BLOCKCHAINS AND THE EUROPEAN DATA PROTECTION AND PRIVACY LAW

LL.M THESIS

UNIVERSITY OF HELSINKI
LAW FACULTY
2019

CAGLA SALMENSUU
014473885

Supervisors

Dr. Tobias Bräutigam

Prof. Ville Pönkä

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta/Osasto Fakultet/Sektion – Faculty LAW FACULTY	Laitos/Institution– Department International Business Law	
Tekijä/Författare – Author CAGLA SALMENSUU		
Työn nimi / Arbetets titel – Title BLOCKCHAINS AND THE EUROPEAN DATA PROTECTION AND PRIVACY LAW		
Oppiaine /Läroämne – Subject GDPR AND REGULATION OF THE BLOCKCHAINS		
Työn laji/Arbetets art – Level MASTER'S THESIS	Aika/Datum – Month and year 14.05.2019	Sivumäärä/ Sidoantal – Number of pages 89

Tiivistelmä/Referat – Abstract

Technology is the application of scientific knowledge. New scientific knowledge produces new technologies and new technologies necessarily expose new vulnerabilities in our laws and legal thinking. Blockchain technology, by allowing us to reduce and even eliminate the role of the middleman in our transactions, triggers a significant paradigm shift in how we deal with *value*. It is often said in online communities that internet democratizes *access to information* and *blockchain democratizes the access to truth.* The aim of this work is to shed light on the unchartered territory of the blockchain with the lenses of the EU data protection and privacy law, and offer an in-depth analysis of the greatest issues the blockchain presents with possible solutions and policy recommendations.

Avainsanat – Nyckelord – Keywords

GDPR, Blockchains, Privacy, Data protection, Regulation, Internet law, IT Law, Cybersecurity, Information security

Säilytyspaikka – Förvaringställe – Where deposited
http://ethesis.helsinki.fi/en/ohjeet/gradu/oikeustieteellinen

Muita tietoja – Övriga uppgifter – Additional information

TABLE OF CONTENTS

SOURCES

OFFICIAL SOURCES

1. *Charter of Fundamental Rights of the European Union*
2. *GDPR*
   General Data Protection Regulation (EU) 2016/ 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC

3. *e-IDAS*
   Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC

4. *NISD*
   Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union

5. *AML Directive*
   Directive (EU) 2015/849 of the European Parliament and of the Council of 20 May 2015 on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing, amending Regulation (EU) No 648/2012 of the European Parliament and of the Council, and repealing Directive 2005/60/EC of the European Parliament and of the Council and Commission Directive 2006/70/EC

6. *Directive 2000/31/EC* of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market

7. *Directive on settlement finality*
   Directive on settlement finality in payment and securities settlement systems, 98/26/EC 19 May 1998

8. *Patrick Breyer v Bundesrepublik Deutschland* 35 ECJ, Case C-582/14.

9. *Malone v. UK*, 2 August 1984, ECHR
10. *Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González.* 13 May 2014, Case C-131/12

SEMI-OFFICIAL SOURCES

1. *WP 2007 Opinion 136*
   Article 29 Working Party 2007 on the concept of personal data, WP 136

2. *WP 2014 Opinion 216*
   Article 29 Working Party 2014 on the Anonymisation Techniques, WP 216

3. *WP 2012 Opinion 196*
   Article 29 Working Party, 'Opinion 05/2012 on Cloud Computing' WP 196

4. *WP 2014 Opinion 215*
   Article 29 Working Party, 'Opinion 04/2014 on surveillance of electronic communications for intelligence and national security purposes' WP 215

5. *WP 2010 Opinion 169*
   Article 29 Working Party, 'Opinion 01/2010 on the concepts of "controller" and "processor"

6. *Opinion of the Hungarian Data Protection Authority*
   The Opinion of the Hungarian National Authority for Data Protection and Freedom of Information on Blockchain Technology in the Context of Data Protection dated 29.01.2018 accessed at:

   https://www.naih.hu/files/Blockchain-Opinion-2018-01-29.pdf

7. *CNIL's Opinion on the Blockchains 2018*
   accessed at: https://www.cnil.fr/en/blockchain-and-gdpr-solutions-responsible-use-blockchain-context-personal-data

8. *ICO Guidance 2014*
   ICO Guideline 2014 on Deleting Personal Data
   accessed at: https://ico.org.uk/media/for-organisations/documents/1475/deleting_personal_data.pdf

9. *Enisa report 2016 Distributed Ledger Technology*
   "Distributed Ledger Technology & Cybersecurity- improving information security in the financial sector"

10. *European Parliament Industry 4.0*
    "Industry 4.0", Study published by the Directorate for Internal Policies Policy Department A: Economic and Scientific Policy, authored by Smit, Kreutzer, Moeller, Carlberg, February 2016, accessed at:
    http://www.europarl.europa.eu/RegData/etudes/STUD/2016/570007/IPOL_STU(2016)570007_EN.pdf

11. *Communication on Critical Information Infrastructure Protection*
European Commission's Communication on Critical Information Infrastructure Protection: "Protecting Europe from large scale cyber-attacks and disruptions: enhancing preparedness, security and resilience", March 2009

12. *Green Paper on a European Programme for Critical Infrastructure Protection*
Commission's Green Paper on a European Programme for Critical Infrastructure Protection, COM(2005) 576 final

13. *Commission Proposal for a Cybersecurity Act 2017*
Regulation of the European Parliament and of the Council on Enisa, the "EU Cybersecurity Agency", and repealing Regulation (EU) 526/2013, and on Information and Communication Technology cybersecurity certification ("Cybersecurity Act")
Accessed at:
https://ec.europa.eu/info/law/better-regulation/initiatives/com-2017-477_en

14. *Cybersecurity Strategy of the European Union 2013*
Joint Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Cybersecurity Strategy of the European Union: An Open, Safe and Secure Cyberspace
Accessed at:
http://eeas.europa.eu/archives/docs/policies/eu-cyber-security/cybsec_comm_en.pdf

15. *UNCTAD Data Protection Regulations and International Data Flows 2016*
accessed at: http://unctad.org/en/PublicationsLibrary/dtlstict2016d1_en.pdf

16. *EU Blockchain Observatory Forum – Blockchain and the GDPR Report 2018*
accessed at: https://www.eublockchainforum.eu/reports

17. *EU Blockchain Observatory Forum – Blockchain and Digital Identity Report 2019*
accessed at: https://www.eublockchainforum.eu/reports

BOOKS

1. *Burchell - Gordon- Miller 1991*
Burchell, Graham; Gordon, Colin; Miller, Peter: "The Foucault Effect - Studies in Governmentality", The University of Chicago Press, 1991
Accessed at:
https://laelectrodomestica.files.wordpress.com/2014/07/the-foucault-effect-studies-in-governmentality.pdf

2. *Brownsword — Yeung 2008*

Brownsword, Roger; Yeung, Karen; "Regulating Technologies Legal Futures, Regulatory Frames and Technological Fixes". Hart Publishing 2008

3. *Brown and Marsden 2013*
   Brown, Ian; Marsden, Christopher T.; "Regulating Code", The MIT Press 2013

4. *Bräutigam —Miettinen 2016*
   Bräutigam, Tobias; Miettinen, Samuli (eds); "Data Protection, Privacy and European regulation in the digital age", Helsinki: Unigrafia 2016

5. *Corrales and others 2019*
   Corrales, Marcelo; Fenwick, Mark; and Haapio, Helena; "Legal Tech, Smart Contracts and Blockchain", Springer, 2019

6. *Foucault 1977*
   Foucault, M.: "Language, Counter-Memory, Practice", Cornell University Press, 1977

7. *Habermas 1996*
   Habermas, Jurgen; "Between Facts and Norms", The MIT Press, Cambridge, Massachusetts, 1996

8. *Hildebrandt — Vries 2013*
   Hildebrandt, Mireille and De Vries, Katja: "Privacy, Due Process and the Computational Turn: The Philosophy of Law Meets the Philosophy of Technology" (New York: Routledge, 2013), 197

9. *Lessig 2006*
   Lessig, Lawrence; "Code and Other Laws of Cyberspace Version 2.0 3", New York, Basic Books, 2006
   accessed at: http://codev2.cc/download+remix/Lessig-Codev2.pdf

10. *Kuner 2007*
    Kuner, Christopher; "European data protection law: corporate compliance and regulation", Oxford University Press 2007, second edition

11. *Millard 2014*
    Millard, Christopher J; Cloud Computing Law, International Review of Law, Computers & Technology Vol. 26, Nos. 2 –3, July –November 2012, 129 –164, Oxford University Press 2014

12. *Murray 2007*
    Murray, Andrew D.: "Regulation of Cyberspace: Control in the Online Environment", Routledge Cavendish 2007

13. *Narayanan et al 2016*

Narayanan, Arvind; Bonneau, Joseph; Felten, Edward; Miller, Andrew; Goldfeder, Steven; "Bitcoin and Cryptocurrency Technologies, A Comprehensive Introduction", Princeton University Press, Princeton and Oxford, 2016

14. *Nissenbaum 2010*
Nissenbaum, Helen: "Privacy in Context: Technology, Policy, and the Integrity of Social Life", Palo Alto: Stanford University Press, 2010

15. *Swan 2015*
Swan M.; Blockchain: Blueprint for a New Economy; O'Reilly Media, Inc. 2015

16. *Personick and Patterson 2003*
Personick, Steward D. and Patterson, Cynthia A., "Critical Information Infrastructure Protection and the Law, An Overview of Key Issues", National Academy of Engineering National Research Council, The National Academies Press, Washington D.C., 2003 available at http://nap.edu/10685

17. *Teubner 1993*
Teubner, Gunther: "Law as an Autopoietic System", Oxford 1993

18. *Teubner 1987*
Teubner, Gunther; "Juridification of the Social Spheres", European University Institute, 1987

19. *Teubner 1986*
Teubner, Gunther; "Dilemmas of Law in the Welfare State", European University Institute, 1986

20. *Turner 2009*
Turner, Bryan S. "The New Blackwell Companion to Social Theory", 2009, Blackwell Publishing Ltd.
Accessed at:
https://is.muni.cz/el/1423/jaro2012/SOC403/um/Turner_New_Blackwell_Companion_to_Social_Theory.pdf#page=156

ACADEMIC ARTICLES

1. *Acquisti et al 2013*
Acquisti, Alessandro; John, Leslie K.; Loewenstein, George; "What Is Privacy Worth?" The Journal of Legal Studies, Vol. 42, Issue 2 (June 2013), pp. 249-274

2. *Ambrose 2014*
Ambrose, Meg Leta; "Speaking of Forgetting: Analysis of Possible Non-EU Responses to the Right to Be Forgotten and Speech Exception" Telecommunications Policy, Vol. 38, issue 8-9, 800- 804 (2014).

3. *Murray* 2008
Murray, Andrew: "Conceptualising the Post-Regulatory (Cyber)state", published in *Brownsword — Yeung* 2008

4. *Atzori 2017*
Atzori, Marcella; "Blockchain Governance and the Role of Trust Service Providers: The TrustedChain® Network", University College London - Center for Blockchain Technologies, May 2017, accessed at:
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2972837

5. *Ambrose 2013*
Ambrose, Meg Leta: "It's about time: privacy, information life cycles, and the right to be forgotten", Stanford Technology Law Review, 16(2), 2013, 101-154.

6. *Baran 1962*
Baran, Paul; "On Distributed Communications Networks" September 1962 accessed at: http://pages.cs.wisc.edu/~akella/CS740/F08/740-Papers/Bar64.pdf

7. *Berberich and Steiner 2016*
Berberich and Steiner, "Blockchain Technology and the GDPR- How to Reconcile Privacy and Distributed Ledger?", 2 Eur. Data Prot. L. Rev. 422 2016.

8. *Black 2001*
Black, Julia: "Decentring regulation: understanding the role of regulation and self-regulation in a post-regulatory world", 54(1) Current Legal Problems (2001), pp. 103-146.

9. *Black 2000- Part I*
Black, Julia: "Proceduralizing Regulation: Part I", Oxford Journal of Legal Studies, Vol. 20, No. 4 (2000), p.597-614.

10. *Black 2001 - Part II*
Black, Julia: "Proceduralizing Regulation Part II", Oxford Journal of Legal Studies, Vol. 21, No.1 (2001), p. 33-58.

11. *Benkler 2000*
Benkler, Y: 'From Consumers to Users: Shifting the Deeper Structures of Regulation Toward Sustainable Commons and User Access' (2000) 52 *Federal Communications Law Journal,* 561

12. *Borgesius 2016*
Borgesius, Frederik J. Zuiderveen; "Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation", Computer law & Security review 32 (2016) 256–271

13. *Bräutigam 2016*
Bräutigam, Tobias, 2016: "The Land of Confusion: International Data Transfers between Schrems and the GDPR" published in *Bräutigam —Miettinen 2016*

14. *Carbone 2015*
Carbone, Chelsea; "To Be or Not to Be Forgotten: Balancing the Right to Know with the Right to Privacy in the Digital Age", 22 Va. J. Soc. Policy & L. 525 2015

15. *Crawford and Schultz 2014*
Crawford, Kate; Schultz, Jason: "Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms", 55 B.C.L. Rev. 93 (2014),
Accessed at: http://lawdigitalcommons.bc.edu/bclr/vol55/iss1/4

16. *Carrapico and Barrinha 2017*
Carrapico, Helena; Barrinha, Andre; "The EU as a Coherent (Cyber)Security Actor?", JCMS 2017 Volume 55, No.6, pp.1254-1272

17. *Christou 2017*
Christou, George; "EU-Japan Security Cooperation: Challenges and Opportunities"; University of Essex, Online paper series, Spring/Summer 2017

18. *De Montjoye 2015*
De Montjoye, Yves-Alexandre; "Computational Privacy: Towards Privacy-Conscientious Uses of Metadata", 2015
Accessed at: https://dam-prod.media.mit.edu/x/files/thesis/2015/yva-phd.pdf

19. *Dewey 1924*
Dewey, John: Logical Method and Law, 10 Cornell L. Rev. 17, 1924
Accessed at: http://scholarship.law.cornell.edu/clr/vol10/iss1/2

20. *De Filippi 2016*
De Filippi, Primavera; "The interplay between decentralization and privacy: the case of blockchain technologies"; Journal of Peer Production, 2016, Alternative Internets, 7. <hal-01382006> accessed at: https://hal.archives-ouvertes.fr/hal-01382006/document

21. *Dori et al 2017*
Dorri, Ali; Kanhere, Salil S.; Jurdak, , Raja; Gauravaram, Praveen: Blockchain for IoT Security and Privacy: The Case Study of a Smart Home, 2nd IEEE Percom Workshop On Security Privacy And Trust In The Internet of Things 2017

22. *Drabiak 2017*
Drabiak, Katherine: "Caveat Emptor: How the Intersection of Big Data and Consumer Genomics Exponentially Increases Informational Privacy Risks" 27 Health Matrix 143 (2017)

23. *Easterbrook 1996*
Easterbrook, Frank H; "Cyberspace and the Law of the Horse" 1996 U. Chi. Legal F. 207 1996

24. *Emam — Alvarez 2015*
Emam, Khaled El; Alvarez, Cecilia; "A critical appraisal of the Article 29 Working Party Opinion 05/2014 on data anonymization techniques", International Data Privacy Law, 2015 Vol. 5 No.1

25. *Esayas 2015*
Esayas, Samson Yoseph; "The Role of Anonymisation and Pseudonymisation under the EU data privacy rules: beyond the 'all or nothing' approach"; European Journal of Law and Technology, Vol 6, No 2 (2015), p. 13

26. *El Khoury 2017*
El Khoury, Alessandro; "Dynamic IP Addresses Can Be Personal Data, Sometimes. A Story of Binary Relations and Schrodinger's Cat"; European Journal of Risk Regulation (EJRR), Vol. 8, Issue 1 (March 2017), pp. 191-197

27. *Fahey 2014*
Fahey, Elaine; The EU's Cybercrime and Cyber-Security Rulemaking: Mapping the Internal and External Dimensions of EU Security, 5 Eur. J. Risk Reg. 46 (2014)

28. *Finck 2017*
Finck; "Blockchains and Data Protection in the European Union", 2017, 10(1) Max Planck Institute for Innovation and Competition Research Paper Series

29. *Gandy — Danna 2002*
Gandy Jr, Oscar H.; Danna, Anthony; "All That Glitters is Not Gold: Digging Beneath the Surface of Data Mining", Journal of Business Ethics 40: 373–386, 2002.
Accessed at:
https://pdfs.semanticscholar.org/59cb/9baed12410bd989e627f5795670d73d5eaf7.pdf

30. *Gonçalves 2017*
Gonçalves, Maria Eduarda (2017); "The EU data protection reform and the challenges of big data: remaining uncertainties and ways forward"; Information & Communications Technology Law, 26:2, 90-115 DOI: 10.1080/13600834.2017.1295838

31. *Serge, De Hert and Sutter 2008*
Gutwirth, Serge; De Hert, Paul; Sutter, De Laurent: "The Trouble with Technology Regulation: Why Lessig's 'Optimal Mix' Will Not Work" in *Brownsword — Yeung 2008*

32. *Hacker 2017*
Hacker, Philipp: "Personalizing EU Private Law: From Disclosures to Nudges and Mandates, European Review of Private Law" 3-2017 [651–678] © 2017 Kluwer Law International BV, The Netherlands

33. *Haber —Stornetta* 1991
Haber, Stuart; and Stornetta, W. Scott; "How to Time-Stamp a Digital Document", Journal of Cryptology, Vol. 3, No. 2, pp. 99{111, 1991.
accessed at:
https://www.anf.es/pdf/Haber_Stornetta.pdf

34. *Habermas 1996*
Habermas, Jurgen: "Three Normative Models of Democracy" in "Democracy and Difference: Contesting the Boundaries of the Political", edited by Seyla Benhabib, 1996

35. *Hendriks 2002*
Hendriks, Aart.: "Genetic Discrimination: How to Anticipate Predictable Problems?", Editorial, European Journal of Health Law. June 2002, Vol. 9 Issue 2, p.87-92.

36. *Khan — Salah 2017*
Khan, Minhaj Ahmad; Salah, Khaled; "IoT security: review, blockchain solutions and open challenges", Future Generation Computer Systems, 2017.

37. *Kshetri 2017*
Kshetru, Nir; "Blockchain's Roles in Strengthening Cybersecurity and Protecting Privacy", Telecommunications Policy 41 (2017) 1027-1038

38. *Kiviat 2015*
Kiviat, Trevor I, "Beyond Bitcoin, issues in regulating Blockchain"; Duke Law Journal. Dec. 2015, Vol. 65 Issue 3, p569-608.

39. *Klimas and Vaiciukaite 2008*
Klimas, Tadas; and Vaiciukaite, Juratee; "The Law of Recitals in European Community Legislation" ILSA Journal of International & Comparative Law (2008)(15), p. 2–33

40. *Kuner et al 2017*
Kuner, Christopher; Svantesson, Dan Jerker B.; Cate, Fred H.; Lynskey, Orla; Millard, Christopher: "The Rise of Cybersecurity and its Impact on Data Protection", International Data Privacy Law, 2017, Vol. 7, No. 2, Editorial

41. *Kokott and Sobotta 2013*
Kokott, Juliane; Sobotta, Christoph: "The Distinction between Privacy and Data Protection in the Jurisprudence of the CJEU and the ECtHR", International Data Privacy Law, 2013, Vol. 3, No.4

42. *Kuner et al 2015*
Kuner, Christopher; Svantesson, Dan Jerker B.; Cate, Fred H.; Lynskey, Orla; Millard, Christopher: "Risk management in data protection", International Data Privacy Law, 2015, Vol. 5, No. 2., Editorial

43. *Koops and Leenes 2014*
Koops, BJ., and Leenes, R., "Privacy regulation cannot be hardcoded. A critical comment on the privacy by design provision in data-protection law" International Review of Law, Computers & Technology, 2014, vol. 28, No. 2, 159-171

44. *Koops 2008*
Koops, Bert-Jaap: "Criteria for Normative Technology: The Acceptability of 'Code as Law' in Light of Democratic and Constitutional Values", published in *Brownsword-Yeung* 2008

45. *Koops 2011*
Koops, Bert-Jaap; "Forgetting Footprints, Shunning Shadows. A Critical Analysis of the "Right To Be Forgotten" in Big Data Practice"; 8 Scripted 229, 236 (2011).

46. *Koops 2013*
Koops, Bert-Jaap: "On decision transparency, or how to enhance data protection after the computational turn" 2013 in: M. Hildebrandt & K. De Vries (eds): "Privacy, Due Process and the Computational Turn", Abingdon: Routledge, p. 196-220

47. *Kuan et al 2011*
H Kuan, C Millard, and I Walden, "The problem of 'personal data' in cloud computing: what information is regulated? — the cloud of unknowing" (2011) 1(4) International Data Privacy 37.

48. *Kuan et al 2012*
Hon, W. Kuan; Hornle, Julia; Millard, Christopher; "Data Protection Jurisdiction and Cloud Computing - When Are Cloud Users and Providers Subject to EU Data Protection Law: The Cloud of Unknowing" International Review of Law, Computers & Technology, Vol. 26, Issue 2-3 (July-November 2012), pp. 129-164

49. *Kuner and others 2018*
Kuner, Christopher; "Blockchain versus Data Protection" 2018, 8, International Data Privacy Law 103

50. *Law 2007*
Law, John: "Actor Network Theory and Material Semiotics" version of 25th April 2007 Accessed at:
http://www.heterogeneities.net/publications/Law2007ANTandMaterialSemiotics.pdf
A publication-version of the same article can be found at *Turner* 2009 p. 141.

51. *Lenard and Rubin 2010*
Lenard, Thomas M., and Paul H. Rubin. 2010; "In Defense of Data: Information and the Costs of Privacy"; Policy and Internet 2:149–83.

