profiling need to be taken into account.[190] How can the controller then take the balancing test and what are the important steps. According to WP 29 first, the interest of the controller needs to be real and the purpose needs to be legitimate. The effects on the data subject need to be considered and additionally the method of processing needs to be assessed. The important question is, whether the goals can be reached in other ways that are less intrusive to the data subject. The activities should be conducted in a way that affects the rights of the data subject as few as possible. The whole balancing process needs to be documented; the data controller must be able to demonstrate that they have in fact

[187]WP 29 Opinion 06/2014, pp. 25 - 26.
[188]WP 29 Opinion 06/2014, pp. 31 - 33.
[189]WP 29 Opinion 06/2014, p. 34.
[190]WP 29 Opinion 06/2014, p .36.

made the balancing of rights and considered alternatives. In addition, the result of such balancing test should never be completely against the reasonable expectations of the data subject.[191] Anonymization is listed as one of the additional safeguards that may tip the scale – meaning that if data may be anonymized it widens the application of the article 7(f), more activities may thus be conducted with data that can be anonymized.[192]

Summarizing the issue of applying article 7(f) as the basis for commercial secondary use of data, I would like to address four points. First, the WP 29 guidance is still unclear in the whole array of data activities; there is a need for more case examples in the area. Second, it would seem that minor secondary use of data might be conducted in accordance with the article; major operations would not likely pass the balancing test. Thirdly, with the proper use of anonymization the scope of these commercial operations can be widened. Fourthly, in many instances WP 29 has highlighted the importance of considering the purpose restriction when using article 7(f) as a basis for data processing. Purpose restriction acts in an even stricter manner when combined with the said basis of legitimate interest. This may completely shut the commercial data activities, if they are not clearly defined prior to collection of data.

In the GDPR, the secondary use of data is further made harder, since the legitimate interest ground does not apply in a similar way. In fact, according to article 6(4), the further processing of personal data cannot be based on the legitimate interest.[193] As personal data legislation can be burdensome from an economic and administrative perspective, and there is no certainty that the new regulation will solve the problems and allow the agile use of personal data as part of data sets, alternate solutions have to be searched.

Following the article 7(f) balancing test and purpose restriction the potential area of secondary activities remains minor. Anonymization might act as a solution to allow more data uses. Anonymization has been already used widely connection with the statistical and research. In the case of research the values against each other is the development of science and privacy. Article 13 of the DPD sets the exemptions for the use of data in research. The next chapter will focus on anonymization in general, and especially on anonymization as a commercial practice as part of commercial data processing. A mainly legal dogmatic approach is adapted; additionally computer science will in this case provide the some input

[191]WP 29 Opinion 06/2014, pp. 39 - 40 and 54 - 55.
[192]WP 29 Opinion 06/2014, p. 42.
[193]GDPR, p. 43 - 44.

as consequentialist arguments. Certain points should be taken into account in relation to anonymization and this chapter. Anonymization is considered as processing of personal data, and it needs thus to fulfill the requirement of legality as well as fit into the purpose restriction. This creates an interesting situation where anonymization can be used as a safeguard, which acts as a scale in the balancing test of article 7(f), while at the same time a similar balancing test would be needed for legitimizing the process of anonymizing data.[194] In my thesis, the role of anonmyization is important, since it currently helps strongly to solve the conflicting fundamental rights of secondary use of data. For this reasons it is important to analyze the legal background for anonymization and the current problems with de-identification and valuation of data. Anonymization also protects the individual's privacy, allowing more data use, without putting the data subject at risk.

3 Anonymizing Data Allows Commercial Secondary Use

3.1 Anonymization as a Privacy Enhancing Technology

Simply put, anonymization is the process of rendering data containing personal information into a form in which it does not identify individuals, but in which the data is still usable.[195]*Prima facie,* it could be stated that anonymization solves most of the problems in connection with secondary use of data.[196] The commercial use of data could be based on the DPD article 7(f) and strong anonymization. Anonymized data is also not considered as personal data and the restrictions of DPD do not thus apply, if the anonymization is conducted in a strong enough manner. However, the technological development is a two-edged sword, at the same time it raises the value of data and thus grows the need for advantaged data processing operations, as well as heightens the risk of anonymization being broken by advanced technologies. As the UKAN states, anonymization is valuable since it allows data to be shared, without endangering privacy.[197] Anonymization is defined as the process of removing, obscuring, aggregating

[194]WP 29 Opinion 05/2014, p. 7 - 8.
[195]ICO, 2012 Anonymisation: managing data protection risk code of practice, p. 48.
[196]Especially Ohm criticizes this view stating that anonymization has let the legislators from answering to different conflicting interests of security, innovation and free flow of information. Ohm, 2009-2010, p. 1736.
[197]Sedayao 2012, p. 2. As stated in a practical sense anonymization allows the use of data in a way that privacy is not challenged. This is off course an over simplification of the issue. UKAN states a similar goal in their homepage – the goal is to maximize value of data and minimize risks to privacy, UK Anonymization Network.

and/or altering the identifiers that establish the data as personally identifiable or simply put as a process of making it impossible to identify data subjects.[198]

Recital 26 of the DPD indicates that anonymization is a possible from of protection of privacy, and that a code of conduct by national authorities may be established.[199] According to the DPD, Recital 26:

> **"**(26) Whereas the principles of protection must apply to any information concerning an identified or identifiable person; whereas, to determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person; whereas the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable; whereas codes of conduct within the meaning of Article 27 may be a useful instrument for providing guidance as to the ways in which data may be rendered anonymous and retained in a form in which identification of the data subject is no longer possible;"

The article 27 of the DPD gives guidance for drafting code of conducts, such as the ICO code of conduct on anonymization. The United Kingdom's DPAs code of practice is the first soft law legal instrument affecting anonymization comprehensively.[200] The new General Regulation has also taken a similar approach. Anonymous data is seen as out of the scope of personal data.[201]

Anonymization has also been treated in the e-Privacy directive (Directive 2002/58/EC), which treats questions regarding electronic communications. In this directive, anonymization is mentioned in two articles, 6(1) 'traffic data' and 9(1) 'location data', as well as in the Recital 26, where it is stated that traffic Data used for marketing should be rendered anonymous. It is important to consider these instances if the anonymization process falls under the scope of the e-Privacy directive.[202]

Anonymization models and techniques fall under the category of the so-called PETs (*privacy enhancing technologies*).[203] Privacy enhancing technologies are technologies, which by implementation protect privacy. According to Acquisti, PETs can protect privacy in a cost-efficient way, which still allowing the exploitation and use of data.[204]

---

[198]Duncan, Elliot and Salazar, p. 171 and Key Information, UK Anonymization Network.
[199] ICO 2012, Anonymisation: managing data protection risk code of practice, p. 10.
[200] In addition, the CNIL guidance mentions anonymization as one of the methods reducing risks in connection with personal data, CNIL 2012a, p. 24.
[201] GDPR, p. 21.
[202] For the purpose of this thesis it is however possible to further analyze the e-Privacy directives approach to anonymization. Although the directive affects many activities conducted especially in the space of operators, it is not in the center of the emphasis on commercial secondary use of data.
[203]ICO 2012, Anonymisation: managing data protection risk code of practice, p.7.
[204]Acquisti, 2010, p. 6.

Anonymization or de-identification is recommended as one of the default activities by the Privacy by Design-movement.[205] It should be noted that encryption is a different activity, although it is also a PET. They key difference is that anonymized data is still useful for a certain purpose, for example, medical research, statistical research, or even marketing, analytics and data sales purposes. Encrypting data, on the other hand, is a process in which data is rendered unreadable, so that it can be decrypted with the correct access code.[206] The definition of truly anonymized data is that it cannot be returned into its previous form, which contains personal identifiers.[207] This is not true with encryption, where it is actually in important to be able to return data into an unencrypted form.

As I have previously stated, the data protection laws apply on personal data as defined in the 3.1 article of the DPD. If data is anonymized, data is not considered personal data and it is considered to be outside the restrictions of the previous chapter. Finnish Personal Data Act does not define anonymous and pseudonymous data.[208] In comparison to anonymized data, pseudonymous data is data where unique identifiers are replaced with a symbol, string of numbers or a letter, for instance, a name is replaced with a connectable pseudonym. Pesoudonymized data and anonymized data have an important difference: pseoudonymized data is still covered under the data protection laws, unlike anonymized data.[209]The concept of pseoudonymized data is also included in the current proposal from the European Commission. According to WP 29, pseyudonymization can still also be a valuable safeguard for the data. Since it lessens the possibility that the data is linked to the data subject, it should not, however, be considered as a method of anonymization.[210]

WP 29 has recently published a complete analysis on anonymization and different techniques of analysis. In the opinion, WP 29 analyzes the risks and requirements of anonymization, further clarifying the use of anonymization techniques in the context of DPD. In the opinion, anonymization techniques have been classified into two families; first of the families or approaches is randomization, second is generalization.[211] *Randomization* means adding elements to the data set, thus reducing the risk of identification; it includes

[205]Cavoukian and El Emam, 2011, p. 1.
[206]Sedayao 2012, p. 6. It should be noted that encryption might be a useful tool when anonymizing, for example a certain part of a data set can be encrypted, which then anonymizes the data set, allowing later to open the anonmyization. This option does however include some additional risks.
[207]WP 29, Opinion 05/2014, p. 3.
[208]In fact, there is no official guidance in Finland about the anonymization of data.
[209]Kuner, 2007, p. 66.
[210] WP 29, Opinion 05/2014, p. 20.
[211] It should be noted that it is possible to choose different classifications for the techniques. Millard 2013, p. 170.

such techniques as *noise addition*, *permutation* and *differential privacy*. *Generalization* is the second approach, which aims to dilute the data or generalize it in a way that it is not possible to single out a data subject.[212] I will not approach further these technical questions or the risks that are associated to individual techniques. The analysis will remain on the more general level of anonymization and the usability of anonymization in the context of secondary use of data. It should be stated that in many cases several of these techniques can be utilized, thus adding extra levels of security to the data subject's privacy. WP 29 also strongly recommends a case-by-case analysis for each of these techniques, it is the data controller's duty to use appropriate safeguards and acknowledge the risks.[213]

WP 29 further analyzes each of these techniques and the risks included on the basis of three categories of risks, which are singling-out, linkability and interference. I will use to some extent the analysis in the following chapter, treating the question of re-identification.

As such, the concept of anonymization seems like godsend, resolving the whole conflict between privacy of the data subjects and fundamental rights of data processors. Anonymized data would fall outside of the concept of personal data and it would thus liberate the commerce from the need to comply with data protection laws, and leave them space for conducting commercial activities on data. At the same time, the principle right of privacy would be secured. It would appear; however, that such a solution might be an oversimplification, as the process of re-identification can nowadays be concluded relatively easily and thus break the anonymization of the data.[214] This has been shown in the cases of AOL and Netflix, which will be later described in more detail. The second problem for this hypothesis arrives from the commercial interest.[215] Anonymized data might not have similar value as non-anonymized data in the context of commercial use of data, which might lower the incentive of using anonymization. The key risks of anonymization as a safeguard for privacy are the failure of anonymization and devaluation of utility of the data anonymized. These two risks will be analyzed in the two following subchapters, starting with the risk of re-identification, which is from the perspective of data subject's fundamental rights more important. The latter analysis of the effects of anonymization on data utility or usability concerns the data processors more.

