Tunnistettavuuskriteeristä henkilötiedon määritelmässä: EU-oikeudellinen tarkastelu kriteerin merkityksestä ja soveltamisesta tietoyhteiskunnan aikakaudella

Lasse Parkkamäki

Pro gradu -tutkielma

Elokuu 2018

Oikeudellinen argumentaatio

Turun yliopisto, oikeustieteellinen tiedekunta

TURUN YLIOPISTO
Oikeustieteellinen tiedekunta

PARKKAMÄKI, LASSE: Tunnistettavuuskriteeristä henkilötiedon määritelmässä: EU-oikeudellinen tarkastelu kriteerin merkityksestä ja soveltamisesta tietoyhteiskunnan aikakaudella
OTM-tutkielma, XIII + 74 s.
Yleinen oikeustiede
Elokuu 2018
Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin Originality Check -järjestelmällä.

Tutkielmani käsittelee henkilötiedon käsitteeseen sisältyvää tunnistettavuuskriteeriä. Sen mukaan tiedot voidaan katsoa henkilötiedoiksi, jos ne liittyvät tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön. Tunnistettavuuskriteeriä soveltamalla voidaan laajentaa tai supistaa henkilötiedoiksi katsottavien tietojen määrää ja tällä on suoria ja epäsuoria vaikutuksia muun muassa markkinoiden toimivuuteen ja yksilöiden perusoikeuksien ja -vapauksien toteutumiseen. Tunnistettavuuden osalta on olemassa kaksi eri koulukuntaa, jotka ovat eri mieltä siitä, miten tunnistettavuuskriteeriä tulisi arvioida. Oikeuskirjallisuudessa esiintyy vastakkaisia näkemyksiä siitä, onko sen määrittämiseksi, onko henkilö tunnistettavissa, nojauduttava ”objektiiviseen” vai ”suhteelliseen” perusteeseen. Tunnistettavuuskriteerin tulkinta on teknologian kehityksen myötä muuttunut sisällöltään erilaiseksi, kuin mitä se on ollut ennen. Tietokoneet mahdollistavat valtavan datamäärän käsittelemisen ja yhdistelemisen hyvinkin nopeassa tahdissa, jolloin eri tietoja yhdistelemällä voidaan monesti pystyä tunnistamaan yksittäinen henkilö sellaisenkin tiedon perusteella, jota ei perinteisesti ole henkilötiedoksi mielletty.

Tutkielmani noudattaa lainopillista tutkimusmenetelmää. Keskeisimpinä lähteinä työssäni olen käyttänyt EU:n yleistä tietosuoja-asetusta (EU 2016/679), EU:n tietosuojadirektiivin 95/46 29 artiklalla perustetun tietosuojatyöryhmän lausuntoja ja unionin tuomioistuimen tuomiota asiassa Patrick Breyer vastaan Saksan liittotasavalta (Asia C-582/14). Lisäksi olen hyödyntänyt ulkomaisia ja kotimaisia artikkeleita sekä kirjallisuutta.

Tutkielman keskeisenä tutkimustuloksena on, että teknologioiden kehittymisen vuoksi tunnistettavuuskriteerin laajalla tulkinnalla henkilötiedoksi voidaan katsoa lähes mikä tahansa tieto. Tämän vuoksi sääntelyyn sisältyvä joustavuus tulisi hyödyntää tunnistettavuuskriteeriä tulkittaessa, jotta vältetään henkilötiedon käsitteen liiallinen laajentaminen, mutta toteutetaan kuitenkin tietosuojasääntelyn perimmäiset tavoitteet eli luonnollisten henkilöiden perusoikeuksien ja -vapauksien suojelu.

Asiasanat: henkilötieto, tietosuoja, tunnistettavuuskriteeri, profilointi, uudelleentunnistaminen, anonymisointi, salaaminen, pseudonymisointi

Sisällys

Sisällys ... II
Lähteet: ... IV
Lyhenteet: ... XIII
1 Johdanto ... 1
1.1 Taustaa ... 1
1.2 Aihe, rajaukset ja tutkimusmenetelmät ... 5
1.3 Keskeiset käsitteet ... 7
2 Henkilötietojen suojan sääntelystä EU:ssa ... 11
2.1 Direktiivistä asetukseen ... 11
2.1.1 Perusta henkilötietojen suojalle EU:ssa ... 11
2.1.2 Syyt sääntelyn muuttamiseen ... 14
2.2 Henkilötietojen suojan merkitys ja suhde muihin perusoikeuksiin ... 16
2.2.1 Henkilötietojen merkitys markkinoilla ... 16
2.2.2 Henkilötietojen suoja ja muut perusoikeudet ... 18
2.3 Henkilötiedon käsitteen osatekijät ... 21
3 Tunnistettavuuskriteerin sisältö ja tulkinta EU:ssa ... 24
3.1 Johdanto ... 24
3.2 Tulkinnan lähtökohdat ... 27
3.2.1 Lähtökohtana laaja tulkinta ... 27
3.2.2 Suora ja epäsuora tunnistaminen ... 29
3.3 Objektiivinen vai suhteellinen peruste ... 30
3.3.1 Objektiivinen peruste ... 30
3.3.2 Suhteellinen peruste ... 31
3.3.3 Arviointia perusteista ... 32
3.4 EUT:n tulkinta tunnistettavuuskriteeristä ... 36
3.4.1 Johdanto ... 36
3.4.2 Tapaus C-582/14 – Patrick Breyer ... 37
3.4.3 Johtopäätökset kriteerin tulkinnasta EU-oikeuskäytännössä ... 42
3.5 Johtopäätökset kriteerin sisällöstä ja tulkinnasta ... 44
4 Tunnistettavuuskriteerin ongelmallisuudesta kehittyvän teknologian aikakaudella ... 45
4.1 Johdanto ... 45
4.2 Henkilötietojen anonymisointi, salaaminen ja pseudonymisointi ... 46

4.2.1 Anonyymi henkilötieto ja anonymisointitekniikat 46
4.2.2 Henkilötietojen salaamisen asema tietosuojasääntelyssä 50
4.2.3 Salatut henkilötiedot – pseudonymisoitua vai anonyymiä tietoa 51
4.3 Vaikutukset henkilön yksityisyyteen ilman tunnistamista 57
4.3.1 Uudelleentunnistamisen ongelma 58
4.3.2 Profilointi 60
4.3.3 Yksilön v. ryhmän suojaaminen 63
4.4 Riskiperusteinen tunnistettavuuden määrittely 66
5 Johtopäätökset 71

Lähteet:

Kirjallisuus:

Arbuckle, Luk – Moher, Ester - J.Bartlett, Susan – Ahmed, Sara - El Emam, Khaled: Montreal Accord on Patient-Reported Outcomes (PROs) use series – Paper 9: anonymization and ethics considerations for capturing and sharing patient reported outcomes. Journal of Clinical Epidemiology 2017, sivut 168-172.

Bosco, Francesca – Creemers, Niklas – Ferraris, Valeria – Guagnin, Daniel – Koops, Bert-Jaap: Profiling Technologies and Fundamental Rights and Values: Regulatory Challenges and Perspectives from European Data Protection Authorities, teoksessa Gutwirth, Serge – Leenes, Ronald - De Hert, Paul: Reforming European Data Protection Law (Springer 2015).

Borges, Georg & Meents Jan Geert: Cloud Computing. Rechtshandbuch, München 2016.

Borgesius, Frederik Zuiderveen: Singling out people without knowing their names. Behavioural targeting, pseudonymous data, and the new Data Protection Regulation' (2016) 32(2) Computer Law & Security Review s. 256-271.

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

Castells, Manuel: The Rise of The Network Society, The Information Age: Economy, Society and Culture. Toinen, uudistettu painos. Blackwell 2000.

De Montjoye, Yves-Alexandre – Radaelli, Laura – Singh, Vivek Kumar, Pentland, Alex: Unique in the shopping mall: On the reidentifiability of credit card metadata. Science 2015, sivut 536-539.

DLA Piper: EU Study on the Legal analysis of a Single Market for the Information Society, New rules for a new age? The future of online privacy and data protection 2009.

ECKHARDT, J.: Commentary on LG Berlin Ruling of 6 September 2007, K&R 2007.

ECKHARDT, J.: Commentary on AG München Ruling of 30 September 2008, K&R 2008.

Eklund, Mia – Lilja Johanna: Kuluttajien profilointi markkinointitarkoituksiin – henkilötietojen käsittelyn ja tietosuojasääntelyn kehityssuuntia. Defensor Legis N:o 2/2013.

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 (EJLT) 6/2015.

Floridi, Luciano: Open Data, Data Protection, and Group Privacy. Philosophy & Technology, 1/2014, sivut 1–3. Saatavissa: https://doi.org/10.1007/s13347-014-0157-8 (katsottu 2.2.2018).

Gutwirth, Serge – Hildebrandt, Mirelle: Profiling the European Citizen – Cross-Disciplinary Perspectives. Springer Science + Business Media 2008.

Hintze, Mike: Viewing the GDPR Through a De-Identification Lens: A Tool for Clarification and Compliance. Forthcoming, International Data Protection Law. Oxford University Press 2017. Saatavissa: https://ssrn.com/abstract=2909121 tai http://dx.doi.org/10.2139/ssrn.2909121 (katsottu 28.4.2018).

Hirvonen, Ari: Mitkä metodit? Opas oikeustieteen metodologiaan. Yleisen oikeustieteen julkaisuja, Helsinki 2011.

Hilbert, Martin: Big Data for Development: From Information- to Knowledge Societies University of California, Davis 2013. Saatavissa: https://ssrn.com/abstract=2205145 (katsottu 8.7.2018).

Hon, Kuan - Millard, Christopher - Walden, Ian: The Problem of 'Personal Data' in Cloud Computing-What Information is Regulated? The Cloud of Unknowing, part 1. Queen Mary University of London, School of Law Legal Studies, 2011.

Keppeler, Lutz M., “Objektive Theorie“ des Personenbezugs und “berechtigtes Interesse“ als Untergang der Rechtssicherheit? Computer und Recht, 2016, s. 360-367.

Kokott, Juliane – Sobotta, Christoph: The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR. International Data Privacy Law, 4/2013, Sivut 222-228. Saatavissa: https://doi.org/10.1093/idpl/ipt017 (katsottu 5.4.2018).

Lindroos-Hovinheimo, Susanna: Henkilötietojen suoja EU-oikeudessa – yksityisyyttä yhteisön kustannuksella? Lakimies 1/2018 s. 52–75.

Lundevall-Unger, Patrick & Tranvik, Tommy: IP Addresses – Just a Number?, International Journal of Law and Information Technology, 1/2011, s. 53–73. Saatavissa: https://doi.org/10.1093/ijlit/eaq013 (katsottu 6.7.2018).

Massey, Rohan, Anonymisation: managing data protection risk – the new UK code, Computer and Telecommunications Law Review (C.T.L.R.), 2013, 19(3), sivut 86-88.

Mantelero, Alessandro: Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection. Computer Law & Security Review 2016.

Mäenpää, Petri: Algoritmi ostaa ja myy pörssissä. Tiede 5/2011.

Ohm, Paul: Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. UCLA Law Review 2010.

Oostveen, Manon: Identifiability and the Applicability of Data Protection to Big Data. International Data Privacy Law 2016. Saatavissa: https://ssrn.com/abstract=2877692 (katsottu 8.7.2018).

Purtova, Nadezhda: The law of everything. Broad concept of personal data and future of EU data protection law. Law, Innovation and Technology 2018.

Pitkänen, Olli - Tiilikka, Päivi - Warma, Eija: Henkilötietojen suoja. Talentum 2013.

Saarenpää, Ahti: Henkilö- ja persoonallisuusoikeus. Teoksessa: Niemi, Marja-Leena (toim.) Oikeus tänään | Osa II. Rovaniemi: Lapin yliopiston oikeustieteellisiä julkaisuja 2015. s. 203–430.

Salokannel, Marjut: Terveystiedot ja EU:n yleinen tietosuoja-asetus. Defensor Legis N:o 4/2016.

Schwartz, Paul & Solove, Daniel: The PII Problem: Privacy and a New Concept of Personally Identifiable Information. New York University Law Review 2011, s. 1814-1894.

Schwartz, Paul & Solove, Daniel: Reconciling Personal Information in the United States and European Union, Cal. L. Rev. 2014. Saatavissa: http://scholarship.law.berkeley.edu/californialawreview/vol102/iss4/7 (katsottu 9.7.2018).

Spindler, Gerald & Schmechel, Philipp: Personal Data and Encryption in the European General Data Protection Regulation. Journal of Intellectual Property, Information Technology and Electronic Commerce Law, 2/2016.

Skouma, Georgia & Leonard, Laura: On-line Behavioral Tracking: What May Change after the Legal Reform on Personal Data Protection, teoksessa Gutwirth, Serge – Leenes, Ronald - De Hert, Paul: Reforming European Data Protection Law (Springer 2015).

Srinivasa, Srinath: Big Data Analytics: First International Conference, BDA 2012, New Delhi, India, December 24-26, 2012.

Sweeney, Latanya: Simple Demographics Often Identify People Uniquely. Carnegie Mellon University, Data Privacy Working Paper 3. Pittsburgh 2000.

Taylor, L., Floridi, L., van der Sloot, B.: Group Privacy: new challenges of data technologies. Dordrecht: Springer 2017.

Tene, Omer & Polonetsky, Jules. "Privacy in the Age of Big Data: A Time for Big Decisions," Stanford Law Review Online 2012, s. 63-69.

Ukkonen, Esko: Mihin algoritmeja tarvitaan? Tieteessä tapahtuu Vol 21 Nro 7/2003, s. 19–22.

Urgessa, Worku Gedefa: The Protective Capacity of the Criterion of Identifiability under EU Data Protection Law, 2 Eur. Data Prot. L. Rev. 521 (2016).

Vanto, Jarno J.: Henkilötietolaki käytännössä. Alma Talent, 2011.

Van der Sloot, Bart – Broeders, Dennis – Schrijvers, Erik: Exploring the Boundaries of Big Data. Amsterdam University Press, Amsterdam 2016.

Yakowitz, Jane: Tragedy of the Data Commons. Harvard Journal of Law & Technology 2011.

Zwitter, Andrej: Big Data ethics. Big Data & Society 2014.

Virallislähteet:

Euroopan komission ehdotus Euroopan parlamentin ja neuvoston asetukseksi: yksilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta (yleinen tietosuoja-asetus). COM(2012) 11 Final, annettu 25.1.2012. Saatavissa:

http://register.consilium.europa.eu/doc/srv?f=ST+5853+2012+INIT&l=fi (katsottu 8.2.2018).

Euroopan komission muutettu ehdotus neuvoston direktiiviksi yksilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta. Bryssel 15.10.1992. COM (92) 422 final, 28.10.1992. Saatavissa: http://aei.pitt.edu/10375/1/10375.pdf (katsottu 11.4.2018).

Euroopan parlamentin kanta, vahvistettu ensimmäisessä käsittelyssä 12. maaliskuuta 2014, Euroopan parlamentin ja neuvoston asetuksen (EU) antamiseksi yksilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta (yleinen tietosuoja-asetus).

European Union Agency for Fundamental Rights: Handbook on European data protection law (Council of Europe 2014).

Julkisasiamiehen ratkaisuehdotus Manuel Campos Sánchez-Bordona 17.1.2018 (1) Asia C-650/16A/S Bevola ja Jens W. Trock ApS vastaan Skatteministeriet.

Julkisasiamiehen ratkaisuehdotus Manuel Campos Sánchez-Bordona 12.5.2016. Asia C‑582/14. ECLI:EU:C:2016:339.

Julkisasiamiehen ratkaisuehdotus 25 päivänä kesäkuuta
2013. Asia C-131/12 Google Spain SL ja Google Inc. vastaan Agencia Española de Protección de Datos (AEPD) ja Mario Costeja González).

WP 169. Working Party, tietosuojatyöryhmä, joka on perustettu direktiivin 95/46/EY 29 artiklalla. Lausunto 1/2010 rekisterinpitäjän ja henkilötietojen käsittelijän käsitteistä. Saatavissa:
http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2010/wp169_fi.pdf
(katsottu 8.2.2018).

WP 136. Working Party, tietosuojatyöryhmä, joka on perustettu direktiivin 95/46/EY 29 artiklalla. Opinion 4/2007 on the concept of personal data.

WP 216. Working Party, tietosuojatyöryhmä, joka on perustettu direktiivin 95/46/EY 29 artiklalla. Lausunto 5/2014 anonymisointitekniikoista.

WP 188. Working Party, tietosuojatyöryhmä, joka on perustettu direktiivin 95/46/EY 29 artiklalla. Opinion 16/2011 on EASA/IAB Best Practice Recommendation on Online Behavioural Advertising.

WP 171. Working Party, tietosuojatyöryhmä, joka on perustettu direktiivin 95/46/EY 29 artiklalla. Opinion 2/2010 on online behavioural advertising.

Internet-lähteet:

Barbaro, Michael & Zeller Jr., Tom: A Face Is Exposed for AOL Searcher No. 4417749. N.Y. Times, 9.8.2006. Saatavissa: https://www.nytimes.com/2006/08/09/technology/09aol.html (katsottu 9.7.2018).

Bäck, Asta - Keränen, Janne: Anonymisointipalvelut – Tarve ja toteutusvaihtoehdot. Liikenne- ja viestintäministeriö 7/2017. Saatavissa: http://urn.fi/URN:ISBN:978-952-243-503-3 (katsottu 4.3.2018).

Evans, Dave: Top 25 Technology Predictions 2009. Saatavissa: https://www.cisco.com/c/dam/en_us/about/ac79/docs/Top_25_Predictions_121409rev.pdf (katsottu 16.2.2018).

ENISA, Privacy by design in big data - An overview of privacy enhancing technologies in the era of big data analytics, 2015. Saatavissa: https://www.enisa.europa.eu/publications/big-data-protection/at_download/fullReport (katsottu 25.3.2018).

Euroopan komission tietosuojayksikkö: Usein kysytyt kysymykset koskien henkilötietojen siirtoa EU/ETA-alueelta kolmansiin maihin. Saatavissa: http://ec.europa.eu/justice/policies/privacy/docs/international_transfers_faq/international_transfers_faq.pdf (katsottu 23.1.2018).

Information Commissioner's Office (UK): Personal Information Online Code of Practice, 2010. Saatavissa: https://ico.org.uk/media/for-organisations/documents/1591/personal_information_online_cop.pdf (katsottu 8.7.2018).

Kuneva, Meglena (European Consumer Commissioner): Keynote Speech (31 March 2009) SPEECH/09/156. Saatavissa: http://europa.eu/rapid/press-release_SPEECH-09-156_en.htm (katsottu 8.4.2018).

World Economic Forum, 'Personal Data: the Emergence of New Asset Class' (2011). Saatavissa: http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf (katsottu 8.4.2018).

Säädökset:

EU:n yleinen tietosuoja-asetus (EUROOPAN PARLAMENTIN JA NEUVOSTON ASETUS (EU) 2016/679, annettu 27 päivänä huhtikuuta 2016, luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta).

EU:n tietosuojadirektiivi (EUROOPAN PARLAMENTIN JA NEUVOSTON DIREKTIIVI 95/46/EY annettu 24 päivänä lokakuuta 1995, yksilöiden suojelusta henkilötietojen käsittelyssä ja näiden tietojen vapaasta liikkuvuudesta).

Euroopan unionin toiminnasta tehdyn sopimuksen konsolidoitu toisinto (C 326/47).

Euroopan unionin perusoikeuskirja (2012). EUVL C 326, 26.10.2012.

Yleissopimus ihmisoikeuksien ja perusvapauksien suojaamiseksi (Euroopan ihmisoikeussopimus, SopS 18-19/1990).

Oikeustapaukset:

C-62/90 Commission v Germany [1992] ECR I2575.

Amann v Switzerland, no. 27798/95, ECHR 2000-II.

Common Services Agency v Scottish Information Commissioner (Scotland) [2008] UKHL 47, [2008] 1 WLR 1550 ('CSA'), [27].

Euroopan unionin tuomioistuin, yhdistetyt tapaukset C-402/05 P ja C-415/05 P Kadi & Al Barakaat International Foundation v Council & Commission, [2008] ECR I-6351.

Google Spain SL, Google Inc v Agencia Española de Protección de Datos and Mario Costeja González (C-131/12) [2014] ECLI:EU:C:2014:317.

Scarlet Extended (C-70/10).

Patrick Breyer v Saksan liittotasavalta (C-582/14) (19 päivänä lokakuuta 2016) ECLI:EU:C:2016:779.

136/79 National Panasonic v Commission [1980] ECR 2033.

Rotaru v Romania [GC] App no 28341/95, ECHR 2000-V.

Lyhenteet:

Asetus	EU:n yleinen tietosuoja-asetus
EIS	Euroopan ihmisoikeussopimus
EIT	Euroopan ihmisoikeustuomioistuin
EU	Euroopan unioni
EUT	Euroopan unionin tuomioistuin
IP-osoite	internetin protokollaosoite
IT	informaatioteknologia
Perusoikeuskirja	Euroopan unionin perusoikeuskirja
SEUT	Euroopan unionin toiminnasta tehty sopimus
Tietosuoja-asetus	EU:n yleinen tietosuoja-asetus
Tietosuojadirektiivi	Euroopan parlamentin ja neuvoston direktiivi 95/46/EC
Tietosuojatyöryhmä	Working Party, tietosuojatyöryhmä, joka on perustettu direktiivin 95/46/EY 29 artiklalla.

1 Johdanto

1.1 Taustaa

Henkilötietojen merkitystä nykyisessä tietotekniikan ja internetin hallitsemassa talouselämässä on kuvattu mm. termein uusi öljy[1] ja jälkiteollisen yhteiskunnan mahdollisuuksien luoja[2]. Henkilötiedot luovat uusia mahdollisuuksia talouskasvuun ja yhteiskunnallisten arvojen luomiseen. Henkilötietoja keräämällä kykenemme paremmin ymmärtämään ja ennustamaan mihin ihmiset kohdistavat huomiotaan ja toimintaansa niin yksilötasolla kuin ryhmissä ja maailmanlaajuisesti.[3] Kun tiedot on kerätty ja analysoitu, yritykset käyttävät niitä yksilöllisten palvelujen tarjoamisen tukena.[4] Myös esimerkiksi viranomaiset ja muu julkinen sektori hyödyntävät henkilötietoja tehokkaampien ja toimivampien palvelujen tarjoamiseksi.[5]

Tietokoneiden mahdollistama suurten tietomäärien kerääminen ja käsitteleminen mahdollistavat monia uudenlaisia innovaatioita ja helpottavat monella tapaa ihmisten arkielämää. Nykyaikana yhteiskunta rakentuu pitkälti tietotekniikan varaan ja uusia automatisoituja prosesseja ja järjestelmiä ryhdytään matalalla kynnyksellä hyödyntämään niin liike-elämässä, teollisuudessa kuin viranomaistoiminnassakin. Teknologian kehitys luo kuitenkin myös uudenlaisia ongelmia ja haasteita.

Yksi tietotekniikan kehityksen mukana muuttuva käsite on henkilötiedon käsite. Erityisesti henkilötiedon määritelmään sisältyvän tunnistettavuuskriteerin tulkinta on teknologian kehityksen myötä muuttunut sisällöltään erilaiseksi, kuin mitä se on ollut ennen. Tietokoneet mahdollistavat valtavan datamäärän käsittelemisen ja yhdistelemisen hyvinkin nopeassa tahdissa, jolloin eri tietoja yhdistelemällä voidaan monesti pystyä tunnistamaan yksittäinen henkilö sellaisenkin tiedon perusteella, jota ei perinteisesti ole henkilötiedoksi mielletty.

---

[1] Kuneva 2009.
[2] World Economic Forum 2011, s. 5.
[3] Ibid.
[4] Tutkimuksissa on osoitettu, kuinka kerättävän tiedon avulla tapahtuva päätöksenteko yrityksissä parantaa tuottavuutta. Ks. esim. Erik Brynjolfsson, Lorin Hitt ja Heekyung Kim: 'Strength in Numbers: How Does Data-Driven Decision-Making Affect Firm Performance?' (2011).
[5] World Economic Forum 2011, s. 5.

Tunnistettavuuskriteerin tulkinnan oikeudellinen arviointi on tarpeellista, koska kriteerin avulla voidaan laajentaa tai supistaa henkilötiedoiksi katsottavien tietojen määrää ja tällä on suoria ja epäsuoria vaikutuksia muun muassa markkinoiden toimivuuteen ja yksilöiden perusoikeuksien ja -vapauksien toteutumiseen. Tunnistettavuuden käsitteen liiallinen laajentaminen voi merkittävästi heikentää sisämarkkinoiden toimivuutta ja tietojen vapaata liikkuvuutta. Erityisesti kriteerin aineellisen soveltamisalan osalta vallitseva epävarmuus voi heikentää oikeusvarmuuden toteutumista ja siten vaikeuttaa tietoturvakäytäntöjen käyttöönottoa. Tunnistettavuuskriteerin liian laajalla tulkinnalla voidaan itse asiassa heikentää yksilöiden perusoikeuksien ja -vapauksien toteutumista. Lisäksi tunnistettavuuskriteerin avulla ei kaikissa tilanteissa onnistuta saattamaan tietosuojasääntelyn piiriin kaikkia sellaisia tietoja, joilla tosiasiassa voi olla vaikutuksia luonnollisten henkilöiden perusoikeuksiin ja -vapauksiin.

Yksinkertaisimmillaan henkilötiedoksi voidaan katsoa esimerkiksi henkilötunnus tai sormenjälki, koska ne voidaan yksiselitteisesti tunnistaa yhtä luonnollista henkilöä koskeviksi tiedoiksi.[6] Henkilötiedon määritelmän mukaan on kuitenkin mahdollistaa luokitella lähes mikä tahansa tieto henkilötiedoksi. Henkilötiedoksi voidaan katsoa esimerkiksi auton rekisterikilpi, internetselailuhistoria, sähköpostiosoite tai tietokoneen IP-osoite (internetin protokollaosoite).[7] Olennaista on, että henkilö pystytään näistä tiedoista tunnistamaan.[8]

EU:n tietosuojatyöryhmä[9] on katsonut, että sen määrittämisessä, onko henkilö tunnistettavissa, olisi otettava huomioon kaikki kohtuullisesti toteutettavissa olevat keinot, joita rekisterinpitäjä tai joku muu voi kyseisen henkilön tunnistamiseksi käyttää, mutta pelkkä hypoteettinen mahdollisuus

[6] Vanto 2011, s. 22.
[7] Vanto 2011, s. 22-23.
[8] Saarenpää 2015, s. 326.
[9] Working Party, tietosuojatyöryhmä, joka on perustettu direktiivin 95/46/EY 29 artiklalla. Tietosuojatyöryhmä on lakkautettu tietosuoja-asetuksella ja sen tilalle on perustettu asetuksen 68 artiklan mukainen Euroopan tietosuojaneuvosto. Käytännössä sen tehtävät säilyvät lähes samankaltaisina kuin aikaisemminkin. Asetuksen 70 (1) artiklan mukaan tietosuojaneuvosto varmistaa, että asetusta sovelletaan yhdenmukaisesti ja antaa tätä tarkoitusta varten mm. neuvoja, suosituksia ja suuntaviivoja henkilötietojen suojaan liittyvissä kysymyksissä.

tunnistamiseen ei riittäisi täyttämään kriteeriä.[10] Henkilötiedon käsite tulisi ymmärtää mahdollisimman laajasti.[11] EU:n tietosuojatyöryhmä on todennut, että henkilötieto koostuu neljästä osatekijästä: 1) kaikenlaisista kielellisistä merkityksistä, jotka 2) koskevat 3) tunnistettavaa 4) luonnollista henkilöä[12]. Vaikka henkilötiedolla on edellä mainitut neljä tunnusmerkkiä, näyttäisi tunnistaminen yleensä olevan ratkaisevassa asemassa ja tulkinnalliset erimielisyydet oikeustieteellisessä tutkimuksessa ja oikeuskäytännössä liittyvät useimmiten siihen. Tulkinnanvaraisia kysymyksiä ovat esimerkiksi, miten tunnistettavuutta tulisi arvioida tietojenkäsittelyn ulkoistustilanteissa ja missä menevät tunnistettavuuden rajat eli milloin henkilö ei enää ole tunnistettavissa tietyn tiedon tai tietojen yhdistelmän perusteella.