52. *Li et al 2014*
Wei Peng, Feng Li, Xukai Zou, and Jie Wu; "A Two-Stage Deanonymization Attack against Anonymized Social Networks" IEEE Transactions on Computers, Vol. 63, No.2, February 2014.

53. *Lessig 1998*
Lessig, Lawrence; "Commentaries, The Law of the Horse: What Cyberlaw Might Teach", Harvard Law Review, Vol. 113:501

54. *Luhmann 1989*
Luhmann, Niklas; "Law As a Social System", 83 Nw. U. L. Rev. 136 (1988-1989)

55. *Lundevall-Unger and Tranvik 2010*
Lundevall-Unger, Patrick; and Tommy, Tranvik; "IP Addresses- Just a Number?" International Journal of Law and Information Technology, Vol. 19 No.1 2010

56. *Luzio et al, 2017*
Di Luzio, A. Mei and J. Stefa, "Consensus Robustness and Transaction De-Anonymization in the Ripple Currency Exchange System," *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, Atlanta, GA, 2017, pp. 140-150.

http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7979962&isnumber=7979941

57. *Mak 2010*
Mak, Vanessa: Standards of Protection: In Search of the 'Average Consumer' of EU Law in the Proposal for a Consumer Rights Directive (June 17, 2010). European Review of Private Law, Vol. 18, 2010. Available at SSRN: https://ssrn.com/abstract=1626115

58. *Mayer-Schönbergert 2010*
Mayer-Schönbergert, Viktor; "Beyond Privacy, beyond Rights - Toward a Systems Theory of Information Governance", 98 Calif. L. Rev. 1853 (2010)

59. *Murray and Scott 2002*
Murray, A and Scott, C 'Controlling the New Media: Hybrid Responses to New Forms of Power', 2002, 65 MLR 491
Accessed at:
http://researchrepository.ucd.ie/bitstream/handle/10197/6754/ControllingNewMediaHybridResponses.pdf?sequence=2

60. *Narayanan and Shmatikov 2010*
Arvind Narayanan, Vitaly Shmatikov; "Myths and fallacies of personally identifiable information" Communications of the ACM, 2010/6/1, volume 53, pp. 24-26

61. *Nissenbaum 2017*
Nissenbaum, Helen: "Deregulating Collection: Must Privacy Give Way to Use Regulation?" (May 1, 2017). Available at SSRN: https://ssrn.com/abstract=3092282

62. *Ohm 2010*
Ohm, Paul; "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization", 57 UCLA Law Review 1701, 1778 (2010)

63. *O'Hara and Shadbolt 2015*
O'Hara, Kieron; Shadbolt, Nigel, "The Right to be Forgotten: Its Potential Role in a Coherent Privacy Regime", 1 Eur. Data Prot. L. Rev. 178, 189 (2015)

64. *Park — Park 2017*
Park, Jin Ho and Park, Jong Hyuk Park; "Blockchain Security in Cloud Computing: Use Cases, Challenges, and Solutions"; Symmetry (20738994), Aug. 2017, Vol. 9 Issue 8, p1-13.

65. *Pazaitis 2017*
Pazaitis A., "Blockchain and value systems in the sharing economy: The illustrative case of Backfeed", Technological Forecasting & Social Change (2017), http://dx.doi.org/10.1016/j.techfore.2017.05.025

66. *Pfitzmann and Hansen 2010*
A.Pfitzmann, M.Hansen; "A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity" accessed at:
http://dud.inf.tu-dresden.de/Anon_Terminology.shtml
the final version:
http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.34.pdf

67. *Polenetsky et al. 2016*
Polonetsky, Jules; Tene, Omer; Finch, Kelsey; "Shades of Gray: Seeing the Full Spectrum of Practical Data De-Identification"; Santa Clara Law Review, Vol. 56, Issue 3 (2016), pp. 593-630

68. *Post and Johnson 1998*
Post, David G.; Johnson, David R.: "Chaos Prevailing on Every Continent: Towards a New Theory of Decentralized Decision-Making in Complex Systems", Chicago-Kent Law Review, Vol. 73, Issue 4 (1998), pp. 1055-1100

69. *Posner 1978*

Posner, Richard: “An Economic Theory of Privacy” 2 Regulation 19, 1978

70. *Popper 1965*
Popper, Karl: Three Views Concerning Human Knowledge. In Karl Popper "Conjectures and Refutations: The Growth of Scientific Knowledge", Chapter 3, London, Routledge, 1965, pp. 97-119.

71. *Raab and De Hert 2008*
Raab, Charles D.; De Hert, Paul: “Tools Regulation: Seeking Analytical Approaches Beyond Lessig and Hood” in Brownsword and Yeung 2008.

72. *Reed and others 2017*
Reed, Chris; Sathyanarayan, Umamahesh; Ruan, Shuhui; Collins, Justine; “Beyond BitCoin – legal impurities and off-chain assets” 2017, Queen Mary School of Law Legal Studies Research Paper No. 260/2017.

73. *Rees 2014*
Rees, Christopher; Heywood, Debbie; “The ‘right to be forgotten’ or the ‘principle that has been remembered”, Computer law & Security review 30 (2014) 574 -578

74. *Reid — Harrigan 2013*
Reid — Harrigan, “An analysis of anonymity in the bitcoin system”, Springer New York 2013

75. *Reidenberg 1998*
Reidenberg, Joel: “Lex Informatica: The Formation of Information Policy Rules Through Technology”, 1998, 76 Texas Law Review 553

76. *Rustad and Kulevska 2015*
Rustad, Michael L; Kulevska, Sanna; “Reconceptualizing the Right to be Forgotten to enable Transatlantic Data Flow”; Harvard Journal of Law & Technology Volume 28, No.2 Spring 2015

77. *Rust 1997*
Rust, John: “Dealing with the Complexity of Economic Calculations”, Yale University, October, 1997 accessed at:
https://pdfs.semanticscholar.org/3a17/a9a56a41f4166eac1874043a3d044bd96440.pdf

78. *Salmensuu 2018*
Salmensuu, Cagla; ”The General Data Protection and the Blockchains”, Liikejuridiikka 1/2018, accessed at:
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3143992

79. *Smolander et al 2016*

Yli-Huumo J, Ko D, Choi S, Park S, Smolander K (2016) "Where Is Current Research on Blockchain Technology?—A Systematic Review"; PLoS ONE 11(10): e0163477. doi:10.1371/journal. Pone.0163477

80. *Solum – Chung 2004*
Solum, Lawrence; Chung, Minn: The Layers Principle: Internet Architecture and the Law, 4-1-2004, Notre Dame Law Review, Vol. 79, Issue 3

81. *Spindler and Schmechel 2016*
Spindler, Gerald and Schmechel, Philipp; "Personal Data and Encryption in the European General Data Protection Regulation" 7 (2016) JIPITEC
Accessed at:

https://www.jipitec.eu/issues/jipitec-7-2-2016/4440/spindler_schmechel_gdpr_encryption_jipitec_7_2_2016_163.pdf

82. *Stalla-Bourdillon — Knight 2017*
Stalla-Bourdillon, Sophie; Knight, Alison; "Anonymous Data v. Personal Data- A False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data" Wisconsin International Law Journal, Vol:34 No:2 2016-2017

83. *Strandburg 2014*
Strandburg, Katherine J.: "Monitoring, Datafication and Consent: Legal Approaches to Privacy in the Big Data Context" in "Privacy, Big Data, and the Public Good: Frameworks for Engagement (pp. 5-43) eds: J. Lane, V Stodden, S. Bender & H. Nissenbaum, Cambridge University Press, 2014

84. *Stigler 1980*
Stigler, George J.; "An Economic Introduction to Privacy in Economics and Politics", 9 J. Legal Stud. 623 1980, p. 625

85. *Susskind 2010*
Susskind, Richard "Rethinking the nature of legal services" (Oxford and New York: Oxford University Press, 2010)

86. *Subramanian 2018*
Subramanian, Hemang; "Decentralized Blockchain-Based Electronic Marketplaces", Magazine Communications of the ACM, Vol. 61 Issue 1, January 2018, pp.: 78-84

87. *Sweeney 2013*
Sweeney, Latanya; "Matching known patients to health records in Washington state data" Available at SSRN 2289850, 2013.

88. *Sweeney 2002*

Sweeney, Latanya; "k-anonymity: A model for protecting privacy", International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10(05):557– 570, 2002. Accessed at: https://epic.org/privacy/reidentification/Sweeney_Article.pdf

89. *Swanson 2015*
Swanson, Tim: Consensus as a Service: A Brief Report on the Emergence of Permissioned, Distributed Ledger Systems, April 6, 2015, accessed at: http://www.ofnumbers.com/wp-content/uploads/2015/04/Permissioned-distributed-ledgers.pdf

90. *Teubner 1983*
Teubner, Gunther: "Substantive and Reflexive Elements in Modern Law", 17 Law & Society Review 239, 1983

91. *Tuori 2002*
Tuori, Kaarlo; "Critical Legal Positivism- (Applied legal philosophy)", 2002 Ashgate Publishing

92. *Vranaki 2017*
Vranaki, Asma A. I.: "Regulating Social Networking Sites: Facebook, Online Behavioral Advertising, Data Protection Laws and Power", 43 Rutgers Computer & Tech. L.J. 168 (2017)

93. *Walker 2002*
Walker, Kent; "The Costs of Privacy"; 25 Harv. J.L. & Public Policy 87, 2001-2002

94. *Whittington and Hoofnagle 2012*
Whittington, Jan & Hoofnagle, Chris Jay: "Unpacking Privacy's Price", 90 N.C. L. Rev. 1327 2011-2012

95. *Wright - Raab 2014*
Wright, David; Raab, Charles: "Privacy principles, risks and harms", International Review of Law, Computers & Technology, 2014, Vol. 28, No. 3, 277–298

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

SYMPOSIUM AND CONFERENCE PAPERS

1. *Rubinstein 2016*
Rubinstein, Ira; Brussels Privacy Symposium on Identifiability: Policy and Practical Solutions for Anonymisation and Pseudonymisation; "Framing the Discussion", 2016 Accessed at: https://fpf.org/wp-content/uploads/2016/11/Rubinstein_framing-paper.pdf

2. *Narayanan & Shmatikov 2009*
Narayanan, Arvind; and Shmatikov, Vitaly; "De-anonymizing Social Networks", 2009 30th IEEE Symposium on Security and Privacy, The University of Texas at Austin

3. *Conti, Kumar E, Lal, Ruj 2017*
Conti, Mauro; Kumar E, Sandeep; Lal, Chhagan; Ruj, Sushmita: "A Survey on Security and Privacy Issues of Bitcoin" July 5, 2017, arXiv:1706.00916 [cs.CR] accessed at: https://arxiv.org/pdf/1706.00916.pdf

4. *Acquisti 2010*
Acquisti, Alessandro: "The Economics of Personal Data and the Economics of Privacy" Background Paper #3 to the OECD Conference 2010, Working Party for Information Security and Privacy (WPISP) Working Party on the Information Economy (WPIE)

NEWSPAPERS AND MAGAZINES AND OTHER REPUTABLE WEB SOURCES

1. Public vs Private Blockchain, PC Magazine. Feb. 2017, p.114-115. 2p.
2. Sams 2015
Sams, Robert: Blockchain Finance, presentation dated March 2015
accessed at: https://www.slideshare.net/rmsams/blockchain-finance
3. Bitcoin and Beyond, Underwood, Sarah. Communications of the ACM. Nov2016, Vol. 59 Issue 11, p15-17. 3p. 1 Color Photograph. DOI: 10.1145/2994581.
4. Birdsall, William F.; "The internet and the ideology of information technology" accessed at: https://www.isoc.org/inet96/proceedings/e3/e3_2.htm
5. Why you can't really anonymise your data
https://www.oreilly.com/ideas/anonymize-data-limits
6. Why anonymous data sometimes isn't
https://www.wired.com/2007/12/why-anonymous-data-sometimes-isnt/
7. Is Bitcoin anonymous?
https://bitcoinmagazine.com/articles/is-bitcoin-anonymous-a-complete-beginner-s-guide-1447875283/
8. So you want to use a Blockchain for that?
https://bitsonblocks.net/2016/07/19/so-you-want-to-use-a-blockchain-for-that/
9. Trend towards privacy
http://sammantics.com/blog/2016/8/23/the-trend-towards-privacy-how-blockchains-plan-to-accomplish-this
10. Legal Analysis of Single Market for the Information Society 2011
European Commission's Information Society and Media Directorate-General, Legal analysis of a Single Market for the Information Society, chapter 4: The future of online privacy and data protection, prepared by DLA Piper 2011 accessed 24 January 2016, pp. 18–21
11. Less Privacy is the Social Norm
Gonsalves, Antone; 2010, accessed at:
https://www.darkreading.com/risk-management/facebook-ceo-less-privacy-is-social-norm/d/d-id/1086052?piddl_msgorder=asc

12. Foggy Thinking About the Right to Oblivion
Fleischer, Peter; March 9, 2011, http://peterfleischer.blogspot.com/2011/03/foggy-thinkingabout-right-to -oblivion.htm.
13. Big Data, Trying to Build Better Workers
Lohr, Steve; N.Y. Times, Apr. 21, 2013 accessed at:
http://www.nytimes.com/2013/04/21/technology/big-data-trying-to-build-better-workers.html
14. Blockchain And Cryptocurrency May Soon Underpin Cloud Storage
Mearian, Lucas; Computerworld, March 26, 2018, accessed at:
https://www.computerworld.com/article/3250274/data-storage/blockchain-and-cryptocurrency-may-soon-underpin-cloud-storage.html

ABSTRACT

Technology is the application of scientific knowledge. New scientific knowledge produces new technologies and new technologies necessarily expose new vulnerabilities in our laws and legal thinking. Blockchain technology, by allowing us to reduce and even eliminate the role of the middleman in our transactions, triggers a significant paradigm shift in how we deal with *value*. It is often said in online communities that internet democratizes *access to information* and *blockchain democratizes the access to truth.* The aim of this work is to shed light on the unchartered territory of the blockchain with the lenses of the EU data protection and privacy law, and offer an in-depth analysis of the greatest issues the blockchain presents with possible solutions and policy recommendations.

1 INTRODUCTION

The influence of the new information technologies on privacy has been debated since the portable camera was developed in the nineteenth century.[1] The genesis of the idea of blockchain is found in the works of Haber and Stornetta in 1991,[2] at a time when securing digital documentation has become a burning need. Their proposal consisted of a system for a secure and immutable timestamping of digital documents. By their method, the date of creation of documents would be provided which would also show order of creation. This method operated on a certification principle whereby a document's owner, upon sending the document to the server, received a certificate with the information.

Möller describes the Fourth Industrial revolution[3] as a both regulatory and a *sui generis* concept.[4] This research locates the utility of blockchain (decentralized platforms) within the

[1] *Brown* 2013 Chapter 3 para. 4
[2] *Haber — Stornetta 1991*
[3] European Parliament Industry 4.0, p. 20
[4] *Bräutigam - Miettiten*, 2016 p.36, Möller's article "Is EU law 'fit to go digital'?"

global Industry 4.0 and regards it a *sui generis* system which needs a *sui generis* approach from the privacy and data protection angles.

Data protection is not the subject of a single, global treaty or agreement. Rather, it is included in a range of international and regional instruments, such as the GDPR, and some are mere voluntary guidelines. What is more, each instrument covers a particular group of countries and vary in scope and application.[5] The Universal Declaration of Human Rights Art. 12 and International Covenant on Civil and Political Rights Art. 17 both govern the right to privacy on a supranational and universal level. Moreover, the UN further published the Statement on the Right to Privacy in the Digital Age[6] supported by the Resolution 68/167 adopted in 2013 to affirm that the "rights held by people offline must also be protected online".[7] The Resolution notes that international human rights law provides the universal framework against which any interference to individual privacy rights must be assessed.[8] The most prominent binding international agreement on data protection is the Council of Europe Data Protection Convention of 1981.[9] Its membership is open to any country, and several non-European countries have signed the Convention or are in the process of becoming members.[10] The second most prominent international document, although not binding, is the OECD Guidelines on the Protection of privacy and Transborder Flows of Personal Data published in 1980 and revised in 2013. Among the OECD members, the US is currently the only country that has not implemented comprehensive data protection laws.[11] The OECD Guidelines contain eight privacy principles that form the backbone of the principles included in most national privacy laws.[12] The final globally influential initiative is the International Data Protection Commissioner's Initiatives which can be summarized as 1) an annual meeting and conference; 2) a system for cooperating in international and cross-border complaints; and 3) a statement on

[5] UNCTAD Data Protection Regulations and International Data Flows 2016 p.24
[6] http://www.ohchr.org/EN/Issues/DigitalAge/Pages/DigitalAgeIndex.aspx
[7] http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/68/167
[8] UNCTAD Data Protection Regulations and International Data Flows 2016 p.26
[9] Also referred to as Convention 108 or the CoE Convention.
[10] Forty-six members out of forty-seven have ratified it and implemented data protection laws that comply with the Convention. Turkey has passed a data protection law compatible with the EU legislation.
[11] UNCTAD Data Protection Regulations and International Data Flows 2016 p. 26
[12] http://www.oecd.org/internet/ieconomy/privacy-guidelines.htm

global privacy principles.[13] The International Data Protection Commissioners made a statement which is cited as the Montreux Declaration in 2005 calling for the development of an international convention on data protection, and it is one of the most significant efforts to harmonize data protection laws around the globe.[14] Germany is the first country that passed a Data Protection Law in 1970. The EU Directive 95/46/EC, the predecessor of the GDPR, was passed in 1995. So, in the EU, too, information law is still at its infancy and we are only at the beginning of a long journey of legal development.[15] The volume of digital data will expand at a compound annual growth rate of 42% over the decade of 2010 to 2020.[16] In the EU data protection law, right to privacy and data protection is defined to be a fundamental right[17] and remains to be the most developed data protection law globally that influences other nations.

The aim of this thesis is to open up new academic debates and deepen the existing ones in relation to the data protection and privacy on the blockchains. With this on mind, the findings and recommendations of this thesis are intended to be critically read, tested and checked against the expertise of the readers for their usability.

1.1 CHOICE OF SUBJECT

At the time of starting this thesis in 2017, there was no other in-depth academic work in the field of data protection and privacy law as applied to the blockchain. The subject matter of the thesis was deliberately selected to lay the foundation of further advanced research in this specific field. In this thesis, the term 'Blockchains' is interchangeably used with the 'Distributed Ledgers". With the opportunity to re-submit this thesis now in 2019, I have had the time to review the new studies done in this field, and include them here. Moreover, I took the time to critically read and assess my own work and apply the feedback I received when I submitted it in 2018.

[13] UNCTAD Data Protection Regulations and International Data Flows 2016 p.27
[14] https://icdppc.org/wp-content/uploads/2015/02/Montreux-Declaration.pdf
[15] *Bräutigam* 2016, p. 169
[16] IDC white paper summarized in https://itknowledgeexchange.techtarget.com/quality-assurance/data-age-2025/
[17] Charter of Fundamental Rights, Art. 7

A large part of this chapter relating to the GDPR and its findings (below in chapters 4-5) was peer-reviewed and published in *Liikejuridiikka* on January 2018.[18]

The research question I have sought to answer in this work are as follows:

> **Research question 1:** What are the challenges and/or benefits of blockchains vis-à-vis privacy and data protection, and cybersecurity legislation in the EU?
>
> **Research question 2:** What are the policy considerations and conclusions in the regards to the regulatory landscape in data protection and privacy vis-à-vis new technologies?

1.2 LIMITATIONS AND CHALLENGES

Blockchain is a newly developing area and has only recently attracted the attention of regulators. Even though it has been used since 2009, technology communities involving start-ups and large corporations ranging from financial institutions to insurance companies are currently researching ways in which they could use the blockchain to optimize their services and increase their profitability. Hence, the real need to undertake a legal academic research has only recently materialized. This research's primary limitation has been that there is rather limited in-depth work investigating the area and working through hypothetical scenarios which may assist or guide the regulators and policy-makers. With this research, it has become clear to me that there is a real and growing need for an empirical study on the European side of cyber-regulation of this technology.[19] Further research is required, perhaps at the PhD. level, to take those missing steps with empirical emphasis which emerged with this study.

The second limitation has been that there is not yet a legal court case that presents a problem under the GDPR as of writing this thesis. The third limitation is the impossibility of avoiding

[18] *Salmensuu 2018*

[19] Building upon the existing regulation theories, Murray has attempted to bring the cyber-regulation discussion which has largely been dominated by American scholars and technologists to Europe (*Murray* 2007). Murray links the debate to systems theory and system dynamics. However, there is still a limited amount of scholarly source from which I was able to draw my framework within the European context.

the technical information about computers and networks in explaining the infrastructure of the blockchain. Technology law is an interdisciplinary area and requires an understanding of the technical aspects of how it works in order to reach a working analysis of the applicable law. This may at times suffocate the text even though the author has tried her best to keep it to minimum. The fourth limitation is related to the language and the international nature of the subject. This research uses sources is English only; as a result, it excludes sources written in Finnish or any other language, in particular, the rich German data protection and privacy literature.

1.3 METHODOLOGY AND STRUCTURE OF THE THESIS

Method of the research is primary source review, literature review and critical analysis. This thesis deals with the research questions in two ways:

1) a legal analysis on the European data protection regulation and other related legal instruments as applied to blockchain architecture, and
2) policy recommendations and conclusions on the basis of the findings of the foregoing two research questions.

In addressing the research questions, this thesis first breaks down the structure of the blockchains and its architecture as a background in Chapter 2. Then in the same Chapter 2, we look at the meaning of regulation of the blockchains as a new digital technology. After setting the scene with those background data, we address the salient issues regarding the EU privacy and data protection law and cybersecurity law on the blockchains in Chapter 3. Finally, in Chapter 4, conclusions and policy recommendations are presented with a critical analytical view, drawing from the legal sociology.