---

[212] *Generalization* consists of such techniques as *aggregation* and *K-anonymity*, and L-*Diversity/T-closeness* See for more information about the techniques in WP 29 Opinion 05/2014, pp. 11 - 18.
[213] WP 29 Opinion 05/2014, p. 24.
[214] There is some controversy upon the fact whether reidentification is actually an easy process or not.
[215] For example, Paul Ohm has claimed that anonymization and commercial use of data cannot be combined since the process of anonymization devaluates data. Ohm, 2009-2010, pp. 1714 - 1715.

3.2 Risk of Re-identification

The biggest risk with anonymized data is re-identification or de-anonymization. De-identification may happen via data intrusions, the mosaic effect or jigsaw identification. The de-anonymization can be done by an advisory that has the sufficient auxiliary information, which can be for example acquired from public data. De-anonymization means that the data is no longer anonymized and an individual may be singled out from the material. Re-identification may be conducted by bringing new information and combining it with the old data, which allows identifying or linking the anonymized data to another revealing the identity of the persons behind the anonymized data.[216] The data used on the intrusion might be publicly available or known by the advisory by other means. In its opinion, WP 29 treats the robustness of each anonymization from three different perspectives:

> "(i) is it still possible to single out an individual
> (ii) is it still possible to link records relating to an individual, and
> (iii) can information be inferred concerning an individual"[217]

From a judicial point of view, the important question is the risk of re-identification.[218] In other words, as put by WP 29, that is the robustness of anonymization.[219] The problem of re-identification is strongly connected with statistical confidentiality, which means that the information is in such a form that it might be used without endangering privacy.[220] In general, there are contradicting views on the risk of re-identification.[221]

There are two competing views on re-identification; the first one is that the anonymization can be easily broken by an advisory[222] the latter view criticizes this view stating that the risk of re-identification is overstated and it is still a useful way of protecting the privacy of data subjects.[223] The first view is supported for example Paul Ohm, and the latter one for

[216]Duncan, Elliot and Salazar 2011, p. 177 and About Anonymisation, UK Anonymization Network, 2014.
[217]WP 29 Opinion 05/2014, p. 3.
[218]Completely risk free regulation cannot be achieved, it should be noted that from a law and economics risk free solutions are rarely necessary.
[219]WP 29, Opinion 05/2014, p. 8.
[220]Duncan, Elliot and Salazar 2011, p. 2.
[221]Most of the discussion regarding anonymization is from the field of medical data and use of medical data, this is because especially in that medical research sector there is a high need for use of data in research, In this thesis the focus is more in the commercial data. The examples of medical data are however used, since not much research exists about anonymization in commercial context.
[222]Advisory is a concept loaned from computer science; it's an imaginary person trying to gain access to data about a certain other person.
[223]Ohm for example calls for the deleting of the word anonymize, since he sees the whole process impossible. Ohm, 2009-2010, p. 1744.

example by Ann Chavoukian.[224] The problem with pseudonymous and anonymous personal data arises from fast evolving technology: even though certain data sets can and are anonymized, in the big data world it might be easily exposed.[225] Also according to Mayer Schönberger and Cukier, big data challenges the technical solution of anonymization of data as well.[226] It could be said that the anonymization techniques are in constant race with the new re-identification technologies. There are multiple cases showing previously anonymous or pseudonymous data converted into personal data via the new models of re-identification.

The problem of re-identification is in connection with aggregation, which means the collection of data in relation to a certain person. Small bits of data collected do not themselves have value, and might not even be considered to be personal data, however, the whole of these parts might be revealing and contain a lot of information. The combined data reveals new private information in comparison to the individual bits of data.[227] The uses of such data can be highly beneficial in a commercial setting, especially when targeting marketing to a certain person for an individually defined price.[228]

Aggregation of data may start with anonymous data, which leads to be de-anonymized in the process; this creates unforeseeable risks for privacy, since the small bits of data are left behind, for example, by using search engines or by visiting web pages.[229]According to the ICO, there are two main ways of achieving re-identification. In the first scenario, the advisory uses already achieved personal data to identify a person from an anonymized data set, and in the second one, the needed data for re-identification is obtained from public records.[230] An important publication on re-identification is Sweeneys k-anonymity, which treats the problems with linking data from multiple sources, and identifying people from such connected data.[231] The focus of the Sweeney's article is in statistic research and protection of privacy with the model of k-anonymity. The major point is that in most of cases, race, sex and zip code are enough to identify a person living in the US meaning that if these are available in a set of anonymized data, the person can be identified.[232] Sweeneys

[224]Cavoukian and El Emam, 2011, p. 1 – 2.
[225]Mayer-Schönberger and Cukier, 2013, p. 155.
[226]Viktor Mayer-Schönberger and Kenneth Cukier, 2013, pp. 151 – 153.
[227]Solove, 2008, p. 118.
[228]Targeted marketing based to profiling may for example use data from the Facebook social graph, e-commerce archives such as previously bought items.
[229]Solove, 2008, p. 119.
[230]ICO, 2012 Anonymisation: managing data protection risk code of practice, p. 19
[231] Re-identification is used as a synonym of de-anonymization.
[232]Sweeney, 2002 and Wu, 2013, 1142.

approach has been challenged by Barth-Jones, who claims that the identification shown is based on the idea of a perfect population registry and also in the idea that the re-identification can be done in such a way that in the end there is only one individual left to whom the data may refer.[233]

There are two famous case examples of anonymization being broken, in both of these cases large data sets of anonymized or pseydonymized data was published in the Internet. The cases *Netflix Prize* and *AOL* show the risks of anonymized data being re-identifiable.[234]Narayanan and Shmatikov point out that datasets that can be linked to publicly available data can be easily de-anonymized.[235] In the case of Netflix, this was done by linking the anonymized data of Netflix database to the data in IMDB-accounts. Netflix published a data set of movie ratings by 500 000 customers containing anonymized movie ratings; the intention was that people could help to develop the Netflix movie recommendation algorithm in a competition. Narayanan and Shamatikov then created an algorithm that demonstrates that with a little of technical knowledge it possible de-identify parts of the data and identify individuals.[236] The algorithm crawls data from IMDB using that data for identify matching movie recommendations, being able to identify individuals easily and with low margin of error. According to Narayanan and Shamatikov the technique can also be used to other data sets of for example transactions data, proving that anonymization is easy to break.[237] The AOL case gives another example of failure of anonymization, AOL decided to publish 20 million search queries for analysis and research purposes. Each individual in the search queries was given an individual number. It was possible to identify individual based to that given pseudonym by combining and analyzing the searches the individual had done.[238]

In my opinion, even though these cases might show weakness in anonymization process, they are not perfect examples of anonymization.[239] In the context of commercial secondary use, this is because in such use anonymized data would unlikely be published for large amounts of people (*publicity of the data)*, and the access to the data would be limited (*limited access*). These additional safeguards might in fact help using data in the

[233]Barth-Jones, 2012, p. 1.
[234] Schneier, 2007.
[235] Narayanan and Shmatikov 2008, p. 1.
[236] Narayanan and Shmatikov 2008, pp 8 - 9.
[237] Narayanan and Shmatikov 2008, p. 12 and 14.
[238]Barbaro and Zeller 2014.
[239] In fact, it could be claimed that the methods used do not fulfil the requirement of anonymization set by WP 29, and that at least in the Netflix price the method was more closely related to pseudonymization.

commercial context, in addition, taking into account that often data used for such activities would not be medical data – but instead mundane activities such as click-streams and information about commercial transactions such as orders. ICO also highlights that different forms of disclosure of data have different risks of re-identification: limiting access allows disclosure of richer data.[240] Additionally, the case of Southern Illinoisan v. Illinois Department of Public Health supports this conclusion.[241] The court heard Dr. Sweeney about the de-anonymization of data and made a decision that the de-anonymization process in that case was based on *special knowledge* of medical area, and could not thus be made by the easy use of Excel, as Dr. Sweeney claimed.[242] I support this approach, since if anonymization were assessed by the standards of a skilled computer scientist or hacker, anonymization would truly be impossible. This would mean that the result of every assessment would be that anonymization might be broken. Jane Yakowitz has also analyzed this case, and made the claim that the easy re-identification is only a myth.[243] Yakowitz also makes an interesting point that the harm of de-anonymization is often overstated. First according to Yakowitz, the possibility of de-identification is overstated, since big portion of the individuals in the data sets are not in risk of being identified, only a small portion of the data set is in the risk of identification.[244] Second not all data acts as a potential identifier, meaning that it is possible to anonymize data.[245] Thirdly, a common man cannot easily de-anonymize data sets; in fact, de-anonymization requires a lot of specialist knowledge.[246] Additionally, in fact, the value of the data the advisory obtains is quite low.[247] The findings of Yakowitz have to be considered when regulating anonymization and considering the benefits and strengths of the technology. The risk of re-identification might actually be much smaller than claimed by Ohm.[248]

Assessing anonymization and the risk of re-identification is an important part of using such techniques. The WP 29 opinion and the ICO guidance, provide tools for such assessment.

---

[240]ICO, Anonymisation: managing data protection risk code of practice, p. 36.
[241]It should be noted that this case is from the relatively different US context – and it is impossible to uses it as a legal source in the European context. The case should be still considered as it has brought anonymization in the courtroom in the judicial analysis of judges.
[242]Even though this case is a US based, it is one of the few court decisions on anonymization. Background, *Southern Illinoisan v. Illinois Department of Public Health.*
[243]Yakowitz, 2011-2012, p. 31.
[244] Yakowitz, 2011-2012, pp. 21 - 22.
[245] I will return to this argument in the next subchapter about the value of data, Yakowitz 2011-2012, p. 23 and 28.
[246] Yakowitz 2011-2012, p. 33.
[247]Yakowitz 2011-2012, 34 - 35.
[248]Ohm, 2009-2010, p. 1705 - 1705.