Tunnistettavuuden osalta on olemassa kaksi eri koulukuntaa, jotka ovat eri mieltä siitä, miten tunnistettavuuskriteeriä tulisi arvioida. Oikeuskirjallisuudessa esiintyy vastakkaisia näkemyksiä siitä, onko sen määrittämiseksi, onko henkilö tunnistettavissa, nojauduttava *objektiiviseen* vai *suhteelliseen* perusteeseen. Euroopan unionin tuomioistuin (EUT) on arvioinut kyseisiä perusteita ratkaisuissaan, joista viimeisin ja paljon keskustelua herättänyt on ennakkoratkaisu asiassa Patrick Breyer v Bundesrepublik Deutschland (C-582/14). Näiden perusteiden arviointi tulee konkreettisesti tarpeelliseksi erityisesti tilanteissa, joissa pohditaan niin sanotun epäsuoran tunnistamisen ulottuvuutta. Tällöin joudutaan arvioimaan sitä, kenen hallussa olevien tietojen perusteella henkilön tulee olla tunnistettavissa, jotta tiedot voidaan katsoa henkilötiedoksi.

Tämän rajauksen tekeminen on tärkeää, koska yritykset ulkoistavat yhä useammin osia toiminnoistaan ulkoisille palveluntarjoajille. Monesti esimerkiksi yrityksen tietoliikenne tai palkkahallinto ulkoistetaan toiselle yritykselle kustannusten karsimiseksi. Yrityksen toimintojen ulkoistaminen on tullut entistä houkuttelevammaksi pilvipalveluiden myötä, sillä ne tarjoavat merkittäviä tehokkuusetuja palveluiden käyttäjille. Ulkoistusta harkitsevien yritysten kannalta

[10] Tietosuojatyöryhmän lausunto 04/2007, s. 15.
[11] Vanto 2011, s. 22.
[12] WP 136. Opinion 4/2007 on the concept of personal data, s. 6.

on tärkeää tietää, miten tunnistettavuutta arvioidaan suhteessa kolmansiin tahoihin, jotta voidaan varmistua tietoturvatoimenpiteiden riittävyydestä.

Euroopan unionin yleisen tietosuoja-asetuksen[13] noudattamista on tehostettu säätämällä hallinnollisesta sakosta, joka voidaan määrätä maksettavaksi asetusta rikkovalle toimijalle. Sakko on taloudellisesti huomattava sanktio, joka yritysten kohdalla pohjautuu organisaation maailmanlaajuiseen liikevaihtoon. Asetuksen tultua sovellettavaksi yritykset todennäköisesti panostavat tietosuojaan entistä enemmän, koska asetuksessa on lisätty uusia velvollisuuksia rekisterinpitäjille ja uusia oikeuksia rekisteröidyille. Markkinoiden toimivuuden kannalta on kuitenkin ensiarvoisen tärkeää, että tunnistettavuuskriteerin sisältö olisi oikeudellisesti selkeä, jotta tietoja käsittelevät toimijat voisivat suunnitella toimintaansa ja asetuksen tavoitteet voisivat toteutua.

EU:n tietosuojanormisto ja sitä koskevat tulkintakannanotot perustuvat pitkälti binääriseen eli kaksijakoiseen lähestymistapaan tunnistettavuudesta. Tiedot ovat joko henkilötietoja ja siten tietosuojalainsäädännön soveltamisalaan kuuluvia, tai anonyymejä, joihin sääntelyä ei ole sovellettava lainkaan. Direktiivillä 95/46/EY perustetun tietosuojatyöryhmän anonymisointitekniikoita koskevasta lausunnosta 05/2014 käy selvästi ilmi, että tietojen anonymisoinnin vaatimukset ovat erittäin suuret. Anonymisoinnin on oltava "peruuttamatonta" ja tiedot on säilytettävä muodossa, jossa rekisteröidyn tunnistaminen "ei ole enää mahdollista".

Tämä kaksijakoinen lähestymistapa voi kuitenkin johtaa epäoptimaalisiin tuloksiin. Esimerkiksi organisaatiolla, joka käyttää toiminnassaan henkilötietoja tarkoituksiin, joita ei voida saavuttaa täydellisesti anonyymeillä tiedoilla, on hyvin vähän kannustimia toteuttaa anonymisointi- tai muita tietoturvatekniikoita. Siksi tietoja on saatettu tähän asti säilyttää täysin tunnistettavassa tilassa, vaikkakin olisi ollut mahdollista toteuttaa joitakin tietoturvatekniikoita säilyttäen tietojen yhteensopivuuden organisaation toiminnan kannalta. Tällä olisi tietenkin myös yksityisyyden suojaa parantavia vaikutuksia. Siten kaksijakoinen lähestymistapa

[13] Euroopan parlamentin ja neuvoston asetus (EU) 2016/679, annettu 27 päivänä huhtikuuta 2016, luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta.

tunnistamiseen voi johtaa siihen, että tietoja suojataan tosiasiassa vähemmän kuin olisi mahdollista.[14]

Tietosuoja-asetuksessa on huomioitu kahtiajaon ongelma ainakin osittain, säätämällä uudesta tunnistamisen ”välimallista” jota kutsutaan pseudonymisoinniksi. Pseudonymisoinnilla tarkoitetaan henkilötietojen käsittelemistä siten, että henkilötietoja ei voida enää yhdistää tiettyyn rekisteröityyn käyttämättä lisätietoja, edellyttäen että tällaiset lisätiedot säilytetään erillään ja niihin sovelletaan teknisiä ja organisatorisia suojatoimenpiteitä. Kuitenkin, pseudonymisoidut tiedot säilyttävät luonteensa henkilötietoina. Tietosuoja-asetuksessa siis säilyy kuitenkin kahtiajako, mutta pseudonymisoinnilla voidaan helpottaa organisaation muiden asetuksen edellyttämien velvoitteiden noudattamista.

Jokaisella on omat näkemyksensä siitä, mitä kaikkea yksityisyyteen kuuluu ja mitä tietoja on pidettävä niin henkilökohtaisina, että niitä tulee suojata. Lindroos-Hovinheimon mukaan rajanvedon tekemiseen suojattavien ja ei-suojattavien tietojen välillä on kaksi vaihtoehtoa. Ensinnäkin yksityisyys voidaan ymmärtää puhtaasti yksilön omien arvostusten ja subjektiivisten mieltymysten mukaisesti. Toiseksi, yksityisyyden mittapuuksi voidaan ottaa yhteisössä jaetut ja yleisesti hyväksytyt käsitykset yksityisyyden laajuudesta ja sisällöstä. Henkilön autonomista persoonaa suojataan, mutta henkilö ei itse saa päättää, miten ja missä määrin niitä suojataan.[15] Tietosuoja-asetuksella ja EU-tuomioistuimen linjauksilla luodaan uutta oikeutta ja täsmennetään sitä, missä henkilön yksityisyyden ja tietosuojan rajat menevät.

1.2 Aihe, rajaukset ja tutkimusmenetelmät

Tutkielmani aiheena on henkilötiedon määritelmään sisältyvä tunnistettavuuskriteeri. Tutkielmani tavoitteena on selvittää, miten tunnistettavuuskriteeriä tällä hetkellä tulkitaan EU:ssa ja millaisia ongelmia kriteerin soveltaminen voi aiheuttaa tietoyhteiskunnan aikakaudella. Tarkastelen

[14] Hintze 2017, s.2.
[15] Lindroos-Hovinheimo, s. 64.

aihetta kehittyvän teknologian ja tietoyhteiskunnan näkökulmasta käsin ja tuon esiin niitä ongelmia, joita nykyinen tunnistettavuuden määritelmä kohtaa uusissa ympäristöissä ja tilanteissa. Kootusti tutkimuskysymykseni ovat 1) *onko tietosuojasääntelyn rakentuminen pitkälti tunnistettavuuden varaan riittävä turvaamaan luonnollisten henkilöiden perusoikeuksia- ja vapauksia*, 2) *millaisia ongelmia tunnistettavuuskriteerin soveltaminen kohtaa tietoyhteiskunnan aikakaudella ja* 3) *miten tunnistettavuuskriteeriä olisi tulkittava, jotta parhaiten turvattaisiin luonnollisten henkilöiden perusoikeudet ja -vapaudet laajentamatta henkilötiedon käsitteen aineellista soveltamisalaa liiaksi*?

Tutkielman tavoitteena on tulkita ja systematisoida uuden tietosuoja-asetuksen, oikeuskäytännön ja oikeuskirjallisuuden avulla vallitsevaa oikeustilaa tunnistettavuuskriteeristä. Tutkielman lähestymistapa on pääosin oikeusdogmaattinen eli lainopillinen lähestymistapa, koska se tarjoaa mahdollisuuden vallitsevan oikeuden sisällön tulkintaan sekä oikeuskäytännössä ja kirjallisuudessa esiin nousseiden argumenttien ja tulkintakriteerien systematisoimiseen. Lainopillinen lähestymistapa tutkielmassa tarkoittaa tutkielman aiheen tarkastelua uuden tietosuoja-asetuksen myötä pätevien oikeusnormien valossa.

Tutkielmassa tulkitaan ja jäsennetään vallitsevaa oikeustilaa ja tuotetaan siten tieteellistä tietoa tarkasteltavista oikeusnormeista. Tarkastelu luo tulkintakannanottoja voimassa olevista oikeusnormeista ja välittää niiden sisältöä lukijalle.[16] Lainopillinen tarkastelu tutkielmassa keskittyy oikeussääntöjen tutkimisen lisäksi oikeusperiaatteiden ja niiden yhteensovittamisen analysointiin sekä oikeusopillisten konstruktioiden tulkintaan. Tarkastelulla pyritään yhtenäistämään ja johdonmukaistamaan tarkasteltavaa aiheitta. Tutkielma pyrkii välittämään lukijalle oikeusnormien, oikeusperiaatteiden ja juridisten konstruktioiden tulkintaa ja systematisointi siten, että tulkintakannanottoja voitaisiin hyödyntää myös käytännön toiminnassa.

Lainoppi antaa kuvauksen voimassa olevasta oikeudesta, mutta lainoppi ei ole kuitenkaan puhdas kuvaus oikeudesta, vaan sen tuottama sisältö on voimassa

[16] Hirvonen 2011, s. 22–25.

olevan oikeuden merkityssisältöä koskeva kannanotto siitä, kuinka oikeutta pitäisi tulkita, punnita ja systematisoida.[17] Tutkielman tarkoituksena onkin lisäksi eritellä niitä ongelmakohtia, joita sisältyy nykyiseen tapaan hahmottaa tunnistettavuuskriteeri ja henkilötiedon käsite ylipäänsä. Esitän myös aihetta koskevia tulkintakannanottoja siitä, kuinka tarkasteltavia käsitteitä olisi tulkittava, jotta saavutettaisiin tietosuojasääntelyn tavoitteet ja riittävä tasapaino henkilötietojen suojan ja muiden perusoikeuksien välillä.

Tutkielmani etenee siten, että luvussa 2 käyn läpi EU:n tietosuojasääntelyn taustaa ja henkilötietojen suojan suhdetta muihin perusoikeuksiin. Luvussa 3 tarkastelen tunnistettavuuskriteerin sisältöä ja tulkintaa. Luvussa 4 käyn läpi ongelmia ja haasteita, joita kehittyvä teknologia asettaa tunnistettavuuskriteerin tulkinnalle. Luvussa 5 esitän johtopäätökset tunnistettavuuskriteerin tämän hetkisestä EU-oikeudellisesta sisällöstä sekä suositukset sille, miten tunnistettavuuskriteeriä tulisi tulkita EU:ssa, jotta tietosuojasääntelyn tavoitteet voisivat toteutua mahdollisimman tehokkaasti ja muut perusoikeudet ja -vapaudet huomioiden.

1.3 Keskeiset käsitteet

Henkilötiedoilla tarkoitetaan tietosuoja-asetuksen 4 (1) artiklan mukaan *”kaikkia tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön (...) liittyviä tietoja*”. Viittauksella kaikkiin tietoihin on tavoiteltu henkilötietolainsäädännön laajaa soveltamisalaa.[18] Yksinkertaisimmillaan henkilötietoja ovat esimerkiksi henkilötunnus tai sormenjälki, koska ne voidaan selvästi tunnistaa yksittäistä luonnollista henkilöä koskeviksi. Henkilötiedon laajan määritelmän mukaan on kuitenkin mahdollista luokitella lähes mikä tahansa tieto henkilötiedoksi. Henkilötiedoksi voidaan katsoa esimerkiksi tietokoneen IP-osoite, auton rekisterikilpi, internetin selailuhistoria tai sähköpostiosoite.[19] Olennaista on, että henkilö pystytään näistä tiedoista tunnistamaan.

[17] Hirvonen 2011, s. 22–25.
[18] Euroopan komission tietosuojayksikkö. Usein kysytyt kysymykset koskien henkilötietojen siirtoa EU/ETA-alueelta kolmansiin maihin, s. 8.
[19] Vanto 2011, s. 22-23.

Henkilötiedon määritelmä on asetuksessa lähes saman sisältöinen kuin tietosuojadirektiivissä. Määritelmään on kuitenkin lisätty joitakin kohtia, jotka saattavat muodostua merkityksellisiksi erityisesti internetissä tapahtuvan asioinnin yhteydessä. Esimerkiksi asetuksen 4 (1) artiklassa verkkotunnistetiedot sisällytetään nimenomaisesti tunnistetiedoksi. Tämän seurauksena monista internetsivustoilla käytettävistä evästeistä tulee henkilötietoja.

Varsinaista tulkintatukea henkilötiedon määritelmä tietosuoja-asetuksessa ei anna sille, tuleeko tunnistettavuuskriteeriä tarkastella objektiivisen vai suhteellisen perusteen avulla. Tietosuoja-asetuksessa nimittäin säädetään tietosuojadirektiivin tavoin tunnistettavuudesta siten, että *”tunnistettavissa olevana pidetään luonnollista henkilöä, joka voidaan suoraan tai epäsuorasti tunnistaa erityisesti tunnistetietojen […] tai yhden tai useamman hänelle tunnusomaisen […] tekijän perusteella”*. Tunnistettavuuden määritelmä on siis jätetty sisällöltään melko avoimeksi ja sitä voidaan täsmentää oikeuskäytännön kautta kuten tähänkin asti.

Rekisterinpitäjällä tarkoitetaan asetuksen 4 (7) artiklan mukaan ”*luonnollista henkilöä tai oikeushenkilöä, viranomaista, virastoa tai muuta elintä, joka yksin tai yhdessä toisten kanssa määrittelee henkilötietojen käsittelyn tarkoitukset ja keinot*”. Rekisterinpitäjä voi olla esimerkiksi yhtiö, joka ylläpitää rekisteriä työntekijöistään tai asiakkaistaan. Tärkeää on tosiasiallinen määräysvalta.[20] Lausunnossaan 1/2010 (WP 169) tietosuojatyöryhmä katsoo, että rekisterinpitäjän käsitteen avulla halutaan jakaa vastuuta sinne, missä henkilötietojenkäsittelyyn tosiasiallisesti vaikutetaan.[21] Rekisterinpitäjän määritelmä on asetuksessa olennaisilta osiltaan sama kuin tietosuojadirektiivissä. Jokainen toimija, joka on rekisterinpitäjä tietosuojadirektiivin mukaan, todennäköisesti tulee olemaan sitä myös tietosuoja-asetuksen mukaan.

[20] Esimerkiksi Euroopan unionin tuomioistuimen julkisasiamies on katsonut ratkaisuehdotuksessaan asiassa C-131/12 (Google), että hakukoneyhtiö ei ole rekisterinpitäjä käsittelemiensä internetsivuilla olevien henkilötietojen osalta. Vaikka yhtiö tarjoaa tiedon paikannusvälineen, se ei tarkoita sitä, että se voisi määrätä sivullisten internetsivuilla olevan tiedon sisällöstä.

[21] Tietosuojatyöryhmän lausunto 1/2010, s. 9. Tietosuojatyöryhmän mukaan rekisterinpitäjän käsite perustuu pikemminkin faktapohjaiseen kuin muodolliseen analyysiin.

Myös *henkilötietojen käsittelijän* määritelmä on asetuksessa pysynyt samana kuin direktiivissä. Tietosuoja-asetuksen 4 (8) artiklan mukaan henkilötietojen käsittelijällä tarkoitetaan ”*luonnollista henkilöä tai oikeushenkilöä, viranomaista, virastoa tai muuta elintä, joka käsittelee henkilötietoja rekisterinpitäjän lukuun*”. Ulkoistettaessa tietojenkäsittelytoimintoja, kuten esimerkiksi yrityksen tietoliikenteen tai palkkahallinnon, toimeksiannon vastaanottavasta osapuolesta tulee henkilötietojen käsittelijä. Jos rekisterinpitäjän lukuun henkilötietoja käsittelevä osapuoli ylittää toimeksiantoon perustuvat valtuutensa ja ryhtyykin käsittelemään tietoja muihin tarkoituksiin kuin on sovittu, henkilötietoja käsittelevästä osapuolesta tulee tältä osin rekisterinpitäjä.[22]

Käsittelyn määritelmä on laaja ja se kattaa käytännössä kaikki mahdolliset henkilötietoihin kohdistuvat toimenpiteet. Tietosuoja-asetuksen 4 (2) artiklan mukaan käsittelyllä tarkoitetaan ”*toimintoa tai toimintoja, joita kohdistetaan henkilötietoihin tai henkilötietoja sisältäviin tietojoukkoihin joko automaattista tietojenkäsittelyä käyttäen tai manuaalisesti, kuten tietojen keräämistä, tallentamista, järjestämistä, jäsentämistä, säilyttämistä, muokkaamista tai muuttamista, hakua, kyselyä, käyttöä, tietojen luovuttamista siirtämällä, levittämällä tai asettamalla ne muutoin saataville, tietojen yhteensovittamista tai yhdistämistä, rajoittamista, poistamista tai tuhoamista*”. Käsittelyä voi yksinkertaisimmillaan olla esimerkiksi henkilötietojen säilyttäminen, kerääminen tai poistaminen. Käsittelyn määritelmä on merkittävä, koska sen mukaisesti EU:n yleinen tietosuoja-asetus tulee todennäköisesti sovellettavaksi aina, kun organisaatio tekee jotain, joka liittyy tai vaikuttaa henkilötietoihin.

Tietosuoja-asetuksessa ei ole määritelty *algoritmin* käsitettä. Tietojenkäsittelytieteessä algoritmi on perinteisesti ollut keskeinen termi ja sillä viitataan eräänlaiseen toimintaohjeeseen, jota tarvitaan tietokoneiden ohjelmointiin. Algoritmit muodostuvat äärellisistä ja täsmällisistä laskenta-askeleista ja niistä muodostuu kokonaisuus, joka kykenee tuottamaan annetuista syöttötiedoista halutun tuloksen.[23] Näiden toimintaohjeiden käyttöala on jatkuvasti laajentunut ja niiden vaikutus on havaittavissa nykypäivänä monilla eri elämän osa-alueilla. Algoritmit pystyvät käsittelemään ja järjestelmään tietoa

[22] Tietosuojatyöryhmän lausunto 1/2010, s. 25.
[23] Ukkonen 2003, s. 19.

sellaisella tavalla ja nopeudella, joihin ihmisten kapasiteetti ei riitä. Algoritmeja hyödyntämällä yritysten toiminnan tehokkuus voi moninkertaistua ja uudet algoritmiset innovaatiot valtaavat jatkuvasti markkinoita. Algoritmeihin liittyvät taloudelliset intressit ovat selkeästi nähtävillä ja ne ohjaavakin yhteiskuntamme toimintaa.[24]

Tietosuoja-asetuksen 4 (4) artiklan määritelmän mukaan *profiloinnilla* tarkoitetaan *"[…] mitä tahansa henkilötietojen automaattista käsittelyä, jossa henkilötietoja käyttämällä arvioidaan luonnollisen henkilön tiettyjä henkilökohtaisia ominaisuuksia, erityisesti analysoidaan tai ennakoidaan piirteitä, jotka liittyvät kyseisen luonnollisen henkilön työsuoritukseen, taloudelliseen tilanteeseen, terveyteen, henkilökohtaisiin mieltymyksiin, kiinnostuksen kohteisiin, luotettavuuteen, käyttäytymiseen, sijaintiin tai liikkeisiin*". Profiloinniksi voidaan kutsua muutakin kuin henkilötietojen automaattista käsittelyä, mutta tietosuojakontekstissa merkittävää on nimenomaisesti tällainen käsittely.[25] Tietosuoja-asetus soveltuu siis vain sellaiseen profilointiin, joka tapahtuu automaattisen tietojenkäsittelyn keinoin. Jotta kyse olisi profiloinnista, toiminnan on sisällettävä sekä automaattista henkilötietojen käsittelyä että henkilön ominaisuuksien arviointia.

Tiedon louhinta (eng. data mining) on prosessi, jossa laajan tietomassan avulla yhdistetään eri tietoja toisiinsa algoritmien avulla. Näin luodaan arvioita ja hypoteeseja esimerkiksi tulevasta käyttäytymisestä jo tapahtuneen toiminnan perusteella.[26]

Anonymisointi tarkoittaa prosessia, jossa data-aineistoa muokataan siten, etteivät yksilöt ole siitä tunnistettavissa edes hyödyntämällä epäsuoria tunnisteita tai taustatietoa. Anonymisoidulla datalla tarkoitetaan yleensä mikrodataa, jolloin yhtä kohdetta koskeva tieto on muusta aineistosta erotettavissa olevana yksikkönä. Tällaista dataa voidaan käyttää data-analyysien tekemiseen. Jos data-aineistosta käy ilmi vain tilastollisia tunnuslukuja ja jakaumatietoja, yksilöiden erottaminen ei ole mahdollista, mutta aineiston

[24] Mäenpää 2011.
[25] Eklund & Lilja 2013.
[26] Esim. Gutwirth – Hildebrandt 2008, s.18.

hyödyntämismahdollisuudet analyyseihin ovat rajalliset.[27] Tietosuoja-asetuksen tietosuojaperiaatteita ei sovelleta anonyymeihin tietoihin, mutta epäselvää on edelleen, millaiset tiedot voidaan luokitella anonyymeiksi ja missä olosuhteissa.

2 Henkilötietojen suojan sääntelystä EU:ssa

2.1 Direktiivistä asetukseen

2.1.1 Perusta henkilötietojen suojalle EU:ssa

Perus- ja ihmisoikeuksien suojelu, joihin myös henkilötietojen suoja kuuluu, on EU:ssa toteutettu kahden erillisen, mutta toisiinsa liittyvän järjestelmän avulla. Ensimmäinen on Euroopan ihmisoikeussopimus, joka on kansainvälinen sopimus ja johon on liittynyt kaikki 47 Euroopan neuvoston jäsenvaltiota. Ihmisoikeussopimukseen ovat liittyneet EU:n jäsenvaltioiden lisäksi esimerkiksi Sveitsi, Venäjä ja Turkki. Sopimuksen noudattamista valvoo Euroopan ihmisoikeustuomioistuin (EIT), mihin jokainen sopimusvaltio tai yksityinen henkilö voi valittaa, jos kokee ihmisoikeussopimuksessa turvattujen oikeuksiensa tulleen loukatuksi.

Toinen järjestelmä perustuu Euroopan unionin tuomioistuimen oikeuskäytäntöön, joka takaa perustavanlaatuisten ihmisoikeuksien suojelun EU:ssa. Näiden oikeuksien kunnioittaminen on osa EU: n perustuslaillisia periaatteita.[28] Aluksi EU:n tuomioistuin on kehittänyt nämä oikeudet EU:n oikeuden yleisinä periaatteina samassa linjassa ihmisoikeussopimuksen kanssa, mutta nykyisin useimmat oikeudet ovat määriteltyinä Euroopan unionin perusoikeuskirjassa. Molemmat järjestelmät liittyvät läheisesti toisiinsa ja nykyään perusoikeuskirjan tulkinta seuraa ihmisoikeussopimuksen tulkintaa[29], vaikka EU ei vieläkään ole osapuolena ihmisoikeussopimuksessa.

[27] Liikenne- ja viestintäministeriö 2017, s. 6.
[28] Euroopan unionin tuomioistuin, yhdistetyt tapaukset C-402/05 P ja C-415/05 P Kadi & Al Barakaat International Foundation v Council & Commission, [2008] ECR I-6351, kohta. 285.
[29] Perusoikeuskirjan 52 artiklan 3 kohdassa säädetään, että siltä osin kuin tämän perusoikeuskirjan oikeudet vastaavat ihmisoikeuksien ja perusvapauksien suojaamiseksi tehdyssä yleissopimuksessa taattuja oikeuksia, niiden merkitys ja ulottuvuus ovat samat kuin

Yksityisyys ja henkilötietojen suoja ovat molemmat tunnustettuina perustuslaillisella tasolla Euroopassa. Sekä kansallisissa perustuslaeissa että Euroopan ihmisoikeussopimuksessa ja perusoikeuskirjassa yksityisyyden suojaa koskeva säännös. Ihmisoikeussopimuksen 8 artiklassa ja perusoikeuskirjan 7 artiklassa säädetään, että jokaisella on oikeus nauttia yksityis- ja perhe-elämäänsä, kotiinsa ja kirjeenvaihtoonsa kohdistuvaa kunnioitusta. Lisäksi oikeutta yksityiselämän kunnioittamiseen on suojeltu ja edelleen suojellaan EU:n lainsäädännön yleisenä periaatteena.[30]

Perusoikeuskirjan 8 artiklassa säädetään erityisesti perusoikeudesta henkilötietojen suojaan. Ihmisoikeussopimuksessa ei ole vastaavaa tietosuojaa koskevaa määräystä. Ihmisoikeustuomioistuin on kuitenkin soveltanut Euroopan ihmisoikeussopimuksen 8 artiklaa antaakseen henkilötiedoille nimenomaista oikeussuojaa.[31]

Perusta henkilötietojen suojalle löytyy Euroopan unionin toiminnasta tehdyn sopimuksen (SEUT) 16 artiklan 1 kohdasta sekä Euroopan unionin perusoikeuskirjan 8 artiklasta. SEUT 16 artiklan 1 kohdan mukaisesti jokaisella on oikeus henkilötietojensa suojaan. EU:n perusoikeuskirjan 8 artiklan 1 kohta on sanamuodoltaan vastaava SEUT 16 artiklan kanssa. Artiklan 2 kohdassa säädetään tarkemmin henkilötietojen käsittelystä: ”*Tällaisten tietojen käsittelyn on oltava asianmukaista ja sen on tapahduttava tiettyä tarkoitusta varten ja asianomaisen henkilön suostumuksella tai muun laissa säädetyn oikeuttavan perusteen nojalla. Jokaisella on oikeus tutustua niihin tietoihin, joita hänestä on kerätty, ja saada ne oikaistuksi.*” Perusoikeuskirjan 8 artiklan 3 kohdan mukaan riippumaton viranomainen valvoo näiden sääntöjen noudattamista.

Perusoikeuskirjan 8 artiklassa siis paitsi erotetaan henkilötietojen suoja yksityisyydestä, lisäksi artiklan kohdissa 2 ja 3 asetetaan erityisiä vaatimuksia

mainitussa yleissopimuksessa. Tämä määräys ei kuitenkaan estä unionia myöntämästä tätä laajempaa suojaa. Näin ollen ihmisoikeussopimuksessa määritellään ne vähimmäisvaatimukset, jotka on taattava perusoikeuskirjan soveltamisen yhteydessä.

[30] Euroopan unionin tuomioistuin, tapaus National Panasonic v Commission, kohta 17 ja tapaus Commission v Germany, kohta. 23.

[31] EIT: Amann v Switzerland, kohta 65 ja Rotaru v Romania, kohta 43.

tällaisten tietojen käsittelylle. Perusoikeuskirjan yksityisyyttä ja tietosuojaa koskevista säännöksistä voidaan nähdä, että nämä kaksi oikeutta eivät ole täysin samankaltaisia.[32] Oikeus henkilötietojen suojaan on siis vahvasti eriytymässä omaksi perusoikeudekseen.

Henkilötietojen suojan erottaminen yksityiselämän suojan piiristä itsenäiseksi oikeudeksi on monella tapaa merkityksellistä. Niiden tarjoama suoja on paljolti erilainen, vaikkakin niiden välillä on myös yhtäläisyyksiä. Yksityis- ja perhe-elämän suoja takaa luonnollisille henkilöille oikeuden yksityiselämän kunnioittamiseen ja estää viranomaisia puuttumasta tämän suojan piiriin kuuluviin asioihin. Henkilötietojen suojalla suojataan myös henkilön yksityisyyttä, mutta eri tavalla. Henkilötietojen suoja koskee sitä tietoa, jota luonnollisesta henkilöstä on kerätty ja suoja ulottuu myös tiedon käyttötapoihin. Aikaisemmin henkilötietojen suojan on katsottu sisältävän lähinnä oikeuden olla rauhassa muilta. Nykyään henkilöä koskevat tiedot itsessään nähdään niin arvokkaana, että niitä suojataan mihin tahansa tiedot kulkevat ja riippumatta henkilön omista subjektiivisista näkemyksistä tiedon tärkeydestä.