I have analysed the meaning of regulating the blockchains using two approaches: the layered approach and the systems theory approach. The layered approach is the natural consequence of the fact that the cyberspace is a more complex and layered environment than the physical

space,[20] while the systems theory approach accounts for the sociology of the law. In the former, Lessig's famous four modes of regulation are used as the primary framework which is developed by Murray.[21] In dealing with the systems theory application, this research has taken guidance from Vranaki.[22] Vranaki roughly divides the discussion into two: earlier debates which gave rise to the argument that the cyber-space could not be regulated by the off-line laws and the Cyber-regulatory Theory ("CRT") school which arose against that notion.[23] CRT literature is dominated by the Lessigian group of scholars who do not necessarily hold the same 'tools-only' perspective as Lessig himself; however, build upon his framework with more emphasis on how the multiple actors interact with each other and such tools in context. [24] The other major source from which this thesis has drawn is Murray, also part of the CRT literature. Murray's guidance is divided into two groups: US scholarship and the European scholarship.[25] Lessig, Benkler and Reidenberg, whom this thesis draws from heavily, are from the other side of the Atlantic. Outside of the guidance of Vranaki and Murray, this thesis has also drawn heavily from another prominent European scholarship who has produced work in English: Koops.

Due to the page limits of the master's thesis, this research does not go further into opening up the CRT school and the systems theory approach apart from benefitting from them in the Chapter 2 background section on regulation and Chapter 4 policy recommendations.

In answering the first research question, this thesis reviewed literature in English language comprehensively under two major sub-headings: privacy and security, and conceives those as the two pillars of data protection. The Charter of Fundamental Rights enshrines the right to privacy and the protection of personal data as a fundamental right respectively in Article 7 (the

[20] *Murray* 2008 p.299, also see the comprehensive thesis: *Solum-Chung* 2004 on the layers approach and law.
[21] *Lessig* 1998 p.506: "behavior is regulated by four kinds of constraints together: *law*, *markets*, *norms* and *architecture.* The sum of the regulatory effects of the four modalities together yield the "net regulation" of any particular policy. The resulting policy is a trade-off among these four regulatory tools based on what works best. These four modalities are applied to the cyberspace just as applied to the real space. The norms regulate behaviour by setting the boundaries of what is appropriate and what is not."
[22] *Vranaki* 2017
[23] *Ibid* p. 175
[24] *Vranaki* 2017 p. 171, *Murray* 2007, *Raab and De Hert* 2008
[25] *Murray* 2007 preface

private life) and Article 8 (the personal data).[26] Distinct statements of such rights indicate that the Charter does not treat privacy and data protection as synonymous.[27] An analysis of this issue based on the jurisprudence concluded that private life does not necessarily include all information on identified or identifiable person while the data protection covers also this type of information therefore the scope of the data protection is broader than the scope of privacy.[28] This classification is central to the structure of this thesis. The relationship of data security with the data protection, however, is made complex by the privacy pillar. In the post-industrial society, it is now established that many measures used to strengthen cybersecurity pose a risk to privacy.[29] With this inherent tension between security and privacy on mind, the modern compliance and risk management in data protection practice is a fact-based act of balancing between the two competing and sometimes overlapping areas. Privacy sub-section is informed by the General Data Protection Regulation while the security sub-section addresses potentials under one new Directive (NSID) and one not-so-old regulation (eIDAS). Moreover, privacy sub-section opens up the GDPR-dominated issues on the basis of the specific questions raised by application of the regulation itself: data controllership, jurisdiction, breadth of the definition of personal data, data subject rights, and finally the potential privacy harms. Due to the nature of those sub-topics, this sub-section draws from interdisciplinary literature.

Swan identifies seven technical challenges and limitations for the adaptation of blockchain technology in the future: throughput, latency, size and bandwidth, *security*, wasted resources, usability, and versioning problems.[30] Smolander et al, under a systematic review of the current stage of the blockchain research using Swan's classification, identifies a new type of limitation: *privacy*.[31] According to the review of the current available research, 34% of the papers were related to the challenges and limitations in Blockchain and Bitcoin security; and 24% of the papers made suggestions related to the privacy issues.[32] This research follows Smolander et al

[26] Respectively, Art. 7 and Art. 8 of the EU Charter of Fundamental Rights
[27] *Kokott and Sobotta* 2015 p. 223
[28] *Ibid* p. 225
[29] *Kuner* et al 2017
[30] *Swan* 2015
[31] *Smolander et al* 2016 p 11
[32] *Ibid* 2016 p. 12 and 17

and uses their eight-challenges parameter[33] and focuses on the two major challenges in way of its use: privacy and security. Correspondingly, this thesis expounds on the issues of *privacy* and *security* under the GDPR in parallel fashion to the classification of the papers by Smolander et al.[34]

In the policy recommendations and conclusions, this research has benefitted from the regulatory – legal sociology review conducted in laying the background, as applied to the problem at hand.

1.4 SOURCES OF THE RESEARCH

The most important and primary source of this thesis is the General Data Protection Regulation that is enforceable from 25th May 2018. The GDPR replaces the currently effective Directive and is directly applicable. In the hierarchy of EU sources, regulations are secondary to the Treaty on the Functioning of the European Union (TFEU) and its protocols. EU laws are supreme to the national laws. The supremacy of the EU law was first doctrinally discussed in *Van Gend en Loos*[35] and established by in *Costa v Enel*.[36] The GDPR is based on Article 16 of the TFEU which specifically introduces the right to protection of personal data, and ascribes the role of establishing the rules related to this area to the European Parliament and the Council and the role of ensuring compliance to the independent authorities.

The Working Party 29 ("WP29") advises on data protection issues and assists interpretation of the Directive and the Regulation. Its role is to act as an independent advisory body of representatives from the Member States. The opinions published by WP 29, however, are not legally binding. Its influential role in data protection nevertheless renders their opinion an important source of information for this thesis on technologically unclear issues.

[33] *Ibid* 2016 p. 20
[34] *Ibid* 2016 p 20 fig. 11
[35] Judgment of the Court of 5 February 1963; NV Algemene Transport- en Expeditie Onderneming van Gend & Loos v Netherlands Inland Revenue Administration.
[36] Judgment dated July 15th 1964, Flaminio Costa v E.N.E.L - Reference for a preliminary ruling: Giudice conciliatore di Milano- Italy. Case no. 256-257

Other sources of this research are academically published articles, books and conference publications as well as non-academic credible web sources which provide the opportunity to capture the current context in which technological developments are presently discussed by the society. It is rather important to note that due to the novel nature of this technology as a target before the legal-regulatory lenses, there exists considerable limitation on the diversity of literature specifically addressing the blockchains, or could be applicable to the problematique posed by this novel target of regulation. This thesis furthermore makes novel arguments in how the legal-regulatory mind-set could potentially approach the topic by drawing from the most relevant sources discussing the ICT regulation.

2 REGULATING THE BLOCKCHAIN

2.1 WHERE ARE THE BLOCKCHAINS AND WHY IS EVERYONE TALKING ABOUT THEM?

Blockchains are popular because they promise a more efficient and more reliable system of things. While it is true that features of the blockchains offer solutions to various problems, as a rule of thumb, all technological developments have a multifaceted relationship with their utilities. Such a complex relationship is defined by the law and politics.[37] An innovative application without an eligible legal environment to support it yields little or no utility to the society. [38]

Blockchain enables a 'decentralised' (see figure 1[39]) way to store data and manage information. In its simplest definition, it is a file, a ledger, which is stored in every computer's node in the network. As a starter for the legal audience, blockchains are *not* the same as cryptocurrencies. Cryptocurrencies, like Bitcoin, are one in many applications *on* the blockchains. There are many use cases of blockchains and yet many more to come in future. Among many other applications are the Ethereum smart contracts, Factom supply-chain management and blockchain-as-a-service (BaaS) products such as offered by Microsoft Azure.[40]

[37] For instance, see *Fairfield* 2015 p. 829: "Technologies are created by narratives, and they are regulated by narratives." See also *Birnhack – Toch – Hadar* 2014 p. 26.
Also for a more general exposition of power dynamics and law, see *Foucault* 1977.
[38] *Swan*'s presentation accessed at:
https://www.slideshare.net/lablogga/blockchain-singularities-65443340
[39] It was a critical pursuit to find an alternative to the traditional, centralized network structure during the Cold War period. A single attack on the center of one of these networks could render the entire system useless thus compromising the national security greatly. As part of that pursuit, Baran proposed "distributing" the critical switching and control equipment around the network. This way, if one part of the network was damaged, the undamaged sections would continue working properly. This organizational redundancy is, today, a primary reason why companies choose to decentralize IT.
[40] For a comprehensive use case illustration see: https://gomedici.com/30-non-financial-use-cases-of-blockchain-technology-infographic/

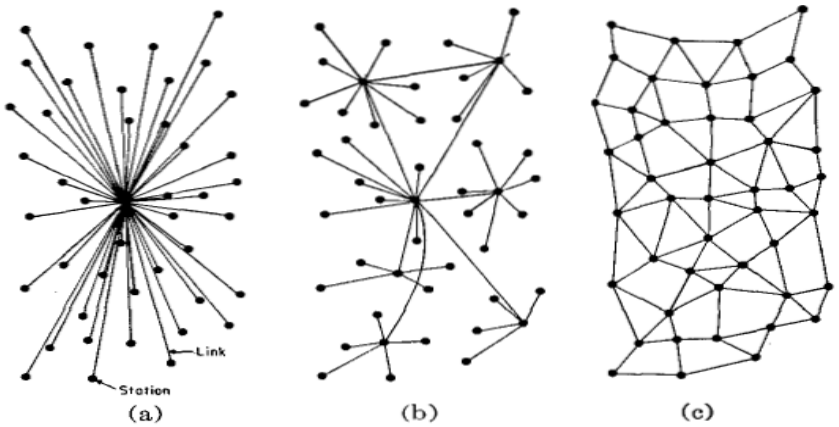


Fig. 1—(a) Centralized. (b) Decentralized. (c) Distributed networks.

source: *Baran* 1962 p.4

From the *personal data* management angle, there are also a few blockchain-based identity management applications. Sovrin, for example, is a global decentralised identity network which allows users and entities to create portable, digital identities controlled solely by themselves.[41]

In Sovrin, the user (identity holder) sits on a secure digital connection, between of the entity requesting information from the user (think of universities, employers, buyers/sellers) and the entity which can give this information to the user about the user (think of government organizations and registries). No personal data is placed on the chain; all data remains off chain. What is placed on a permissionless blockchain is decentralized identifiers (DIDs) offering high security. The personal data, however, remains off-chain therefore the application is untouched by some of the GDPR's specific obligations such as the right to erasure. Such application is conceivably popular with personally sensitive data such as health data or financial data. Maintaining data off-chain may or may not be a solution for every conceivable blockchain application; for this reason, this research handles the landscape of personal data vis-à-vis the blockchains with all possibilities on mind.

[41] Sovrin White Paper

2.2 STRUCTURE OF THE BLOCKCHAINS

Blockchain works by the power of the computers connected to its network. Different to other distributed systems, the data integrity is ensured by a trust that over 51% of those computers connected to the network are not going to attack the data integrity.

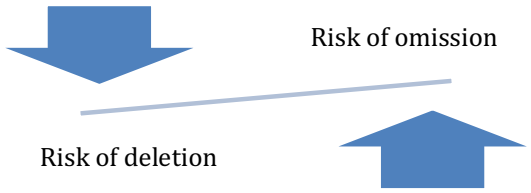
The core issue of the blockchains in relation to the GDPR stems from the diverse architectures of the blockchains. Consequently, there is not one kind of a blockchain; there are a few. The diversity of blockchain architectures is central to the analysis under the GDPR because they have different *features* giving rise to different pictures in relation to the notion of the "controller" and transparency. To be more specific, blockchain applications can track assets both on-chain and off-chain. This difference is reflected to what features they may offer to the users. For example, the property titles, intellectual property, shares, valuable papers (e.g. bills of lading), debts, medical equipment, pharmaceuticals and commercial aircrafts may be tracked off-chain. What this means is that those assets have a legal existence in the tangible reality outside of the blockchains and will continue to exist irrespective of whether the ledger that proves their existence is destroyed. This is not the case for the on-chain assets, however. On-chain assets exist only on the ledger for as long as the ledger exists e.g. Bitcoin.[42]

While a simple database is prone to *replication errors* and *delays*, blockchains are not.[43] A model to fully understand the blockchains vis-a-vis a simple ledger can be borrowed from Robert Sams' presentation.[44] A simple database is prone to *replication errors* and *delays.*[45] Another model can be borrowed from Robert Sams' presentation.[46] *The three ills of centralised ledgers* which the blockchains are capable of remedying are:

[42] *Reed and others 2017*
[43] *Swanson* 2015 p. 24.
[44] *Sams* 2015 with explanations supported from Bitfury White Paper p.6
[45] *Swanson* 2015 p. 24.
[46] *Sams* 2015.

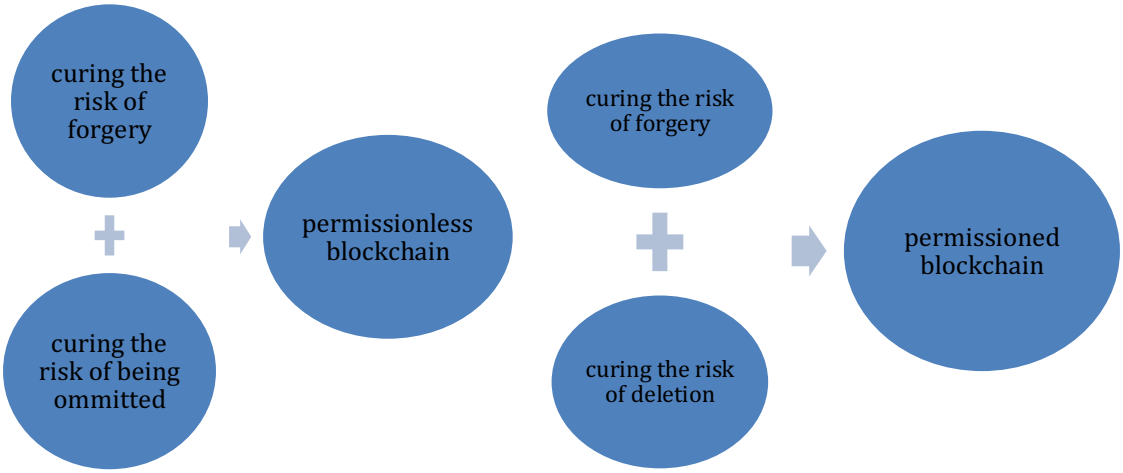
1. Sin of Commission - forgery of transaction
2. Sin of Omission- censorship of transaction
3. Sin of Deletion - reversal of transaction

As both Sams and Swanson argue[47], there is no one size fits all solution with a blockchain application. While all blockchain applications are capable of remedying ill #1 by ensuring the security of data, the ability to remedy ill #2 and #3 are mutually exclusive. In other words, a design can either cure #1 and #2 or #1 and #3 but not all three. Obviously, priorities will depend on use cases; thus, designing an application on blockchains necessitates a sound judgment on how much it makes sense to use a particular design for a particular problem in a particular industry.



In each case, the designer will settle on a trade-off among different features of the blockchain based on the service it intends to provide. The degree to which any of those three features could be achieved optimally largely depends on whether the blockchains are "permissioned" or "permissionless".[48]

[47] *Sams* 2015 and *Swanson* 2015 pp. 24– 25.
[48] *Swanson* 2015 p. 25.



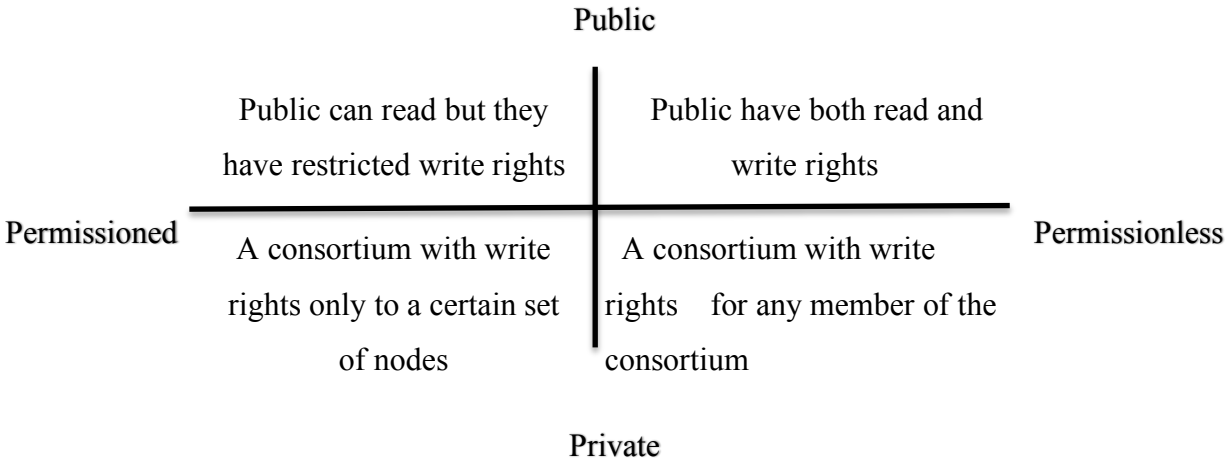
2.2.1 WHY DO WE USE THE CONCEPT OF "PERMISSION"?

The openness of a blockchain is on a continuum: on the one extreme, there is the permissionless blockchains (e.g. Bitcoin blockchain) and on the other extreme, there is the fully permissioned blockchains.[49] In the centre of the continuum, there are the *consortium blockchains* which have the attributes of both extremes.

What makes a blockchain permissioned or permissionless is their *consensus mechanism*. In the permissionless blockchains, each node in the network participates in *validating* the transaction and must together reach a *consensus* on such validity.[50] All of those nodes have the right to read and write a transaction. In the permissioned blockchains, only *a selected group* of nodes do this job and only they have *the right to write*. In other words, no one other than the selected group have the permission to write a transaction. The consortium blockchains, as the name suggests, have a consortium of e.g. financial institutions and a certain number of them must sign in order

[49] This distinction in the community is known as 'public versus private' blockchain.
[50] Transactions are validated by way of consensus because the database is distributed across a peer-to-peer network and operates without a central authority. A consensus is attained by way of the "mining" process and the mining process is done by engaging in mathematically complex, electrically resource-heavy computational equations that yield an "expensive" reward. Those rewards are called the mines and they are the equivalent of Bitcoin. Mining then creates the "proof of work" which is a piece of data that is capable of being verified by others. A valid transaction on the blockchain must possess a proof of work that consensus was achieved. Without the help of mining a reward e.g. coins as an incentive, the distributed computational power and time needed to create the proof of work would be difficult to ensure.

to validate a block. The consortium thus has the write rights. R3 is one famous consortium as such.[51] The consortium thus has the 'write rights'. The read rights may or may not be public, that is to say, it *is* possible to design the consortium blockchains to allow everyone to read the blocks as in the permissionless blockchains.



In reference to the blockchains tracking off-chain assets and blockchains tracking on-chain assets, the permissioned and permissionless blockchains offer different levels of solutions. Swanson explains that because of the possibility of reversibility in permissionless blockchains, off-chain assets are best tracked on permissioned blockchains.[52] Management of land registry cannot afford the risk of *reversibility* for example. Similarly, international trade finance, global capital markets or land registry are able to work due to the norm of *finality of a settlement*. This norm is enshrined in the EU Directive on the finality of settlements. For this reason, permissioned blockchains are the preferred kind of blockchains in the finance industry, land registry and potentially in tax administration. They are capable of curing the two ills of the centralized ledgers (forgery and reversibility) with more efficiency despite remaining potentially prone to "censorship".[53] Moreover, permissioned blockchains, as explained above, validate the transactions by using a selected group of validators. Although in a highly secured

[51] See https://www.r3.com/
[52] *Swanson* 2015 pp. 21 and 25
[53] See: https://blog.ethereum.org/2015/08/07/on-public-and-private-blockchains/

blockchain network the number of such selected validators will inevitably be high, they will remain *identifiable*, which is not possible with the permissionless blockchains. Legal enforceability of contractual or property rights depend on the *identifiability* of not only the claimant but also the debtor. Whatever one's thoughts and feelings may be in relation to the political ideologies underlying our economic systems, *identifiability* of persons and things is a basic feature of a system that ensures any legal course of action for the protection of citizens' private properties. Without identifiability, there is no accountability. Without accountability, there is no legal enforceability. Without the legal enforceability, the *value* of a contract or deal is significantly diminished. Why would anyone wish to build and use a financial product that is not capable of being legally enforced in the "real world"?[54]

It goes without saying that there are implications of being vulnerable to censorship and of entrusting the validation power to a selected group of nodes.[55] The libertarian narrative is that as long as the decisions are made by a group of people, the structures will inevitably centralize power.[56] However, this research finds that the discussion on *centralization of power* is much more nuanced than the *technical centralization* of nodes.[57] Sector- specific solutions to sector-specific inefficiencies *may* be placed in a broader context of problems related to the existing economic structures which are in turn the products of ideologies. An in-depth discussion on this topic in included in the final chapter of this thesis. From the data protection and privacy angle, different designs of blockchains application matter in how the responsibilities and obligations may be assigned.

---

[54] *Swanson* 2015 p. 22– 25.
[55] On the possibility of increased surveillance and centralization of the blockchain: see *Wright – De Filippi* 2015 p. 53.
[56] Some examples: https://www.theatlantic.com/technology/archive/2017/05/blockchain-of-command/528543/ https://halshs.archives-ouvertes.fr/halshs-01524440/file/17020.pdf http://hackeducation.com/2016/04/14/blockchain-ideology
[57] *Swanson* 2015 p. 28.
On the point of concentration of power by way of mining pools: https://www.buybitcoinworldwide.com/mining/pools/on a study on the Bitcoin mining reward systems see *Rosenfeld* 2011.