ICO has created the *motivated intruder test* for analyzing such situations. The motivated intruder test consists of two parts, firstly, it should be assessed whether disclosing the data is likely to result in re-identification, and secondly, whether anyone would be motivated to achieve re-identification. The motivated intruder is a *motivated common man*, who does not possess any unique skills in connection to de-anonymization, such as hacking. The approach is thus similar to the one adopted by Yakowitz and Southern Illinoisan. Sources of motivation are classified to contain, for example, financial gains, gaining newsworthy information or the motivation to embarrass the data subject.[249] Steps that ICO advises to take as a part of assessing the risk of re-identification of individuals, include, among others, seeing if it possible to combine social media data to an individual in the data set, and seeing local resources such as electoral register and library resources in connection with the data.[250] The WP 29 approaches the risk in the form of a more detailed analysis, providing analysis of different anonymization techniques, noting also that since anonymization techniques are under constant research, it might be hard to assess all the risks currently.[251] This also causes the risk of re-identification to vary over time, which should also be noted by the data controller.[252]

In the context of commercial secondary use, I would claim that anonymization of data is a viable solution, and that the risks are lower than in the medical context. This is due to two factors. Firstly, the data collected is not often as sensitive in nature; secondly, it is possible to restrict access to the data, since the purpose is not to do research that would require public access.[253] Cavoukian and El Emam have also maintained a similar position, protecting the value of de-identification, even though it is not a completely certain measure.[254] The motivated intruder test leads to some steps that the data processor must consider. In the commercial context, the restriction of access is the most important solution, since if the anonymized data is accessed only by a low number of people; the risk of re-identification is lowered greatly. For the data controller, this means having an internal governance model or privacy program, and the necessary technical and organizational safeguards for restricting and monitoring access. For example, the concrete steps could consist of limiting the data use for only certain projects, or for only certain

[249]ICO 2012, Anonymisation: managing data protection risk code of practice, p. 22 - 23.
[250]ICO 2012, Anonymisation: managing data protection risk code of practice 2012, p. 23 - 24.
[251]WP, Opinion 05/2014, p. 12.
[252]WP, Opinion 05/2014, p. 24.
[253]Also Ohm notes that the non-public disclosure might in fact be less risky. Ohm, 2009-2010, p. 1729 – 1730.
[254]Cavoukian and El Emam, 2011, p. 4.

individuals, as well as limiting the copying of data and disclosure of the data. In the organizational scale, measures taken could include training of staff and organizational security.[255] WP 29 has also advised taking into account the disclosure and control measures, when analyzing the robustness of anonymization.[256]

I suggest that anonymization rules should base on a risk-based approach. Anonymization should be encouraged to allow the secondary use of data, whenever the risks are low enough, this approach is explored more in the chapter 5 of this thesis. Simply put the risk-based anonymization in marketing context could for instance mean that data could be used for creating anonymized profiles, even though there would be a slight risk of re-identification and loss of privacy – this is due the fact that the commercial benefits would in my opinion weight more than the minuscule risk of loss of privacy. The risk of re-identification should, however, be taken into consideration when regulating anonymization. In the context of research, the risks as well as the benefits of the data use are different. In commercial use of data, the benefits are of economic nature. My view is that the legislators should not completely change the approach to anonymization, the gains of anonymization in different fields are simply too great to be lost for the increased privacy. The next subchapter analyzes devaluation utility, which from the perspective of commerce is an even greater problem than re-identification. The loss of utility is in connection with the usability of data that has been anonymized. Data utility is part of the greater *superquestion* of this thesis; the values balanced against each other are privacy and the commercial interests of the controller.

There are highly varying opinions on re-identification as well as on the value of data after anonymization. There is a definitive need to have new guidance for anonymization, EU-widely, activities such as data-mining can be extremely valuable, and thus the legal framework for the techniques should be in place. Already a lot of activities in commerce depend on the use of anonymized data. A good example of such practices is Telefonica Insights that collects mobile data, anonymizes it and then uses the data for analytics. The data can be used for monitoring crowd movements, this can be used analyzing how interested is the crowd of a store or how does competition affect the movements.[257]

[255]ICO 2012, Anonymisation: managing data protection risk code of practice, pp. 37 - 39
[256]WP 29, Opinion 05/2014, p. 25.
[257] Telefonica 2014.

Another example is SiSense that provides retailers tools for analytics, that combine data from various sources and then visualizes the data for analysis.[258]

3.3 Devaluation of Data Utility

Data utility means the usability of the data; this definition includes the quality of the data and its analytical value. In the commercial secondary use of data, this could be, for example, usability in marketing, profiling or customer research.

> "A summary term describing the value of a given data release as an analytical resource. This comprises the data's analytical completeness and its analytical validity. Disclosure control methods usually have an adverse effect on data utility. Ideally, the goal of any disclosure control regime should be to maximize data utility whilst minimizing disclosure risk. In practice disclosure control decisions are a trade-off between utility and disclosure risk."[259]

The definition from the OECD summarizes some of the conflicting elements in data utility; lowering the risk of re-identification also often lowers the data utility. From an economics of law perspective, the level of anonymization should be set to a point where privacy is protected, but the value of data is not lost. In the level of fundamental rights, the data utility discussion contains a collision; the values against each are utility (protected by *effective markets or the fundamental rights of the controller*) and privacy. From the perspective of the data subjects, it would be perfect if the risk of re-identification would be zero, as the data utility does not concern the data subjects – the controller's interest would be to have as good quality data available as possible.

Although the risks of re-identification could be overstated, it must be analyzed whether the anonymized data still has value for commercial purposes. If the anonymized data is not valuable anonymization might not work as the key to secondary use of data. Another option is that the commercial player will not anonymize data if the value drops too much. The case might also be that in statistical research, anonymized data is valuable, but in commercial sense, for example, for marketing activities anonymization devaluates the data, thus lowering the incentive to use the data or anonymization models.

Businesses that uses data could be roughly divided to data collectors – which are the parties that collect our data, data marketers (or aggregators), data users, for example, advertising companies; and data protectors, which are businesses that make money by

[258] SiSense 2014
[259] OECD Statistics Glossary 2005, Data Utility.

providing protection for privacy.[260] There are several large data broker firms, for example, US based Acxiom, which is one of the largest data-relying marketers, and as stated by New York Times, they have high capacities of data and extensive knowledge about consumers.[261] Other large data broker firms include Experian and Epsilon.[262] In the European context, for example, Telefonica has founded a separate company Telefonica Insights, for selling anonymous and aggregated subscriber data.[263]

From a European point of view, the question is whether data can be anonymized in a way that it is in compliance with the DPD or in the future, the General Data Protection Regulation, and at the same time be valuable. This question is entwined with the previously explored debate about the possibility of reidentification. Ohm views that achieving anonymity destroys the value of data in most cases. The statement is that utility and privacy are concepts in war, and achieving the other requires lowering the other. Modest gains in privacy might result to complete destruction of utility in certain models.[264] Yakowitz has criticized Ohms view. Yakowitz claims that Ohms assumption is erroneous, especially when it comes to the value of anonymized data sets in the context of research.[265] This other perspective is that the devaluation of anonymized data is not as drastic as presented in Ohms work.[266] In my opinion anonymization acts as a valuable too, even though some of the value of data is apparently removed. This is only natural since for the protection of privacy some of the identifiers need to be removed. However in many fields the data can still be used to comprise profiles of consumption habits in certain geographic areas and concentrate marketing to those locations. Additionally the data masses may also be used to improve the processes and, for example, inventory control – if anonymized data sets show that there is a demand for a certain color of cars that may be taken into account already in the production in real time.

Can perfect anonymity and marketing be combined? At least in connection of marketing or individual price discrimination this is not possible, however if the market segment is larger and the individuals cannot be likely be identified – the process of anonymization and

[260]Craig and Lundloff 2011, p. 51.
[261] Acxiom holds data of 190 million individuals and 126 million households in the United States. It has for example also worked with the government after the September 2001 terrorist attacks, providing information about 11 of the 19 hijackers. This shows their identification capacity. Singer 2014
[262]Epsilon 2014 and Experian 2014
[263]Mayer-Schönberger and Cukier 2013, p. 107 and Telefonica 2014.
[264]Ohm, 2009-2010, p. 1752 - 1753.
[265]Yakowitz, 2011-2012, pp. 8-10 and 62 - 64.
[266]Yakowitz, 2011-2012, 30 - 31.

marketing can be combined, especially if using the access controls to further ensure the fundamental rights. Although there is great controversy on the issue of utility and privacy, since re-identification arises great problems to anonymization techniques, my view is that we should regulate anonymization taking into account the risk factors. Anonymization process is currently the best way of achieving usability for data and the cost and benefits for commerce are enormous. Anonymization does affect the value of data, that is certain. However, without the use of anonymization it would not be possible to use data in many cases, as the use would not be legitimate or within the previously purpose for data collection. This is due the role anonymization plays both as a risk-redactor allowing the data processing in some DPD article 7(f) scenarios and additionally because if done properly the data protection laws do not apply to the processing of truly anonymized data. This strengthens the role of anonymization in the era of big data, even though, there are concerns about both devaluation and reidentification.

The excursions to re-identification and value of data give a rather dualist view on anonymization. The following chapter treats enforcement of privacy, since anonymization does not provide us certain solutions for resolving the principal right oriented conflict; a re-emphasis on enforcement and sanctions is needed. This is needed, since use of secondary data is needed and if there are no technical methods for the use and the legal framework is too strict, there is a high probability for non-compliance.[267] Compliance and anonymization may give certain remedies for commerce that wishes to use data for secondary purposes. With the strict purpose restriction and legitimacy requirements on secondary use of data and limited applicability of anonymization, there is a clear incentive for non-compliance. The enforcement practices, especially low sanctions, are behind this. The current situation leads to a double loss of rights – commerce safeguarding fundamental rights loses to the commerce that is non-compliantly using data for economic benefit. Both economic loss and loss of fundamental rights could be prevented by increasing the level of sanctions. In the next chapter, I will explore the European approach to sanctions, as well as show how the US approach to sanctions is much more efficient in multiple areas. Additional emphasis is placed on the regulation proposal by the European commission. *Prima facie* the next chapter about sanctions treats the relationship between DPAs and data controllers, and the data subject fades behind. I would emphasize that the loss of rights by non-compliant data use violates especially the privacy of the data subjects.

[267]Brinhack and Elkin-Koren, 2011 and SPECIAL EUROBAROMETER 359 p. 2. More than 70% of people are concerned by the use of data to other purposes than the presented one.

4 Sanctioning Unauthorized Data Use

4.1 Protection of Privacy Requires Sanctions

The value of data creates an incentive for crime and non-compliant behavior in relation with the data protection laws. According to a major study conducted by Brinhack and Elkin-Koren, big portion of internet activities does not comply with norms in connection with collection of personal data. Especially the collection of data without giving the user sufficient notice or controls is a widespread problem.[268] The lack of credible sanctions is in fact one of the factors that has been noted in the impact assessment that was conducted as a background material for the GDPR.[269] For this reason, it is vital to provide insights into the shortcomings of current enforcement, even though the draft regulation will strengthen and harmonize enforcement.[270] There is the risk that the GDPR will not actually solve the problem in relation with sanctions. Prior venturing to the current levels of fines for unlawful commercial sector data processing in EU and US, it is important to analyze the theories of sanctions and enforcement from the perspective of law and economics.

Enforcement has been described as the act of ensuring compliance with legislation. This definition sees enforcement as modification of behavior.[271] In this chapter, enforcement is analyzed from the sanctions perspective, sanctions are only a part of enforcement, and there are also other aspects In the context of data protection law, enforcement in my opinion means a wide array of possible tools to control the data controller, for instance, auditing by DPA or the US FTC consent order. The claim is that sanctions create the incentive for data controllers to safeguard privacy of the data subjects.[272] This is also well in line with the opinion of Fundamental Rights Agency of EU.[273] Too low sanctions benefit parties that do not consider privacy. Law and economics shows clearly that externalities should be controlled. The importance of sanctions also gain extra weight, if the theories on privacy as a commons or collective are taken into account (i.e. *privacy commons*[274] view or *privacy as a resource*[275]*)* the importance of sanctions is highlighted.