Tällä hetkellä henkilötietojen suoja on murrosvaiheessa. Euroopan unionin yleistä tietosuoja-asetusta on alettu soveltaa kaikissa EU:n jäsenvaltioissa 25.5.2018. Varsinkin yritykset ovat nyt entistä kiinnostuneempia tietosuojaan liittyvistä velvoitteista, koska tietosuoja-asetuksen rikkomisesta tuomittavat hallinnolliset sakot ovat rahallisesti huomattavia eikä organisaatioilla ole varaa kärsiä imagotappioita riittämättömän tietosuojan vuoksi. Tietosuoja-asetuksessa on kiinnitetty erityistä huomiota rekisteröidyn oikeuksiin ja näitä oikeuksia ollaankin asetuksessa vahvistettu ja jopa luotu uusia oikeuksia, kuten oikeus siirtää tiedot tietojärjestelmästä toiseen (asetuksen 20 artikla) ja oikeus olla joutumatta automatisoitujen päätösten kohteeksi (asetuksen 11 artikla).

Kaiken henkilötietojen suojaa koskevan sääntelyn keskiössä on kuitenkin henkilötiedon käsite, koska sen avulla rajataan henkilötietojen suojaa koskevan sääntelyn soveltamisalaa. Normit tulevat sovellettavaksi vain, mikäli yksittäinen käsitelty tieto on henkilötieto ja jos se ei sitä ole, tietosuojanormisto ei sovellu

[32] Kokott & Sobotta, s. 223.

käsittelyyn. Henkilötietojen suojan sääntelyn tavoitteena on luonnollisten henkilöiden perusoikeuksien ja –vapauksien suojelu ja heidän oikeus henkilötietojensa suojaan[33]. Sääntelyllä tavoitellaan kuitenkin myös sisämarkkinoiden toimivuuden parantamista erityisesti digitaalitalouden osalta. Lisäksi tavoitteena mainitaan oikeusvarmuuden ja luottamuksen vahvistaminen käytännön toiminnan sujuvuuteen.[34] Näiden tavoitteiden toteutuminen on paljolti riippuvainen siitä, mikä tieto ylipäänsä katsotaan henkilötiedoksi.

2.1.2 Syyt sääntelyn muuttamiseen

Aikaisemmin EU:n alueella tietosuojasääntely perustui vuoden 1995 tietosuojadirektiiviin[35]. Direktiivissä säädettiin niistä periaatteista ja velvollisuuksista, jotka jäsenvaltioiden on implementoitava omaan lainsäädäntöönsä. Direktiivi ei ole lainsäädäntöinstrumenttina jäsenvaltioissa suoraan sovellettavaa oikeutta, vaan jäsenvaltioiden on saatettava direktiivin vaatimukset osaksi kansallista lainsäädäntöään. Direktiivissä säädetään ne yleiset tavoitteet, jotka jäsenvaltioiden tulee saavuttaa, mutta jäsenvaltiot voivat laajalti itse päättää niistä keinoista, joilla tavoitteet saavutetaan. Koska jäsenvaltioilla on suuri harkintavalta käytettävistä keinoista, eri maiden lainsäädäntöihin voi muodostua merkittäviä eroavaisuuksia. Näin on käynyt myös tietosuojadirektiivin käyttöönotossa. Direktiivillä ei ole onnistuttu estämään tietosuojantäytäntöönpanon hajanaisuutta ja oikeudellista epävarmuutta. Jäsenvaltioiden väliset eroavuudet henkilötietojen käsittelyssä voivat muodostua esteeksi unionin taloudelliselle toiminnalle, vääristää kilpailua ja estää viranomaisia suorittamasta unionin oikeuden mukaisia velvollisuuksiaan.[36]

Unionin tasolla ollaan siis havahduttu siihen, että direktiivi ei ole oikea instrumentti säätää tietosuojasta. Sillä on saavutettu henkilötietojen suojan

[33] Tietosuoja-asetus 1 (2) artikla.
[34] Tietosuoja-asetuksen johdanto-osan kohta 7. Asetusten johdanto-osat eivät ole juridisesti sitovia, mutta ne otetaan kuitenkin huomioon tulkittaessa artikloja. Johdanto-osan kappaleet voivat kutenkin tulkintavaikutuksen lisäksi tietyissä olosuhteissa antaa yksittäisille henkilöille aihetta oikeutettuihin odotuksiin tai toisaalta rajoittaa tällaisia odotuksia. Ks. esim. Euroopan unionin neuvoston pääsihteeristö, Lingvistijuristiyksikkö: Euroopan unionin neuvoston säädöskäsikirja. Neljäs laitos, 2002 s. 83.
[35] Euroopan parlamentin ja neuvoston direktiivi 95/46/EY annettu 24 päivänä lokakuuta 1995, yksilöiden suojelusta henkilötietojen käsittelyssä ja näiden tietojen vapaasta liikkuvuudesta
[36] Tietosuoja-asetuksen johdanto-osan kohta 9.

korkea taso, mutta sääntelyvälineenä se on aiheuttanut odottamattomia esteitä markkinoille. Euroopan komissio onkin todennut, että toimivin säädöstyyppi henkilötietojen suojan sääntelyyn unionissa on asetus.[37]

Tietosuoja-asetus on jäsenvaltioissa suoraan sovellettavaa oikeutta eikä jäsenvaltioiden tarvitse implementoida sitä osaksi kansallista lainsäädäntöä. Asetuksen odotetaan harmonisoivan tietosuojasääntelyä ja helpottavan markkinoiden toimintaa unionin alueella. Yhtenäisen tietosuojasääntelyn uskotaan poistavan monikansallisten yritysten tarpeen suunnitella ja toteuttaa erilaisia prosesseja ja käytänteitä täyttääkseen tietosuojavaatimukset eri Euroopan maissa. Asetuksen on tarkoitus luoda johdonmukaisempi tietosuojakehys niin, että digitaalitalous voi kehittyä koko sisämarkkinoiden alueella ja yksilöt voivat valvoa omia tietojaan. ”*Näin myös vahvistetaan oikeusvarmuutta ja luottamusta käytännön toiminnan sujuvuuteen talouden toimijoiden ja viranomaisten kannalta*.”[38]

Tietosuojauudistuksen toteuttamiseen on vaikuttanut myös se, että henkilötietoja koskevalta sääntelyltä vaaditaan nykyisin eri asioita kuin tietosuojadirektiivin säätämishetkellä. Sääntely oli siis pitkälti vanhentunutta eikä enää soveltunut nykyisiin tietojen käyttöympäristöihin. Yhteiskunta on muuttunut aidosti verkottuneeksi yhteiskunnaksi, jossa tietoverkko on vakiintunut keskeiseksi arkipäivän toimintojen väylästöksi.[39]

Aikaisemmin ihmiset ovat osanneet ja kyenneet melko hyvin hallitsemaan omaa yksityisyyttään ja sosiaalisia roolejaan erilaisissa tilanteissa, kuten julkisilla paikoilla, töissä ja tavatessa ystäviään. Tämä on ollut helppoa siksi, että keskustellessaan toisen henkilön kanssa on voinut samalla nähdä, ketkä muut kuulevat tai näkevät keskustelijat. Tilanne on kuitenkin muuttunut internetissä toimivien sosiaalisten verkostojen myötä, joihin kuuluminen on nykyään osa monen ihmisen identiteettiä.[40]

---

[37] Euroopan komission ehdotus Euroopan parlamentin ja neuvoston asetukseksi (yleinen tietosuoja-asetus), s. 6.
[38] Euroopan komission ehdotus Euroopan parlamentin ja neuvoston asetukseksi (yleinen tietosuoja-asetus), s. 2.
[39] Castells 2000.
[40] Pitkänen – Tiilikka - Warma 2013, s. 4.

Internetissä toimivien sosiaalisen median palveluiden ja sovellusten avulla mahdollisuus tietojen levittämiseen on helpottunut huomattavasti. Samaan aikaan ihmisten kontrolli omien tietojen levittämisen osalta on hyvin vaihtelevaa palvelusta riippuen. Jossakin määrin henkilö voi itse määrittää millaista tietoa hän jakaa, milloin hän haluaa tietojansa poistettavan tai kenelle hän haluaa tietojaan jakaa. Kuitenkin internetin välityksellä toimivissa kommunikaatiopalveluissa tapahtuva viestintä tekee haastavaksi ymmärtää sitä, ketkä voivat tosiasiassa seurata tätä viestintää ja mihin tarkoituksiin. Monien sosiaalisen median palveluiden käyttäminen edellyttää käyttöehtojen hyväksymistä, jolloin henkilö antaa palvelun tarjoajalle laajat oikeudet käyttää henkilön jakamia tietoja eri tarkoituksiin. Internetin käytön kasvaminen, uusien teknologioiden käyttöön ottaminen ja digitalisoituminen ovat luoneet uusia tietosuojariskejä, joita palveluja käyttävät henkilöt eivät aina edes ymmärrä.

Tietosuojasääntelyn soveltamisympäristöön ovat vaikuttaneet myös internetin räjähdysmäisesti lisääntyneen käytön luomat uudenlaiset tietosuojaongelmat, joita ei ole voitu ottaa huomioon tietosuojadirektiiviä säädettäessä. Henkilötietojen keräämisen ja käyttämisen tapa on muuttunut teknologian kehittymisen ja globalisaation vuoksi perinpohjaisesti.[41] Tämä taas on luonut aivan uudenlaisia oikeudellisia ongelmia. Internet on lisännyt mahdollisuuksia viestintään ja siten laajentanut myös sananvapauden käyttämisen mahdollisuuksia. Kuitenkin uudenlaisten tietosuojauhkien vuoksi, mikäli riittävää tietosuojaa luonnollisille henkilöille ei taata, saattaa tietojen kerääminen pahimmassa tapauksessa rajoittaa ihmisten halukkuutta käyttää tällaisia palveluita. Uudenlaisten tietosuojauhkien ja vaikeammin ymmärrettävien käyttöympäristöjen vuoksi tietosuojasääntelyn uudistaminen on ollut välttämätöntä.

2.2 Henkilötietojen suojan merkitys ja suhde muihin perusoikeuksiin

2.2.1 Henkilötietojen merkitys markkinoilla

[41] Pitkänen – Tiilikka - Warma 2013, s. 6.

Yhä useamman yrityksen ydinliiketoiminta perustuu erilaisten tietojen keräämiseen ja käsittelyyn. Hyvin usein ainakin osa näistä kerätyistä tiedoista on luokiteltavissa henkilötiedoiksi. Myös julkishallinto kerää ihmisistä enenevissä määrin tietoa erilaisiin rekistereihin talletettavaksi. Myös ihmisten oma rooli aktiivisena tiedon tuottajana on merkittävä. Sosiaalisessa mediassa ja mobiilisovelluksissa jaettava tiedon määrä kasvaa jatkuvasti, eikä ihmisten yksityiselämääkään koskevien tietojen jakamiseen vaikuta olevan yhtä suurta kynnystä kuin aikaisemmin, sillä tietojen jakaminen internetin välityksellä on helppoa ja nopeaa.

Henkilötietojen käsittely tulee todennäköisesti tulevina vuosina mullistumaan entisestään. Esimerkiksi teknologiayhtiö CISCO on ennusteissaan[42] arvioinut, että vuonna 2020 keskivertoihminen tulee ylläpitämään 130 terabittiä tietoa, kun vastaava määrä vuonna 2009 oli 128 gigabittiä.

Kun yhä useamman yrityksen liiketoiminta tällä tavoin perustuu erilaisten tietojen käsittelyyn, on rekisterinpitäjinä toimivien organisaatioiden kannalta ensiarvoisen tärkeää tietää yleisen tietosuoja-asetuksen aineellisen soveltamisalan laajuus. Toisin sanoen rekisterinpitäjien on tiedettävä, onko heidän kulloinkin käsittelemänsä tiedot luokiteltavissa henkilötiedoksi vai vapauttaako esimerkiksi tietojen salaaminen salausavaimella heidät tietosuojasääntelyn asettamilta velvoitteilta.[43] Tietosuoja-asetuksen myötä asian merkitys korostuu, sillä sääntely vaikuttaa uudella tavalla myös sellaisiin rekisterinpitäjiin, joiden kotipaikka ei ole EU:n alueella. Tietosuoja-asetuksen alueellista soveltamisalaa koskevassa 3 (1) artiklassa nimittäin säädetään, että *”[…] asetusta sovelletaan henkilötietojen käsittelyyn, jota suoritetaan unionin alueella sijaitsevassa rekisterinpitäjän tai henkilötietojen käsittelijän toimipaikassa toiminnan yhteydessä, riippumatta siitä, suoritetaanko käsittely unionin alueella vai ei*”. Esimerkiksi henkilötietojen käsittelyn ulkoistustilanteissa tiedot saattavat helposti siirtyä käsiteltäväksi unionin alueelle, vaikka rekisterinpitäjä itse toimisi unionin ulkopuolelta käsin.

Lisäksi tietosuoja-asetuksen 3 (2) artiklassa säädetään, että asetusta sovelletaan ”*[…]henkilötietojen käsittelyyn, milloin käsittelytoimet liittyvät tavaroiden tai*

[42] Evans 2009.
[43] Spindler & Schmechel, s. 164.

palvelujen tarjoamiseen unionissa tai rekisteröityjen käyttäytymisen seurantaan, kunhan tämä käyttäytyminen tapahtuu unionin alueella”. ”*Seurannalla*” tarkoitetaan esimerkiksi rekisteröidyn toimien seurantaa internetin välityksellä, minkä perusteella yksilöstä voidaan luoda profiili erityisesti häntä koskevien päätösten tekemistä varten tai hänen henkilökohtaisten mieltymystensä, käyttäytymisensä ja asenteidensa analysointia tai ennakoimista varten.[44]

Yleisen tietosuoja-asetuksen laaja alueellinen soveltamisala siis velvoittaa monissa tapauksissa myös EU:n ulkopuolelle sijoittautuneet rekisteripitäjät huomioimaan uudet vaatimukset omassa toiminnassaan, kun ne käsittelevät henkilötietoja. Tietojen salauksen ja muiden teknologioiden, joiden avulla henkilötietojen määrä voidaan minimoida ja välttää tietosuoja-asetuksen soveltaminen, merkitys tulee kasvamaan entisestään.[45] Vastaavasti rekisterinpitäjiä siis tulee entistä enemmän kiinnostamaan se, mikä ylipäätään katsotaan henkilötiedoksi ja millaisin perustein tämä rajanveto ratkaistaan.

2.2.2 Henkilötietojen suoja ja muut perusoikeudet

Tietosuoja-asetuksen mukaan henkilötietojen käsittely olisi suunniteltava niin, että se palvelee ihmistä. Oikeus henkilötietojen suojaan ei ole absoluuttinen, vaan sitä on tarkasteltava suhteessa sen tehtävään yhteiskunnassa ja sen on suhteellisuusperiaatteen mukaisesti oltava oikeassa suhteessa muihin perusoikeuksiin. Tietosuoja-asetuksessa kunnioitetaan kaikkia perusoikeuksia ja otetaan huomioon perusoikeuskirjassa tunnustetut vapaudet ja periaatteet sellaisina kuin ne ovat vahvistettuina perussopimuksissa.[46] Tietosuoja-asetuksen tavoite ei ole ainoastaan suojata luonnollisia henkilöitä vaan myös mahdollistaa henkilötietojen vapaa liikkuvuus toimivien sisämarkkinoiden luomiseksi.[47] Yksilöiden ja tietoja käsittelevien tahojen etujen välille on siis löydettävä tasapaino, jotta tavoitteet voivat toteutua.

[44] Tietosuoja-asetuksen johdanto-osan kohta 24.
[45] Spindler & Schmechel, s. 165.
[46] Tietosuoja-asetuksen johdanto-osan kohta 4.
[47] Tietosuoja-asetuksen johdanto-osan kohta 166. Myös direktiivissä samat tavoitteet 1 (2) artiklassa ja johdanto-osan kohdassa 3 (5).

Unionin oikeudessa on selvästi havaittavissa pyrkimys korostaa yksityisyyden suojaa suhteessa yhteisöllisiin arvoihin. Tarkoituksena on luoda yksilöille suoja, jonka avulla heidän ei tarvitse sietää toisia yhteisön jäseniä tai muita tahoja henkilökohtaisen sfäärinsä sisällä.[48] Unionin tuomioistuin on korottanut henkilötietojen suojan perusoikeusasemaa EIT:tä selkeämmin ja aiemmin. Esimerkiksi jo vuonna 2010 unionin tuomioistuin argumentoi Bavarian Lager -tapauksessa[49], että yksityisyys voi sivuuttaa unionin hallinnon läpinäkyvyyden tarpeet. Viimeaikaisissa EU-tuomioistuimen ratkaisuissa on todistettavissa yhä aktiivisempi ote sen suhteen, miten yksityisen ja yleisen edun välinen raja tulisi hahmottaa.[50]

Euroopan unionin julkisasiamies on useasti ollut yksilön suojan laajentamista vastaan ja argumentoinut yhteisöllisten etujen puolesta. Julkisasiamies lausui yleisestä oikeudesta tulla unohdetuksi seuraavaa:

> ”[D]irektiivissä ei säädetä yleisestä oikeudesta tulla unohdetuksi siten, että rekisteröidyllä olisi oikeus rajoittaa itselleen haitallisiksi katsomiensa tai omien intressiensä vastaisten henkilötietojen levittämistä tai lopettaa niiden levittäminen, ja yhdyn tähän näkemykseen.”[51]

Lausunnossaan julkisasiamies siis katsoo, että yksilöiden määräysvaltaa tiedoistaan on tietosuojadirektiivissä rajoitettu ja yksilöiden oikeutta henkilötietojensa hallintaan tulisi tulkita suppeasti. Hänen mukaansa henkilötietojen suojan taso ja laajuus eivät ole riippuvaisia esimerkiksi yksilöiden subjektiivisista mieltymyksistä:

> ”Käsiteltäessä tietoja ilman asianomaisen suostumusta on sovellettava perusteita, jotka koskevat käsittelyn tarkoitusta ja siihen liittyviä intressejä verrattuina rekisteröidyn intresseihin, eikä rekisteröidyn subjektiivisiin mieltymyksiin liittyviä perusteita. Subjektiiviset mieltymykset eivät ole direktiivin 14 artiklan a alakohdassa tarkoitettuja huomattavan tärkeitä ja perusteltuja syitä.[52]”

Lausunnossaan julkisasiamies myös korosti niitä etuja, joita yhteiskunnalle koituu tietojen vapaasta ja helposta saatavuudesta:

> ”Tiedonvälityksen vapaudelle on mielestäni annettava erityinen suoja unionin lainsäädännössä erityisesti, kun otetaan huomioon se, että muualla

[48] Lindroos-Hovinheimo 2018, s. 63.
[49] C-28/08, Commission v. Bavarian Lager Co.
[50] Lindroos-Hovinheimo 2018, s. 65.
[51] Julkisasiamiehen lausunto asiassa C-131/12 Google Spain, kohta 108.
[52] Ibid.

> maailmassa autoritaariset järjestelmät rajoittavat yhä enemmän internetiin pääsyä tai sensuroivat sen kautta välitettäviä tietoja.[53]”

Julkisasiamiehen mukaan henkilötietojen suojaa ei olisi tullut laajentaa entisestään, koska tällöin jouduttaisiin uhraamaan muita keskeisiä oikeuksia ja vapauksia.[54] Tuomioistuin päätyi ratkaisussaan kuitenkin erilaiseen näkemykseen. Tuomion perusteluissa todetaan, että yksilön henkilötietojen suoja on tässä tapauksessa painavampi perusoikeus kuin yrityksen elinkeinonharjoittamisen vapaus tai yleisön tiedonsaantioikeus.

Tuomiossa korostetaan, että tietosuojadirektiiviä on tulkittava osana perusoikeuksien suojaa, kuten jo aiemmassa oikeuskäytännössä oli linjattu. Henkilötietojen käsittelyä, joka saattaa loukata perusoikeuksia ja varsinkin oikeutta yksityisyyteen, on välttämättä tulkittava perusoikeuksien valossa, sillä nämä perusoikeudet ovat vakiintuneen oikeuskäytännön mukaan erottamaton osa yleisiä oikeusperiaatteita, joiden noudattamista unionin tuomioistuin valvoo ja jotka sisältyvät nykyään perusoikeuskirjaan.[55] Tuomion perusteluista suuri osa keskittyy korostamaan henkilötietojen suojan merkitystä. Yksityisyyden suoja menee tuomioistuimen perustelujen mukaan ainakin jossakin määrin muiden arvojen edelle.

Perinteisesti henkilötietojen suojan on katsottu olevan osa yksityisyyden suojaa. Henkilötietojen suojan tavoitteena on tämän näkökannan mukaan yksinomaisesti suojella niitä samoja arvoja, joita yksityisyyden suoja jo valmiiksi kattaa. Juliane Kokott ja Christoph Sobotta ovat kuitenkin artikkelissaan[56] osoittaneet, että näiden oikeuksien erottaminen EU:n perusoikeuskirjassa ei ole pelkästään symbolinen, vaan oikeuksien sisällöt poikkeavat toisistaan. Euroopan ihmisoikeustuomioistuimen ja EU:n tuomioistuimen ratkaisukäytäntöä tarkastelemalla, näiden molempien oikeuksien laajuus ja rajoittaminen eroavat toisistaan useilla tavoilla, huolimatta merkittävistä päällekkäisyyksistä.[57]

[53] Julkisasiamiehen lausunto asiassa C-131/12 Google Spain, kohta 121.
[54] Julkisasiamiehen lausunto asiassa C-131/12 Google Spain, kohta 133.
[55] EUT:n tuomio asiassa C-131/12 Google Spain SL, Google Inc v Agencia Española de Protección de Datos ja Mario Costeja González, kohta 68.
[56] Kokott & Sobotta 2013: The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR.
[57] Kokott & Sobotta 2013, s. 222.

Kokott ja Sobotta esittävät, että yksityisyyden suoja ja henkilötietojen suoja liittyvät läheisesti toisiinsa Euroopan ihmisoikeustuomioistuimen ja Euroopan unionin tuomioistuimen oikeuskäytännössä, mutta niitä ei pidä pitää samanlaisina.

Oikeuksien soveltamisalaan liittyy paljon päällekkäisyyksiä, mutta ne myös eroavat toisistaan monissa tilanteissa niin henkilöllisen kuin aineellisenkin soveltamisalansa osalta. Lisäksi näiden oikeuksien yhteisvaikutuksella voidaan nähdä olevan positiivisia vaikutuksia perusoikeuksien toteutumisen kannalta. Nimittäin vaatimus siitä, että henkilötietoja on käsiteltävä oikeudenmukaisesti ja vain tiettyä tarkoitusta varten, kattaa monia sellaisia tilanteita, joissa yksityisyyteen puuttuminen on kyettävä oikeudellisesti perustelemaan. Nämä henkilötietojen suojaa koskevat käsittelyn vaatimukset auttavat kiinnittämään huomiotamme niihin tekijöihin, jotka ovat alttiita vaikuttamaan perusoikeuksiimme.[58]

2.3 Henkilötiedon käsitteen osatekijät

Henkilötiedoista säädetään tietosuoja-asetuksen 4 (1) artiklassa seuraavasti:

> *”[H]enkilötiedoilla* [tarkoitetaan] kaikkia tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön, jäljempänä ’rekisteröity’, liittyviä tietoja; tunnistettavissa olevana pidetään luonnollista henkilöä, joka voidaan suoraan tai epäsuorasti tunnistaa erityisesti tunnistetietojen, kuten nimen, henkilötunnuksen, sijaintitiedon, verkkotunnistetietojen taikka yhden tai useamman hänelle tunnusomaisen fyysisen, fysiologisen, geneettisen, psyykkisen, taloudellisen, kulttuurillisen tai sosiaalisen tekijän perusteella.”

Artiklan perusteella voidaan todeta, että *tunnistettu* ja *tunnistettavissa oleva* ovat samassa asemassa eli toisin sanoen tiedot ovat henkilötietoja, jos luonnollinen henkilö *tunnistetaan* tai *on tunnistettavissa* näistä tiedoista. Jotkut ovat nähneet ongelmalliseksi, että tunnistetun ja tunnistettavissa olevan käsitteiden sisältöä ja oikeusvaikutuksia ei ole eriytetty.[59]

[58] Kokott & Sobotta 2013, s. 228.
[59] Esim. Schwartz & Solove 2011, s. 1874.

Tietosuojatyöryhmä on katsonut lausunnossaan[60], että henkilötieto koostuu neljästä osatekijästä: 1) kaikki tiedot 2) liittyä 3) tunnistettuun tai tunnistettavissa olevaan 4) luonnolliseen henkilöön. Kaikkien osatekijöiden yhtäaikainen olemassaolo on välttämätöntä, jotta tieto voidaan katsoa henkilötiedoksi. Kaikkia näitä osatekijöitä käytetään henkilötiedon käsitteen soveltamisalan laajentamiseen, erityisesti viittaamalla epäsuoraan tunnistamiseen. Käsitteen laaja tulkinta aiheuttaa herkästi ongelmia varsinkin tietoverkoissa tapahtuvan henkilötietojen käsittelyn osalta.[61] Käsittelen tässä lyhyesti kaikkien henkilötiedon käsitteen osatekijöiden ydinsisällön tunnistettavuuskriteeriä lukuun ottamatta. Muut osatekijät ovat tärkeässä asemassa myös tunnistettavuuskriteerin kannalta, sillä niiden tulkintakäytäntö ohjaa osittain myös tunnistettavuuden tulkintaa, usein laajentavaan suuntaan.

Ensimmäinen osatekijä ”*kaikkia tietoja*” viittaa selkeästi EU:n lainsäätäjän tarkoitukseen määritellä henkilötiedon käsite mahdollisimman laajasti. Tietojen *luonteen* osalta henkilötietojen käsitteeseen sisältyvät kaikenlaiset henkilöä koskevat lausunnot. Se kattaa objektiiviset tiedot, kuten tiedon tietyn aineen olemassaolosta henkilön veressä verikokeiden tuloksissa. Lisäksi se kattaa subjektiiviset tiedot, kuten mielipiteet tai arvioinnit. Subjektiiviset tiedot kattavat merkittävän osan tiedoista, joita käsitellään esimerkiksi pankkialalla, kun arvioidaan henkilön luotettavuutta tai työelämässä arvioitaessa työntekijän työsuoritusta.[62] Myös objektiivisia tietoja hyödynnetään kattavasti liiketoiminnassa, esimerkiksi kun kerätään tietoa henkilön mielenkiinnon kohteista tai aktiivisuudesta eri verkkosivuilla.

Tietojen ei tarvitse olla tosia tai oikeaksi todistettuja ollakseen henkilötietoja.[63] Itse asiassa tietosuojasääntelyyn sisältyy jo mahdollisuus, että tietojen ollessa virheellisiä rekisteröidyllä on oikeus tutustua kyseisiin tietoihin ja vaatia niitä korjattavaksi asianmukaisilla korjaustoimenpiteillä.[64]

Tietojen *sisällön* osalta henkilötietojen käsitteeseen sisältyvät tiedot, jotka tarjoavat mitä tahansa informaatiota. Tämä kattaa paitsi arkaluonteiseksi

[60] Tietosuojatyöryhmän lausunto 04/2007.
[61] DLA Piper 2009, s. 18.
[62] Tietosuojatyöryhmän lausunto 04/2007 s. 6.
[63] Ibid.
[64] Tietosuoja-asetuksen 15 ja 16 artiklat.

katsottavat tiedot, myös yleisemmäksi katsottavat tiedot. Tiedot voivat sisältää esimerkiksi informaatiota siitä, minkä tyyppistä toimintaa yksittäinen henkilö tekee, kuten työsuhteita tai yksilön taloudellista tai sosiaalista käyttäytymistä koskevaa tietoa.[65] Tietosuojatyöryhmän lausunnon perusteella voidaankin todeta, että *kaikki on tai ainakin sisältää tietoa*.[66]

Toinen osatekijä ”*liittyä*” on tietosuojatyöryhmän mukaan ratkaisevan tärkeä, koska sen avulla voidaan selvittää, mitkä ovat ne linkit tai liitynnät, joilla on asian kannalta merkitystä ja miten ne voidaan erottaa muunlaisista liitynnöistä. Yleisesti ottaen tietoja voidaan pitää yksilöön ”liittyvänä”, kun ne koskevat kyseistä henkilöä. Monissa tilanteissa tämä suhde voidaan helposti todeta. Tällaisia ovat esimerkiksi tiedot potilaan lääkärintarkastuksesta, jotka sisältyvät hänen potilasrekisteriinsä tai kuva henkilöstä, joka on kuvattu kyseisen henkilön videohaastattelussa. Joissakin tilanteissa ei ole aina yhtä itsestään selvää, että tiedot "liittyvät" yksilöön. Joissakin tilanteissa tiedoista saatava informaatio koskee ensi vaiheessa esineitä, eikä yksittäistä henkilöä. Tämä sama tieto voi kuitenkin eri asiayhteydessä muuttua luonnolliseen henkilöön liittyväksi.[67] Esimerkiksi kiinteistön arvo myyntiesitteessä ei itsessään liity luonnolliseen henkilöön, eikä sitä näin ollen olisi katsottava henkilötiedoksi. Kuitenkin sama tieto, kiinteistön arvo, esimerkiksi verovelvollisuuden tai veron määrän selvittämiseksi voidaan katsoa liittyvän henkilöön.