2.2.2 WHOSE PRIVACY ARE WE TALKING ABOUT?

It may seem obvious to some that the question is the privacy of the users-citizens participating in a blockchain network in the context of the GDPR. However, in a completely opposite direction, the GDPR is also indirectly concerned about the privacy of the *nodes* that validate the transaction. Paradoxically, the more the privacy of users need protection, the less privacy the validators ought to have and the more identifiable those nodes must be. This is where the anti-money laundering directive and know your customer ("KYC") requirements diverge with the GDPR. As of writing this article, the European Commission has agreed to revise the current anti money-laundering directive to require the cryptocurrency exchanges to identify its *users*.[58] KYC requirements necessitate full disclosure of things and people; GDPR requires a point of responsibility handling the data.[59] Needless to say that the relationship between the GDPR and KYC vis-à-vis blockchains may seem only relevant to the use cases of the blockchains to the extent the cryptocurrencies or tokens are necessary.[60] But it must be remembered that that the identities of the parties transacting are always possible to trace *unless* specific mixing techniques are applied.[61] Monero project, for instance, uses ring signatures that mix the spender's address with a group of others, making it harder to establish a link between transactions.[62] Pinpointing the persons responsible for taking care of data protection and privacy of real persons on a permissionless blockchain, however, is not at all a realistic goal. A very large number of nodes are connected to the network and every node on a permissionless blockchain keeps a replica of

[58] https://www.reuters.com/article/uk-eu-moneylaundering/eu-agrees-clampdown-on-bitcoin-platforms-to-tackle-money-laundering-idUSKBN1E928M

[59] IBM has partnered with Crédit Mutuel Arkéa group and created a blockchain pilot to *centralize* KYC information. -https://www.ibm.com/blockchain/use-cases/ It may seem odd to employ a *decentralized* platform to create a centralized system; however, this use example is certainly a good one showing how highly personal information can be placed on the blockchains and its relationship to the KYC rules.

[60] The relationship between the GDPR and the KYC rules vis-à-vis the blockchain applications is outside of this article and merits another research.

[61] *Wright – De Filippi* 2015 p. 21*, De Filippi* 2016 pp. 10 and 14, *Reid – Harrigan* 2013.

[62] http://monero.org/

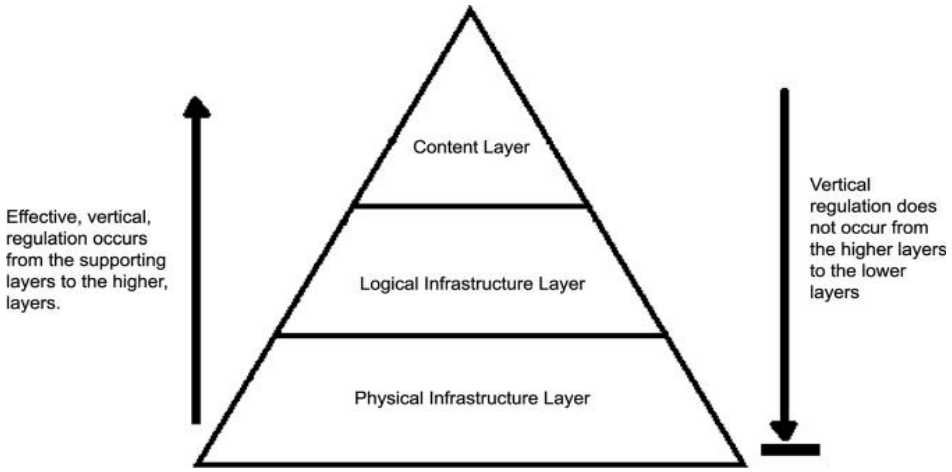
the ledger.[63] What is more, they may or may not be the same node working on validating the transactions each time, and they may or may not keep the ledger in their computer once they complete their role in the network.

2.3 REGULABILITY OF THE BLOCKCHAINS

Internet was originally designed for the machines to communicate with each other.[64] It was not architected to enable *trusted* interactions which we find growingly needed today while using the internet. This was not an inevitable consequence of the nature of technology; this was an intentional decision to have a network perform a wide range of functions.[65] Even though some scholars hold the view that analysing the issue of the impossibility of externally regulating the cyberspace is now an archaic attempt emulating the Cyber-libertarian narrative of the previous decades[66], the current, popular discourse around the blockchains motivate us to revisit the question of *regulation of internet.* As seen in *Table 1*, blockchains are protocols that run on the internet; they are *not* the next internet. Lessig identifies three questions for the state to answer before it can step into an act of regulation: who, what and where.[67] *Who* is the primary problem faced by the regulators in dealing with the blockchains. Internet does not have a way of authenticating who its users say they are. In that respect, the architecture of the blockchains present both a challenge and an opportunity in service of the regulators: the necessary authentication could be procured by the self-sovereign blockchains.[68] The regulatory problems in relation to the internet are carried over to the applications that run on that medium. Due to the architecture of the internet which was *designed*, *authentication* was a problem long before the blockchain became popular. Blockchains are thus regulated by not only their own peculiar

[63] *Swan* 2015, *Narayanan– Bonneau – Felten – Miller – Goldfeder* 2016, *De Filippi* 2016 and *Wright – De Filippi* 2015.
[64] *Lessig* 2006 pp. 38, 43- 44
[65] *Lessig* 2008 p. 44 citing *Jerome H. Saltzer* et al., "End-to-End Arguments in System Design," in Integrated Broadband Networks, edited by *Amit Bhargava* (Norwood, Mass.: Artech House, 1991), pp. 30–41
[66] *Murray* 2008 p. 295
[67] *Lessig* 2008 p.39
[68] *Lessig* 2006 p.45. Lessig actually gives the example of the single sign-on services ("SSO") as the identity authentication layer which has a widespread use nowadays with e.g. Facebook log in details. A prominent example is Microsoft Passport. This thesis, however, uses his analysis on the blockchain example.

architecture; but also, the architecture of the internet.[69] The internet, in theory, *could have been* designed to require personal identification in order to be accessed and used. But it was not designed that way.[70] Instead, it was designed to require no more than the packets of data to be labeled with their destination address.[71] The consequence of that design is a certain liberty with which applications using protocols including the ones with which a personal identification is not required are made. Blockchains are one such protocols. It is therefore possible to posit that the blockchains are enabled and *regulated* by the architecture of the internet.



[69] *Reidenberg* 1998 finds that the rules which are imposed by the technology and communication networks constitute a 'Lex Informatica' and that the internet is not *unregulable*; instead regulated by its own architecture. Also see Lessig 2008 p. 34: "Regulability is a function of design".

[70] *Murray* 2007 p. 73 "Internet's open architecture makes it very hard to create regulatory structures at either the internetworking (or TCP/IP) layer or the content (or presentation HTTP) layer. These open designs help ensure that at the heart of the modern network retains Bob Kahn's rule that, 'there would be no global control at the operation level.'

[71] Those protocols are collectively referred to as the "TCP/IP" for the exchange of data packets between machines on the internet. That protocol as it currently is does not have a technology for identifying the content carried in the data packets and report it. In other words, internet's original design is silent on the '*what'* element of regulability. Applications (softwares) layered on top of the internet such as "iProtectYou", however, provide the sort of content control which parents want in order to filter out harmful or obscene content. Similar filters could be implemented by employers to keep their employees from using social media during work hours. Softwares are not the only ways which implement filters successfully; filters can also be implemented on a network infrastructure (proxy servers, DNS servers or firewalls). Such filtering, however, would not be successful if the data in the packets are encrypted. There are other ways to work-around the filtering which are referred to as "internet censorship circumvention". Another method that which addressed the third pillar of the regulability, the *where* element, is the geographical mapping of the IP addresses. Originally, the IP addresses did not have a geographical information. However, commercial interest (*Cyril Houri*) has made it develop a technology for mapping the IP addresses. Nevertheless, even if such attempts to regulate the *what* pillar may count for something, the *who* element in service of regulation of internet remains. See *Lessig* 2006 p. 58-59.

Table 2 Benkler's model[72]

Once the internet is understood as the designed medium it is, the narrative around the problems start to change. Fahey explains that the cyber-regulation seems "fragmentary, multi-sourced and ostensibly unfocused" for two reasons: 1) it aims to also regulate the new and emerging technologies, and 2) it inherently necessitates *multi-level* risk regulation approach drawing from both international and supranational components as well as the local enforcement.[73] This research, too, finds that an effort to regulate cyber-space is too broad, unspecified and futile without nuanced and granular targets within the cyber-space to regulate. What does the cyber-regulation exactly want to regulate? As can be seen from Benkler's model, which broadly indicates that the physical layer is made up of cables and the logical layer is made up of the codes, the effective regulation can only take place at the lower layers of the internet.[74] In other words, regulating the technological applications closer to the top layer would yield much less effective regulation than regulating the architecture of the internet and its layers. Murray suggests that weak point within the network need to be identified in order to achieve any regulatory control in the architecture of the cyberspace.[75] He calls them "pinch points". Pinch

[72] *Murray* 2008 p. 299 citing from *Benkler''s regulatory model in Benkler* 2000 p. 568
[73] *Fahey* 2014 p.46-47
[74] *Benkler* argues that the physical infrastructure of the internet - nowadays the fibre optic cables- does not economically pose a barrier to entry hence it is possible to conceive the internet as a distributed infrastructure. The logical infrastructure and content layers, however, entail making of a wide range of choices which are relevant to how the power is structured in the ICT, more specifically the media. A prominent example to such choices are related to the intellectual property law. Protection of copyright at the logical layer leads to the cultural commons to be less and less available as a resource for personal experience, public discourse and creativity (see *Benkler* p. 568-577). It is nevertheless important to note that the 'internet backbone' is owned by a handful of companies in the world (please see the Wikipedia page for a comprehensive list: https://en.wikipedia.org/wiki/Tier_1_network). There are obvious competition law implications of this fact which are beyond the topic of this thesis. However, strictly in the context of 'concentration of power' in the ICT vis-a-vis the Benkler model, it is worth taking stock of the necessity of an in-depth research in this area. As a started in that research could be: 'Who Controls the Internet' by Jack Goldsmith and Tim Wu, Oxford University.
Also see: http://icaruswept.com/2016/06/28/who-owns-the-internet/
Murray 2007 p.45
[75] *Murray* 2007 p.74

points, Murray says, are found at the *transition points* between layers.[76] He recommends a study of the *controllers* of each layer in Benkler's model (*Table 2*).[77]

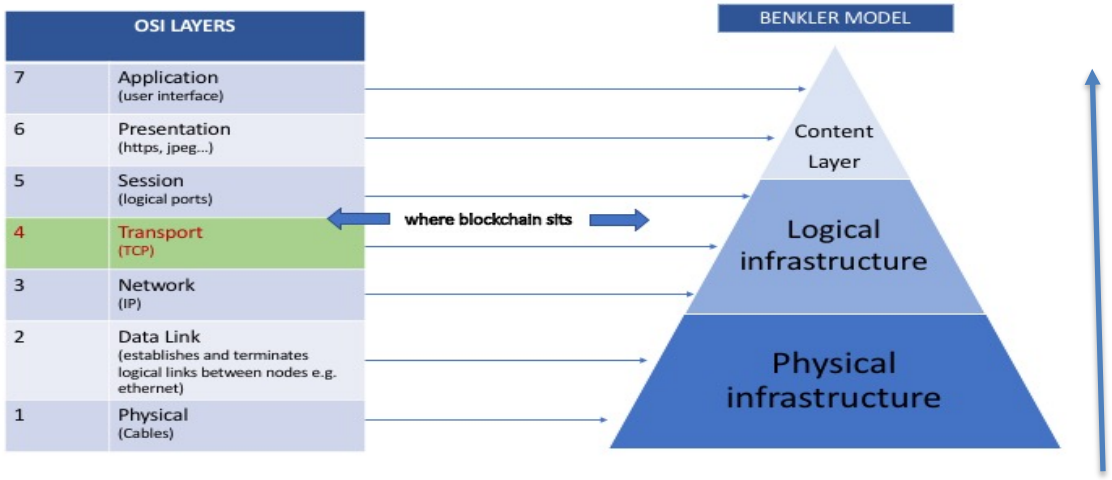


Table 3: Matching Benkler's model with the OSI layers to show the Murray's "pinch points"

On the basis of Benkler's and Murray's logics respectively, 1) regulation has to begin from the lowest layer possible upward in order to become effective, 2) regulation has to focus or target the specifics of process taking place at that point. For the blockchains, this occurs on top of the OSI Layer 4 -TCP. As explained above, Layer 4 – TCP is the protocol sending packets of data between computers. Intervention with the blockchain applications thus requires interference with the workings of the TCP. The only attempt at regulating the internet at this level has been "packet filtering" to censor content whereby certain keywords are detected and communication between computers is disabled. Naturally, the geography-dependence of such censorships mean that the more tech-savvy users turn to use VPN and TLS/SSL in order to escape the regulation. Without universally agreed and enforced TCP designs, attempts -if that were to happen- aimed at regulating the TCP layer within the EU would also fail.[78]

[76] *Ibid*
[77] For the sake of completeness, it is important to note that network stratification can be modelled in other ways, too: The most famous ones are the OSI seven-layer model, and Tim-Berners-Lee's four-layer model. Please see below *Table 3* for an example of the OSI model.
[78] Please see *Solum – Chung 2004* for a comprehensive research and legal analysis on this point.

Similarly, on the lower, physical level, it is theoretically possible to require the network providers to block actions by non-identified individuals although it is 'politically' almost impossible.[79] Another politically more realistic alternative is to regulate the intermediaries (ISPs) instead of the layers. In the EU, there are two precedents of regulating the IPSs: the e-Commerce Directive and the GDPR by virtue of Article 17 the right to erasure.[80] Within the context of the data protection and privacy regime, assignment of duties and responsibilities to a data controller is the functional equivalent of the same mind-set of easier regulation of the intermediaries rather than imposing identification requirement to access the internet. The world has moved from a place where programming was a marginal job within the hands of a select libertarian heroes. Market regulates the coders' choices by offering them the option to work for high-paying companies. In turn, the high-paying companies (or indeed any company which wishes to stay in the competition) fits the bill issued by the government regulations. Is it possible to earn money from blockchain applications outside the government purview? It is. Bitcoin and Ethereum are excellent examples of that. But at a *social* cost.

Lacking the third pillar for the regulability, namely the *who* pillar, internet opens up a variety of doors to liberally create and innovate.[81] On the flip side, however, the immediate legal consequence of this is the lack of accountability. The relationship between regulation and accountability is a symbiotic one.[82] Architecture of the permissionless blockchains (as enabled by the architecture of the internet) remove the accountability of the users. Moreover, by way of making tracking *almost* impossible or otherwise very costly by using different technologies such as the Tor project,[83] the users of permissionless blockchain applications such as Bitcoin can hide their footsteps in illegal trade activity. In that sense, not only the *who* pillar but also the *what* and *where* pillars can be rather evasive if the users want them to be on a permissionless

[79] Murray is adamant that if any or all of the controllers at the physical infrastructure level attempt to take over it, the market would find its way around this anomaly "in the same way the network routes around damaged nodes" in *Murray* 2007 p. 85.
[80] See Chapter 3 below
[81] *Lessig* 2006 p. 111-112
[82] *Black* 2001 p.143
[83] See https://www.torproject.org/

blockchain. The question is, then, *how much* regulatory intervention is justified to make it much harder if not impossible for those pillars, in particular the vital *who* pillar, to evade oversight.

3 EUROPEAN DATA PROTECTION LAW, CYBER-SECURITY LAW AND THE BLOCKCHAINS

3.1 APPLICABILITY OF THE GDPR TO BLOCKCHAIN

3.1.2 WHO IS THE DATA CONTROLLER IN A MULTI-NODE APPLICATION?

Under the GDPR, controlling the data means *making a decision* about why and how a particular data processing activity takes place.[84] Article 29 Working Party ("WP29") defines such *determining* capacity to be the preliminary element in assigning the role of the data controller. Even though such capacity may be conferred by way of law, it would usually be based on a factual analysis of the circumstances in each case and does not depend on a designation of any party as the data controller. The concept of controller is thus a functional concept based on the *factual influence* which may require an in-depth and lengthy investigation. [85]

The GDPR presumes where personal data is processed, there *is* either a processor or controller, or both. In other words, the new Regulation is not designed to serve a world where data is processed without anyone unidentifiable processing it.[86]

Permissionless blockchains are *decentralised* as per their architecture. It is not clear if the reference to function can be interpreted to include the meaning of architecture (as in "decentralised as per their architecture"). In any case, however, the GDPR presumes *technology neutrality* in its application which, by definition, can reasonably expected to disregard the significant architectural differences in data processing.[87]

As explained above, the lack of an identifiable entity or entities in the position of 'control' on the permissionless blockchain applications is a feature of this technology. Naturally, this feature

[84] WP 2010 Opinion 169 p. 8.
[85] WP 2006 Opinion 128.
[86] There is only one reference to a "decentralised system" in the Definitions Article. Accordingly, the GDPR defines the "filing system" to be: "any structures set of personal data which are accessible according to specific criteria, whether centralised, *decentralised* or dispersed *on a functional or geographical basis.*
[87] Recital 15 of the GDPR: "the protection of natural persons should be technologically neutral and should not depend on the techniques used".

of the blockchains has provoked the question of whether this technology is subject to the GDPR at all. Berberich and Steiner propose that if the notion of data controller implies any actual control over the information, two outcomes are possible: either no node would qualify as a data controller within the meaning of the GDPR or every node where the copies of the distributed ledger exist.[88] They suggest that neither of those outcomes are meaningful and due to the inherent uncertainty of the regulators' approach on this; consequently, the entities that use the blockchain as the infrastructure run the risk of walking on thin ice. De Filippi suggest that "the responsibility of keeping data private merely shifts from the operator to the individual user" in multi-node systems.[89] Assigning the liability for one's own data to oneself is, by definition, consistent with the EU's consumer empowerment trend within the data protection and privacy.[90] In practice, however, this notion necessarily presumes well-informed users with sophisticated knowledge of available technologies to protect themselves from the potential harms which may find them in the wild depths of the World Wide Web. It is therefore open to debate if this characterisation of a consumer is compatible with the *average consumer* in the EU.[91] As mentioned above, this situation is likely to yield an accountability gap which is not ideal for the regulators or citizens, and which violates the accountability principles under the GDPR. For this reason, at least for the time being, it is more likely for the regulators to allocate at least some degree of liability to the entities using decentralized architectures in protection of data and privacy initially. Indeed, the EDPS specifically asks for an investigation of the privacy and security implications of blockchain due to the difficulty of determining the liability issues in the permissionless system, lawfulness principle of data protection and data subject rights'.[92] This is understandably so since the role of a data controller and data processor are crucial for implementing protective and preventive measures as well as in assigning liability for data protection and privacy within the EU.

[88] *Berberich – Steiner* 2016 p. 425.
[89] *De Filippi* 2016; p. 15
[90] Recital 7 explicitly states that the GDPR framework is based on the ideal of giving natural persons the control of their own personal data.
[91] *Mak* 2010.
[92] European Data Protection Supervisor Annual Report 2016 p. 111-112.

Even in industries where the irreversibility and immutability are the desired features and permissioned ledger is thus used, there may still be some room for legitimate privacy expectations of the users which may *prima facie* be supported by the GDPR. To illustrate this point, we can think about the personal data which will be entered into the land registry, tax office or financial settlement institutions. The system in those examples work on the basis that no entry can be edited or deleted without authorization. Subsequent entries can change the subsequent course of events (i.e. the transfer of ownership of property, entering third party interest on land or tax returns) however the current entries cannot be edited. So this system will inevitably pose two problems under the GDPR: *irreversibility* of personal data (a piece of data which is no longer relevant, which may subsequently have been changed by way of entering new transactions will still stay on the block) and *radical transparency*. Regulation of those two features will largely depend on whether the blockchains subject to regulation are permissionless or permissioned. In the applications based on permissionless blockchains, regulators and authorities will be challenged with the lack of an identifiable entity making decisions about how the data will be processed. From strictly the point of view of the legal norms, we find some precedent in the EU law for finding liability in a party *who does nothing but provide the space* for others to publish or store information. Those examples may potentially provide a framework for considering a solution to the lack of a central controller. For example, the intermediary service providers (ISPs) are regulated both by the e-Commerce Directive and the GDPR by virtue of Article 17 the right to erasure.[93] Under both instruments, the ISPs are ascribed a level of *accountability* regardless of their participating in deciding how the data will be processed. Using the e-Commerce Directive to justify a finding of liability in a service provider goes both ways, however. ISPs are *exempted* from liability under the e-Commerce Directive if the "*activity is of a mere technical, automatic and passive nature*".[94] Arguably, *the mere technical, automatic*

[93] Article 17 "The right to erasure" is the embodiment of the seminal Google v Spain case where Google is the ISP.

[94] Recital 42 of the e-Commerce Directive: "ISPs are exempted from liability under the e-Commerce Directive if the: "information society service provider is limited to the technical process of operating and giving access to a communication network over which information made available by third parties is transmitted or temporarily stored, for the sole purpose of making the transmission more efficient; this activity is of a mere technical, automatic and passive nature, which implies that the information society service provider has neither knowledge of nor control over the information which is transmitted or stored".
Also see the Joined Cases C-236/08, C-237/08 and C-238/08 Google France and Google v. Louis Vuitton [2010].

and passive nature of data processing performed by the nodes on a blockchain may find some basis to be excluded from the liability if this logic is applied.

Similarly, the seminal ECJ decision on the right to erasure[95] which is now enshrined in the GDPR Article 17 compels the search engines to delete personal data upon request even if the original decision to put the data on the web may have been the data subject's own. The GDPR Article 17 is analysed in more detail below. Other examples for finding liability in the ISP include data processor obligations of the cloud computing providers.