[268]Brinhack and Elkin-Koren 2011, p. 363 and 366
[269] Commission 2010, Impact Assessment GDPR , p. 30.
[270] GDPR, p. 92 – 93.
[271]DREAM framework can be used as an analysis model for enforcement. Baldwin, Cave and Lodge, 2012, p. 227
[272]Equity of sanctions is an equally important part of defining the correct level of sanctions.
[273] Data Protection in the European Union: the role of National Data Protection Authorities 2010 p, 8.
[274]Yakowitz, 2011-2012

The privacy commons view also allows deriving from R. H. Coase, who has addressed the *problem of social cost*, which means, for example, the harm caused by a factory polluting, and the measures that can be taken to protect people from this harm.[276] Coase analyzes the question of whether the party causing damage is liable, and how the responsibility of caused damage is allocated efficiently between the parties.[277] The theory of social cost is well in line with the fact that there are potential economic gains in exploitation of personal data. In addition to Coases theorem, the *theory of rational crime* shows the importance of sanctioning activities that violate rights and may benefit the violator. The theory also highlights the fact that if the combination of risk being caught and the cost of being caught is lower there is an incentive for committing the crime.[278] In the context of data protection, this means that if the cost of a privacy violation is lower than the possibility of sanctions (fines, audit) and public relations damages, the controller has no incentive for safeguarding privacy, which is especially problematic, noting the dual role of the controller as a protector of fundamental rights. Also in the history of regulation, there are good examples about how regulation alone is not enough. For example, environmental pollution and practices such as dumping give good examples of this. Prior environmental regulation, sanctions and enforcement it was commonplace to dump waste to rivers. Similarly also creation of monopolies, cartels and information exchange between competitors was common and it did not stop prior the anti-trust legislation entered into force with high sanctions.[279] In both of these fields the sanctions for non-compliant behavior is much higher than it is for data protection violations or breaches of data protection law.

The next subchapter will focus on the FTC enforcement of privacy in the US. It acts as a benchmark to the enforcement in the Member States, and can thus gives further input for the analysis of sanctions in the EU.[280] Prior to venturing into the FTC enforcement of privacy laws, it is important to address the difference of privacy regulation in the US. The

---

[275]Karl Muth argues that privacy should be seen from a natural resources point of view, building an allusion to preservation of forests and timber industry. Muth 2009 pp. 346 – 347 and 351.
[276]Coase 1960, p. 1.
[277]Coase 1960, p.5 -6.
[278]In addition the theory of rational crime argues for punishment because, if only compensation is paid there is an incentive for violations. Cooter and Ulen 2014, pp. , 458 - 459.
[279]This happened first in the US and later on in the Europe via the actions of EU. The sanctions have been set high since the gains for monopolies and cartels are also high, the consumer loses when such activities are done. See more about the history of competition law Massimo Motta 2004 Competition Policy, p. 2 – 15 and about consumer welfare p. 19 - 20.
[280]Oker-Blom 2009, p. 192. As stated by Oker-Blom the different legal realities are in competition, and in a way they compete for clients, from this perspective it can be highly useful to analyze competing legal realities.

purpose is not to give a detailed analysis of US privacy law, nor to analyze the differences of the law this would not serve the research question of this thesis.

4.2 Federal Trade Commission as a DPA

US approach varies greatly from the European approach, starting from the view to privacy, which in Europe is more understood from the view of personal integrity, and the US approach more about the right to be left alone. This will focus on the enforcement of FTC, giving perspective to the European enforcement. The enforcement in connection with commercial secondary use falls under the regimen of FTC, which is the federal agency for safeguarding the consumer's rights.[281] In the USA, there is no universal privacy legislation; privacy is regulated in sectoral laws, as a part of consumer protection in the consumer protection laws, or in separate sector laws, such as the Fair Credit Reporting Act or The Children's Online Privacy Act. There are certain areas of privacy in the US that are heavily regulated by federal regulation, such as the Health Privacy Insurance Portability and Accountability Act (HIPAA), which regulates the healthcare sector data.[282] It has been claimed that the US has a much more liberal approach, and less legislation, in connection with commercial secondary use of personal data.[283] In a recent article by Solove & Hartzog, the authors claim that this claim is outdated, since the FTC has developed widespread jurisprudence around privacy via enforcing the privacy policies of the companies as their promises.[284] Solove & Hartzog bring out the fact, that even though there are many areas uncovered by the privacy regulation on the federal level, the FTC regulation covers these areas[285] by taking action against companies that do not keep their privacy promises:

> "Although these enormous areas are for the most part unregulated by any industry-specific statute, they are nevertheless regulated. A substantial number of companies today nearly every large company, have privacy policies, and privacy policies are enforced by the FTC. The FTC can bring an action against a company for breaching a promise in its privacy policy –and, even more broadly, for any deceptive or unfair act of practice. This fact has efficiently given the FTC a sprawling jurisdiction to enforce privacy in addition to the statutory jurisdiction."[286]

[281] See Federal Trade Commission, Enforcing Privacy Promises.
[282] For a detailed presentation on the different sector laws that exist in US in connection with Privacy, see Solove and Schwartz, 2009.
[283] This claim has been made by for example Yu, McLaughlin and Levy, 2014, p.11.
[284] Solove and Hartzog 2014, p. 586.
[285] Solove and Hartzog 2014, p. 587.
[286] Solove and Hartzog 2014, p. 588.

This could be concluded by saying that the FTC acts as the *de facto* data protection authority in the US, with widespread powers to enforce privacy. In comparison to the European level of sanctions, the commercial sector US enforcement has shown much higher sanctions.[287] There are, however, controversies around the FTC enforcement; it has, been described as 'toothless and low-tech' by Peter Maass.[288] My view is that even though FTC has its shortcomings and room for critique, the European legislators should take a closer look at the FTC practices, especially since it has shown great capabilities with adapting self-regulating schemes and since the sanction sums are generally higher.[289]

What are the powers of the FTC then, and what makes it so different of our European DPAs. FTC has different powers, these have been categorized into three categories by Solove and Hartzog: the power of investigation, enforcement and litigation.[290] The typical privacy issue starts from an FTC investigation, which usually stems from activities from either complaining consumers or press. When the FTC decides to bring action, it will raise a complaint; the reaction to a complaint can be either settled or disputed in front of an administrative or federal district court judge. The FTC typically settles cases by a consent order, which does not require admission of guilt by the targeted company. These consent orders act as the future basis for enforcement actions. The future sanction for each violation of the consent order is up to $16 000 per violation.[291] In the light of this description, FTC enforcement could be classified to be act-based: the enforcement actions are taken when an act is noticed.[292] In addition, the settlement-nature of the consent orders adds an aspect of soft law, or the so-called compliance approach as described by Baldwin.[293] The consent orders duration can go up to 20 years, and it may contain orders for audits and organizational measures for taking care of privacy. Normally a consent order contains three aspects: financial penalties, bans for activities and recommendations for corrective actions.[294] The variation in monetary sanctions has been high; the fines range from $1 000 to 35 million dollars. I will not analyze the nature of the fines, but instead

[287]It should be noted that the sanctions given by FTC are not *per se* administrative fines, but agreements or settlements that are enforced in courts. The FTC consent order gives obligations to the company that accepts to it and if it is broken, it can be then enforced in the court.
[288]Maass 2014 .
[289]The FTC approach has especially from a historical perspective been strongly in connection with self-regulation and adding extra punch by enforcing the promises given by companies, see more Solove and Hartzog, 2014, p. 594 and 598.
[290]Solove & Hartzog 2014, p. 608.
[291]Solove and Hartzog 2014, pp. 609 - 610.
[292]Baldwin, Cave and Lodge 2012, p. 245.
[293] Baldwin, Cave and Lodge 2012, p. 239.
[294]This is not a complete analysis of the aspects of consent orders; see more Solove and Hartzog 2014, p. 614.

concentrate on the monetary amounts. The following cases shed light to the FTC enforcement, as well as analyzing the sanctions amounts that have been issued. The goal is not to offer a comprehensive list of FTC enforcement cases.[295]

Google Safari (Decision and Later Enforcement), the FTC enforcement of 22,5m$ shows that Federal Trade Commission is clearly in a better position for giving sanctions than the European DPAs.[296] In the FTC case of Google v. US District Court, Google was forced to pay 22,5 million USD in fines for the opt-out practices in connection with the misinformation provided to Safari users about marketing cookies of DoubleClick Advertising and the use of Google Buzz.[297] Google expressly told Safari users that they do not need to opt-out from targeted advertising on Safari. Despite its representations, however, Google conducted such marketing, having the cookie to store the user information for targeted DoubleClick marketing.[298]

As previously stated the strength of FTC enforcement is not merely due to the fines the commission might set. The case of *Vision I Properties* was about the fact that *Vision I* rented and sold personal data collected via their service of e-commerce shopping carts, which they provided to online vendors, without mention of the possibility of such activities.[299] FTC merely ordered a $9 000 fine as a disgorgement and prohibited the activities. In addition, the order included auditing and a prohibition to rent or sell personal data that had been collected prior to the order to change their privacy policy.[300] The FTC strength lies in this flexibility of consent orders, they contain rules for the future operations and the violation then triggers the sanctioning mechanism. The high sanctions FTC may issue are only one of the remedies; a more common approach is a settlement made, which may for example include external auditing of privacy, or changes in policies. In my view, the high penalties act as a "boost" to the data controllers to safeguard privacy.[301] In general, it could be stated that FTC enforcement takes care of the fact that companies need to keep their promises regarding privacy; for example, this includes the recent enforcement in cases where the parties have not been properly safe harbor-certified.

[295]FTC 'Enforcing Privacy Promises' and Solove & Hartzog 2014, pp. 612 - 615.
[296]For the theory of optimal punishment, see e.g. Friedman 2000, pp. 227 - 229.
[297]B. Adequacy of the Civil Penalty, District Order Approving Stipulated Order for Permanent Injunction - Google Safari and Statement of The Commission United States of America v. Google Inc., 2012.
[298]Factual Background, District Order Approving Stipulated Order for Permanent Injunction -Google Safari
[299]Vision I Properties
[300]Solove and Schwartz 2009, p. 423 - 424.
[301]The FTC has enforcement jurisdiction especially in the cases in which the companies do not fulfil their promises given in the privacy policies; this is in more limited in comparison to the European authorities.

However, it should be noted, that in comparison to the European data protection authorities, the jurisdiction of FTC is much vaguer, and the fragmented nature of US data protection laws (or lack of such laws) creates some problematic situations. Raab and Bennet have seen that the lack of private sector regulation has especially problematic privacy.[302] The US enforcement regimen is not perfect either; the sectoral approach to law creates different enforcement across sectors. In some pockets of the law, there is strong enforcement. The FTC takes privacy among other consumer rights, and safeguards the consumer's right to trust the promises of privacy given to them. The clear advantage of the European system is that enforcement actions are brought cross-sector by a clearly appointed administrative enforcer. However, in my opinion, EU DPAs would have much to learn in two aspects, which are the amount of sanctions and the flexibility of consent orders in bringing forth corrective actions. *De lege ferenda* it should be considered whether the flexibility of consent orders could be bought to Europe. In addition, the problem in US seems to be also the fact the FTC does not always catch the wrongdoers, due the limited resources. This creates incentive for non-compliance as the economic theory of rational crime proves.