Neljäs osatekijä ”*luonnollinen henkilö*” viittaa yleisesti ihmisiin ja oikeus henkilötietojen suojaan on tässä mielessä yleinen eikä se rajoitu tietyn maan kansalaisiin tai asukkaisiin. Lähtökohtaisesti henkilötiedoksi ei siis katsota yritystä tai muuta organisaatiota tai kuollutta henkilöä koskevia tietoja. Poikkeuksia kuitenkin on, mutta en käsittele niitä tässä tutkielmassa.[68]

Edellä esitettyjen henkilötiedon käsitteen osatekijöiden perusteella on nähtävissä selvä pyrkimys henkilötiedon käsitteen mahdollisimman laajan tulkinnan soveltamiseen. Tietosuojatyöryhmä on todennut, että EU:n lainsäätäjän

[65] Tietosuojatyöryhmän lausunto 04/2007 s. 6.
[66] Purtova, s. 50.
[67] Tietosuojatyöryhmän lausunto 04/2007 s. 9.
[68] Esimerkiksi yrityksen toiminimitieto voi sisältää luonnollisen henkilön nimen, jolloin myös yrityksen toiminimitieto voisi olla henkilötietoa.

nimenomaisena tarkoituksena on ollut ottaa käyttöön laaja merkitys henkilötiedoista, mutta tämä merkitys ei voi olla rajoittamaton.[69]

Tämän tutkielman kannalta mielenkiintoinen on tietosuojatyöryhmän toteamus tietosuojadirektiivin ja yleisemminkin tietosuojasääntelyn tarkoituksesta ja laajuudesta:

> ”On aina pidettävä mielessä, että direktiivissä säädettyjen sääntöjen tavoitteena on suojata henkilöiden perusoikeuksia ja -vapauksia, erityisesti heidän oikeuttaan yksityisyyteen henkilötietojen käsittelyn osalta. Sääntöjen oli tarkoitettu tulevan sovellettavaksi tilanteisiin, joissa yksilön oikeudet olisivat vaarassa ja siten suojelun tarpeessa. Tietosuojasääntöjen soveltamisalaa ei pidä lieventää, mutta myös henkilötietojen käsitteen rajoittamista aiheettomasti olisi vältettävä.”[70]

Kuten jäljempänä huomataan, näiden tavoitteiden toteuttaminen on haastavaa uudenlaisten tietojen käsittely-ympäristöjen yhteydessä laajentamatta aiheettomasti tietosuojasäännösten soveltamisalaa. Tietosuojasääntely ei myöskään välttämättä tule sovellettavaksi aina sellaisissa tilanteissa, joissa yksilön oikeudet ovat vaarassa. Toisaalta sääntely mahdollistaa myös sellaisten tietojen määrittämisen henkilötiedoiksi, joilla ei lähtökohtaisesti ole lainkaan sellaista liityntää luonnolliseen henkilöön, että oikeussuojaa olisi aiheellista antaa.

3 Tunnistettavuuskriteerin sisältö ja tulkinta EU:ssa

3.1 Johdanto

Henkilötiedolla tarkoitetaan tietosuoja-asetuksessa ”kaikkia *tunnistettuun* tai *tunnistettavissa olevaan* luonnolliseen henkilöön […] liittyviä tietoja”.[71] Määritelmä on sama kuin tietosuojadirektiivissä.[72] Tämän lisäksi asetuksen määritelmä sisältää kuitenkin myös uutta, kun se tarkentaa tunnistettavissa olevan luonnollisen henkilön käsitettä. Tietosuoja-asetuksen mukaan tunnistettavissa olevana pidetään luonnollista henkilöä, joka voidaan suoraan tai

[69] Tietosuojatyöryhmän lausunto 04/2007 s. 25.
[70] Tietosuojatyöryhmän lausunto 04/2007 s. 25.
[71] Tietosuoja-asetuksen 4 (1) artikla.
[72] Tietosuojadirektiivin 2 (a) artikla.

epäsuorasti tunnistaa erityisesti tunnistetietojen, kuten nimen, henkilötunnuksen, sijaintitiedon, verkkotunnistetietojen taikka yhden tai useamman hänelle tunnusomaisen fyysisen, fysiologisen, geneettisen, psyykkisen, taloudellisen, kulttuurillisen tai sosiaalisen tekijän perusteella.[73] Siten *tunnistettavissa olevan* määritelmä eroaa *tunnistetun* määritelmästä siten, että siinä viitataan tunnistetietoihin ja henkilökohtaisiin ominaisuuksiin, joiden perusteella voidaan selvästi yksilöidä luonnollinen henkilö.

Yleisen tietosuoja-asetuksen henkilötiedon määritelmässä *tunnistettu* ja *tunnistettavissa oleva* ovat samassa asemassa eli toisin sanoen tiedot ovat henkilötietoja, jos luonnollinen henkilö *tunnistetaan* tai *on tunnistettavissa* näistä tiedoista. Käsitteiden samaistaminen on peräisin Saksan tietosuojalainsäädännöstä.[74] Niiden asettamista samalle tasolle EU:n tietosuojalainsäädännössä on kritisoitu erityisesti vertaamalla sitä Yhdysvaltojen malliin, jossa tietosuojavelvoitteiden soveltaminen vaihtelee sen mukaan, onko henkilö tiedoista selkeästi tunnistettu vai ainoastaan tunnistettavissa. Kritiikki kohdistuu muun muassa siihen, että tällä tavoin määritelty tunnistettavuuskriteeri soveltuu suureen määrään erilaisia tietoja, joiden osalta riskit tunnistamiseen ovat eri tasoisia, mutta näitä tietoja kohdellaan silti lainsäädännössä ikään kuin ne olisivat samankaltaisia.[75] Todellista riskiä yksilön tunnistamiseksi ei siis oteta huomioon.

Kaikki yleisen tietosuoja-asetuksen mukaiset velvoitteet soveltuvat yhtä lailla esimerkiksi henkilötunnukseen kuin IP-osoitteeseenkin, vaikka jälkimmäisessä tapauksessa henkilö olisi tiedoista tunnistettavissa ainoastaan kolmannen tahon hallussa olevien lisätietojen avulla. Voitaisiin kuitenkin väittää, että henkilötunnuksen osalta riski tunnistamiseen on suurempi, koska sen avulla henkilön tunnistaminen on helpompaa ja tunnistamisen voi tehdä käytännössä kuka tahansa. Näin ollen olisi perusteltua suojata henkilötunnusta laajemmin, kuin IP-osoitetta.

[73] Tietosuoja-asetuksen 4 (1) artikla.
[74] Schwartz & Solove 2011, s. 1874.
[75] Schwartz & Solove 2011, s. 1876.

Tunnistettu ja tunnistettavissa oleva ovat siis saman arvoisia tietosuojasääntelyn soveltumisen kannalta. Kun tunnistettavissa olevan määritelmän täyttymiseen riittää se, että on olemassa *mahdollisuus* tunnistaa henkilö, tunnistettavuuskriteerin täyttymisen arvioimiseksi ei ole tarpeen tarkastella, onko henkilö tunnistettu, vaan onko henkilö tunnistettavissa oleva.

Yleinen tietosuoja-asetus ei tuo mukanaan merkittäviä muutoksia henkilötiedon käsitteeseen verrattuna aikaisempaan tietosuojadirektiiviin. Direktiivin tavoin uusi asetus noudattaa kaksijakoista määrittelyä eli tietty tieto joko on tai ei ole henkilötietoa. Tiedon ollessa henkilötietoa, kaikki asetuksen velvoitteet tulevat sovellettavaksi ja tiedon ollessa ei-henkilötietoa, toisin sanoen anonyymiä tietoa, asetuksen mitkään velvoitteet eivät sovellu. Tietosuoja-asetuksen aineellista soveltamisalaa koskevan 2 (1) artiklan mukaan ”*Tätä asetusta sovelletaan henkilötietojen käsittelyyn, joka on osittain tai kokonaan automaattista, sekä sellaisten henkilötietojen käsittelyyn muussa kuin automaattisessa muodossa, jotka muodostavat rekisterin osan tai joiden on tarkoitus muodostaa rekisterin osa*”.

Tunnistettavuuskriteeri on merkittävä tai jopa merkittävin erottava tekijä, kun arvioidaan tietyn tiedon luokittelemista joko henkilötiedoksi tai anonyymiksi tiedoksi. Kriteerin avulla käytännössä siis arvioidaan sitä, tuleeko tietosuojanormisto sovellettavaksi vai ei.[76] Termi *tunnistettu* viittaa siihen, että tiedot kuvailevat henkilöä sellaisella tavalla, että henkilö on erotettavissa kaikista muista henkilöistä ja tunnistettavissa yksilönä.[77] *Tunnistettavissa oleva* sen sijaan viittaa siihen, että tiedot sisältävät sellaista informaatiota, joiden avulla henkilö *voidaan* tunnistaa suoraan tai epäsuorasti.[78] Tunnistettavuus määritelmänä tarkoittaa siis sitä, että tietojen avulla on mahdollista erottaa yksittäinen henkilö joko suoraan ennalta kerätyn ja olemassa olevan tietojoukon avulla tai epäsuorasti yhdistämällä olemassa olevaan tietoon lisätietoja, jotka mahdollistavat tunnistamisen.[79]

[76] Urgessa 2016, s. 521.
[77] European Union Agency for Fundamental Rights, Handbook on European data protection law (Council of Europe 2014) s. 39.
[78] Ibid. 40.
[79] Urgessa 2016, s. 521.

Tietosuojatyöryhmän mukaan yksilön erottaminen muista tapahtuu yleensä hyödyntämällä sellaisia tietoja, jotka ovat erityisen läheisesti yksilöön liitynnäisiä.[80] Tällaisina mainitaan esimerkiksi nimi, henkilötunnus, paikkatieto tai verkkotunnistetiedot sekä muut epäsuorat tunnisteet kuten fyysiset, psyykkiset tai taloudelliset tiedot.[81]

3.2 Tulkinnan lähtökohdat

3.2.1 Lähtökohtana laaja tulkinta

Jotta voidaan tarkastella sitä, kuinka tunnistettavuuskriteeriä olisi tulkittava, on tarkasteltava myös henkilötiedon käsitteen tulkintaa yleisesti. Koska tunnistettavuuskriteeri on osa henkilötiedon määritelmää, vaikuttavat henkilötiedon käsitteen tulkintaohjeet suoraan myös tunnistettavuuskriteerin tulkintaan.

EU:n henkilötietolainsäädännön tavoitteena on henkilötiedon tulkitseminen laajasti. Tietosuojatyöryhmä on katsonut lausunnossaan, että henkilötiedon määritelmä tietosuojadirektiivissä kuvastaa EU:n lainsäätäjän tahtoa määrittää käsite ”henkilötieto” laajasti.[82] Vastaavasti Euroopan komissio on katsonut ehdotuksessaan, että henkilötiedon määritelmän tulisi olla mahdollisimman kaiken kattava ja yleinen, jotta siihen sisältyisi kaikki sellainen tieto, joka koskee yksittäistä tunnistettavissa olevaa henkilöä.[83] Tietosuojatyöryhmän mukaan aiheeton henkilötiedon käsitteen tulkinnan rajoittaminen ei ole toivottavaa.[84] Tämän laajentavan tulkinnan tarkoitus on turvata tietosuojalainsäädännön perimmäisten tavoitteiden mahdollisimman tehokas suojaaminen. Nämä tavoitteet on kuvattu tietosuojadirektiivissä ja tietosuoja-asetuksessa olevan luonnollisten henkilöiden perusoikeuksien ja vapauksien suojaaminen sekä erityisesti heidän oikeus yksityisyyteen henkilötietojen käsittelyn yhteydessä.

[80] Tietosuojatyöryhmän lausunto 04/2007, s. 12.
[81] Ibid. s. 12-15; maininnat sisältyvät myös tietosuojadirektiivin 2 (a) artiklaan ja tietosuoja-asetuksen 4 (1) artiklaan.
[82] Tietosuojatyöryhmän lausunto 04/2007, s. 4.
[83] Euroopan komission muutettu ehdotus neuvoston direktiiviksi yksilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta, s. 10.
[84] Tietosuojatyöryhmän lausunto 04/2007, s. 5.

Tavoitteet on otettava huomioon tietosuojalakeja sovellettaessa ja tulkittaessa, sillä ne ovat tärkeässä roolissa, kun määritellään yksilöiden suojan tarvetta kussakin tilanteessa.[85]

Työryhmän mukaan tietosuojasäännösten soveltamisalaa ei tulisi myöskään liiaksi laajentaa. Henkilötiedon käsitettä liikaa laajentavia tulkintoja tulisi välttää yksittäisessä soveltamistilanteessa hyödyntämällä tietosuojasääntelyyn sisältyvää joustavuutta.[86] Työryhmän mukaan tietosuojasäännöksiä sovellettaessa tulisi siis aina tapauskohtaisesti harkita, voidaanko tiedot yhdistää jollakin tavalla luonnolliseen henkilöön. Tietosuojasäännösten olisi aina sovelluttava silloin, kun käsittely vaikuttaa jollakin tavalla luonnollisen henkilön oikeuksiin tai vapauksiin. Kuitenkin henkilötiedon määritelmään sisältyvä joustavuus olisi tarpeen mukaan otettava käyttöön, jotta voidaan välttää henkilötiedon käsitteen liiallinen laajentaminen.

Sekä tietosuojadirektiivissä että tietosuoja-asetuksessa[87] avataan tunnistettavuuden määritelmää seuraavasti:

> *”Jotta voidaan määrittää, onko luonnollinen henkilö tunnistettavissa, olisi otettava huomioon kaikki keinot, joita joko rekisterinpitäjä tai muu henkilö voi kohtuullisen todennäköisesti käyttää mainitun luonnollisen henkilön tunnistamiseen suoraan tai välillisesti, kuten kyseisen henkilön erottaminen muista.”*

Tietosuojatyöryhmä arvioi lausunnossaan[88] tunnistamisen keinoja ja erityisesti sitä, minkälaisiin seikkoihin tulee kiinnittää huomiota, kun arvioidaan *kohtuullisen todennäköisiä keinoja*, joita rekisterinpitäjä tai muu henkilö voi käyttää tunnistamiseen. Pelkkä hypoteettinen mahdollisuus yksittäisen luonnollisen henkilön tunnistamiseen ei riitä katsomaan tätä henkilöä tunnistettavissa olevaksi. Jos tällaisia mahdollisuuksia ei ole tai ne ovat merkityksettömän pienet, tällöin tietoja ei katsottaisi henkilötiedoiksi. Arvioinnissa tulee kiinnittää huomiota kaikkiin kyseisessä tapauksessa käsillä oleviin seikkoihin. Työryhmä nostaa esiin arviointiin vaikuttavina seikkoina tunnistamiseen liittyvät kustannukset, tietojen käyttötarkoituksen, tietojenkäsittelyn rakenteen, rekisterinpitäjän tavoitteleman

[85] Ibid, s. 4.
[86] Ibid, s. 4-6.
[87] Molemmissa johdanto-osan kohta 26.
[88] Tietosuojatyöryhmän lausunto 04/2007.

edun, luonnollista henkilöä koskevat intressit ja tietotekniset riskitekijät kuten tietoturvaloukkaukset ja tekniset häiriöt.

Tietosuojatyöryhmä korostaa, että tiedon luokitteleminen henkilötiedoksi on riippuvainen asiayhteydestä.[89] Esimerkiksi hyvin yleinen sukunimi ei riitä yksittäisen henkilön tunnistamiseen koko maan väestöstä, mutta todennäköisesti hänet voidaan tunnistaa osana yksittäisen koululuokan oppilaita.[90] Tunnistettavuuskriteerin tulkinnassa olisi siis huomioitava tapauskohtaisesti se todennäköisyys, jolla henkilö voidaan tietojen perusteella tunnistaa.

3.2.2 Suora ja epäsuora tunnistaminen

Tietosuoja-asetuksen määritelmän mukaan tunnistettavissa olevana pidetään luonnollista henkilöä, joka voidaan *suoraan* tai *epäsuorasti* tunnistaa. Tämä määritelmä tekee tunnistettavuuskriteeristä verrattain laajan. Epäsuoran tunnistamisen nimenomainen sisällyttäminen tietosuoja-asetuksen tekstiin mahdollistaa sen, että kaikkien tietojen, joiden avulla yksittäinen henkilö voidaan tunnistaa, ei tarvitse olla saman henkilön hallussa. Epäsuoran tunnistamisen johdosta henkilötiedon käsitteen soveltamisalaa voidaan laajentaa lähes rajattomasti ja siksi yksittäisissä soveltamistilanteissa tulisi välttää liian pitkälle meneviä tulkintoja. Sen sijaan soveltamisessa tulisi hyödyntää tietosuojasäädöksiin jätettyä joustavuutta tapauskohtaisessa harkinnassa.

Henkilö on suoraan tunnistettu, kun tieto johtaa selvään ja varmaan johtopäätökseen henkilön identiteetistä.[91] Esimerkiksi henkilön nimi voi olla sellainen tieto, josta yleensä voidaan johtaa tieto henkilön identiteetistä. Tämä ei kuitenkaan aina pidä paikkaansa ja huomiota joudutaankin kiinnittämään kyseessä olevan tilanteen olosuhteisiin. Jos esimerkiksi tarkastellaan jotakin yleistä nimeä tietyllä asuinalueella, tällainen tieto ei välttämättä erota tiettyä henkilöä riittävällä tarkkuudella muista saman nimisistä henkilöistä. Sama henkilö voi kuitenkin olla tunnistettavissa muussa yhteydessä, kuten esimerkiksi

[89] Henkilötiedon käsitteen riippuvaisuus asiayhteydestä on vahvistettu tapauksessa Common Services Agency v Scottish Information Commissioner.
[90] Tietosuojatyöryhmän lausunto 04/2007, s. 13.
[91] Esayas 2015, s. 2.

pienemmällä asuinalueella. Ensin mainitussa tapauksessa nimi saattaa silti olla tieto, jonka avulla henkilö voidaan tunnistaa, kun yhdistetään nimeen myös muita tietoja. Tällöin on kyseessä epäsuora tunnistaminen.

Euroopan komission mukaan mahdollisimman laaja määritelmä henkilötiedosta on tärkeää, jotta katetaan kaikki sellaiset tiedot, jotka voidaan mahdollisesti yhdistää yksittäistä luonnollista henkilöä koskeviksi.[92] Henkilö on tunnistettavissa, vaikka ei ole suoraa ja selkeää viittausta tiettyyn henkilöön, mutta on mahdollisuus tunnistaa tämä henkilö tiettyjen tietojen "ainutlaatuisella yhdistelmällä"[93], joka mahdollistaa yksilön erottamisen muista. Täten siis pelkkä mahdollisuus tiettyjen tietojen yhdistämiseksi yksittäistä henkilöä koskeviksi riittää siihen, että henkilötiedoksi katsomisen edellytykset täyttyvät.[94]

On kuitenkin edelleen kiistanalaista, onko arvioinnin tekemistä varten käytettävä niin sanottua objektiivista vai suhteellista perustetta[95], jolla arvioidaan rekisterinpitäjän kykyä tunnistaa luonnollinen henkilö.

3.3 Objektiivinen vai suhteellinen peruste

3.3.1 Objektiivinen peruste

Objektiivisen perusteen soveltaminen tarkoittaa sitä, että tiettyjen tietojen voitaisiin katsoa olevan henkilötietoja siitä huolimatta, että ainoastaan sivullinen kykenee näistä tiedoista määrittämään rekisteröidyn henkilöllisyyden. Rekisterinpitäjällä tai muulla tietoja käsittelevällä taholla ei siis tämän perusteen mukaisesti tarvitse itsellään olla käytettävissään kaikkia niitä tietoja, jotka tekevät rekisteröidyn tunnistamisen mahdolliseksi ja silti tietojen voidaan katsoa olevan henkilötietoja nimenomaisesti kyseessä olevaan rekisterinpitäjään tai muuhun nähden. Kyse on siis eräänlaisesta yhteisvaikutuksesta, jossa henkilön tunnistamisen mahdollistavat tiedot ovat eri toimijoiden hallussa ja vasta näiden

[92] Tietosuojatyöryhmän lausunto 04/2007, s.4; Komission ehdotus neuvoston direktiiviksi koskien yksilöiden suojelua henkilötietojen käsittelyssä COM(90) 314 final, 13.9.1990, s.19.
[93] Tietosuojatyöryhmän lausunto 04/2007, s.13.
[94] Massey 2013, s. 87; Esayas 2015, s. 2.
[95] Termit objektiivinen ja suhteellinen peruste esiintyvät esimerkiksi EUT:n oikeuskäytännössä Breyer tuomiossa (C-582/14). Englanninkielessä vastaavat yleisesti käytetyt termit ovat *absolute and relative approach*, ruotsiksi *absolut eller relativ* bedömning.

tietojen yhdisteleminen mahdollistaa tunnistamisen. Objektiivisen perusteen mukaan, jos tietoja yhdistelemällä on mahdollista tunnistaa henkilö, tulee kaikista näistä ns. osatiedoista tällöin henkilötietoja suhteessa kaikkiin niihin toimijoihin, joilla kyseisiä tietoja on.

Objektiivinen peruste korostaa ajatusta, että huomioon on otettava kaikki keinot, joita *muu henkilö* voi käyttää henkilön tunnistamiseksi. Siten kaikissa tilanteissa, joissa tiettyjen tietojen yhdisteleminen muun henkilön hallussa olevien tietojen kanssa mahdollistaa luonnollisen henkilön tunnistamisen näiden tietojen avulla, katsottaisiin kyseiset tiedot henkilötiedoiksi. Arviointi on riippumaton siitä, liittyvätkö nämä muun henkilön hallussa olevat lisätiedot rekisterinpitäjään tai tämän toimintaan vai ei. Toisin sanoen henkilötiedon käsite on riippumaton kussakin tilanteessa tarkasteltavana olevista henkilöistä. Käytännössä objektiivisen perusteen mukaisesti siis, jos tieto katsotaan henkilötiedoksi yhden henkilön kannalta, on tuo sama tieto myös jonkun muun henkilön hallussa määriteltävä henkilötiedoksi.[96]

3.3.2 Suhteellinen peruste

Kaikki eivät kuitenkaan hyväksy objektiivisen perusteen mukaista tulkintaa tunnistettavuudesta. Joidenkin mukaan tunnistettavuutta on arvioitava suhteellisen perusteen mukaisesti siten, että sama tieto voi olla toisen henkilön hallussa anonyymiä ja toisen henkilön hallussa taas henkilötietoa.[97] Suhteellista perustetta kannattavien mukaan täytyisi enemmän huomiota kiinnittää *keinoihin, joita kohtuullisen todennäköisesti* käytetään henkilön tunnistamiseksi.[98] Tämä tarkoittaa sitä, että tiedot ovat luokiteltavissa henkilötiedoiksi sellaisen henkilön hallussa, joka pystyy tunnistamaan niistä luonnollisen henkilön, mutta eivät sellaisen henkilön hallussa, jolla ei ole realistista mahdollisuutta tunnistaa luonnollista henkilöä.

Suhteellisen perusteen mukaan tietoja voidaan pitää henkilötietoina ainoastaan sellaiseen tahoon nähden, jolla on itsellään käytettävissä kaikki sellaiset tiedot,

[96] DLA Piper 2009, s. 19.
[97] Esim. Eckhardt 2007, s. 603 ja Eckhardt 2008, s. 769.
[98] DLA Piper 2009, s. 19.

joiden avulla henkilö on mahdollista tunnistaa täsmällisesti.[99] Tietoja ei sen sijaan voida pitää henkilötietoina sellaiseen tahoon nähden, jolla ei ole määrätyllä hetkellä ilman suhteetonta vaivaa käytettävissään henkilön tunnistamiseksi tarvittavia tietoja, vaikka tällaiset tiedot olisi kuitenkin jostakin saatavissa. Arvioinnissa joudutaan siis kiinnittämään huomiota mm. siihen, millaisia lisätoimia tietoja käsittelevältä taholta vaaditaan, jotta henkilön tunnistaminen olisi mahdollista.

3.3.3 Arviointia perusteista

EU:n yleisessä tietosuoja-asetuksessa käytetään laajaa määritelmää tunnistettavissa olevasta luonnollisesta henkilöstä, mikä on ollut tavoite jo tietosuojadirektiivin soveltamisen aikana. Laajan määritelmän voidaan katsoa ilmentävän objektiivista perustetta tunnistettavuuskriteerin määrittelyyn. Kuitenkin jotkin asetuksen kohdat ja muu tulkinta-aineisto näyttävät ilmentävän myös suhteellista perustetta.[100] Kuten EUT:n tuomioistuinkäytännöstä, myöskään tietosuoja-asetuksesta ei voida vetää lopullista johtopäätöstä sen suhteen, tulisiko tunnistettavuuden arvioinnissa nojautua objektiiviseen vai suhteelliseen perusteeseen.

Tietosuoja-asetuksen johdanto-osan kohdan 26 tarkastelusta voidaan tunnistaa elementtejä molemmista perusteista: ”Jotta voidaan määrittää, onko luonnollinen henkilö tunnistettavissa, olisi otettava huomioon kaikki keinot, joita joko rekisterinpitäjä tai *muu henkilö* voi kohtuullisen todennäköisesti käyttää mainitun luonnollisen henkilön tunnistamiseen suoraan tai välillisesti, kuten kyseisen henkilön erottaminen muista.” Maininnan ”muu henkilö” voidaan katsoa viittaavan objektiiviseen perusteeseen, koska tämä muu henkilö voi olla käytännössä kuka tahansa.[101] Tällainen tulkinta olisi yhteensopiva myös EU:n perusoikeuskirjan näkökulmasta, sillä sen mukaisesti tunnistettavuutta tulisi tarkastella laajentavasti.[102]

[99] Ks. tältä osin EUT:n tuomio asiassa Scarlet Extended (C-70/10), kohta 51.
[100] Spindler & Schmechel, s. 166.
[101] Spindler & Schmechel, s. 166.
[102] Perusoikeuskirjan 8 artikla.

Asetuksen johdanto-osan kohdan 26 maininnan luonnollisen henkilön suorasta tai epäsuorasta tunnistamisesta erottamalla kyseinen henkilö muista voidaan katsoa myös ilmentävän objektiivista perustetta. Henkilö voidaan siis katsoa olevan tunnistettavissa, kun hänet voidaan erottaa muista, vaikka olisikin epätodennäköistä, että hänen nimensä voitaisiin yhdistää tähän tietoon. Henkilön erottaminen muista katsotaan siis itsessään riskiksi henkilön yksityisyyden vaarantumiseen.[103]

Toisaalta maininta kohtuullisen todennäköisistä keinoista näyttäisi viittaavan suhteellisen perusteen mukaiseen tulkintaan, varsinkin termin *kohtuullisen* osalta.[104] Lisäksi, jos noudatettaisiin tulkintaa, jonka mukaan ei tarvitsisi olla olemassa lainkaan todellista riskiä henkilön tunnistamiseksi, mikään olemassa oleva tekniikka ei pystyisi takaamaan vaadittavaa anonymisoinnin tasoa.[105] Tietosuojatyöryhmän mukaan pelkkä hypoteettinen mahdollisuus yksilön erottamiseen muista ei tekisi yksilöä tunnistettavissa olevaksi.[106]

Tietosuoja-asetuksen johdanto-osan kohdassa 26 luetellaan kriteerejä, joiden valossa arvioidaan yksilön tunnistamiseksi käytettävien keinojen tulkintaa:

> ”Jotta voidaan varmistaa, voidaanko keinoja kohtuullisen todennäköisesti käyttää luonnollisen henkilön tunnistamiseen, olisi otettava huomioon kaikki objektiiviset tekijät, kuten tunnistamisesta aiheutuvat kulut ja tunnistamiseen tarvittava aika sekä käsittelyajankohtana käytettävissä oleva teknologia ja tekninen kehitys.”