3.1.3 JURISDICTION, LIABILITY AND ENFORCEMENT IMPLICATIONS

At regular intervals, society wakes up to discovering that information is fluid and it moves from one place to another easily and fast.[96] Enforcement of rules, however, depend on legally and socially distinct jurisdictions of countries. World Wide Web is a *deterritorialised* space and data is available, accessible and attainable simultaneously from anywhere in the world. Data protection regulations are people's effort to *territorilise* the web. However, data is produced and processed independently from the physical territory, rendering the physicality unnatural and ill-fitting to the reality of the web. Consequently, traditional territoriality no longer serves a useful tool in explaining and justifying jurisdiction on the web. In practice, more than one state's citizens could be affected by an action taken in that space triggering multiple jurisdictions. The GDPR sets its territorial scope under Article 3. Accordingly, it applies to data controllers and processors (1) that have an "establishment" in the EU and (2) where the data processing takes places "in the context of the activities of such an establishment" whether in the EU or not.[97]
Tasked with defining the concept of "establishment", Weltimmo Court built upon the Google v Spain Court's reasoning that the words 'in the context of the activities of an establishment' cannot be interpreted restrictively; and the Advocate General's opinion that the concept has to be flexible and depart from a formalistic approach whereby undertakings are established solely

[95] Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González
[96] *Bräutigam 2016 p. 143*
[97] GDPR Art. 3(1)

in the place where they are registered.[98] Accordingly, "the concept of 'establishment', within the meaning of Directive 95/46, extends to any real and effective activity — even a minimal one — exercised through stable arrangements.[99]

Weltimmo Court advised that a two-prong test be applied in order to ascertain whether this is the case: (i) is the activity of the controller in respect of that processing *mainly or entirely directed at that Member State*, and (ii) does that controller have a representative in that Member State, who is responsible for that activity and for representing the controller in the administrative and judicial proceedings relating to the processing of the data concerned?[100]

The Regulation also applies to "the processing of personal data of data subjects who are in the Union by a controller or processor not established in the Union, where the processing activities are related to:

1. the offering of goods or services, irrespective of whether a payment of the data subject is required, to such data subjects in the Union; or
2. the monitoring of their behaviour as far as their behaviour takes place within the Union."[101]

Art. 3 is not the only means by which the GDPR finds jurisdiction to regulate: General principles for transfer in Article 44 also exerts jurisdiction to a degree extraterritorially by regulating data transfers outside of the EU. The current (effective as of writing this thesis) Directive[102] required that the data controllers used equipment in establishing jurisdiction. The GDPR removes the difficulties faced by the national authorities and courts in their effort to make a broad reading of the word "equipment" to include e.g. use of cookies and other technologies without a physical device.

It is important to be reminded that the distinct roles of the concepts of data controller and data processor matter to the extent it helps assigning responsibility. Given that the issue of liability

[98] Weltimmo v NAIH (C-230/14) para. 29
[99] *Ibid* para. 31
[100] *Ibid* paras. 41 and 66
[101] Art. 3(2) of the GDPR
[102] Directive 95/46/EC

is intertwined with the issue of jurisdiction, this section begins with an in-depth analysis of the two actors.

3.1.4 LOCATION OF DATA

The data is not stored in one location: some of the information is stored on the computer running the software, some is stored in the public blockchain. The information on the computer is stored in the "wallet" file. Users store their transactions in which they are interested e.g. the transactions which have coins that belong to them or the newest blocks. Some users store the whole blockchain. As a result, there are multiple copies of the same ledger across the internet, at various computers. That ledger carries the record of every transaction ever made.

The GDPR is reasonably expected to apply to data controllers utilizing blockchains. This is going to happen if the data controller has an 'establishment' in the EU such as financial institutions and insurance companies, governments regardless of where the computers that maintain the blockchains are located (e.g. the USA). So as far as the GDPR is concerned, the data may be in a blockchain ledger across the internet, tied to a great number of computers located in from Japan to Russia and USA as well as the computers in the EU, and the applicability of the GDPR is going to remain as long as there is personal data related to the residents or citizens of the EU. It should, however, be noted that within the limited context of the blockchain, the GDPR will not be exerting *extraterritorial jurisdiction* in the strict sense of the term as long as there is one computer maintaining the blockchain ledger that is located in the EU. That is because regardless of where every other computer is located, there is *only one* ledger and the GDPR will be applying to that one ledger which also happens to be in the EU. [103]

3.2 DATA PROCESSING AND BLOCKCHAIN

The GDPR defines the "processing" as **"**any operation or set of operations which is *performed on* personal data or on sets of *personal data*, whether or not by automated means, such as

[103] For a discussion in relation to a similar architecture, namely the cloud computing, see: *Kuan et al 2011 and Kuan et al 2012*

collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction."[104] The broad array of activities listed, in particular, the "recording, storage, use, disclosure by making available of personal data" are descriptive of the data activity on the blockchain, rendering its scope of activity subject to compliance by the GDPR.

3.3 PERSONAL DATA AND THE BLOCKCHAIN

The GDPR enhances the definition of "personal data" found in the current Directive[105] and introduces three more elements which contribute to the "identifiability" to the natural person's personal data. An "identifiable natural person" is someone "who can be identified, directly or indirectly, in particular by reference to *an identifier* such as a name, an identification number, *location* data, an online identifier or to one or more factors specific to the physical, physiological, *genetic*, mental, economic, cultural or social identity of that natural person".[106] A piece of data that is *capable of identifying* a natural person is within the purview of GDPR. Moreover, the class of sensitive personal data, which require additional protections and restrictions, is expanded to include the genetic and biometric data.[107] It is only the anonymised data to which the GDPR does not apply:

> "Principles of data protection should (…) not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes."[108]

[104] GDPR Article 4(2)
[105] Directive 95/46/EC on the Protection of Individuals with regards to the processing of personal data and on the free movement of such data.
[106] GDPR Article 4.
[107] GDPR Article 9(1).
[108] Recital 26 Articles 4 and 5 of the GDPR.

The definition of anonymity is not provided by the GDPR, however, it can be logically deduced that anonymous data is the type of data which cannot render its data subject identifiable.[109] The level of identifiability is a question of degree and judgment: whether a piece of data renders a data subject identifiable will vary according to the sophisticated methods and circumstances. A method of judgment is thus naturally required in assessing whether the identifiability standard is met. There are two ways in assessing the identifiability: *absolute* approach and *relative* approach.

Under the *absolute* approach, all possibilities and chances in which the data controller would be able to identify the data subject individually are taken into account, irrespective of the expenses needed to be undertaken by the data controller in order to do so.[110] In other words, the data controller or processor becomes responsible under the GDPR "as long as *someone* can decrypt the data set" even though the data controller or processor themselves do not possess the key for decryption.[111]

Under the relative approach, however, the necessary effort or the expense required by the data controller in order to identify the data subject is taken into consideration.[112] In that sense, the legislation is applicable if the data controller is able to decrypt a certain data set or at least has *reasonable chances* of obtaining the decryption key.[113]

Recital 26.3 offers some ambiguous insight into the method of judgment:

> "(…) to determine whether a natural person is identifiable, account should be taken of *all the means reasonably likely to be used*, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly."

Although the Recitals are not per se legally binding, they will be used in interpreting an otherwise ambiguous provision.[114] By making a reference of what a third person may use, the

[109] *Spindler – Schmechel* 2016 p. 3.
[110] *Ibid* p. 165.
[111] *Ibid*
[112] *Ibid* p. 165 and *Lundevall-Unger – Tranvik* 2010.
[113] *Spindler – Schmechel* 2016 p. 165.
[114] *Klimas – Vaiciukaite* 2008 p. 33.

Recital article's wording seem to keep wide the possibilities of persons in the world. Borgesius argues that the GDPR demonstrates a tendency favouring the *absolute approach* because such a person could be *anybody* in the world.[115] Conversely, Spindler and Schmechel views the language of the Recital to imply a mix of both approaches due to the use of "the means reasonably likely".[116] In Esayas' view, the legislator's use of "the means reasonably likely" imparts the *relative* approach, thus where the identification risk is remote or highly theoretical, the data should not be considered personal.[117] The legislator's qualification of the objective factors to be taken into consideration such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments[118] could also be seen as pointing at an attempt to limit the broad and *absolute* elements of the GDPR's scope.[119] An important case guiding us here is the *Breyer case* where the Court of Justice ruled that a dynamic IP address is personal data in relation to a certain internet service provider.[120] Recognizing that an IP address alone cannot provide information about the identity of the person operating the device that is connected to a network, the Court qualified this ruling that "an IP address is only personal data where the internet service provider had the *legal means* which would enable it to identify the data subject".[121] This is interpreted by El Khoury as a *de facto recognition of a grey zone* by the Court of Justice where data can be personal and non-personal at the same time.[122] By not categorically declaring the IP addresses as personal data on the sheer possibility of identification, the Court seemed to have taken a balanced approach in resisting a ruling that would broaden the regulatory burden on data-processing entities thereby avoided an outcome which may have been disproportionate considering the actual risks to the privacy of data subjects. The Court's approach is also validated by Recital 4 of the GDPR: "The right to the protection of personal data is not an absolute right; it must be considered in relation to its function in society and be balanced against

[115] *Borgeisus* 2016 p. 9.
[116] *Spindler – Schmechel* 2016 p. 166.
[117] *Esayas* 2015 p. 6.
[118] Recital 26 Art. 4 and 5 GDPR.
[119] *Spindler – Schmechel* 2016 p. 167.
[120] *Patrick Breyer v Bundesrepublik Deutschland*
[121] *Ibid* para. 49.
[122] *El Khoury* 2017.

other fundamental rights, in accordance with the principle of proportionality".[123] In that respect, Court can be seen to have implicitly agreed that the binary notion of personal data is too simplistic and not particularly useful in the grand and complex scheme of data collection and information flows.[124]

Internet protocol addresses and the infrastructure of blockchain applications are clearly different; however, the case presents the legal reasoning which could be applied to the same class of personal data and identifiability dynamics under other technological applications. In the case, Mr Breyer sought to have an injunction against the Federal Republic of Germany to stop from registering and storing the IP addresses of the pages run by the government he accessed alongside the dates of such visits. The Court of Justice refers to the "legal channels" that exist "so that the online media services provider is able to contact the competent authority in order to obtain information from the internet service provider" with a view to combat cyber-attacks.[125] So in ruling that the IP addresses are personal data as long as there are legal means to obtain more information to make the IP addresses meaningful in identifying the data subject, the Court effectively defined the "legal means" to be "a possible channel not prohibited by law".

The WP 29 underlines that *identification* should not be conceived solely in relation to "the possibility of retrieving a person's name and/or address, but also includes potential identifiability by singling out, linkability and inference."[126] European law understands "the data about data" to be personal data by virtue of Article 8 of the Charter of Fundamental Rights.[127] In other words, metadata is personal data.[128] A data subject can be identified without necessarily finding her/his name and address. Therefore, following the legal reasoning of the Court as to what amounts to personal data, we can conclude that data on blockchain may be fully encrypted and it may not be possible to link it to a data subject *per se*. However, the data it contains (meta-data) may *still turn out to be personal data* capable of identifying the data subject as long as:

[123] Recital 4 and Art. 2 GDPR.
[124] *El Khoury* 2017 p. 5.
[125] ECJ ruling "IP Address as Personal Data" para. 47.
[126] WP 2014 Opinion 216.
[127] *Malone v. UK*
[128] WP 2014 Opinion 215.

(1) the means of access to the necessary information about the subject that makes the otherwise non-personal data personal is *not* prohibited by law, and

(2) the process by which such information is obtained is *not* particularly complex.

El Khoury appropriately points, however, that with the advanced and accessible technology, *the quality of the parameters* used in assessing the risk of identification (the 'legal means' test) may rapidly change. For example, the cloud computing technology makes it possible to have access to a wide variety of complex computing services whereby the risk of identification could be seen as both (1) achieved by legal means, and (2) using not so complex technique.[129] Online identifiers provided by the devices is yet another catch-all parameter that affects the risk assessment significantly.

Recital 30 of the GDPR states that:

> "(n)atural persons may be associated with *online identifiers* provided by their devices, applications, tools and protocols, such as internet protocol addresses, cookie identifiers or other identifiers such as radio frequency identification tags. This may leave traces which, in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them."

By virtue of the logic applied in the IP addresses as well as other online identifiers listed in Recital 30 of the GDPR, a lot of the data produced by Internet of Things technologies will become personal data even if they were 'attribute data' e.g. sheer machine data.[130]

In the permissionless blockchains, there are two *layers* of data: content data and protocol data.[131] Decentralized architectures are capable of privacy at the content level by way of encryption. However, the secondary data that is the metadata remains publicly disclosed as an essential feature of such architectures.[132] Protocol transparency refers to the disclosure of metadata or other kinds of administrative information. In decentralized platforms, protocol transparency is

[129] *El Khoury* 2017 p.7.
[130] *Spindler – Schmechel* 2016 p. 168
[131] *De Filippi* 2016
[132] *Ibid* pp. 1, 3, 10.

an essential feature even if content transparency may be forgone. In Bitcoin blockchain, for example, the data is inside the Bitcoin on the Blockchain, coupled with any other data which the user may wish to store on the block.

3.4 RISK OF IDENTIFICATION

As described above, the problem with encryption is that it always leaves the meta-data accessible. Metadata poses the risk of unintended disclosure of personal data e.g. of the form of the data controller entity, time and date. In some cases such as health information, this may be enough to single out individual persons in combination with other information such as those provided by camera logs.[133] This research has found that multiple researchers agree on the difficulty in maintaining anonymity where network data on *user behaviour* is available.[134] Sweeney shows that re-identification does not even require sophisticated expertise in her experiment on matching publicly available patient information with the information available through local newspapers.[135] Sweeney, also in another experiment, demonstrates that an anonymous medical database could be combined with a voters' list to extract the health record of the governor of Massachusetts.[136] Still, one important example is from Reid – Harrigan who ran a demonstration by way of constructing two network structures: (1) the transaction network and (2) the user network using the publicly available transaction history.[137] Their demonstration includes integrating off-network information to show types of information leakage that can contribute to the de-anonymization of the system's users. With use of sufficient associations and combinations of those with the network structures they have created, they attained a serious threat to "anonymity" of Bitcoin. Off-network information may include e-mail addresses, shipping addresses, credit card or bank account details, IP addresses which may be accessible through entities which accept Bitcoin as payment. In their illustration to make this point, Reid – Harrigan found that it is possible to associate the IP addresses with the Bitcoin recipient's

[133] *Enisa* 2015 Data Protection and Big Data p. 44.
[134] *Reid – Harrigan* 2013, *Narayanan – Shmatikov* 2009*, Wei – Li – Zou – Wu* 2014 and *Borgesius* 2016.
[135] *Sweeney* 2013 p. 10.
[136] *Sweeney* 2002.
[137] *Reid – Harrigan* 2013.

public keys which were also capable of revealing the IP addresses related to previously used transactions and obtain geolocation of the users. The conclusion is that using outside information, it is possible to associate public-keys with each other and that strong anonymity is not a feature of the Bitcoin.[138] What is understood is that Bitcoin is not anonymous to the same level as cash is since clever algorithms may be able to link create patterns and link outside information to those patterns to attain unique, personally identifying information about a data subject.[139]

3.5 ANONYMITY AND PSEUDONYMITY

The distinction between when a piece of data is considered pseudonymous or anonymous is vital in establishing the applicability of the GDPR. Most data reduced to a seemingly non-personal form may be conceived as pseudonymous if it is using data to "single out a person regardless of whether a name can be tied to the data,"[140] attracting the GDPR rights and obligations in relation to such data. Thus, the degree of the ability *to single someone out* may be determinant in classifying a piece of data anonymous or pseudonymous.[141]

Pfitzmann and Hansen define the anonymity of a subject means that the subject is not identifiable within a set of subjects, the anonymity set. Not identifiable means the subject is "not distinguishable from the other subjects within the anonymity set" and that the anonymity set refers to the set of what the authors humorously called as the "usual suspects".[142] It is, however, difficult to know exactly in what situations the usual suspects will behave in a similar fashion. De Montjoye highlights that if the *patterns of* mobility traces of humans is unique, presumably anonymous datasets may yield personal data related to data subjects.[143] This means that in cases where the "suspects" are not behaving in a *usual* or uniform way, metadata about them can easily lead to personal information about them even in a sparse, large-scale and coarse mobility

[138] *Ibid* 2013 p. 26.
[139] *Narayanan – Bonneau – Felten – Miller – Goldfeder* 2016, foreword p. 14.
[140] WP 2007 Opinion 136.
[141] *Borgesius* 2016.
[142] *Pfitzmann – Hansen* 2010 p. 9.
[143] *De Montjoye* 2015 pp. 19-23.

dataset.[144] Similarly, knowledge of time-stamped transaction, the shop at which such transaction took place and the price of the transaction yield the data subjects using credit cards as reidentifiable as mobile phone.[145] Such metadata could include information about the transaction amount, the assets being transferred and the time of transaction, which are unique enough to narrow down the class of "usual suspects" for the purpose of accurate identification. Indeed, the large amount of data associated with a public key is publicly available and offers as much information as the identity of the entities transacting with the original key holder.

The GDPR does not define the anonymous data and we have to refer to the semi-official sources of the WP 29 opinions and Enisa reports for an authoritative legal guidance on this point. WP 29's understanding of anonymous data requires *irreversible anonymisation*. Irreversibility, in data processing sense, means a form of erasure rendering it *impossible* to process personal data. That kind of impossibility achieves the "effective anonymisation" as a solution preventing "all parties from *singling out an individual in a dataset*, from linking two records within a dataset and from inferring any information in such dataset".[146] In evaluating possible anonymization techniques, the WP 29 judges the methods against three types of risks: 1. Singling out, 2. Linkability, 3. Inference.[147] Interestingly and perhaps encouragingly, however, the WP 29 leaves room for a risk-based approach and recognizes the context-boundness of the risks of identification in subsequent paragraphs.[148] Indeed, the WP29 makes explicit reference to an unacceptable risk of identification of data subjects *without* specifying under what circumstances the risk of identification is considered to be acceptable, or what an acceptability means. This lack of clarity on the point of a risk threshold has attracted criticism given that the WP29 sets a high standard of near-zero probability for identification.[149] In a similar vein, Stalla-Bourdillon & Knight find such contradictory approaches in the same Opinion problematic.[150] However, while acknowledging the need for a clarification on the acceptability of levels of risk, this

[144] *Ibid* p. 27.
[145] *De Montjoye* 2015 pp. 32–40.
[146] WP 2014 Opinion 216 p. 9.
[147] WP 2014 Opinion 216 p. 10-11.
[148] WP 2014 Opinion 216 pp. 6- 10 and 23–25.
[149] *Esayas* 2015, *Emam – Avarez* 2015.
[150] *Stalla-Bourdillon – Knight* pp. 297–298.

research interprets the concept of anonymisation as used by the WP 29 as an *ideal* which data controllers and processors must aspire to attain and not an unrealistic zero-risk requirement. Emam - Avarez explain that zero-risk is *not* practically achievable, and argues that a more precise way to describe anonymous data would be that which "has very small risk of re-identification".[151]

In the light of this analysis, rethinking the above illustrations against the standards of the WP29 raises the question whether any piece of data can ever be *fully* and *irreversibly* anonymised. Stalla– Bourdillon- Knight suggest that the dynamic state of anonymised data must be accepted by the policy-makers.[152] None of the known anonymisation techniques can provide such an assurance while at the same time preserving the utility of the data sets.[153] Moreover, the risk of identification increases with the number of databases and possible correlations yielding "accretion problem" in data anonymization[154]. On permissionless blockchains, it is *not* possible to fully anonymise data irreversibly while simultaneously preserving the ability of the nodes to "understand" the transaction which they are required to verify in order to yield the consensus. The data on the blockchains are and will be pseudonymous within the meaning of the GDPR.

The GDPR provides that pseudonymity "means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject *without the use of additional information*, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person"[155] In the Opinion of the WP29, while pseudonymisation reduces the linkability of a dataset with the original identity of a data subject, a natural person is still likely to be identified indirectly.[156] Encryption, hash-function, keyed-hash function with stored key, deterministic encryption or keyed-hash function with deletion of the key and tokenization are listed as pseudonymisation techniques.[157]

---

[151] *Emam – Avarez* 2015 p. 76.
[152] *Stalla – Bourdillon – Knight* 2017.
[153] *Emam – Alvarez* 2015 p. 84.
[154] *Narayanan – Shmatikov* 2008; see also *Ohm* 2010 pp. 1701 and 1746.
[155] Article 4(5) of the GDPR.
[156] *Ibid* p. 21.
[157] *Ibid* p. 21.

The Opinion offers guidance in counteracting the common mistakes when using and managing the key. Those are:

1. Using the same key in different databases,
2. Using different keys for different users, and
3. Keeping the key alongside the data.[158]

It comes as no surprise that the WP 29's guidelines for ensuring the appropriate privacy safeguards mean having high-level security in managing the access control and authentication. Traditionally a computer security subject, *access control and authentication* thus become an essential topic in the GDPR compliance. "Access control and authentication" mean the recipient of information has the authority to receive that information.[159] In her remarkable work, Sweeney warns that while such protections "can safeguard against direct disclosures, they do not address disclosures based on inferences that can be drawn from released data."[160] Nonetheless, the GDPR compliance requires implementation of the state-of-art security measures and privacy-preserving techniques with a view to fortifying the privacy and data integrity.[161]

3.6 APPLICATION OF DATA SUBJECTS' RIGHTS UNDER THE GDPR AND BLOCKCHAINS

Blockchain applications which store on-chain personally identifiable data are at odds with two of the data subject's rights under the GDPR due to the immutable nature of decentralized platforms: The right to erasure[162] and the right to rectification[163] both of which are supplemented by Recital 65. Under its Article 17, the GDPR requires that individuals have a right to request the deletion or removal of personal data *whether there is no compelling reason for its continued processing*. Whilst the immutability of the data on blockchain is a key feature of the new technology, such feature stands at odds with the Regulation.