As a concluding note on the FTC enforcement, it should be stated that even though FTCs role as the privacy enforcer can be critiqued, we could use the practice as a good benchmark when defining a European approach to fines. According to Baldwin a higher level of sanctions can increase the preventive nature of the sanctions.[303] This preventive nature would heighten the importance of data subject's right to privacy and thus provide the necessary safeguards for the e-consumers fading privacy. In the same time, the sanctions would promote healthy law-obliging competition in the electronic markets. The next two subchapters will analyze the sanctions in European Union. The emphasis of the next subchapter is on the problems that the fragmented or disharmonized approach creates.

4.3 European Approach to Sanctions

In the first subchapter of this chapter on sanctions, it was established that both the *theory of rational crime* and the *Coases* theory of social cost explain why non-enforced or non-regulated economic environments are problematic. An unenforced statute, law or regulation creates an incentive for non-compliant behavior, since it is more profitable than

[302]Bennett and Raab 2006, p. 131.
[303]Baldwin, Cave and Lodge 2012, pp. 244 - 245.

complying with the rule. This is true for all of the areas of data protection that contain possibilities for monetary gains. As an example, data transfers are conducted commonly without the proper authorization.[304] It is only logical that secondary use of data would also be done against the statue. First, it is unlikely to be caught and second the consequences for getting caught are low, the most likely penalty would be soft law advice from the DPA or less than million euros in sanctions. From the perspective of data subject's privacy, the lack of enforcement and levels of fines further worsen this across EU.

No uniform approach to sanctions on privacy violations exist within the European single market.[305] The FRA has summarized the problem into three aspects. Firstly, in some jurisdiction, sanctions are limited or non-existing, secondly DPAs across EU generally prefer soft instruments and thirdly there are certain rules and mechanisms that limit seeking compensation via the courts.[306] Additionally many DPAs across Member States are under-resourced.[307]

The article 26 of the DPD states that an individual suffering harm due to unlawful processing of personal data is entitled to receive sufficient compensation. In the DPD article 27, EU sets an obligation to set necessary sanctions and enforcement processes on a national level. The varying implementation of the DPD has caused great variation across the European jurisdictions on the issue of enforcement. In reality in many areas, the reality of European Data Protection does not meet the requirements set in the DPD and the data subject's fundamental rights are not fulfilled.[308] The variation in enforcement means that some Member States have strong DPAs while others have DPAs with a softer approach.[309] Lee Bygrave has claimed that the DPD enforcement is done in the 'spirit of soft law', without much using sanctions and official enforcement as part of the administrative actions.[310] In addition, FRA has criticized the soft approach to the issues.[311]

There are problems with enforcement in multiple Member States. For example, Finland is a good example of a country with a softer approach. The DPA has no power for giving

[304]Kuner 2013 pp, 144 - 145.
[305]Commission 2010, GDPR Impact Assessment, p. 49 and 52.
[306] These include, for example, quantifying of damage and burden of proof in data protection violations cases FRA 2010, Data Protection in the European Union: the role of National Data Protection Authorities, p. 6.
[307]FRA 2010, p. 42 and Kuner 2013, p. 4.
[308]Commission 2010, Impact Assessment GDPR, p. 25 and FRA 2010 pp, 42 - 43.
[309]This is understandable since two distinct styles of enforcement can be seen. These have be called compliance approach which supports more soft methods and deterrence or sanctioning that supports stronger means to achieve the same end, Baldwin, Cave and Lodge, 2012, p. 239.
[310]Bygrave, 2002, p. 79.
[311]FRA 2010, p. 20 and 43.

administrative fines on privacy violations, and the *de facto* potential remedies are within the civil and criminal law.[312] Section 47 of the Personal Data Act sets the possibility of tort liability in cases where the statute is violated.[313] This covers economic and other losses. Section 9 of Chapter 38 of the Finnish Criminal Code sets the possibility of criminal charges for individuals intentionally or grossly negligently performing activities violating the Personal Data Act. In practice, no cases of data controller liability exist in Finland. The tort liability may occur, even if there was no negligence in part of the data controller.[314] Even though in theory a data controller could be liable for data protection practices, to this date no such cases exist.

In UK, France and Spain there are monetary sanctions in some scenarios, although the sanctions are modest in their monetary value. For example, in Spain the LOPD (*Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal*) contains three categories of severity for data protection violations article allowing sanctions ranging from 900€ to 600 000€. Also in the UK sanctions go up to £500 000, in France the sanctions can go up only to 150 000€; which is the amount imposed to for example Google for recent non-compliant behavior in relation with Google Buzz.[315] The monetary value of these sanctions is considerably low when considering the potential value of secondary uses of data. It is quite clear that they do not prevent non-compliant behavior or efficiently punish such activities. As Kuner states additional problems arise with jurisdictional questions, these questions include the questions of establishment and choosing the law that applies.[316]

The problems of current Member State sanctions can be crystallized into three areas. The sanctions monetary effect is either weak due monetary reasons, the soft approach is practiced and no sanctions are issued or the DPA does not have the sufficient resources for detecting and enforcing non-compliant behavior. In some cases it might be so that all of the three causes co-exist, making the enforcement of the national implementation of the DPD virtually impossible. For these reasons, I see that the Commissions proposal in GDPR to increase the fines and harmonize DPAs activities is a good starting point for fixing the enforcement of data protection.

[312]FRA notes that there are problems in relation with enforcement in the following Member States: Finland, Hungary, Lithuania, Denmark, Belgium, Poland, Austria, United Kingdom, and Ireland. FRA 2010, p. 43.
[313] Torts however have the problem of classification of sanctions.
[314]For a further reading on sanctions in Finnish context, see e.g., Vanto, 2011, p. 175.
[315] CNIL, Délibération n. 2013-420, p. 27.
[316]In the current state, there is the risk for forum shopping for the lowest possible enforcement. Kuner, 2007, p. 50.

4.4 Commission Proposal on Sanctions

The proposed new European Regulation addresses many of the problems in connection with the monetary amount of the sanctions. In the new Proposal for European Data Protection Regulation, sanctions are higher and sanctioning would be uniform and EU-wide. This would solve the problems in connection with non-uniform sanctions and affect the incentive to process data without complying with the regulation. In connection with the anonymization of data, it would seem that in this area the proposed GDPR would solve the problems. The problem is however, that too high sanctions in combination with strict norms on purpose restriction may cause problems.

The proposed GDPR would in article 47 establish independent and well-resourced data protection authorities that have the all the required powers. According to article 46 the authorities would also collaborate EU-wide.[317] The intention of the regulation is to create a one-stop shop system that would benefit companies across Europe, since they could take care of all data protection matters with only one DPA in one Member State.[318] Article 79 treats the administrative sanctions, according to the article the sanctions shall:

> "The administrative sanction shall be in each individual case effective, proportionate and dissuasive. The amount of the administrative fine shall be fixed with due regard to the nature, gravity and duration of the breach, the intentional or negligent character of the infringement, the degree of responsibility of the natural or legal person and of previous breaches by this person, the technical and organisational measures and procedures implemented pursuant to Article 23 and the degree of cooperation with the supervisory authority in order to remedy the breach."[319]

Commission has set extra measures that individuals processing data as well as small and medium enterprises would not be affected negatively by the sanctions.[320] The sanctions would clearly increase the powers of DPAs. The initial proposal by European Commission was from 250 000€ to one million or 0,5 – 2% the annual global turnover.[321] The EU Parliament amended this to from 1 – 5% of the annual global turnover.[322] To demonstrate the whole change of paradigm in European fines, it is good to see what would this mean for example, if Google violated the GDPR. A bad violation would cost Google between 600 million and 3 billion USD or in euros approximately 450 million to 2.3 billion

[317] European Commission, 2012General Data Protection Regulation, p. 75 - 76.
[318] General Data Protection Regulation , 32.
[319] GDPR, p. 92.
[320] GDPR, p. 92.
[321] GDPR, pp. 92 - 93.
[322] Commission 2014, MEMO 14/186.

euros.[323] This is a huge increase from the current level of fines, which are in maximum 0.5 million euros in the UK, or the 150 000 € fines Google got from CNIL.[324] The change is a drastic one, especially compared to the current situation of enforcement. However, as previously shown, change to the European enforcement is needed. The sanctions would also be high in comparison with the FTC sanctions, however as stated by Solove and Hartzog the FTC sanctions are relatively small when viewing the global turnover of the violators.[325]

The fact is in fact, there is surprisingly little material or economic research behind setting the levels of fines. No complete economic analysis has been conducted about the potential effects that the sanctions have. With an analogous comparison to sanctions in the area of antitrust law, the Parliament proposal of 5% sanctions seems highly overstated. In for example the antitrust case against Microsoft the commission gave out 560 million euros for failing to uphold an antitrust settlement.[326] In an antitrust case against Intel, the commission ordered sanctions 1.06 billion euros to Intel for abuse of dominant market position.[327] Mirroring the cost of a data protection violation to violations of antitrust law, I would state that the original commission proposal of 1-2% sanctions would be sufficient.

Whether in the final regulation the level of the sanctions will be 1 – 2% of the global turnover or even 5%, it is clear that there are big fundamental right and equality questions when issuing such sanctions. It should be asked how the DPAs would take their new role as enforcers, and whether the administrative enforcement is the correct *forum* for the enforcement. In addition, the potential high amount of the sanctions should be notified when reforming data protection authorities. Enforcing authorities should be properly resourced and they should as well be trained to take the new big responsibility. This creates a great need for principles of good administration and arises the question whether DPAs are the correct enforcers for sanctions.

Calculation of the sanctions under the new regulation might also be un-proportional, for example, in a case in which a subcontractor or affiliate of a global information technology company acts against the data protection regulation; the sanction will be calculated in the basis of the *global annual turnover*. This creates problems, since not always are the major

[323]Google revenue for 2013 was approximately 60 billion US$. See for the revenue of Google http://www.statista.com/statistics/266206/googles-annual-global-revenue/.
[324]CNIL, Délibération n. 2013-420, p. 27
[325] Solove & Hartzog 2014, p. 605.
[326] Spencer Kimball, 2013.
[327] Commission MEMO/09/235.

entities in full control of their affiliates and for example of the marketing activities done locally. Other issues that need to be resolved arise in connection with forum shopping; what if a company sets into a location in which enforcement is relaxed – will a DPA in another EU country set sanctions? The problem is that in some jurisdictions the DPAs have clearly adopted an approach of soft law and even though they would have the resources and powers to give out sanctions, they would not do so.