Näiden tekijöiden voidaan katsoa olevan pyrkimys kaventaa tietosuoja-asetuksen objektiivisia elementtejä suhteessa tunnistettavuuteen.[107]

Tietosuoja-asetuksen aineellisen soveltamisalan liiallinen laajentaminen voi potentiaalisesti johtaa turhan kuormittavaan sääntelyyn tietoja käsittelevien tahojen näkökulmasta. Lisäksi tämä raskas sääntely olisi epäsuhteessa rekisteröityjen todellisen yksityisyyden suojan tarpeen kanssa.[108] On myös esitetty näkemyksiä, että liian laaja tietosuojasääntelyn soveltaminen olisi itse

[103] Spindler & Schmechel, s. 166.
[104] Ibid.
[105] Ibid.
[106] Tietosuojatyöryhmän lausunto 04/2007 s. 15.
[107] Spindler & Schmechel, s. 167.
[108] Schwartz & Solove 2014, s. 887.

asiassa tietosuojasääntelyn tavoitteiden vastaista. Tätä on usein perusteltu viittaamalla tietosuoja-asetuksen johdanto-osan kohtaan 4:

> ”Oikeus henkilötietojen suojaan ei ole absoluuttinen; sitä on tarkasteltava suhteessa sen tehtävään yhteiskunnassa ja sen on suhteellisuusperiaatteen mukaisesti oltava oikeassa suhteessa muihin perusoikeuksiin.”

Jos unionin tuomioistuin noudattaisi laajaa, lähes objektiivisen perusteen mukaista tulkintaa tunnistettavuudesta, käytännössä lähes kaikki tieto olisi katsottavissa henkilötiedoksi, mikä lopulta heikentäisi tietosuojasääntelyn vaikuttavuutta ja saattaisi tehdä siitä toimimattoman.[109] Nimittäin mikäli tietosuojasääntely tulee lähes automaattisesti sovellettavaksi mihin tahansa tietoon sen luonteeseen katsomatta, tämä luo ei-toivottuja kannustimia olla suojaamatta tietoja esimerkiksi anonymisoimalla ja täten riskit yksityisyydelle todellisuudessa kasvaisivat.[110] Lopputulos olisi täysin päinvastainen kuin mitä tietosuojasääntelyllä pyritään saavuttamaan.

On myös esitetty, että tällainen liian laaja tulkinta heikentäisi informoidun suostumuksen asemaa tietojen käsittelyn oikeusperusteena. Kun henkilötiedoksi katsottavien tietojen joukko kasvaa, samalla myös suostumusten hankkimisen määrä kasvaa, sillä rekisterinpitäjät ja muut tietoja käsittelevät tahot joutuvat osoittamaan oikeusperusteen henkilötietojen käsittelylle. Yksi oikeusperuste henkilötietojen käsittelylle on rekisteröidyn suostumus ja kun suostumuksia näin ryhdyttäisiin hankkimaan rekisteröidyiltä lähes rutiininomaisesti tutkimatta tietojen luonnetta sen tarkemmin, suostumuksen todellinen merkitys kärsisi inflaatiota. Rekisteröidyt eivät enää kiinnittäisi tarkkaa huomiota suostumuksen sisältöön.[111]

Objektiivisen ja suhteellisen perusteen eroavaisuuksia voidaan havainnollistaa tarkastelemalla IP-osoitteita. Voidaanko IP-osoitteen katsoa olevan henkilötieto? Objektiivisen perusteen mukaan IP-osoite on kiistatta henkilötieto, sillä internet-yhteyden tarjoaja pystyy helposti tunnistamaan luonnollisen henkilön IP-osoitteen

[109] Tene & Polonetsky 2012, s. 66.
[110] Tene & Polonetsky 2012, s. 66.
[111] Keppeler, s. 360.

perusteella vertaamalla sitä omaan asiakasrekisteriinsä.[112] Suhteellisen perusteen mukaan IP-osoitteet ovat henkilötietoja ainoastaan internet-yhteyden tarjoajaan ja joihinkin viranomaisiin[113] nähden. Muihin nähden IP-osoitteet eivät olisi henkilötietoja, koska niillä ei ole kohtuullisen todennäköisesti keinoja käytettävissään luonnollisen henkilön tunnistamiseksi IP-osoitteen perusteella. Olettamus jälkimmäisessä on se, että internet-operaattori ei vapaaehtoisesti luovuta käyttäjiensä henkilötietoja kolmansille osapuolille ja mahdollinen oikeusteitse tuomioistuimesta hankittava lupa tällaisen tiedon saamiseksi ei enää ole ”kohtuullisen todennäköisten” keinojen piirissä.[114]

EU-tasolla ei vallitse yksimielisyyttä siitä, kumpaa perustetta tässä arvioinnissa olisi noudatettava. Näin ollen henkilötiedon käsitteen laajuudesta vallitsee epävarmuus EU:ssa ja siten myös sen jäsenvaltioissa. Tämä epävarmuus tarkoittaa myös sitä, että yleisen tietosuoja-asetuksen aineellinen soveltamisala jää avoimeksi. EU:n jäsenvaltioiden velvollisuus on tulkita kansallista lainsäädäntöään EU-oikeuden mukaisesti. Erityisesti jäsenvaltioiden on huomioitava EU-oikeuskäytäntö, koska sen avulla usein täsmennetään avointen lainsäännösten tulkintaa. Tällä hetkellä kuitenkaan EU-oikeuskäytäntökään ei tunnu tarjoavan suuresti tulkinta-apua tunnistettavuuskriteerin arviointiin ja kansalliset tuomioistuimet ovatkin omissa ratkaisuissaan päätyneet toisistaan poikkeaviin ratkaisuihin kriteerin tulkinnasta.

Usein yritykset ja muut tietoja käsittelevät toimijat joutuvatkin niin sanotusti varmuuden vuoksi suunnittelemaan tietojenkäsittelytoimensa yhteensopiviksi henkilötietojen käsittelyä koskevan EU-lainsäädännön kanssa, koska henkilötiedon käsitteestä ei ole muodostunut harmonisoitua tulkintaa. Nämä toimijat eivät voi siis olla varmoja, tulkitaanko heidän olevan rekisterinpitäjiä tai henkilötiedon käsittelijöitä vai ei.

[112] Tässä tosin on huomattava, että sama ei välttämättä päde yleisessä käytössä oleviin tietokoneisiin ja niiden IP-osoitteisiin, koska tällöin varsinaisena käyttäjänä ei ole internet-yhteyden tilaaja.
[113] Viranomaiset, jotka voivat pyytää internet-yhteyden tarjoajalta tarvittavat lisätiedot luonnollisen henkilön tunnistamiseksi.
[114] DLA Piper 2009, s. 19.

3.4 EUT:n tulkinta tunnistettavuuskriteeristä

3.4.1 Johdanto

Euroopan unionin tuomioistuimen (EUT) ratkaisukäytännöstä voidaan todeta, että se on perusteluissaan painottanut henkilötiedon mahdollisimman laajaa tulkintaa. Esimerkiksi tuomiossaan asiassa Google Spain[115], se torjui julkisasiamiehen ratkaisuehdotuksessaan esittämän näkökannan suhteellisuusperiaatteen huomioimisesta tietosuojadirektiivin soveltamisalan kaventamiseksi. Julkisasiamiehen ehdotus oli seuraava:

> ”Teknisen kehityksen vuoksi henkilötietojen, henkilötietojen käsittelyn ja rekisterinpitäjän laajat määritelmät kattavat nykytilanteessa todennäköisesti ennenkuulumattoman laajan joukon uusia tosiseikastoja. […] Unionin tuomioistuimen on tämän vuoksi direktiivin soveltamisalaa tulkitessaan sovellettava […] suhteellisuusperiaatetta, jotta kohtuuttomat ja liialliset oikeusvaikutukset voidaan välttää”[116]

EUT kuitenkin katsoi, ettei suhteellisuusperiaatetta ole otettava huomioon direktiivin soveltamisalaa määritettäessä. EUT perusteli kantaansa sillä, että direktiivin säännöksiä on ”*välttämättä tulkittava perusoikeuksien valossa, sillä nämä perusoikeudet ovat vakiintuneen oikeuskäytännön mukaan erottamaton osa yleisiä oikeusperiaatteita, joiden noudattamista unionin tuomioistuin valvoo ja jotka sisältyvät nykyään perusoikeuskirjaan*”. EUT:n kanta on toisin sanoen ollut se, että mahdollisia tietosuojalainsäädännön laaja-alaisen soveltamisen ei-toivottuja vaikutuksia ei lievennetä soveltamisalan suppeamman tulkinnan kautta, vaan pikemminkin yksittäisten tietosuojasäännösten oikeasuhtaisen soveltamisen kautta.

EUT:n ratkaisukäytännöstä ei ole saatavissa selkeitä tulkintaohjeita tunnistettavuuskriteerille. Lokakuussa 2016 se antoi kuitenkin ratkaisunsa asiassa C-582/14 Patrick Breyer v Saksan liittotasavalta, joka on täsmentänyt jossain määrin niitä perusteita, joilla tunnistettavuutta tulisi EU:n alueella tulkita. Kirjallisuudessa tämän tuomion perusteluista on kuitenkin esitetty erilaisia tulkintakannanottoja. Käyn seuraavassa alaluvussa läpi tuomion perustelut ja

[115] Asia C-131/12 Google Spain SL, Google Inc v Agencia Española de Protección de Datos ja Mario Costeja González [2014] ECLI:EU:C:2014:317.
[116] Julkisasiamiehen Niilo Jääskisen ratkaisuehdotus asiassa C‑131/12, kohta 30.

esitän omat näkemykseni siitä, miten EUT on arvioinut tunnistettavuuskriteeriä ja millaisia johtopäätöksiä tästä voidaan tehdä.

Ennen Breyer -tuomiota EUT on katsonut niin sanottujen staattisten IP-osoitteiden[117] olevan henkilötietoja. EUT katsoi tuomiossaan Scarlet Extended (C-70/10), että internetin käyttäjien IP-osoitteet, joiden avulla nämä käyttäjät on mahdollista tunnistaa täsmällisesti, ovat henkilötietoja. Linjaus koski kuitenkin ainoastaan tilannetta, jossa käyttäjä on mahdollista tunnistaa IP-osoitteen perusteella täsmällisesti ja jossa internetin käyttäjien IP-osoitteiden keräämisen ja tunnistamisen suorittivat internet-yhteyden tarjoajat.

Tämän EUT:n tuomion perusteella kysymys IP-osoitteiden henkilötietoluonteesta jäi kuitenkin käytännössä hyvinkin avoimeksi. Siinä ei ole jouduttu ottamaan kantaa epäsuoraan tunnistamiseen tai sivullisilla olevien lisätietojen merkitykseen, toisin kuin seuraavana käsiteltävässä Breyer -tapauksessa.

3.4.2 Tapaus C-582/14 – Patrick Breyer

Euroopan unionin tuomioistuimen (EUT) tuomiossa asiassa Patrick Breyer v. Saksan liittotasavalta[118] on pääosin kysymys siitä, onko verkkosivulla vierailevan henkilön dynaaminen IP-osoite katsottava henkilötiedoksi verkkosivun julkaisijaan nähden, kun toisella taholla, tässä tapauksessa internetyhteyden tarjoajalla, on mahdollisuus tunnistaa kyseinen henkilö IP-osoitteen perusteella. Tuomioistuin katsoi, että IP-osoite on henkilötieto verkkosivun julkaisijaan nähden, jos tällä julkaisijalla on käytettävissään oikeudelliset keinot hankkia internetyhteyden tarjoajalta verkkosivulla vierailleesta henkilöstä tarvittavat lisätiedot, jotka mahdollistavat henkilön tunnistamisen.

Patrick Breyer vieraili useilla Saksan liittovaltion laitosten ylläpitämillä verkkosivuilla. Useimmat näistä verkkosivuista tallentavat vierailijoistaan niin

[117] Staattinen IP-osoite tarkoittaa kiinteää IP-osoitetta, joka on muuttumaton ja jonka perusteella verkkoyhteydessä oleva laite voidaan tunnistaa pysyvästi.
[118] Asia C-582/14 Patrick Breyer v Saksan liittotasavalta (19 päivänä lokakuuta 2016) ECLI:EU:C:2016:779.

sanottuja lokitietoja, jotka sisältävät ainakin käynnin päätyttyä haetun sivuston tai tiedoston nimen, hakukentissä käytetyt hakusanat, käynnin päivämäärän ja kellonajan, siirrettyjen tietojen määrän, ilmoitukset sivustolle pääsyn onnistumisesta ja sen tietokoneen IP-osoitteen, jolta sivustolle on pyritty.[119] Sivustot tallentavat kyseisiä tietoja verkkohyökkäyksiltä suojautumista ja hyökkääjien rikosoikeudelliseen vastuuseen saattamista varten.

EU:n tuomioistuin kuvailee IP-osoitteiden olevan ”*verkossa oleville tietokoneille niiden internetin kautta tapahtuvan viestinnän helpottamiseksi annettuja numerosarjoja*”[120]. IP-osoitteet voivat olla staattisia tai dynaamisia. Staattiset IP-osoitteet ovat muuttumattomia ja joiden perusteella verkkoyhteydessä oleva laite voidaan tunnistaa pysyvästi. Dynaamiset IP-osoitteet taas ovat tilapäisiä osoitteita, jotka annetaan kunkin internetyhteyden yhteydessä ja jotka vaihtuvat myöhempien internetyhteyksien yhteydessä.[121] Verkkoyhteyksien tarjoajilla on rekisteri, josta selviää tietylle laitteelle kullakin hetkellä annettu IP-osoite.[122] Toisin kuin staattiset IP-osoitteet, dynaamiset IP-osoitteet eivät mahdollista yleisesti saatavilla olevien tiedostojen avulla yhteyden toteamista tietyn tietokoneen ja internetpalvelun tarjoajan käyttämän fyysisen verkkoliittymän välillä.[123]

Breyer nosti saksalaisessa hallintotuomioistuimessa kanteen, jossa hän vaati, että Saksan liittotasavaltaa kielletään tallentamasta tai antamasta sivullisen tallennettavaksi Breyerin isäntäjärjestelmän IP-osoitetta Saksan liittovaltion laitosten verkkomediapalveluiden yleisön saatavilla olevilla sivustoilla käynnin päätyttyä, jollei tallentaminen ole tarpeen verkkomediapalvelun käytettävyyden palauttamiseksi häiriön ilmetessä. Breyerin kanne hylättin ensimmäisessä oikeusasteessa ja hän haki muutosta hylkäävään päätökseen. Toisen oikeusasteen tuomioistuin katsoi, että Breyerin IP-osoitteen tallentaminen ei tehnyt hänestä tunnistettavissa olevaa verkkosivun julkaisijaan eli Saksan liittotasavaltaan nähden ja siten IP-osoitetta ei katsottaisi henkilötiedoksi.[124]

[119] Tuomion kohta 14.
[120] Ibid kohta 15.
[121] Ibid kohta 36.
[122] Julkisasiamiehen ratkaisuehdotus Manuel Campos Sánchez-Bordona 12 päivänä toukokuuta 2016 (1), kohta 2.
[123] Tuomion kohta 16.
[124] Tuomion kohta 21.

Tässä toisen oikeusasteen tuomioistuin kuitenkin huomautti, että dynaaminen IP-osoite yhdessä sivulla käynnin ajankohdan kanssa on silloin, kun asianomaisen internetsivuston käyttäjä on ilmaissut henkilöllisyytensä käynnin yhteydessä, henkilötieto, koska sivuston ylläpitäjä voi tunnistaa tämän käyttäjän yhdistämällä hänen nimensä hänen tietokoneensa IP-osoitteeseen.[125]

Sekä Breyer että Saksan liittotasavalta tekivät toisen oikeusasteen tuomioistuimen ratkaisusta valituksen Saksan liittotasavallan ylimpään tuomioistuimeen Bundesgerichtshofiin. On kiistatta selvää, kuten Bundesgerichtshofkin katsoi, että Breyerin henkilöllisyyttä ei voida tunnistaa suoraan näin tallennettujen IP-osoitetietojen perusteella. Verkkosivun tarjoaja voi tunnistaa Breyerin ainoastaan, jos se saa internetyhteyden tarjoajalta tämän tunnistamiseksi tarvittavat lisätiedot. Kysymykseksi muodostuikin näin ollen se, onko Breyer ollut *tunnistettavissa* tietosuojadirektiivin tarkoittamassa merkityksessä.

Bundesgerichtshof tuo perusteluissaan esiin oikeuskirjallisuudessa esiintyvät vastakkaiset näkemykset siitä, onko sen määrittämiseksi, onko henkilö tunnistettavissa, nojauduttava objektiiviseen vai suhteelliseen perusteeseen. Objektiivista perustetta sovellettaessa päädytään siihen, että kyseessä olevan IP-osoitteen kaltaisten tietojen voitaisiin internetsivustoilla käyntikerran päätyttyä katsoa olevan henkilötietoja siitä huolimatta, että ainoastaan sivullinen kykenee määrittämään rekisteröidyn henkilöllisyyden. Tämä sivullinen on tässä tapauksessa Breyerin internetyhteyden tarjoaja, joka on tallentanut lisätietoja, jotka mahdollistavat hänen tunnistamisensa mainittujen IP-osoitteiden avulla.

Suhteellisen perusteen mukaan näitä tietoja voidaan pitää henkilötietoina ainoastaan Breyerin internetyhteyden tarjoajan kaltaiseen elimeen nähden, koska niiden avulla käyttäjä on mahdollista tunnistaa täsmällisesti, mutta niitä ei voida pitää henkilötietoina niiden internetsivustojen ylläpitäjän, joilla Breyer on käynyt, kaltaiseen toiseen elimeen nähden, koska tällä ylläpitäjällä ei ole silloin, kun Breyer ei ole ilmaissut henkilöllisyyttään sivustoilla käyntien yhteydessä,

[125] Tuomion kohta 20.

ilman suhteetonta vaivaa käytettävissään hänen tunnistamisekseen tarvittavia tietoja.[126]

Bundesgerichtshof pyysi ennakkoratkaisua EUT:lta ja lausui kysymyksenään[127] seuraavaa:

> ”Onko direktiivin 95/46 2 artiklan a alakohtaa tulkittava siten, että dynaaminen IP-osoite, jonka verkkomediapalvelun tarjoaja tallentaa henkilön käydessä tämän palveluntarjoajan yleisön saataville asettamalla internetsivustolla, on tähän palveluntarjoajaan nähden kyseisessä säännöksessä tarkoitettu henkilötieto, jos vain sivullisella, tässä tapauksessa tämän henkilön internetyhteyden tarjoajalla, on käytettävissään hänen tunnistamisekseen tarvittavia lisätietoja[?]”.[128]

Ennakkoratkaisua pyytäneen tuomioistuimen kysymys perustuu kahteen lähtökohtaiseen olettamaan:

> 1. tiedot, jotka koostuvat dynaamisesta IP-osoitteesta ja tästä IP-osoitteesta tapahtuneen internetsivustolla käynnin päivämäärästä ja kellonajasta ja jotka verkkomediapalvelujen tarjoaja tallentaa, eivät yksinään anna tälle palveluntarjoajalle mahdollisuutta tunnistaa kyseistä tällä internetsivustolla käynyttä käyttäjää
>
> 2. internetyhteyden tarjoajalla on käytettävissään lisätietoja, joiden perusteella on mahdollista tunnistaa tämä käyttäjä, kun niihin yhdistetään kyseinen IP-osoite.[129]

EUT totesi, että sen ratkaisemiseksi, onko dynaaminen IP-osoite kyseenä olevassa tilanteessa direktiivin 96/45 2 artiklan a alakohdassa tarkoitettu henkilötieto verkkomediapalvelujen tarjoajiin nähden, on tutkittava, voidaanko tällaisen palveluntarjoajan tallentamaa dynaamista IP-osoitetta pitää tunnistettavissa olevaa luonnollista henkilöä koskevana tietona, kun tämän palveluntarjoajan yleisön saataville asettaman internetsivuston käyttäjän tunnistamiseksi tarvittavat lisätiedot ovat kyseisen käyttäjän internetyhteyden tarjoajan hallussa.[130]

Ratkaisussa sovellettavana olleen tietosuojadirektiivin 2 (a) artiklan alakohdassa säädetään, että tunnistettavissa olevana pidetään henkilöä, joka voidaan tunnistaa paitsi suoraan myös epäsuorasti. EUT tulkitsi, että unionin lainsäätäjä

[126] Tuomion kohta 25.
[127] Ensimmäisenä kysymyksenään. Tuomioistuin esitti EUT:lle myös toisen kysymyksen Saksan kansallisen lain direktiivin mukaisuudesta, mitä ei käsitellä tässä esityksessä.
[128] Tuomion kohta 31.
[129] Tuomion kohta 37.
[130] Tuomion kohta 39.

on halunnut maininnalla epäsuorasta tunnistamisesta tehdä selväksi, että tiedon luokitteleminen henkilötiedoksi ei edellytä, että tämä tieto yksin mahdollistaa rekisteröidyn tunnistamisen.[131] Tämän jälkeen EUT totesi tietosuojadirektiivin johdanto-osan 26 perustelukappaleeseen viitaten, että sen määrittämiseksi, onko henkilö tunnistettavissa, olisi otettava huomioon kaikki kohtuullisesti toteutettavissa olevat keinot, joita joko rekisterinpitäjä tai joku muu voi kyseisen henkilön tunnistamiseksi käyttää.[132]

Kaikkien tietojen, joiden perusteella rekisteröity voidaan tunnistaa, ei tarvitsisi olla yhden tahon hallussa, jolloin verkkomediapalvelujen tarjoajan osalta ei voida sulkea pois kyseisen IP-osoitteen luokittelemista henkilötiedoksi.[133] EUT:n ratkaisussa tärkeimmäksi näyttääkin muodostuneen sen arvioiminen, onko mahdollisuus yhdistää dynaaminen IP-osoite internetyhteyden tarjoajan hallussa oleviin lisätietoihin rekisteröidyn tunnistamiseksi *kohtuullisesti toteutettavissa oleva keino*. Tässä arvioinnissa EUT viittaa julkisasiamiehen ratkaisuehdotukseen[134], jonka mukaan keinot eivät ole kohtuullisesti toteutettavissa ainakaan, kun rekisteröidyn tunnistaminen on laissa kielletty tai se ei ole käytännössä toteutettavissa esimerkiksi siitä syystä, että se veisi suhteettomasti aikaa ja aiheuttaisi suhteettomasti kustannuksia ja työtä, minkä seurauksena tunnistamisen riski näyttäytyy käytännössä merkityksettömänä.[135] EUT on siis perusteluissaan ottanut huomioon sekä objektiivisen että suhteellisen perusteen mukaiset näkökannat.

EUT päätyy toteamaan, että kyseisessä tapauksessa näyttää olevan olemassa laillisia ja kohtuullisesti toteutettavissa olevia keinoja rekisteröidyn tunnistamiseksi. Näin on erityisesti siksi, että verkkomediapalvelujen tarjoaja voisi esimerkiksi kyberhyökkäystilanteessa kääntyä toimivaltaisen viranomaisen puoleen ja pyytää tätä kautta käytettäväkseen rekisteröidyn tunnistamisen kannalta tarpeelliset lisätiedot.[136] Lopuksi EUT vastaa ennakkoratkaisukysymykseen kaiken edellä esitetyn mukaan seuraavasti:

---

[131] Tuomion kohta 41.
[132] Tuomion kohta 42.
[133] Ibid. 43 ja 44.
[134] Julkisasiamiehen ratkaisuehdotus Manuel Campos Sánchez-Bordona 12 päivänä toukokuuta 2016 (1), kohta 68.
[135] Tuomion kohta 46.
[136] Tuomion kohta 47.

> *”direktiivin 95/46 2 artiklan a alakohtaa on tulkittava siten, että dynaaminen IP-osoite, jonka verkkomediapalvelujen tarjoaja tallentaa henkilön käydessä tämän palveluntarjoajan yleisön saataville asettamalla internetsivustolla, on tähän palveluntarjoajaan nähden kyseisessä säännöksessä tarkoitettu henkilötieto, jos palveluntarjoajalla on käytettävissään oikeudelliset keinot, joiden perusteella se voi tunnistaa kyseisen henkilön sellaisten lisätietojen avulla, jotka ovat tämän henkilön internetyhteyden tarjoajan käytettävissä.”*[137]

EUT ei ole perusteluissaan arvioinut kohtuullisuuden kriteeriä erikseen, mutta näyttäisi siltä, että ainakin tuomioistuinteitse hankittava lupa saada internetyhteyden tarjoajalta lisätietoja täyttäisi kohtuullisen keinon määritelmän. Perusteluissa mainitaan nimenomaisesti, että mikäli tunnistamisen tekemiseksi on olemassa oikeudellisia keinoja, tunnistettavuuskriteeri täyttyy. EUT on nähdäkseni epäonnistunut perustelemaan niitä kohtuullisen todennäköisiä keinoja, joilla tunnistamisen täytyy tapahtua. Perustelujen mukaan nimittäin henkilö olisi aina tunnistettavissa, ellei tunnistamista ole laissa nimenomaisesti kielletty ja tunnistamisen mahdollistavia lisätietoja olisi saatavilla muualta. Erityisesti uusien tietojenkäsittelytekniikoiden ja tietojen määrän valtavan kasvun vuoksi voitaisiin väittää, että tällaisia lisätietoja olisi aina saatavissa jostakin muualta. EUT:n olisi pitänyt arvioida seikkaperäisemmin tunnistamisen toteuttamiseen liittyviä objektiivisia tekijöitä, kuten tunnistamisesta aiheutuvia kuluja ja tunnistamiseen tarvittavaa aikaa.

3.4.3 Johtopäätökset kriteerin tulkinnasta EU-oikeuskäytännössä

EUT:n melko vähäisestä kriteeriä koskevasta oikeuskäytännöstä on vaikeaa vetää pitkälle meneviä johtopäätöksiä siitä, miten kriteeriä käytännössä tulisi tulkita. Tämä johtuu ainakin siitä, että tilanteet, joissa tunnistettavuuskriteerin täyttymistä arvioidaan, vaihtelevat huomattavasti keskenään ja jokaista tapausta on tarkasteltava nimenomaisesti siltä kannalta, mitä kulloinkin kyseessä olevan tiedon haltijan voidaan olettaa kykenevän päättelemään jostakin tiedosta ja millaisia lisätietoja sillä on mahdollisesti hankittavissa kohtuullisin keinoin muualta. Tiedon määrän jatkuvasti kasvaessa yhteiskunnassa ja etenkin yritysten hyödyntämien valtavien tietovarantojen avulla voi olla vaikeaa osoittaa,

[137] Tuomion kohta 49.

että sellaisia lisätietoja, joiden avulla yksilö kyettäisiin tunnistamaan, ei olisi ilman kohtuutonta vaivaa saatavilla jostakin muualta.

Kirjallisuudessa on esitetty eriäviä kantoja sen suhteen, nojautuuko EUT tulkinnassaan objektiiviseen vai suhteelliseen perusteeseen, kun se määrittelee dynaamisen IP-osoitteen henkilötiedoksi. Tulkinnan voidaan katsoa ilmentävän objektiivisen perusteen käyttöä, koska EUT:n perustelut korostavat sitä, että henkilö voi olla tunnistettavissa myös jonkun muun kuin rekisterinpitäjän toimesta. Toisaalta voidaan myös esittää, että tulkinta on molempien perusteiden yhdistelmä, jossa hyödynnetään sekä objektiivisen että suhteellisen perusteen sisältöjä punninnalla.