[158] WP 2014 Opinion 216 p. 21.
[159] *Sweeney* 2002 p. 5.
[160] *Ibid* p. 5 where she famously argues "Computer security is not privacy protection".
[161] Article 32 of the GDPR, Recitals 76–80.
[162] Article 17 of GDPR.
[163] Article 16 of GDPR.

Data controller has an obligation to process data subjects' privacy requests including a request to be removed from search results even if initially lawful processing of accurate data may, in the course of time, become incompatible with the directive where those data are no longer necessary in the light of the purposes for which they were collected or processed.[164] Even though the right to erasure is not an absolute right, individuals have a right to have personal data erased and to prevent processing in specific circumstances.[165] Under the GDPR, the individuals are also entitled to have personal data *rectified* if it is inaccurate or incomplete.[166] With the access and authorization techniques, it is possible to design systems in which users are the only person who have access to their own data. Arguably, then, if the user is the only person who gets to access her own data, there is little room for anyone else to feel concerned as she is her own data controller and she is capable of "disabling" access to anyone to whom she has allowed access. It is not possible to "erase" data from the block, however, it is possible to disable access (read and write) to the data on the block. The question is, then, whether the "disabling access of others" will amount to "erasure" at law. Applying the logic of functional equivalents, the regulators and authorities should have no problem recognizing the act of *disabling access* as *erasure* at law. Indeed, the UK DPA has previously provided guideline that a more realistic approach would be "putting information 'beyond use' and for data protection compliance issues to be 'suspended' provided certain safeguards are in place."[167] This necessitates a discussion on how the right to erasure can be enforced besides the obvious "delisting" requests from Google and how could the technology cater to the socio-legal expectations. A similar discussion was addressed to by O'Hara on the potential utility of the Semantic Web or the Linked Data Web.[168]

---

[164] Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González.

[165] Under Article 17(1) GDPR, such specific circumstances are listed:

1. Where the personal data is no longer necessary in relation to the purpose for which it was originally collected/processed,
2. When the individual withdraws consent,
3. When the individual objects to the processing and there is no overriding legitimate interest for continuing the processing,
4. The personal data was unlawfully processed (i.e. otherwise in breach of the GDPR),
5. The personal data has to be erased in order to comply with a legal obligation,
6. The personal data is processed in relation to the offer of information society services to a child.

[166] Article 16 GDPR.

[167] ICO Guideline 2014 on Deleting Personal Data p.4

[168] *O'Hara* 2012.

Similarly but more comprehensively, results of the research program PrimeLife[169] present detailed and technical analysis of possible privacy-enhancing techniques. In particular, the results published under the category of Identity Management protocols ("IdM") is remarkable for providing a sound basis for treating credential systems that give the users full control over their data.[170] Anonymous credential systems are found to provide privacy "*in the strongest possible sense: issuing and relying organisations cannot learn anything beyond what users choose to disclose when presenting their tokens, even if they collude and have unlimited resources to analyse protocol data*".[171] Moreover, the Commission promotes the eIDAS on its website as a solution to reducing *security* and *privacy* concerns of the citizens.[172] In a similar vein, the reports published by ENISA[173] present security and accountability controls to include *granular access control*.[174] It is therefore not unreasonable to expect the EU regulators and authorities to recognize such methods as fulfilling the GDPR obligations for the rights of individuals. Moreover, this approach is compatible with the consumer-empowering spirit of the GDPR. [175] The real interest of this part of the thesis's is in conveying the message that the solution does not have to come from blockchain's own architecture.

Although not exactly a perfectly fitting example to the above, Zyskind et al presented a possible design to use the blockchains as a way to secure personal data.[176] Their design uses a combination of two blockchains: one for an access-control and the other with an off-chain storage for data. The off-chain data storage is maintained by a permissioned blockchain and only the user has control over his/her data.[177] The on-chain data contains only "hashed pointers",

---

[169] PrimeLife Identity Management Report.
[170] Privacy and Identity Management for Europe: a project supported by the European Commission's 6th Framework Programme and the Swiss Federal Office for Education and Science, and
PrimeLife: Bringing sustainable privacy and identity management to future networks and services, a research project funded by the European Commission's 7th Framework Programme.
Please see PrimeLife Results of Privacy and Identity Management Research.
[171] PrimeLife Identity Management Report Chapter 4 p. 35.
[172] Electronic Identification and Trust Services (eIDAS) Regulation Q&A
[173] European Union Agency for Network and Information Security.
[174] ENISA report on privacy by design in big data p.42.
[175] Recital 7 explicitly states that the GDPR framework is based on the ideal of giving natural persons the control of their own personal data; also see "A New Consumer Empowering Agenda 2012" and the WP 2016 Opinion 242.
[176] *Zyskind – Nathan – Pentland* 2015. Another example is *Dorri– Kanhere – Jurdak– Gauravaram* 2017.
[177] *Zyskind – Nathan – Pentland* 2015 p. 181 and p.183.

which would be pseudonymous data under the GDPR. Zyskind and Nathan argue that such data is meaningless to an adversary and therefore does not assist in revealing the identity of the data subject. It is still conceivable that the regulators may ascribe at least some shared responsibility to the infrastructure provider simply because the user is receiving a service which *guarantees* to her that her data will solely be controlled by her.

Additionally, recognition of access authentication systems based on anonymous credential systems brings other obligations under the eIDAS as long as the user wishes to transact with public bodies in addition to the private ones. This is particularly important in designing permissioned blockchain solutions for land registry, tax administration and healthcare systems. At this junction, the GDPR, the eIDAS and the NISD together control the techno-legal design of solutions. Diversity of technical and legal approaches to the protection and management of electronic identities by the member states is a challenge in identifying a universal solution.[178] Thus, the challenges in an EU-wide harmonizing of the electronic IdM become relevant to the problems of a harmonized application of the GDPR across the EU in respect to the blockchains. Ensuring compliance of the use of one particular technology by employing another type of technology is precisely how technologies act as both the regulatory tools and as the target of regulations.[179]

Without such IdM solution enabling the users to control their own data, *the selected group of validators* will likely take on the responsibility of data processor in the permissioned blockchains, which would not be very practical and efficient. As a starter, such group of validators would need to specify the processing as 'lawful' within the meaning of Article 6 of the GDPR. A processing is lawful if the data subject has given her *consent* or entered into a *contract* (e.g. contract of use) with the data controller/processor. Other bases could be relied on for the legality of processing by the public institutions: *legal obligation* and/or *performance of a task carried out in the public interest or in the exercise of official authority*.[180] The four bases

[178] *De Andrade* 2012, p. 291.
[179] *Brownsword – Yeung* 2008.
[180] Art. 6(c) and Art.6(e) of the GDPR

have different implications for the data controller/processor. If the permissioned validators (potentially public institutions or institutions providing public service) rely on the legal obligation or the performance of a task carried out in the public interest or in the exercise of official authority, derogations from the right to erasure are available, relying on the same classification. In other words, the public institutions may have the opportunity to insist on processing despite a request for erasure on the basis of the legal obligation or the performance of a task carried out in the public interest or in the exercise of official authority, or for having legitimate interest in continued processing.[181]

If the permissioned blockchain validators rely on the consent of the data subject (user), then there is little scope to benefit from one of the derogations given by the GDPR if the user later on wishes to erase all his/her data from one of the blocks or wishes to rectify some of the data there. This is because a consent is easily revocable and there are no derogations which could be invoked to overcome a request of the delete or rectify data by the data subject if a consent is the legal basis of processing. However, if the legal basis of the processing is a contract, Berberich and Steiner offer another potential solution: retention of data is necessary in order to comply with a legal obligation being the contract.[182] The problem with this approach, however, would be the possibility of the data subject to rescind the contract on compelling reasons. Another solution offered by Berberich and Steiner is to rely on the "necessity" derogation of Article 17(i)(a). The architecture of the blockchains *require* or *mandate* immutable records.[183] In other words, it is the modus operandi of the blockchains that the data is not capable of being deleted or altered. Based on that qualification, it may be possible to conceive this core functioning principle of the blockchains as a 'continuing necessity' arising from a *legitimate interest* of the controller[184] with a view to trump respectively Article 17(i)(a) or (b) if invoked.[185] Finding a

[181] Respectively: Art. 17(3)(b), Art. 17(3)(c) and Recital 69 of the GDPR
[182] Recital 65 of the GDPR: "(...) the further retention of the personal data should be lawful where it is necessary, (...) for compliance with a legal obligation".
[183] *Berberich – Steiner* 2016 p. 426.
[184] Recital 69: "(...) It should be for the controller to demonstrate that its compelling legitimate interest overrides the interests or the fundamental rights and freedoms of the data subject (in order to trump the right of objection of a data subject for processing of his/her data).
[185] Article 17(1):"(...) the controller shall have the obligation to erase the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed".

legitimate interest is a balancing act between the policy objectives and fundamental rights, and it is more likely to succeed in defense of public entities rather than private companies. Finally, whether this application would be successful in the eyes of the Court is yet to be seen.

In the permissionless blockchains, on the other hand, the accountability gap may leave the infrastructure/ application providers in a difficult situation. In that case, the solution is likely to turn on the regulators and courts embracing a revised notion of *data controller* to include the data subject herself.

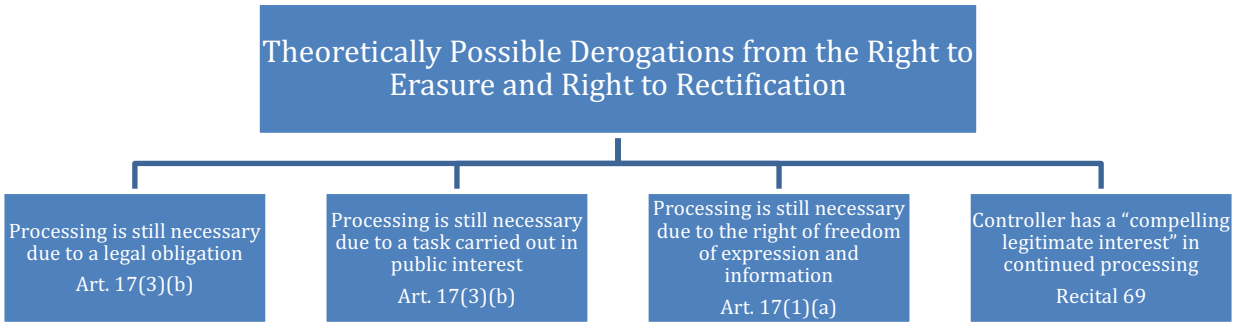


Table 4: *A mini mind-flow for lawyers working with the GDPR*

3.7 PRIVACY THREATS POSED BY THE BLOCKCHAIN

Koops argues that the relationship between the privacy risks and the concept of personal data is unclear.[186] Kuner, too, suggest that of the two pillars of data protection, privacy and security, harms from security breaches are generally well understood while the consensus on what constitutes harm in privacy is still being developed.[187] This research, however, has identified literature making it a sufficiently plausible premise that unchecked use of personal data can result in actual harm.[188] The most current data protection instrument, the GDPR, too, gives clear

[186] *Koops* 2013, also see Ohm 2010 p.1728 for a similar argument.
[187] *Kuner* et al 2015 p. 97
[188] For example, Drabiak analysed arguably the most sensitive and personal information available: the genomic sequence. She posits that individuals are most likely not aware of the range of *subsequent uses* for their genomic and personal information [*Drabiak* 2017]. Similarly, in the EU, regulation of potential risks e.g. discrimination associated with genetic information go as far back as 1997 and it follows that there is a reasonable anticipation that unregulated use of genetic information will likely cause harm in the form of discrimination [The Council of

insight into the nature of harms intended to be avoided: physical, material or moral damage with particular emphasis on "discrimination, identity theft or fraud, financial loss, damage to the reputation, loss of confidentiality of data protected by professional secrecy, unauthorized reversal of pseudonymization, or any other significant economic or social disadvantage."[189] Moreover, three high-risk examples are identified under Article 33 of the GDPR: (1) "systematic and extensive" automated profiling that "significantly affects" individuals, (2) large-scale processing of special categories of data, and (3) large-scale, "systematic monitoring of a publicly accessible area." It is not difficult to see that when used in their *permissionless* architecture, the blockchains are going to maintain metadata transparently and are likely to be a nest of valuable data mines about our lives which may be harmful to us if used with disregard to our protection. By analysing metadata and combining previously discrete data sets, 'Big Data' is able to create novel personal data.[190] Despite a constant feed of how our lives are going to be improved if there is less regulation on the Big Data, there is more than reasonable doubt that the society may also be harmed due to some unintended damage from unchecked use of personal data.[191]

Europe's Convention on Human Rights and Biomedicine (1997) contains, besides privacy provisions, a special provision prohibiting discrimination on grounds of 'genetic heritage' (Article 11)].

Along the same lines, *Crawford and Schultz* argue that "personal harms emerge from the inappropriate inclusion and predictive analysis of an individual's personal data without their knowledge or express consent". In what they call "predictive privacy harms", Crawford and Schultz describe the situations in which harm is inflicted on data subjects by way of predicting personal data about them which is not yet publicly available. [*Crawford and Schultz* 2014 p. 94-95 a remarkable example is taken from the New York Times article which revealed that a retail chain mined data to predict which female customers were pregnant. This activity then resulted in unauthorized disclosure of personal data to marketers: Charles Duhigg, Psst, You in Aisle 5, N.Y. TIMES, Feb. 19, 2012, § 6 (Magazine), at 30]. In what they call "predictive privacy harms", Crawford and Schultz describe the situations in which harm is inflicted on data subjects by way of predicting personal data about them which is not yet publicly available. Predictive privacy harms include i. *discriminatory practices* across industries including but not limited to the real estate and credit loan, ii. *health analytics and personalized medicine,* and iii. *predictive policing*. [*Crawford and Schultz* 2014 p. 104];

Gandy and Danna 2002 where they examine the ways in which data mining and the use of consumer profiles may exclude classes of consumers from full participation in the market-place, and may limit their access to information essential to their full participation as citizens in the public sphere; Other examples include *Hendriks* 2002, *Wright - Raab* 2014, and *Vranaki* 2017 p. 208-209 where she explains that "mass dataveillance can be restrictive as such categorization can at times maintain or produce social inequalities".

[189] Recital 60 of the GDPR

[190] *Crawford and Schultz* 2014 p. 94

[191] *Nissenbaum* 2017 p. 26: it is also worth noting that the most agreed upon benefits of big data usually flow from *smart regulation*. In other words, it is the upgrading of regulation vis-a-vis the big data that enables the individuals to reap the benefits.

As for the privacy harms implicated in the use of the blockchains, the nature of the data matters. As analysed above in Chapter 3, the greatest risk of unintended transparency pertains to the metadata on the blockchains. Presence of metadata on the blockchains is inherent in the modus operandi of the blockchains. In other words, storage of metadata on the blockchains is *inevitable*. Although not specifically addressing the blockchain architectures, Nissenbaum is critical of characterization of data collection as an *inevitable* feature of technology.[192] Accordingly, passive storage must be understood to be clearly divorced from *higher value collection*.[193] Interpretation and analysis of metadata requires complex designs and systems that "capture, transformation, channeling and pooling of data impressions" are required. None of that is, however, inevitable and systems could be *designed* to disallow high value collection. That distinction between passive storage and the continuum of collection to use is fundamental in classifying the data processing activity on the blockchains as well. In an attempt to dissect the concepts of collection and use in great detail, Nissenbaum argues that what makes data is our *collection* of it. In other words, 'meaningfulness' is inherent in the concept of data; and in order for data to be meaningful, it must be *constructed* or *created*.[194] This process, dubbed as the 'datafication'[195], involves affirmative action. From a minimalist approach, collection thus entails *datafication* prior to storage. Mere existence of personally identifiable metadata poses no harm to individuals. It is when such metadata is *datafied* that the actual harm becomes a real possibility. As explained above, the blockchains store metadata which has not been *datafied* yet.

3.8 SECURITY

Security in the ICT can be understood in two ways: (1) security of the data, or (2) security of the information network. However, these two ways are certainly not mutually exclusive and not collectively exhaustive. Moreover, both of those set the limits of the privacy. What is good for privacy is often good for security as well.[196] Ensuring privacy is not possible without also

[192] *Nissenbaum* 2017
[193] *Ibid*
[194] *Ibid* p. 8-9
[195] *Strandburg* 2014 p. 10
[196] *Kuner et al* 2017 p.74

ensuring that the data to be protected are accessed or stolen by unauthorized third parties.[197] The place of robust security measures in data protection has been affirmed in a number of documents at the advent of the internet age.[198] For example, the influential 1980 Guidelines on the Protection of Privacy and Trans-border Flows of Personal Data, adopted by the Committee of Ministers of the Organization for Economic Cooperation and Development (OECD) in 1980 included the Security Safeguards Principle as one of the eight foundational principles of data protection. Security has been recognized in every significant codification of data protection law since then, including the EU Data Protection Directive, the U.S. Federal Trade Commission's fair information practice principles, the APEC Privacy Framework, and the EU General Data Protection Regulation.

Both the European Data Protection Directive (Directive 95/46/EC) and the GDPR deal with 'security of processing' personal data.[199] The WP 29's guidelines for ensuring the appropriate privacy safeguards mean having *high-level security* in managing the access control and authentication. This symbiotic relationship is further emphasized by the NISD:

> "Personal data are in many cases compromised as a result of incidents. In this context, competent authorities and data protection authorities should cooperate and exchange information on all relevant matters to tackle any personal data breaches resulting from incidents."[200]

Today, information infrastructure can be considered to be particularly important one among the critical infrastructures which could include "those physical resources, services, information technology facilities, networks and infrastructure assets, which, if disrupted or destroyed would have a serious impact on the health, safety, security, economic or social well-being of either two

[197] *Kuner et al* 2017 p.73
[198] Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, adopted by the Committee of Ministers of the Organization for Economic Cooperation and Development (OECD) in 1980, also in the revised 2013 version of the OECD Privacy Framework
[199] Art. 16-17 of the Directive and Art. 32 of the GDPR
[200] NISD Recital 63

or more member states or involve three or more member states."[201] Information infrastructure is analogous to *power* because none of the other infrastructures could function effectively without the information infrastructure. Included in the information infrastructure are the public telephone network, the internet, and terrestrial and satellite wireless networks.[202] This understanding is shared by the European Parliament. The fragmented nature of the member states' preparedness for security incidents on critical infrastructures has motivated the Parliament which adopted the NIS Directive as the first EU-wide legislation on cybersecurity in 2016 and the European Commission which has proposed to put in place an ICT cybersecurity certification and to grant permanent mandate to ENISA in the matters of cybersecurity.[203] That proposal builds on the 2013 EU cybersecurity strategy and the Digital Single Market Strategy both of which identify cybersecurity as one of the three key areas of action.[204] The EU legislators thus acknowledge the cybersecurity as essential to economic and societal activities, in particular to the functioning of the internal market.[205]

NISD stipulates that each Member State should have a national strategy on the security of network and information systems defining the strategic objectives and concrete policy actions to be implemented.[206] The effect of cybersecurity breaches on society is becoming more and more visible.[207] Member states are expected to transpose it to national legislation by May 9th, 2018. NISD is a minimum harmonization instrument therefore the stakeholders are free to impose stricter requirements for security than provided by the NISD.[208] NIS applies to both operators of 'essential services' and 'digital service providers.'[209] In relation to sector-specific

[201] The Green Paper on European Critical Information Infrastructure p. 7 also see *Personick and Patterson* 2003, preface and p.1 for a comparison on the US definition of critical infrastructures that only include finance, transport, water and energy.
[202] *Personick and Patterson* 2003, preface and p.1
[203] Commission Proposal for a Cybersecurity Act 2017
[204] Cybersecurity Strategy of the European Union 2013
[205] NISD Recital 1
[206] NISD Recital 29
[207] https://www.telegraph.co.uk/technology/0/worst-meltdowns-time/
[208] NISD Recital 6, Art. 3
[209] NISD Recital 7: "However, the obligations on operators of essential services and digital service providers should not apply to undertakings providing *public communication networks* or publicly available electronic communication services." Security and integrity requirements of public communication networks are regulated by Directive 2002/21/EC of the European Parliament and of the Council (the "Framework Directive") and eIDAS.

security requirements which are provided in sector-specific instruments such as the water transport sector, incident reporting requirement of the NISD is *lex specialis*.[210] Similarly, the NISD expects that the security requirements of other sectors may also be regulated by the member states in future in which case the member states can apply national laws as long as the requirements are at least equivalent in effect, and that they provide information to the Commission on the application of such *lex specialis* provisions.[211] NISD primarily requires the member states to designate a national single point of contact responsible for coordinating issues related to the security of network and information systems and cross-border cooperation at Union level called "the computer security incident response teams ('CSIRTs') also known as computer emergency response teams ('CERTs')".[212] Security requirements of the Directive apply only to those public administrations which are identified as 'operators of essential services' which the member states are responsible for determining according to the criteria set by the Directive.[213] Accordingly, Art. 14(1) of the NISD provides that:

> "Member States shall ensure that *operators of essential services* take appropriate and proportionate technical and organisational measures to manage the risks posed to the security of network and information systems which they use in their operations. Having regard to *the state of the art*, those measures shall ensure a level of security of network and information systems appropriate to the risk posed."