Although, the analysis is mainly in connection with the fines, I see that other instruments should also be mentioned, since they could also help to protect privacy of the data subjects. For example class action suits, could act as an incentive for complying with data protection laws, even in scenarios where monetary gains are big. In most of the cases the problem is however establishing a monetary damage value for the loss of privacy. The whole doctrine in this area should change so that class actions could be an effective solution.[328] Privacy torts could also be a possible solution, but similarly to class actions, they have the problem of valuation of privacy.[329] In my opinion, these court-based solutions do not give sufficient safeguards to the data subjects, who as noted by Lessig, are often without much representation.[330] If privacy class action suits could however be led by the DPAs or NGOs, this problem would be solved. However the problem of valuating privacy in monetary terms would remain. For this reason, administrative sanctions that are given out in a cross-European way, eliminating the possibility of forum shopping, are an advisable measure that should be taken.[331] Illegal secondary use of data, the sanction under the GDPR would for example be 2%, or as the EU Parliament proposes as high as 5% of global annual turnover. The basis in these cases could be that the processing has been done without sufficient legal basis.[332]

After this analysis of sanctions in both EU and US it becomes apparent that they both have problems balancing privacy with economic interests. In the EU there are big problems with complying with the DPD, the GDPR will increase the sanction but only future will show if it helps. In the US the privacy laws are fragmented across sector. Even though the current

[328]There is, for example, a European class actions against Facebook, Lomas 2014.
[329]Solove and Hartzog, 2014 p. 590 – 592 and Solove and Schwartz 2009, for privacy torts.
[330]The problem in this aspect is that there are not so many monetary interests behind the protection of privacy, Lessig 2006, p. 200.
[331]An interesting question regarding the secondary use of data is in connection with the situations where the secondary use of data is conducted by external parties. This would be especially when sanctions would be due the actions of an external 3-party; conducting analytics for example on basis of an agreement. This theme will not be further analyzed however, since it falls outside of the scope of this thesis.
[332] GDPR, p. 93.

FTC monetary sanctions trump their European equivalents, even these sanctions and enforcement actions have problems.

As stated, there are some pitfalls when strengthening the enforcement of data protection. First the DPAs do not most likely have the resources for enforcing all no-compliant behavior, high sanctions in a selective manner is unlikely to solve the problem of non-compliant secondary use of data. This creates a problem with good administration, since all should be equal in front of the officials. Second, the sanctions could affect small-business companies that try to enter the market, a big repeat-player company, could most likely survive from the sudden change of regulatory environment. Thirdly, the enforcement costs will most likely be pulverized down to consumers. For these reasons, I see that the proposed GDPR and the sanctions increase the need for risk-based regulation.

The next chapter will analyze the regulation of privacy from a risk-based perspective, which would in collaboration with sanctions effectively solve the fundamental rights conflict. The first subchapter will analyze regulation of privacy more generally, and provide insights to privacy regulation from three different perspectives; from the economic perspective of privacy, from the perspective of privacy as a common right and from the idea of law as code. These ideas will be explored briefly to frame the effectivity of the risk-based approach.

5 Risk-based Regulation of Data Use

5.1 Economic Theory of Privacy

Regulation of cyberspace is problematic, since legislation is often national; the World Wide Web has created problems, challenges, risks and opportunities that are international. It is important to notice certain economic factors behind privacy when analyzing the regulation of data protection. Without understanding the economic aspects of privacy, it is impossible to navigate towards satisfactory privacy regulation, which could answer the different needs of different interest groups, and balance the fundamental rights in conflict. I will next analyze how economists have viewed privacy; the analysis is used for bringing economic input to the discussion of regulating privacy. The economic view of privacy differs highly from the view of privacy as a fundamental right.[333]

[333] The right to privacy as a fundamental right was explored in the subchapter 2.1.

According to the analysis of Acquisti, the most traditional neoclassic economics literature on privacy states that consumers are rationally informed agents with stable privacy preferences.[334] The traditional models also state that privacy is not itself a value, but an economic benefit that it confers.[335] In the economic model of the Chicago school, privacy can be seen as a mere cause for inefficiency, since people likely disclose only information that is beneficial to them. Especially Posner views privacy in highly critical light: his perspective is that privacy should only have value when it is essential for the protection of trade secrets. In addition Posner also views privacy as a harmful right – since it allows consumers to hide facts that might affect, for example, the decisions of insurers.[336] Calzolari and Pavan have studied the effects of information disclosure about consumers, and concluded that the consumer does not necessarily suffer from such practices and that in the end the disclosure of data between parties may add to the total welfare of all parties involved.[337] Varian has claimed that if privacy is regulated without taking the economic aspects into account, this may lead to overly strict solutions that lead to the loss of welfare. This is well in line with the fact that Internet services are currently financed mainly by data-driven advertising.[338] The risk-based approach to privacy would better include these economic factors to the regulatory process. It is important to understand the role personal data plays in the online economy. Especially if the sanctions proposed in the GDPR enter into force, the effects in the online economy and welfare of the e-consumers might be surprising. The increased sanctions and their costs might be driven down to the consumers, which could cause monetary losses to the consumers.

Privacy as a collective right or privacy commons is another idea that should be considered *de lege ferenda*. Karl Muth states that privacy can be seen as a collective right, which can be understood similarly as a natural resource. This view of Karl Muths is especially useful when using enforcement examples from, for example, environmental regulation.[339] Risk-based regulation is already used widely in different jurisdictions and different fields of regulation, especially in environmental law and financial markets regulation. The risk-based regulation approach is well in line with the social theory that we the current modern

[334]Acquisti, 2010, p. 4
[335]Acquisti, Leslie and Loewenstein 2009, p.5.
[336]Posner, 1978, p. 22 and Posner 1998, p. 44 - 45.
[337] I use the term welfare here to mean the economic benefit in total to all parties involved in the online economy, the consumer or buyer (or data subject) and the different parties selling. Calzolari and Pavan 2006, p. 2 - 4, 22.
[338] Deighton and Quelch, 2009, pp. 23 – 25 and Varian, 1996
[339]Muth 2009, pp. 346 - 348 and Bennett and Raab 2006, p. 11.

society could be classified as a risk society.[340] As Muth also states, there is a clear analogy to be made between environmental law and privacy legislation.[341] Thus, I see it as only logical to use a similar risk-based approach on the regulation of privacy. Several problematic issues, however, arise from a risk-based approach to legislation.[342]

The third theoretical part behind the *de lege ferenda* analysis is the Lawrence Lessig's idea of regulating cyberspace. Lessig's idea is that the regulation of the Internet needs to take into account the code, and that the regulation of cyberspace can be built into the code.[343] This idea should especially be taken into account when considering how to regulate anonymization. In anonymization, the question is precisely about that; code is made so that the personal data saved is in a form where privacy is not at harm. The view also concerns the secondary use of data, which is of technical nature. In extreme, the privacy decisions could be included with the data used as metadata, all data could technically contain a label that informs about the risk-classification and the legality of the data use.[344] This data labeling would be a technical solution that would facilitate the use of data, while protecting privacy. The discussion of regulating code is well in line with the statement of Baldwin; it is hard to distinguish technical and social risks, since in the end, people have created the technical risks in cyberspace.[345]

The next subchapter will explore risk-based regulation. The difference between the concepts of *risk-based approach and risk-based regulation* need to be clarified. The first means purely taking the risk as the primary point of regulation and regulating to control a certain risk, for example, against harm caused to the data subjects. Risk-based regulation, on the other hand, is a more complex concept, that also includes other aspects such as monitoring the effectiveness of regulation by seeing what are the identified risks and how does the applied regulation then affect them.[346] The goal is to show why the regulator should take ideas from the risk-based regulation and then apply the risk-based approach to

[340]Beck 1992, p. 1.
[341]Muth 2009, pp. 350 - 352.
[342]Baldwin, Cave and Lodge, 2012, pp. 91 - 93.
[343]Lessig discusses the possibility of regulating internet in a way that anyone there could be identified, and thus increasing the ability to regulate cyberspace. Lessig, cop. 2006, p. 62, 66 - 69. See also Rule, 2007, p. 19.
[344]Lessig brings out the possibility that technical solutions could help treating privacy in similar way to a property right. A technical solution could allow users consenting for the data uses. This idea has a lot of beneficial aspects, but it might be from economic and politic perspective to re-engineer cyberspace in such way. Lessig 2006, pp. 227 - 230.
[345]Baldwin, Cave and Lodge, 2012, p. 83.
[346]This is how I use these concepts in this chapter, for the different aspects of risk-based regulation see Black, 2012a, p. 1, chapter 14.1.

secondary use of data. The 5.2 subchapter explores the risk-based regulation from the perspective of the regulator. The 5.3 then applies the lessons learned and creates a data classification model that would push risk-management duties to the data controllers and in return liberate certain parts of secondary data use. My analysis is done with taking the risk as the starting point for regulatory actions, however, the risk-based approach to regulation has immense benefits and most of these would be useful in the field of privacy.[347]

5.2 Risk-based Regulation

The risk-based regulation of data protection, means seeing the regulator in the role of a risk-manager or allocator of risks. This includes deciding which risks the data controllers may take and how do the safeguards applied affect the potential use of data. Risk-based regulation is an approach of applying risk-management processes both to the regulation as activity, and using risk as the key concept for understanding the regulated subject.[348] According to Pounds, the risk-based approach is characterized by understanding that not every risk can be regulated, and the legislator should acknowledge and analyze different risks.[349] The legislator or the state can be seen as the ultimate risk manager. Power has claimed that the state, as the regulator, has not been aware of this duty for long.[350] Control of risks can be seen as the object of regulation.[351] Societal risk is a scenario where a party has the possibility of losing their right; as such, the loss of privacy is a risk for data subjects.[352] The essence of risk-based regulation is the prioritization of regulatory actions in accordance with the risks associated; the goal for risk-based frameworks is to control relevant risks, instead of looking for compliance with a set of rules.[353]

The risk-based regulation affects all of the steps in regulation, starting from the planning of new regulation and ending in the enforcement actions that can be accordingly targeted in the highest areas of risk. According to Black and Baldwin, there are five common aspects in the risk-based frameworks. Firstly is the control of risk. Secondly, the regulators scale

[347]It is important to make an important distinction in this point, in the corporate privacy management there are multiple risk-assessments processes for controlling privacy and security risks – the intention of this chapter is not to analyze those, they are left outside the scope of this thesis. These processes are internal measures of the data controllers and not in the field of regulatory analysis.
[348]Baldwin, Cave and Lodge, p. 281, Black, 2010b, p. 187 and Bounds, 2010, p. 32.
[349] Gregory Bounds, 2010, pp. 16 - 17.
[350]Power, 2004, p. 17.
[351]Baldwin, Cave and Lodge, 2012, p. 83.
[352]See 1.3 for the role of risk in the methodology of this thesis.
[353] Baldwin, Cave and Lodge, 2012, p. 281.

for accepting risks. Thirdly, the frameworks involve assessing the hazards or adverse effects. Two broad categories of risks are classified; inherent risks and management and control risks. Fourthly, the risks are scored so they can be assessed. Lastly, the risk-based framework provides means for linking the risk assessment with enforcement – since resource allocation is an important part of risk frameworks. It is however important to notice that according to Baldwin and Black, no risk-based system is identical and that the creation of such frameworks is not merely a technical process. Meaning that the creation of such frameworks includes decisions that later affect the whole system created by the use of the frameworks.[354]

A good example of Risk-Based Frameworks for regulation is the IRGC framework. It is intended to be used as a general tool for multiple areas of regulation that include technology or natural risks. The framework contains five areas; 1) Risk Pre-Assessment 2) Risk Appraisal 3) Characterization and evaluation 4) Risk Management and 5) Risk Communication.[355] Practical examples of applying risk-based regulation can also be found from the areas of the financial sector and environmental regulation. In the UK risk-based regulation is widely used especially after the Hampton Review recommended such measures to be taken in the areas of enforcement.[356] After the early adaptation of risk-based regulation, its use has multiplied, especially in the financial markets regulation and environmental regulation.[357] Fisher has noted that in these areas the regulator needs knowledge from multiple areas to first assess the risks and later to regulate and enforce the subject.[358] This would also apply to the regulation of data use and data protection, expertise is required to various fields.