Tuomion perusteluissa on selvästi viitteitä molempien perusteiden käytöstä. EUT käyttää tulkinnassaan objektiivista perustetta siltä osin, kun se toteaa, että tunnistaminen voi tapahtua myös muun henkilön toimesta yhdistelemällä tietoja rekisterinpitäjän hallussa oleviin tietoihin. Suhteellista perustetta se käyttää perustelemalla, että esimerkiksi tunnistaminen, joka on tullut mahdolliseksi lain vastaisesti hankittujen tietojen ansiosta, ei voi täyttää tunnistettavuuskriteeriä. Tämä ei voisi kuulua niihin *kohtuullisen todennäköisiin keinoihin*, joita joku muu voi käyttää henkilön tunnistamiseen. Tällainen tulkinta on myös linjassa tietosuojatyöryhmän kannan kanssa, jonka mukaan pelkkä hypoteettinen mahdollisuus yksilön erottamiseen muista ei tekisi yksilöä tunnistettavissa olevaksi.[138]

Nähdäkseni EUT hyödyntää perusteluissaan sekä objektiivisen että suhteellisen perusteen mukaista tulkintaa tunnistettavuudesta, mutta suurempi painoarvo on annettu objektiiviselle perusteelle. Suhteellista perustetta EUT on käyttänyt lähinnä vain rajatakseen pois sen, että henkilö voisi olla tunnistettavissa internetsivuston ylläpitäjään nähden, jos joku muu olisi hankkinut tunnistamiseen tarvittavat lisätiedot lainvastaisesti. Kaikki laillisin keinoin hankitut lisätiedot kenen tahansa muun toimesta olisi siis otettava huomioon. Tämän vuoksi näen, että EUT:n käytännössä suurempi painoarvo on annettu objektiivisen perusteen mukaiselle tulkinnalle ja suhteellista perustetta käytetään lähinnä

[138] Tietosuojatyöryhmän lausunto 04/2007 s. 15.

tapauskohtaisesti rajaamaan lainvastaisesti hankitut tiedot pois *kohtuullisen todennäköisten keinojen* kategoriasta. Tulkinnassa on kuitenkin otettava huomioon myös tunnistamisen todennäköisyys eli liian epätodennäköisiä tunnistamisen keinoja ei tulisi huomioida kriteerin täyttymistä arvioitaessa. EUT:n perustelujen valossa epäselväksi kuitenkin jää, miten tätä tunnistamisen todennäköisyyttä tulee arvioida.

3.5 Johtopäätökset kriteerin sisällöstä ja tulkinnasta

Edellä esitetyn johdosta voidaan todeta, että tunnistettavuuskriteeriä olisi tulkittava siten, että huomioon on otettava kaikki kohtuullisen todennäköiset keinot, joita rekisterinpitäjä tai joku muu voi käyttää henkilön tunnistamiseksi, mutta ainoastaan siinä määrin kuin objektiivisten tekijöiden avulla voidaan osoittaa, ettei tunnistaminen ole käytännöllisesti mahdotonta tai vaadi suhteetonta vaivaa eikä tunnistamisen tekemiseksi ole laillista estettä. Yleisen tietosuoja-asetuksen sanamuodon ja EUT:n antamien tulkintaohjeiden perusteella tunnistettavuuskriteeriä tulkitaan hyödyntäen sekä objektiivisen että suhteellisen perusteen mukaista tulkintaa, kuitenkin niin, että objektiiviselle perusteelle annetaan suurempi painoarvo. Henkilötiedon käsite on tarkoitettu laajaksi ja näin ollen myös tunnistettavuuskriteeriä tulee tulkita laajentavasti siten, että henkilötiedoksi katsotaan kaikki tiedot, joilla voi olla vaikutuksia luonnolliseen henkilöön ja tämän oikeuksiin.

Suhteellisen perusteen mukaista tulkintaa voidaan kuitenkin tapauskohtaisesti hyödyntää tapauksissa, joissa tiedon luokitteleminen henkilötiedoksi olisi selkeästi perusteetonta. Tällaisia tapauksia voisivat olla esimerkiksi tilanteet, joissa henkilö olisi tunnistettavissa ainoastaan suhteettoman vaikeasti toteutettavin keinoin tai tunnistaminen muutoin olisi erittäin epätodennäköistä.

Tällainen tulkinta on linjassa myös tietosuojatyöryhmän esittämien näkökantojen kanssa. Työryhmä on lausunnoissaan nimenomaisesti painottanut henkilötiedon käsitteen laajan tulkitsemisen tärkeyttä, mutta myös sitä, että käsitteen soveltamisalaa ei laajennettaisi liikaa. Tasapaino olisi siis löydettävä

tapauskohtaisella harkinnalla. Tunnistettavuuskriteerin tulkinta kohtaa kuitenkin aivan uudenlaisia ongelmia teknologioiden kehittyessä, kun tietoja käsitellään eri tavoilla ja uudenlaisiin tarkoituksiin. Seuraavassa luvussa tarkastelen tunnistettavuuskriteerin ongelmia näissä uudenlaisissa tilanteissa.

4 Tunnistettavuuskriteerin ongelmallisuudesta kehittyvän teknologian aikakaudella

4.1 Johdanto

Eurooppalainen tietosuojasääntely on sidottu henkilötiedon käsitteeseen. Tunnistettavuuskriteeri on yksi keskeisimmistä tietosuojasääntelyyn liittyvistä käsitteistä, jonka avulla arvioidaan, onko kyse henkilötiedosta. Henkilötietojen suojaa koskevien säädösten soveltamisala määräytyy sen perusteella, onko asiassa kysymys henkilötiedoista vai ei. Lähtökohtana tietosuojasääntelyn taustalla on olettamus siitä, että mikäli kyse ei ole henkilötiedoista, ei ole olemassa myöskään minkäänlaista tietosuojauhkaa tai suojan tarvetta. Kuitenkaan tunnistettavuudesta ei ole olemassa mitään yksiselitteistä määritelmää tietosuojaa koskevassa lainsäädännössä, oikeuskäytännössä tai oikeuskirjallisuudessa.

Tietojenkäsittelytieteissä on osoitettu[139], että monissa tapauksissa niin kutsuttu ei-henkilötieto voidaan kuitenkin yhdistää tiettyä henkilöä koskevaksi tiedoksi tietotekniikan avulla. Myös jo kerran pseudonymisoitu tai anonymisoitu tieto voidaan tietoja yhdistelemällä yhdistää tiettyyn henkilöön.

Big datalla viitataan erittäin suurten, järjestelemättömien, jatkuvasti lisääntyvien tietomassojen keräämistä, säilyttämistä, jakamista, etsimistä, analysointia sekä esittämistä tilastotiedettä ja tietotekniikkaa hyödyntäen.[140] Big data on siis yhteisnimitys valtaisille datamäärille, joiden yhteydessä ei voida soveltaa perinteisiä datan hallinnointitapoja.[141] Big datan hyödyntäminen luo positiivisia mahdollisuuksia yhteiskunnassa, mutta se voi vaikuttaa myös haitallisesti

[139] Schwartz & Solove 2011, s. 1841.
[140] Hilbert 2013.
[141] Srinivasa 2012.

yksilöiden oikeuksiin, erityisesti yksityisyyden suojaan ja henkilötietojen suojaan. Tietosuojalainsäädäntöä pidetään yleisesti keinona, jolla tällaisia negatiivisia vaikutuksia voidaan torjua. EU:n tietosuojasääntelyn lähempi tarkastelu kuitenkin osoittaa, että sen lähestymistavoissa on huomattavia puutteita koskien big dataa ja sen käsittelyä. EU:n tietosuojasääntely rakentuu tunnistettavuuden varaan, jolloin se tulee sovellettavaksi vain osaan tällaista tietojenkäsittelyä, joka tapahtuu big data -ympäristössä.[142]

Tässä luvussa käsittelen tunnistettavuuskriteerin haasteita erityisesti tietojen salaamisen, anonymisoinnin ja pseudonymisoinnin kannalta. Erityisesti pidän ongelmallisena sitä, että nykyisen sääntelyn ja sitä koskevan tulkinnan johdosta tietosuojasääntely saattaa tulla sovellettavaksi liian laajaan joukkoon tietoja. Lisäksi jatkossa teknologian kehittyessä tämä tietojen joukko tulee todennäköisesti vain kasvamaan. Tietoturvatoimenpiteiden, kuten tietojen salaamisen toteuttaminen voi osoittautua lähes mahdottomaksi, mikäli tunnistettavuuskriteerin tulkintaa laajennetaan entisestään. Lisäksi tarkastelen tilanteita, joissa luonnollisen henkilön perusoikeuksiin ja -vapauksiin puututaan ilman, että henkilö olisi tiedoista tunnistettavissa. Katson tämän olevan, tunnistettavuuskriteerin liiallisen laajentamisen lisäksi, kriteerin suurimpia ongelmakohtia.

4.2 Henkilötietojen anonymisointi, salaaminen ja pseudonymisointi

4.2.1 Anonyymi henkilötieto ja anonymisointitekniikat

EU:n tietosuojasääntely rakentuu olettamalle, että tiedot ovat joko henkilötietoja tai anonyymejä. Tietosuojanormisto tulee kokonaisuudessaan sovellettavaksi kaikkiin henkilötiedoksi katsottavien tietojen käsittelyyn, kun taas ei-henkilötietoihin normistoa ei sovelleta lainkaan.[143] Tämä kahtiajako todetaan tietosuoja-asetuksen johdanto-osan kohdassa 26:

[142] Oostveen 2016, s. 299.
[143] Urgessa 2016, s. 528.

> Tietosuojaperiaatteita olisi sovellettava kaikkiin tietoihin, jotka koskevat tunnistettua tai tunnistettavissa olevaa luonnollista henkilöä. […] Tämä asetus ei […] koske […] anonyymien […] tietojen käsittelyä.[144]

Olettama siitä, että jokin tieto voidaan aina kategorisoida joko henkilötiedoksi tai anonyymiksi, on tietoyhteiskunnan aikakaudella ongelmallinen. Ensinnäkin, sama tietoaineisto voi olla toisen henkilön hallussa anonyymiä, kun taas toisen hallussa henkilötietoa. Toiseksi, tietotekniikan kehitys mahdollista enenevissä määrin anonyymin tiedon uudelleentunnistamisen. Tiedon anonymiteetti riippuu tarkastelun ajankohdasta ja kontekstista. Tiedot voivat olla tänään anonyymejä, mutta huomenna henkilötietoja. Lisäksi tietojen luonne muuttuu, kun niitä tarkastellaan eri olosuhteissa.[145]

Anonymisoidun henkilötiedon määritelmä löytyy EU:n yleisen tietosuoja-asetuksen johdanto-osan 26 kohdasta. Sen mukaan anonyymeillä tiedoilla tarkoitetaan kaikkia niitä tietoja, jotka eivät liity tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön, tai henkilötietoja, joiden tunnistettavuus on poistettu siten, ettei rekisteröidyn tunnistaminen ole tai ole enää mahdollista. Näin ollen anonymisoidut henkilötiedot jäävät EU:n yleisen tietosuoja-asetuksen soveltamisalan ulkopuolelle, eikä tietosuojaperiaatteita siten sovelleta tällaisiin tietoihin.

Sen määrittämiseksi, onko jokin tieto anonyymiä vai ei, on tietoja jälleen tarkasteltava tunnistettavuuskriteerin kautta. Huomioon on siis otettava kaikki keinot, joita joko rekisterinpitäjä tai muu henkilö voi kohtuullisen todennäköisesti käyttää luonnollisen henkilön tunnistamiseen suoraan tai välillisesti. Huomioon olisi otettava kaikki objektiiviset tekijät, esimerkiksi tunnistamisesta aiheutuvat kulut ja tunnistamiseen tarvittava aika sekä käytettävissä oleva teknologia sekä tekninen kehitys. Huomioon otettavien tekijöiden laajuus ja varsinkin vaatimus käytettävän teknologian ja teknisen kehityksen huomioon ottamisesta muodostuu ankaraksi vaatimukseksi, kun asiaa tarkastellaan yritysten ja muiden taloudellisesti tietoja hyödyntävien organisaatioiden käytännöissä.

[144] Tietosuojadirektiivin johdanto-osan kohtaan 23 sisältyy lähes samansisältöinen kahtiajaon määrittely.
[145] Urgessa 2016, s. 528.

Henkilötietojen anonymisointi siis vapauttaa rekisterinpitäjän tai muun henkilön kaikista henkilötietojen käsittelyyn liittyvistä vaatimuksista. Anonymisointi on kuitenkin haastavaa toteuttaa siten, että yksilöiden tunnistettavuus katoaa, mutta samalla mahdollistetaan tietojen myöhempi hyödynnettävyys. Tämä johtuu siitä, että anonymisoinnin on oltava *peruuttamatonta* sulkeakseen pois tietosuoja-asetuksen soveltamisen.[146] Kirjallisuudessa on monesti esitetty ja todistettu, että anonymisoinnin vaikeustaso riippuu keskeisesti siitä, minkä tyyppisiä tietoja aineistossa on ja miten montaa eri muuttujaa se sisältää. Laajat ja yleisesti saatavissa olevat taustatiedot ja jatkuvasti kehittyvät tiedonlouhintatyökalut muodostavat haasteen luotettavalle anonymisoinnille.[147]

Viimeaikaisessa tieteellisessä keskustelussa on esitetty vastakkaisia näkemyksiä sen suhteen, onko tällaisen täydellisen anonymisoinnin tason saavuttaminen ylipäätään mahdollista. Monet ovat olleet sitä mieltä, että ainakaan ilman suhteetonta vaivaa ja kustannuksia täydellinen anonymisointi on nykypäivänä lähes aina mahdotonta.[148] Tietojen anonyymius on ennemminkin oikeudellinen kuin todellinen käsite. Luottokorttitietoja koskeneessa tutkimuksessa tutkijat onnistuivat osoittamaan, että jo neljän satunnaisen aika- ja paikkatiedon tunteminen mahdollisti identifioida 90 % henkilöistä ja täten paljastamaan kaikki heidän tietonsa. Yksinkertaisesti anonymisoitu metadata on siis helposti uudelleentunnistettavissa.[149]

Kun tietosuoja-asetuksen aineellinen soveltamisala määräytyy edellä kuvatusti henkilötiedon ja anonyymin tiedon kahtiajakoon perustuvan määrittelyn kautta, on suuri mahdollisuus yksityisyyden ja henkilötietojen suojan loukkauksille olemassa anonymisoitujen tietojen uudelleentunnistamisen vuoksi. Oikean tasapainon löytäminen yksityisyyden suojan ja tietojen hyödynnettävyyden välillä on ensiarvoisen tärkeää, jotta voidaan realisoida metadatan suuri potentiaali.[150] Mikäli EU:n viimeaikainen henkilötiedon käsitteen laajentavan tulkinnan trendi jatkuu, on mielestäni hyvin todennäköistä, että tietojen hyödynnettävyys vaikeutuu ja samalla tehdään itse asetuksen tavoitteiden toteutumisesta

[146] Tietosuojatyöryhmän lausunto 5/2014, s. 6.
[147] Bäck & Keränen 2017, s. 3.
[148] Salokannel 2016, s. 544.
[149] de Montjoye ym. 2015, s. 537.
[150] de Montjoye ym. 2015, s. 539.

hankalampaa. Tällöin ei nimittäin voida esimerkiksi lisätä talouden toimijoiden luottamusta tietosuojasääntelyn toimivuuteen tai lisätä oikeusvarmuutta. Lisäksi on muistettava, että oikeus henkilötietojen suojaan ei ole absoluuttinen, joten myös tietojen hyödynnettävyys on otettava vakavana tekijänä sen arvioimisessa, kuinka pitkälle EU:ssa ollaan valmiit menemään yksityisyyden ja henkilötietojen suojaamisessa. Liian suuri yksityisyyden ja individualismin korostaminen itse asiassa vaikeuttaa tietosuojasääntelyn tavoitteiden toteuttamista ja samalla haittaa taloudellisten toimijoiden toimintaedellytyksiä EU-alueella.

Yksi vaihtoehtoinen lähestymistapa anonymisointiin on riskiperusteinen tietoturva-arviointi. Riskiperusteisessa tietoturva-arvioinnissa anonyymiutta tulee tulkita laajemmin:

> It is not whether data subjects are anonymous or not, but to what degree they are anonymous. The interpretation of anonymity must also account for the context of data sharing, such as who will be working with the data, who will have access to the data, how this will be ensured (via security and privacy protocols), and what contracts or data-sharing agreements must be signed.[151]

Riittävän anonymiteetin taso saattaa vaihdella anonymisoidun tiedon käyttöyhteydestä ja sen taustalla olevan henkilötiedon luonteesta riippuen. Esimerkiksi sosiaali- ja terveystietojen luonne arkaluonteisina henkilötietoina tulisi ottaa huomioon niitä anonymisoidessa. Onkin mielestäni erikoista, että tietosuoja-asetuksessa korostetaan arkaluonteisten tietojen erityisasemaa eli tehdään sinänsä ero eri tyyppisten tietojen välille suojan tarpeen näkökulmasta, mutta anonymisoinnin sääntelyyn ei ole nähty tarpeelliseksi sisällyttää tällaista riskiarviointia.

Esimerkiksi tietosuoja-asetuksen johdanto-osan kohdassa 37 säädetään, että *”[h]enkilötiedot, jotka ovat erityisen arkaluonteisia perusoikeuksien ja -vapauksien kannalta, ansaitsevat erityistä suojelua, koska niiden käsittelyn asiayhteys voisi aiheuttaa huomattavia riskejä perusoikeuksille ja -vapauksille*”. Samalla tavoin voitaisiin edellyttää vahvempaa anonymiteettiä tällaisilta arkaluonteisilta henkilötiedoilta ollakseen anonyymejä tietosuoja-asetuksen tarkoittamassa merkityksessä. Tietosuoja-asetuksen johdanto-osan 26 kohdassa mainittujen kohtuullisen todennäköisesti käytettävissä olevien

[151] Arbuckle ym. 2017, s. 169.

tunnistamiskeinojen lisäksi arkaluonteisia henkilötietoja anonymisoitaessa ja edelleen luovutettaessa tulisi anonymiteettiä arvioida laajasti yhteydessä siihen kontekstiin, jossa näitä tietoja tullaan hyödyntämään.

Paul Ohm on esittänyt, että kaikki tiedot ovat uudelleentunnistettavissa, osa vaikeammin ja osa helpommin. Hänen mukaansa anonymisoinnin tuoma suoja yksityisyydelle on vain harhakuva ja ajatus, jonka mukaan voisimme erottaa tietojoukosta ne yksittäiset tiedot, jotka ovat helpommin yhdistettävissä johonkin yksilöön, on menettänyt tieteellisen pohjan ja siten tämä ajattelumalli olisi unohdettava.[152] Ohmin mukaan riskiä yksityisyyden suojan loukkaukselle tulisi arvioida tapauskohtaisesti tietoja luovutettaessa tai tällaista tietoa sisältävää tietoaineistoa julkaistaessa. Ohmin niin sanotusta riskitestistä enemmän jaksossa 4.4.

EU:n tietosuojatyöryhmän lausunnossa korostetaan, että vaikka tietosuojalakeja ei sovelleta anonyymeihin tietoaineistoihin, voi niistä silti seurata yksityisyyden loukkaus. Tämän vuoksi anonyymeihin tietoihin on suhtauduttava erityisen varovasti aina, kun tehdään yksilöihin välillisestikin vaikuttavia päätöksiä.[153] Henkilötietojen anonyymius on myös dynaamista: se mikä eilen oli anonyymiä, ei sitä välttämättä tänään ole. Tietojen haltijoiden tulee siten säännöllisin väliajoin arvioida kriittisesti käsittelemiensä tietomassojen status eli voidaanko tiedot laajasti arvioiden lukea edelleen riittävällä tasolla anonyymeiksi.[154]

4.2.2 Henkilötietojen salaamisen asema tietosuojasääntelyssä

Henkilötietojen salaaminen[155] on tietoturvatekniikka, jonka avulla tiedot muunnetaan ei-ymmärrettävään muotoon ja vain osapuolet, joilla on hallussaan koodauksen purun mahdollistava mekanismi ja salausavain pääsevät käyttämään tietoja.[156] Tietojen salaus voi muodostua merkittäväksi keinoksi

[152] Ohm 2010, s. 1732.
[153] Tietosuojatyöryhmän lausunto 5/2014, s. 11.
[154] Salokannel 2016, s. 535. Ks. myös EU:n tietosuojatyöryhmän lausunto 5/2014, s. 4.
[155] Käytetään myös esimerkiksi nimitystä kryptaaminen.
[156] Enisa 2015, s. 38.

varmistua riittävästä tietosuojan tasosta, varsinkin pilvipalveluiden yhteydessä. Kun rekisterinpitäjä salaa tiedot ennen niiden siirtämistä pilvipalveluun, tietoja pidetään henkilötietoina todennäköisesti vain rekisterinpitäjään nähden, sillä vain tällä on hallussaan tietojen purkuun vaadittava salausavain.[157] Henkilötietojen salaus varmistaa, että tietoihin ei pääse käsiksi ulkopuoliset tahot ja täten rekisterinpitäjä säilyy ainoana tahona, joka kykenee tunnistamaan yksittäiset luonnolliset henkilöt tietomassasta.

Tietosuoja-asetuksen lopullinen versio ei sisällä tarkkaa määritelmää henkilötietojen salaamisesta. Euroopan parlamentin ehdotuksessa[158] salatuilla tiedoilla tarkoitetaan ”*henkilötietoja, jotka on teknisten suojatoimenpiteiden avulla muutettu sellaiseen muotoon, että ne eivät ole sellaisten henkilöiden ymmärrettävissä, joilla ei ole lupaa päästä tietoihin*.” Tietosuoja-asetuksessa salaaminen mainitaan kuitenkin useassa kohdassa keinona varmistaa tietosuojasääntelyn määräysten mukaisuus. Asetuksen artiklassa 32 (1 a) säädetään, että salaus on yksi asianmukainen tekninen ja organisatorinen toimenpide varmistaa tietojen käsittelyn turvallisuus. Tämä ei kuitenkaan määritä tietosuoja-asetuksen soveltamisalan laajuutta suhteessa henkilötietojen salaamiseen, vaan kyseessä on henkilötietojen suojaamisen kannalta tarpeellisten toimien määrittelystä.

Salauksesta todetaan myös asetuksen 34 (3 a) artiklassa, että rekisterinpitäjän ei tarvitse ilmoittaa rekisteröidylle tapahtuneesta tietoturvaloukkauksesta, mikäli se on toteuttanut henkilötietojen salauksen teknisenä ja organisatorisena suojatoimenpiteenä. Lisäksi salaus mainitaan yhtenä asianmukaisena suojatoimena, kun arvioidaan sellaisen käsittelyn sääntelyn mukaisuutta, joka tapahtuu muuta tarkoitusta varten kuin sitä, jonka vuoksi tiedot on alun perin kerätty.[159]

4.2.3 Salatut henkilötiedot – pseudonymisoitua vai anonyymiä tietoa

[157] Spindler & Schmechel, s. 169.
[158] Euroopan parlamentin ehdotus yleiseksi tietosuoja-asetukseksi 4 (2 b) artikla. Ehdotuksessa käytetään termiä ”suojatut tiedot”.
[159] Tietosuoja-asetuksen 6 (4 e) artikla.

Tietosuoja-asetus ei määrittele mitä salatulla tiedolla tarkoitetaan, mutta asiaa voidaan valaista vertaamalla salattua tietoa pseudonymisoituun tietoon, joka sen sijaan on määritelty asetuksessa. Tietojen salauksen määrittäminen on tärkeää siksi, että tiedetään, onko henkilötietojen salaus sellainen suojatoimi, joka anonymisoi tiedot vai ainoastaan pseudonymisoi ne.[160] Mikäli tietojen salaaminen katsottaisiin vain yhdeksi pseudonymisoinnin keinoksi, se tarkoittaisi, että salaaminen ei estäisi tietosuojasääntelyn soveltumista kyseisiin tietoihin. Tämä taas poistaisi rekisterinpitäjiltä ja muilta tietoja käsitteleviltä tahoilta kannustimet salata tietoja lainkaan, jolloin myös tietosuojasääntelyn tavoitteiden toteutuminen kokonaisuudessaan vaikeutuisi.

Anonymisointia pidetään tärkeänä turvatoimenpiteenä luonnollisten henkilöiden yksityisyyden suojan kannalta tietojen käsittelyssä[161], mutta tietosuoja-asetuksessa ei ole avattu käsitteen merkityssisältöä millään tavalla ja muutoinkin maininnat siitä ovat jääneet hyvin vähäisiksi. Asetuksen ainoa anonyymiä tietoa käsittelevä kohta on johdanto-osassa kohdassa 26:

> ”Tietosuojaperiaatteita ei […] pitäisi soveltaa anonyymeihin tietoihin eli tietoihin, jotka eivät liity tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön, tai henkilötietoihin, joiden tunnistettavuus on poistettu siten, ettei rekisteröidyn tunnistaminen ole tai ei ole enää mahdollista. Tämä asetus ei tämän vuoksi koske tällaisten anonyymien, muun muassa tilasto- tai tutkimustarkoituksia varten käytettävien tietojen käsittelyä.”

Asetusta ja tietosuojaperiaatteita ei siis sovelleta anonyymeihin tietoihin, mutta ongelmaksi muodostuu jälleen tunnistettavuuskriteerin sisältö ja tulkinta, koska tieto on anonyymiä ainoastaan, kun henkilö ei enää ole tiedoista tunnistettavissa. Maininnan tietojen anonymisoinnista on katsottu merkitsevän suhteellisen perusteen mukaista tunnistettavuuden arvioimista. Nimittäin täysin anonyymiä tietoa ei olisi olemassa jos noudatettaisiin täysin objektiivisen perusteen mukaista arviointia, kun otetaan huomioon nykyteknologian mahdollistamat uudelleentunnistamisen mahdollisuudet ja tietojen yhdistelemisen big data –ympäristössä.[162]

[160] Spindler & Schmechel, s. 170.
[161] Spindler & Schmechel, s. 170.
[162] Ibid.

Pseudonymisoiduista tiedoista säädetään tietosuoja-asetuksen 4 (5) artiklassa seuraavasti[163]:

> "'pseudonymisoimisella' [tarkoitetaan] henkilötietojen käsittelemistä siten, että henkilötietoja ei voida enää yhdistää tiettyyn rekisteröityyn käyttämättä lisätietoja, edellyttäen että tällaiset lisätiedot säilytetään erillään ja niihin sovelletaan teknisiä ja organisatorisia toimenpiteitä, joilla varmistetaan, ettei henkilötietojen yhdistämistä tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön tapahdu."

Pseudonymisoinnin tarkempaa merkitystä avaa myös asetuksen johdanto-osan kohta 28:

> "Pseudonymisoinnin soveltaminen henkilötietoihin voi vähentää asianomaisiin rekisteröityihin kohdistuvia riskejä sekä auttaa rekisterinpitäjiä ja henkilötietojen käsittelijöitä noudattamaan tietosuojavelvoitteitaan. "Pseudonymisoinnin" nimenomaisella sisällyttämisellä tähän asetukseen ei ole tarkoitus sulkea pois mitään muita tietosuojatoimenpiteitä."

Tämän perusteella pseudonymisoitu tieto on siis edelleen henkilötietoa, koska yhteys luonnollisen henkilön ja tiedon välillä säilyy.[164] Pseudonymisointi on siis keino, jolla voidaan vähentää todennäköisyyttä sille, että henkilö on tunnistettavissa tietojen perusteella, mutta tietosuoja-asetus kaikkine velvoitteineen edelleen soveltuu myös pseudonymisoituihin tietoihin.[165] Pseudonymisointia pidetään lähinnä tietoturvatoimenpiteenä[166], jota asetuksen monessa kohdassa suositellaan käytettäväksi. Esimerkiksi asetuksen 32 (1 a) artiklassa sitä pidetään yhtenä teknisenä ja organisatorisena toimenpiteenä varmistamaan riskiä vastaava turvallisuustaso henkilötietojen käsittelyssä.

Ensi näkemältä pseudonymisoinnin määritelmä vaikuttaa hyvin samankaltaiselta kuin salauksen määritelmä esimerkiksi Euroopan parlamentin valiokunnan ehdotuksessa. Onko henkilötietojen salaamisella mitään merkitystä tietosuojasääntelyn sovellettavuuden kannalta vai samaistetaanko salatut tiedot pseudonymisoituihin tietoihin siten, että tietosuojasääntely soveltuu myös niihin täysimääräisesti?