NISD lists the 'operators of essential services' in Annex II. Accordingly, undertakings operating in energy, transport, banking, financial market infrastructures, health sector, drinking water supply and distribution and digital infrastructures (IXPs, DNS service providers and TLD name registries) are prima facie falling into the category of operators of essential services. Digital service providers are listed in Annex III to be online marketplace, online search engine and cloud computing service. Because some digital services with no alternatives available could be an important resource for their users, including for the operators of essential services, therefore

[210] NISD Recital 10: In water transport industry, security requirements for companies, ships, port facilities, ports and vessel traffic services under Union legal acts cover all operations, including radio and telecommunication systems, computer systems and networks.
[211] NISD Recital 9
[212] NISD Recital 31-32, 34, Art. 7, Art. 9 and Art. 12
[213] NISD Recital 45

this Directive should also apply to providers of such services.[214] Hardware and software manufacturers, however, are subject to product liability rules which are outside of the scope of the Directive. From the practical point of view, many businesses and undertakings which fall under either of the categories presently maintain adequate ISO standards to be cyber-secure. However, the incident-reporting obligation under the NISD may bring additional costs to businesses which they are not fully prepared to undertake.[215] The full effect and implications of the Directive is yet to be seen once all member states transpose and enforce it. NISD is nevertheless a long-due harmonization effort by the legislators in the EU that highlights the merger of security of data, security of network and privacy. The US has had a similar law in place, the Cybersecurity Information Sharing Act, since 2015 with no significant differences to the NSID.[216]

3.8.1 SECURITY BY DESIGN: THE CASE FOR BLOCKCHAINS

Security by design is one major component of the traditional Privacy by Design concept.[217] The GDPR, for the first time, has rendered the Privacy by Design a legal obligation. Kshetri argues that blockchains may prove to be a nightmare for cybercriminals, data manipulators and others who mishandling personal data.[218] Dubbing the blockchain and cloud computing as the "kissing cousins" from security and privacy angles, he presents a comparison of the two.[219] He concludes that even though the newness of the blockchains mean that external security mechanisms have not yet been developed for some systems, some of the key security challenges associated with the cloud can potentially be addressed by the nature of the blockchains.[220] As there is no single point of failure or vulnerability, blockchains offer a substitute for outdated systems such as the SWIFT from security point of view.[221] Maersk' decision to apply blockchain

[214] NISD Recital 48-49
[215] https://www.telegraph.co.uk/business/2016/03/02/businesses-keep-quiet-over-cyber-attacks-as-eu-cracks-down-on-un/
[216] https://www.congress.gov/bill/114th-congress/senate-bill/754/text
[217] Art. 25 of the GDPR
[218] *Kshetri* 2017 p. 1036
[219] *Ibid*
[220] Also see: *Park — Park 2017*
[221] Bangladesh Central Bank suffered a large loss in 2016 and recently India's City Union Bank suffered a smaller loss by same methods in 2018:
https://www.wired.com/2016/05/insane-81m-bangladesh-bank-heist-heres-know/

to its supply chain following the June 2017 cyberattack[222] is surely not a coincidence.[223] Similarly, the EU Digital Single Market project also problematizes the *security* and *verifiability* of *identity* being a major obstacle to interacting devices online.[224] A research conducted by Veracode in 2015 shows that existing IoT devices could suffer from major security vulnerabilities including, among others, lack of encryption, failure to enforce strong passwords and lack of transport layer security ("TLS") encryption or lack of proper certificate validation.[225] These are severe risks in relation to IoT devices which could potentially mean an open invitation to hackers and criminals. Khan and Salah identify five solutions provided by blockchains in IoT security.[226] These are: 1) address space, 2) identity and governance, 3) data authentication and integrity, 4) authentication, authorization and privacy, 5) secure communications.

In sum, blockchains' superiority in ensuring security is the result of:

1. Disintermediating the transactions whereby removing the possibility of single point of failure and ensuring the major security requirement of the "CIA" namely the confidentiality, integrity and availability,[227]
2. The ability to deploy specific member targets in the chain such as regulators and auditors,
3. Use of the cryptographic hash functions, and the public-private key cryptography, and
4. Its ability to authenticate and verify *identity and origins* of things and persons.

This is not to say that blockchain applications are without invincible. In 2016, the Decentralized Autonomous Organization set out the largest crowd-funding project undertaken at the time. The funding had a 28-day window during which period someone exploited vulnerability in the DAO's code and stole USD 3.6 million ether from the fund. That amount was equivalent of a

https://www.reuters.com/article/us-city-union-bank-swift/india-bank-hack-similar-to-81-million-bangladesh-central-bank-heist-idUSKCN1G319K

[222] https://www.bloomberg.com/news/articles/2017-08-16/maersk-misses-estimates-as-cyberattack-set-to-hurt-third-quarter

[223] https://www.maersk.com/press/press-release-archive/maersk-and-ibm-to-form-joint-venture

[224] https://ec.europa.eu/digital-single-market/en/internet-of-things

[225] *Kshetri* 2017 p.1031, for the report: https://info.veracode.com/whitepaper-the-internet-of-things-poses-cybersecurity-risk.html

[226] *Khan and Salah 2017*

[227] CIA is a security concept also referred to in Art. 32(1)(b) of the GDPR

value between USD 64 million – 101 million.[228] Such an attack doubtlessly raised serious questions on the level of security the blockchains could offer.[229] Nevertheless, due to the unprecedented security of proof of identity function offered by the blockchains, including the confidentiality, integrity and availability (the "CIA" in the information security parlance) the existing vulnerabilities are arguably merely a work-in-progress. For instance, since the blockchains are capable of decentralizing the DNS, the contents can be distributed to a high number of nodes which can in turn prevent a DDoS attack. This is because for a DDoS attack to be successful, the attack needs to be able to completely eradicate the data on each node at the same time which is possible due to the centralized nature of the DNS but nearly impossible on a fully decentralized blockchain.[230]

Using a blockchain infrastructure in order to enhance the security of online marketplaces (as mentioned in the NSID) is a good example of how this technology could be utilised vis-à-vis the cybersecurity legislation.[231] In a decentralized marketplace, identity of users is not disclosed on the network. Tracking transactions can only be facilitated with difficulty. It is also possible to conceal transactional details behind layers of encryption by way of using mixing techniques even though it is never possible to hide the meta-data completely. In traditional online marketplaces, on the other hand, the security is ensured as much as the network's own components. Moreover, it is highly costly to distort the consensus mechanism on the blockchains.[232] With this on mind, it would not be an exaggeration to state that the blockchains offer a valuable *state-of-art* mechanism as advised under the NISD for compliance. Similarly, use cases where distributed ledger technology is utilized as a "blockchain infrastructure as a service" in the same way as the cloud computing services, the NISD advice can be maximally achieved.[233]

---

[228] https://www.coindesk.com/understanding-dao-hack-journalists/
[229] https://www.ft.com/content/05b5efa4-7382-11e6-bf48-b372cdb1043a
[230] https://www.infosecurity-magazine.com/next-gen-infosec/blockchain-cybersecurity/
[231] *Subramanian* 2018 p. 81
[232] *Ibid* p. 83
[233] Blockchain And Cryptocurrency May Soon Underpin Cloud Storage, Blockchain-based systems challenge AWS, Dropbox

Utilizing the blockchains to combat fraud with maximum security is in particular desired in industries where authenticity is of paramount important. Banking and finance industry are already using the blockchain for fulfilling their Anti-Money Laundering – Countering the Financing of Terrorism (AML-CFT) compliance obligations. Indeed, AML Directive requires a minimum of five years of personal data retention "in order to be able to cooperate fully and comply swiftly with information requests from competent authorities for the purposes of the prevention, detection or investigation of money laundering and terrorist financing"[234] while at the same time stipulating in no uncertain terms that:

> "Member States should require that specific safeguards be put in place to ensure the security of data and should determine which persons, categories of persons or authorities should have exclusive access to the data retained."[235]

While using blockchains for storing personal data is not a good idea for the reasons explained in Chapter 3 vis-à-vis the GDPR, at least certainly not on a permissionless blockchain, there are initiatives utilizing blockchains as a compliance tool for the KYC obligations.[236] Normative technology regulates behaviour because it is used intentionally as an instrument to influence human behaviour. Technologies, however, are not always necessarily invented with the intention of controlling or regulating behaviour. Such use may simply be a secondary outcome of a particular technology. Even though blockchains are not (fully) instrumentalised as a tool for regulation in the same way e.g. DRM systems, filtering systems, PETs and terminator technologies are,[237] using them as a compliance tool for the KYC obligation is certainly an important step in that direction. This type of utilization of blockchains is consistent with the Reidenbergian argument that policymakers ought to embrace the *Lex Informatica* and utilize it

[234] AML Directive Recital 44
[235] AML Directive Recital 44
[236] http://fintechnews.sg/14420/blockchain/ibm-completes-poc-blockchain-based-shared-kyc-deutsche-bank-hsbc-mufg-cargill-ibm-treasuries/
[237] *Koops* 2008, p.157: "Technology has always had a certain normative element—it is never neutral. Notable examples are Digital Rights Management (DRM) systems (enforcing—or extending—copyright), filtering systems (which block 'harmful' content), Privacy Enhancing Technologies (PETs, which allow citizens the control over personal data that they are losing in the digital age), and terminator technologies (which prevent genetically modified foods to multiply, forcing farmers to buy new crops each year)."

as a *tool* for controlling information flow on the global networks.[238] There is a strategic implication in that: through effective channeling, *Lex Informatica* facilitates the government interference to be characterized more as an indirect influence rather than the direct regulation. While it is not clear if such applications are storing personal data which is due to be erased from the system after a number of years, there are other ways to take the benefit of the security feature of the blockchains in the banking and finance industry without risking placement of personal data on the chains.

3.8.2 THE WHO QUESTION

The Cambridge Analytica scandal has brought the concept of *data portability* to public awareness with force. While much of the debate has been going around businesses that exploit personal data, a more fundamental change enforced by the GDPR Art. 20 on data portability as a data subject's right is to bring the consumer into the focus of data business. In order to fully reach that goal, however, the *who* pillar of the internet as the target of regulation has to be answered and interoperable systems have to be secured.[239] Proof of identity has been a topic of focus in the blockchain uses in particular where the KYC/AML checks on natural persons are required. Along the same line, self-sovereign personal data management has been debated in the tech-communities such as the MyData initiative for the last a few years. The problem is it is not so easy to duplicate the process of paper document authentication done by humans in a digital environment. That is due to two problems: 1) lack of a standard format, and 2) lack of a standard way of verification. The first problem is within the domain of the question of *interoperability*, a topic also pinned by the GDPR Recital 68,[240] and the second question is within the domain of 'digital credentials'. While digital signatures are already legally valid, they require two keys: private and public keys. Private key is the key with which the owner signs a document. That is

[238] *Reidenberg* 1998 p.586
[239] See Chapter 2 of this thesis on the *Regulability of Blockchains* for a reminder on the question of the 'who' pillar.
[240] The problematique of interoperability is dealt with by the revised Directive on the reuse of Public Sector Information, the INSPIRE Directive as well as the new EU initiatives such as the European Cloud Initiative, the EU eGovernment Action Plan 2016-2020 and the envisaged Single Digital Gateway Regulation lay the regulatory background to information infrastructure landscape in the EU. However, a comprehensive analysis of those instruments is outside the scope of this research and merits a separate research.

kept secret. The public key is used in order to 1) verify the signature, and 2) ensure that the document in question is not tampered with. However, there exists no standard way to verify the public key of the issuer which functions as proving the authenticity of the credential. W3C Verifiable Claims Working Group Charter states in their Charter that it is currently difficult to express banking account, education qualifications, healthcare data, and other sorts of machine-readable personal information ("claims") that has been verified by a third party on the web.[241] As explained in Chapter 2, internet is the consequent design of standardized network packets. The Web is the consequent design of 'standardized hypertext.' [242] Thus, standardizing digital credentials just as having standardized the hypertext can enable a system of credential issuers, owners and verifiers all exchanging interoperable verifiable claims.[243] To do that, the security afforded by public blockchains can be utilized as the foundation on which the decentralized identifiers can be maintained publicly with very little security risk of being compromised. Self-sovereign identity applications such as Sovrin or uPort do precisely that. In that model, the data subject takes the role of his/her own controller from the regulatory point of view in a similar way that was discussed in above and also in Chapter 4.

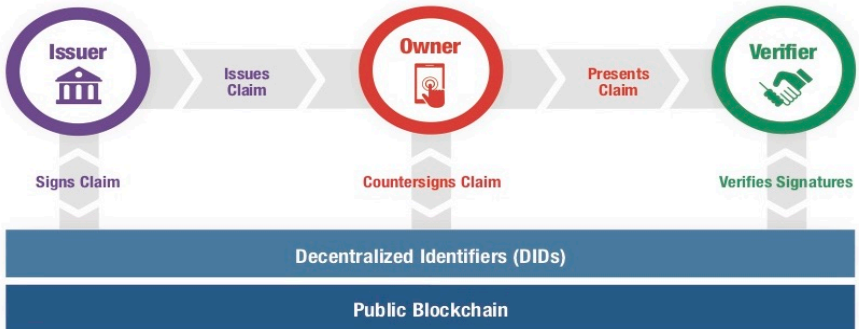


Table 5: Sovrin digital self-sovereign identity infrastructure

3.8.3 HARMONIZING DIGITAL TRANSACTIONS ACROSS THE EU

Another way the blockchain infrastructure could be used for consistently with the European legislation as both an object of regulation and a regulatory tool is the 'Trusted Service

[241] https://www.w3.org/2017/vc/charter.html
[242] The HTTP provides a standard for web browsers and servers to communicate and the HTTP is an application layer network protocol built on top of the TCP.
[243] Sovrin White Paper p.6

Providers'. In that model, TSPs sit in the role of a central controller from the regulatory point of view.

Streamlining digital authentication has been one of the goals of the European Commission as part of the Single Digital Market. To that effect, the Commission passed the eIDAS in 2016. EIDAS lays down the rules on electronic identification and trust services for electronic transactions in the internal market. It defines trust services for supporting electronic signatures, electronic seals, electronic time stamps, electronic registered delivery services and website authentication. In a bid to reduce barriers to digital single market, the legislation is designed to have member states recognize the electronic identification schemes of one another for the purposes of cross-border authentication for a public service online. The idea is to allow citizens, businesses and public administrators to use the eID means and 'qualified trust services' of their choice for any cross-border transaction across the Union. To do that, EIDAS requires that the European governmental agencies appoint the Trust Service Providers, highly qualified market operators with 'EU trust mark', after a strict conformity assessment[244] "*where an electronic identification using an electronic identification means and authentication is required in order to access a service provided by a public-sector body online."* [245] TSPs are typically supposed to procure services including timestamping, electronic seals, document storage and archiving. EIDAS requires TSPs to meet specific requirements in delivery of those services. Those requirements relate to *high security*, use of *trustworthy systems*, performance audit, *legal certainty* and consumer protection. TSPs are excluded from the application of NISD security requirements since they have their own requirements under the eIDAS Art. 19. Moreover, TSPs can be held liable for failing to hold to those standards.

By way of highly regulated TSPs, eIDAS thus readily affords the possibility of legal compliance by using blockchains and it is highly likely that by utilising blockchains, duties of TSPs could be facilitated.[246] As seen above, for the regulatory effort to make sense, a stable and static anchor

[244] In Finland, the Population Registry is the sole qualified trust service provider by the Finnish Communications Regulatory Authority. Please see: http://vrk.fi/sahkoinen-henkilollisyys-ja-varmenteet for more information.
[245] Art. 6 of eIDAS
[246] A*tzori* 2017 p.10

is needed in the object of regulation in order to monitor standards and for the accountability principle to exist. This is why the roles of data controller and data processor are quite significant and meaningful for the protection of consumer rights. Therefore, legally compliant blockchain use cases all need that anchor of accountability. Establishing interoperable, high security electronic identity schemes with TSPs as the accountability points are a realistic use case in which the blockchain is both the object of regulation and a tool for regulation.[247]

Today's cybersecurity reality and the nature of data protection and privacy compel the regulatory mindset to take greater stock of the technology. Even though cybersecurity has been driven with the motivation to save costs in terms of the financial losses suffered by the breached entity, Kuner et al suggests that "a greater understanding that information security, as a component of data protection, is not just a financial obligation, but a human rights obligation might contribute to a broader accounting of the harms that may be caused by breaches and the range of parties who may be injured. [248] In that regard, "the human rights foundations of data protection law could benefit efforts to improve cybersecurity as well."[249] On the EU level, however, the institutionalization of the cybersecurity efforts should be taken with caution. Even though the EU could be conceived as an emerging soft power in cybersecurity,[250] *Carrapico and Barrinha* suggest that the normative assumptions underlying such institutionalization should be questioned and not taken for granted.[251]

[247] *EU Blockchain Observatory Forum – Blockchain and Digital Identity Report 2019* p.19-21
[248] *Kuner* et al 2017 p. 75
[249] *Ibid*
[250] *Christou* 2017 p. 9
[251] *Carrapico and Barrinha* 2017 p.1267. conclude that this institutionalization operates from the presumptions that: 1) it is better for the EU to act as a unitary actor, and 2) only in a more coherent Union is can the best possible toolkit of action be offered in the security field. A broader analysis of EU's cybersecurity policy is, however, beyond the scope of this thesis.

4. CONCLUSIONS AND POLICY CONSIDERATIONS

4.1 GENERAL REGULATORY POLICY

The European privacy and data protection regulation of the blockchains should be left at a high level, into its natural course of evolution.

With all the challenges and benefits they offer, blockchains can be thought of as a *disruptive event causing an external flux*.[252] Our thinking of blockchains require a *foundational shift* in the way we have always done things in order to make a difference. Developments and discussions around economies and value-creation models as an alternative to the existing system based on fiat money,[253] for instance, compel us to question our economic and ideological premises. Similarly, a total disintermediation of transactions directs us revisit the concept of 'trust' and why we have needed the middle man in the first place. Part of the reason for which blockchain is referred to as a revolutionary tool is that it is capable of distributing the control among users and restoring the power to the *ordinary* man of internet.[254] Therefore, implementation of blockchain in various settings has the potential of affecting the existing power dynamics between online operators and their users.[255] Although in the recent decades the consumers' consumption habits have perhaps radically changed, the ideology that instrumentalizes the information technology to perpetuate its dominant characteristics remain the same.[256] Blockchain, however, does represent a radical break from the existing ideology in our culture. Regulating a democratizing tool will imply regulating power relationships to a certain extent, and hence the ideological superstructures of technological development must be accounted for in a thorough analysis of policy. Therefore, it is possible to deduce that among the other technological, economic or environmental developments, blockchains compel not only a legal

[252] *Murray* 2008 p. 289: the natural evolution of the regulatory settlements takes the form of external and internal fluxes. It is the disruptive events that cause the external flux which in turn compel an internal flux in the form of a response to the moral or social development that ensue.
[253] *Pazaitis* 2017
[254] *De Filippi*, Primavera "The interplay between decentralization and privacy: the case of blockchain technologies. Journal of Peer Production, 2016, Alternative Internets, 7. <hal-01382006>" p.2
[255] *De Filippi* 2015
[256] *Birdsall* "The internet and the ideology of information technology"

but also a moral and social response that translates as an *internal flux*. In the external flux, we have the opportunity to look at the way the regulations are going to apply to the technological change; and in the internal flux, we can look at the values, ideologies and philosophies underpinning our expectations and limitations as societies. It is at this deeper layer where the normative values are highlighted.[257] Technology can be both norm-enforcing and norm-establishing[258] depending on the kinds of values the society chooses to uphold.[259] In this context, blockchains can be used for enhancing the autonomy of the data subjects and improving their ability to protect their personal data (see below in Conclusion 3). Blockchains can also be used in order to facilitate compliance for KYC. In effect, those choices are *political*.[260]

In order for the society to settle on the kinds of values on which it wants the cyber regulation to premise, it needs ample leeway and time. In other words, the regulators will have the opportunity to understand the object of regulation and its limits as well as the societal expectations in time.[261] The regulatory environment is yet to experience the outcomes of the blockchain's broad

---

[257] *Lessig* 1998: "the law's relationship to the behaviour is two-fold: by affecting the other three constraints and by directly affecting the behaviour itself."; *Lessig* 2006 p. 111-112: "Internet's architecture embed certain values which can change as the features of the architecture changes. The kind of architecture we encourage is a choice about the policy we encourage."; *Lessig* 2006, p. 70: "the central to regulation of the internet is not "whether governments could induce an ID-rich internet." Central to the discussion is "*the kind* of ID-rich internet which the governments induce." *Lessig* 2006 p. 70- 78: "a regulatory environment must be informed by the kind of values we want our world to have."

cf. *Serge, De Hert and Sutter 2008 Ibid p. 195* whereby Lessig was heavily criticized for its "optimal mix" that it "actually does not work because regulation cannot be a form of activism, and that regulation is not only a legal matter, it embeds a political and ideological process. However, this thesis disagrees with that reading of Lessig's regulatory approach. On the contrary, Lessig makes no hard distinction between the law and politics purely because it is implicit in the legislative process that what becomes law is society's values.

[258] *Koops* 2008, p.159; also see *Brownsword 2008* p. 158: "technology, increasingly, enforces or supplements law as an important regulatory instrument. (...) however tentatively, technology is increasingly used intentionally to enforce or establish norms. Technology that sets new norms clearly raises questions about the acceptability of the norms, but also if technology is used 'only' to enforce existing legal norms, its acceptability can be questioned since the reduction of 'ought' or 'ought not' to 'can' or 'cannot' threatens the flexibility and human interpretation of norms that are fundamental elements of law in practice."

[259] *Lessig* 2006 p. 78, for the choices of layers also see *Benkler* 2000

[260] *Lessig* 2006 p. 78, also *Koops* 2008: "Rather than only look at normative technology from the perspective of safeguarding the democratic constitutional state, we should thus also look at democratic and constitutional values from the perspective of normative technology."