Benefits of the risk-based approach for the regulator are significant. First, the regulator has the possibility for savings since the regulatory activities are better aimed and second, the regulation may have less effect on the efficiency of markets. The third argument is that risk-based regulation adds accountability to the regulatory activities since, as the goals are clearly defined, it is easy to notice whether the regulatory systems work.[359] These however are big promises, since in essence the risk-based approach to regulation promises of

[354]Black and Baldwin, 2010, pp. 184 - 185.
[355] IRGC 2008, Risk Governance Framework, p. 8 - 14.
[356] Hampton 2005, p. 9 - 10.
[357] For example Canadian Banking Regulator, the Office of Superintendent of Financial Institutions (OSFI), Portugal's environmental regulator IFAOT has adopted the risk-based models developed in the UK. Black and Baldwin, 2010, pp. 183 - 184.
[358]Elizabeth Dr Fisher, 2010, p. 51.
[359]Black 2010b, pp. 188 - 189 and Bounds, 2010, p. 32.

converting regulation, which has been often seen as an art-like operation to something that can be monitored and assessed.[360] It is only natural that the risk-based approach to regulation has also been criticized, and even described as simply dressing old regulatory processes in new clothes.[361] The critique affects multiple areas of the approach. According to Black & Baldwin the creation of risk-based frameworks require a lot of information about the regulated subjects.[362]

Several factors should be considered when creating a risk-based regulatory system. First of all the level of risk tolerance needs to be considered – this varies often sector by sector. For example, in the food safety regulation there is a zero tolerance for risks, in the financial sectors regulation the appetite of risks is much higher.[363] In my opinion the risk tolerance for data protection regulation falls somewhere in middle, most likely nearer the food regulation, however in the area of data protection there are several different areas in which the risk tolerance and goals are different. The second aspect that needs to be considered are the risks and how they can be identified. The relevant risks need to be noted and assessed; the regulator has to concentrate on concrete risks that are identifiable.[364] In the area of data protection the loss of data subjects privacy acts as the core risk, however, the quantification of the risk can be done by analyzing the quality of data and then by the amount of data. If the loss of data would lead to discrimination of the individual or to losses of democracy, the risk is considerable.[365] Additionally if the loss of data can lead to psychic or physical harm, the loss of the data creates considerable risks. After the risks have been identified and the risk tolerance has been chosen, the regulator needs to set risk indicators and assess the probability of the risks.[366] This helps deciding which of the risks can be transferred to the data controllers via legislation.

The risk-based approach to data protection would better solve the fundamental rights conflict caused by the dual nature of data protection laws. Data protection laws have the aim of solving the fundamental rights conflict. The economic side of the issue has only emerged recently, and for that reason a new risk-based approach would benefit the

[360]Part of the critique states that the financial crisis proves that risk-assessment and management practices are unreliable and that they create the false sense of security; Black and Baldwin, 2010, pp. 203 - 204.
[361]Black, 2010b, 188.
[362]Baldwin and Black offer a really responsive risk-based system, which responds to problems by using a variation of the risk-based regulation, Black and Baldwin 2010, pp. 198 - 199.
[363]Black 2010b, 193
[364]Black 2010b, 194
[365]Ojanen, 2010, p. 43 - 44.
[366]Black 2010b, p. 195 -196

legislator. As the data controller carries the profits of data processing, the controller should also carry the risk. The risk-based approach, in connection with the strong enforcement proposed by the GDPR, would create a system that is has fewer administrative burdens and allow data subjects' fundamental right to privacy to be fulfilled. It should be noted that in an ideal scenario, regulation is something that allows and encourages activities such as commercial use of data, instead of only putting limits on behavior. This works both ways, for example by potentially empowering individuals for privacy activism, which can be done by granting access rights.[367] The Center for Information Policy Leadership (CIPL), has already created initial proposals for adopting the risk-based approach to privacy, however the initial draft frameworks do not completely fulfil the idea of risk-based regulation.[368] The CIPL frameworks aim to include the control of privacy risks to privacy regulation. They do not, for example, include other aspects of a risk-based regulation, such as the increased accountability of regulation.

The goal of the next subchapter is to view how a risk-based regulatory approach could be applied to design data protection regulation of data use *de lege ferenda.* Such a model is needed since the regulation of cyberspace has to take into account the conflicting fundamental rights and the realities of the online economy.

5.3 Risk-based Model for Data Use

I previously have explored the problems current regulation has in relation with the purpose limitation, anonymization and its possible failure and the failures in the field of data protection enforcement.[369] The use of data could be regulated by using a risk-based approach and by classifying data in several risk-categories. This categorization could be used for setting the obligations and guidance's for the data controllers. The model would partly the fundamental rights conflict and fit the ideas of law as code and the regulation of privacy commons together. The fact that privacy regulation has economic effects has to be taken into account. In the model that I explore and propose, the goal is that data could be used as freely as possible and that at the same time the individual's privacy would be

[367]Bennett and Raab 2006, p. 121 - 122.
[368] CIPL, 2014a A Risk-based Approach to Privacy? An Initial Issues Paper for Privacy Risk Framework and CIPL, 2014b, Risk-based Approach to Privacy Project and A Risk-based Approach to Privacy: Improving Effectiveness in Practice.
[369]Subchapters 2.4, 2.5 and 3.1.

safeguarded – this would both take care of the fundamental rights concerns and also ensure the maximum economic welfare.

CIPL has already drafted a preliminary risk-based approach. Their risk-based approach aims to reduce the harm that might be caused to data subjects, instead of only protecting privacy as a fundamental right. CIPL classifies harm as all potential damages, injuries or negative impacts to the data subjects; this classification includes for example monetary or reputational damages.[370] CIPL further classifies the harms to three categories; tangible damages, intangible distress and societal harm.[371] The CIPL framework is intended for controlling the significant privacy risks and it is meant to be used alongside the current data protection laws and legal instruments.[372] WP 29 has commented on the recent discussion on the risk-based approach, raising the point that the current regulation should not be abandoned, and that the risk-based elements are already involved in the regulatory process.[373] WP 29 has also commented that in all scenarios, fundamental rights of data subjects need to be safeguarded even though the risks in question would be small.[374] In my opinion, the WP 29 does not fully take into the account the economic needs for data use, also ignoring the fact that the current system does not protect the fundamental rights due failures in enforcement. After the GDPR fixes the problems in the enforcement system, the *de facto* balancing of privacy and economic interests needs new tools, and the risk-based approach could be one of these tools.[375] I would not however advocate for abandoning the current systems completely, the risk-based approach could be tried out in several areas and the testing could lead to expanding to other areas as well. The data controllers have their fundamental rights and there are economic benefits in the use of data for both the economy in general and consumers alike.[376] The risk-based approach works well in certain areas of data protection, and the secondary use of data is one of those areas. Correct levels of sanctions in connection with a risk-based framework will also further ensure compliance with the rules in connection with the data.

[370] CIPL 2014b, pp. 1 -2 and CIPLa, pp. 1 - 3.
[371] Tangible is physical or economic harm, intangible distress is, for example, reputational harm or discrimination, societal harm is damage to the democratic institutions or social trust, see more, CIPL 2014b, 7.
[372] CIPL 2014b, p. 2.
[373] WP 29 2014, Statement on the role of a risk-based approach, p. 1.
[374] WP 29 2014, Statement, p. 2.
[375] See chapter 4.4 for the future sanctions that will have big economic impacts to data processing.
[376] As discussed previously in subchapter 2.1.

The next goal of this subchapter is to create a model for classifying the risks and regulating data use according to this model. The previously explored CIPL material gives good input for understanding the harm in question.[377] In addition, the risk-based regulation provides input to the data use mode. According to Bounds, the first phase of risk assessment is framing and forecasting the probabilities of risk. In this context, there are two risks: the first, being the loss of data subjects and the second being the loss of commercial benefits to the data controllers and consumers alike. The second phase of the assessment is then managing the risks by taking action. The control of risk can be classified in four categories: risk avoidance, reduction, retention and transfer of risks. Risk avoidance, in practice, means prohibiting certain activities, retention means accepting the risk, transfer means the transfer of the risk to private parties and reduction is the strategic measures to reduce risk.[378]Also as CIPL has noted, it is important to take into account the fact that the privacy risks of different phases of data processing are different.[379] At this point, it is important to note that enforcement and sanctions have an important function ensuring that the risks taken are included in the framework.

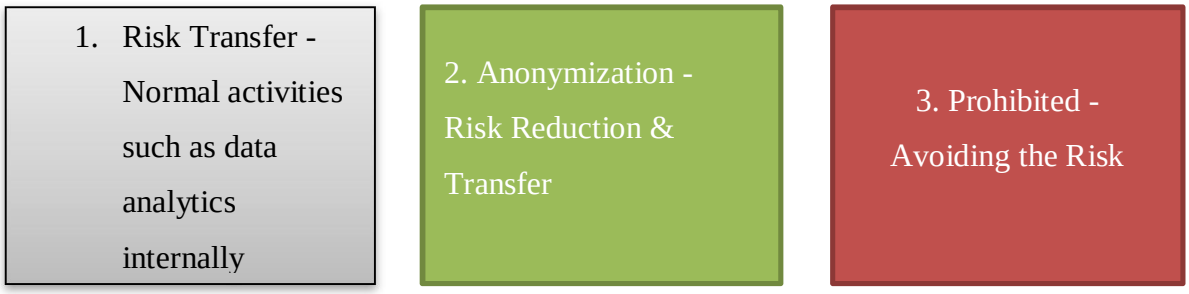
As the second phase of the assessment process proposed by Bounds. The legislator should classify categories of personal data. Certain types of data do not contain potential harm for data subjects, in this first category; the commercial use of data would be free and acceptable. Data that would fall into this category would be, for example, ordinary data of the buyer such as age and address. In the second category, the data use would be limited and anonymization of the data would be required as a prerequisite for the data use. In addition, the regulator would then give guidance in the possible techniques of anonymization. The measures of control could in this category include: limiting disclosure of data to only parties that are deemed as safe, access controls and certain internal procedures.[380] In the third category, the data use would be prohibited due to the high risks. These high risks would apply to data that is so highly confidential to the individual or of

[377]Interestingly the CIPL also notes that the different privacy risks are not classified well. CIPL, 2014b, p. 4 – 5. This in my opinion calls for additional research in the area of different risks, since if regulation is done without knowing the risks the resources are focused in wrong areas, Hampton 2006, p. 43.
[378] Bounds 2010, p. 19.
[379] CIPL 2014b, p. 5.
[380]See the safeguards proposed by the working party for the legitimate interests balancing test, WP 29, Opinion 06/2014, p. 42. In addition, CNIL 2014a, p. 12 – 18 on assessing the possible risks to the collected data in the role of data controller and on the measures the controller should take CNIL 2014b, p. 4 - 24.

sensitive nature. The classification could be made EU-wide in the regulation, which would then also benefit the free flow of data.[381]

The following picture demonstrates the categories of data in the model:

Picture 4.1



The risk classification in this model would follow and fulfil the fundamental rights protected by data protection regimens. For example, in the first category, the loss of rights is considered low, and in fact allowing the data use does not cause risk of rights loss to the data subjects. In the second category the additional safeguards guarantee the rights of the data subjects, while simultaneously allowing some economic activities.[382] In the third category there would be high risk of loss of privacy and for this reason the commercial data use is prohibited. This category would include the commercial uses of sensitive data, which would have high risks for discrimination of the individuals. The model would especially clarify the situation with anonymization and ordinary data, making it possible to use more data while simultaneously mitigating the risks in connection with the data use. Enforcement would be directed so that the sanctions would be highest for the third category data use; also, if safeguards would not be used in the second category, the DPA could fine the data controller.