Tunnistettavuuskriteerin tulkinnan rajanveto objektiivisen ja suhteellisen perusteen välillä tulee jälleen arvioitavaksi, kun tarkastellaan salattujen tietojen

[163] Tietosuoja-asetusta edeltävässä tietosuojadirektiivissä ei ollut mainintaa pseudonymisoiduista tiedoista.
[164] Spindler & Schmechel, s. 171,
[165] Ibid.
[166] Spindler & Schmechel, s. 171; Borgesius, s. 256.

asemaa tietosuojasääntelyn kontekstissa. Kun henkilötietoja salataan, luodaan samalla yleensä salausavain, joka mahdollistaa tietojen käsittelyn ja ymmärrettävyyden ainoastaan salausavaimen haltijalle. Salausavaimen haltijaan nähden salattu tieto kiistattomasti säilyy henkilötietona[167], koska tämä kykenee purkamaan salauksen ja paljastamaan tiedot, joiden avulla yksittäisten henkilöiden tunnistaminen on mahdollista.

Salausavaimen voidaan katsoa olevan asetuksen 4 (5) artiklan tarkoittama erillään säilytetty lisätieto, johon sovelletaan teknisiä ja organisatorisia toimenpiteitä, joilla varmistetaan, ettei henkilötietojen yhdistämistä tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön tapahdu.[168] Tietojen salaamisen on siis katsottava olevan ainakin keino pseudonymisoida tietoja.[169] Tulkinnanvaraista kuitenkin on, voidaanko salattujen tietojen katsoa olevan myös anonyymiä tietoa.

Kuten edellä on todettu, salatut tiedot ovat ainakin salausavaimen haltijalle edelleen henkilötietoja. Salattujen tietojen luonteen arvioiminen kuitenkin muuttuu hankalammaksi, kun arvioidaan niitä sellaisen henkilön hallussa, jolla ei ole käytettävissään salausavainta, eikä siten mahdollisuutta ainakaan suoraan tunnistaa niistä luonnollista henkilöä. Oikeuskirjallisuudessa on esitetty kantoja, joiden mukaan salatut tiedot olisivat henkilötietoja myös sellaisiin nähden, joilla ei ole käytettävissään salausavainta.[170]

Tietosuojatyöryhmä on katsonut lausunnossaan, että on yleinen virhe luulla pseudonymisoituja tietoja anonyymeiksi tiedoiksi.[171] Työryhmä lausuu, että monissa tapauksissa voi olla yhtä helppoa tunnistaa yksilö peitenimillä suojatusta tietoaineistosta kuin alkuperäisistä tiedoista ja että tietoaineiston anonymisointi vaatii lisätoimia. Lisäksi tietosuoja-asetuksen johdanto-osan kohdassa 26 todetaan, että ”*[p]seudonymisoidut henkilötiedot, jotka voitaisiin yhdistää*

[167] Spindler & Schmechel, s. 171; Tietosuojatyöryhmän lausunto 5/2014, s. 29.
[168] Spindler & Schmechel, s. 171.
[169] Spindler & Schmechel, s. 171; Tietosuojatyöryhmän lausunto 5/2014, s. 21; Esayas, s. 8.
[170] Esayas, s. 8.
[171] Tietosuojatyöryhmän lausunto 5/2014 s. 21. Suomenkielisessä versiossa käytetään nimitystä ”peitenimellä suojattu tietoaineisto”.

luonnolliseen henkilöön lisätietoja käyttämällä, olisi katsottava tiedoiksi, jotka koskevat tunnistettavissa olevaa luonnollista henkilöä".

Objektiivista perustetta noudattaen salattu tietoaineisto olisi aina katsottava henkilötiedoksi, koska joku pystyisi aina purkamaan salauksen ja pääsemään käsiksi henkilön tunnistamisen mahdollistaviin lisätietoihin. Näin pystyisi tekemään ainakin salausavaimen haltija ja kuka tahansa, jolla on käytettävissään riittävät ajalliset ja taloudelliset resurssit sekä laskentateho salauksen purkamiseksi, sillä mikään nykyinen salaustekniikka ei ole täysin turvallinen kaikilta osin. Objektiivista perustetta noudattaen tietojen salaamisen olisi katsottava olevan ainoastaan tekninen ja organisatorinen toimenpide, jolla voidaan osaltaan varmistua siitä, että ulkopuoliset eivät pääse käsiksi tietoihin. Salaaminen ei kuitenkaan muuttaisi tietojen asemaa henkilötietoina tietosuoja-asetuksen merkityksessä.[172]

Suhteellista perustetta noudattaen tilanne voi näyttäytyä toiselta ja salatut tiedot voitaisiin katsoa anonyymeiksi tiedoiksi ainakin tiettyihin tahoihin nähden. Suhteellisen perusteen mukaan mahdollisuudet yksilön tunnistamiseen pitäisi tarkastella suhteessa jokaiseen tietojen haltijaan erikseen. Tulee siis arvioitavaksi kysymys siitä, millaiset keinot sisältyvät kohtuullisen todennäköisiin keinoihin salauksen purkamiseksi ja täten rekisteröityjen tunnistamiseksi. Käytännössä on siis arvioitava, millainen salauksen taso on riittävä, jotta sen purkamista ei voida pitää kohtuullisen todennäköisenä keinona sellaiseen tahoon nähden, jolla ei ole käytettävissään salausavainta.

Salatun tiedon turvallisuutta suhteessa salauksen purkamiseen voidaan tarkastella kolmen objektiivisen tekijän valossa: 1. salausalgoritmin vahvuus 2. salausavaimen pituus ja 3. salausavaimen säilyttämisen järjestäminen.[173] Yleisin tapa salauksen purkamiseen on käyttää tyhjentävää avainhakua, eli arvata salausavaimia niin kauan, kunnes oikea avain löytyy.[174] Tällaista keinoa ei

[172] Spindler & Schmechel, s. 171.
[173] Hon - Millard - Walden 2011, s. 22.
[174] Gfirses & Preenel, teoksessa van der Sloot – Broeders – Schrijvers: Exploring the Boundaries of Big Data, s. 62.

kuitenkaan todennäköisesti voitaisi pitää kohtuullisen todennäköisenä keinona, mikäli tiedot on salattu käyttäen turvallista salaustekniikkaa.[175]

Ulkopuolisilla saattaa olla käytettävissään salausavaimen hankkimiseksi myös muita keinoja, joiden osalta on arvioitava, ovatko keinot kohtuullisen todennäköisesti käytettävissä. Avaimen luovuttamiseksi voi olla mahdollista saada tuomioistuimen päätös, avain voidaan poimia ohjelmistosta tai laitteistosta sekä voidaan hyödyntää salaustekniikassa olevia virheellisyyksiä tai ns. takaovia, joiden kautta salausavain saadaan selville. Näitä keinoja voidaan kuitenkin pitää kohtuullisen todennäköisinä vain, jos ne eivät ole lainvastaisia tai jos tietyllä hetkellä käytettävissä olevalla laskentateholla voidaan todennäköisesti ohittaa salaus.[176]

Tietojen käsittelyn hetkellä käytettävissä oleva salaustekniikka on otettava arvioinnissa huomioon. Julkisasiamies on lausunnossaan todennut, että salauksen purkamista ei voida pitää kohtuullisen todennäköisenä, jos purkaminen on käytännössä lähes mahdotonta toteuttaa. Mikäli salaus toteutetaan käyttämällä viimeisintä ja kehittyneintä salausteknologiaa, suurimmassa osassa tapauksia salauksen purkaminen olisi mahdotonta.[177]

Kuitenkin jos pelkästään potentiaalinen mahdollisuus salausavaimen saamiseksi toiselta laillisin keinoin olisi riittävä tunnistettavuuskriteerin täyttymiseen, mahdollisuudet välttyä tietosuojanormien sovellettavuudelta olisivat erittäin rajoitetut.[178]

Arvioinnissa tulee huomioida salauksen purkamisen mahdollistavien teknologioiden kehitys, kuten tietokoneiden laskentatehon lisääntyminen ja kehittyvät algoritmit.[179] Tietojen haltijan tulee ottaa huomioon teknologinen kehitys, joka tulee tapahtumaan sinä aikana, kun tietoja on tarkoitus käsitellä. Mikäli tietoja on tarkoitus käsitellä kymmenen vuoden ajan, myös teknologinen

[175] Cahsor & Sorge, teoksessa Borges & Meents, jotka väittävät, että käytettäessä 128-bittistä salausavainta, salauksen purkaminen olisi lähes mahdotonta ja siten epätodennäköistä.
[176] Spindler & Schmechel, s. 172.
[177] Julkisasiamiehen ratkaisuehdotus manuel campos sánchez-bordona 17.1.2018, johdanto-osan kohta 66.
[178] Spindler & Schmechel, s. 172.
[179] Tietosuoja-asetuksen johdanto-osan kohta 26.

kehitys ja sen mahdollisuudet on huomioitava koko tältä ajalta. Esimerkiksi jos jokin salattu tietoaineisto on teknologisten muutosten myötä mahdollista purkaa yhdeksäntenä vuotena, tiedoista tulisi henkilötietoja vasta siitä päivämäärästä lukien.[180]

Tietojen salauksen tasoa on siis arvioitava jatkuvasti uudelleen ottamalla huomioon teknologinen kehitys. Tiedot ovat anonyymejä vain tietyn ajanjakson verran, eikä tietojen haltijan ole riittävää arvioida salauksen tasoa ainoastaan tietojen käsittelyn aloittamishetkellä.[181] Vielä ongelmallisemmaksi tilanne muuttuu, jos luovutetaan valmiiksi salattuja tietoja edelleen. Tällöin tietojen vastaanottajan tulee varmistua siitä, sisältävätkö kyseiset tiedot alun perin henkilötietoja ja mikäli sisältävät, on tämän samalla tavoin jatkuvasti arvioitava salauksen tason riittävyyttä. Tässä tosin on huomattava, että vastuun laajuus vaihtelee eri henkilörelaatioissa. Rekisterinpitäjän vastuu verrattuna henkilötietojen käsittelijän vastuuseen tietosuojavelvoitteista on jossain määrin erilainen, riippuen muun muassa käsittelyn tarkoituksesta.

Mikäli tietojen haltijalla ei ole käytettävissään salausavainta tai muita keinoja päästä käsiksi salattuun tietoaineistoon, suurimmassa osassa tapauksia voidaan katsoa, että on kohtuullisen todennäköistä, että tämä ei pääse käsiksi henkilötietoihin. Tällöin salatut tiedot katsottaisiin siis anonyymeiksi tiedoiksi, joihin ei sovelleta EU:n tietosuoja-asetusta. Tämä edellyttää käytettävän viimeisintä salausteknologiaa ja huomioon on joka tapauksessa aina otettava myös kolmansien osapuolten potentiaaliset mahdollisuudet salausavaimen hankkimiseksi tai muutoin salauksen purkamiseksi. Näiden keinojen tulee kuitenkin olla kohtuullisen todennäköisiä, muussa tapauksessa salatun tiedon anonymiteetti säilyy.[182]

4.3 Vaikutukset henkilön yksityisyyteen ilman tunnistamista

[180] Tietosuojatyöryhmän lausunto 04/2007 s. 15.
[181] Lundevall-Unger & Tranvik kutsuvat tätä velvollisuutta rekisterinpitäjien taakaksi, jota ne eivät todennäköisesti pysty kantamaan ja tällainen vaatimus ei ainakaan lisää rekisterinpitäjien halukkuutta noudattaa EU:n tietosuojavelvoitteita (s. 71).
[182] Spindler & Schmechel, s. 173.

4.3.1 Uudelleentunnistamisen ongelma

Sellaiset tiedot, jotka eivät viittaa millään tapaa yksittäiseen henkilöön, kuten puhdas konetieto, eivät kuulu tietosuoja-asetuksen aineelliseen soveltamisalaan. Esimerkiksi sensorit, jotka keräävät dataa sovelluksiin, joiden tarkoituksena on analysoida ilmastoa tai tuotantolaitosten koneiden toimintaa eivät käsittele henkilötietoja missään vaiheessa.[183]

Tällainenkin tieto voi kuitenkin jossakin vaiheessa muuttua henkilötiedoksi, kun se yhdistetään muuhun tietoon. Tiedot, joita kerätään *esineiden internetin* kontekstissa esimerkiksi autoista, kodinkoneista ja niin sanotuista älykodeista voidaan todennäköisesti monessa tapauksessa yhdistää luonnollista henkilöä koskeviksi. Kehittyvät teknologiat luovat siten uusia haasteita tunnistettavuuskriteerin soveltamiselle. Tietojenkäsittelytieteissä löydetään jatkuvasti uusia tapoja, joilla aiemmin ei-henkilötiedoksi luokiteltuja tietoja yhdistelemällä tiedoista voidaan kuitenkin tunnistaa luonnollinen henkilö.[184] Tämä pätee myös sellaisiin tietoihin, joiden osalta tunnistettavuus on pyritty aktiivisin toimenpitein poistamaan, esimerkiksi anonymisoimalla.

Esimerkiksi internet-palveluntarjoaja AOL oli vuonna 2006 julkaissut 20 miljoonaa internet-hakua tieteelliseen tutkimukseen käytettäväksi. Näiden tietojen katsottiin olevan täysin anonymisoituja. New York Timesin toimittajat osoittivat kuitenkin pystyvänsä nopeasti tunnistamaan luonnollisen henkilön näiden hakujen perusteella ainakin osissa tapauksista.[185]

Uudelleentunnistamisen helppouden näyttämiseksi tehty tutkimus osoitti, että pelkästään postinumeron, syntymäajan ja sukupuolen perusteella kyettiin tunnistamaan 87 % Yhdysvaltain kansalaisista, vaikka näitä tietoja ei perinteisesti ole luettu henkilötiedoksi.[186] Tällaiset tiedot eivät myöskään ole erityisen arkaluonteisia tai muutoinkaan läheisesti henkilön identiteettiin liittyviä. Kuitenkin niiden yhdisteleminen uusien teknologioiden avulla mahdollistaa tunnistamisen, jolloin tietosuojasääntelyn kaikki velvoitteet tulee huomioida niiden käsittelyssä.

[183] Spindler & Schmechel, s. 168.
[184] Schwartz & Solove 2011, s. 1841.
[185] Barbaro & Zeller 2006.
[186] Sweeney 2000.

Teknologia siis mahdollistaa yhä useammassa tapauksessa tunnistamisen sellaisenkin tiedon perusteella, joka ei aluksi vaikuta lainkaan yksilöön liittyväksi tiedoksi.

Uudelleentunnistamisen kannalta toinen merkittävä tekijä on se, että ihmisistä saatavilla oleva tietomäärä on kasvanut räjähdysmäisesti. Mitä enemmän tietoa henkilöstä on saatavilla, sitä todennäköisempää on, että näitä tietoja voidaan hyödyntää tämän henkilön tunnistamiseksi tai luomaan uutta tietoa henkilöstä. Tietoja yhdistelemällä tuotetaan lisää tietoa, jolloin myös tietojen anonymisointi vaikeutuu huomattavasti.

Tietosuoja-asetuksessa on huomioitu erityisten tunnistetietojen vaikutus yksilön tunnistamiseen:

> ”Luonnolliset henkilöt voidaan yhdistää heidän käyttämiensä laitteiden, sovellusten, työkalujen ja protokollien verkkotunnistetietoihin, kuten IP-osoitteisiin, evästeisiin tai muihin tunnisteisiin, esimerkiksi radiotaajuustunnisteisiin. Näin käyttäjästä voi jäädä jälkiä, joita voidaan käyttää luonnollisten henkilöiden profilointiin ja tunnistamiseen etenkin, kun niitä yhdistetään yksilöllisiin tunnisteisiin ja muihin palvelimille toimitettuihin tietoihin.”[187]

Tietosuoja-asetuksessa on täten nimenomaisesti huomioitu uudenlaiset ympäristöt, joissa tietoja käsitellään ja että tällaisten tietojen perusteella tunnistaminen voi olla mahdollista.

Aikaisemmin ei-henkilötiedoksi katsottu tieto tulee monissa tapauksissa muuttumaan henkilötiedoksi esineiden internetin, Big Datan, verkkotunnistetietojen ja entistä tehokkaampien algoritmien vaikutuksesta. Lisäksi luonnollinen henkilö voi olla tunnistettavissa sen perusteella, että tätä koskevat tiedot on erotettu muista tiedoista, vaikka henkilön nimi ei niistä ilmenisikään. Kaiken kaikkiaan vaikuttaa siltä, että tulevaisuudessa tulee olemaan entistä vaikeampaa käsitellä tietoa siten, ettei sillä olisi mitään yhteyttä yksittäiseen henkilöön.[188]

Jos tunnistettavuuskriteerin tulkinnassa sovelletaan puhtaasti objektiivisen perusteen mukaista tulkintaa, on vaikeaa osoittaa, että tunnistamisen

[187] Tietosuoja-asetuksen johdanto-osan kohta 30.
[188] Spindler & Schmechel, s. 169.

mahdollistamia lisätietoja ei olisi jonkun muun henkilön hallussa. Objektiivisen perusteen mukaan tunnistamisen mahdollisuutta arvioidaan paitsi rekisterinpitäjän, myös kenellä tahansa muulla olevien tietojen perusteella. Voidaankin mielestäni väittää, että jokaisessa tapauksessa jollakin taholla on hallussaan sellaisia tietoja, joiden perusteella tai joita yhdistelemällä henkilö voidaan tunnistaa. Tämän vuoksi onkin tärkeää, että tulkinnassa hyödynnetään myös suhteellista perustetta siten, että tunnistamiseen vaadittava työmäärä ja tunnistamisen todennäköisyys otetaan huomioon. Tunnistettavuuden tulee olla kontekstiriippuvainen käsite, jota arvioidaan aina tapauskohtaisesti.

4.3.2 Profilointi

Tietojen hankkimisesta ja käsittelystä saatavat hyödyt eivät kuitenkaan aina ole samassa linjassa muiden tärkeiden yhteiskunnallisten intressien kanssa. Tietojen massamittainen käsittely saattaa loukata keskeisiä arvojamme ja perusoikeuksiamme, kuten yksityisyyttä, koskemattomuutta, autonomiaa ja muita vastaavia. Tietosuojalainsäädännöllä pyritään sovittamaan yhteen tietojenkäsittelyn mahdollistamia taloudellisia ja sosiaalisia etuja sekä perusoikeuksiamme ja arvojamme.[189] Koska tietosuojanormisto tulee kuitenkin sovellettavaksi ainoastaan silloin, kun tietojen perusteella on tunnistettavissa luonnollinen henkilö, monet sellaiset tiedot, jotka vaikuttavat näihin perusoikeuksiin ja -arvoihin, saattavat kuitenkin jäädä sääntelyn ulkopuolelle. Näin käy ainakin silloin, kun tunnistettavuutta arvioidaan supistavan tulkinnan kautta. Erityisesti suuret toimijat, jotka käsittelevät huomattavia määriä tietoja ja joiden liiketoiminta on riippuvainen tiedoista, pyrkivät vaikuttamaan siihen, että tunnistettavuuden oikeudellista arviointia ei laajenneta, vaan pikemminkin supistetaan.[190]

Esimerkiksi profilointia varten kerätty data, jota voidaan käyttää useisiin eri tarkoituksiin, mukaan lukien käyttäytymisen analysoinnin perusteella tapahtuvaan mainontaan, voi liittyä läheisestikin yksittäiseen henkilöön, mutta ei välttämättä kaikissa tilanteissa täytä tunnistettavuuden vaatimusta eikä siten ole

[189] Urgessa 2016, s. 523.
[190] Ibid.

henkilötietoa.[191] Sama pätee myös esimerkiksi dynaamisiin IP-osoitteisiin, vaikkakin Unionin tuomioistuin joltakin osin onnistui selkiyttämään oikeustilaa niiden osalta Breyer -tuomiossaan.

Profilointia käytetään useimmiten hyödyllisten tietojen poimimiseen valtavista tietomassoista käyttämällä tehokkaita tiedon louhimistekniikoita. Näillä tekniikoilla saadaan suuristakin tietomassoista nopeasti ja helposti poimittua relevantit osat, mihin ihmisten tietojenkäsittelykapasiteetti ei riittäisi.[192] Esimerkiksi internetin käytön seurannan tekniset mekanismit kehittyvät jatkuvasti, eikä keskivertokäyttäjä välttämättä pysty niitä edes havaitsemaan tai välttämään. Tällaisia seurantamekanismeja ovat esimerkiksi evästeet, nk. superevästeet, paikkatietojen ja sosiaalisen median seuranta.[193] Tiedoista luotavaa profiilia voidaan käyttää useisiin eri tarkoituksiin, mutta yleisimmin taloudellisiin tarkoituksiin kuten kohdennettuun mainontaan.

Profiloinnin avulla kertyneiden tietojen luonne muuttuu ongelmalliseksi silloin, jos ne eivät ole tietosuoja-asetuksen merkityksessä henkilötietoja, mutta kiistatta sisältävät elementtejä yksilöinnistä ja täten voivat potentiaalisesti vaikuttaa luonnollisten henkilöiden perusoikeuksiin ja -vapauksiin. Näin käy esimerkiksi silloin, kun profilointiprosessit käyttävät vain tietoja, jotka liittyvät koneisiin tai muihin ei-ihmiskohteisiin.[194] Näiden nk. ”harmaan alueen tietojen”[195] ei katsota täyttävän tunnistettavuuskriteeriä, koska nämä tiedot tunnistavat vain koneita, eivät luonnollisia henkilöitä[196]. Lisäksi, vaikka näiden tietojen avulla olisikin mahdollista tunnistaa luonnollinen henkilö, tunnistaminen tapahtuu vasta huomattavien resurssien käytön jälkeen, jolloin asetuksen kohtuullisen todennäköisesti -kriteeri jää kuitenkin täyttymättä.

Tällaiset argumentit voidaan kuitenkin helposti haastaa viittaamalla teknologiseen kehitykseen ja siihen, kuinka helposti sellaisia lisätietoja on

[191] Ibid.
[192] Bosco ym. s. 4.
[193] Skouma & Leonard 2015, s. 38-44.
[194] Bygrave 2002, s. 315.
[195] "Harmaalla alueella olevilla tiedoilla" viitataan niihin tietoihin, jotka eivät välttämättä ole yksiselitteisesti luokiteltavissa henkilötiedoksi, mutta silti kykenevät jollain tasolla yksilöimään niistä henkilöitä. Ks. Urgessa 2016, s. 524.
[196] Esimerkiksi Google on ilmaissut tämän kannan käytänteitään koskevassa blogissaan https://publicpolicy.googleblog.com/2008/02/are-ip-addresses-personal.html.

saatavilla, joiden avulla henkilö voidaan kuitenkin tunnistaa myös näiden vain koneita koskevien tietojen perusteella.[197] Tulevaisuudessa vastaavien argumenttien painoarvo tulee todennäköisesti pienenemään entisestään, kun uusia tiedonkeruumenetelmiä syntyy ja tietojenkäsittelykapasiteetti kasvaa.

Joka tapauksessa, yritykset saattavat voida hyödyntää tällaisia profiloinnin kautta kertyviä tietoja esimerkiksi internetmainonnassa joutumatta ikinä varsinaisesti tunnistamaan niistä yksittäistä henkilöä.[198] Niin kauan kun tietosuojasääntely EU:ssa rakentuu tunnistettavuuskriteerin varaan, ei kaikkien tietojen osalta voida ikinä olla yksiselitteisen varmoja niiden kuulumisesta henkilötietojen piiriin. Uusia tiedon kategorioita syntyy jatkuvasti uusia ja joudummekin todennäköisesti haastavien tulkintatilanteiden eteen myös jatkossa.

Tiedon hyödyntämistä kaupallisessa tarkoituksessa ilman tunnistamista voidaan konkretisoida käyttäytymiseen perustuvan mainonnan avulla. Evästeiden avulla verkkosivujen vierailijoista kerätään tietoa heidän käyttäytymisestään kyseisellä sivustolla ja tätä tietoa käytetään myöhemmin hyödyksi kohdennettujen mainosten tarjoamisen muodossa. Käyttäjistä kerätään tavallisesti niin kutsuttu ”virtuaalinen polku”, jonka käyttäjä jättää sivustolla vieraillessaan, johon lukeutuu jokainen verkkosivusto ja jokaisen verkkosivuston yksittäisen sivu, jolla käyttäjä käy, kuinka kauan käyttäjä on sivulla tai sivustolla, missä järjestyksessä sivuilla käytiin ja niin edelleen. Verkkosivuston ylläpitäjä, joka kerää näitä tietoja, antaa kullekin sivuston vierailijalle tietyn tunnistenumeron, jonka avulla ylläpitäjä tunnistaa tämän. Tämän sivuston ylläpitäjän näkökulmasta ei ole tarpeen tietää kyseisen käyttäjän todellista identiteettiä.

Tällöin ratkaisevaksi tekijäksi muodostuukin se, millä tavoin henkilön tulee olla tiedoista tunnistettavissa, jotta tieto on henkilötietoa. Riittääkö siis se, että henkilön tällainen ”virtuaalinen identiteetti” on tunnistettavissa vai pitääkö henkilön reaalinen, todellinen identiteetti paljastua? EU:n tietosuojatyöryhmä on lausunnossaan katsonut, että yleensä tiedot, joita käsitellään käyttäytymiseen perustuvan mainonnan tuottamiseen, olisi katsottava henkilötiedoiksi.[199] Samaan

[197] Ohm 2010.
[198] Schwartz & Solove 2011, s. 1848-1862.
[199] Tietosuojatyöryhmän lausunto 16/2011, s. 8. Katso lisäksi tietosuojatyöryhmän lausunto 2/2010.

päätelmään on tullut myös esimerkiksi Iso-Britannian tietosuojaviranomainen.[200] Tästä huolimatta on mahdollista väittää, että tunnistettavuuskriteerin täyttyminen edellyttää henkilön todellisen identiteetin, ei virtuaalisen identiteetin paljastumista.[201] EU:n tietosuojasääntelyssä on selvästi havaittavissa trendi sääntelyn soveltamisalan laajentamiseen ja tämä koskee myös tunnistettavuuskriteeriä. Oikeustila kuitenkin säilyy edelleen epäselvänä tällaisten ”harmaan alueen” tietojen osalta. Tämän vuoksi on mielestäni tärkeää, että EUT avaisi perusteluissaan tunnistettavuuskriteerin tulkintaa laajemmin.

Jos henkilötietojen määritelmä laajennetaan kattamaan lähes kaiken tiedon, ei yrityksillä ja muilla tietoja käsittelevillä tahoilla ole enää kannustimia panostaa anonymisointi- ja muihin tietoturvatekniikoihin. Tällöin sääntelyn perimmäisen tavoitteen toteutuminen tosiasiassa vaikeutuisi. Kuitenkin myös liian suppea tulkinta aiheuttaa edellä kuvatusti ongelmia yksityisyyden ja henkilötietojen suojan näkökulmasta.

Niin kauan kuin oikeustila säilyy epäselvänä tällaisten ”harmaan alueen” tietojen osalta, se vaikuttaa koko ajan perusoikeuksiimme ja -vapauksiimme. Kun yksilöt kokevat heidän olevan tarkkailun kohteena internetsivustoilla ja muussa verkkoasioinnissa, heidän halukkuutensa käyttää kyseisiä palveluita voi heikentyä. Tämä vuorostaan vaikuttaa haitallisesti kansalaisten yleiseen osallistumiseen ja panokseen demokraattisessa yhteiskunnassa.[202] Suurista ennakko-odotuksista huolimatta tietosuoja-asetus ei ole onnistunut poistamaan näitä epävarmuuksia, vaikkakin asetuksen valmistelussa on huomioitu tietotekniikan kehitys kauttaaltaan.