[261] *Law* 2007 p.8 - 12: Law draws from the Deleuzian logic that "there is no overall social, natural or conceptual framework or scale within which events take place: as webs grow they tend to grow their own metrics" thus "we need to trace how the webs of heterogeneous material and social practices produce them."

implementation across public and private sectors. Murray, both building from Lessig's theory and distinguishing it, illustrates this point by use of the 'gardener's dilemma.'[262] In summary, the gardener's dilemma is the challenge identifying the optimum system configuration which produces the maximum yield from the garden as a whole.[263] A regulator can never be sure, in any complex system, what affect their actions will have because the problems laid out by the gardener's dilemma are "computationally intractable.[264]" Lack of predictability in the ways which other sub-systems will be disrupted is likely to slow the evolution of the regulatory law on new technologies such as the blockchains. This is no different in the case of regulation of privacy and data protection. Therefore, it can be said that the optimal privacy and data protection regulatory environment for the blockchains is arguably one without narrow proscriptions focused on well-defined problems.[265] The language of the GDPR, the statements of the EDPD and various guidelines from the national DPAs appear to support this position.[266] It appears that the privacy and data protection regulatory culture in the EU is adaptive and non-rigid; thus, the language creates an elastic system that is open to adapting to the form set by technology or capital investment.[267] As a result, regulatory settlement between the external flux of blockchains and its internal flux can be achieved in its own natural evolutionary process.[268]

At present, privacy and data protection is regulated by the national data protection authorities in the EU and the data subjects can enforce their rights the national courts. Where a new technology

---

[262] *Murray* 2007 Chapter 2: Regulatory Competition and Webs of Regulation, in particular from p. 25. Systems theory was first proposed by the biologist Ludwig von Bertalanffy. It focuses on the arrangement of and relations between the parts that connect into a whole. Moreover, Murray suggests four modalities of regulation which they call 1) hierarchical control, 2) competition-based control, 3) community-based control and 4) design-based control. (*Murray - Scott* 2002)

[263] *Murray* 2007 p. 26. Also see *Rust* 1997 for the original source of the gardener's dilemma used by *Murray*.

[264] This argument is based on "The Law of Requisite Variety" developed by *W. Ross Ashby* (An Introduction to Cybernetics, 1956, London: Chapman & Hall) cited by *Murray* 2007 p.26. 'Chaos Theory' also dubs an identical effect as the 'butterfly effect'.

[265] cf. *Walker* 2002 "Resolving conflicts between particular technologies and the right to privacy require "narrow proscriptions focused on well-defined problems."

[266] For example see Opinion of the Hungarian Data Protection Authority; EU Blockchain Observatory Forum – Blockchain and the GDPR Report 2018, *CNIL's Opinion on the Blockchains 2018*

[267] *Luhmann 1989* p.144 - 146

[268] *Law* 2007 p.8, quoting Deleuze, Gilles, and Félix Guattari (1988), A Thousand Plateaus: Capitalism and Schizophrenia, London: Athlone. "There is no overall social, natural or conceptual framework or scale within which events take place: as webs grow they tend to grow their own metrics."

is employed and there is a likelihood that the data subjects' rights and freedoms can be compromised, a data protection impact assessment is required to be made.[269] If such impact assessment returns a high risk, then a consultation with the DPA is needed.[270] It is likely that at least some use cases of blockchains will be processed as part of Art. 36, "prior consultation." Thus, the blockchains can reasonably be expected to be regulated on an ad-hoc, case-by-case basis until the foundational social and legal shifts have achieved stability. The unpredictability of the disruptive effect of the blockchains as well as the interactions between the systems based on such effect create a dynamic regulatory spirit which cannot be proscriptively fixed.[271]

This is obviously the preferred way to regulate the new technologies even if at times it may not be the most effective design of regulation as it gives rise to uncertainties. Nevertheless, focusing on the most effective design to regulate the privacy and protection of the personal data could potentially overshadow the discussion on the *values* that are pursued.[272] It is, thus, better to allow the leeway and time for the shifts to be experienced, societal settlements to take place and the nature to take its course.

4.2 FROM EX ANTE TO EX POST REGULATION

The privacy and personal data protection on the blockchains call for a "shift from ex ante to ex-post regulation."

Crawford and Schultz propose that a *procedural data due process* could reach the places where the harm takes place.[273] Instead of regulating ex-ante, the procedural data due process can regulate the 'fairness' of Big Data's analytical processes with regard to how the personal data is *used,* in other words, ex-ante. In a similar vein, Black defines *proceduralization* as a

[269] Art. 35 of the GDPR
[270] Art. 36 of the GDPR
[271] *Mayer-Schönbergert 2010:* "privacy regulation can only be achieved if the underlying mechanisms of the information governance can be understood"; also see *Vranaki* 2017 p. 210. Vranaki shows that "the protection and/or violation of personal data is an effect generated from specific *socio-technical-legal assemblages* rather than the outcome of a single actant" in other words, regulation is an outcome of interactions.
[272] *Black* 2000-1 p.598
[273] *Crawford and Schultz* 2014 p. 109

regulatory effort to optimize the 'design' and 'implementation' of a policy rather than with normative concerns of what that policy should be.[274] The GDPR's risk-based approach can be seen as one step towards that vision. This thesis argues that one further step should be for the regulators and authorities to recognize the nuances of data processing and treating 'storage' and 'collection' and 'use' with varying degrees of force proportionate to the harm they may cause, based on whether the data is or can be datafied or not (not whether they are personal data or not, see below). And in order to legitimate recognition of such distinction, a third further step could be that the next generation legislative effort can and should make a distinction among 'storage' and 'collection' and 'use'.

Crawford and Schultz are joined by Koops[275] and Nissenbaum[276] in the *proceduralization* approach. According to Koops, the data protection regulation should be done *ex post* on the use of the data in decision-making. By this way, the desired outcome of the protection of the citizens would be more likely to be achieved[277] because the protection is placed on the individuals rather than their personal data. In the *ex post* regulation of the use of data, a recalibration of transparency is suggested by Koops. Such a recalibration would take place in a matter of vertical scale whereby the transparency would diminish upwards and lessen downwards. The further down the scale, the more transparent the data would be. The more transparent the data, the more mechanisms would be employed on a case-by-case or (near) real-time basis.[278] The transparent metadata revealed on the blockchains as discussed in Chapter 3 of this thesis can be put in context of this theory.[279] The privacy harm caused by the blockchains pertains to the unintended transparency of metadata on the blockchains. However, until that metadata is *datafied*, an actual

[274] *Black* 2000-1 p.598 where Black builds on the Teubnerian reflexive law approach of 'deliberation' and 'participation.'
[275] *Koops* 2013
[276] *Nissenbaum 2017* : the 'regulatory effort' must be redirected from the collection of data to the use of data. She calls this approach the Big Data Exceptionalism ('BDE'). Nissenbaum suggests that regulation should take the form of "application of the techniques through which the desired outcomes can be achieved."
[277] *Koops* 2013 p.2 – 5. To illustrate his point about moving to an ex post regulation Koops explains that that the GDPR's requirement for the data controllers to inform the data subjects about the existence of processing for an automated decision offers little utility since it merely ensures *process transparency* and not *outcome transparency.* Knowledge of an automated decision-making is less likely to yield any protection to the data subject unless the data subject also understands the implications and risks of such decision-making.
[278] *Koops* 2017 p. 9
[279] See Chapter 3 of this thesis for a discussion on the radical transparency of blockchains.

harm is arguably not possible. Until a negative outcome of meta-data transparency materializes, there is no actual harm. In that case, it is unreasonable to argue that the radical transparency of data on blockchains is harmful per se. Applying Koops' vertical scale, the *undatafied* data revealed by the blockchains should be considered as an individual case, and not be considered to be processing within the meaning of the GDPR and should not attract the GDPR obligations.[280] Such a harm-based view on defining the scope of personal data processing mandates an understanding of the risks of harm. To that end, Esayas recommends that different *risks of harm* should be considered on the basis of (1) the *sensitivity* of the data, and (2) the context of its *usage*.[281] Indeed, the GDPR proscribes a risk-based approach which informs the basis of, among others, the Art. 35 and Art. 36 obligations.

The question of regulating the mere storage of unencrypted meta-data is intertwined with the question of anonymization. Even if it were feasible to encrypt the meta-data, it is nearly impossible to achieve the 'adequate level of legal protection' by way of *anonymization* at the standard expected by the WP 29 without destroying the *utility* of data.[282] Once again, as it is currently not possible to predict at what stage of encryption the unquantifiable privacy harms could be alleviated, this research finds that a more nuanced approach is needed also vis-à-vis potential linkability/ identifiability of personal data. Famously, Ohm has argued that the scientific basis of the promise of perfect privacy by way anonymization is lost because the computer science has proven that "all anonymized data could be re-identified."[283] He recommends abandoning treating anonymization as a 'silver bullet' in privacy challenges and removing the distinction between personal data and non-personal data as "it no longer serves a

[280] Art. 4(2) of the GDPR
[281] *Esayas* 2015 p. 8:
1. the anonymisation is employed to sensitive data and would be publicly available,
2. the anonymisation is applied to sensitive data and would be available with limited access,
3. the anonymisation is applied to non-sensitive data and would be publicly available, and
4. the anonymisation is applied to non-sensitive data and would be available with limited access.

On that model, the fourth level would require less strict anonymisation than the previous levels.
[282] Moreover, there are disadvantages and societal costs of requiring destruction of data in that way, see *Emam - Avarez* 2015 p. 81.
[283] *Ohm 2010* p.1736

function."[284] This thesis agrees with Ohm's finding and argues that the societal expectation of complete camouflage is a false ideal that should not be nurtured by the regulators and academics. Possibility of strong pseudonymisation may soon become irrelevant in the face of the ever-developing methods to link data and identify persons. With this on mind, a risk-based approach without also factoring into the possibility of actual harm to the data subjects is not helpful.

El Khoury, along the same lines, questions the possibility of a correct *risk assessment* in the face of the uncertainty of which piece of data is capable of becoming personal data due its role in to re-identification. Could *every* piece of information then be potentially personal data? He contends that there is no univocal answer to this question. The likelihood of identification is high and grows with time so it is possible to conceive every piece of data as potentially personal data. On the basis of a strict interpretation, however, not every piece of information is personal data because not every piece of data can reveal factual circumstances or behaviors of data subjects.[285] Both points of views are presented as valid by El Khoury. Personal data is necessarily a *relative* concept, and safeguards should be applicable *beyond* the notion of personal data. El Khoury argues that this approach would necessarily mean acknowledging *a grey zone* where certain provisions of the DPD are not applicable, namely, the ones linked with the exercise of the data subject's rights.[286] Another criticism to sweeping privacy regulations comes from Ohm. Ohm argues that instead of universally applicable data protection laws, *sector-based privacy regulations* should be enforced if it is evidently shown that harm outweighs benefits of unfettered information flow.[287] This research also agrees with El Khoury' solution and defends the position that a 'grey zone' should also be recognized in relation to the blockchain applications vis-à-vis the GDPR.[288]

---

[284] It is worth noting at this point that Ohm's work has been widely criticized, for example, by *Stalla-Bourdillon & Knight* 2017 who argued that Ohm's approach would be incompatible with the GDPR stating that "robust anonymization practices in satisfaction of *an adequate level of legal protection* of individual's privacy and other rights" should still be allowed [p. 307]. This research disagrees with this interpretation of Ohm's work. Ohm's revelation does not imply that robust anonymization practices should be 'disallowed'; rather, people to correctly 'understand' anonymization with its limitations.

[285] *El Khoury* 2017 p. 7

[286] *Ibid*

[287] *Ohm 2010* p. 1759

[288] As previously discussed, there are architectural similarities between the cloud and blockchain. In that respect, a lot of the criticism and analysis in relation to cloud guidelines previously bear significant relevance in shaping our

Thus, a shift in the regulatory focus on the use of metadata could indeed cumulatively deliver more benefits to the society than an obscure and single-minded application of a rule practically rendering the (permissionless, in particular) blockchains illegal under the data protection law for uncontrolled transparency of meta-data.[289] At the consultation stage, the authorities should recognize a difference between the undatafied data and datafied data, with an added harm-based approach to inform the existing risk-based approach. This is especially important in cases where it is not possible to further mitigate the risks associated by a particular technology.

4.3 DATA SUBJECTS AS DATA CONTROLLERS

There is no controller on the permissionless blockchains within the meaning of the GDPR and that the data subjects can be considered sovereign controllers

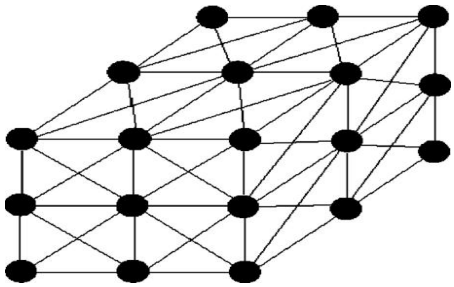
A consequence of the paradigm shift promised by the blockchains is tied to a number of key social changes as laid out in conclusion #1 above.[290] In particular, one key social change linked to the missing "who" pillar of the internet is intertwined with the expanded data subject rights under the GDPR. The expansion of the data subject rights and emphasized transparency on the processing of personal data are a result of the European consumer empowerment trend in digital world. Especially today when the internet users are questioning the premises of the data-driven world, the ethical and financial limits of technological conveniences and development and would like to be more empowered.

approach to blockchain, too. See WP196 and *Millard* 2014: if each piece of data is potentially personal data due to identifiabillity and linkability, it would be impossible "for data controllers to use public cloud computing for processing personal data. The highly customized service that regulators demand would not be 'cloud', and would deliver neither the efficiencies nor the process improvements that make cloud computing attractive. Requirements designed for traditional outsourcing simply do not fit the public cloud model."

[289] *Reidenberg* 1998 p.587

[290] *Murray* 2008 p. 287 where Murray builds on Black's work (*Black* 2001) and argues that the socio-legal order has moved from the regulatory state to the post-regulatory state. In analyzing the socio-legal order a decade ago, he found that "paradigm shifts reflect key social changes; incremental shifts represent society in evolution."

While the claims of interference in democratic elections by unethically built algorithms and the "fake news" debates have made the citizens feel less empowered and more like the Lessigian "pathetic dot,"[291] the conversation is far from final. Murray, opposing Lessig's characterisation of the role of the individual in the regulatory system as a 'pathetic dot', argued in 2008 that we are witnessing a "formalisation of the power of the community."[292] Individuals are now more and more empowered agents in the complex ICT environment as part of web of community of individuals. In Murray's 'Active Dot' Matrix, (*see Figure 2 below*) any action by any one member of the regulatory matrix (either as regulator or regulatee) has an effect on the actions of the others thus together shaping the environment.[293] The possibility of sovereign digital identities using the blockchains may allow the consumers to become their own decision-makers for how their personal data can be used.[294]



This would create an interesting outcome from the regulatory point of view whereby the decentralized architecture of the blockchains would result in "decentring the regulation."[295] and free the individuals from being mere

[291] *Lessig* 2006 p.123
[292] *Murray* 2008 p. 302
[293] *Ibid* where Murray argues that the regulation of cyberspace can be achieved on the pinch-points. Also see *Reidenberg* 1998, *Lessig* 2006 p. 122-123; cf. *Vranaki* 2017 whereby cyberspace regulation is argued to be an outcome of numerous interactions among various actants that produce 'relational power effects'. Overall, Vranaki argues that "the protection and/or violation of personal data is an effect generated from specific socio-technical-legal assemblages rather than the outcome of a single actant" in reference to Lessig's code. (p.210)
[294] EU Blockchain Observatory Forum – Blockchain and Digital Identity Report 2019 p.14; also see See for a discussion on self-sovereign identities:
https://sovrin.org/wp-content/uploads/2017/06/The-Inevitable-Rise-of-Self-Sovereign-Identity.pdf
As a foundation to self-sovereign identity projects of today, Lessig's point about the efficiencies of *virtual wallets* in 2006 is enlightening. Lessig argues that due to the inefficiencies of real-space technologies, a large quantity of personal data that is not necessary for the specific authentication purpose is made available to interested entities. For example, if you want to authenticate that you are over 18 in order to purchase alcoholic beverage, you need to show your official identity card to the seller which reveals more data than necessary. By use of a virtual wallet, whichever piece of fact that is necessary could be authenticated without revealing extra-facts about the person (*Lessig* 2006 p.51)
[295] *Black* 2000-2 p.34: Black argues that decentring the regulation "entails displacement of the state from the centre of *the activity of the regulation* and distributively granting that task to the *agents* in the system" which is explained by Benker in *Benkler* 2000 p. 562-563 as individuals becoming "users-participants in the production of their information environment". The regulators in this new legal paradigm are then to move from the *hierarchical*

Figure 2: Murray's Active Dot Matrix

pathetic dots, towards being into active dots. The regulators would thus become involved in helping the users-citizens who are now the self-controllers to make *informed* decisions in this new role.[296]

It is not difficult to see that transition of the hierarchical to heterarchical relationship is tantamount to removal of government and administration from the *conceptual* centre of society"[297] and "*fragmentation* of exercise of power and control".[298] Law, economics, politics and administrations are all self-referentially closed sub-systems which construct information about other systems according to their own limited viewpoints and binary oppositions. Blockchains, with their decentralized architectures, necessarily require a serious shift in the perception of the regulator vis-a-vis the regulatee. At one extreme of the permissionless blockchains, there is no one to make decisions other than the user/data subject, and no one to have any power to enforce any legal sanctions. This novel architecture without a regulator disrupts the traditional meaning of "regulation" as something done with threat of sanctions and redefines it as something that happens in the absence of formal legal sanction. Regulation on a decentralized architecture thus becomes "the product of interactions (webs of influence) and not of the exercise of the formal, constitutionally recognized authority of government."[299] In other words, by the decentralized architecture of the blockchains, the power and control which was once monopolized in the hands of state or state-accredited institutions (e.g. banks) can be divided among newer actors that are autonomous. With all this fragmentation of power, the *autonomy* of social actors should be recognized[300] even of it may be experienced as a

relationship to a *heterarchical* one. The precise nature of the heterarchical role would entail a 'translatory' and 'facilitatory' capacity required by the *deliberations*.

[296] The problems related to the individuals making an informed choice are discussed by *Barocas- Nissenbaum* 2009. *O'Hara - Shadbolt* 2015 also greatly emphasize the importance of the informed choice in an autonomous regulatory context which the consent mechanism is supposed to deliver to the citizens: "The autonomy which is supposedly preserved by the regime is undermined as the choice cannot be seen as informed, or sometimes even uncoerced."

[297] *Black* 2001 p.104

[298] *Ibid* p.108

[299] *Ibid* p.110-111

[300] *Ibid* p.108

compromise of power on the part of the state and regulators.[301] From the perspective of the privacy and personal data protection, such autonomy is manifested through the users being able to perform their own transactions and handle their own personal data on the blockchains. They are going to be able to autonomously give or deny access to their data to the interested parties. A comprehensive analysis of the effect of this sort of autonomy on the regulation of privacy and personal data protection is beyond the scope of this thesis; however, it can confidently be said that internet users are moving in the direction of self-regulation through autonomy by the gradually broadening adoption of blockchains.[302] The regulatory law, in the field of privacy and personal data protection, thus becomes the "reflexive, procedural or post-regulatory law"[303] whereby the regulatory processed is democratized [304] by way of enabling the participation of the subject in her/his own regulation or in the regulation of her/his rights.[305] Indeed, this seems to be the position of other researchers in this field as well as the EU Blockchain Observatory Forum and at least one DPA.[306] From a practical point of view, it is would be in order to see a deeper discussion by and among the European DPAs before even a court ruling is attained. In absence of more official views, it is possible to find blockchain users and operators being stuck in a dilemma of controller versus processor.[307]

[301] *Black* 2000-1 p.610, *Habermas* 1996 p. 406: The material and formal law assumes the state and the citizen are in a *zero-sum game*: the implication of granting autonomy to the citizen is that state's competence is diminished.
[302] *Teubne*r 1993 p. 32-34, and *Black* 2001 p.109
[303] *Black* 2001 p.126
[304] *Black* 2000-1 pp.597-598
[305] *Habermas* 1996
[306] Opinion of the Hungarian Data Protection Authority; EU Blockchain Observatory Forum – Blockchain and the GDPR Report 2018 p. 18; *Finck 2017, Salmensuu 2018*
[307] Wrigley's article in *Corrales and others 2019, p. 233*

5. FINAL REMARKS

Lessig says that "we are not usually trained to think about all the different ways technology could achieve the same ends through different means."[308] This has been an inspiring view for me in the conduct of this research. In a fast-digitalising world through the IoT, AI and blockchains, our first duty as the jurists is to understand the beast we are confronting. With a humble attempt in this work, I have thus shown that the common use of blockchains is not such an impossible ideal in the light of the GDPR; on the contrary, the blockchains can contribute to the making of an all the more empowered consumers/ data subjects with the added security benefits and shift of power in the legal regulatory balance.

Sweeping regulations which are applied indiscriminately to different architectures of ICT may cost the society in the form of *loss of utility*. Resolving conflicts between particular technologies and the right to privacy require "narrow proscriptions focused on well-defined problems."[309] What is absolutely required is to avoid a formalistic approach and adopt a unified, permissive approach from the regulators to encourage innovation, build-up of experience and widespread dissemination of information.[310] EU Blockchain Observatory is, without a doubt, giving hope towards that goal. Further effort should also be made by all DPAs across the EU beyond just one or two.

Remaining within the confines of this thesis has been a tough task. The immediate further research theme would be to delve deeper into understanding the harms which may be caused by the *undatafied* data, which would require a comprehensive qualitative and quantitative research that covers both privacy and cybersecurity aspects.
For now, however, these conclude this thesis.

[308] *Lessig* 2006 p. 32
[309] *Walker* 2002
[310] *Berberich – Steiner* 2016 p. 426; and for an understanding of the instrumentalist approach, see *Popper* 1965 and *Dewey* 1924.