The model interacts in two ways with anonymization. First of all, the anonymization process can be used as part of the framework to reduce risk, thus allowing certain data uses, which, considering the data subjects fundamental rights, would otherwise be classified as too risky. Additionally risk-based approach can be used to assess the anonymization practices, to see whether they are effective or not. This would serve the

[381] There are of course types of data that are considered more confidential in some EU countries than in others. For example, membership in labor unions is such information. The model would require harmonization of data protection across EU.
[382] This is also, in line with what has been discussed about the devaluation of data in subchapter 3.3.

better regulation of anonymization practices. Firstly, anonymization can be used as a solution to reduce risk; secondly, anonymization practices can be recommended by regulators to be used in certain situations and thirdly, it a may be required that a certain anonymization method reduces a certain amount of risk. Although risk-based approach to anonymization standards would probably be the best possible option, there are several problems with this approach. It may even be impossible to assess the risk of re-identification correctly.[383] ICO already stresses the importance of risk analysis as a part of the anonymization process.[384] The risk assessment is done by conducting the motivated intruder test, which is used as mean to assess whether anonymization is conducted properly.[385] *De lege ferenda,* the correct way of regulating anonymization should be as technology-neutral as possible. The correct level of anonymization is set in a point where the value of data can be kept as high as possible and re-identification of risk as low as possible. When setting the minimum for anonymization *de lege ferenda*, the legislator should consider the amount that the risks decrease by as well as the increased cost of implementing the technology and the lowered value of data. From the perspective of the risk-based framework, the best regulatory solution would set the guidance for selecting anonymization techniques and assessing the risks, the data controllers could then choose the techniques that they may want to apply. If the GDPR enters into force, this task would be best set to an advisory body consisting of WP 29 and the required technical expertise.[386]

The risk-based approach could be also designed to interact with the purpose restriction. The problem with purpose restriction is that it is both too vague and strict. It does not provide real safeguards for the data subjects.[387] *De lege ferenda*, risk could be included as part of the purpose restriction or a as a new data protection principle; "Risk based processing". The risk based processing principle would add that data may be processed for other than areas that are within the purpose, if they are low risk or considered as such after the successful anonymization process. As a hot-fix I would propose the following modifications to the language of the purpose restriction:

---

[383] It is general hard to quantify risk and there is always the possibility that not all risks are taken to account.
[384]ICO 2012, Anonymisation: managing data protection risk, p. 18. In general, it should be noted that assessing risks is nearly impossible, since some situations are so called Black Swans that might be impossible to predict. About "black swans" and highly improbable situations, Taleb, 2008.
[385]The motivated intruder test is simple; the question should be asked whether a motivated intruder could achieve re-identification, if they are motivated enough. ICO 2012, Anonymisation: managing data protection risk code of practice, pp. 22 - 24.
[386] Much like in the UK, where UKAN assists with anonymization relating questions UK Anonymization Network 2014, About Anonymisation: for data about people.
[387] The chapter 2.4 has provided the problems to in connection with purpose restriction.

GDPR article 5 (bold section added by the author):

> "(b) collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes, **further processing of low-risk and medium-risk data categories may be accepted if necessary safeguards are employed**;"

The fix would implement the concept of risk to the GDPR. The risk assessment could be modelled after the balancing guidelines that WP 29 has already given in relation with the article 7(f).[388] In fact, the WP 29 has proposed similar measures for amending the purpose restriction, also concentrating on the safe guards of the data. The intention of WP 29 is similar; the approach would allow the use of data and more flexibility with the purpose limitation.[389] In my opinion, the risk-based approach would fulfill the same goal. In this context, it should be noted that the DPD already contains elements that could easily be modified to be even more responsive for the potential risks and the control of risks. The article 7(f) balancing test introduced by the WP 29 is good example of this.[390]

The risk-based approach to data use cannot really survive without the functioning enforcement coupled with strict sanctions. This is because there should always be regulation in place to limit taking risks that are not acceptable. Without the reform of data protection enforcement, the risk-based approach is doomed to fail in safeguarding the data subject's privacy. Without proper enforcement it will act only as a justification for going against the fundamental rights of data subjects. The risk based-approach to data protection has problems, and if such systems would be put on practical use as a regulation even more problems would probably be found. As also Baldwin states, it is often hard to predict and measure the efficiency of regulatory solutions; regulatory assessment is easier after the effects have been seen in practice.[391] I would however see it important that the current system of regulating the secondary use of data would be reconsidered. In my opinion, this should be done at the same time as changing the sanction mechanisms and data protection enforcement. Since as proven, the current purpose restriction and legitimate interests ground, even in combination with anonymization, does not solve the conflict of fundamental rights in this area and allow data operations.[392]

---

[388] WP 29, 06/2014, p. 55.
[389] WP 29, Opinion 03/2013, p. 42.
[390] WP 29, Opinion 06/2014, pp. 33 and 55 and CIPL 2014c, p. 7.
[391] For measuring regulatory quality see, Baldwin, Cave and Lodge 2012, pp. 34 - 38.
[392] See chapters 2.4, 2.5, 3.2 and 3.3.

6 Conclusions

Data protection legislation is at a crossroads. The choices are either evolving or staying put and becoming a dead letter of law. The current regulation does not satisfy the economic needs of the era of big data. Neither does it sufficiently safeguard the privacy of consumers, since there are major shortcomings in the area of enforcement and sanctions. The Regulators ignore both the fundamental rights conflict and the economic realities in connection with the online economy and the need for data use. Currently the commercial secondary use of data is done majorly via anonymization, which allows escape from burdensome restrictions. The increasing value of data further heightens these problems and conflicts of interest. There is a high risk that without proper enforcement consumers suffer losses of privacy. Additionally the risk of re-identification looms above the use of anonymization techniques. These times of fast change call for new perspectives in the regulation of privacy as analyzed in chapter 1 and subchapter 2.1. The fundamental rights conflict, which is worsened due to fast changes in technology and online economy, is not currently solved by the DPD.

I have treated the commercial secondary use of data primarily from the perspective of purpose restriction and the legitimacy of data use. The thesis has analyzed the DPD article 6 and DPD article 7(a) and 7(f). In addition, the Member State implementations have been viewed. Anonymization has also been explored as a potential safeguard and a solution to the value of data. The enforcement of data protection has gained much attention since the economic analysis of law shows that high rewards of non-compliant behavior increase non-compliance. Additionally, enforcement has an important role as part of the solution for data use. In Europe, the sanctions are weak and they are only randomly imposed. Even though the proposed GDPR repairs many of the areas with sanctions, it leaves the rules for data use nearly unchanged, or even further restricting the use of data. This is especially problematic and calls for new approaches with privacy regulation. An ideal privacy regulation takes into account the different needs and the economic effects of such regulation to the welfare of the society. The regulation should safeguard individuals' privacy – without imposing too excessive restrictions to the ability to use data. This thesis has explored how the risk-based approach would provide answers in the various problematic areas of the regulation. The risk-based approach would *prima facie* solve a large portion of the problems in the area of commercial secondary use of data.

In the second chapter of this thesis, I treated the problems for potential secondary use of data caused by the current DPD. A main problem in relation with big data and the commercial secondary use of data is the inflexibility of purpose limitation. The principle limits the use of data in many cases where there is no potential risk for harm to the data subjects. This is especially problematic, since a lot of data falls under the scope of personal data and thus it is governed by the DPD. This also causes problems with DPD article 7, especially in relation with consent and legitimate interest for data use. Even though WP 29 has given out guidance for utilizing the balancing test of article 7(f) and adding some flexibility for repurposing of data, this is not alone enough to solve the problems. In a scenario of high sanctions and effective enforcement, this inflexibility of the regulation would be especially problematic. Consent is an especially problematic basis for secondary use of data.

I have analyzed anonymization deriving from the DPA guidance's and WP 29 guidance in chapter 3. Anonymization has been, and will in future be, a good solution for facilitating secondary use of data. It also plays an important role in the risk-based approach. However, there are multiple questions in relation with anonymization. First, the risk of re-identification challenges the usability of many techniques, and second, there is also the risk that robust anonymization techniques decrease the value too much. The risk of reidentification is hard to control, since the technology around anonymization changes frequently. Even though there is a risk of reidentification and the value of data is devaluated after anonymization, I see that the role of anonymization will not diminish. Many uses of personal data would be impossible without the use of anonymization and it is possible to add additional safety to ordinary use and disclosure of personal data by carefully selecting anonymization techniques. Organizational safeguards, such as access control and limiting the disclosure of the data may be the solution for anonymization in the context of secondary use of data. Current and future sanctions have been analyzed in chapter 4 of this thesis. The sanctions were found to be insufficient, considering the high value of data. The FTC enforcement was explored, and especially the use of consent order could be explored in the European concept. The monetary effectivity of sanctions should be improved, the future GDPR harmonization will improve the situation in regards with enforcement and sanctions.

The risk-based regulation has been explored in chapter 5 of this thesis. Risk-based approach to privacy regulation would *de lege ferenda* support the regulator in creating

regulatory solutions that balance the fundamental rights conflict. A risk-based approach to privacy would focus the regulatory actions to the correct points, where the privacy of the data subjects can be best safeguarded, without losing the potential and value of data. The risk-based model for data use would allow certain uses of data, allowing the restrictions caused by the purpose limitation principle to be ignored. In turn, the model would also take the power of anonymization into consideration and use it to allow data activities that only have mitigable risks. Some commercial activities with data would be prohibited in cases where the risks to data subjects are too high. The risk-based approach would act in combination with the proposed sanctions in the GDPR. This would balance the conflict of fundamental rights; also taking into account the societal and welfare benefits data use apply across the economy and societies.

The risk-based approach to commercial secondary use could act as a new regulatory tool for solving the fundamental rights conflict of the big data era. Such new approaches are needed so that the privacy of individuals remains as a functioning fundamental right. Future research would be needed for assessing the correct areas where the framework could be applied. The risk-based approach might not be the solution for all data protection, nor should it be advocated as the solution for everything. Future studies would also be needed in the analysis of data protection sanctions. It would be especially beneficial to analyze the monetary effects even more.

I see that data protection regulation is more important than ever as the consumers have the right to privacy. At the same time, an effective data protection regulation safeguards also the general welfare and acts as a basis for economic activities. New legal innovations should be put to use for the protection of privacy during this era of fast technological change – privacy is not dead, it is evolving to the next level.