4.3.3 Yksilön v. ryhmän suojaaminen

Nykyinen tietosuojasääntely keskittyy suojaamaan yksittäistä luonnollista henkilöä ja tämän perusoikeuksia ja -vapauksia. Nykypäivän Big Data-aikakaudella uudet teknologiat ja tehokas analytiikka mahdollistavat suurten tietomäärien keräämisen ja analysoinnin, joiden avulla voidaan tunnistaa

[200] Information Commissioner's Office (UK) 2010, s. 22.
[201] Urgessa 2016, s. 525.
[202] Borgesius, s. 267-268.

yksilöistä muodostuvien ryhmien käyttäytymismalleja, joita hyödynnetään eri tarkoituksiin.[203] Tietoja analysoivat tahot käyttävät Big Dataa selvittääkseen ostokäyttäytymistämme, terveydentilaamme, unirytmiämme, liikkumistamme, ystävyyssuhteitamme ja niin edelleen. Ainoastaan muutamissa tapauksissa tällaisesta tiedosta tunnistetaan tai voidaan tunnistaa luonnollisia henkilöitä.[204] Näin ollen tällaiset tiedot eivät välttämättä ole henkilötietoja, koska ne eivät välttämättä tee tiettyä henkilöä tunnistettavissa olevaksi. Vaikka niihin ei siten sovellettaisikaan tietosuojasääntelyä, pelkästään Big Datan sisältämä tietojen määrä itsessään on erittäin suuri riski yksilön oikeuksille. Lisäksi tällainen määrä tietoa voi muodostaa niin sanottuja satunnaisia yhteyksiä, jolloin tietojen yhdistelemisen vaikutuksesta tunnistaminen voi tulla mahdolliseksi satunnaisissa tapauksissa. Toisin sanoen näin suuret tietomäärät tekevät satunnaisten yhteyksien todennäköisyyden suuremmaksi.[205]

Yksilön suojaamisen sijaan tulisi *Luciano Floridin* mukaan sääntelyssä kiinnittää enemmän huomiota myös ryhmän suojaamiseen, sillä aina ihmisistä ei olla kiinnostuneita yksilöinä vaan osana jotakin ryhmää. Todellisia kiinnostuksen kohteita, oikeuksien ja arvojen haltijoita sekä potentiaalisten riskien kantajia ovat ryhmät, eivät yksilöt.[206] Avoin data on omiaan luokittelemaan tyyppejä tai ryhmiä ennemmin kuin yksilöitä. Tämän vuoksi ei voida ajatella yksilöön keskittyvän tietosuojan suojaavan samalla myös jotakin ryhmää yksilön suojan kautta.[207] Floridi on havainnollistanut ryhmän yksityisyyden suojan tarvetta seuraavasti:

> *"There are very few Moby-Dicks. Most of us are sardines. The individual sardine may believe that the encircling net is trying to catch it. It is not. It is trying to catch the whole shoal. It is therefore the shoal that needs to be protected if the sardine is to be saved. […] Sometimes the only way to protect the individual is to protect the group to which the individual belongs. Preferably before any disaster happens."*

Tietosuojasääntelyn rakentuminen yksilökeskeisyydelle ei välttämättä suojaa luonnollisten henkilöiden perusoikeuksia ja -vapauksia toivotulla tavalla. Monesti oikeus yksityisyyteen, josta myös oikeus henkilötietojen suojaan on johdettu, nähdään henkilökohtaisen vapauden keskeisimpänä oikeudellisena takeena. Se

[203] Mantelero 2016, s. 8.
[204] Zwitter 2014, s. 4.
[205] Zwitter 2014, s. 5.
[206] Floridi 2014, s. 1.
[207] Ibid.

liitetään usein yksilön itsemääräämisoikeuteen ja autonomiaan. Yksityisyyden merkitys EU:ssa määritellään siis pitkälti yksilöllisyyden kautta.[208] Tietoyhteiskunnan aikakaudella on mahdollista kerätä big datan avulla kaupallisesti hyödynnettävissä olevaa tietoa tarkastelemalla ryhmien käyttäytymistä, jolloin yksittäisten henkilöiden yksityisyyteen ei ole välttämätöntä kajota, ainakaan perinteisesti ymmärretyllä tavalla. Myöskään tietosuoja-asetus ei tunnusta ryhmille minkäänlaista henkilötietojen suojaa, vaan keskittyy suojaamaan yksittäisiä tunnistettavissa olevia luonnollisia henkilöitä.

Uudenlaisissa tiedon analysointimenetelmissä ei ole enää keskeistä tunnistaa yksittäistä henkilöä. Tällaisissa menetelmissä sen sijaan tietoa kerätään laajasta ja määrittämättömästä joukosta.[209] Tietoa kerätään luomalla erilaisia kaavoja ja ryhmäprofiileja, joita voidaan hyödyntää laajasti esimerkiksi yritysten toimintamallien parantamiseen ja tuotekehitykseen. Yksilöt ovat edelleen keskeisessä asemassa tällaisen tiedon tuottamisessa, mutta EU:n tietosuojasääntely ei tunnusta niille suojaa, mikäli heitä ei voida ryhmien joukosta tunnistaa yksilöinä.

Ryhmien yksityisyyden käsite ei ole sinänsä uutta, mutta Big Datan avulla ryhmistä kerättävien tietojen suuren määrän vuoksi ryhmien yksityisyys saa uudenlaisen merkityksen. Perinteisesti ryhmien profilointiin ja kategorisointiin on käytetty vain muutamia eri muuttujia, kuten esimerkiksi sukupuoli, ikä ja siviilisääty. Big Data sen sijaan hyödyntää usein satoja eri muuttujia ryhmien tietojen analysointiin, minkä vuoksi se eroaa selkeästi perinteisestä mallista. Lisäksi analytiikan avulla luotujen ryhmien jäsenet eivät yleensä edes ole tietoisia kuulumisestaan tiettyyn ryhmään tai siitä johtuvista seurauksista.[210]

Kun suuria tietomääriä ryhmien jäsenistä analysoidaan tehokkaiden algoritmien avulla, ei yksittäisen luonnollisen henkilön tunnistettavuudella ole enää merkitystä. Tunnistettavissa on ainoastaan yksilöistä muodostuva ryhmä. Merkitystä olisi tämän vuoksi annettava tällaiselle ryhmäidentiteetille.[211] Täten

[208] Lindroos-Hovinheimo s. 58.
[209] Taylor ym. 2017, s. 20.
[210] Mantelero 2016, s. 2.
[211] Ibid.

myös ryhmien yksityisyys olisi ymmärrettävä uudella tavalla ja ryhmien yksityisyyteen kiinnittää enemmän huomiota myös lainsäädännön tasolla.

4.4 Riskiperusteinen tunnistettavuuden määrittely

Edellä olen käynyt läpi niitä haasteita, joita EU:n tietosuojasääntelyn soveltamisalan määrittely tunnistettavuuden kautta aiheuttaa, kun tietoja käsitellään uudenlaisten teknologioiden avulla ja erilaisissa käyttöympäristöissä. Oikeuskirjallisuudessa on aiheellisesti nostettu esiin nykyisen sääntelyn potentiaalisia riskejä yksilöiden perusoikeuksien ja -vapauksien kannalta. Teknologian kehittyessä tunnistettavuuskriteeriä olisi tulkittava siten, että kaikki sellainen tieto, joka liittyy tai vaikuttaa jollakin tavalla luonnollisen henkilön yksityisyyteen, tulisi tietosuojasääntelyn soveltamisalan piiriin. Kuitenkaan kriteerin soveltamisalaa ei voida laajentaa liiaksi, koska muutoin ajaudutaan tilanteeseen, jossa lähes kaiken tiedon on katsottava olevan henkilötietoa. Tällä taas olisi haitallisia vaikutuksia taloudellisten toimijoiden toimintavapauteen ja se heikentäisi tietojen vapaata liikkuvuutta huomattavasti. Lisäksi henkilötietoja käsittelevillä organisaatioilla ei olisi kannustimia toteuttaa tietoturvaa parantavia toimenpiteitä, joka taas heikentäisi tietosuojasääntelyn tavoitteiden toteutumista.

Tässä alaluvussa tarkastelen Paul Ohmin esittämää riskiperusteista mallia tunnistettavuuden määrittelyyn ja tarkastelen, millä tavoin tällaisia näkökohtia olisi huomioitava tunnistettavuuskriteeriä arvioitaessa. Tarkastelen asiaa erityisesti anonymisoinnin ja uudelleentunnistamisen näkökulmista, jotka liittyvät suoraan tunnistettavuuskriteerin arviointiin.

Ohmin riskiperusteinen lähestymistapa henkilötietojen suojaan pyrkii osoittamaan, että eri tilanteissa yksilön tarve oikeussuojalle on eriasteinen. Sen mukaan tietosuojasääntelyn tulisi ottaa huomioon se todellinen riskitaso, jonka tietojen käsittely aiheuttaa yksilön oikeuksille.

Ohm on luonut nk. ”testin”, jossa viiden tekijän avulla voidaan arvioida mahdollista riskiä yksityisyyden suojan loukkaukselle. Nämä viisi tekijää ovat

käytettävissä oleva tietojenkäsittelytekniikka, henkilötietoja sisältävän julkaisun yksityisyys tai julkisuus, kerättävien tietojen määrä, motiivi ja luottamus.[212]

Ensimmäinen arviointikriteeri eli käytettävissä oleva tietojenkäsittelytekniikka on otettava huomioon arvioitaessa sitä, onko henkilö tunnistettavissa jostakin tietystä tietoaineistosta. Algoritmien kehittyessä ja tietojenkäsittelynopeuden kasvaminen merkitsevät sitä, että tulevaisuudessa voi olla mahdollista tunnistaa henkilö sellaisen tiedon perusteella, joka tänä päivänä katsotaan anonyymiksi tiedoksi. Vastaavasti tietojenkäsittelytekniikan kehittyminen vaikuttaa myös toiseen suuntaan eli myös tietojen suojaaminen voi tulevaisuudessa olla luotettavampaa. Käytettävissä olevan tietotekniikan avulla voidaan luoda anonymisointimalleja, jotka takaavat entistä vahvemman suojan henkilötiedoille.

EU:n yleisen tietosuoja-asetuksen johdanto-osan 26 kohdan mukaan anonymiteetin arvioinnissa tulee ottaa huomioon muun muassa käsittelyajankohtana käytettävissä oleva teknologia ja tekninen kehitys. Tietosuoja-asetuksen soveltamisessa voidaan siis ottaa huomioon tietojenkäsittelytekniikan taso ja kehittyminen, ainakin arvioitaessa tiedon anonyymiutta. Vastaavasti se voitaisiin huomioida tulkittaessa yksilön oikeussuojan tarvetta muutoinkin. Teknologian huomioiminen on ongelmallista silloin, kun joudutaan arvioimaan sen kehittymistä tulevaisuudessa. Sellainen tieto, jolla ei tänä päivänä todennäköisesti ole vaikutuksia yksilön oikeuksiin tai vapauksiin, voi myöhemmin muuttua suoraan tähän yksilöön liittyväksi tiedoksi. Tällöin myös tietosuojanormiston olisi automaattisesti tultava sovellettavaksi tällaisten tietojen käsittelyyn.

EU-tuomioistuimen ei kuitenkaan tulisi mielestäni tehdä sellaisia tulevaisuuteen pohjautuvia arvioita tietojen luonteesta, vaan tarkastella niitä ratkaisuhetken mukaisen tietämyksen valossa. Olisi jokseenkin erikoista, jos tuomioistuin perustelisi tietosuojasääntelyn soveltamista johonkin tietojoukkoon sillä, että todennäköisesti myöhemmin tulevaisuudessa tiedot vaikuttavat yksilön oikeuksiin ja vapauksiin, vaikka ratkaisuhetkellä näin ei olisi. Sen sijaan yritysten ja muiden tietoja käsittelevien organisaatioiden tulisi ottaa myös tulevaisuuden

[212] Ohm 2010, s. 1765-1768.

näkymät huomioon tietoja käsitellessään ja tietoturvatoimenpiteitä toteuttaessaan.

Toisena riskiarviointikriteerinä Ohm mainitsee ”yksityinen v. julkinen” -jaottelun. Merkitystä tulee siis antaa sille, onko jokin henkilötietoja sisältävä julkaisu suunnattu suurelle yleisölle vai onko se yksityinen, esimerkiksi vain kahden tahon välinen tietojenvaihto. Julkaisut, jotka on tarkoitettu julkisiksi, tulisi Ohmin mukaan tutkia tarkemmin kuin yksityisten välinen tietojen luovutus, sillä suurelle yleisölle suunnattuun tietoaineistoon on vaivattomampi pääsy. Tämän hetkinen tunnistettavuuden tulkinta EU:ssa ei näyttäisi ottavan tällaista jaottelua huomioon. Esimerkiksi Breyer -tuomion perusteluita mukaillen voitaisiin päätyä lopputulokseen, että kahden välisessä tietojenvaihdossa siirtyvä tietoaineisto joko on tai ei ole henkilötietoa, riippuen siitä, onko jollakin muulla mahdollisesti kyky, toisin sanoen oikeudelliset keinot, tunnistaa niistä luonnollinen henkilö. Perusteluissa ei ole otettu kantaa siihen todelliseen riskiin, joka yksityisyydelle aiheutuu tapauksen olosuhteissa, joissa oli kysymys yksityisen tahon omia tarpeita varten säilyttämistä tiedoista.

Jaottelua voitaisiin nähdäkseni hyödyntää tunnistettavuuskriteerin tulkinnassa ainakin arvioitaessa anonymisointitekniikoiden luotettavuutta. Tietojen anonymisoimiseksi käytetty teknologia tulisi olla sitä vahvempi, mitä julkisempia tiedot ovat. Esimerkiksi yksityiseen tutkimuskäyttöön tarkoitettujen tietojen osalta voitaisiin katsoa alhaisempi anonymisoinnin taso riittäväksi kuin täysin julkiseksi tarkoitettujen tietojen osalta. Kummassakin tapauksessa anonymisointitekniikan olisi kuitenkin oltava sillä tavoin luotettava, että henkilön tunnistaminen ei ole tunnettujen menetelmien valossa mahdollista ilman suhteetonta vaivaa. Jaottelua voitaisiin siten hyödyntää erityisesti arvioitaessa tietosuoja-asetuksen johdanto-osan kohdan 26 mukaisia kohtuullisen todennäköisiä keinoja henkilön tunnistamiseksi.

Kolmanneksi Ohm mainitsee tietojen keräämisen sallitun määrän. Hänen mukaansa tulisi myös säännellä sitä, kuinka paljon rekisterinpitäjä voi ylipäänsä kerätä tietoja. Tietovarantojen koko on ollut merkittävä avustava tekijä, kun ollaan tutkittu uudelleentunnistamisen mahdollisuutta eli mitä enemmän tietoa on ollut

käytettävissä, sitä todennäköisempää on myös uudelleentunnistaminen.[213] Tietosuoja-asetukseen sisältyy periaate tietojen minimoinnista[214], mutta mitään määrällistä rajoitusta sinänsä ei ole pyrittykään määrittämään lainsäädännön tasolla. Periaatteen mukaan henkilötietojen on oltava asianmukaisia ja olennaisia ja rajoitettuja siihen, mikä on tarpeellista suhteessa niihin tarkoituksiin, joita varten niitä käsitellään.

Anonymisoitujen henkilötietojen kohdalla ongelmalliseksi muodostuu kuitenkin se, ettei yleisen tietosuoja-asetuksen periaatteita sovelleta anonymisoituun tietoon. Tällaista tietoa voi yksittäisellä taholla olla siis hallussaan kuinka paljon tahansa. Mielestäni tietojen keräämisen määrällinen rajoittaminen ei ole henkilötietojen suojan kannalta olennaista, koska siinä vaiheessa, kun tietoja on kertynyt yksittäiselle toimijalle niin paljon, että luonnollisen henkilön tunnistaminen on mahdollista, tulee asetus ja sen periaatteet kaikilta osin sovellettavaksi, jolloin erillistä suojamekanismia ei tarvita. Tietojen määrä voidaan kuitenkin ottaa huomioon yhtenä arviointikriteerinä osana riskiperusteista arviointia, mutta määrällisiä rajoituksia ei pidä mielestäni sisällyttää lainsäädäntöön.

Tietojen määrä voitaisiin huomioida esimerkiksi siten, että suuri määrä tietoja voi osoittaa, että henkilön tunnistaminen tietoja yhdistelemällä on todennäköisempää kuin silloin, jos tietoja on vain hyvin vähän. Tietojen määrälle ei voida kuitenkaan nähdäkseni antaa kovinkaan suurta painoarvoa tunnistettavuuskriteerin arvioinnissa. Vaikka uudelleentunnistamisen mahdollisuus kasvaakin tiedon määrän lisääntyessä, tämä ei aina pidä paikkaansa ja riski tunnistamiseen riippuu paljolti tietojen sisällöstä. Pelkästään tietojen määrän osalta ei voida siis päätellä sitä riskiä, joka yksilön henkilötietojen suojalle käsittelystä aiheutuu.

Neljäntenä arviointikriteerinä Ohm nostaa esiin motiivin. Tällä hän tarkoittaa motiivia uudelleentunnistamiseen eli huomioon olisi otettava yksittäisen toimijan tunnistamisesta mahdollisesti saama hyöty. Hänen mukaansa esimerkiksi arkaluonteinen tieto on usein sellaisten toimijoiden hallussa, joilta puuttuu motiivi

[213] Ohm 2010, s. 1766-1767.
[214] Tietosuoja-asetuksen 5 (1 c) artikla.

uudelleentunnistaa niistä luonnollisia henkilöitä.[215] Sääntelyssä tulisikin kiinnittää huomiota siihen, millaisilla tahoilla tällainen motivaatio voisi todennäköisesti olla. Yhtenä, joskin ei ainoana, tekijänä tulisi ottaa huomioon uudelleentunnistamisesta saatava taloudellinen hyöty. Esimerkiksi anonyymin tietoaineiston avaaminen kohdennettua mainontaa varten olisi tällainen taloudellinen motiivi. Motiivin tarkastelu muistuttaa läheisesti tietosuoja-asetukseen sisältyvän käyttötarkoitussidonnaisuuden sisältöä siinä mielessä, että molemmissa huomiota pitää kiinnittää tietojen käyttötarkoitukseen. Toinen asia on se, että anonyymien tietojen haltijan käyttötarkoitusta voi olla haastavaa määrittää, koska mikään säännös ei velvoita tätä yksilöimään tätä tarkoitusta, toisin kuin henkilötietojen osalta.

Motiivia voitaisiin hyödyntää tunnistettavuuskriteerin arvioinnissa erityisesti siten, että sellaista organisaatiota, joka käsittelee henkilötietoja pääosin kaupallisiin tarkoituksiin, koskisi tiukemmat vaatimukset tietojen anonymisoinnin osalta kuin vaikkapa tieteelliseen tutkimukseen tietoja käyttävää tahoa. Tällaisessa arvioinnissa tulisi huomioitua ne todelliset riskit, jotka luonnollisen henkilön oikeuksille ja vapauksille aiheutuu tietojen käsittelystä. Toisaalta, henkilötietojen suojan merkitys on viime aikoina kehittynyt siihen suuntaan, että henkilöä koskevia tietoja suojataan riippumatta siitä, mihin tarkoitukseen niitä käsitellään. Motiivin tarkastelu voisi kuitenkin olla hyödyllistä sellaisissa rajanvetotapauksissa, joissa ei voida osoittaa varmaksi, että henkilö olisi tiedoista tunnistettavissa. Mikäli tietoja käsittelevän organisaation motiivina voitaisiin osoittaa olevan henkilön tunnistaminen, voitaisiin tulkita tunnistettavuuskriteerin täyttyvän. Erityisen hyödyllistä tämä olisi arvioitaessa sellaista tietojenkäsittelyä, joka perustuu profilointiin.

Viimeisenä riskiarviointikriteerinä Ohm mainitsee luottamuksen. Sääntelyssä tulisi kiinnittää huomiota yksityisten ihmisten tai instituutioiden väliseen luottamukseen. Käytännössä Ohm siis ehdottaa, että lainsäädännön tasolla tulisi määritellä, mitkä tahot ovat luotettavia tietojen haltijoita ja mitkä eivät. Yksityisyyden suojan tarve olisi erilainen perustuen johonkin lähtökohtaiseen olettamaan tiedot vastaanottavan tahon identiteetistä. Esimerkiksi tieteelliseen

[215] Ohm 2010, s. 1767.

tutkimukseen voitaisiin luovuttaa tietoja vapaammin kuin kaupallisesti tietoja hyödyntäville yrityksille.

Tietosuoja-asetuksen 5 (1 b) artiklassa säädetään, että ”*myöhempää [henkilötietojen] käsittelyä yleisen edun mukaisia arkistointitarkoituksia taikka tieteellisiä tai historiallisia tutkimustarkoituksia tai tilastollisia tarkoituksia varten ei katsota […] yhteensopimattomaksi alkuperäisten tarkoitusten kanssa*”. Tietosuoja-asetuksessa siis huomioidaan ainakin jossakin määrin se luottamus, joka vallitsee koskien tietojen käsittelyä tällaisia tarkoituksia varten. Tietosuoja-asetuksen 89 (1) artiklan mukaan tällaiseenkin käsittelyyn kuitenkin on sovellettava rekisteröidyn oikeuksia ja asianmukaisia suojatoimia asetuksen mukaisesti. Nähdäkseni luottamusta sellaisenaan ei voida ottaa arviointikriteeriksi tunnistettavuuden osalta. Sillä on kuitenkin oltava merkitystä henkilötietojen käsittelyn oikeusperusteiden osalta, eli käsittely tällaisiin tarkoituksiin tulee sallia laajemmin, kuin henkilötietojen käsittely muutoin. Näin ollaan myös tietosuoja-asetuksessa säädettykin.

Ohmin näkemyksiä on vastustettu esimerkiksi väittämällä, että tietojenkäsittelytieteellinen kirjallisuus ymmärretään väärin ja tämän seurauksena liioitellaan anonymisoinnin toimimattomuutta ja anonymisoinnista syntyvät todelliset riskit ovat vähäisiä tai jopa merkityksettömiä.[216] Mielestäni kuitenkin erityisesti uudelleentunnistamisen helppouden ja teknologioiden kehittymisen vuoksi anonymisoinnin tehokkuutta tunnistettavuuskriteerin kannalta on arvioitava kriittisesti ja arvioinnissa on huomioitava kaikki käytettävissä oleva teknologia. Esimerkiksi tänä päivänä luotettavan anonymisoinnin tarjoava salaustekniikka ei välttämättä vuoden päästä ole enää lainkaan tehokas.

5 Johtopäätökset

Henkilötiedon käsitteen osatekijöistä tunnistettavuus on yleensä ongelmallisin ja eniten tulkinnanvaraa sisältävä tekijä. Oikeuskirjallisuudessa esiintyy hyvin erilaisia kannanottoja siitä, kuinka tätä kriteeriä olisi tulkittava. Tärkeä esille

[216] Yakowitz 2011, s. 4.

nostettu ongelma koskee sitä, että tunnistettavuuskriteerin liian laajalla tulkinnalla voidaan ajautua tilanteeseen, jossa kaikki tieto katsotaan lopulta henkilötiedoksi. Tämä taas poistaisi rekisterinpitäjiltä kannustimet toteuttaa tietoturvatoimenpiteitä. Lisäksi nykyinen EU:n yleinen tietosuoja-asetus olisi seikkaperäisyytensä ja sen sisältämien velvoitteiden vuoksi liian raskas noudatettavaksi, mikäli sen soveltaminen koskisi kaikkea tietoa.

Tietosuoja-asetuksen mukainen tunnistettavuuskriteeri koskee myös tilanteita, joissa henkilö voidaan epäsuorasti tunnistaa. Käytännössä tämä tarkoittaa sitä, että henkilön tunnistamiseksi tarvittavien tietojen ei tarvitse olla saman henkilön hallussa. Kriteeri täyttyy myös siinä tapauksessa, että eri henkilöillä olevia tietoja yhdistelemällä voidaan tunnistaa yksittäinen henkilö. Tällaisten toiselta henkilöltä saatavien lisätietojen tulee kuitenkin olla hankittavissa kohtuullisen todennäköisten keinojen avulla. Vaikka EUT on Breyer -tuomiossaan jossakin määrin avannut tällaisten kohtuullisen todennäköisten keinojen tulkintaa, ei tarkkoja tulkintalinjoja voida esittää. Varsinkin Big Data -ympäristöissä tapahtuva käsittely aiheuttaa ongelmia sen suhteen, millaisia keinoja voidaan pitää kohtuullisen todennäköisinä. Kun algoritmit ja muu tietojenkäsittelyteknologia jatkuvasti kehittyy, voi jatkossa olla vaikeaa perustella, ettei henkilön tunnistaminen olisi lainkaan mahdollista jostakin muualta saatavien tietojen avulla.

Ennen internetin käytön ja erilaisten teknologioiden kehittymisen räjähdysmäistä kasvua tunnistettavuuskriteerin tulkinta ei ole ollut kovinkaan ongelmallista. Yksittäisten tietojen osalta on ollut mahdollista päätellä, onko luonnollinen henkilö ollut niistä tunnistettavissa. Kun käytettävissä olevan tiedon määrä on ollut vähäinen, ei tunnistamisen mahdollisuutta olla aiemmin jouduttu tarkastelemaan yhtä seikkaperäisesti kuin nykyisin. Kun tietokoneet ja kehittyneet algoritmit ovat mahdollistaneet valtavan datamäärän käsittelemisen ja yhdistelemisen hyvinkin nopeassa tahdissa, luonnollisen henkilön tunnistaminen on tullut mahdolliseksi eri tietoja yhdistelemällä sellaisenkin tiedon perusteella, jota ei perinteisesti ole henkilötiedoksi mielletty.

EUT näyttää tulkitsevan tunnistettavuuskriteeriä soveltamalla sekä objektiivisen että suhteellisen perusteen mukaista tulkintaa, painottaen kuitenkin enemmän

objektiivista perustetta. EUT:n mukaan tiedon luokitteleminen henkilötiedoksi ei edellytä, että tämä tieto yksin mahdollistaa rekisteröidyn tunnistamisen, eikä kaikkien tietojen, joiden perusteella henkilö voidaan tunnistaa, tarvitse olla yhden ainoan tahon hallussa. Lisätietojen yhdistäminen on kuitenkin oltava kohtuullisesti toteutettavissa oleva keino. EUT:n perusteluista ei ole löydettävissä seikkaperäistä arviointia siitä, milloin tällainen yhdistäminen olisi katsottava kohtuulliseksi keinoksi. Se ainoastaan toteaa, julkisasiamiehen ratkaisuehdotukseen viitaten, että keino ei ole kohtuullisesti toteutettavissa, kun rekisteröidyn tunnistaminen on kielletty laissa tai kun se ei ole käytännössä toteutettavissa esimerkiksi siitä syystä, että se veisi suhteettomasti aikaa ja aiheuttaisi suhteettomasti kustannuksia ja työtä. Kaiken kaikkiaan EUT:n perustelut viittaavat siihen, että tunnistettavuuskriteeriä olisi sovellettava mahdollisimman laajasti hyödyntäen objektiivisen perusteen mukaista tulkintaa. Kuitenkin poikkeuksellisissa tapauksissa suhteellista perustetta voidaan käyttää rajaamaan sellaiset tiedot pois henkilötiedon piiristä, joissa tunnistamiseksi vaaditut lisätiedot olisivat erittäin vaikeasti saatavissa tai tunnistaminen olisi laissa kiellettyä.

Tietosuoja-asetuksen tavoitteiden toteutumisen kannalta on tärkeää, että rekisterinpitäjille ja muille henkilötietoja käsitteleville tahoille on olemassa kannustimet luoda ja ylläpitää erilaisia tietoturvatekniikoita. Tällaisia kannustimia ovat esimerkiksi edellä käsitellyt tietojen anonymisointi, salaaminen ja pseudonymisointi. Kannustimien painoarvo on pitkälti riippuvainen siitä, mihin suuntaan tunnistettavuuskriteerin tulkinnassa EU:ssa jatketaan. Julkisasiamies ja tietosuojatyöryhmä ovat pyrkineet osaltaan ainakin jossain määrin vaikuttamaan siihen, että henkilötietojen määritelmää ei laajennettaisi liian laajaksi. EUT:llä ei ole ollut velvollisuutta perustaa ratkaisuaan näiden esittämiin tulkintoihin ja se onkin omilla ratkaisuillaan nähdäkseni pyrkinyt ylläpitämään henkilötiedon mahdollisimman laajaa tulkintaa.

Nähdäkseni tunnistettavuuskriteerin tulkinnassa olisi huomioitava käytettävissä oleva tietojenkäsittelytekniikka, tietojen määrä, tietojen käsittelyn tarkoitukset ja se, millaiselle yleisölle tiedot on tarkoitettu. Näin arvioinnissa voitaisiin paremmin huomioida se tosiasiallinen riski, joka henkilön yksityisyydelle ja tietosuojalle tietojen käsittelystä aiheutuu yksittäisessä tilanteessa. Vaikka EU:ssa

henkilötieto määritellään kaksijakoisella joko-tai periaatteella, voidaan sääntelyyn sisältyvä joustavuus kuitenkin hyödyntää nimenomaan tunnistettavuuskriteeriä tulkittaessa. Näin voidaan välttää henkilötiedon käsitteen liiallinen laajentaminen, mutta kuitenkin toteuttaa tietosuojasääntelyn perimmäinen tavoite, luonnollisten henkilöiden perusoikeuksien ja -vapauksien suojelu.